

SplatAudit: Reference-Free Geometric Auditing of Floater Artifacts in 3D Gaussian Splatting

Ian Wai Si
Tsinghua University
Beijing, China

srw24@mails.tsinghua.edu.cn

Elaine Huang
Mission San Jose High School
Fremont, California, USA
elahua352@fusdk12.net

Abstract

3D Gaussian Splatting (3DGS) represents a scene as a collection of explicit Gaussian primitives, yet reconstruction quality is usually judged from rendered images. Metrics such as PSNR, SSIM, and LPIPS measure image fidelity without directly examining whether the learned primitive field forms a coherent 3D representation. This leaves a data-quality gap in which geometric failures such as floater artifacts remain difficult to assess without reference geometry.

We present SplatAudit, a post-hoc, reference-free framework for auditing the geometric coherence of pretrained 3DGS scenes. Each Gaussian defines an anisotropic support neighbourhood, within which surface variation and mutual coverage measure local surface evidence and reciprocal embedding. Their combination produces primitive-level confidence, while projection through the 3DGS compositor yields per-view heatmaps and an opacity-normalised Floater Artifact Score (FAS). For compression, we combine the ranked geometric confidence with PointSplat’s opacity–volume importance through a harmonic mean, favouring primitives supported by both geometric and intrinsic evidence.

Experiments on all seven Mip-NeRF 360 scenes under a sparse 24-view setting show that, at 20% retention, SplatAudit obtains the lowest FAS in every scene. The harmonic ranking improves all three mean image metrics over geometry alone while retaining a lower mean FAS than PUP and PointSplat. At 10% retention, it achieves the best PSNR and SSIM on five scenes among the compressed methods and remains below both baselines in FAS throughout. These results show that geometric coherence and rendering importance capture complementary properties of explicit learned representations.

CCS Concepts

• **Computing methodologies** → **Rendering; Point-based models; Anomaly detection.**

Keywords

3D Gaussian Splatting, floater artifacts, geometric auditing, reference-free evaluation, primitive pruning

1 Introduction

A recurring challenge in data-driven modelling is that standard evaluation metrics do not always capture the failure modes that matter most in practice. A model can perform well on established benchmarks while retaining structurally questionable internal components. In such cases, evaluating its outputs is only part of the problem; the learned representation is itself an object of evaluation. From a data-centric perspective [33], aggregate performance describes model outputs, whereas an audit examines the structure encoded within the model.

This problem appears clearly in 3D Gaussian Splatting (3DGS) [10]. A reconstructed scene is stored as an explicit collection of anisotropic Gaussian primitives, yet its quality is usually judged from rendered images. Measures such as PSNR, SSIM, and LPIPS quantify agreement with held-out views, but do not directly test whether the primitive field forms a coherent 3D representation. This limitation is especially visible in *floater artifacts*: Gaussians that occupy weakly supported locations away from scene surfaces, often where sparse observations leave the photometric objective under-constrained [30, 36]. Recent artifact taxonomies similarly distinguish geometric failures from ordinary rendering distortions [25].

Existing evaluation tools only partially address this gap. Photometric metrics measure rendered images rather than the 3D structure that produced them. Reference-based geometric distances, such as Chamfer distance, require ground-truth geometry that is unavailable in standard novel-view synthesis benchmarks and measure alignment to that reference rather than coherence within the primitive field. Neither approach directly audits the stored Gaussians. A complementary diagnostic should therefore assess primitive-level geometric consistency from the representation itself, without reference geometry or retraining.

We therefore introduce **SplatAudit**, a post-hoc, reference-free geometric audit of pretrained 3DGS scenes. Each Gaussian’s 3σ support defines a directed neighbourhood in which surface variation captures local structure and mutual coverage tests the reciprocity of support. Together they define primitive-level geometric confidence. Its projection through the 3DGS compositor produces per-view heatmaps and an opacity-normalised Floater Artifact Score (FAS) that localise and summarise geometric evidence.

Geometric confidence also induces a primitive-pruning ranking. We combine its scene-wise rank with the opacity–volume importance of PointSplat [20] through a harmonic mean, favouring primitives supported by both criteria. Across seven Mip-NeRF 360 scenes [1], geometry-guided selection consistently produces lower FAS than the importance-based baselines, while harmonic fusion preserves more rendering quality than geometry alone. This contrast supports the central premise that geometric coherence and rendering importance describe complementary properties of the learned representation.

The main contributions of this paper are:

- a reference-free geometric audit based on representation-aligned support, surface variation, and mutual coverage;
- primitive confidence, FAS, and per-view heatmaps for examining and localising geometric inconsistency; and
- a harmonic pruning score combining geometric confidence with PointSplat importance.

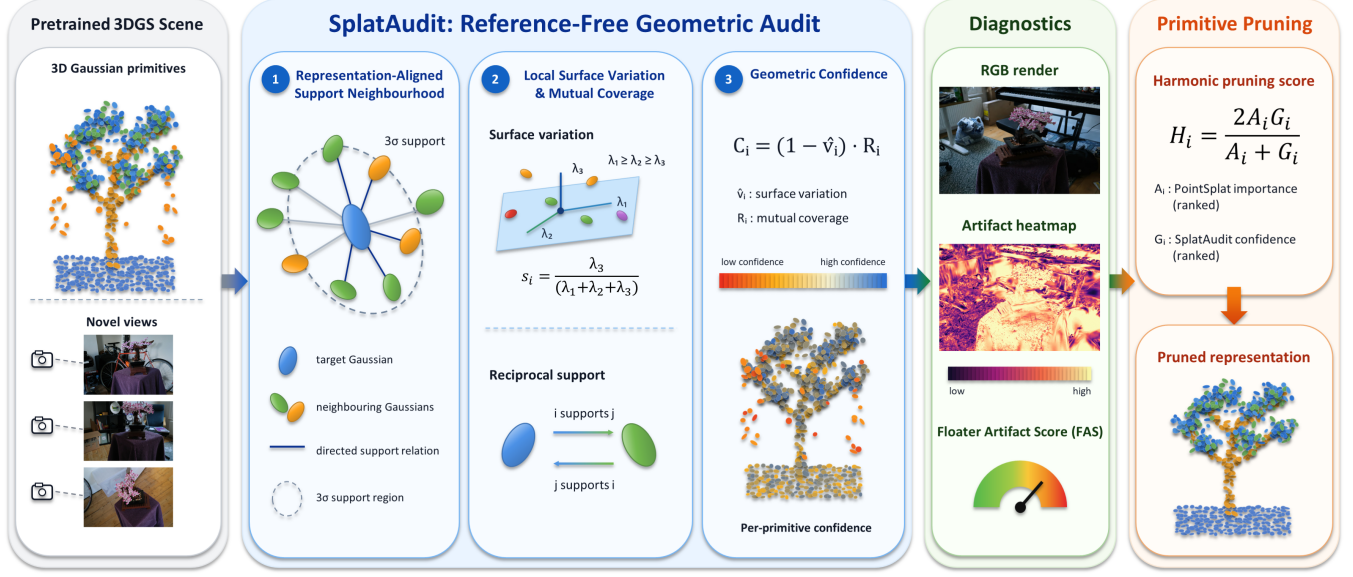


Figure 1: SplatAudit pipeline. Representation-aligned support, surface variation, and mutual coverage define geometric confidence, which supports view-dependent diagnostics and confidence-guided primitive pruning.

2 Related Work

2.1 Artifacts and Geometric Regularisation in 3DGS

Most work on 3DGS artifacts treats them as failures to prevent during optimisation. Sparse-view methods compensate for weak multiview constraints through depth priors, few-shot regularisation, or additional observations [13, 21, 30, 36]. Others intervene in primitive creation, density control, and removal [17, 24, 35]. Recent approaches increasingly combine reliability cues: TIDI-GS couples pruning with depth supervision [31], while DOC-GS calibrates evidence across optimisation and observation domains [12]. ArtifactWorld instead organises artifact types to learn restoration of their rendered effects [25]. These methods prevent or repair artifacts by changing training or learning a restoration model; they do not directly audit an arbitrary pretrained primitive field.

A parallel line treats learned Gaussian geometry as useful structure. FeatureGS and EntON derive eigenvalue-based local features for regularisation or densification [8, 9], while 2DGS, GSRec, SuGaR, and Gaussian Opacity Fields recover or align surfaces from Gaussian representations [5, 7, 28, 32]. Together, they establish that primitive geometry carries information beyond rendered appearance, although their objectives remain tied to training or surface reconstruction. SplatAudit instead forms a post-hoc neighbourhood from each Gaussian’s anisotropic support and examines both local surface variation and reciprocal support.

2.2 Primitive Pruning and Compression

The compression literature asks which primitives can be removed while preserving rendered appearance. Existing approaches combine importance pruning, learned masks, quantisation, entropy models, structured anchors, and constrained primitive counts [3,

4, 11, 15, 29]; more recent work couples selection to new compact primitive constructions [23]. Together, these methods reveal substantial redundancy in trained Gaussian fields, but they optimise rendering utility rather than geometric consistency. This objective is appropriate for compact rendering, but does not distinguish a visually dispensable primitive from a geometrically anomalous one.

PUP and PointSplat expose this distinction most clearly. PUP estimates how spatial parameters affect reconstruction error through a second-order photometric sensitivity score [6]; PointSplat ranks Gaussians from intrinsic opacity and volume [20]. Both are relevant post-hoc comparators, yet neither tests whether a primitive is coherently embedded in neighbouring Gaussian support. This distinction motivates combining SplatAudit with PointSplat rather than treating either criterion as a substitute for the other.

2.3 Reference-Free Quality Assessment

Reference-free quality assessment spans learned features from rendered novel views [19], subjective quality prediction from native Gaussian attributes [22], and geometric assessment of point clouds [14]. These methods principally predict aggregate quality; removing the reference requirement does not by itself make an assessment local or interpretable. An aggregate prediction may indicate degradation without identifying the primitives or spatial relations responsible.

SplatAudit instead views the explicit Gaussian field as structured model data whose local consistency can be inspected [33]. This shares the local-relative principle of outlier analysis [2], but specialises it to anisotropic Gaussian support. Its non-learned primitive score and projected evidence remain traceable to the geometric relations that produced them.

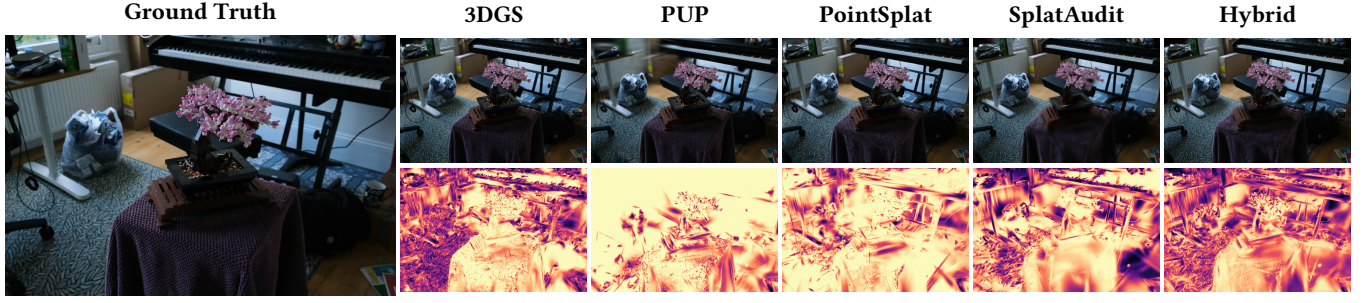


Figure 2: Qualitative comparison on the bonsai scene. The upper row shows RGB renders and the lower row shows the corresponding FAS maps.

3 Methodology

Let $\mathcal{G} = \{g_i\}_{i=1}^N$ denote a pretrained 3DGS scene. SplatAudit evaluates whether each primitive is coherently embedded in the learned Gaussian field. Its design rests on two geometric properties expected of a surface primitive: the surrounding centres should exhibit locally surface-like structure, and the primitive should share reciprocal support with its neighbours. Surface variation characterises the former, while mutual coverage characterises the latter. Their combination defines primitive confidence, whose projection through the 3DGS compositor yields a view-dependent measure of floater evidence. Figure 1 summarises the audit, diagnostic, and pruning stages.

3.1 3D Gaussian Splatting Preliminaries

In 3DGS [10], primitive g_i has centre $\mu_i \in \mathbb{R}^3$, covariance Σ_i , opacity α_i , and view-dependent colour. Its unnormalised spatial density is

$$\mathcal{G}_i(\mathbf{x}) = \exp\left[-\frac{1}{2}(\mathbf{x} - \mu_i)^\top \Sigma_i^{-1}(\mathbf{x} - \mu_i)\right]. \quad (1)$$

Following the image-formation model shared by neural volumetric rendering and 3DGS [10, 16], the colour of pixel p is obtained by depth-ordered alpha blending:

$$C(p) = \sum_{i=1}^N T_i(p) \alpha_i(p) c_i, \quad T_i(p) = \prod_{j=1}^{i-1} (1 - \alpha_j(p)). \quad (2)$$

Here $\alpha_i(p)$ denotes the projected 2D Gaussian weighted by its learned opacity α_i . The spatial density in Eq. (1) determines which primitives provide geometric support, while the compositing weights in Eq. (2) determine how that support is expressed in an observed view.

3.2 Representation-Aligned Geometric Confidence

Support Neighbourhood. Unlike an unstructured point, a Gaussian primitive has a learned anisotropic extent that determines the region over which it represents scene content. Its covariance factorisation is

$$\Sigma_j = \mathbf{U}_j \mathbf{S}_j \mathbf{S}_j^\top \mathbf{U}_j^\top,$$

where $\mathbf{U}_j$ is the rotation and $\mathbf{S}_j$ is the diagonal matrix of positive axis scales. These attributes define an anisotropic ellipsoid and provide a representation-aligned notion of proximity. Accordingly,

the neighbourhood of g_i is formed by the primitives whose 3σ support contains its centre μ_i :

$$\mathcal{N}_i = \{j \neq i : (\mu_i - \mu_j)^\top \Sigma_j^{-1}(\mu_i - \mu_j) < 9\}. \quad (3)$$

The resulting relation is directed: a broad Gaussian may cover a nearby narrow primitive even when the reverse relation does not hold. This asymmetry is informative rather than incidental, since reciprocal support distinguishes primitives embedded within a locally compatible field from those that are only enclosed by larger neighbours. In contrast to fixed- k Euclidean neighbourhoods, the support relation follows the scale, shape, and orientation learned by each primitive and requires no separately chosen k or spatial radius.

Surface Variation. Centres associated with a local surface vary primarily along two tangent directions and only weakly along the surface normal. For each query with $|\mathcal{N}_i| \geq 4$, let $\bar{\mu}_i$ be the opacity-weighted centroid of the covering centres. Four centred samples are the minimum required for a full-rank covariance in $\mathbb{R}^3$; a smaller set forces the least eigenvalue to zero independently of the local surface structure. The weighted covariance is

$$\mathbf{L}_i = \frac{\sum_{j \in \mathcal{N}_i} \alpha_j (\mu_j - \bar{\mu}_i)(\mu_j - \bar{\mu}_i)^\top}{\sum_{j \in \mathcal{N}_i} \alpha_j}. \quad (4)$$

Let $\lambda_{i,1} \geq \lambda_{i,2} \geq \lambda_{i,3} \geq 0$ denote the eigenvalues of $\mathbf{L}_i$. The standard surface-variation feature [18, 27] is

$$s_i = \frac{\lambda_{i,3}}{\lambda_{i,1} + \lambda_{i,2} + \lambda_{i,3}} \in \left[0, \frac{1}{3}\right], \quad (5)$$

which is small for a planar neighbourhood and increases as its spatial distribution becomes less surface-like.

$\mathbf{L}_i$ is induced by the primitives covering g_i , so its variation characterises the coverers rather than the query in isolation. The value s_i is therefore attributed to every $g_j \in \mathcal{N}_i$. Let $\mathcal{I}_j = \{i : j \in \mathcal{N}_i, |\mathcal{N}_i| \geq 4\}$. Its aggregated surface variation is

$$v_j = \frac{1}{|\mathcal{I}_j|} \sum_{i \in \mathcal{I}_j} s_i. \quad (6)$$

This attribution prevents a displaced primitive from inheriting the apparent planarity of a nearby surface: its score instead reflects the collection of support neighbourhoods to which it contributes. The resulting variation is normalised as $\hat{v}_j = 3v_j \in [0, 1]$.

Mutual Coverage. Surface variation does not establish whether local support is reciprocal. A support relation between g_i and g_j is reciprocal when $j \in \mathcal{N}_i$ and $i \in \mathcal{N}_j$. A small displaced primitive may lie within several broad surface Gaussians and therefore inherit locally planar evidence, although its own support does not extend back to those Gaussians. Let

$$d_i = |\mathcal{N}_i|, \quad m_i = |\{j \in \mathcal{N}_i : i \in \mathcal{N}_j\}|$$

denote the number of primitives supporting g_i and the subset whose support is reciprocal, respectively. Their regularised ratio defines mutual coverage:

$$R_i = \frac{m_i}{d_i + 1} \in [0, 1). \quad (7)$$

High mutual coverage indicates that g_i participates in the same local support structure as its neighbours; low mutual coverage exposes a predominantly one-sided relation. The unit term prevents a single reciprocal relation from constituting complete evidence of embeddedness.

Primitive Confidence. Surface variation and mutual coverage encode complementary requirements for a coherent surface primitive: locally surface-like structure and reciprocal embedding. Their product defines primitive confidence:

$$C_i = (1 - \widehat{v}_i)R_i. \quad (8)$$

Multiplication treats the two signals as joint requirements rather than interchangeable evidence. Low C_i consequently identifies primitives that are locally non-surface-like, weakly embedded, or both. A primitive without sufficient surface evidence receives zero confidence.

3.3 Artifact Projection and Floater Artifact Score

Primitive confidence characterises the stored 3D representation, whereas the visible effect of a floater depends on the view. Projecting confidence with the same transmittance and alpha weights as Eq. (2) preserves the visibility, opacity, and occlusion structure of the rendered scene. Replacing colour by the primitive artifact evidence $1 - C_i$ and normalising by the accumulated contribution gives

$$H(p) = \frac{\sum_{i=1}^N T_i(p) \alpha_i(p) (1 - C_i)}{\sum_{i=1}^N T_i(p) \alpha_i(p)}. \quad (9)$$

$H(p)$ therefore measures the visibility-weighted artifact evidence among the primitives contributing to p , independent of the total accumulated opacity.

To separate representation quality from the amount of empty background in a particular camera framing, FAS is defined over the foreground domain $\mathcal{M}$ supported by the unpruned representation. This domain remains common to all variants of the same scene. A pixel in $\mathcal{M}$ that loses all support after pruning is assigned $H(p) = 1$, ensuring that removed content is not interpreted as an absence of geometric artifacts. FAS is the mean projected evidence over this common domain:

$$\text{FAS} = \frac{1}{|\mathcal{M}|} \sum_{p \in \mathcal{M}} H(p). \quad (10)$$

3.4 Confidence-Guided Pruning

The confidence C_i induces a geometry-guided ranking based on local surface structure and reciprocal embedding. This is inherently a relational criterion: a primitive is assessed through its position within the surrounding Gaussian field. Opacity and spatial extent describe a different property, namely the amount of scene content associated with the primitive, and are not directly represented by geometric confidence.

PointSplat supplies a complementary primitive-wise criterion [20] by combining opacity, which determines blending strength, with ellipsoid volume, which characterises spatial extent. For axis scales $(s_{x,i}, s_{y,i}, s_{z,i})$,

$$V_i = \frac{4\pi}{3} s_{x,i} s_{y,i} s_{z,i}, \quad P_i = \lambda z(\alpha_i) + (1 - \lambda) z(V_i). \quad (11)$$

Here, $z(\cdot)$ denotes scene-wise standardisation. We retain PointSplat's published setting $\lambda = 0.3$, corresponding to opacity and volume weights of 0.3 and 0.7. The resulting scores capture distinct evidence: P_i can favour a large or opaque primitive without regard to its local embedding, while C_i can favour a geometrically coherent primitive whose opacity and volume are moderate.

To place the two criteria on a common ordinal scale, let A_i and G_i denote the scene-wise percentile ranks of P_i and C_i , representing intrinsic importance and geometric coherence, respectively. Their harmonic mean defines the combined pruning score:

$$H_i^{\text{prune}} = \frac{2A_i G_i}{A_i + G_i}. \quad (12)$$

We take $H_i^{\text{prune}} = 0$ when $A_i = G_i = 0$. This formulation gives preference to primitives supported by both criteria, rather than allowing strong evidence under one criterion to dominate weak evidence under the other. The geometric audit remains defined by C_i ; the combined score serves only as the ordering used for pruning.

4 Experiments and Results

4.1 Experimental Setup

Dataset and setup. We evaluate on all seven Mip-NeRF 360 scenes used in this study: *bicycle*, *bonsai*, *counter*, *garden*, *kitchen*, *room*, and *stump* [1]. Every eighth camera in the sorted image sequence is reserved for evaluation, and 24 evenly spaced cameras from the remainder are used for training. Each 3DGS model is trained for 30,000 iterations with random seed zero.

Compared methods. We compare the unpruned 3DGS reference with four post-hoc selectors: the photometric sensitivity ranking of PUP [6], the published opacity-volume score of PointSplat [20], SplatAudit confidence, and the harmonic score in Eq. (12). PointSplat is evaluated as a score-only selector, without its learned transformer. Each ranking is applied to a common pretrained scene and followed by standard 3DGS refinement. We retain 20% and 10% of the primitives in one pass, followed by 5,000 and 10,000 refinement iterations, respectively, with densification disabled.

Metrics and implementation. Rendering quality is measured on the held-out cameras using PSNR, SSIM [26], and LPIPS [34]; FAS evaluates the projected geometric evidence defined in Eq. (10). LPIPS uses non-overlapping tiles of at most 1024×1024 pixels with area-weighted aggregation. FAS uses the foreground domain

Table 1: Per-scene held-out results at 100%, 20%, and 10% retention. Best and second-best compressed results are bold and underlined within each retention level.

Ret.	Method	Bonsai				Counter				Kitchen				Room			
		FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
100%	3DGS	0.922	22.861	0.840	0.345	0.933	23.227	0.820	0.338	0.937	23.167	0.825	0.272	0.896	23.879	0.821	0.347
20%	PUP	0.961	<u>22.613</u>	0.841	0.362	0.948	22.930	<u>0.821</u>	0.362	0.916	23.531	<u>0.839</u>	0.289	0.902	23.782	0.823	0.356
	PointSplat	0.946	23.316	<u>0.835</u>	<u>0.378</u>	0.948	<u>23.252</u>	0.817	<u>0.370</u>	0.936	<u>23.542</u>	0.827	0.316	0.894	24.403	0.832	<u>0.363</u>
	SplatAudit	0.863	22.046	0.809	0.415	0.892	23.148	0.812	0.405	0.868	23.361	0.824	0.332	0.820	<u>24.417</u>	0.828	0.382
	Hybrid	<u>0.879</u>	22.575	0.822	0.396	<u>0.908</u>	23.448	0.823	0.382	<u>0.886</u>	23.652	0.840	<u>0.300</u>	<u>0.848</u>	24.550	<u>0.829</u>	0.371
10%	PUP	0.976	21.261	<u>0.816</u>	0.396	0.950	22.222	0.804	0.389	0.943	22.754	<u>0.821</u>	0.326	0.928	23.612	0.825	0.367
	PointSplat	0.954	22.891	0.818	<u>0.403</u>	0.966	<u>23.051</u>	<u>0.806</u>	<u>0.391</u>	0.945	<u>23.175</u>	0.801	0.363	0.908	<u>24.382</u>	<u>0.829</u>	<u>0.373</u>
	SplatAudit	<u>0.897</u>	21.491	0.789	0.442	0.926	22.599	0.795	0.431	0.895	22.262	0.782	0.400	0.842	24.154	0.824	0.393
	Hybrid	0.891	<u>22.356</u>	0.810	0.414	<u>0.940</u>	23.169	0.811	0.402	<u>0.906</u>	23.639	0.828	<u>0.333</u>	<u>0.871</u>	24.506	0.830	0.377
Ret.	Method	Bicycle				Garden				Stump							
		FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓				
100%	3DGS	0.959	18.745	0.506	0.474	0.960	20.025	0.596	0.409	0.968	20.971	0.641	0.470				
20%	PUP	0.971	19.679	0.556	0.467	0.967	22.089	0.656	0.395	0.985	21.634	0.681	0.472				
	PointSplat	0.952	19.796	0.568	<u>0.479</u>	0.960	21.870	<u>0.642</u>	<u>0.420</u>	0.985	21.351	0.677	<u>0.495</u>				
	SplatAudit	0.907	20.035	<u>0.574</u>	0.504	0.934	21.841	0.628	0.445	0.929	21.356	0.675	0.522				
	Hybrid	<u>0.935</u>	<u>20.029</u>	0.576	0.486	<u>0.936</u>	<u>21.972</u>	0.641	0.423	<u>0.977</u>	<u>21.396</u>	<u>0.680</u>	0.502				
10%	PUP	0.980	19.386	0.550	0.480	0.969	<u>21.931</u>	0.640	0.424	0.980	21.473	<u>0.676</u>	0.489				
	PointSplat	0.962	19.785	<u>0.569</u>	<u>0.492</u>	0.967	21.876	0.630	<u>0.444</u>	0.989	21.140	0.671	<u>0.517</u>				
	SplatAudit	0.909	<u>19.811</u>	0.569	0.516	0.940	21.519	0.605	0.476	0.947	20.545	0.660	0.551				
	Hybrid	<u>0.936</u>	20.145	0.582	0.506	<u>0.945</u>	21.996	<u>0.631</u>	0.451	<u>0.978</u>	<u>21.342</u>	0.678	0.530				

and alpha reference of the corresponding unpruned model. All experiments run on a single NVIDIA RTX 4090.

4.2 One-Pass Pruning Results

Table 1 reports the per-scene held-out results at both retention levels.

Moderate compression. At 20% retention, SplatAudit obtains the lowest FAS in every scene, and the harmonic score ranks second among the compressed methods throughout. This ordering reflects the relational consistency encoded by the geometric confidence. Averaged across scenes, its FAS is 0.8876, compared with 0.9458 for PointSplat and 0.9500 for PUP. The consistency across indoor and outdoor scenes is more informative than any single reduction: a ranking derived only from local support relations repeatedly selects a field with stronger projected geometric confidence.

The appearance metrics expose the complementary role of primitive-wise importance. Pure SplatAudit has mean PSNR 22.315, SSIM 0.7356, and LPIPS 0.4293. Harmonic fusion changes these values to 22.517, 0.7444, and 0.4085, respectively, while retaining a substantially lower FAS than either PUP or PointSplat. The improvement is also consistent per scene: SSIM and LPIPS improve over SplatAudit in all seven cases, and PSNR improves in six. PUP remains strongest in LPIPS, indicating that photometric sensitivity is particularly effective at preserving perceptual features, whereas the hybrid attains the highest mean PSNR and nearly matches PUP in mean SSIM. The moderate setting therefore separates two effects that aggregate rendering quality alone would conflate: SplatAudit favours coherent local structure, while the opacity-volume term restores primitives whose visual contribution is not implied by that structure.

Aggressive compression. At 10% retention, the value of combining the two rankings becomes clearer. Relative to SplatAudit, the hybrid improves mean PSNR by 0.682 dB, mean SSIM by 0.0209, and mean LPIPS by 0.0281. These gains are larger than at 20%, although the mean FAS difference between the two methods narrows from 0.0224 to 0.0161. The fusion therefore becomes more effective, rather than less relevant, as the budget tightens: intrinsic importance prevents the geometric ranking from removing a small set of appearance-bearing primitives, while the geometric rank continues to reject much of the weakly embedded structure favoured by PUP and PointSplat.

This balance is also evident in Table 1. SplatAudit gives the lowest FAS on six scenes and the hybrid on the remaining scene; the hybrid remains below PUP and PointSplat in FAS throughout. At the same time, it achieves the highest PSNR and SSIM on five scenes each. PUP still gives the lowest LPIPS on every scene, so no single signal dominates all notions of quality. *Kitchen* illustrates the mechanism particularly clearly: adding intrinsic importance to SplatAudit raises PSNR from 22.262 to 23.639 and SSIM from 0.782 to 0.828, while FAS changes only from 0.895 to 0.906. The result supports the intended division of labour between the scores: geometric confidence defines the audit, whereas their harmonic combination is the more effective ordering for severe compression.

4.3 Geometric Ablation

We isolate the two confidence components at 20% retention and compare their representation-aligned support with a fixed- k Euclidean neighbourhood ($k = 50$). Table 2 reports the seven-scene mean under the same one-pass setting.

The fixed- k and surface-variation rows obtain similar FAS, PSNR, and SSIM; representation-aligned surface variation alone primarily improves LPIPS. The mutual-coverage-only score accounts for the largest change in geometric discrimination, reducing mean FAS from 0.9458 to 0.8880 while also improving PSNR and SSIM. The complete confidence product gives the best mean result for every reported metric: 0.8876 FAS, 22.315 PSNR, 0.7356 SSIM, and 0.4293 LPIPS. Its margin over mutual coverage is modest, indicating that reciprocal embedding supplies most of the discrimination, while surface variation contributes a complementary refinement rather than an equally dominant signal. The comparison also shows that replacing a fixed- k neighbourhood with anisotropic support is not, by itself, sufficient. Its main value is to form the directed relation on which mutual coverage operates; the gain arises from testing reciprocity within that representation-aligned relation.

Table 2: Seven-scene average geometric ablation at 20% retention.

Variant	FAS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Fixed- k ($k = 50$)	0.9478	21.914	0.7300	0.4339
Surface variation only	0.9458	21.804	0.7297	<u>0.4297</u>
Mutual coverage only	<u>0.8880</u>	<u>22.295</u>	<u>0.7347</u>	0.4307
SplatAudit	0.8876	22.315	0.7356	0.4293

5 Discussion

SplatAudit is intentionally a geometric audit. Its confidence is derived from the position, extent, and opacity of the stored primitives, leaving their view-dependent colour unused. It therefore asks whether a Gaussian is coherently embedded in the learned field, whereas PUP and PointSplat estimate primitive importance from photometric sensitivity or intrinsic attributes. A primitive may satisfy one criterion without satisfying the other, which explains why geometry-guided pruning and importance-based rankings favour different subsets of the same scene. The harmonic score does not replace this distinction; it uses it. SplatAudit confidence retains its role as the audit, while fusion with intrinsic importance provides a more balanced ordering when the objective is compression.

FAS inherits the same geometric scope. By projecting primitive confidence through the 3DGS compositor, it localises where weakly embedded structure contributes to a particular view, rather than introducing a separate notion of image quality. It is therefore most informative when read alongside PSNR, SSIM, and LPIPS: the former examines the internal organisation of the representation, while the latter describe agreement in rendered appearance. The current study does not connect this internal evidence to reference geometry or independently annotated artifacts. Such evidence would make it possible to study how support inconsistency relates to physical displacement in the reconstructed scene.

The representation-aligned neighbourhood is also the main source of computational cost. Exact support relations must be established before either surface variation or mutual coverage can be evaluated. Table 3 reports the corresponding scoring cost across the seven scenes.

Table 3: SplatAudit scoring cost on the sparse-24 scenes. N denotes input primitives and $\bar{d}$ mean support degree.

Scene	N (M)	$\bar{d}$	Neigh. (s)	Total (s)
Bicycle	3.336	50.2	128.9	308.3
Bonsai	0.723	10.1	30.0	37.6
Counter	0.885	20.4	34.5	39.8
Garden	2.706	48.2	108.4	150.7
Kitchen	1.032	21.0	72.1	92.6
Room	1.200	36.1	84.4	96.8
Stump	1.806	10.0	33.0	39.4
Mean	1.670	28.0	70.2	109.3

Runtime is dominated by neighbourhood construction, whose cost depends on primitive count and graph density. Hierarchical or streamed construction offers a direct path towards scaling the audit.

The experiments focus on one-pass pruning of sparse-view Mip-NeRF 360 scenes. In this setting, limited observations leave the Gaussian field weakly constrained and make geometric inconsistency especially relevant. The same focus also limits the conclusions: the behaviour of the audit under dense observations, other scene domains, and different 3DGS parameterisations remains open, while the timing measurements reflect a single implementation and hardware setting. Support relations inherit the scales learned by the underlying model, so broadly scaled or poorly calibrated primitives may alter the neighbourhood before surface variation and mutual coverage are evaluated. Beyond wider evaluation, the confidence could be used for targeted refinement rather than binary removal, or introduced during training as a signal for detecting the emergence of weakly supported geometry.

6 Conclusion

SplatAudit shows that a pretrained 3DGS scene can be examined as geometric data rather than judged only through its rendered images. Representation-aligned support, surface variation, and mutual coverage provide a primitive-level confidence derived from the Gaussian field itself. Its projection produces view-dependent heatmaps and FAS, while the same confidence supplies a geometry-guided ordering for one-pass pruning. Combining that ordering with PointSplat importance yields a pruning criterion that balances geometric coherence with rendering quality under a restrictive primitive budget.

Across seven sparse-view scenes, the results reveal that geometric coherence and rendering importance describe different aspects of a learned representation. Their complementarity is central to the method: SplatAudit exposes structure that image metrics do not directly inspect, while the harmonic extension reconnects this audit to the practical demands of compression. More broadly, explicit learned representations make it possible to evaluate not only what a model renders, but also the structure it has learned to store.

References

- [1] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. 2022. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. *CVPR* (2022).
- [2] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 2000. LOF: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data* (Dallas, Texas, USA) (SIGMOD '00). Association for Computing Machinery, New York, NY, USA, 93–104. doi:10.1145/342009.335388
- [3] Zhiwen Fan, Kevin Wang, Kairun Wen, Zehao Zhu, Dejia Xu, and Zhangyang Wang. 2024. LightGaussian: Unbounded 3D Gaussian Compression with 15x Reduction and 200+ FPS. In *Advances in Neural Information Processing Systems*, Vol. 37. doi:10.52202/079017-4447
- [4] Guangchi Fang and Bing Wang. 2024. Mini-Splatting: Representing Scenes with a Constrained Number of Gaussians. In *Computer Vision – ECCV 2024 (Lecture Notes in Computer Science, Vol. 15135)*. 165–181. doi:10.1007/978-3-031-72980-5_10
- [5] Antoine Guédon and Vincent Lepetit. 2024. SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering. *CVPR* (2024).
- [6] Alex Hanson, Allen Tu, Vasu Singla, Mayuka Jayawardhana, Matthias Zwicker, and Tom Goldstein. 2025. PUP 3D-GS: Principled Uncertainty Pruning for 3D Gaussian Splatting. In *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*. 5949–5958. <https://pup3dgs.github.io/>
- [7] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2024. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. In *ACM SIGGRAPH 2024 Conference Papers*. doi:10.1145/3641519.3657428
- [8] Miriam Jäger, Markus Hillemann, and Boris Jutzi. 2025. FeatureGS: Eigenvalue-feature optimization in 3D Gaussian Splatting for geometrically accurate and artifact-reduced reconstruction. *ISPRS Open Journal of Photogrammetry and Remote Sensing* 17 (2025), 100100. doi:10.1016/j.ophoto.2025.100100
- [9] Miriam Jäger and Boris Jutzi. 2026. EntON: Eigenentropy-Optimized Neighborhood Densification in 3D Gaussian Splatting. arXiv:2603.06216 [cs.CV] <https://arxiv.org/abs/2603.06216>
- [10] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics* 42, 4 (July 2023). <https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/>
- [11] Joo Chan Lee, Daniel Rho, Xiangyu Sun, Jong Hwan Ko, and Eunbyung Park. 2024. Compact 3D Gaussian Representation for Radiance Field. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 21719–21728. doi:10.1109/CVPR52733.2024.02052
- [12] Hantang Li, Qiang Zhu, Xiangdong Meng, Debin Zhao, and Xiaopeng Fan. 2026. DOC-GS: Dual-Domain Observation and Calibration for Reliable Sparse-View Gaussian Splatting. arXiv:2604.06739 [cs.CV] <https://arxiv.org/abs/2604.06739>
- [13] Jiahe Li, Jiawei Zhang, Xiao Bai, Jin Zheng, Xin Ning, Jun Zhou, and Lin Gu. 2024. DNGaussian: Optimizing Sparse-View 3D Gaussian Radiance Fields with Global-Local Depth Normalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 20775–20785. https://openaccess.thecvf.com/content/CVPR2024/html/Li_DNGaussian_Optimizing_Sparse-View_3D_Gaussian_Radiance_Fields_with_Global-Local_Depth_CVPR_2024_paper.html
- [14] Zheng Li, Bingxu Xie, Chao Chu, Weiqing Li, and Zhiyong Su. 2025. No-reference geometry quality assessment for colorless point clouds via list-wise rank learning. *Computers & Graphics* 127 (April 2025), 104176. doi:10.1016/j.cag.2025.104176
- [15] Tao Lu, Mulin Yu, Linning Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. 2024. Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 20654–20664. https://openaccess.thecvf.com/content/CVPR2024/html/Lu_Scaffold-GS_Structured_3D_Gaussians_for_View-Adaptive_Rendering_CVPR_2024_paper.html
- [16] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *ECCV*.
- [17] Hyunwoo Park, Gun Ryu, and Wonjun Kim. 2025. Dropgaussian: Structural regularization for sparse-view gaussian splatting. In *Proceedings of the Computer Vision and Pattern Recognition Conference*. 21600–21609.
- [18] Mark Pauly, Markus Gross, and Leif P. Kobbelt. 2002. Efficient Simplification of Point-Sampled Surfaces. In *Proceedings Visualization 2002*. IEEE Computer Society, 163–170. <https://www.graphics.rwth-aachen.de/publication/03122/>
- [19] Qiang Qu, Yiran Shen, Yuk Ying Chung, Weidong Cai, Xiaoming Chen, and Tongliang Liu. 2025. NVS-SQA: Exploring Self-Supervised Quality Representation Learning for Neurally Synthesized Scenes without References. *arXiv preprint arXiv:2501.06488* (2025).
- [20] Anh Thuan Tran and Jana Kosecka. 2026. PointSplat: Efficient Geometry-Driven Pruning and Transformer Refinement for 3D Gaussian Splatting. Accepted to CVPRW 2026 (3DMV). arXiv:2604.09903 [cs.CV] doi:10.48550/arXiv.2604.09903
- [21] Matias Turkulainen, Xuqian Ren, Iaroslav Melekhov, Otto Seiskari, Esa Rahtu, and Juho Kannala. 2025. DN-Splatter: Depth and Normal Priors for Gaussian Splatting and Meshing. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2421–2431. https://openaccess.thecvf.com/content/WACV2025/html/Turkulainen_DN-Splatter_Depth_and_Normal_Priors_for_Gaussian_Splatting_and_Meshing_WACV_2025_paper.html
- [22] Zhaolin Wan, Yining Diao, Jingqi Xu, Hao Wang, Zhiyang Li, Xiaopeng Fan, Wangmeng Zuo, and Debin Zhao. 2026. Perceptual Quality Assessment of 3D Gaussian Splatting: A Subjective Dataset and Prediction Metric. *Proceedings of the AAAI Conference on Artificial Intelligence* 40, 12 (2026), 9657–9665. doi:10.1609/aaai.v40i12.37928
- [23] Haoran Wang, Guoxi Huang, Fan Zhang, David Bull, and Nantheera Anantrasirichai. 2026. Prune Wisely, Reconstruct Sharply: Compact 3D Gaussian Splatting via Adaptive Pruning and Difference-of-Gaussian Primitives. arXiv:2602.24136 [cs.CV] <https://arxiv.org/abs/2602.24136>
- [24] Luchao Wang, Qian Ren, Kaimin Liao, Hua Wang, Zhi Chen, and Yaohua Tang. 2025. StableGS: A Floater-Free Framework for 3D Gaussian Splatting. arXiv:2503.18458 [cs.CV] <https://arxiv.org/abs/2503.18458>
- [25] Xinliang Wang, Yifeng Shi, and Zhenyu Wu. 2026. ArtifactWorld: Scaling 3D Gaussian Splatting Artifact Restoration via Video Generation Models. arXiv:2604.12251 [cs.CV] <https://arxiv.org/abs/2604.12251>
- [26] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* 13, 4 (April 2004), 600–612. doi:10.1109/TIP.2003.819861
- [27] Martin Weinmann, Boris Jutzi, Stefan Hinz, and Clément Mallet. 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* 105 (2015), 286–304. doi:10.1016/j.isprsjprs.2015.01.016
- [28] Qianyi Wu, Jianmin Zheng, and Jianfei Cai. 2024. Surface Reconstruction from 3D Gaussian Splatting via Local Structural Hints. In *European Conference on Computer Vision*.
- [29] Shuzhao Xie, Weixiang Zhang, Chen Tang, Yunpeng Bai, Rongwei Lu, Shijia Ge, and Zhi Wang. 2024. MesonGS: Post-Training Compression of 3D Gaussians via Efficient Attribute Transformation. In *Computer Vision – ECCV 2024*. https://www.ecva.net/papers/eccv_2024/papers_ECCV/html/4927_ECCV_2024_paper.php
- [30] Haolin Xiong, Sairisheek Muttukuru, Rishi Upadhyay, Pradyumna Chari, and Achuta Kadambi. 2023. SparseGS: Real-Time 360° Sparse View Synthesis using Gaussian Splatting. *Arxiv* (2023).
- [31] Sooyeon Yang, Cheyul Im, Jee Won Lee, and Jongseong Brad Choi. 2026. TIDI-GS: Floater Suppression in 3D Gaussian Splatting for Enhanced Indoor Scene Fidelity. arXiv:2601.09291 [cs.GR] <https://arxiv.org/abs/2601.09291>
- [32] Zehao Yu, Torsten Sattler, and Andreas Geiger. 2024. Gaussian Opacity Fields: Efficient Adaptive Surface Reconstruction in Unbounded Scenes. *ACM Transactions on Graphics* (2024).
- [33] Daochen Zha, Zaid Pervaiz Bhat, Kwei-Herng Lai, Fan Yang, Zhimeng Jiang, Shaochen Zhong, and Xia Hu. 2023. Data-centric Artificial Intelligence: A Survey. arXiv:2303.10158 [cs.LG] <https://arxiv.org/abs/2303.10158>
- [34] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. 2018. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In *CVPR*.
- [35] Zheng Zhang, Wenbo Hu, Yixing Lao, Tong He, and Hengshuang Zhao. 2024. Pixel-GS: Density Control with Pixel-aware Gradient for 3D Gaussian Splatting. In *ECCV*.
- [36] Zehao Zhu, Zhiwen Fan, Yifan Jiang, and Zhangyang Wang. 2023. FSGS: Real-Time Few-Shot View Synthesis using Gaussian Splatting. arXiv:2312.00451 [cs.CV]

A Reproducibility Supplement

A.1 Reproducibility Details

Configuration. Code and the complete configuration are available at <https://github.com/Gearoid522/SplatAudit>. On each Mip-NeRF 360 scene, every eighth sorted camera is held out and 24 evenly spaced remaining cameras are used for training. The shared 3DGS model is trained for 30,000 iterations with seed zero. The resulting split and pretrained scene are fixed across all selectors.

Pruning and evaluation. One-pass 20% and 10% selections are refined for 5,000 and 10,000 iterations, respectively, without densification, with matched retention counts across methods. PointSplat uses $\lambda = 0.3$, and the hybrid uses Eq. (12). Held-out PSNR, SSIM, LPIPS, and FAS follow Section 4.1. Experiments use a single NVIDIA RTX 4090.

A.2 Qualitative Results



Figure 3: Held-out views after one-pass 10% retention and 10,000 refinement iterations. Each inset shows the corresponding FAS map for the reconstructed scene.