Do Flat Minima Improve Sparse Novel View Synthesis?

Youngsik Yun, Dongjun Gu, and Youngjung Uh*

Yonsei University, Seoul, Korea
{bbangsik, djku1020, yj.uh}@yonsei.ac.kr

Abstract. Despite the success of recent novel view synthesis methods, they tend to struggle in sparse-view settings. This poor generalization to unseen viewpoints is an inherent challenge when training with limited data. To address this, we investigate the relationship between loss sharpness and generalization in novel view synthesis—an underexplored direction. Interestingly, while pursuing flatter minima is widely known to improve generalization in deep learning, reducing loss sharpness is not always beneficial in novel view synthesis. We demonstrate that this difference arises because high-detail regions inherently require a sharp loss landscape for accurate reconstruction, whereas low-detail regions benefit from a flat loss landscape for improving generalization. Based on this insight, we introduce structure-aware sharpness, defined within structure-adaptive neighborhoods, and propose to adaptively adjust the sharpness regularization weight according to the local image structure. This strategy encourages flatter minima for generalization while preserving the loss sharpness necessary to reconstruct fine details. Across various datasets and configurations, our strategy consistently improves a wide range of baselines. Code is available at https://bbangsik13.github.io/FASR.

Keywords: Sparse novel view synthesis · Loss sharpness · Generalization

1 Introduction

Novel view synthesis from multi-view images has been a long-standing problem of interest. While Neural Radiance Fields (NeRFs) [45] achieve high-quality novel view synthesis, their slow rendering has motivated a shift towards 3D Gaussian Splatting (3DGS) [27]. However, it requires densely captured input views, which are costly to obtain. In sparse-view settings, models often produce unresolved details and floating artifacts in novel viewpoints while correctly reconstructing training views. This overfitting is due to limited data, leading to poor generalization. To address this challenging problem, previous approaches have adopted various strategies, such as integrating geometric priors [9, 10] and leveraging

* Corresponding author

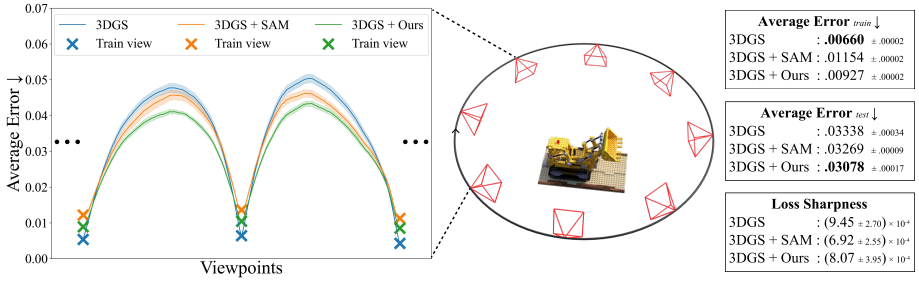

Fig. 1: Overview. While 3DGS exhibits overfitting, our method improves its generalization, achieving lower test errors. Interestingly, loss sharpness and generalization are not strictly correlated in novel view synthesis—although Sharpness-Aware Minimization (SAM) [13], a representative flat minima optimization method, yields the lowest loss sharpness [42, 61], it results in a higher test error than ours—challenging the common belief that flatter minima generalize better. We evaluate using eight training views and 800 interpolated test views generated from the `lego` scene of Blender synthetic dataset [45]. Plots are means and standard deviations over ten runs.

dense correspondence prediction models [18, 22]. Although these methods show effectiveness, research on improving the optimization procedure of such scene representations itself has not been thoroughly explored.

Rather than introducing additional model priors, we investigate the role of local loss sharpness in novel view synthesis. In deep learning, flatter minima[1] are widely known to improve generalization [20, 28]. However, unlike the domain of previous work, achieving flatter minima *does not always guarantee* better generalization in novel view synthesis (Fig. 1). In Fig. 2, we demonstrate that this discrepancy arises because the importance of loss sharpness for accurate reconstruction varies across regions: model parameters representing high-detail regions (e.g., edges) inherently induce a sharp loss landscape, whereas those corresponding to low-detail regions favor flatter minima. In this sense, pursuing flat minima for all regions is problematic; sharp minima in high-detail areas are desirable for accurate reconstruction, and flat minima in low-detail regions are preferable for better generalization. Specifically, it will over-penalize high-detail regions while under-penalize low-detail regions, failing to sufficiently improve generalization.

To this end, we propose Structure-Aware Sharpness Regularization, an optimization strategy that calculates and penalizes loss sharpness based on the local image structure. Specifically, we define structure-aware sharpness and adaptively adjust the sharpness regularization: in high-detail regions, we both reduce the regularization weight and the neighborhood radius used for sharpness estimation, while in low-detail regions, we increase them. This strategy encourages a flatter loss landscape while *retaining the loss sharpness necessary for fine details*. For example, our strategy efficiently mitigates floating artifacts that inher-

[1] We refer to flatter *local* minima as flatter minima for brevity.

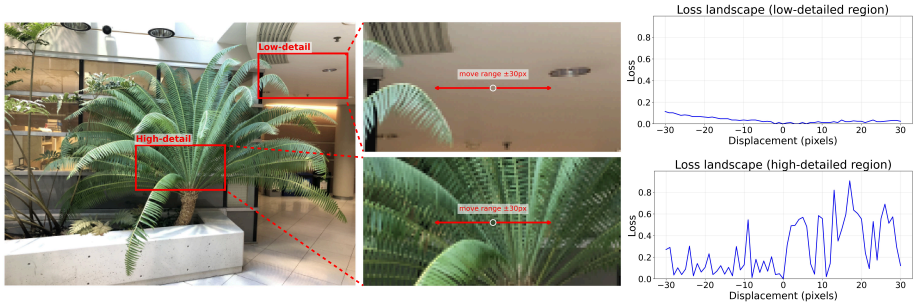

Fig. 2: Problem setting. The desirable degree of loss sharpness varies across the local structure of the input images. We visualize the loss landscape by perturbing a perfectly fitted pixel in high-detail and low-detail regions, respectively.

ently induce sharp minima, because their loss changes dramatically even with a slight perturbation[2]. Our strategy guides these unexpected sharp minima to flat minima, successfully narrowing the generalization gap, thus suppressing floaters without penalizing high-detail regions. Crucially, we demonstrate that our strategy serves as a general principle for novel view synthesis, yielding complementary improvements whether applied to a wide range of 3DGS-based frameworks, implicit NeRF, various flat minima optimization methods, and scene dynamics.

Our core contributions are as follows:

- **Fundamental insight.** To the best of our knowledge, we present the first fundamental investigation into the relationship between loss landscape and generalization in novel view synthesis. Notably, we reveal that flat minima across all Gaussian parameters are suboptimal because the optimal loss sharpness varies with the local structure of the input images.
- **Novel optimization strategy.** Based on this insight, we propose a strategy that introduces structure-aware sharpness and adaptively adjusts the regularization weight to encourage flatness for generalization while preserving the necessary loss sharpness for fine-detail reconstruction.
- **Generality.** We demonstrate that our strategy serves as a broadly applicable optimization principle rather than a method-specific adaptation, consistently improving various reconstruction pipelines.

2 Related Work

2.1 Sparse Novel View Synthesis

Reconstructing scenes from highly sparse inputs remains a fundamental challenge, as limited supervision often leads models to overfit training views, resulting in insufficient generalization in unseen views. Early efforts sought to extend

[2] Camera perturbations can be interpreted as model parameter perturbations, e.g., parallel movement of 3D Gaussians is equivalent to moving the camera.

neural radiance fields (NeRFs) [45] by incorporating additional regularizations or auxiliary cues. For example, these studies [10, 47, 58, 64] introduce geometry- or depth-based constraints and frequency regularization to alleviate the underconstrained nature of sparse-view optimization. Although these approaches demonstrate effectiveness, they inherit the slow training and rendering processes of NeRF.

Recent studies have shifted to use 3D Gaussian Splatting (3DGS) [27], aiming for real-time rendering efficiency. These studies [9, 22, 34, 62, 70, 72] similarly leverage external priors such as depth [6, 49], correspondence [33], or flow [53], while others focuses on ensemble-like regularization [48, 67, 69]. Sharing a key insight with our method, Sparfels [23] aims to minimize a worst-case loss for robustness; however, it does so by freezing the Gaussian means and approximating the objective with an upper bound, which simplifies to a color variance regularization along each ray. Although this proxy effectively improves details, it may overlook color-matched floaters. In contrast, our method directly optimizes the worst-case loss with respect to all Gaussian parameters, interpreting it as a sharpness regularization.

Importantly, our optimization strategy is complementary: rather than altering the representation or adding priors, we fundamentally guide novel view synthesis methods to converge on a general solution in under-constrained sparse supervision.

2.2 Flatness and Generalization

The relationship between the flatness of the loss landscape and model generalization has been extensively investigated in prior research. Keskar *et al.* [28] shows that converging to sharp minima leads to poor generalization, while Jiang *et al.* [26] identifies that loss sharpness is the most correlated indicator of generalization. Subsequent work [7,20] demonstrated that averaging parameters along training trajectories can lead to flatter minima with better generalization.

Beyond averaging strategies, Sharpness-Aware Minimization (SAM) [13] explicitly estimates and reduces sharpness, inspiring subsequent work [4, 37, 46, 55] to analyze and improve upon its formulation. Moreover, some works have analyzed the accuracy of estimated sharpness as a measure of generalization. They show that sharpness changes with parameter rescalings [11], and the appropriate sharpness estimation differs across different training setups and tasks [2], which hinders the correlation between sharpness and generalization. Consequently, recent work introduces normalization and invariant formulations to appropriately estimate sharpness, enabling a more reliable link between loss sharpness and generalization [17, 21, 30, 31, 57]. On the other hand, some works report counterexamples where sharper models generalize well, indicating that flatter minima are not always the optimal strategy for generalization [3, 11, 60].

Building on this perspective, we revisit the loss landscape sharpness in the context of novel view synthesis. We hypothesize that the optimal loss sharpness is not uniform but varies with the local detail of the input images. Furthermore, the radius for estimating the local loss sharpness should also vary. Therefore,

Algorithm 1 Overview of our proposed algorithm

Input: Multi-view images $\tilde{I}^v$, where camera $v \in \mathcal{V}$
Output: Optimized $\mathcal{G} = (\boldsymbol{\mu}, \mathbf{q}, \mathbf{s}, \sigma, \mathbf{Y}^{\mathrm{DC}}, \mathbf{Y}^{\mathrm{AC}})$

$\Gamma^v \leftarrow \mathrm{ToleranceMap}(\tilde{I}^v)$ // *Quantify local image structure (Sec. 3.2)*
while $\mathcal{G}$ not converged **do**
  $L \leftarrow \mathrm{Loss}(\mathrm{Render}(\mathcal{G}, v), \tilde{I}^v)$ // *Empirical loss*
  for all $\boldsymbol{\theta} \in \mathcal{G}$ **do**
    $\hat{\boldsymbol{\theta}} \leftarrow \mathrm{StructureAwarePerturb}(\boldsymbol{\theta}, L, \Gamma)$ // *Calculate perturbation (Sec. 3.3)*
  end for
  $\hat{\mathcal{G}} \leftarrow (\hat{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \mathcal{G})$ // *Local maximum*
  $\hat{L} \leftarrow \mathrm{Loss}(\mathrm{Render}(\hat{\mathcal{G}}, v), \tilde{I}^v)$ // *Worst-case loss*
  $\mathcal{L} \leftarrow \mathrm{StructureAwareWeight}(L, \hat{L}, \Gamma)$ // *Scale regularization weight (Sec. 3.4)*
  $\mathcal{G} \leftarrow \mathcal{G} - \mathrm{Adam}(\nabla_{\hat{\mathcal{G}}} \mathcal{L})$ // *Gradient descent at original parameter*
end while

instead of pursuing flat minima across all model parameters, we introduce an optimization strategy that appropriately regularizes loss sharpness for the reconstruction task, allowing sharp minima where they benefit fine details.

2.3 Reconstructing with Random Noise

Adapting random noise into radiance field training has been introduced for two primary purposes.

Unlike our approach, which aims to improve generalization, some studies employ perturbation for uncertainty quantification rather than relying on a deterministic approach. Stochastic or Bayesian formulations of NeRFs have explicitly modeled radiance or density distributions [32, 52]. Subsequent work perturbs trained models to estimate epistemic uncertainty [16]. Similar ideas have been extended to 3DGS, where Gaussian parameters are sampled from learned distributions to render both images and calibrated uncertainty maps [1].

Some works [8,14,29,40,50] achieve robustness by injecting random noise into Gaussian parameters or query locations, which can be interpreted as randomly choosing parameters in a neighborhood radius. This random perturbation serves as an inefficient proxy for finding the worst-case loss, and thus may flatten the loss landscape. In contrast, our work fundamentally investigates the direct link between the loss landscape and generalization, allowing us to rethink the success of these prior works.

3 Method

To implement our fundamental insight—optimal loss sharpness for accurate reconstruction correlates with the local structure of the input images—into practice, we apply our strategy within Sharpness-Aware Minimization (SAM) [13] and 3D Gaussian Splatting (3DGS) [27], as representative choices of flat-minima

optimization and scene representation, respectively. First, we provide preliminaries on SAM and 3DGS (Sec. 3.1). Then, we introduce the geometric tolerance map to quantify local image structures (Sec. 3.2). Finally, based on this image structure, we define the structure-aware sharpness (Sec. 3.3) and adaptively adjust the regularization weight (Sec. 3.4)[3]. Algorithm 1 summarizes our method.

3.1 Preliminaries

Sharpness-Aware Minimization (SAM) [13] is an optimization method that improves generalization by encouraging solutions to lie in flat regions of the loss landscape. Along with minimizing the empirical loss at a parameter point $\boldsymbol{w}$, SAM estimates the loss sharpness within a neighborhood of radius ρ. The optimization then jointly minimizes both the empirical loss and this estimated sharpness:

$$L^{\text{SAM}} \triangleq \underbrace{\max_{\|\boldsymbol{\epsilon}\|_2 \leq \rho} L(\boldsymbol{w}+\boldsymbol{\epsilon}) - L(\boldsymbol{w})}_{\text{loss sharpness}} + \underbrace{L(\boldsymbol{w})}_{\text{empirical loss}} = \underbrace{\max_{\|\boldsymbol{\epsilon}\|_2 \leq \rho} L(\boldsymbol{w}+\boldsymbol{\epsilon})}_{\text{worst-case loss}}, \tag{1}$$

which is equivalent to the worst-case loss.

In practice, it is approximated via first-order Taylor expansion for the loss function around the current parameters, perturbing the parameters in the gradient direction with magnitude ρ:

$$\hat{\boldsymbol{\epsilon}}(\boldsymbol{w}) \triangleq \arg\max_{\|\boldsymbol{\epsilon}\|_2 \leq \rho} L(\boldsymbol{w}+\boldsymbol{\epsilon}) \approx \rho \cdot \frac{\nabla_{\boldsymbol{w}} L(\boldsymbol{w})}{\|\nabla_{\boldsymbol{w}} L(\boldsymbol{w})\|_2}. \tag{2}$$

This method improves generalization by tightening the generalization bound based on sharpness derived from the PAC-Bayesian framework [12, 43, 51]. In practice, ρ is a hyperparameter. See Foret *et al.* [13] for details.

3D Gaussian Splatting (3DGS) [27] represents a scene as a set of 3D Gaussians $\mathcal{G} = \{\mathcal{G}_i\}_{i=1}^{N}$. Each Gaussian $\mathcal{G}_i$ is parameterized by its mean $\boldsymbol{\mu}_i$, rotation $\mathbf{q}_i$, scale $\mathbf{s}_i$, opacity σ_i, and spherical harmonics (SH) coefficients $\mathbf{Y}_i^{\text{DC}}, \mathbf{Y}_i^{\text{AC}}$ that model view-dependent color. Given a camera view v, the renderer projects each Gaussian onto the image plane and produces a rendered image $I^v = \text{Render}(\mathcal{G}, v)$. See Kerbl *et al.* [27] for details.

3.2 Geometric Tolerance Map

To quantify local structure of input images, we introduce a geometric tolerance map $\Gamma^v(\mathbf{p})$ defined over each training image $\tilde{I}^v$ with view v, where $\mathbf{p}$ denotes a 2D image coordinate. The map encodes the distance from each pixel to the

[3] We apply this strategy to the mean $\boldsymbol{\mu}_i$, rotation $\mathbf{q}_i$, and scale $\mathbf{s}_i$, which are geometric attributes of the Gaussian.

nearest high-detail structure, such as fine-scale textures. Let $\mathcal{H}^v$ denote the set of pixels with large image gradient in view v. Then, the map is defined as:

$$\Gamma^v(\mathbf{p}) = \min_{\mathbf{h} \in \mathcal{H}^v} \|\mathbf{p} - \mathbf{h}\|_2.$$

As a result, pixels located in low-detail structure exhibit large distance values, whereas pixels near or within high-detail structure have small values.

Near high-detail structure, even small geometric changes of projected primitives can produce noticeable variations in the rendered image. Conversely, in low-detail structure, similar changes tend to have a limited visual impact. Therefore, we interpret this distance as a *geometric tolerance* for the primitive projected at location $\mathbf{p}$, with respect to rendering sensitivity.

We describe the implementation of the geometric tolerance map in the Appendix.

3.3 Structure-Aware Sharpness

We adapt the loss sharpness of each Gaussian based on the geometric tolerance defined in Sec. 3.2. To do this, we begin by computing the gradient on the per-Gaussian attribute $\boldsymbol{\theta}_i \in \mathcal{G}_i$ separately rather than the entire parameter set and use it to compute the perturbation independently (Eq. (2)).

Then, for each Gaussian, we query the value from the geometric tolerance map at the projected center coordinate μ_i' on the 2D camera plane, $\boldsymbol{\gamma}_i = \Gamma^v(\mu_i')$. Since this geometric tolerance is defined in the 2D image space, we convert it into a 3D space using perspective geometry. Under perspective projection, a 2D displacement of $\boldsymbol{\gamma}_i$ corresponds to a 3D displacement scaled by $\frac{d_i}{f}$, where d_i is the depth of the corresponding Gaussian and f is the focal length. We interpret this 3D geometric tolerance as an admissible perturbation of the 3D parameters of the Gaussian. Accordingly, the structure-aware sharpness for the Gaussian attribute $\boldsymbol{\theta}_i$ is defined as:

$$\max_{\|\boldsymbol{\epsilon}_i\|_2 \leq \boldsymbol{\gamma}_i \frac{d_i}{f} \rho_{\boldsymbol{\theta}}} L(\boldsymbol{\theta}_i + \boldsymbol{\epsilon}_i) - L(\boldsymbol{\theta}_i).$$

In the SAM implementation, the perturbation Algorithm 1.5 becomes:

$$\hat{\boldsymbol{\theta}}_i \leftarrow \boldsymbol{\theta}_i + \boldsymbol{\gamma}_i \frac{d_i}{f} \rho_{\boldsymbol{\theta}} \frac{\nabla_{\boldsymbol{\theta}_i} L}{\|\nabla_{\boldsymbol{\theta}_i} L\|_2}.$$

This structure-aware formulation resolves the imbalance in the perturbation magnitude. Specifically, it induces slight perturbations in high-detail regions to prevent overshoot of the intended local maximum, while applying stronger perturbations in low-detail regions to effectively ascend toward the local maximum.

We provide a deeper analysis of this derivation in the Appendix.

Table 1: Quantitative comparison. Our method consistently improves all baselines across both LLFF and Mip-NeRF360, achieving better results on all evaluation metrics.

	LLFF (3 views)			MipNeRF-360 (12 views)		
Method	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
3DGS [ACM TOG'23]	19.810 ± .339	.6790 ± .0078	.2145 ± .0065	18.903 ± .179	.5499 ± .0036	.3734 ± .0042
+ Ours	**20.783** ± .300	**.7197** ± .0032	**.1965** ± .0034	**19.303** ± .185	**.5622** ± .0051	**.3552** ± .0047
CoR-GS [ECCV'24]	20.185 ± .142	.7015 ± .0040	.2029 ± .0035	19.515 ± .243	.5733 ± .0049	.3741 ± .0066
+ Ours	**20.862** ± .154	**.7283** ± .0042	**.1932** ± .0030	**19.805** ± .233	**.5833** ± .0058	**.3681** ± .0069
DropGaussian [CVPR'25]	20.461 ± .212	.7070 ± .0045	.2064 ± .0047	19.514 ± .199	.5722 ± .0042	.3657 ± .0036
+ Ours	**20.853** ± .227	**.7295** ± .0045	**.1969** ± .0040	**19.625** ± .254	**.5750** ± .0053	**.3627** ± .0051
NexusGS [CVPR'25]	21.048 ± .049	.7382 ± .0008	.1776 ± .0009	18.506 ± .098	.5222 ± .0031	.3587 ± .0021
+ Ours	**21.348** ± .078	**.7511** ± .0012	**.1714** ± .0011	**18.736** ± .103	**.5316** ± .0034	**.3522** ± .0024
SE-GS [ICCV'25]	20.725 ± .217	.7203 ± .0049	.1861 ± .0058	19.931 ± .288	.5930 ± .0063	.3702 ± .0054
+ Ours	**21.141** ± .223	**.7403** ± .0042	**.1803** ± .0038	**20.135** ± .232	**.5960** ± .0059	**.3644** ± .0052

3.4 Structure-Aware Regularization Weighting

In addition to structure-aware sharpness, we also modulate the weight of sharpness regularization. As described in Eq. (1), SAM can be interpreted as minimizing empirical loss together with a sharpness term. To control the relative contribution of these two terms, we introduce a structure-aware weighting scheme.

Inspired by weighted variants of SAM [65], we consider a weighted objective of the following form. Therefore, Algorithm 1.9 becomes:

$$\mathcal{L} = L + \frac{\bar{\gamma}_i}{1-\bar{\gamma}_i}(\hat{L} - L) = \frac{1-2\bar{\gamma}_i}{1-\bar{\gamma}_i}L + \frac{\bar{\gamma}_i}{1-\bar{\gamma}_i}\hat{L}.$$

Let the regularization weight $\bar{\gamma}_i = 0.95\boldsymbol{\gamma}_i/\boldsymbol{\gamma}_{\max}$, where $\boldsymbol{\gamma}_{\max}$ is the maximum distance, and 0.95 is determined empirically. When the regularization weight $\bar{\gamma}i$ is close to 0, the sharpness term vanishes, and the objective reduces to minimizing only the empirical loss. In contrast, when it is close to 0.95, the sharpness term is emphasized.

As a result, the sharpness penalty is reduced in high-detail regions, preserving necessary loss sharpness for accurate reconstruction, while applying a stronger penalty in low-detail regions to facilitate better regularization.

4 Experiments

Dataset. We evaluate our method on LLFF [44] and MipNeRF-360 [5], following previous works [48, 67, 70, 72], where the input resolution is 8× downsampled, and 3 and 12 input views are split for LLFF and MipNeRF-360, respectively.

Implementation. We choose the publicly available 3DGS [27] and its follow-up works as baselines. Specifically, we choose the state-of-the-art NexusGS [70], which leverages foundation models, and CoR-GS [67], DropGaussian [48], and SE-GS [69], which do not.

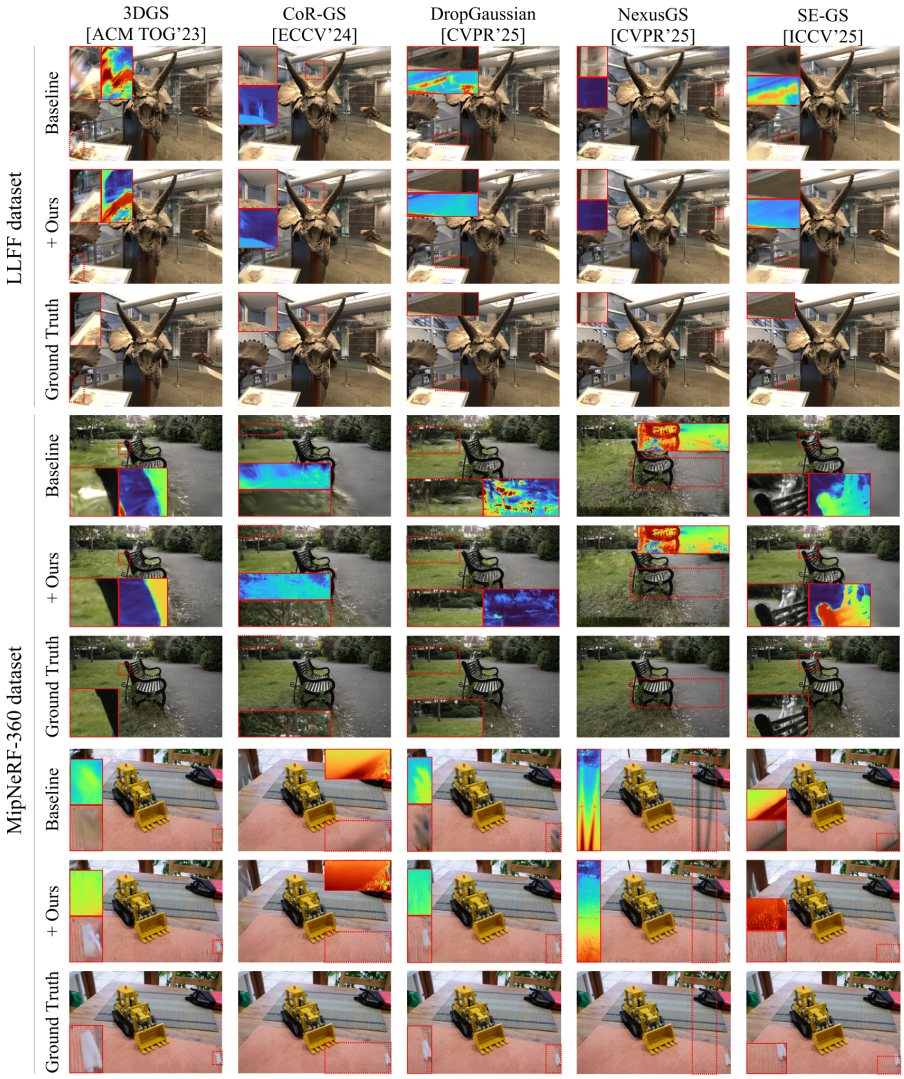

Fig. 3: Qualitative comparison. Our method can be applied to various baselines in a plug-and-play manner, improving reconstruction quality by removing floaters while preserving fine details.

Metrics. We use PSNR, SSIM [59], and LPIPS [68] as evaluation metrics, where LPIPS is computed with a VGG network [54]. Additionally, we use Average Error (AVGE) [47], which is the geometric mean of PSNR, SSIM, and LPIPS. Considering the randomness of 3DGS, we conduct ten runs on the LLFF dataset and five runs on the MipNeRF-360 dataset. We provide an analysis of computational cost in the Appendix.

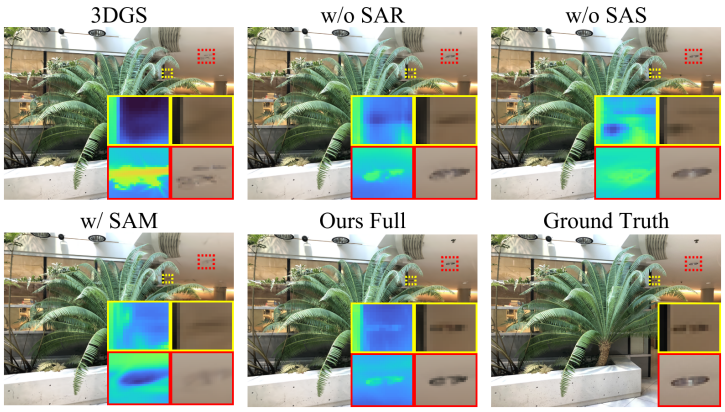

Fig. 4: Ablation study on LLFF dataset. SAR and SAS denote structure-aware regularization weight and structure-aware sharpness, respectively. "3DGS" and "SAM" produce inaccurate geometry (red box). All except "Ours Full" show blurry results (yellow box).

Table 2: Ablation study on LLFF dataset. SAS and SAR denote structure-aware sharpness and structure-aware regularization weight, respectively.

SAM	SAS	SAR	PSNR ↑	SSIM ↑	LPIPS ↓
✗	✗	✗	19.810 ±.339	.6790 ±.0078	.2145 ±.0065
✓	✗	✗	20.198 ±.218	.6958 ±.0034	.2095 ±.0036
✓	✗	✓	20.570 ±.161	.6980 ±.0037	.2142 ±.0042
✓	✓	✗	20.560 ±.259	.7116 ±.0038	.2023 ±.0032
✓	✓	✓	**20.783** ±.300	**.7197** ±.0032	**.1965** ±.0034

4.1 Reconstruction Quality

As shown in Tab. 1, our method achieves clear gains in all metrics, datasets, and baselines. Importantly, these improvements are achieved without any architectural changes or additional priors, but only with our optimization algorithm. Fig. 3 shows visual comparisons. Applying our method consistently improves the baselines by correcting geometric inaccuracies and reducing floating artifacts in novel viewpoints. These results reveal that our proposed optimization algorithm can be seamlessly integrated into current and future 3DGS-based frameworks, providing complementary enhancements.

4.2 Analysis

Ablation study. As shown in Fig. 4 and Tab. 2, directly applying SAM [13] to 3DGS leads to degraded performance. This naive approach uniformly perturbs

Table 3: Performance by covisibility level on LLFF dataset. Our method shows greater improvement with higher view sparsity.

Covisibility level	3DGS	+ Ours	Δ
Covisibility 3	.0483 ±.0093	.0417 ±.0080	-.0066 ±.0043
Covisibility 2	.0757 ±.0156	.0644 ±.0137	-.0114 ±.0067
Covisibility 1	.0948 ±.0256	.0806 ±.0232	**-.0142** ±.0074

Gaussians regardless of image structure, resulting in blurry reconstructions. Structure-aware regularization weighting preserves sharpness in high-detail regions. Meanwhile, structure-aware sharpness enables more faithful estimation by adaptively adjusting perturbation magnitude. Applying either component individually shows performance gains. However, removing either component degrades performance. Using both achieves better generalization, balancing between sharpness reduction and detail preservation.

Loss landscape visualization. To analyze the convergence behavior of our method, we visualize the reconstruction loss landscape and the corresponding optimization trajectories.

SAM converges to flatter minima than 3DGS. In Fig. 5a, the loss range between the local maximum and minimum is 2.37×, and the measured loss sharpness $\lambda_{\max}$ is 1.49× smaller. SAM also achieves a test loss of 0.0129 lower than 3DGS, narrowing the generalization gap from 0.0985 to 0.0809. However, our method converges to less flat minima than SAM. As shown in Fig. 5b, the local loss range and measured loss sharpness of ours are 1.11× smaller and 1.01× smaller than 3DGS, respectively. Interestingly, our method achieves test loss of 0.0141 lower than 3DGS and further narrows the generalization gap to 0.0805, outperforming SAM in generalization.

These results suggest that loss sharpness is not strictly correlated with generalization, supporting our hypothesis that loss sharpness of high-detail regions should be preserved. This aligns with our finding that SAM tends to over-penalize fine details, leading to blurry results (Fig. 4).

Performance improvement by covisibility level. The improvement from our method is more substantial in regions observed by fewer training views. Following CoMapGS [22], we compute covisibility maps using MASt3R [33]. As shown in Tab. 3, the improvement of our method over the baseline 3DGS increases progressively as the covisibility decreases. This behavior aligns with our hypothesis that our method enhances generalization, particularly in under-constrained regions.

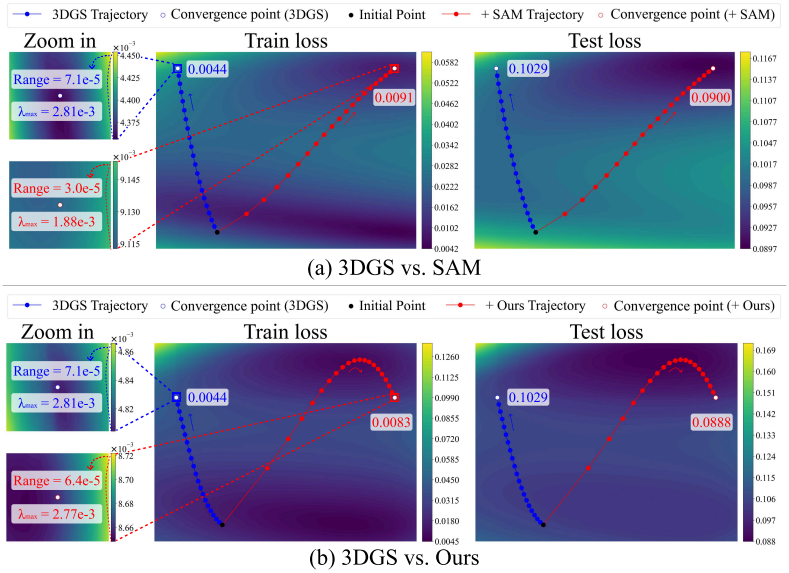

Fig. 5: Loss landscape visualization on the LLFF `fern` scene. We compare the convergence behaviors of 3DGS, SAM, and Ours. To do this, we train 3DGS for 5k iterations, and continue training for an additional 5k iterations without densification, under three distinct settings: 3DGS optimization, SAM, and our proposed method. For visualization, we project the high-dimensional parameter trajectories onto a 2D plane using Principal Component Analysis with parameters from both trajectories. Because the visualization produces a smoothed loss landscape, we provide a zoomed-in view near the convergence points. We measure loss sharpness as the maximum eigenvalue $\lambda_{\max}$ of the Hessian matrix [42, 61].

4.3 Generality of Our Insight

While our primary evaluation focuses on SAM and 3DGS, our core insight—the optimal loss sharpness varies with local detail—has broader implications for reconstruction tasks. Below, we demonstrate its generality with another flat minima optimizer, dynamic scenes, and the NeRFs representation.

Compatibility with another flat minima optimization method. To validate the broader applicability, we apply our core insight to another flat minima optimization technique, Random Weight Perturbation (RWP) [36] instead of SAM. Specifically, within the mixed-RWP loss objective [36], we set both perturbation magnitude and the balance coefficient to be adaptive to local image structure.

As shown in Tab. 4, integrating our approach into RWP yields consistent performance improvements, echoing the quantitative gains in Tab. 2 (third and sixth rows). Furthermore, Fig. 6 illustrates that applying our approach to RWP

Table 4: Quantitative comparison. Our approach generalizes well to other flat minima optimization method.

	LLFF (3 views)			MipNeRF-360 (12 views)		
Method	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
3DGS + RWP	20.079 ± .278	.6971 ± .0050	.1984 ± .0039	19.090 ± .187	.5544 ± .0047	.3767 ± .0042
+ Ours	**20.650** ± .280	**.7157** ± .0043	**.1912** ± .0039	**19.284** ± .193	**.5664** ± .0043	**.3567** ± .0034

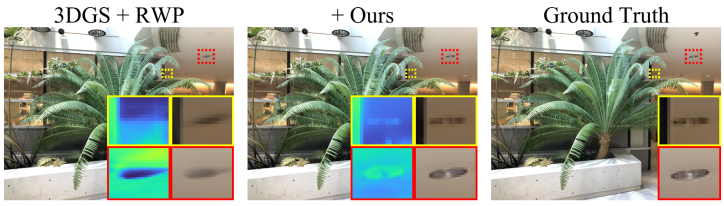

Fig. 6: Qualitative results demonstrating the effect of our approach on another flat minima optimization method. RWP itself produces blurry results, whereas applying our approach improves fine details.

Table 5: Applying FASR to online dynamic 3D Gaussians. Our method is effective in dynamic scenarios under temporal sparsity, notably improving temporal consistency by reducing mTV.

Method	PSNR ↑	SSIM ↑	mTV ↓
Yun *et al.* [66]	32.542	.9486	.1109
+ Ours	**32.622**	**.9497**	**.0989**

effectively improves fine details, mirroring the results presented in Fig. 4 (second row). These results align with our main experiments, further suggesting that flat minima optimization techniques become effective for reconstruction tasks when applied with our insight.

Improving generalization in temporal sparsity. We demonstrate the extensibility of our insight to dynamic scenes. Specifically, we conduct experiments on the Neural 3D Video dataset [38] with an online configuration [15, 19, 35, 63], where observations are spatially dense but temporally sparse, meaning that we can only access the current frame in a sequentially processed video stream. We applied our method to Yun *et al.* [66] with the 3DGStream [56] backbone. Following their protocol, we select the first frame 3D Gaussians with the highest PSNR for initialization and compute the masked total variation (mTV) to measure temporal consistency.

Our method improves temporal consistency and visual quality compared to the baseline (Tab. 5). This shows that our approach enhances generalization not only in the spatial domain but also in the temporal domain. Furthermore, Yun

Table 6: Quantitative comparison with FreeNeRF on LLFF dataset. Applying our approach to the NeRF baseline provides additional performance gain.

Method	PSNR ↑	SSIM ↑	LPIPS ↓
FreeNeRF	19.523	.6063	.3103
+ Ours	**19.584**	**.6182**	**.2983**

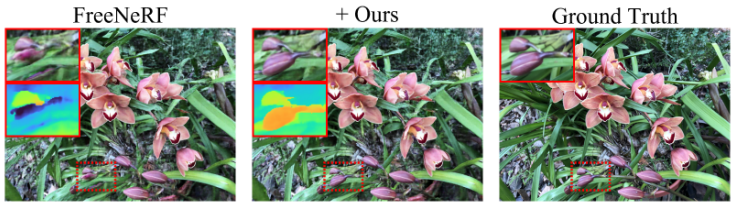

Fig. 7: Qualitative comparison with FreeNeRF on LLFF dataset. Applying our approach corrects the inaccurate geometry of FreeNeRF.

et al. [66] claim that one cause of temporal jittering is the inevitable noise in training datasets. Since Foret *et al*. [13] demonstrate robustness on noisy training data, our finding aligns with this explanation.

Applying our insight to NeRFs. To validate applicability, we apply our core insight to another scene representation, NeRFs [45]. As NeRFs are an implicit representation where each parameter influences the entire scene, we apply our insight to baselines that employ a coarse-to-fine procedure [40, 64]. Specifically, we set a strong perturbation magnitude and regularization weight when the model parameters learn low-frequency components, and gradually decrease these values as the model learns high-frequency details.

We apply this approach to FreeNeRF [64] and demonstrate that it improves both visual quality in novel view synthesis (Fig. 7) and quantitative results (Tab. 6). Although the improvement over the NeRF baseline is smaller than that in the 3DGS baselines due to inherent differences in the representation, these additional gains suggest that our core insight remains effective regardless of the specific implementation or underlying scene representation.

5 Conclusion

In this work, we present the first fundamental investigation linking loss landscape and generalization to sparse novel view synthesis. Notably, we provide valuable insight that optimal loss sharpness in novel view synthesis is highly correlated with local image structure: both flat and sharp minima are essential for accurate reconstruction, which challenges the typical perspective that flatter minima generalize better. To implement this insight, we introduce Structure-Aware Sharpness Regularization, an optimization algorithm that adapts both

the neighborhood radius when calculating loss sharpness and the regularization weight, considering the local image structure. Our approach improves baseline performance by correcting geometric inaccuracies and suppressing floating artifacts. We demonstrate these consistent improvements across a wide range of 3DGS-based frameworks, various flat minima optimization methods, implicit and explicit representations, and static and dynamic scenarios. These results suggest that our strategy provides a general and practical principle for novel view synthesis rather than a narrow implementation. We hope this work inspires the research community to explore the link between loss sharpness and generalization in novel view synthesis, opening a new research direction.

Acknowledgements

This work was supported by an Institute for Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2017-II170072, Development of Audio/Video Coding and Light Field Media Fundamental Technologies for Ultra Realistic Tera-media)

References

1. Aira, L.S., Valsesia, D., Magli, E.: Modeling uncertainty for gaussian splatting. IEEE Transactions on Neural Networks and Learning Systems **36**(6), 11657–11663 (2025). https://doi.org/10.1109/TNNLS.2025.3553582
2. Andriushchenko, M., Croce, F., Müller, M., Hein, M., Flammarion, N.: A modern look at the relationship between sharpness and generalization. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (eds.) Proceedings of the 40th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 202, pp. 840–902. PMLR (23–29 Jul 2023), https://proceedings.mlr.press/v202/andriushchenko23a.html
3. Andriushchenko, M., Croce, F., Müller, M., Hein, M., Flammarion, N.: A modern look at the relationship between sharpness and generalization. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (eds.) Proceedings of the 40th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 202, pp. 840–902. PMLR (23–29 Jul 2023), https://proceedings.mlr.press/v202/andriushchenko23a.html
4. Andriushchenko, M., Flammarion, N.: Towards understanding sharpness-aware minimization. In: Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., Sabato, S. (eds.) Proceedings of the 39th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 162, pp. 639–668. PMLR (17–23 Jul 2022), https://proceedings.mlr.press/v162/andriushchenko22a.html
5. 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 (CVPR). pp. 5470–5479 (June 2022)
6. Bhat, S.F., Birkl, R., Wofk, D., Wonka, P., Müller, M.: Zoedepth: Zero-shot transfer by combining relative and metric depth (2023), https://arxiv.org/abs/2302.12288

7. Cha, J., Chun, S., Lee, K., Cho, H.C., Park, S., Lee, Y., Park, S.: Swad: Domain generalization by seeking flat minima. In: Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., Vaughan, J.W. (eds.) Advances in Neural Information Processing Systems. vol. 34, pp. 22405–22418. Curran Associates, Inc. (2021), https://proceedings.neurips.cc/paper_files/paper/2021/file/bcb41ccdc4363c6848a1d760f26c28a0-Paper.pdf
8. Chen, K., Zhong, Y., Li, Z., Lin, J., Chen, Y., Qin, M., Wang, H.: Quantifying and alleviating co-adaptation in sparse-view 3d gaussian splatting. In: The Thirty-ninth Annual Conference on Neural Information Processing Systems (2025), https://openreview.net/forum?id=GrPo8NTtzK
9. Dai, D., Xing, Y.: Eap-gs: Efficient augmentation of pointcloud for 3d gaussian splatting in few-shot scene reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 16498–16507 (June 2025)
10. Deng, K., Liu, A., Zhu, J.Y., Ramanan, D.: Depth-supervised nerf: Fewer views and faster training for free. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 12882–12891 (June 2022)
11. Dinh, L., Pascanu, R., Bengio, S., Bengio, Y.: Sharp minima can generalize for deep nets. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 70, pp. 1019–1028. PMLR (06–11 Aug 2017), https://proceedings.mlr.press/v70/dinh17b.html
12. Dziugaite, G.K., Roy, D.M.: Computing nonvacuous generalization bounds for deep (stochastic) neural networks with many more parameters than training data. In: Proceedings of the 33rd Annual Conference on Uncertainty in Artificial Intelligence (UAI) (2017), http://auai.org/uai2017/proceedings/papers/173.pdf
13. Foret, P., Kleiner, A., Mobahi, H., Neyshabur, B.: Sharpness-aware minimization for efficiently improving generalization. In: International Conference on Learning Representations (2021), https://openreview.net/forum?id=6Tm1mposlrM
14. Gao, Q., Meng, J., Wen, C., Chen, J., Zhang, J.: Hicom: Hierarchical coherent motion for dynamic streamable scenes with 3d gaussian splatting. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 80609–80633. Curran Associates, Inc. (2024), https://proceedings.neurips.cc/paper_files/paper/2024/file/9370cc8d438b9ed8cb4344818a251d9a-Paper-Conference.pdf
15. Girish, S., Li, T., Mazumdar, A., Shrivastava, A., Luebke, D., De Mello, S.: Queen: Quantized efficient encoding of dynamic gaussians for streaming free-viewpoint videos. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 43435–43467. Curran Associates, Inc. (2024), https://proceedings.neurips.cc/paper_files/paper/2024/file/4c9477b9e2c7ec0ad3f4f15077aaf85a-Paper-Conference.pdf
16. Goli, L., Reading, C., Sellán, S., Jacobson, A., Tagliasacchi, A.: Bayes' rays: Uncertainty quantification for neural radiance fields. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 20061–20070 (June 2024)
17. Haas, M., Xu, J., Cevher, V., Chennuru Vankadara, L.: μp2: Effective sharpness aware minimization requires layerwise perturbation scaling. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 38888–38959.

Curran Associates, Inc. (2024), https://proceedings.neurips.cc/paper_files/paper/2024/file/449a016a6ce6fba3fe50d05482abf836-Paper-Conference.pdf
18. Han, L., Zhou, J., Liu, Y.S., Han, Z.: Binocular-guided 3d gaussian splatting with view consistency for sparse view synthesis. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 68595–68621. Curran Associates, Inc. (2024). https://doi.org/10.52202/079017-2191, https://proceedings.neurips.cc/paper_files/paper/2024/file/7eca17ef54789b0663cab421f2e9dbf5-Paper-Conference.pdf
19. Hu, Q., Zheng, Z., Zhong, H., Fu, S., Song, L., Zhang, X., Zhai, G., Wang, Y.: 4dgc: Rate-aware 4d gaussian compression for efficient streamable free-viewpoint video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 875–885 (June 2025)
20. Izmailov, P., Podoprikhin, D., Garipov, T., Vetrov, D., Wilson, A.G.: Averaging weights leads to wider optima and better generalization. In: Proceedings of the 34th Annual Conference on Uncertainty in Artificial Intelligence (UAI) (2018), http://auai.org/uai2018/proceedings/papers/313.pdf
21. Jang, C., Lee, S., Park, F., Noh, Y.K.: A reparametrization-invariant sharpness measure based on information geometry. Advances in neural information processing systems **35**, 27893–27905 (2022)
22. Jang, Y., Pérez-Pellitero, E.: Comapgs: Covisibility map-based gaussian splatting for sparse novel view synthesis. In: Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR). pp. 26779–26788 (June 2025)
23. Jena, S., Ouasfi, A., Younes, M., Boukhayma, A.: Sparfels: Fast reconstruction from sparse unposed imagery. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 27476–27487 (October 2025)
24. Jensen, R., Dahl, A., Vogiatzis, G., Tola, E., Aanæs, H.: Large scale multi-view stereopsis evaluation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. pp. 406–413. IEEE (2014)
25. Jiang, L., Mao, Y., Xu, L., Lu, T., Ren, K., Jin, Y., Xu, X., Yu, M., Pang, J., Zhao, F., Lin, D., Dai, B.: Anysplat: Feed-forward 3d gaussian splatting from unconstrained views. ACM Trans. Graph. **44**(6) (Dec 2025). https://doi.org/10.1145/3763326, https://doi.org/10.1145/3763326
26. Jiang, Y., Neyshabur, B., Mobahi, H., Krishnan, D., Bengio, S.: Fantastic generalization measures and where to find them. In: International Conference on Learning Representations (2020), https://openreview.net/forum?id=SJgIPJBFvH
27. Kerbl, B., Kopanas, G., Leimkuehler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. **42**(4) (Jul 2023). https://doi.org/10.1145/3592433, https://doi.org/10.1145/3592433
28. Keskar, N.S., Mudigere, D., Nocedal, J., Smelyanskiy, M., Tang, P.T.P.: On large-batch training for deep learning: Generalization gap and sharp minima. In: International Conference on Learning Representations (2017), https://openreview.net/forum?id=H1oyRlYgg
29. Kheradmand, S., Rebain, D., Sharma, G., Sun, W., Tseng, Y.C., Isack, H., Kar, A., Tagliasacchi, A., Yi, K.M.: 3d gaussian splatting as markov chain monte carlo. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (eds.) Advances in Neural Information Processing Systems. vol. 37, pp. 80965–80986. Curran Associates, Inc. (2024), https://proceedings.neurips.cc/paper_files/paper/2024/file/93be245fce00a9bb2333c17ceae4b732-Paper-Conference.pdf

30. Kim, M., Li, D., Hu, S.X., Hospedales, T.: Fisher SAM: Information geometry and sharpness aware minimisation. In: Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., Sabato, S. (eds.) Proceedings of the 39th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 162, pp. 11148–11161. PMLR (17–23 Jul 2022), https://proceedings.mlr.press/v162/kim22f.html
31. Kwon, J., Kim, J., Park, H., Choi, I.K.: Asam: Adaptive sharpness-aware minimization for scale-invariant learning of deep neural networks. In: Meila, M., Zhang, T. (eds.) Proceedings of the 38th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 139, pp. 5905–5914. PMLR (18–24 Jul 2021), https://proceedings.mlr.press/v139/kwon21b.html
32. Lee, S., Kang, K., Ha, S., Yu, H.: Bayesian nerf: Quantifying uncertainty with volume density for neural implicit fields. IEEE Robotics and Automation Letters **10**(3), 2144–2151 (2025). https://doi.org/10.1109/LRA.2025.3526572
33. Leroy, V., Cabon, Y., Revaud, J.: Grounding image matching in 3d with mast3r. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024. pp. 71–91. Springer Nature Switzerland, Cham (2025)
34. Li, J., Zhang, J., Bai, X., Zheng, J., Ning, X., Zhou, J., Gu, L.: 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 (CVPR). pp. 20775–20785 (June 2024)
35. LI, L., Shen, Z., Wang, Z., Shen, L., Tan, P.: Streaming radiance fields for 3d video synthesis. In: Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A. (eds.) Advances in Neural Information Processing Systems. vol. 35, pp. 13485–13498. Curran Associates, Inc. (2022), https://proceedings.neurips.cc/paper_files/paper/2022/file/57c2cc952f388f6185db98f441351c96-Paper-Conference.pdf
36. Li, T., Tao, Q., Yan, W., Wu, Y., Lei, Z., Fang, K., He, M., Huang, X.: Revisiting random weight perturbation for efficiently improving generalization. Transactions on Machine Learning Research (2024), https://openreview.net/forum?id=WbbgOHpoPX
37. Li, T., Zhou, P., He, Z., Cheng, X., Huang, X.: Friendly sharpness-aware minimization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5631–5640 (June 2024)
38. Li, T., Slavcheva, M., Zollhöfer, M., Green, S., Lassner, C., Kim, C., Schmidt, T., Lovegrove, S., Goesele, M., Newcombe, R., Lv, Z.: Neural 3d video synthesis from multi-view video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5521–5531 (June 2022)
39. Lindeberg, T.: Scale selection properties of generalized scale-space interest point detectors. Journal of Mathematical Imaging and Vision **46**(2), 177–210 (2013). https://doi.org/10.1007/s10851-012-0378-3, https://doi.org/10.1007/s10851-012-0378-3
40. Ling, S., Nimier-David, M., Jacobson, A., Sharp, N.: Stochastic preconditioning for neural field optimization. ACM Trans. Graph. **44**(4) (Jul 2025). https://doi.org/10.1145/3731161, https://doi.org/10.1145/3731161
41. Lu, C., Yin, F., Chen, X., Liu, W., Chen, T., Yu, G., Fan, J.: A large-scale outdoor multi-modal dataset and benchmark for novel view synthesis and implicit scene reconstruction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 7557–7567 (October 2023)

42. Luo, H., Truong, T., Pham, T., Harandi, M., Phung, D., Le, T.: Explicit eigenvalue regularization improves sharpness-aware minimization. In: The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024), https://openreview.net/forum?id=JFUhBY34SC
43. McAllester, D.A.: Some pac-bayesian theorems. In: Proceedings of the Eleventh Annual Conference on Computational Learning Theory. p. 230–234. COLT' 98, Association for Computing Machinery, New York, NY, USA (1998). https://doi.org/10.1145/279943.279989, https://doi.org/10.1145/279943.279989
44. Mildenhall, B., Srinivasan, P.P., Ortiz-Cayon, R., Kalantari, N.K., Ramamoorthi, R., Ng, R., Kar, A.: Local light field fusion: practical view synthesis with prescriptive sampling guidelines. ACM Trans. Graph. **38**(4) (Jul 2019). https://doi.org/10.1145/3306346.3322980, https://doi.org/10.1145/3306346.3322980
45. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M. (eds.) Computer Vision – ECCV 2020. pp. 405–421. Springer International Publishing, Cham (2020)
46. Mueller, M., Vlaar, T., Rolnick, D., Hein, M.: Normalization layers are all that sharpness-aware minimization needs. In: Oh, A., Naumann, T., Globerson, A., Saenko, K., Hardt, M., Levine, S. (eds.) Advances in Neural Information Processing Systems. vol. 36, pp. 69228–69252. Curran Associates, Inc. (2023), https://proceedings.neurips.cc/paper_files/paper/2023/file/da909fc3893d272f26fd9db82e09d954-Paper-Conference.pdf
47. Niemeyer, M., Barron, J.T., Mildenhall, B., Sajjadi, M.S.M., Geiger, A., Radwan, N.: Regnerf: Regularizing neural radiance fields for view synthesis from sparse inputs. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5480–5490 (June 2022)
48. Park, H., Ryu, G., Kim, W.: Dropgaussian: Structural regularization for sparse-view gaussian splatting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 21600–21609 (June 2025)
49. Ranftl, R., Bochkovskiy, A., Koltun, V.: Vision transformers for dense prediction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 12179–12188 (October 2021)
50. Seo, K., Hyun, S., Lee, M., Heo, J.P.: Improving sparse-view 3DGS generalization via flat minima optimization (2026), https://openreview.net/forum?id=eH9Wlahibz
51. Shawe-Taylor, J., Williamson, R.C.: A pac analysis of a bayesian estimator. In: Proceedings of the Tenth Annual Conference on Computational Learning Theory. p. 2–9. COLT '97, Association for Computing Machinery, New York, NY, USA (1997). https://doi.org/10.1145/267460.267466, https://doi.org/10.1145/267460.267466
52. Shen, J., Ruiz, A., Agudo, A., Moreno-Noguer, F.: Stochastic neural radiance fields: Quantifying uncertainty in implicit 3d representations. In: 2021 International Conference on 3D Vision (3DV). pp. 972–981 (2021). https://doi.org/10.1109/3DV53792.2021.00105
53. Shi, X., Huang, Z., Li, D., Zhang, M., Cheung, K.C., See, S., Qin, H., Dai, J., Li, H.: Flowformer++: Masked cost volume autoencoding for pretraining optical flow estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1599–1610 (June 2023)
54. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (2015), https://arxiv.org/abs/1409.1556

55. Sun, H., Shen, L., Zhong, Q., Ding, L., Chen, S., Sun, J., Li, J., Sun, G., Tao, D.: Adasam: Boosting sharpness-aware minimization with adaptive learning rate and momentum for training deep neural networks. Neural Networks **169**, 506–519 (2024). https://doi.org/https://doi.org/10.1016/j.neunet.2023.10.044, https://www.sciencedirect.com/science/article/pii/S0893608023006068
56. Sun, J., Jiao, H., Li, G., Zhang, Z., Zhao, L., Xing, W.: 3dgstream: On-the-fly training of 3d gaussians for efficient streaming of photo-realistic free-viewpoint videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 20675–20685 (June 2024)
57. Tsuzuku, Y., Sato, I., Sugiyama, M.: Normalized flat minima: Exploring scale invariant definition of flat minima for neural networks using pac-bayesian analysis. In: International Conference on Machine Learning. pp. 9636–9647. PMLR (2020)
58. Wang, G., Chen, Z., Loy, C.C., Liu, Z.: Sparsenerf: Distilling depth ranking for few-shot novel view synthesis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 9065–9076 (October 2023)
59. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004). https://doi.org/10.1109/TIP.2003.819861
60. Wen, K., Li, Z., Ma, T.: Sharpness minimization algorithms do not only minimize sharpness to achieve better generalization. Advances in Neural Information Processing Systems **36**, 1024–1035 (2023)
61. Wen, K., Ma, T., Li, Z.: How sharpness-aware minimization minimizes sharpness? In: The Eleventh International Conference on Learning Representations (2023), https://openreview.net/forum?id=5spDgWmpY6x
62. Xu, Y., Wang, L., Chen, M., Ao, S., Li, L., Guo, Y.: Dropoutgs: Dropping out gaussians for better sparse-view rendering. In: Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR). pp. 701–710 (June 2025)
63. Yan, J., Peng, R., Wang, Z., Tang, L., Yang, J., Liang, J., Wu, J., Wang, R.: Instant gaussian stream: Fast and generalizable streaming of dynamic scene reconstruction via gaussian splatting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 16520–16531 (June 2025)
64. Yang, J., Pavone, M., Wang, Y.: Freenerf: Improving few-shot neural rendering with free frequency regularization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 8254–8263 (June 2023)
65. Yue, Y., Jiang, J., Ye, Z., Gao, N., Liu, Y., Zhang, K.: Sharpness-aware minimization revisited: Weighted sharpness as a regularization term. In: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. p. 3185–3194. KDD '23, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3580305.3599501, https://doi.org/10.1145/3580305.3599501
66. Yun, Y., Bae, J., Son, H., Kim, S., Lee, H., Bang, G., Uh, Y.: Compensating spatiotemporally inconsistent observations for online dynamic 3d gaussian splatting. In: Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers. SIGGRAPH Conference Papers '25, Association for Computing Machinery, New York, NY, USA (2025). https://doi.org/10.1145/3721238.3730678, https://doi.org/10.1145/3721238.3730678
67. Zhang, J., Li, J., Yu, X., Huang, L., Gu, L., Zheng, J., Bai, X.: Cor-gs: Sparse-view 3d gaussian splatting via co-regularization. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024. pp. 335–352. Springer Nature Switzerland, Cham (2025)

68. 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 (CVPR) (June 2018)
69. Zhao, C., Wang, X., Zhang, T., Javed, S., Salzmann, M.: Self-ensembling gaussian splatting for few-shot novel view synthesis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 4940–4950 (October 2025)
70. Zheng, Y., Jiang, Z., He, S., Sun, Y., Dong, J., Zhang, H., Du, Y.: Nexusgs: Sparse view synthesis with epipolar depth priors in 3d gaussian splatting. In: Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR). pp. 26800–26809 (June 2025)
71. Zhou, Z., Wang, M., Mao, Y., Li, B., Yan, J.: Sharpness-aware minimization efficiently selects flatter minima late in training. In: The Thirteenth International Conference on Learning Representations (2025), https://openreview.net/forum?id=aD2uwhLbnA
72. Zhu, Z., Fan, Z., Jiang, Y., Wang, Z.: Fsgs: Real-time few-shot view synthesis using gaussian splatting. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024. pp. 145–163. Springer Nature Switzerland, Cham (2025)