

Frequency-Aware Density Control via Reparameterization for High-Quality Rendering of 3D Gaussian Splatting

Zhaojie Zeng¹, Yuesong Wang^{1*}, Lili Ju², Tao Guan¹

¹School of Computer Science & Technology, Huazhong University of Science and Technology, China

²Department of Mathematics, University of South Carolina, USA

zhaojiezeng@hust.edu.cn, yuesongwang@hust.edu.cn, ju@math.sc.edu, qd_gt@hust.edu.cn

Abstract

By adaptively controlling the density and generating more Gaussians in regions with high-frequency information, 3D Gaussian Splatting (3DGS) can better represent scene details. From the signal processing perspective, representing details usually needs more Gaussians with relatively smaller scales. However, 3DGS currently lacks an explicit constraint linking the density and scale of 3D Gaussians across the domain, leading to 3DGS using improper-scale Gaussians to express frequency information, resulting in the loss of accuracy. In this paper, we propose to establish a direct relation between density and scale through the reparameterization of the scaling parameters and ensure the consistency between them via explicit constraints (i.e., density responds well to changes in frequency). Furthermore, we develop a frequency-aware density control strategy, consisting of densification and deletion, to improve representation quality with fewer Gaussians. A dynamic threshold encourages densification in high-frequency regions, while a scale-based filter deletes Gaussians with improper scale. Experimental results on various datasets demonstrate that our method outperforms existing state-of-the-art methods quantitatively and qualitatively.

Code — <https://github.com/whoiszzj/FDS-GS>

Introduction

Novel View Synthesis (NVS) is a significant task in computer vision, aiming to generate images from arbitrary viewpoints based on scene representations reconstructed from captured images. Neural Field Rendering (NeRF) (Mildenhall et al. 2020) has emerged as a leading solution for NVS, renowned for its realistic rendering quality, and has inspired numerous subsequent works (Barron et al. 2021, 2022, 2023; Zhang et al. 2020; Chen et al. 2023b). However, 3D Gaussian Splatting (3DGS) (Kerbl et al. 2023) has recently gained attention for its combination of high-quality rendering and speed, offering a novel point-based real-time rendering solution for NVS applications in the VR and AR industries (Xu et al. 2023a; Deng et al. 2022b; Li et al. 2018).

In addition to the flexibility offered by organizing a set of 3D Gaussian primitives (termed Gaussians in the following), each defined by position, opacity, spherical harmonic,

*Corresponding author.

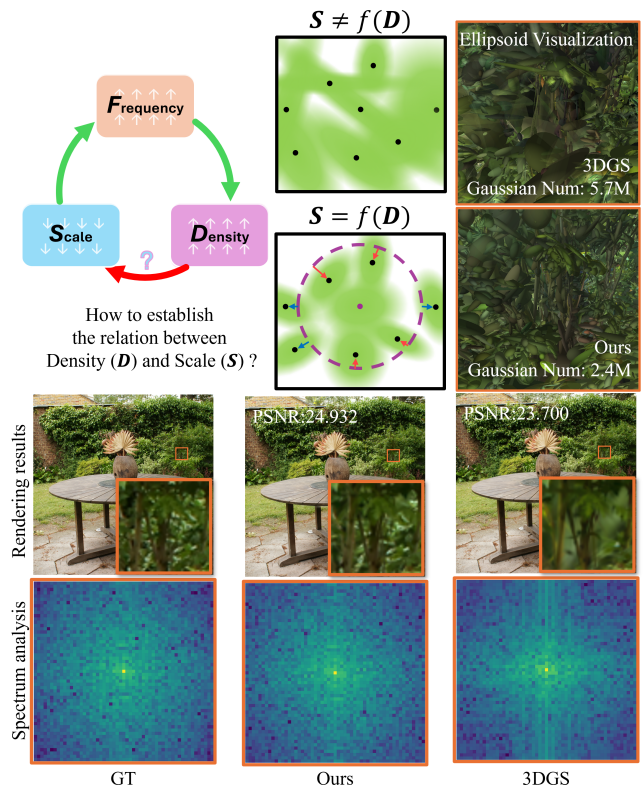


Figure 1: By linking the density and scale (the purple circle corresponds to the scale, constrained by the density), our FDS-GS can better reflect the frequency changes than 3DGS via density control (more similar to the GT’s spectrum), achieving better render quality with fewer Gaussians.

and covariance matrix decomposed to scale and rotation, another pivotal innovation of 3DGS is its adaptive density control. As practical NVS demands accurate representation of scene details and sparse points are commonly insufficient for scene coverage and detail representation, using dense point cloud as an initial input is necessary for many point-based methods (Kopanas et al. 2021; Zhang et al. 2022; Franke et al. 2024). The adaptive density control liberates point-based methods from the necessity of relying on dense point clouds as input. Consequently, 3DGS can produce high-

quality results by leveraging only sparse points generated for free as part of the Structure-from-Motion (SfM) process (Schönberger and Frahm 2016).

The underlying principle of density control can be analyzed within the signal processing framework. According to the Fourier transform, complex signals can be represented as a superposition of sine and cosine waves at different frequencies. Sufficient high-frequency components are necessary to capture complex signals, especially abrupt changes, accurately. Similarly, 3DGS uses alpha blending to render scenes by blending Gaussians of different colors to produce the final output. It requires the superposition of many Gaussians to represent complex color variations and sufficient smaller-scale ones to capture high-frequency details.

In practice, 3DGS typically generates more Gaussians through view space gradients and performs split operations on larger Gaussians. However, due to the extensive parameter space in 3DGS, this process does not always result in appropriate scale, sometimes producing overly large Gaussians in regions with details (i.e., local optimal solution), as illustrated in Fig. 1. Moreover, the decision to split a Gaussian is often made based on its current scale but without understanding the region it represents. This inflexibility will prevent the generation of sufficiently small Gaussians needed to capture fine details as high-frequency components. From the above discussions and comparison results shown in Fig. 1, we can identify the following inconsistencies in 3DGS. When dealing with high-frequency regions that represent fine details, 3DGS increases the density of Gaussians through densification. Although there is a tendency for the scale to decrease, the lack of explicit and direct regularization constraints leads to that the Gaussians generated through densification do not fully contribute to representing high-frequency information.

To ensure consistency in Frequency, Density, and Scale in 3DGS, we propose to establish a direct relation between density and scale across the domain by reparameterization. Specifically, we first quantify the density information at each position as the number of Gaussians in a unit volume centered at the position. We then reparameterize the scale and decompose it into four variables, including an absolute scalar and a relative scale vector which respond to its anisotropy in three dimensions. Following the principle that higher-density areas will have smaller scales, making them act as high-frequency components, we establish an explicit constraint between the absolute scale scalar and the density. We further introduce a frequency-aware density control to enhance the density quality, i.e., using fewer Gaussians to get better rendering quality. Following the framework of 3DGS, we use both densification and deletion strategies for density control. First, we propose a dynamic threshold for densification: we sort the gradients of the Gaussians and adjust the threshold based on their distribution. This adaptive approach provides finer control over densification compared to the fixed threshold used in vanilla 3DGS, making the density more responsive to changes in frequency. Subsequently, we design a scale-based confidence filter to effectively remove overfitting Gaussians with improper scales, further ensuring consistency in frequency, density, and scale. We will

call our proposed method “FDS-GS” for brevity.

In summary, our main contributions are as follows:

- We quantify the density and reparameterize the scale of the Gaussian to establish a direct relation between density and scale, based on which we use the density to constrain the covariance of the Gaussian to ensure that the density can better respond to changes in frequency.
- We develop a frequency-aware density control strategy consisting of densification with a dynamic threshold and a scale confidence filter, optimizing the density to be more expressive of frequency variations to improve the accuracy of the rendering.
- Our method, FDS-GS, outperforms existing state-of-the-art methods by achieving higher rendering quality with a reasonable number of Gaussians.

Related Works

NeRF-based Rendering. Different from traditional methods that synthesize novel views based on explicit geometry representations such as meshes or points, recently emerged neural rendering methods utilize a set of MLPs to represent the scenes implicitly. Neural Radiance Field (NeRF) (Mildenhall et al. 2020) can be seen as a landmark work in this direction. By encoding a continuous volume with a view-dependent appearance using the parameters of a deep fully-connected neural network, NeRF can achieve both photorealistic rendering quality and low storage pressure. Many subsequent works have been proposed to further advance the progress. Mip-NeRF (Barron et al. 2021) renders the scene using conical frustums instead of rays to reduce aliasing artifacts. Mip-NeRF360 (Barron et al. 2022) presents an extension of mip-NeRF to handle unbounded scenes better. Martin-Brualla et al. (Martin-Brualla et al. 2021) introduce a series of extensions to NeRF to make it possible to synthesize novel views using only in-the-wild photographs. On the other hand, NeRF-based methods generally suffer from high consumption for training and rendering times due to the deep MLP structure. To tackle this drawback, KiloNeRF (Reiser et al. 2021) utilizes thousands of tiny MLPs to replace the deep MLP, which can greatly increase the rendering speed. FastNeRF (Garbin et al. 2021) applies a graphics-inspired factorization to speed up the computation of the pixel values in the rendered image. Deng et al. (Deng et al. 2022a) introduce depth supervision to guide geometry learning, which allows faster training. Instant-NGP (Müller et al. 2022) replaces the original encoding with multi-resolution hash encoding and only adopts a smaller MLP, consequently leading to a dramatic reduction of time consumption for both training and rendering. Despite the above efforts, it is currently still hard for NeRF-based methods to achieve both low time consumption and high rendering quality.

3DGS-based Rendering. In order to ensure high rendering quality while reducing the time overhead, researchers also have adopted different structures (Fridovich-Keil et al. 2022; Chen et al. 2022) other than NeRF, but the results were still unsatisfactory enough until the advent of 3D Gaussian

Splatting (Kerbl et al. 2023). 3DGS is developed upon differentiable point-based rendering methods (Zhang et al. 2022; Rückert, Franke, and Stamminger 2022) and treat each 3D point as a 3D Gaussian which allows a more flexible representation regime. Then the novel view is rendered using a splatting strategy (Zwicker et al. 2001). Moreover, 3DGS has designed an adaptive density control strategy, which makes it possible to use as input only sparse point clouds that are generated for free when using SfM to obtain camera parameters. Such feature of 3DGS can greatly reduce the computational cost, allowing it to replace NeRF-based methods in almost any scenario (Kocabas et al. 2023; Chen et al. 2023a; Zhou et al. 2023; Matsuki et al. 2024; Keetha et al. 2023). Given the great potential of 3DGS, more studies have been invested to further improve its performance. Chung et al. (Chung, Oh, and Lee 2023) introduce monocular depth estimation to supervise additionally to improve the performance given sparse-view inputs. Scaffold-GS (Lu et al. 2023) divides the space into voxels, which serve as anchors for nearby 3D Gaussians so that these 3D Gaussians have a soft geometric constraint and the probability of generating outliers is reduced. GaussianPro (Cheng et al. 2024) explicitly constrains the positions of the newly generated 3D Gaussians by introducing the idea of depth propagation in Multi-View Stereo. While these works mainly focus on improving the rendering quality by constraining the geometry of the 3DGS to reduce the floaters, other researches (Zhang et al. 2024; Yu et al. 2024) have noted the importance of frequency constraints for 3DGS, given that nerf-based methods (Xu et al. 2023b) have demonstrated the importance of learning in the frequency space. Mip-Splatting (Yu et al. 2024) goes for alias-free 3DGS with frequency constraints, which mainly addresses the problem of inconsistent rendering at different resolutions but does not pay attention to improving the rendering at the original resolution.

FreGS (Zhang et al. 2024) attempted to interpret and optimize 3DGS from a frequency perspective by transferring the image into the frequency domain via a Fourier transform and designing a frequency loss function that encourages the rendered image to be consistent with the training image in the frequency domain. This approach can partially solve the problem of fitting high-frequency information with large-scale Gaussians. Due to the large frequency loss in these regions, the density control will make the large-scale Gaussians further split into smaller ones, thus jumping out of the pit of local optima. On the other hand, further densification in this approach will simultaneously lead to a huge rise in Gaussian numbers, causing giant rendering and storage pressure. In addition, the same spatial region will result in large differences in image-space frequency if it is seen from different viewpoints, i.e., the image-space frequency does not have multi-view consistency. For instance, a region with mid-frequency information might appear as high-frequency from certain viewpoints, resulting in an excessive number of small-scale Gaussians, further causing redundancy. Therefore, improving the rendering quality of 3DGS from the frequency perspective still remains to be delved into.

Proposed Method

Overview

In the following section, we will first briefly review the vanilla 3DGS method. Then, we will delve into the motivations behind building a direct relation between density and scale and explain how it is implemented via reparameterization. After linking the density and scale, we introduce a frequency-aware density control to better capture frequency variations with density changes. This control strategy involves densification with a dynamic threshold to enable the method to respond more swiftly to high-frequency information and a scale-based confidence filter that computes confidence scores for Gaussians and filters them accordingly. The frequency-aware density control process allows us to significantly reduce the number of Gaussians while still achieving high reconstruction quality. Fig. 2 illustrates the pipeline.

Preliminary

3DGS (Kerbl et al. 2023) represents scenes through a series of anisotropic Gaussian primitives and obtains the rendered images for the corresponding viewpoint through the splatting algorithm. Each primitive is parameterized by a position $\mu \in \mathcal{R}^3$, an opacity $\alpha \in [0, 1]$, a covariance matrix $\Sigma \in \mathcal{R}^{3 \times 3}$ and spherical harmonic (SH) coefficients. The probability distribution for each Gaussian G is defined as:

$$G(\mathbf{x}) = e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)}, \quad (1)$$

where $\mathbf{x} \in \mathcal{R}^3$ is an arbitrary position. To maintain the semi-positive definiteness of the covariance matrix Σ , it is often to decompose it into a scaling matrix \mathbf{S} and a rotation matrix \mathbf{R} such that:

$$\Sigma = \mathbf{R} \mathbf{S} \mathbf{S}^T \mathbf{R}^T, \quad (2)$$

where \mathbf{S} is a diagonal matrix with positive entries, which can be parameterized by a vector $\mathbf{s} \in (\mathcal{R}^+)^3$, and \mathbf{R} is a rotation matrix, which can be parameterized by a quaternion $\mathbf{q} \in \mathcal{R}^4$. To render an image for a given viewpoint, the 3D Gaussians G are first projected to get a 2D Gaussians G' according to (Zwicker et al. 2001), and the spherical harmonics are converted to view-dependent color \mathbf{c} . Then, the image is rendered by alpha-blending according to the Gaussians' depth order based on the tile. For each pixel with position \mathbf{x}' on the image, the color is acquired by:

$$\mathbf{C}(\mathbf{x}') = \sum_{k=1}^K \mathbf{c}_k \alpha_k G'_k(\mathbf{x}') T_k, \quad (3)$$

with

$$T_k = \prod_{j=1}^{k-1} (1 - \alpha_j G'_j(\mathbf{x}')), \quad (4)$$

where K denotes the number of Gaussians used to render the pixel. Since 3DGS takes sparse points from SfM as input, usually inadequate for scene details representation, a densification strategy is then introduced to create new Gaussians. A key criterion for densification involves accumulating view space gradients. If the gradient exceeds a preset parameter τ_{pos} , the Gaussian will be cloned or split based

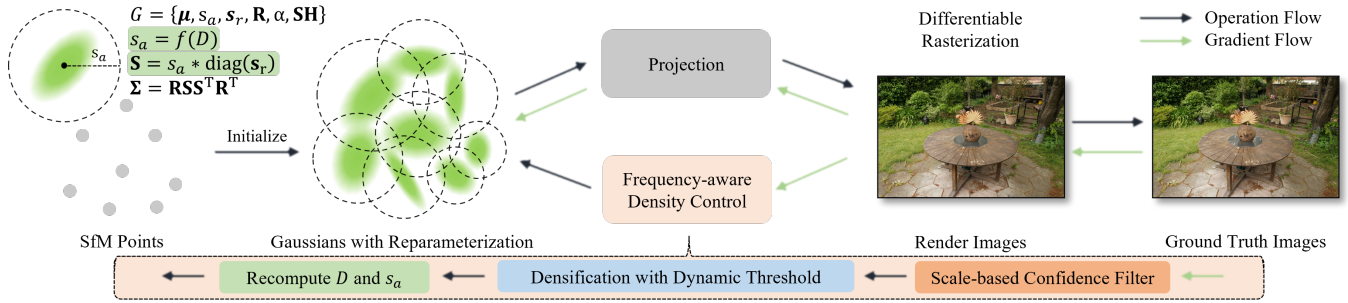


Figure 2: The pipeline of the proposed FDS-GS. FDS-GS is an improvement upon the 3DGS. By reparameterizing the scale \mathbf{S} of the Gaussian and quantifying the density as D , we establish a direct relation between density and scale and constrain the scale according to the density to better express the frequency changes of the scene. To better control the density to achieve better rendering with fewer Gaussians, every few iterations, we compute the confidence of each Gaussian, delete the Gaussians with low confidence, and then compute a dynamic threshold for densification based on the gradient of the remaining Gaussians. Once the Gaussian distribution has changed, we will then recompute the values of D and s_a .

on the scale of the Gaussian. Cloning duplicates Gaussians while splitting divides large Gaussians into smaller intersecting ones. Newly generated Gaussians inherit most attributes from the original Gaussian seamlessly. Additionally, to control the number of Gaussians, 3DGS suggests resetting opacity every 3000 iterations and periodically removing Gaussians with opacity below a preset parameter ϵ_α or with excessively large scales.

Linking the Density and Scale

Density Estimation and Scale Reparametrization As mentioned before, when rendering high-frequency details, 3DGS often increases density in these areas, allowing a greater number of Gaussians to capture the details and reduce discrepancies between the ground truth and rendered images. However, without explicit constraint on the scale of each Gaussian, the extensive solution space associated with the 3DGS parameterization can result in an inappropriate distribution of improper-scale Gaussians, thereby hindering accurate representation of the scene details as demonstrated in Fig. 1. To address this problem, we introduce an explicit constraint by defining a relation between density and scale.

Our first step is to quantify density information based on the current Gaussian distribution. Here, we let the density refer to the number of points per unit volume at each position. Specifically, for a set of Gaussians $\{G\}$, the density D at a position \mathbf{x} is estimated as follows:

$$D(\mathbf{x}) = \frac{K(\mathbf{x}, \Delta V, G)}{\Delta V}, \quad (5)$$

where K denotes the number of Gaussians in $\{G\}$ which intersect the region described by a unit volume ΔV . The unit volume ΔV is a small region in 3D space, centered at a position \mathbf{x} . The size of ΔV is determined by a scale parameter, which could represent the radius of a sphere.

Next, we try to establish a direct relation between density and scale. By reparameterizing the scale vector \mathbf{s} into an absolute scale scalar $s_a > 0$ and a relative scale vector $\mathbf{s}_r \in (0, 1)^3$, all of which are learnable, we impose a constraint – the absolute scale scalar s_a is defined as a function

of density, expressed as:

$$s_a = f(D). \quad (6)$$

Then the scaling matrix in Eq. (2) is rewritten as:

$$\mathbf{S} = \text{diag}(\mathbf{s}), \quad \mathbf{s} = s_a(\mathbf{s}_{r_1}, \mathbf{s}_{r_2}, \mathbf{s}_{r_3}). \quad (7)$$

Relation Model and Scale Constraint To apply the above definitions to the training optimization process, we first need to compute the density at the center μ of each Gaussian and then compute the absolute scale scalar s_a corresponding to this density by defining a relation function f . Instead of defining a fixed search range for density estimation, which is difficult to be generalizable for all scenarios and computationally expensive, we employ a CUDA-based KDTree (Grandits et al. 2021) to find the K nearest neighbors of each Gaussian. The weighted distance to these neighbors is used to estimate the local density D for each Gaussian as follows:

$$D(\mu) = \frac{K}{\frac{4}{3}\pi\tilde{R}^3}, \quad (8)$$

with

$$\tilde{R} = \frac{\sum_{k=1}^K \omega_k d_{\mathbf{q}_k \mu}}{\sum_{k=1}^K \omega_k}, \quad (9)$$

where K denotes a preset fixed number of neighbor Gaussians for density calculation, $d_{\mathbf{q}_k \mu}$ is the distance of the k -th neighbor Gaussian to the current Gaussian's center μ , $\omega_k > 0$ is a weighting parameter defined by:

$$\omega_k = \exp\left(-\left(\frac{d_{\mathbf{q}_k \mu} - d_{\mathbf{q}_1 \mu}}{\text{median}_\mu\{d_{\mathbf{q}_1 \mu}\}}\right)^2\right), \quad (10)$$

where $d_{\mathbf{q}_1 \mu}$ is the nearest neighbor distance, which represents the distance between the center point of G 's nearest neighbor and μ , and $\text{median}_\mu\{d_{\mathbf{q}_1 \mu}\}$ denotes the median of the nearest neighbor distances for all Gaussians, which corresponds to a scale scalar factor for this scene. With the use of ω_k , \tilde{R} can be adaptively adjusted according to the scale

of the scene and the Gaussian distribution to ensure generalization.

Having quantified the density, we need to calculate the corresponding absolute scale scalar. Intuitively, higher-density regions should have smaller-scale Gaussians and vice versa, indicating a negative correlation. From Eq. (8), we observe a negative correlation between \tilde{R} and density. Thus, it infers that the \tilde{R} and the absolute scale scalar s_a are positively correlated. Furthermore, \tilde{R} and s_a should belong to the same domain in terms of their physical meaning. Therefore, here we simply define the mapping between D and s_a as follows:

$$s_a = f(D(\mu)) := \tilde{\theta}D(\mu)^{-\frac{1}{3}} = \theta\tilde{R}, \quad (11)$$

where $\theta = \tilde{\theta}(\frac{4\pi}{3K})^{1/3} > 0$ is a predefined constant parameter related to K . To optimize computational efficiency, we recompute the density D only when the Gaussians undergo structural changes, i.e., adding or removing Gaussians. We then compute the corresponding s_a based on f to reset the scale of each Gaussian to impose the explicit constraints. Note that s_r will not be reset, preserving the original shape of the Gaussian.

Frequency-aware Density Control

By linking the density and scale and imposing explicit constraints on the scale, we now expect to better express the changes in scene frequency by controlling the density, achieving better rendering quality. However, the density control of vanilla 3DGS is not designed from the frequency perspective, so the final density distribution isn't the optimal one to express frequency changes. To resolve this issue, i.e., to use fewer Gaussians to get better rendering quality as much as possible, we design a frequency-aware density control approach. As with 3DGS, our frequency-aware density control also consists of densification and deletion.

Densification with Dynamic Threshold The complexity of scenes varies, and achieving the best rendering effect often requires meticulous attention to detail. While large-scale Gaussians struggle to represent regions with high-frequency information, a sufficient number of small-scale Gaussians can readily convey both low-frequency and high-frequency information. Given a fixed threshold for densification, it is safest to set it low enough to encourage over-densification of the Gaussians, allowing numerous small-scale Gaussians to fit arbitrary scenarios. However, this naive strategy would cause the explosion of Gaussians and thus is impractical considering resource limitations. Therefore, we need a wiser threshold that not only ensures enough Gaussian to express the scene details but also avoids redundancy from over-densification. We propose to use a dynamic gradient threshold approach to resolve the above issue. In each densification process, we calculate the minimum gradient value $grad_{min}$ and the mean gradient value $grad_{mean}$ and then perform histogram statistics between $grad_{min}$ and $3 \cdot grad_{mean}$. From largest to smallest, we compute the gradients $grad_{25\%}$ of the top 25% Gaussian points, compare them with the preset threshold $grad_{preset}$, and obtain the final densification gradient threshold $\tau_{pos} = \max(grad_{25\%}, grad_{preset})$.

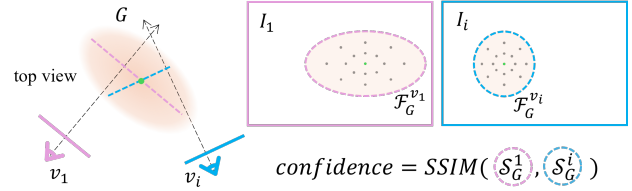


Figure 3: Confidence Calculation Example. For a given Gaussian G , project it onto observable images to obtain the corresponding footprints. We sample several points within the footprints to calculate the SSIM as the confidence.

In the early stage of optimization, we mainly rely on $grad_{25\%}$ to determine the threshold of densification. This scene-dependent threshold encourages higher density in regions currently with higher rendering discrepancy while simultaneously ensuring a controlled and gradual increase in the number of Gaussians. This way prevents an unmanageable explosion in Gaussians, which could impose significant computational burdens. On the other hand, it also could avoid optimization stagnation due to a mismatch between the threshold setting and the scene. After the optimization has stabilized, $grad_{preset}$ is gradually considered more to prevent excessive densification.

Scale-based Confidence Filter Although our method can link scale and density to better represent frequency changes, due to the stochastic nature of the splitting operation in the densification and the complexity of the optimization, the Gaussians may not always be closely distributed on the real scene surface. This could result in the density possibly not matching well with the scene, negatively affecting our scale constraints in some regions. Thus, we need to correct the density of these regions. Inspired by the PatchMatch (Bleyer, Rhemann, and Rother 2011) method, we design below a scale-based confidence filter, which calculates the photometric consistency of the Gaussian on different views according to the scale, uses it as a basis for determining the Gaussian's confidence and filters out the ones with low confidence.

Firstly, we need to select several observable views for each Gaussian. For a given view v_i , we obtain the contribution c of the Gaussian G under this view by:

$$c(G, v_i) = \sum_{\mathbf{x}' \in \mathcal{F}_G^{v_i}} \alpha G'(\mathbf{x}'), \quad (12)$$

where $\mathcal{F}_G^{v_i}$ is the footprint of G at view v_i and G' denotes its 2D Gaussian distribution. Then, we employ a CUDA-based Heap Sort to retrieve the top $M > 1$ views with the largest contributions for the Gaussian G as the observable views $V^{(M)}$:

$$V^{(M)} = \text{HeapSort}_M(c(G, v_i) | i = 1, 2, \dots, N_v), \quad (13)$$

where N_v is the total number of views. According to the 2D Gaussian distribution obtained from each projection, we sample 49 points within the footprint and compute their weight, denoted as \mathcal{S}_G^i and Ω_G^i , as shown in Fig. 3. Please refer to the supplementary material for more details.

Dataset	MipNeRF-360			Tanks&Temples			Deep Blending		
Method — Metrics	SSIM↑	PSNR↑	LPIPS↓	SSIM↑	PSNR↑	LPIPS↓	SSIM↑	PSNR↑	LPIPS↓
Instant-NGP	0.699	25.59	0.331	0.745	21.92	0.305	0.817	24.96	0.390
Mip-NeRF360	0.792	27.69	0.237	0.759	22.22	0.257	0.901	29.40	0.245
3DGS	0.815	27.21	0.214	0.841	23.14	0.183	0.903	29.41	0.243
3DGS†	0.813	27.48	0.221	0.848	23.89	0.181	0.900	29.48	0.247
Scaffold-GS†	0.806	27.57	0.237	0.852	24.51	0.189	0.907	29.98	0.257
MipSplatting†	0.827	27.73	0.190	0.859	23.93	0.161	0.903	29.34	0.239
3D-HGS†	0.813	27.84	0.218	0.859	24.65	0.173	0.904	29.44	0.244
FreGS	0.826	27.85	0.209	0.849	23.96	0.178	0.904	29.93	0.240
Ours	0.835	27.95	0.180	0.863	24.27	0.141	0.906	29.98	0.231

Table 1: Quantitative comparison of our FDS-GS method with other state-of-the-art methods. The results reproduced by ourselves using the official code are marked with †. The best, second best, state-of-the-art best are marked in red, orange, and yellow, respectively.

After obtaining the sampling points for the Gaussian G across the observable views $V^{(M)}$, we calculate the SSIM (Wang et al. 2004) between these sampling points:

$$confidence = \frac{1}{M-1} \sum_{i=2}^M SSIM(\mathcal{S}_G^1, \Omega_G^1, \mathcal{S}_G^i, \Omega_G^i). \quad (14)$$

where \mathcal{S}_G^1 and Ω_G^1 are from the view with the largest contribution. Please refer to the supplementary material for more details. After obtaining the confidence for each Gaussian (ranging from $[0, 1]$), we filter out those with extremely poor confidence using a fixed threshold τ_c .

Experimental Results

Implementation

We select 3DGS (Kerbl et al. 2023) as the baseline and implement our FDS-GS method based on its open-source code. Following 3DGS, we utilize the SfM point cloud obtained from Colmap (Schönberger and Frahm 2016) for initialization. Except for the first 1000 iterations, densification is executed every 500 iterations, followed by recomputing density and s_a . The scale confidence filter is executed every 1000 iterations. The frequency-aware density control persists until reaching 15000 iterations. In all our experiments, we set $K = 50$, $\theta = 1.2$ for linking the density and scale, $grad_{preset} = 0.0005$ for dynamic threshold, and $M = 2$, $\tau_c = 0.2$ for scale confidence filter.

Datasets

Following 3DGS, we utilize the full set of scenes presented in *Mip-NeRF360* (Barron et al. 2022), two scenes from the *Tanks&Temples* (Knapitsch et al. 2017) and two scenes provided by the *Deep Blending* (Hedman et al. 2018) to test our method. We use common metrics (PSNR, SSIM (Wang et al. 2004), LPIPS (Zhang et al. 2018)) to quantify the performance. All experiments were performed on an A100 GPU.

Quantitative and Qualitative Comparison

We compare our method with other state-of-the-art methods, including Instant-NGP (Müller et al. 2022), MipNeRF360 (Barron et al. 2022), 3DGS (Kerbl et al. 2023), Scaffold-GS (Lu et al. 2023), MipSplatting (Yu et al. 2024),

3D-HGS (Li et al. 2024) and FreGS (Zhang et al. 2024) on the *MipNeRF360*, *Tanks&Temples*, and *Deep Blending* datasets. Since some of them do not provide experimental results for the full set of *Mip-NeRF360* in their papers, we have reproduced those on the datasets using the official code provided by the authors with the same coding environment. Quantitative results are shown in Tab. 1. Our method outperforms other methods on all datasets, further proving its robustness. Some qualitative results are depicted in Fig. 4. The scenes shown in Fig. 4 are all complex with rich details, which highly tests the method’s ability to represent scene details. Due to the lack of explicit scale constraints in 3DGS, it is often difficult to cope with complex frequency changes by controlling the density alone, and thus, the details of the scene are difficult to preserve. On the other hand, benefiting from linking density and scale, our FDS-GS method can explicitly encourage consistency between them, which reduces the optimization difficulty and leads to better performance in detail restoration. While FreGS also aims to preserve these scene details, as shown in Tab. 1, its overall performance is not as good as ours. In addition, it usually introduces a large number of redundant Gaussians (as mentioned in its paper, the number of Gaussians on the *stump* set shoots up to 8.1 million, compared with the 4.6 million of the vanilla Gaussian). At the same time, our method has reasonable control over density with the Gaussian number 1 million less than the vanilla 3DGS (3.2 million on *stump*). More specific data will be given in the supplementary material.

Ablation Study

We conduct ablation studies on the *MipNeRF360* full set to demonstrate the effectiveness of the modules we propose.

Linking the Density and Scale In this ablation experiment, we eliminate the scale constraint provided by linking density and scale to prove its effectiveness. From Tab. 2, we can observe that the number of Gaussians is slightly increased compared to the full-version method, yet the rendering quality is significantly degraded. This is a good indication that solely controlling the density can not maintain a good representation of the frequency changes in the scene. With our scale constraint, we can obtain a better rendering result by simultaneously controlling both density and scale.

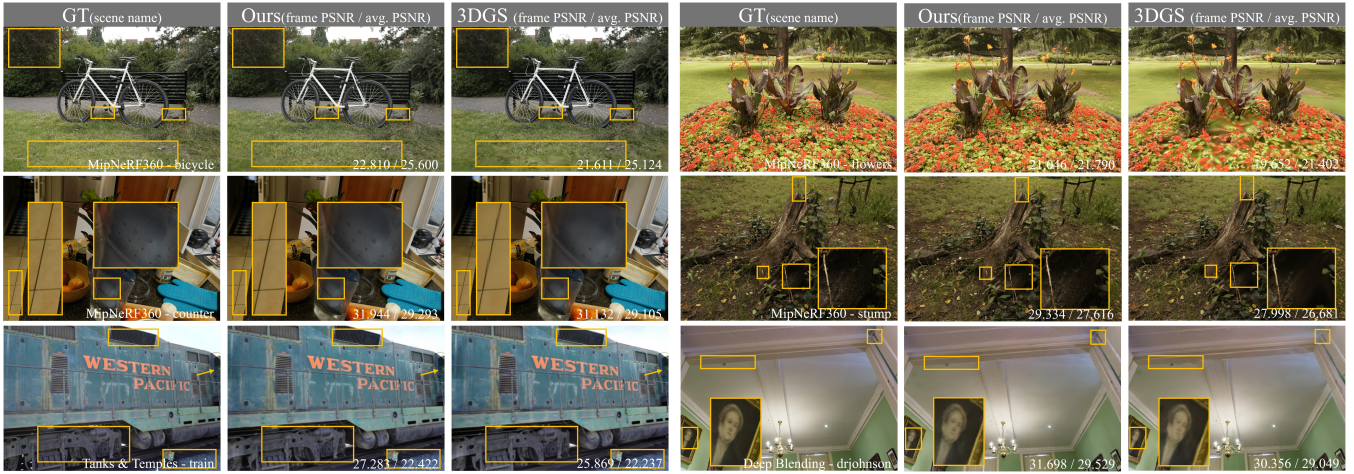


Figure 4: Qualitative comparisons of our FDS-GS method with vanilla 3DGS. Our method provides a superior restoration of scene details. GT means the ground truth. Zoom in for more details.

As illustrated in Fig. 5, we visualized the relation between density and scale through a scatter plot. It can be observed that the *Density* and *Scale* are in negative correlation. Our FDS-GS method and 3DGS both have this trend, but our method displays a tighter distribution, indicating the effectiveness of the constraint between scale and density.

Densification with Dynamic Threshold We now test the effectiveness of the dynamic threshold in densification. We compare it with the fixed gradient threshold, $\tau_{pos} = 0.0005$. As seen from Tab. 2, the rendering quality achieved with a fixed threshold is nearly the same as with our dynamic threshold, aside from minor fluctuations. However, the number of Gaussians increases by almost 1 million, indicating significant over-densification. In other words, Gaussians with excessive density are used to represent regions with minimal frequency changes. This redundant representation, caused by the mismatch between density and frequency, leads to unnecessary rendering and storage overhead.

Scale-based Confidence Filter In this experiment, we replace the scale-based confidence filter with the original filter used in the vanilla 3DGS. As shown in Tab. 2, using the original filter causes the number of Gaussians to skyrocket, which not only fails to improve rendering quality but also degrades it since incorrect density misguides the scale. Our proposed filter more effectively controls density and, when combined with scale, leads to better rendering results.

Conclusion

In this paper, we develop FDS-GS, which can significantly enhance the performance of 3DGS by linking the density and scale to better capture frequency changes. We reparameterize scale, quantify density, and establish a direct relation between them, using density to constrain scale for consistency. This approach generates more Gaussians with small scales when high-frequency information is required, improving rendering quality. Additionally, we design

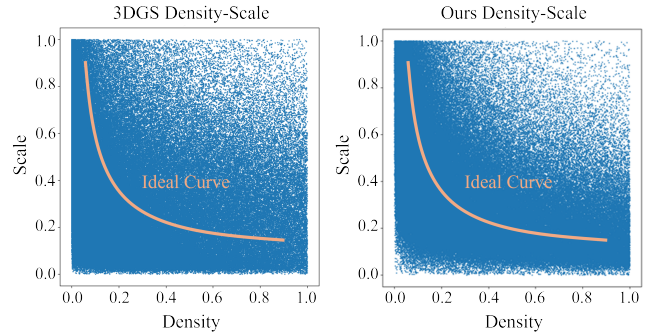


Figure 5: The relation between density and scale. We randomly sample 25,000 Gaussians from the point cloud at iteration 30,000 (*MipNeRF360—bicycle*). The *Density* is computed according to Eq. (8), and the *Scale* corresponds to the volume of the Gaussian calculated using the scaling matrix. The ideal curve is obtained by fitting the result. The Gaussians of our methods are distributed around the ideal curve, while those of 3DGS show a more diffuse distribution.

Method – Metrics	SSIM \uparrow	PSNR \uparrow	LPIPS \downarrow	Gauss.Num. \downarrow
3DGS	0.813	27.48	0.221	3174k
+ D&S linking	0.827	27.65	0.178	5442k
+ D&S linking + dyna. thres.	0.831	27.75	0.178	3770k
+ D&S linking + conf. filter	0.836	27.97	0.175	3123k
Full	0.835	27.95	0.180	2370k

Table 2: Ablation study. We sequentially add scale-constraint (+ D&S linking), dynamic threshold for densification (+ dyna. thres.), and scale-based confidence filter (+ conf. filter) to demonstrate their effectiveness.

a frequency-aware density control mechanism to optimize density for accurate scene representation. Experimental results on various datasets show that FDS-GS performs better than existing 3DGS-based approaches.

Acknowledgements

This work is supported by the Natural Science Foundation of China under Grant 62302174. The computation is completed in the HPC Platform of Huazhong University of Science and Technology. We also thank Farsee2 Technology Ltd for providing devices to support the validation of our method.

References

- Barron, J. T.; Mildenhall, B.; Tancik, M.; Hedman, P.; Martin-Brualla, R.; and Srinivasan, P. P. 2021. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5855–5864.
- Barron, J. T.; Mildenhall, B.; Verbin, D.; Srinivasan, P. P.; and Hedman, P. 2022. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5470–5479.
- Barron, J. T.; Mildenhall, B.; Verbin, D.; Srinivasan, P. P.; and Hedman, P. 2023. Zip-NeRF: Anti-Aliased Grid-Based Neural Radiance Fields. *ICCV*.
- Bleyer, M.; Rhemann, C.; and Rother, C. 2011. Patchmatch stereo-stereo matching with slanted support windows. In *Bmvc*, volume 11, 1–11.
- Chen, A.; Xu, Z.; Geiger, A.; Yu, J.; and Su, H. 2022. Tensorf: Tensorial radiance fields. In *European Conference on Computer Vision*, 333–350. Springer.
- Chen, Y.; Chen, Z.; Zhang, C.; Wang, F.; Yang, X.; Wang, Y.; Cai, Z.; Yang, L.; Liu, H.; and Lin, G. 2023a. Gaussianeditor: Swift and controllable 3d editing with gaussian splatting. *arXiv preprint arXiv:2311.14521*.
- Chen, Z.; Funkhouser, T.; Hedman, P.; and Tagliasacchi, A. 2023b. Mobilenerf: Exploiting the polygon rasterization pipeline for efficient neural field rendering on mobile architectures. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16569–16578.
- Cheng, K.; Long, X.; Yang, K.; Yao, Y.; Yin, W.; Ma, Y.; Wang, W.; and Chen, X. 2024. GaussianPro: 3D Gaussian Splatting with Progressive Propagation. *arXiv preprint arXiv:2402.14650*.
- Chung, J.; Oh, J.; and Lee, K. M. 2023. Depth-regularized optimization for 3d gaussian splatting in few-shot images. *arXiv preprint arXiv:2311.13398*.
- Deng, K.; Liu, A.; Zhu, J.-Y.; and Ramanan, D. 2022a. Depth-supervised nerf: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12882–12891.
- Deng, N.; He, Z.; Ye, J.; Duinkharjav, B.; Chakravarthula, P.; Yang, X.; and Sun, Q. 2022b. Fov-nerf: Foveated neural radiance fields for virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11): 3854–3864.
- Franke, L.; Rückert, D.; Fink, L.; and Stamminger, M. 2024. TRIPS: Trilinear Point Splatting for Real-Time Radiance Field Rendering. *arXiv preprint arXiv:2401.06003*.
- Fridovich-Keil, S.; Yu, A.; Tancik, M.; Chen, Q.; Recht, B.; and Kanazawa, A. 2022. Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5501–5510.
- Garbin, S. J.; Kowalski, M.; Johnson, M.; Shotton, J.; and Valentin, J. 2021. Fastnerf: High-fidelity neural rendering at 200fps. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14346–14355.
- Grandits, T.; Effland, A.; Pock, T.; Krause, R.; Plank, G.; and Pezzuto, S. 2021. GEASI: Geodesic-based earliest activation sites identification in cardiac models. *International Journal for Numerical Methods in Biomedical Engineering*, 37(8): e3505.
- Hedman, P.; Philip, J.; Price, T.; Frahm, J.-M.; Drettakis, G.; and Brostow, G. 2018. Deep blending for free-viewpoint image-based rendering. *ACM Transactions on Graphics (ToG)*, 37(6): 1–15.
- Keetha, N.; Karhade, J.; Jatavallabhula, K. M.; Yang, G.; Scherer, S.; Ramanan, D.; and Luiten, J. 2023. Splatam: Splat, track & map 3d gaussians for dense rgb-d slam. *arXiv preprint arXiv:2312.02126*.
- Kerbl, B.; Kopanas, G.; Leimkühler, T.; and Drettakis, G. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 42(4).
- Knapitsch, A.; Park, J.; Zhou, Q.-Y.; and Koltun, V. 2017. Tanks and Temples: Benchmarking Large-Scale Scene Reconstruction. *ACM Transactions on Graphics*, 36(4).
- Kocabas, M.; Chang, J.-H. R.; Gabriel, J.; Tuzel, O.; and Ranjan, A. 2023. Hugs: Human gaussian splats. *arXiv preprint arXiv:2311.17910*.
- Kopanas, G.; Philip, J.; Leimkühler, T.; and Drettakis, G. 2021. Point-Based Neural Rendering with Per-View Optimization. In *Computer Graphics Forum*, volume 40, 29–43. Wiley Online Library.
- Li, H.; Liu, J.; Sznajder, M.; and Camps, O. 2024. 3D-HGS: 3D Half-Gaussian Splatting. *arXiv preprint arXiv:2406.02720*.
- Li, X.; Yi, W.; Chi, H.-L.; Wang, X.; and Chan, A. P. 2018. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in construction*, 86: 150–162.
- Lu, T.; Yu, M.; Xu, L.; Xiangli, Y.; Wang, L.; Lin, D.; and Dai, B. 2023. Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering. *arXiv:2312.00109*.
- Martin-Brualla, R.; Radwan, N.; Sajjadi, M. S.; Barron, J. T.; Dosovitskiy, A.; and Duckworth, D. 2021. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7210–7219.
- Matsuki, H.; Murai, R.; Kelly, P. H. J.; and Davison, A. J. 2024. Gaussian Splatting SLAM. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
- Mildenhall, B.; Srinivasan, P.; Tancik, M.; Barron, J.; Ramamoorthi, R.; and Ng, R. 2020. Nerf: Representing scenes

as neural radiance fields for view synthesis. In *European conference on computer vision*.

Müller, T.; Evans, A.; Schied, C.; and Keller, A. 2022. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. *ACM Trans. Graph.*, 41(4): 102:1–102:15.

Reiser, C.; Peng, S.; Liao, Y.; and Geiger, A. 2021. Kilonerf: Speeding up neural radiance fields with thousands of tiny mlps. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14335–14345.

Rückert, D.; Franke, L.; and Stamminger, M. 2022. Adop: Approximate differentiable one-pixel point rendering. *ACM Transactions on Graphics (ToG)*, 41(4): 1–14.

Schönberger, J. L.; and Frahm, J.-M. 2016. Structure-from-Motion Revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.

Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612.

Xu, L.; Agrawal, V.; Laney, W.; Garcia, T.; Bansal, A.; Kim, C.; Rota Bulò, S.; Porzi, L.; Kotschieder, P.; Božič, A.; et al. 2023a. VR-NeRF: High-fidelity virtualized walkable spaces. In *SIGGRAPH Asia 2023 Conference Papers*, 1–12.

Xu, M.; Zhan, F.; Zhang, J.; Yu, Y.; Zhang, X.; Theobalt, C.; Shao, L.; and Lu, S. 2023b. Wavenerf: Wavelet-based generalizable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 18195–18204.

Yu, Z.; Chen, A.; Huang, B.; Sattler, T.; and Geiger, A. 2024. Mip-splatting: Alias-free 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19447–19456.

Zhang, J.; Zhan, F.; Xu, M.; Lu, S.; and Xing, E. 2024. Fregs: 3d gaussian splatting with progressive frequency regularization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 21424–21433.

Zhang, K.; Riegler, G.; Snavely, N.; and Koltun, V. 2020. NeRF++: Analyzing and Improving Neural Radiance Fields. *arXiv:2010.07492*.

Zhang, Q.; Baek, S.-H.; Rusinkiewicz, S.; and Heide, F. 2022. Differentiable point-based radiance fields for efficient view synthesis. In *SIGGRAPH Asia 2022 Conference Papers*, 1–12.

Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; and Wang, O. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 586–595.

Zhou, X.; Lin, Z.; Shan, X.; Wang, Y.; Sun, D.; and Yang, M.-H. 2023. Drivinggaussian: Composite gaussian splatting for surrounding dynamic autonomous driving scenes. *arXiv preprint arXiv:2312.07920*.

Zwicker, M.; Pfister, H.; Van Baar, J.; and Gross, M. 2001. EWA volume splatting. In *Proceedings Visualization, 2001. VIS'01.*, 29–538. IEEE.