中图法分类号:TP18;TP391.4　　**文献标识码**:A　　**文章编号**:1006-8961(2026)09-3397-15

论文引用格式:Ma D L, Hui X T and Zhao H. 2026. Edge cluster-driven Gaussian splatting for detail enhancement in sparse view synthesis. Journal of Image and Graphics, 31(9):3397-3411(马栋林, 惠旭同, 赵宏. 2026. 边缘簇驱动的高斯泼溅稀疏视图重建细节增强. 中国图象图形学报, 31(9):3397-3411)[DOI:10.11834/jig.250538]

边缘簇驱动的高斯泼溅稀疏视图重建细节增强

马栋林,惠旭同*,赵宏

兰州理工大学计算机与人工智能学院,兰州　730050

摘　要:**目的**　稀疏视图下的新视角合成是三维视觉领域的关键挑战。现有方法多依赖复杂外部先验弥补信息缺失,在场景边缘等细节区域重建质量不佳。本文旨在探索一条不依赖外部先验、从表征内部结构入手的优化路径,提出一种边缘簇驱动的三维高斯泼溅方法,以实现稀疏输入下边缘细节增强的高质量视图合成。**方法**　本研究核心在于构建一个"识别—增强—约束"的内生优化框架。首先,引入聚类分析,将离散三维高斯点组织为语义簇,并提出多准则融合的边缘簇识别算法,综合评估簇中心梯度、平均梯度幅值、簇空间规模及簇间邻近度,精准定位表征边缘结构的关键簇。其次,设计局部邻近引导的高斯解池策略,在识别出的边缘簇内部依据局部邻近图智能增补高斯,针对性地强化边缘轮廓的连续性与完整性。最终,提出协同监督的深度正则化方法,通过合成伪视图增强数据,并融合基于SfM(structure from motion)点云尺度归一化的绝对深度损失与对尺度偏移不敏感的皮尔逊相关深度损失,为高斯位置优化提供鲁棒的几何约束。三者协同,形成从内部结构感知到针对性增强与几何一致性保障的闭环。**结果**　在Mip-NeRF360(unbounded anti-aliased neural radiance fields)和LLFF(local light field fusion)数据集共17个场景上的实验表明,在稀疏输入下,本文方法在峰值信噪比(peak signal-to-noise ratio,PSNR)、结构相似性指数(structural similarity index,SSIM)和学习感知图像块相似度(learned perceptual image patch similarity,LPIPS)指标上显著优于主流对比方法。在Mip-NeRF360数据集(1/8分辨率)上,3项指标平均达到23.82 dB、0.753和0.201;在LLFF数据集(1/8分辨率)上平均达到20.48 dB、0.662和0.268。在Deep Blending、Tanks&Temples和DTU(Technical University of Denmark robot image datasets)数据集上的实验进一步验证了方法的有效性与泛化能力,其性能达到或逼近当前最优方法。视觉对比显示,本文方法在自行车辐条、植物茎叶等复杂边缘的重建上更清晰锐利,有效减少模糊与断裂伪影。渲染速度保持在275帧/s以上,满足实时性要求。**结论**　本文提出的边缘簇驱动方法,通过内部结构感知与针对性增强,有效提升稀疏视图下三维高斯泼溅对场景边缘及各向异性结构的重建质量,为高质量、实时的新视角合成提供一条轻量化且自包含的技术路径。

关键词:三维高斯泼溅(3DGS);稀疏视图合成;边缘簇;聚类分析;深度正则化

Edge cluster-driven Gaussian splatting for detail enhancement in sparse view synthesis

Ma Donglin, Hui Xutong*, Zhao Hong

School of Computing and Artificial Intelligence, Lanzhou University of Technology, Lanzhou　730050, China

Abstract: **Objective**　Novel view synthesis from sparse views presents a critical and ill-posed challenge in 3D computer

收稿日期:2025-11-05;**修回日期**:2026-01-25;**预印本日期**:2026-02-01

*通信作者:惠旭同　logandahui@163.com

基金项目:国家自然科学基金项目(62166025)

Supported by:National Natural Science Foundation of China(62166025)

vision, primarily due to the severe underconstraint inherent in limited input imagery. Current state-of-the-art approaches can be broadly categorized into two streams: One relies on complex external prior models, such as monocular depth estimators or pretrained diffusion models, to compensate for the information deficit; the other, including methods such as few-shot Gaussian splatting (FSGS), operates within a more self-contained framework. However, a common limitation across these methods is their frequently inadequate reconstruction of detailed regions characterized by high-frequency geometry and thin structures, such as scene edges. Methods dependent on external priors can be misled by their generic nature, while self-contained strategies often lack targeted mechanisms in their densification and regularization processes to model these anisotropic areas effectively. To address these limitations, this study introduces a novel approach for sparse-view synthesis based on 3D Gaussian splatting, which is driven by the concept of edge clusters. The core idea is an "inside-out" optimization paradigm that deliberately avoids heavy external priors, instead seeking to mine and leverage structural information directly from the internal distribution of the 3D Gaussians. The objective is to achieve high-quality novel view synthesis with notably enhanced edge detail fidelity under challenging sparse-input conditions, exploring the potential of a fully self-contained and efficient framework. **Method** The proposed framework is built upon the standard 3D Gaussian splatting pipeline, integrating three key components tailored for sparse view optimization. First, the concept of clustering analysis is introduced into the optimization process. Candidate points for potential edge regions are selected on the basis of their average spatial position gradient norm. These points are subsequently partitioned into clusters using the K-means algorithm. A multicriteria fusion algorithm is then proposed to identify "edge clusters" critical for representing scene structures robustly. This algorithm evaluates four distinct criteria per cluster: cluster-center gradient, average gradient magnitude, cluster spatial scale, and intercluster proximity. A voting mechanism is employed to finalize the edge cluster set, enhancing robustness against the failure of any single criterion. Second, a locally proximity-guided Gaussian unpooling strategy is designed to address under-representation within the identified edge clusters. In contrast to global proximity guidance, this strategy constructs a local proximity graph within each edge cluster. For Gaussians exhibiting insufficient local coverage, new Gaussians are spawned along the connections to their K-nearest neighbors within the cluster, specifically at the midpoints. This approach strategically reinforces the continuity and geometric integrity of edge contours. Third, a collaboratively supervised depth regularization mechanism is introduced to provide stronger geometric constraints, which is particularly beneficial for newly added Gaussians in underconstrained edge regions. This mechanism involves synthesizing pseudoviews for data augmentation and imposing a hybrid depth loss. The hybrid loss combines an absolute depth loss, which uses monocular depth estimates normalized via SfM point clouds to address scale ambiguity, and a relative depth loss based on the Pearson correlation coefficient to preserve depth structure consistency. The depth loss weight follows an exponential decay schedule. The overall loss function integrates standard photometric losses with the proposed absolute and relative depth losses. **Result** Extensive experiments were conducted on the Mip-NeRF360, LLFF, Deep Blending, Tanks&Temples, and DTU datasets under standardized sparse-view settings. The proposed method was quantitatively compared with several state-of-the-art methods, including Mip-NeRF360, RegNeRF, SparseNeRF, vanilla 3DGS, and FSGS. Results demonstrate favorable performance of our approach. On the Mip-NeRF360 dataset at 1/8 resolution, our method achieved average PSNR, SSIM, and LPIPS values of 23.82 dB, 0.753, and 0.201, respectively, showing improvements over the FSGS baseline. On the LLFF dataset at 1/8 resolution, it achieved 20.48 dB, 0.662, and 0.268, indicating consistent gains. These trends were observed across most tested scenarios and at 1/4 resolution. Qualitatively, visual comparisons reveal that the proposed method successfully recovers fine edge details, such as bicycle spokes and plant stems, with notable clarity and continuity, whereas baseline methods often exhibit blurriness or fragmentation in these areas. The rendering speed remains high, exceeding 275 frames/s, confirming real-time capability. Ablation studies validate the complementary contributions of each proposed module, and further analysis confirms the robustness of the multicriteria clustering to its key parameters. **Conclusion** This study presents a novel edge cluster-driven approach for 3D Gaussian splatting that enhances sparse-view reconstruction quality through an internally structural-aware methodology. By leveraging clustering analysis for targeted edge cluster identification and augmentation, coupled with a collaboratively supervised regularization scheme, the method achieves measurable gains in quantitative metrics and improved visual fidelity for detail-rich regions. It provides a self-contained and efficient pathway for high-quality, real-time synthesis from sparse inputs.

Key words: 3D Gaussian splatting (3DGS); sparse view synthesis; edge clusters; clustering analysis; depth regularization

0 引　言

虚拟现实、增强现实及自动驾驶等三维视觉应用的快速发展，对高质量新视角合成技术提出迫切需求。该技术旨在从二维输入图像中重建三维场景，并在未知视角下渲染出具有照片级真实感的图像。然而，现有高性能方法通常严重依赖密集视图作为强先验（Wu等，2024b），这在实际应用中因成本或条件限制而难以满足，极大限制了其适用范围。

为平衡重建质量与效率，新视角合成领域先后涌现出隐式与显式两类代表性方法。神经辐射场（neural radiance field，NeRF）（Mildenhall等，2022）及其后续工作（Niemeyer等，2022）通过隐式场景表示实现出色的渲染质量，然而隐式表示的固有瓶颈与高昂计算成本推动领域向高效的显式表示探索。三维高斯泼溅（3d Gaussian splatting，3DGS）（Kerbl等，2023）的提出标志着一次重要的范式转变。该方法将场景建模为一系列显式、可学习的三维高斯椭球，并通过可微分泼溅渲染器进行光栅化，在保持高保真度的同时，实现训练速度的飞跃与实时渲染能力。这一显式范式迅速催生大量改进工作，主要集中于以下方向：1）模型压缩与渲染加速，如CompGS（Navaneet等，2024）利用矢量量化压缩高斯属性，LightGaussian（Fan等，2024）通过神经网络剪枝与蒸馏策略降低计算与存储开销；2）质量提升，MS3DGS（Yan等，2024）引入多尺度高斯泼溅表示以抑制锯齿，程龙昊等人（2025）通过解耦高斯的颜色与光照提升低光条件下的合成质量；3）3D生成与重建，Splatter Image（Szymanowicz等，2024）探索从单幅图像直接回归3D高斯表示，MVSplat（Chen等，2024）利用多视图立体（multi-view stereo，MVS）技术优化高斯初始化，DreamGaussian（Tang等，2024）结合扩散模型进行3D内容生成；4）动态场景建模，如4DGS（Wu等，2024a）将3DGS扩展至动态4D场景重建。

尽管3DGS效率卓越，但其性能高度依赖密集输入。在稀疏视图条件下，从运动恢复结构（structure from motion，SfM）获得的初始点云过于稀疏，导致高斯分布初始化不足，进而引发严重的过拟合与伪影，表现为整体纹理模糊与几何细节丢失。为应对这一挑战，现有研究多引入外部先验以弥补信息缺失，或采用正则化进行约束。在NeRF框架下，主流方法主要包括3类：1）几何监督，如Depth-NeRF（Deng等，2022）利用深度图作为额外监督，刘晓楠等人（2024）利用密集深度图缓解SfM不准确导致的质量下降，肖强等人（2024）通过深度监督与形变损失增强稀疏全景图的重建鲁棒性；2）正则化约束，如RegNeRF（regularizing neural radiance fields）（Niemeyer等，2022）在未观测视角施加深度平滑约束，SparseNeRF（Wang等，2023）提出空间连续性损失以规范空白区域优化；3）合成伪视图，如Deceptive-NeRF（Liu等，2024）利用预训练扩散模型生成高质量伪视图，以增强NeRF在稀疏视角下的训练稳定性。随着3DGS的兴起，稀疏视图合成方法沿袭类似思路并进一步发展：1）深度引导，如DepthRegGS（Chung等，2024）与DNGaussian（Li等，2024）分别引入单目深度作为软硬约束，SparseGS（Xiong等，2025）进一步结合深度先验、深度渲染与剪枝策略，系统抑制漂浮伪影并缓解背景坍塌；2）协同优化，如CoR-GS（Zhang等，2024）通过邻近视角协同正则化以提升几何一致性；3）初始化与修复，如Gaussian-Object（Yang等，2024a）通过视觉外壳建立初始高斯表示，并利用扩散模型修复缺失物体；4）伪视图增强，如Liu等人（2024）进一步将伪视图合成思路拓展至3DGS框架，提出deceptive-3DGS的方法，通过扩散模型生成的伪视角补充训练数据。FSGS（few-shot Gaussian splatting）（Zhu等，2024）是系统地将3DGS适配于极端稀疏视图的代表性工作，该方法通过邻近引导的高斯解池、深度正则化与伪视图合成，取得显著进展。

然而，包括FSGS在内的现有方法，在场景边缘及各向异性精细结构的重建上仍存在明显不足。其根本原因在于，通用的致密化与正则化策略缺乏对边缘区域的针对性感知与处理：稀疏视图导致边缘区域初始高斯覆盖不足；标准的致密化（Kerbl等，2023）或现有的邻近引导机制（Zhu等，2024）虽然能提升整体质量，但其全局性优化缺乏对局部几何结构的显式建模；此外，全局性正则化约束也难以对边

缘等高频细节区域的优化形成精确引导。这凸显出一个明确的研究空白：如何对场景内部结构，尤其是边缘区域，进行针对性感知与增强。

针对上述问题，本文探索了一条不依赖复杂外部先验的“由内而外”优化路径，提出一种边缘簇驱动的三维高斯泼溅方法。该方法的核心创新在于以边缘结构感知为指导，形成“识别—增强—约束”的协同优化流程。具体而言，首先通过无监督聚类将离散高斯点组织为语义“簇”，并利用多准则融合算法精准识别表征场景轮廓的“边缘簇”；在此基础上，针对识别出的边缘簇，设计局部邻近引导的高斯解池策略，依据局部邻近关系智能增补高斯，以弥补几何连续性缺口；同时，为引导新增高斯的合理放置并抑制伪影，引入一种协同监督的深度正则化方法，提供鲁棒的几何约束以确保一致性。上述环节环环相扣，共同构成一个针对边缘细节重建的完整优化体系，实现边缘细节的高质量重建。

本文主要贡献如下：1）提出一种新颖的边缘簇驱动三维高斯泼溅方法，首次将聚类分析思想引入稀疏视图下的高斯优化，通过内部结构感知显著提升边缘细节重建质量；2）提出多准则融合的边缘簇识别算法与局部邻近引导的高斯解池策略，实现对边缘区域高斯分布的精准定位与针对性增强，克服全局策略的局限性；3）提出一种协同监督的深度正则化方法，结合绝对与相对深度约束，为边缘区域优化提供鲁棒几何指导，有效缓解过拟合并提升几何一致性。

1 本文方法

本文提出一种边缘簇驱动的三维高斯泼溅方法，其整体框架如图1所示。给定一组稀疏输入图像，先通过SfM获取相机姿态与稀疏点云，并依此初始化三维高斯场景表示。核心优化流程依次为：1）通过聚类分析与多准则融合算法，自动定位表征场景边缘结构的关键高斯簇；2）在识别出的边缘簇内部执行局部邻近引导的高斯解池策略，于覆盖不足区域智能增补高斯，强化边缘结构的连续性与完整性；3）引入混合深度损失约束高斯椭球的位置优化，抑制过拟合并确保几何一致性。上述优化步骤与3DGS主训练循环交替进行，最终获得细节丰富且几何准确的高斯场景模型，实现高质量稀疏视图新视角合成。

1.1 边缘高斯簇的感知与识别

在3DGS初始化得到高斯场景表示后，本文核心思路是将离散高斯点依据其几何属性聚合为“簇”，并从中识别对边缘结构至关重要的“边缘簇”。边缘区域的高斯点需刻画几何快速变化，因此在训练中通常表现出较大视图空间位置梯度幅值（Kerbl

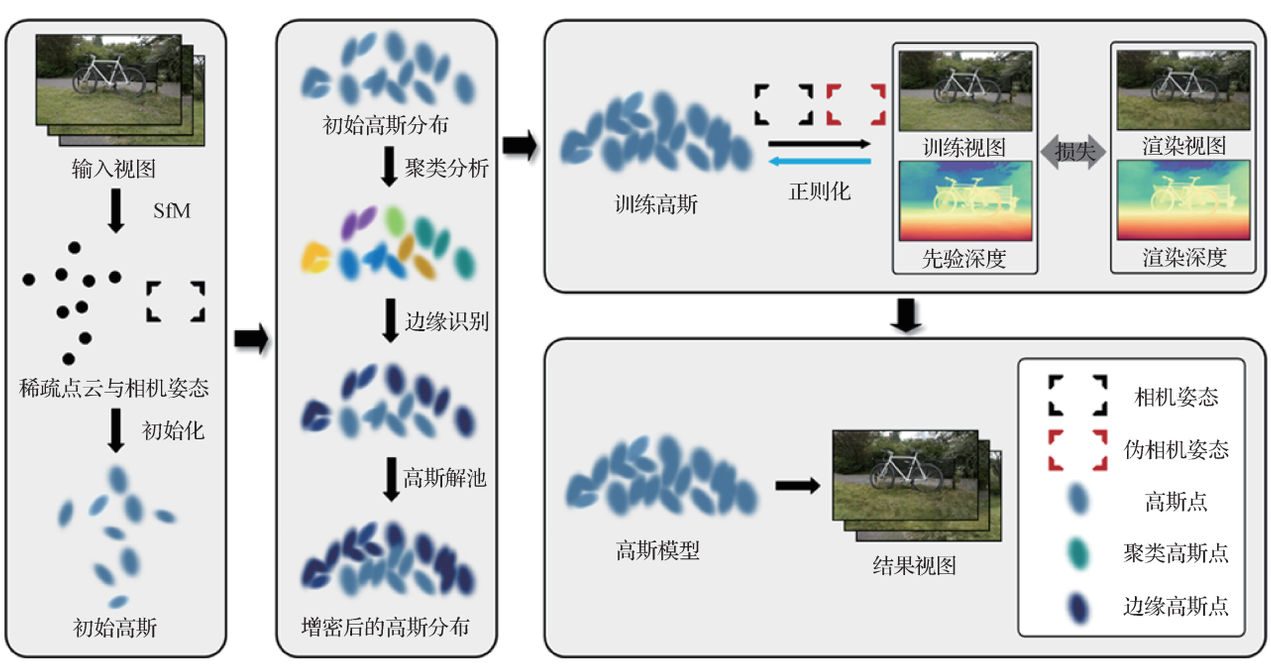


图1 边缘簇驱动的三维高斯泼溅方法整体框架

Fig. 1 The overall framework of the proposed edge cluster-driven 3D Gaussian splatting method

等，2023），这为边缘定位提供了有效线索。受点云分析中聚类识别结构（杨庆鑫 等，2025）启发，本文提出基于聚类的高斯边缘簇感知方法。

1. 1. 1　候选点筛选与高斯簇划分

首先，计算每个高斯点 $\boldsymbol{p}$ 在训练视图下的平均空间位置梯度范数 $\|\nabla f(\boldsymbol{p})\|$，筛选梯度幅值大于阈值 θ 的点，构成候选边缘高斯点集 $\boldsymbol{P}$。

随后，采用 K-means 算法将点集 $\boldsymbol{P}$ 划分为 k 个高斯簇 $\boldsymbol{C}=\{\boldsymbol{C}_1,\boldsymbol{C}_2,\cdots,\boldsymbol{C}_k\}$，算法优化目标为最小化簇内误差平方和（sum of squared error，SSE），计算为

$$SSE=\sum_{i=1}^{k}\sum_{\boldsymbol{x}\in\boldsymbol{C}_i}\left\|\boldsymbol{x}-\boldsymbol{\mu}_i\right\|_2^2 \tag{1}$$

式中，$\boldsymbol{x}$ 为高斯点位置向量，$\boldsymbol{\mu}_i$ 是簇 $\boldsymbol{C}_i$ 的均值中心（质心），计算为

$$\boldsymbol{\mu}_i=\frac{1}{|\boldsymbol{C}_i|}\sum_{\boldsymbol{x}\in\boldsymbol{C}_i}\boldsymbol{x} \tag{2}$$

SSE 值越小，表明簇内点空间聚集性越强，越能代表连续几何结构片段。

1. 1. 2　多准则边缘簇识别

完成簇划分后，需从 $\boldsymbol{C}$ 中精准识别边缘簇 $\boldsymbol{E}$。基于边缘结构在梯度、规模及空间关系上的特性，本文提出一种多准则融合识别算法，从 4 个维度综合评估各簇的“边缘性”。

1）簇中心梯度准则。梯度信息是经典图像边缘检测的核心依据（Canny，1986）。场景边缘的几何或颜色剧变会使该区域高斯点具有较高空间梯度。簇中心梯度 $\boldsymbol{g}(\boldsymbol{c}_i)$ 表征该簇区域整体变化趋势，是判断边缘区域的强特征。计算簇 $\boldsymbol{C}_i$ 几何中心 $\boldsymbol{c}_i$（即式（2）中 $\boldsymbol{\mu}_i$ 处的平均梯度）作为其中心梯度，计算为

$$\boldsymbol{g}(\boldsymbol{c}_i)=\frac{1}{|\boldsymbol{C}_i|}\sum_{\boldsymbol{p}\in\boldsymbol{C}_i}\nabla f(\boldsymbol{p}) \tag{3}$$

式中，$\nabla f(\boldsymbol{p})$ 是点 $\boldsymbol{p}$ 的视图空间位置梯度。若 $\|\boldsymbol{g}(\boldsymbol{c}_i)\|$ 超过阈值 θ_g，则判定该簇 $\boldsymbol{C}_i$ 处于变化剧烈区域。

2）平均梯度幅值准则。为避免簇内梯度方向抵消导致误判，补充计算簇内所有点的平均梯度幅值，计算为

$$\bar{\boldsymbol{g}}_i=\frac{1}{|\boldsymbol{C}_i|}\sum_{\boldsymbol{p}\in\boldsymbol{C}_i}\left\|\nabla f(\boldsymbol{p})\right\| \tag{4}$$

较高的平均梯度幅值（$>\theta_{\text{avg}}$）表明，簇内含大量变化剧烈的点。

3）簇空间规模准则。在形态学处理中，有意义的边缘结构通常具有连续且成一定规模的几何特征（Soille，1999），过小的簇更可能为噪声。本文引入簇规模度量，计算所有簇的平均大小，具体为

$$\bar{s}=\frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{C}_i| \tag{5}$$

若簇大小 $|\boldsymbol{C}_i|$ 显著大于该平均值，则认为其更能代表显著边缘结构。

4）簇间邻近度准则。受 FSGS 利用邻近度控制高斯密度的启发（Zhu 等，2024），本文将类似思想从点级别推广至簇级别。属于同一语义边缘的要素在空间上往往邻近且连续（Felzenszwalb 和 Huttenlocher，2004），因此同一边缘的簇间距离应较小。对于簇 $\boldsymbol{C}_i$，计算其质心 $\boldsymbol{\mu}_i$ 到其他所有簇质心的平均距离作为簇邻近度得分，计算为

$$Proximity(\boldsymbol{C}_i)=\frac{1}{N-1}\sum_{j\neq i}\left\|\boldsymbol{\mu}_i-\boldsymbol{\mu}_j\right\| \tag{6}$$

较低的平均距离（$Proximity(\boldsymbol{C}_i)<\theta_{\text{proximity}}$）表明，该簇周围存在其他簇，属于连续边缘结构的一部分。

最终边缘簇集合 $\boldsymbol{E}$ 通过投票机制确定：若某簇 $\boldsymbol{C}_i$ 满足至少 k_{vote} 项上述准则，则判定为边缘簇。该多准则融合策略有效提升了识别鲁棒性，避免了单一准则失效导致的误判或漏检。

1. 2　局部邻近引导的高斯解池

识别边缘簇 $\boldsymbol{E}$ 后，需对其进行针对性致密化以弥补稀疏视图带来的几何细节缺失。3DGS 的传统自适应密度控制（Kerbl 等，2023）及 FSGS（Zhu 等，2024）采用的邻近引导高斯解池策略在边缘区域存在局限：其分裂与克隆策略缺乏对局部几何关系的显式建模。为此，本文在边缘簇内部设计一种局部邻近引导的高斯解池法，通过构建局部邻近图智能引导新高斯生成，以强化边缘结构的连续性与完整性，如图 2 所示。

1）局部邻近图构建。对于边缘簇 $\boldsymbol{E}$ 中每个高斯

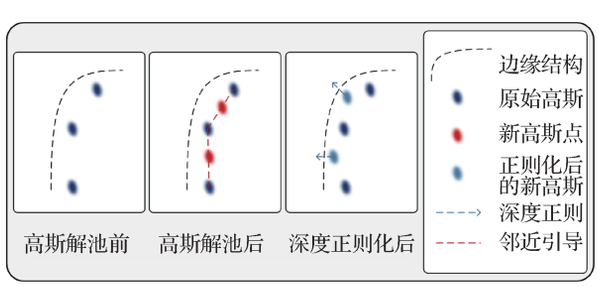


图 2　局部邻近引导的高斯解池策略示意图

Fig. 2　Illustration of the locally proximity-guided Gaussian unpooling strategy

点$\boldsymbol{p}_i$，确定其K个最近邻点，构建局部有向邻近图，具体为

$$\boldsymbol{D}_i^K = K - \min(d_{ij}) \tag{7}$$

式中，对于每个高斯点$\boldsymbol{p}_i$计算其到簇内其他点的欧氏距离$d_{ij}=\|\boldsymbol{\mu}_i-\boldsymbol{\mu}_j\|$。点$\boldsymbol{p}_i$的邻近度得分$P_i$定义为其与$K$个近邻的平均距离，具体为

$$P_i = \frac{1}{K}\sum_{j=1}^{K} D_i^K \tag{8}$$

优化过程中，簇内邻近度图在增密或修剪后更新。

2）边缘簇的高斯解池。对于边缘簇内每个高斯点$\boldsymbol{p}_i$，若其邻近度得分P_i超过预设阈值t_{prox}，则表明该点所处区域覆盖不足。此时，在该点与其K个近邻点间的连线上，分别生成一个新高斯点。新高斯位置$\boldsymbol{p}_{\text{new}}$位于连线中心，具体为

$$\boldsymbol{p}_{\text{new}} = \frac{\boldsymbol{\mu}_i + \boldsymbol{\mu}_j}{2} \tag{9}$$

新高斯的尺度属性$\boldsymbol{S}_{\text{new}}$与不透明度$\alpha_{\text{new}}$继承自“目的地”高斯$\boldsymbol{p}_j$，旋转、球谐系数等属性则初始化为零。该策略确保新增高斯点有效放置在现有高斯分布稀疏的区域，从而精准增强边缘结构重建。

1.3 协同监督的深度正则化

在3DGS传统训练中，新增高斯的位置优化缺乏强几何约束，在稀疏视图下易因监督信号不足而陷入局部最优，导致几何结构失真。FSGS（Zhu等，2024）虽然通过伪视图合成与深度正则化部分缓解了该问题，但其对边缘区域优化仍显不足。为此，本文提出一种协同监督的深度正则化策略，以更有效地约束高斯位置优化。

1.3.1 伪视图合成

为增强训练视角多样性，本文采用伪视图合成模拟未观测视点。每次训练迭代时，在欧氏空间中选择距离最近的两个训练视图，计算其平均相机旋转四元数$\boldsymbol{q}$，并对相机位置$\boldsymbol{t}$施加随机噪声扰动$\boldsymbol{\varepsilon}$，以生成虚拟相机参数，计算式为

$$\boldsymbol{P}' = (\boldsymbol{t} + \boldsymbol{\varepsilon}, \boldsymbol{q}), \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \delta) \tag{10}$$

式中，δ控制噪声强度。

1.3.2 几何驱动的深度损失

本文设计一种结合绝对深度约束与相对结构约束的混合正则化方案，为高斯优化提供更鲁棒的几何引导。

1）绝对深度损失。为解决单目深度估计器的尺度模糊性（Ranftl等，2021），本文采用COLMAP（Schönberger和Frahm，2016）提供的SfM点云对估计深度进行尺度归一化。首先计算深度图$\boldsymbol{D}$的尺度$t(\boldsymbol{D})$和偏移$s(\boldsymbol{D})$，具体为

$$t(\boldsymbol{D}) = median(\boldsymbol{D}) \tag{11}$$

$$s(\boldsymbol{D}) = \frac{1}{M}\sum_{i \in \boldsymbol{SFM}} \left| \boldsymbol{D}_i - t(\boldsymbol{D}) \right| \tag{12}$$

式中，$\boldsymbol{SFM}$为SfM点的索引集合，M是SfM点数量。进而将单目深度估计值$\boldsymbol{D}$归一化至真实尺度，具体为

$$\hat{\boldsymbol{D}}_{\text{est}} = \frac{s(\boldsymbol{D}_{\text{SFM}})}{s(\boldsymbol{D})}\boldsymbol{D} + t(\boldsymbol{D}_{\text{SFM}}) - t(\boldsymbol{D})\frac{s(\boldsymbol{D}_{\text{SFM}})}{s(\boldsymbol{D})} \tag{13}$$

通过对各高斯点中心位置处的深度d_i进行Alpha混合来渲染深度图，具体为

$$\boldsymbol{D}_{\text{ras}} = \sum_{i=1}^{N} T_i \alpha_i d_i \tag{14}$$

同时，定义绝对深度损失为

$$L_{\text{D}} = \left| \boldsymbol{D}_{\text{ras}} - \hat{\boldsymbol{D}}_{\text{est}} \right| \tag{15}$$

2）相对深度损失。为增强深度结构一致性，引入皮尔逊相关系数作为相对深度损失，该损失对全局尺度和偏移不敏感，侧重于保持深度结构的相对关系，计算式为

$$Corr(\boldsymbol{D}_{\text{ras}}, \hat{\boldsymbol{D}}_{\text{est}}) = \frac{Cov(\boldsymbol{D}_{\text{ras}}, \hat{\boldsymbol{D}}_{\text{est}})}{\sqrt{Var(\boldsymbol{D}_{\text{ras}})Var(\hat{\boldsymbol{D}}_{\text{est}})}} \tag{16}$$

$$L_{\text{Pearson}} = 1 - Corr(\boldsymbol{D}_{\text{ras}}, \hat{\boldsymbol{D}}_{\text{est}}) \tag{17}$$

为进一步降低深度先验中的噪声影响，深度损失总体权重采用指数衰减策略，具体为

$$W = W_0 \cdot \exp(-k \cdot iter) \tag{18}$$

式中，W_0为初始权重，k为衰减系数，$iter$为当前训练迭代次数。

1.4 损失函数

本文方法的整体损失函数L由以下4部分组成，共同指导模型优化，具体为

$$L = \lambda_1 L_1 + \lambda_2 L_{\text{SSIM}} + \lambda_3 L_{\text{D}} + \lambda_4 L_{\text{Pearson}} \tag{19}$$

式中，L_1和L_{SSIM}为源自3DGS的光度重建损失，用于保证颜色与纹理的渲染质量。L_{D}为式（15）定义的绝对深度损失，通过强约束促使渲染深度与归一化估计深度数值一致，为高斯位置提供精确几何监督。L_{Pearson}为式（17）定义的皮尔逊相关损失，通过弱约束确保渲染与估计深度在结构上相似，增强对复杂几何结构的重建鲁棒性。该组合损失充分利用不同监

督信号的互补优势：L_{D}保障几何重建的绝对精度，L_{Pearson}提升对边缘结构的适应性，二者共同引导高斯点优化至合理空间位置，从而在稀疏输入下实现更准确、可靠的场景重建。

2　实验设计与结果分析

2.1　数据集

为全面评估算法在稀疏视图下的性能，本文实验在5个广泛使用的真实场景数据集上进行。Mip-NeRF360（unbounded anti-aliased neural radiance fields）（Barron 等，2022）与 LLFF（local light field fusion）（Mildenhall 等，2019）作为核心基准，用于评估新视角合成的综合质量；为验证方法泛化能力与边缘重建优势，补充 Deep Blending（Hedman 等，2018）、Tanks&Temples（Knapitsch 等，2017）和 DTU（Technical University of Denmark robot image datasets）（Jensen 等，2014）3个各具挑战的数据集。这些数据集涵盖有界/无界、室内/室外、物体级/场景级等多种类型，包含丰富几何细节与复杂纹理，可系统检验方法的有效性、鲁棒性与泛化性。

1）Mip-NeRF360数据集。包含9个复杂大规模无界场景。稀疏视图下，其复杂几何、严重遮挡与广阔空间范围对算法鲁棒性与细节重建能力构成严苛考验，适合验证方法处理极端稀疏输入与复杂场景的能力。

2）LLFF数据集。包含8个有界前向场景，视角变化相对受限。实验遵循极度稀疏设定，旨在验证方法在信息量有限条件下维持场景整体结构与合成质量的鲁棒性。

3）Deep Blending数据集。侧重具有复杂真实光照与反射材质的室内场景，含大量光滑表面、透明物体及细节装饰，对高频光影、镜面高光与物体边缘的精确重建挑战显著，适合验证方法在复杂光照下保持细节与边缘的能力。

4）Tanks&Temples数据集。包含大规模室内外场景，具有复杂几何结构、重复纹理和深远景深。其大范围场景内容与剧烈视角变化对算法几何一致性、大尺度场景理解及大视角外推能力要求较高，适合评估方法在视角差距较大时的合成效果。

5）DTU数据集。作为高精度多视图立体重建基准，提供高分辨率图像、精确相机参数及真实三维扫描模型，支持基于几何真值的定量对比，适用于评估稀疏视图下恢复准确几何结构的能力。

所有数据集的划分与预处理均遵循FSGS（Zhu 等，2024）标准流程。对于 Mip-NeRF360、Deep Blending 和 Tanks&Temples，每隔8幅图像选取1幅作为测试集，并从剩余图像中均匀采样24幅作为训练视图。对于LLFF，测试集划分方式同上，训练集固定选取3幅视图。对于DTU数据集，遵循RegNeRF（Niemeyer 等，2022）的实验设置。

2.2　评估指标

本文采用新视角合成领域公认的3个全参考图像质量评估指标：峰值信噪比（peak signal-to-noise ratio，PSNR）、结构相似性指数（structural similarity index，SSIM）和学习感知图像块相似度（learned perceptual image patch similarity，LPIPS），从像素精度、结构相似性和感知质量3个维度评估重建效果。

PSNR值越高代表像素级重建精度越高。SSIM衡量两幅图像在亮度、对比度和结构上的相似性，取值范围为[0，1]，值越接近1表示结构一致性越好。LPIPS通过预训练深度神经网络提取特征并计算感知差异，更贴近人类视觉判断，值越小表示感知质量越好。此外，本文以帧率（frames per second，FPS）为单位评估渲染速度，衡量方法在特定硬件上的实时渲染能力，FPS值越高代表渲染效率越高。

2.3　实验细节

所有实验均在NVIDIA A6000 GPU上完成，基于PyTorch框架实现，使用Adam优化器。三维高斯初始化所需相机姿态与稀疏点云由COLMAP（Schönberger 和 Frahm，2016）从输入视图计算得到。关键参数与训练策略如下：

1）致密化。边缘簇致密化每100次迭代执行一次，训练开始2 000次迭代后停止，邻近值$K=3$，邻近度得分预设阈值$t_{\mathrm{prox}}=10$。

2）优化。总迭代次数为10 000，500次迭代后开始对冗余高斯进行剪枝。

3）深度监督。采用Depth Anything V2（Yang 等，2024b）模型提供深度先验，2 000次迭代后启用深度损失，损失权重指数衰减的初始权重$W_0=0.1$，衰减系数$k=0.000\,1$。

4）损失权重。损失函数各分量权重设为$\lambda_1=0.8$，$\lambda_2=0.2$，$\lambda_3=0.05$，$\lambda_4=0.05$。

2.4 实验结果分析

2.4.1 对比实验

本文方法在Mip-NeRF360、LLFF、Deep Blending、Tanks&Temples和DTU数据集上与多种前沿方法进行比较，包括Mip-NeRF360（Barron等，2022）、3DGS（Kerbl等，2023）、RegNeRF（Niemeyer等，2022）和SparseNeRF（Wang等，2023），并与基于3DGS的稀疏视图合成方法FSGS（Zhu等，2024）及SparseGS（Xiong等，2025）重点对比。所有对比方法在相同稀疏输入视图、相机参数及SfM初始化点云下训练与测试。定量结果如表1—表3所示。

分析可知，本文方法在所有数据集和分辨率下，在PSNR、SSIM和LPIPS 3个核心指标上均取得一致且显著的提升。例如，在Mip-NeRF360数据集（1/8分辨率）上，与FSGS相比，PSNR从23.70 dB提升至23.82 dB，SSIM从0.745提升至0.753，LPIPS从0.220降低至0.201。这表明在相同输入信息下，本文提出的边缘簇驱动策略能更有效利用有限数据提升重建质量。效率方面，本文方法的FPS相比FSGS略有下降，但在1/8分辨率Mip-NeRF360数据集上仍保持较高渲染速度（= 275 帧/s），远高于实时渲染通用标准，体现重建质量与渲染效率间的良好平衡。

定性结果如图3所示，进一步印证了定量分析结论。可观察到，3DGS在稀疏视图下重建效果差，产生大量模糊和浮游伪影；FSGS能重建基本场景结构，但在边缘区域仍存在明显模糊、断裂或丢失；SparseGS整体质量更优，但对边缘区域仍缺乏有效重建手段。相比之下，本文方法能更精准重建这些纤细复杂边缘结构，恢复更清晰锐利的轮廓，视觉质量提升显著。

表1 Mip-NeRF360数据集上的新视角合成定量结果对比

Table 1 Quantitative comparison of novel view synthesis on the Mip-NeRF360 dataset

方法	1/8分辨率				1/4分辨率			
	FPS/(帧/s)↑	PSNR/dB↑	SSIM↑	LPIPS↓	FPS/(帧/s)↑	PSNR/dB↑	SSIM↑	LPIPS↓
Mip-NeRF360	0.12	21.23	0.613	0.351	0.07	19.78	0.530	0.431
3DGS	223	20.89	0.633	0.317	145	19.93	0.588	0.401
RegNeRF	0.07	22.19	0.643	0.335	0.04	20.55	0.546	0.398
SparseNeRF	0.07	22.85	0.693	0.315	0.04	21.13	0.600	0.389
FSGS	**290**	23.70	0.745	0.220	**203**	22.82	0.693	0.293
SparseGS	–	23.80	**0.759**	0.216	–	**23.02**	**0.713**	0.290
本文	275	**23.82**	0.753	**0.201**	193	22.95	0.701	**0.275**

注：加粗字体表示各列最优结果。“–”表示该指标未进行测试。“↑”表示值越大越好，“↓”表示值越小越好。

表2 LLFF数据集上的新视角合成定量结果对比

Table 2 Quantitative comparison of novel view synthesis on the LLFF dataset

方法	1/8分辨率				1/4分辨率			
	FPS/(帧/s)↑	PSNR/dB↑	SSIM↑	LPIPS↓	FPS/(帧/s)↑	PSNR/dB↑	SSIM↑	LPIPS↓
Mip-NeRF360	0.21	16.11	0.401	0.460	0.14	15.22	0.351	0.540
3DGS	385	17.43	0.522	0.321	312	16.94	0.488	0.402
RegNeRF	0.21	19.08	0.587	0.336	0.14	18.06	0.535	0.411
SparseNeRF	0.21	19.86	0.624	0.328	0.14	19.07	0.564	0.401
FSGS	**458**	20.31	0.652	0.288	**351**	19.88	0.612	0.340
SparseGS	–	20.39	**0.706**	0.271	–	19.86	**0.668**	0.322
本文	435	**20.48**	0.662	**0.268**	333	**20.05**	0.623	**0.320**

注：加粗字体表示各列最优结果。“–”表示该指标未进行测试 。“↑”表示值越大越好，“↓”表示值越小越好。

表 3　Deep Blending、Tanks&Temples 和 DTU 数据集上的新视角合成定量结果对比

Table 3　Quantitative comparison of novel view synthesis on Deep Blending, Tanks&Temples and DTU datasets

方法	Deep Blending 数据集			Tanks&Temples 数据集			DTU 数据集		
	PSNR/dB ↑	SSIM ↑	LPIPS ↓	PSNR/dB ↑	SSIM ↑	LPIPS ↓	PSNR/dB ↑	SSIM ↑	LPIPS ↓
Mip-NeRF360	20.13	0.655	0.521	17.07	0.562	0.425	16.71	0.757	0.239
3DGS	19.71	0.672	0.485	16.53	0.586	0.391	14.18	0.628	0.301
RegNeRF	21.08	0.688	0.505	17.87	0.593	0.416	18.89	0.745	0.190
SparseNeRF	21.70	0.733	0.481	18.55	0.640	0.385	**19.55**	0.769	0.201
FSGS	22.81	0.796	0.375	19.60	0.703	0.280	18.36	0.707	0.232
SparseGS	22.90	**0.809**	0.371	19.75	**0.720**	0.272	18.89	**0.834**	**0.178**
本文	**22.98**	0.802	**0.362**	**20.03**	0.712	**0.261**	18.81	0.751	0.198

注：加粗字体表示各列最优结果。“↑”表示值越大越好，“↓”表示值越小越好。

2.4.2　消融实验

本文通过消融实验系统评估各核心模块贡献。如表 4 与图 4 所示，在 Mip-NeRF360 数据集（1/8 分辨率）上，以 FSGS 作为基准模型（实验 A），依次引入本文核心模块进行分析。

单独引入局部邻近引导的高斯解池（实验 B）后，PSNR 提升至 23.75 dB，SSIM 提升至 0.748，LPIPS 降至 0.215。视觉上边缘区域密度增加，但因几何约束不足而伴随不规则噪声点。

单独引入协同监督深度正则化（实验 C）后，PSNR 提升至 23.72 dB，SSIM 提升至 0.746，LPIPS 降至 0.218。视觉上整体结构更清晰规整，但边缘细节重建有限，难以主动创造高频特征。

当两个模块协同工作时（实验 D），所有定量指标均达最优：PSNR 提升至 23.82 dB，SSIM 提升至 0.753，LPIPS 降至 0.201。视觉上同时实现丰富的边缘细节和清晰准确的几何结构，有效消除了单一模块的缺陷。该实验可验证模块间功能互补性：解池策略负责细节增强，正则化策略保障几何合理，二者协同显著提升重建质量。

2.4.3　多准则融合策略分析

本节验证多准则融合策略相对于单一准则的优越性。为清晰展示核心结论，选择最具代表性的两类准则对比：反映局部变化的梯度准则（包含簇中心梯度与平均梯度幅值）和反映空间结构的空间准则（包含簇规模与邻近度）。结果如表 5 所示。

分析可知，任何单一准则的引入均能带来性能提升，但四准则融合策略实现最佳性能，证明多准则融合能更鲁棒地识别边缘结构，避免单一准则局限性。

2.4.4　极端视角差下的效果分析

为验证方法在训练与测试视角间存在极端差异条件下的合成鲁棒性与泛化能力，本实验在 Deep Blending 数据集中选取视角变化剧烈的测试序列，构建相机位置跨度大、与训练视角差异显著的测试集。在此挑战性设定下，对比本文方法与 3DGS、FSGS 的合成效果。

实验结果如图 5 所示。极端视角差导致所有方法重建质量下降，整体结构出现模糊，局部细节丢失明显。但在边缘结构保持方面，本文方法重建结果展现出更完整的几何连续性与更少的视觉断裂，尤其在轮廓清晰度上优于对比方法。

定量结果如表 6 所示，在极端视角差条件下，所有方法的 PSNR 与 SSIM 指标均下降，LPIPS 值上升。尽管如此，本文方法在 3 项指标上仍优于 FSGS 与 3DGS，表明其在应对大幅视角变化时，仍能保持更优的感知质量与结构一致性。

2.4.5　重建质量的区域差异性分析

为验证本文方法在针对性增强边缘细节的同时，是否保持平滑区域原有的高质量重建，本实验选取同时包含大面积平滑表面与丰富边缘结构的场景进行对比分析。定性对比如图 6 所示。

在平滑区域，本文方法与 FSGS 渲染结果视觉上基本一致，未出现明显纹理失真或结构扭曲；在边缘区域，本文方法呈现更清晰、连续的轮廓结构，而 FSGS 存在模糊与几何断裂。该对比表明，方法的优

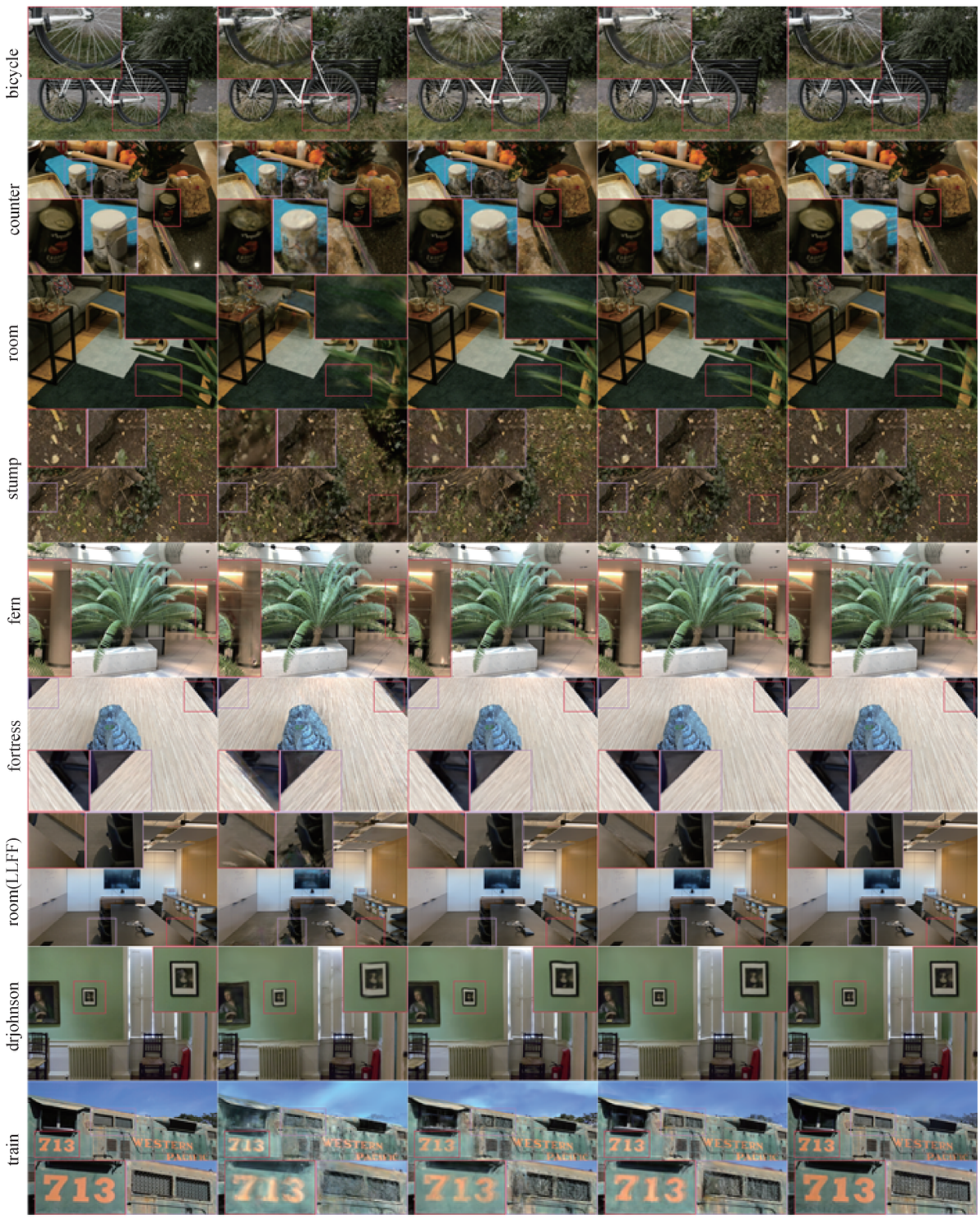


图3　不同方法在稀疏视图下的新视角合成视觉对比

Fig. 3　Visual comparison of novel view synthesis under sparse views by different methods

((a) ground truth; (b) 3DGS; (c) FSGS; (d) SparseGS; (e) ours)

化集中于待增强的边缘区域。

定量分析如表7所示。本文提取部分平滑区域局部图像块，计算FSGS与本文方法在PSNR、SSIM与LPIPS指标上的绝对差值。表7数据显示，所有指标差异均极小，表明本文方法在显著提升边缘质量的同时，完全保持了与基线模型相当的平滑区域重建精度，其增强策略具有区域性。

表 4　Mip-NeRF360 数据集上对各核心模块的消融实验

Table 4　Ablation study on the core modules on the Mip-NeRF360 dataset

实验	模型	Mip-NeRF360 数据集（1/8 分辨率）			
		FPS/(帧/s)	PSNR/dB	SSIM	LPIPS
A	FSGS	**290**	23.70	0.745	0.220
B	w/ 高斯解池	280	23.75	0.748	0.215
C	w/ 深度正则	285	23.72	0.746	0.218
D	本文	275	**23.82**	**0.753**	**0.201**

注：加粗字体表示各列最优结果。w/ X 代表单独启用 X 模块。



(a) FSGS　(b) w/ 高斯解池　(c) w/ 深度正则　(d) 本文

图 4　消融实验视觉结果对比

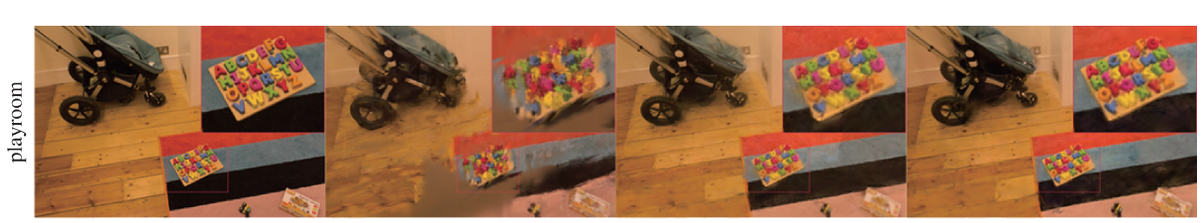
Fig. 4　Visual comparison of ablation studies ((a) FSGS; (b) w/ Gaussian unpooling; (c) w/ depth regularization; (d) ours)

表 5　Mip-NeRF360 数据集上对多准则融合策略的消融实验

Table 5　Ablation study on multi-criteria fusion strategies on the Mip-NeRF360 dataset

方法	FPS/(帧/s)	PSNR/dB	SSIM	LPIPS
FSGS	**290**	23.70	0.745	0.220
仅梯度准则	285	23.80	0.750	0.210
仅空间准则	288	23.75	0.748	0.215
四准则融合	275	**23.82**	**0.753**	**0.201**

注：加粗字体表示各列最优结果。



(a) 真实场景　(b) 3DGS　(c) FSGS　(d) 本文

图 5　极端视角差下的新视角合成视觉对比

Fig. 5　Visual comparison of novel view synthesis under extreme viewpoint differences ((a) ground truth; (b) 3DGS; (c) FSGS; (d) ours)

表 6　极端视角差下的定量结果对比

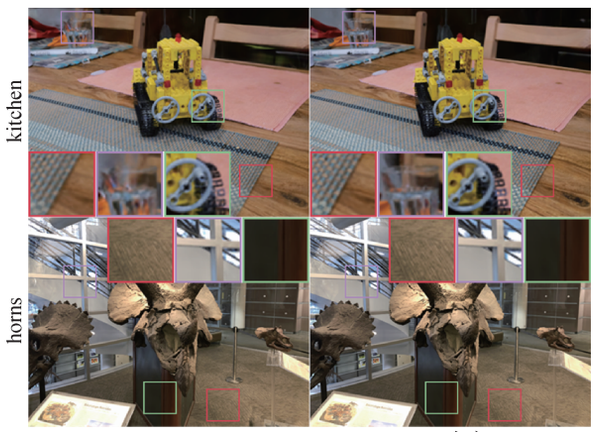
Table 6　Quantitative comparison under extreme viewpoint differences

方法	PSNR/dB	SSIM	LPIPS
3DGS	19.29	0.656	0.485
FSGS	22.30	0.789	0.373
本文	**22.59**	**0.796**	**0.363**

注：加粗字体表示各列最优结果。

2.4.6　参数敏感性分析与确定

本文方法在边缘簇识别模块涉及聚类数 k、投票阈值 k_{vote}、梯度阈值 θ_{g}、平均梯度阈值 θ_{avg} 以及邻近度阈值 $\theta_{\text{proximity}}$ 共 5 项关键参数。为系统验证算法对参数设置的鲁棒性，在 Mip-NeRF360 与 LLFF 数据集上进行参数敏感性分析。实验在 1/8 分辨率下采用控制变量法，依次在上述参数合理波动范围内测试，记录 PSNR、SSIM 和 LPIPS 指标变化。



(a) FSGS (b) 本文

图6 平滑区域与边缘区域渲染质量的视觉效果对比

Fig. 6 Visual comparison of rendering quality in smooth versus edge regions ((a) FSGS; (b) ours)

参数敏感性分析曲线如图7和图8所示，分别展示了两个数据集的5个关键参数，在对应测试范围内变化时3项评估指标的响应曲线。可以观察到，所有参数在各自合理区间内波动时，3项指标均保持平稳，未出现剧烈振荡或性能陡降，初步表明算法对参数变化不敏感。

表7 平滑区域重建质量差异分析

Table 7 Analysis of reconstruction quality difference in smooth regions

数据集	ΔPSNR/dB	ΔSSIM	ΔLPIPS
Mip-NeRF360	0.03	0.006	0.003
LLFF	0.02	0.003	0.001

注：Δ值表示平滑区域内本文方法与FSGS对应指标的绝对差值。

为定量评估鲁棒性，表8汇总了各参数在测试范围内引起的最大性能波动。结果显示，在Mip-NeRF360数据集上，所有参数引起的PSNR波动未超过0.03 dB，SSIM波动未超过0.003，LPIPS波动未超过0.002；在LLFF数据集上，波动范围同样维持在极低水平。如此微小波动表明，算法性能提升主要源于整体框架设计，而非对个别参数的过度依赖，显示了较强的鲁棒性与实用性。

基于上述系统参数鲁棒性分析，确定了边缘簇识别各关键参数在取得最佳性能时的取值，如表9所示，该组参数作为方法默认设置。

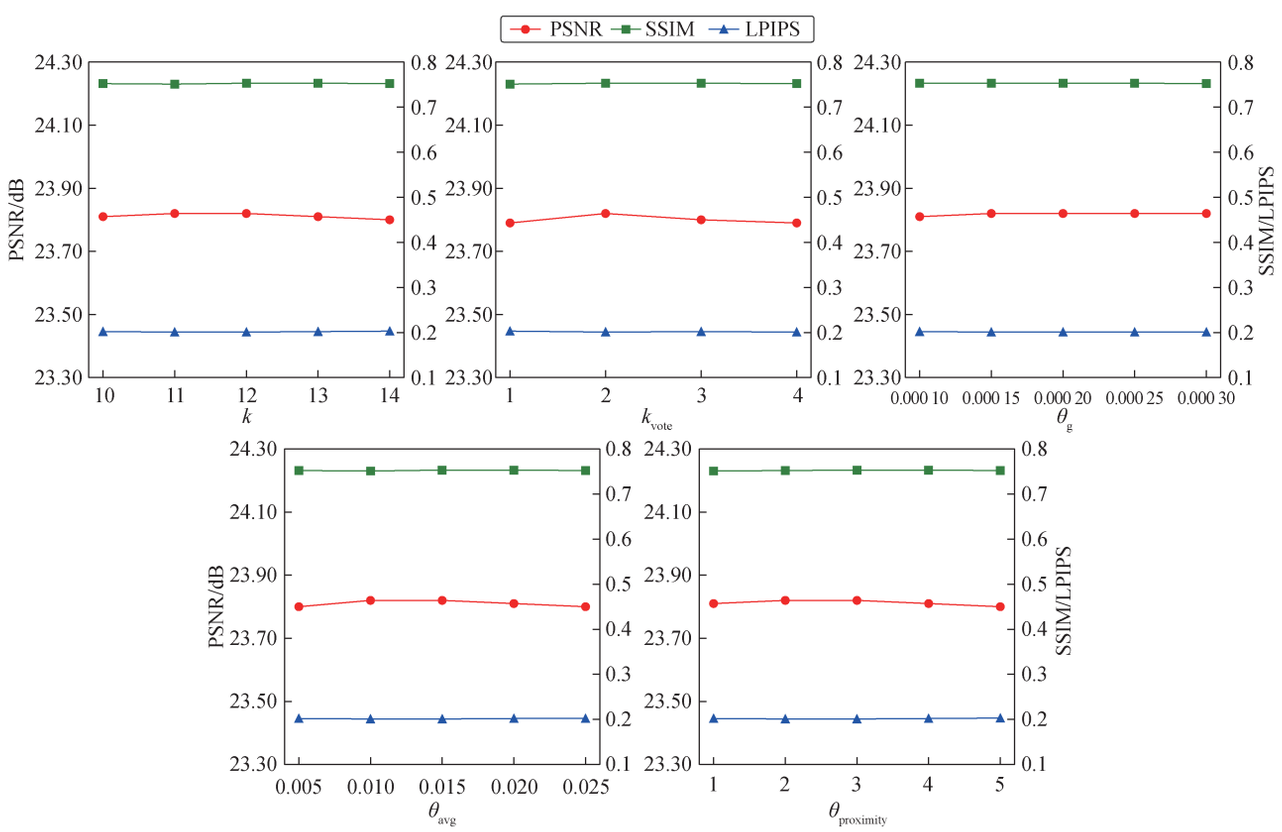


图7 Mip-NeRF360数据集上的边缘簇识别关键参数的敏感性分析曲线

Fig. 7 Sensitivity curves for key parameters in edge cluster identification on the Mip-NeRF360 dataset

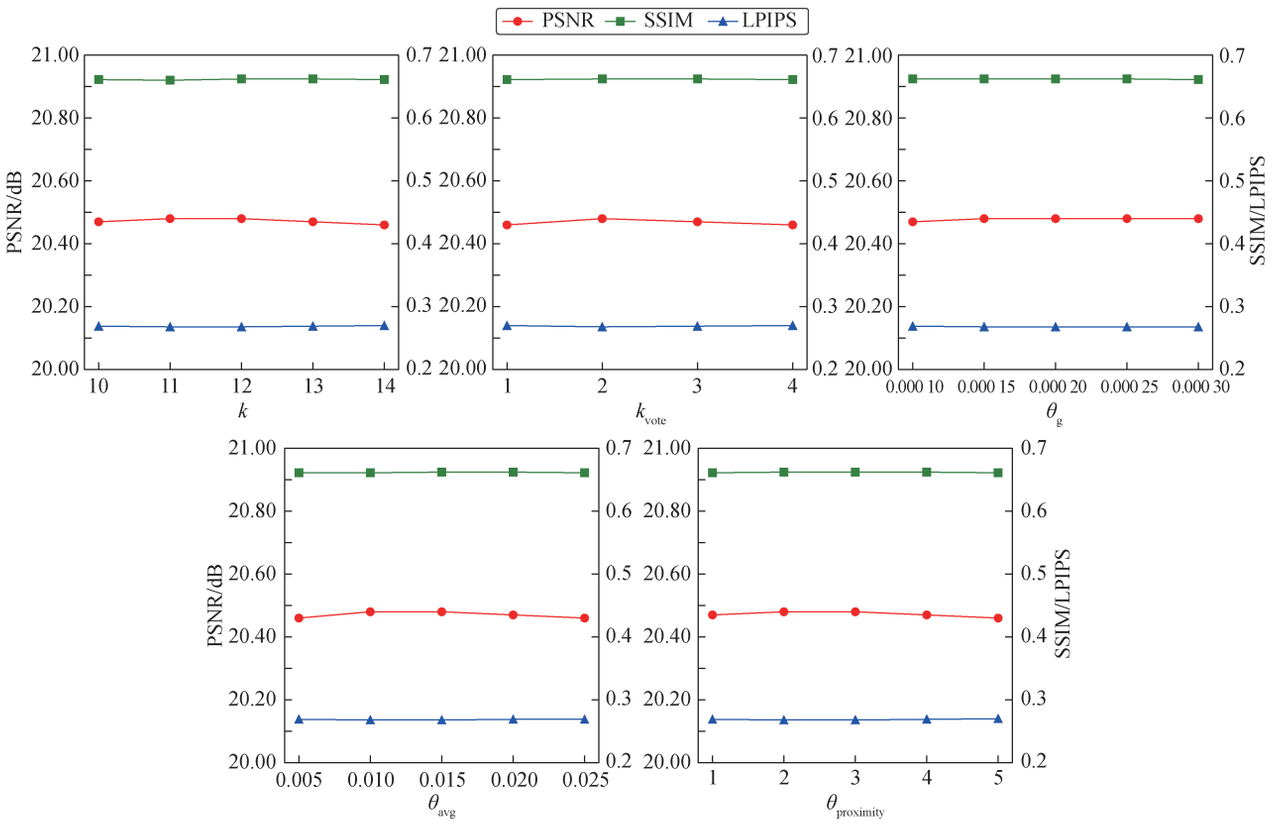


图 8　LLFF 数据集上的边缘簇识别关键参数的敏感性分析曲线

Fig. 8　Sensitivity curves for key parameters in edge cluster identification on the LLFF dataset

表 8　关键参数波动引起的性能变化范围

Table 8　Performance variation ranges induced by key parameter fluctuations

参数	测试范围	Mip-NeRF360 数据集			LLFF 数据集		
		ΔPSNR/dB	ΔSSIM	ΔLPIPS	ΔPSNR/dB	ΔSSIM	ΔLPIPS
聚类数 k	10、11、12、13、14	0.02	0.002	0.002	0.02	0.002	0.002
投票阈值 k_{vote}	1、2、3、4	0.03	0.002	0.002	0.03	0.001	0.002
梯度阈值 θ_g	0.000 10、0.000 15、0.000 20、0.000 25、0.000 30	0.02	0.003	0.002	0.02	0.003	0.003
平均梯度阈值 θ_{avg}	0.005、0.010、0.015、0.020、0.025	0.02	0.002	0.001	0.02	0.001	0.001
邻近度阈值 $\theta_{proximity}$	1、2、3、4、5	0.02	0.002	0.002	0.02	0.001	0.002

注：Δ 值表示对应指标在该参数测试范围内的最大值与最小值之差。

表 9　边缘簇识别关键参数的最优取值

Table 9　Optimal values of key parameters for edge cluster identification

参数	最优值
k	12
k_{vote}	2
θ_g	0.000 20
θ_{avg}	0.015
$\theta_{proximity}$	3

3　结　论

本文针对稀疏视图下三维高斯泼溅在场景边缘区域重建质量不足的问题，提出一种边缘簇驱动的三维高斯泼溅方法。该方法摒弃依赖复杂外部先验的传统路径，转向从三维高斯内部结构分布中挖掘信息，通过“由内而外”的优化机制实现边缘细节增强。本文主要贡献包括：1）提出基于聚类分析的边

缘高斯簇感知方法，通过多准则融合算法精准识别场景边缘结构；2）设计局部邻近引导的高斯解池策略，在识别出的边缘簇内部依据局部邻近关系智能增密，强化轮廓表征；3）引入协同监督的深度正则化方法，结合绝对与相对深度约束，保障几何一致性。上述环节形成从结构感知、针对性增强到几何约束的完整优化回路。在Mip-NeRF360、LLFF、Deep Blending、Tanks&Temples和DTU数据集上的实验表明，本文方法在稀疏输入下，于PSNR、SSIM和LPIPS指标上均取得稳定提升或具有竞争优势。特别是在边缘细节重建上取得显著视觉改善，同时维持了实时渲染效率。消融实验验证了各模块的有效性与互补性，参数分析也证实方法对关键参数变化不敏感，具备较强鲁棒性。然而，本研究仍存在局限：方法在极端稀疏条件下的稳定性有待进一步验证；深度先验质量对正则化效果存在影响；当前方法主要针对静态场景，对动态场景的适应性仍需探索。未来工作将围绕以下方向展开：探索更高效的边缘簇识别算法以降低计算开销；研究深度先验质量的自适应评估机制；将方法扩展至动态场景重建；结合语义信息进一步提升结构合理性。

参考文献（References）

Barron J T, Mildenhall B, Verbin D, Srinivasan P P and Hedman P. 2022. Mip-NeRF 360: unbounded anti-aliased neural radiance fields//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 5460-5469 [DOI: 10.1109/CVPR52688.2022.00539]

Canny J. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6): 679-698 [DOI: 10.1109/TPAMI.1986.4767851]

Chen Y D, Xu H F, Zheng C X, Zhuang B H, Pollefeys M, Geiger A, et al. 2024. MVSplat: efficient 3D Gaussian splatting from sparse multi-view images//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 370-386 [DOI: 10.1007/978-3-031-72664-4_21]

Cheng L H, Li C H, Hu R Z and Liu L G. 2025. 3D Gaussian splatting model for low-light enhancement. Journal of Image and Graphics, 30(10): 3289-3301 (程龙昊, 李常颢, 胡瑞珍, 刘利刚. 2025. 面向低光增强的三维高斯泼溅模型. 中国图象图形学报, 30(10): 3289-3301) [DOI: 10.11834/jig.240598]

Chung J, Oh J and Lee K M. 2024. Depth-regularized optimization for 3D Gaussian splatting in few-shot images//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 811-820 [DOI: 10.1109/CVPRW63382.2024.00086]

Deng K L, Liu A, Zhu J Y and Ramanan D. 2022. Depth-supervised NeRF: fewer views and faster training for free//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 12872-12881 [DOI: 10.1109/CVPR52688.2022.01254]

Fan Z W, Wang K, Wen K R, Zhu Z H, Xu D J and Wang Z Y. 2024. LightGaussian: unbounded 3D Gaussian compression with 15 × reduction and 200+ FPS//Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #4447 [DOI: 10.5555/3737916.3742363]

Felzenszwalb P F and Huttenlocher D P. 2004. Efficient graph-based image segmentation. International Journal of Computer Vision, 59(2): 167-181 [DOI: 10.1023/B:VISI.0000022288.19776.77]

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): #257 [DOI: 10.1145/3272127.3275084]

Jensen R, Dahl A, Vogiatzis G, Tola E and Aanæs H. 2014. Large scale multi-view stereopsis evaluation//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 406-413 [DOI: 10.1109/CVPR.2014.59]

Kerbl B, Kopanas G, Leimkuehler T and Drettakis G. 2023. 3D Gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics, 42(4): #139 [DOI: 10.1145/3592433]

Knapitsch A, Park J, Zhou Q Y and Koltun V. 2017. Tanks and temples: benchmarking large-scale scene reconstruction. ACM Transactions on Graphics (TOG), 36(4): #78 [DOI: 10.1145/3072959.3073599]

Li J H, Zhang J W, Bai X, Zheng J, Ning X, Zhou J, et al. 2024. DNGaussian: optimizing sparse-view 3D Gaussian radiance fields with global-local depth normalization//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 20775-20785 [DOI: 10.1109/CVPR52733.2024.01963]

Liu X H, Chen J B, Kao S H, Tai Y W and Tang C K. 2024. Deceptive-NeRF/3DGS: diffusion-generated pseudo-observations for high-quality sparse-view reconstruction//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 337-355 [DOI: 10.1007/978-3-031-72640-8_19]

Liu X N, Chen C Y, Hu X J and Yu H Y. 2024. Virtual viewpoint image synthesis using neural radiance fields with depth information supervision. Journal of Image and Graphics, 29(7): 2035-2045 (刘晓楠, 陈纯毅, 胡小娟, 于海洋. 2024. 带深度信息监督的神经辐射场虚拟视点画面合成. 中国图象图形学报, 29(7): 2035-2045) [DOI: 10.11834/jig.221188]

Mildenhall B, Srinivasan P P, Ortiz-Cayon R, Kalantari N K, Ramamoorthi R, Ng R, et al. 2019. Local light field fusion: practical

view synthesis with prescriptive sampling guidelines. ACM Transactions on Graphics (TOG), 38(4): #29 [DOI: 10.1145/3306346.3322980]

Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2022. NeRF: representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1): 99-106 [DOI: 10.1145/3503250]

Navaneet K L, Pourahmadi Meibodi K, Abbasi Koohpayegani S and Pirsiavash H. 2024. CompGS: smaller and faster Gaussian splatting with vector quantization//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 330-349 [DOI: 10.1007/978-3-031-73411-3_19]

Niemeyer M, Barron J T, Mildenhall B, Sajjadi M S M, Geiger A and Radwan N. 2022. RegNeRF: regularizing neural radiance fields for view synthesis from sparse inputs//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 5470-5480 [DOI: 10.1109/CVPR52688.2022.00540]

Ranftl R, Bochkovskiy A and Koltun V. 2021. Vision transformers for dense prediction//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 12159-12168 [DOI: 10.1109/ICCV48922.2021.01196]

Schönberger J L and Frahm J M. 2016. Structure-from-motion revisited//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 4104-4113 [DOI: 10.1109/CVPR.2016.445]

Soille P. 1999. Morphological Image Analysis: Principles and Applications. Berlin, Germany: Springer

Szymanowicz S, Rupprecht C and Vedaldi A. 2024. Splatter image: ultra-fast single-view 3D reconstruction//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 10208-10217 [DOI: 10.1109/CVPR52733.2024.00972]

Tang J X, Ren J W, Zhou H, Liu Z W and Zeng G. 2024. DreamGaussian: generative Gaussian splatting for efficient 3D content creation//Proceedings of the 12th International Conference on Learning Representations. Vienna, Austria: OpenReview.net

Wang G C, Chen Z X, Loy C C and Liu Z W. 2023. SparseNeRF: distilling depth ranking for few-shot novel view synthesis//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 9031-9042 [DOI: 10.1109/ICCV51070.2023.00832]

Wu G J, Yi T R, Fang J M, Xie L X, Zhang X P, Wei W, et al. 2024a. 4D Gaussian splatting for real-time dynamic scene rendering//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 20310-20320 [DOI: 10.1109/CVPR52733.2024.01920]

Wu T, Yuan Y J, Zhang L X, Yang J, Cao Y P, Yan L Q, et al. 2024b. Recent advances in 3D Gaussian splatting. Computational Visual Media, 10(4): 613-642 [DOI: 10.1007/s41095-024-0436-y]

Xiao Q, Chen M L, Zhang Y and Huang X H. 2024. Neural radiance field reconstruction for sparse indoor panoramas. Journal of Image and Graphics, 29(9): 2596-2609 (肖强, 陈铭林, 张晔, 黄小红. 2024. 室内稀疏全景图的神经辐射场重建. 中国图象图形学报, 29(9): 2596-2609) [DOI: 10.11834/jig.230643]

Xiong H L, Muttukuru S, Xiao H Y, Upadhyay R, Chari P, Zhao Y J, et al. 2025. Sparsegs: sparse view synthesis using 3D Gaussian splatting//Proceedings of 2025 International Conference on 3D Vision (3DV). Singapore, Singapore: IEEE: 1032-1041 [DOI: 10.1109/3DV66043.2025.00100]

Yan Z W, Low W F, Chen Y and Lee G H. 2024. Multi-scale 3D Gaussian splatting for anti-aliased rendering//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 20923-20931 [DOI: 10.1109/CVPR52733.2024.01977]

Yang C, Li S K, Fang J M, Liang R F, Xie L X, Zhang X P, et al. 2024a. GaussianObject: high-quality 3D object reconstruction from four views with Gaussian splatting. ACM Transactions on Graphics (TOG), 43(6): #199 [DOI: 10.1145/3687759]

Yang L H, Kang B Y, Huang Z L, Zhao Z, Xu X G, Feng J S, et al. 2024b. Depth anything V2//Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #688 [DOI: 10.5555/3737916.3738604]

Yang Q X, Kong D M, Chen J, Li X W and Shen Y. 2025. Enhancing PointPillars three-dimensional object detection with density clustering and dual attention mechanisms. Laser and Optoelectronics Progress, 62(10): #1012001 (杨庆鑫, 孔德明, 陈晶, 李晓伟, 沈阅. 2025. 基于密度聚类和双重注意力机制的PointPillars改进方法. 激光与光电子学进展, 62(10): #1012001) [DOI: 10.3788/LOP240732]

Zhang J W, Li J H, Yu X H, Huang L, Gu L, Zheng J, et al. 2024. CoR-GS: sparse-view 3D Gaussian splatting via co-regularization//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 335-352 [DOI: 10.1007/978-3-031-73232-4_19]

Zhu Z H, Fan Z W, Jiang Y F and Wang Z Y. 2024. FSGS: real-time few-shot view synthesis using Gaussian splatting//Proceedings of 2024 European Conference on Computer Vision. Milan, Italy: Springer: 145-163 [DOI: 10.1007/978-3-031-72933-1_9]

作者简介

马栋林，男，副教授，主要研究方向为模式识别、深度学习和智能信息处理。E-mail: mdonglin@163.com

惠旭同，通信作者，男，硕士研究生，主要研究方向为三维重建和新视角合成。E-mail: logandahui@163.com

赵宏，男，教授，主要研究方向为计算机视觉、深度学习和自然语言处理。E-mail: zhaoh@lut.edu.cn