



激光与光电子学进展
Laser & Optoelectronics Progress
ISSN 1006-4125, CN 31-1690/TN

《激光与光电子学进展》网络首发论文

题目：基于三维高斯溅射技术的可微分渲染研究进展
作者：高建，陈林卓，沈秋，曹汛，姚遥
收稿日期：2024-05-27
网络首发日期：2024-07-12
引用格式：高建，陈林卓，沈秋，曹汛，姚遥. 基于三维高斯溅射技术的可微分渲染研究进展[J/OL]. 激光与光电子学进展.
<https://link.cnki.net/urlid/31.1690.TN.20240708.1614.028>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于三维高斯溅射技术的可微分渲染研究进展

高建¹, 陈林卓¹, 沈秋², 曹汛², 姚遥^{1*}

¹ 南京大学智能科学与技术学院, 江苏 苏州 215163;

² 南京大学电子科学与工程学院, 江苏 南京 210023

摘要 近来, 三维高斯溅射技术 (3D Gaussian Splatting, 3DGS) 的兴起为可微分渲染技术带来了革命性的变化。尽管神经辐射场 (Neural Radiance Field, NeRF) 作为可微分渲染技术的里程碑, 已经在高真实感的新视角合成方面取得了突破性进展, 但它依然存在隐式表达不直观、训练效率低等问题。相比之下, 3DGS 技术通过显式点云表达和高度并行化的可微分光栅化管线, 成功解决了 NeRF 的痛点, 显著提升了训练速度和渲染效率, 并提供了更高的场景控制性。3DGS 技术的出现, 将可微分渲染推向了新的高度, 已经在新视角渲染、动态场景重建等多个领域取得了显著成果, 展现出广阔的应用前景和巨大潜力。本文旨在展示 3DGS 技术的最新发展动态, 激发对可微分渲染技术的深入探索, 为未来技术的进一步发展提供启示。

关键词 三维高斯溅射, 可微分渲染, 三维重建, 综述

中图分类号 TP37 **文献标志码** A

Advances in Differentiable Rendering Based on 3D Gaussian Splatting

Gao Jian¹, Chen Linzhuo¹, Shen Qiu², Cao Xun², Yao Yao^{1*}

¹ School of Intelligence Science and Technology, Nanjing University, Suzhou, Jiangsu 215163, China;

² School of Electronic Science and Engineering, Nanjing University, Nanjing, Jiangsu 210023, China

Abstract Recently, the advent of 3D Gaussian Splatting (3DGS) technology has brought revolutionary changes to the development of differentiable rendering techniques. While Neural Radiance Field (NeRF) has been a milestone in differentiable rendering, achieving breakthroughs in high-fidelity novel view synthesis, it still suffers from issues such as implicit expression and low training efficiency. In contrast, 3DGS technology addresses the pain points of NeRF by employing explicit point cloud representation and a highly parallelizable differentiable rasterization pipeline. Compared to NeRF, 3DGS notably improves training speed and rendering efficiency while offering greater scene control. The advent of 3DGS technology propels differentiable rendering to new heights, achieving remarkable results in areas such as novel view synthesis, dynamic scene reconstruction and etc. 3DGS has demonstrated broad application prospects and tremendous potential. This paper aims to showcase the latest developments in 3DGS technology, inspire in-depth exploration of differentiable rendering techniques, with the hope of providing insights for technological advancements in the future.

Key words 3D Gaussian Splatting, Differentiable Rendering, 3D Reconstruction, Review

OCIS codes 150.6910; 150.1135; 150.0155

基金项目: 国家自然科学基金专项项目 (62441204)

通信作者: *E-mail: yaoyao@nju.edu.cn

1 引言

三维计算机视觉的核心挑战之一是如何有效地重建三维场景的信息。近年来,随着可微分渲染技术的涌现,这一领域取得了显著进展^[1-6]。其中,神经辐射场 (Neural Radiance Field, NeRF)^[7]的出现,是可微分渲染技术 (Differentiable Rendering) 的里程碑式节点。NeRF 使用多层感知机 (Multi-Layer Perceptron, MLP) 将坐标与视角方向映射为颜色和密度,通过体渲染 (Volume Rendering) 的方式,在新视角合成任务上展现出了极具真实感的渲染结果。然而,NeRF 并非没有缺点:其一,其隐式表达不直观,可控性不高;其二,其训练效率低,难以进行高分辨率的实时渲染。最近,三维高斯溅射技术 (3D Gaussian Splatting, 3DGS)^[8] 的出现,则从根本上解决了 NeRF 的这些痛点,带来了一场技术上的变革。3DGS 将场景表达为显示的点云,通过高度并行化的可微分光栅化管线,在保留了体渲染高质量优势的同时,大大提高了训练速度和渲染效率,也为场景的控制提供了条件。

在可微分渲染技术之前,传统的三维重建流程采用运动恢复结构 (Structure from Motion, SfM)^[9-14]、多视图立体匹配 (Multi-View Stereo, MVS)^[15-20] 等技术来实现场景的三维几何重建。然而,这些方法建立在场景物体表面的朗伯体假设之上,存在一定的局限性。相比之下,基于可微渲染的三维重建技术通过可微分化渲染管线,仅通过最小化渲染图像与真实观测图像之间的光度误差,就可实现对三维场景信息的优化。

相较于传统的三维重建方法,可微渲染技术能够更自然地建模场景的光照信息,并且不受朗伯体假设的限制,在自由视点渲染^[7,8]、反射面物体^[21]、物体材质与光照解耦^[22,23]等任务中展现出显著优势。神经辐射场 (Neural Radiance Field, NeRF)^[7]作为一种代表性的基于神经隐式表达的可微渲染技术,受到了广泛关注。然而,NeRF 在实时渲染和场景编辑等方面仍存在挑战。最近,以 3DGS 为代表的点云可微渲染优化方法取得了新的突破。与神经隐式表达不同,3DGS 利用显式的点云表达方式来精确捕捉和展现三维场景的信息。在优化速度、渲染速度以及精度方面,3DGS 均明显优于 NeRF 技术,将可微分渲染技术推向了一个新的高度。3DGS 技术已成为引领相关领域发展的变革性力量。随着基于三维高斯点云的可微渲染优化技术在新视角渲染、动态场景重建等领域取得的令人瞩目的成果,其研究热度不断提升(如图 1 所示)。

本文旨在向读者展示 3DGS 技术在多个领域中的最新发展动态,凸显其广阔的应用前景与巨大潜力。本文的章节安排如图 2 所示:第 2 节介绍 3DGS 技术的基本原理;第 3 节探讨针对 3DGS 渲染的优化方案,包括对 3DGS 模块的优化(第 3.1 节)和在 3DGS 中

应用抗锯齿技术（第 3.2 节）；第 4 节详述对 3DGS 技术重建的场景进行几何和材质提取的发展，包括表面格网提取（第 4.1 节）和材质解耦（第 4.2 节）；第 5 节聚焦于复杂场景下的 3DGS 技术，包括动态场景重建（第 5.1 节）、大场景重建（第 5.2 节）以及稀疏视角重建（第 5.3 节）；最后，在第 6 节进行总结，以期激发对可微分渲染技术的深入思考与探索，并为未来技术的发展提供有益的启示。

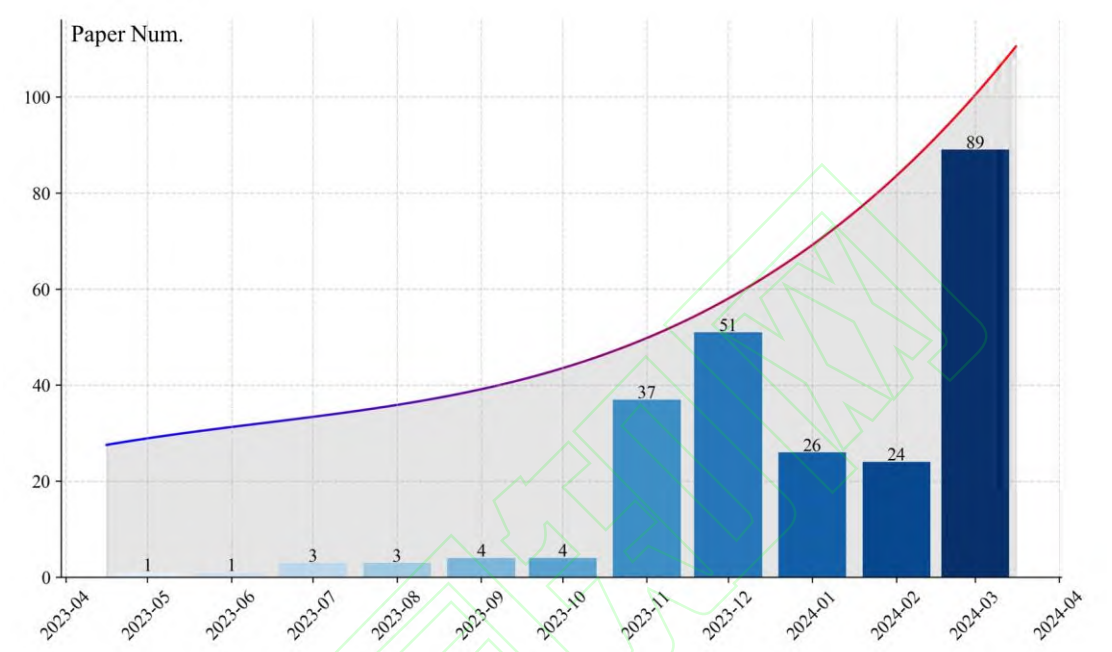


图1 基于三维高斯溅射技术的研究趋势
Fig. 1 Research trends on 3D Gaussian Splatting techniques

| 3D Gaussian Splatting (Sec. 2) | | | | | | |
|--|---|--|---|---|--|---|
| Optimized Rendering (Sec. 3) | | Geometric and Material (Sec. 4) | | Advanced Scene Reconstruction (Sec. 5) | | |
| Components Optimization (Sec. 3.1) | Anti-Aliasing Techniques (Sec. 3.2) | Surface Mesh Extraction (Sec. 4.1) | Material Decomposition (Sec. 4.2) | Dynamic Scene Reconstruction (Sec. 5.1) | Large-Scale Reconstruction (Sec. 5.2) | Sparse View Reconstruction (Sec. 5.3) |
| O-GS GaussianPro Pixel-GS MiniSplatting FreGS Scaffold-GS GES Spec-Gaussian | Mip-Splatting Analytic-Splatting MS-3DGS SA-GS | SuGar Mesh-GS GaMes 3DGSR NeuSG DN-Splatter 2DGS | GaussianShader GIR GS-IR Relightable 3DG | 4DGS Control4D GauFre SC-GS DynMF NPG CoGS Dy3DG Spacetime Gaussian GaussianFlow E3DGS 4DG | Fred3DGS CityGaussian VastGaussian | SparseGS FSGS CoherentGS DNGaussian CoR-GS InstantSplat Pixelsplat MVSplat LatentSplat Gamba |

图2 基于三维高斯溅射技术的研究进展总览
Fig. 2 Overview of advances on 3D Gaussian Splatting technology

2 三维高斯溅射技术基本原理

三维高斯溅射(3D Gaussians Splatting, 3DGS)^[8] 技术将三维场景中的辐射场定义在离散的三维高斯点云上, 以实现可微分的体渲染 (Volume Rendering), 其技术流程如图 3 所示。在 3DGS 中, 每一个点被表达为独立的三维高斯分布, 其数学表述为:

$$G(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}(\mathbf{x} - \boldsymbol{\mu})\right) \quad (1)$$

其中, $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 分别表示了三维高斯分布的均值和协方差矩阵。每一个高斯点 \mathbf{P} 还被赋予了一个不透明度 α 以及一个颜色值 \mathbf{c} , 用于表达三维场景的辐射场。由于镜面反射、高光等现象的存在, 同一物体在不同观测角度下的颜色是不一致的。为了模拟颜色值随观察视角变换的特性, 3DGS 采用球谐系数 (Spherical Harmonics, SHs) 来进行建模。

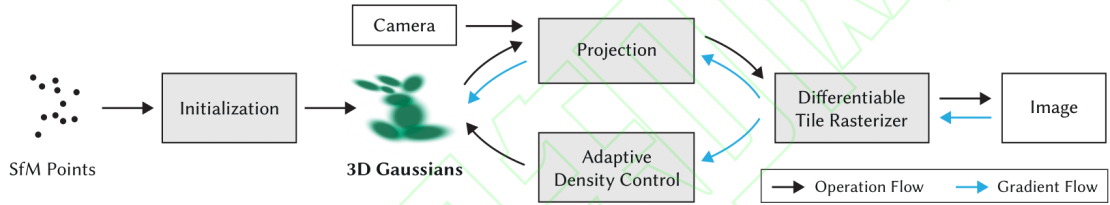


图3 三维高斯溅射技术总览^[8]

Fig. 3 Overview of 3D Gaussian Splatting technology^[8]

3DGS 技术的渲染过程, 即可微分渲染优化过程中的前向, 采用的是物序 (Object-Order) 法。其基本思想是将所有的高斯点 $\{\mathbf{P}\}$ 根据相机参数投影到影像平面上, 形成一系列的二维平面高斯 $\{\mathbf{p}\}$, 再通过二维平面上高斯的叠加混合, 最终合成影像。然而, 将三维高斯分布准确投射为二维高斯分布并非易事。为了解决这一问题, 3DGS 采用了椭圆加权平均 (Elliptical Weighted Average, EWA) 技术^[24], 使用局部仿射变换模型来近似三维高斯分布到二维平面的投影过程:

$$\boldsymbol{\Sigma}' = \mathbf{J} \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T \mathbf{J}^T \quad (2)$$

其中, \mathbf{J} 和 \mathbf{W} 分别表示透视投影变换的一阶仿射近似和视角变换矩阵。

对于每一个像素, 共存在 N 个二维高斯分布覆盖其上。根据成像的基本原理, 靠近相机的物体会对远离相机的物体造成遮挡。在体渲染, 这一原理表现为: 对这 N 个二维高斯分布进行深度上的排序, 再从前到后执行透明度混合 (alpha-blending), 从而合成该像素的颜色值:

$$\mathcal{C} = \sum_{i \in N} T_i \alpha_i \mathbf{c}_i \quad \text{with} \quad T_i = \prod_{j=1}^{i-1} (1 - \alpha_j) \quad (3)$$

其中, α 是通过高斯点的不透明度 α 和其二维分布计算得到。3DGS 技术的后向阶段中, 通过计算渲染影像和真实观测影像之间的差异来评估损失。3DGS 技术采用最小绝对值偏差 L_1 和结构相似损失 L_{D-SSIM} 的加权和来评估两者之间的差异:

$$L = (1 - \lambda)L_1 + \lambda L_{D-SSIM} \quad (4)$$

之后, 通过反向传播算法进行梯度的反传, 对三维高斯点云中每个点的参数进行更新, 从而重建出整个三维场景中的辐射场。

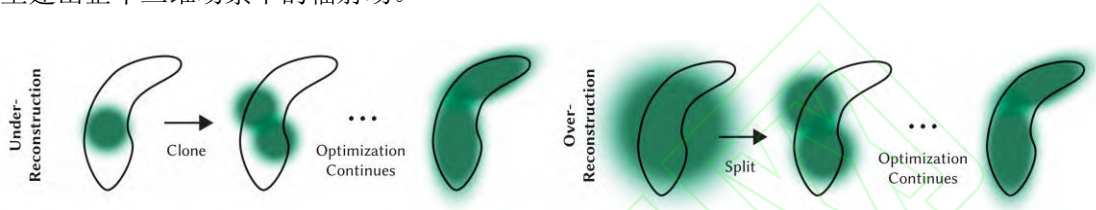


图4 三维高斯喷溅技术使用自适应密度控制来进行增稠^[8]

Fig. 4 3D Gaussian splatting technique utilizes adaptive density control for densification^[8]

3DGS 在构建初始的三维高斯点云时, 仅仅依赖于完全随机的点或通过运动恢复结构 (Structure from Motion, SfM) 得到的稀疏点云。三维高斯点的数量直接关系到最终的渲染质量。因此, 仅仅依靠可微分的渲染管道可以对三维点云的参数进行持续的更新优化, 还不足以达到其现有的卓越性能。3DGS 成功的另一关键要素在于: 对三维高斯点云的自适应密度控制 (Adaptive Control of Gaussians), 如图 4 所示。当渲染影像中某些区域的渲染结果与真实观测存在显著差异时, 这些区域便成为了点云增稠的重点。3DGS 通过影像平面上二维高斯点位置的平均累积梯度信息, 判断出哪些高斯点需要进行增稠处理。根据二维高斯点在平面上所占据范围的大小, 3DGS 进一步将待加密的高斯点划分为“过度重建” (Over-Reconstruction) 和“重建不足” (Under-Reconstruction) 两类。对于过度重建的高斯点, 3DGS 采用克隆 (clone) 操作, 而对于重建不足的高斯点, 则实施分裂 (split) 操作。通过这种处理方式, 3DGS 有效地增加了三维场景中高斯点的数量, 显著提升了渲染的质量。

3 渲染质量优化

3.1 模块优化

经过对 3DGS 技术的深入研究, 学者们发现其存在一些设计上的不足, 这些问题制约

了其性能进一步提升的可能性。为了解决这些问题，他们进行了一系列的改进。这些改进涉及到了多个方面，包括投影方式、密度控制策略以及对三维场景的表达。

O-GS^[25] 意识到 3DGS 技术中的局部仿射近似引入了误差，导致“伪影”(artifacts) 和渲染质量下降。为了解决这一问题，他们提出了一种最优的投影方式：将高斯点投影到单位球的切平面上，并在单位球上进行透明度混合，根据像素发出的光线与单位球的交点检索颜色值，最终渲染得到影像。一些学者认为 3DGS 技术中的自适应密度控制策略存在缺陷，并着手进行一些改进。GaussianPro^[26] 从传统的多视图立体几何技术中汲取灵感，以 PatchMatch 的思想设计了一种渐进的传播策略来引导高斯点的增稠，为 3DGS 的密度控制引入了准确的场景几何信息。Pixel-GS^[27] 发现 3DGS 技术的渲染结果在初始点不足的区域中会出现模糊和针状伪影，并将这一原因归结于过于简单的密度控制策略。为此，他们提出将二维高斯覆盖的像素数作为权重来动态地聚集不同视角的梯度，从而促进初始点不足的区域中高斯点的增稠。MiniSplatting^[28] 认为高斯点云表达的空间分布低效，限制了模型的性能。为此，他们提出了“模糊分裂”、“重新初始化”的增稠策略以及“高斯二值化”和“采样”的简化策略，规整了三维高斯点云的空间位置分布，使用更少的点数更高效地表达了整个重建场景。FreGS^[29] 从频率域的角度考虑，采用渐进式的频域正则化，实现由粗到细的高斯场景增稠方案，达到了更高的渲染精度。一些学者则认为三维高斯这种场景表达仍有缺陷。Scaffold-GS^[30] 发现 3DGS 倾向于使用冗余的点来拟合每一个训练视角，而忽略场景的内在几何结构，导致模型难以应对显著的视角变化和无纹理区域。因此，Scaffold-GS 为三维高斯点云表达引入锚点(anchor)，利用“锚点-局部高斯点”这种双层次结构来表示场景。在 Scaffold-GS 中，每个锚点都可以利用多层感知机生成一组局部高斯点，从而动态适应各种观察角度和距离，如图 5。GES^[31] 将三维高斯替换为广义指数函数 (Generalized Exponential Function, GEF)。对于同一个场景，GEF 的表达所需的点数更少，使得模型效率更高和存储空间更少，并展现出了更好的边缘细节。3DGS 仅使用三阶球谐系数来表达辐射场，不足以对高光区域进行拟合。为此，Spec-Gaussian^[32] 引入了各向异性球形高斯 (Anisotropic Spherical Gaussian, ASG) 外观场来对每一个高斯点的辐射进行建模，在高光区域上取得了更好的效果。

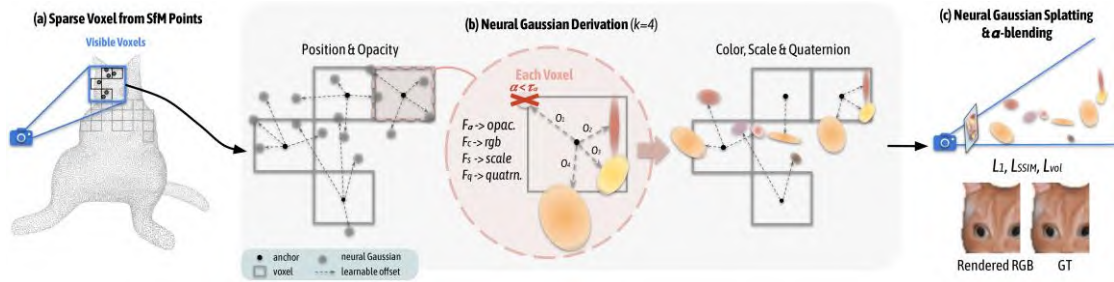


图5 Scaffold-GS技术路线图^[30]

Fig. 5 Roadmap for Scaffold-GS^[30]

本文对上述各项方法在 Mip-NeRF 360 数据集^[33]上的评估结果进行了整理，如表 1 所示。其中，渲染精度指标包括：峰值信噪比 (Peak Signal to Noise Ratio, PSNR)、结构相似性指数 (Structural Similarity Index, SSIM) 以及图像感知相似度指标 (Learned Perceptual Image Patch Similarity, LPIPS)^[34]。此外，表中还整理出了渲染帧率 (FPS) 和模型的存储大小。每个指标中最优和次优分别用粗体、下划线标出。需要注意的是，表中 MiniSplatting 的渲染精度和渲染帧率均摘录的是原论文中 Mini-Splatting-D 的结果，而存储大小摘录的是 Mini-Splatting-C 的结果。前者着重于更高质量的渲染，而后者则注重于场景的压缩。

表 1 3DGS 的改进方案在 Mip-NeRF360 数据集上的定量对比结果

Table 1 Quantitative comparison results of the improved scheme of 3DGS on the Mip-NeRF360 dataset

| | 3DGS | O-GS | GaussianPro | Pixel-GS | MiniSplatting* | FreGS | Scaffold-GS | GES | Spec-Gaussian |
|--------|------------|--------------|-------------|--------------|----------------|-------|--------------|-------|---------------|
| PSNR↑ | 27.21 | 27.17 | 27.92 | 27.88 | 27.54 | 27.85 | 28.84 | 26.91 | <u>28.01</u> |
| SSIM↑ | 0.815 | <u>0.836</u> | 0.825 | 0.834 | 0.832 | 0.826 | 0.848 | 0.794 | 0.812 |
| LPIPS↓ | 0.214 | 0.210 | 0.208 | <u>0.176</u> | 0.175 | 0.209 | 0.220 | 0.250 | 0.222 |
| FPS↑ | <u>134</u> | - | 108 | 89 | 83 | - | 102 | 186 | 70 |
| Mem↓ | 734 | - | - | - | <50 | - | <u>156</u> | 377 | 245 |

3.2 抗锯齿技术

尽管 3DGS 已经能够得到高度真实的新视角合成结果，然而在进行不同分辨率的渲染时可能会产生严重的模糊或锯齿。这一现象被归结为“混叠” (alias)，如图 6 所示。具体而言，三维高斯溅射技术通过在屏幕空间将二维高斯进行固定窗口的膨胀以实现三维高斯点的近似投影，如图 6(a) 所示。然而，固定的窗口会导致退化现象，如图 6(b)。当通过焦距或相机距离改变采样率时，3DGS 的渲染结果会出现图 6(c)所示的扩张效应和图 6(d) 所示的高频伪影。针对这一问题，一些研究者在 3DGS 上引入“抗混叠” (anti-aliasing) 技术。

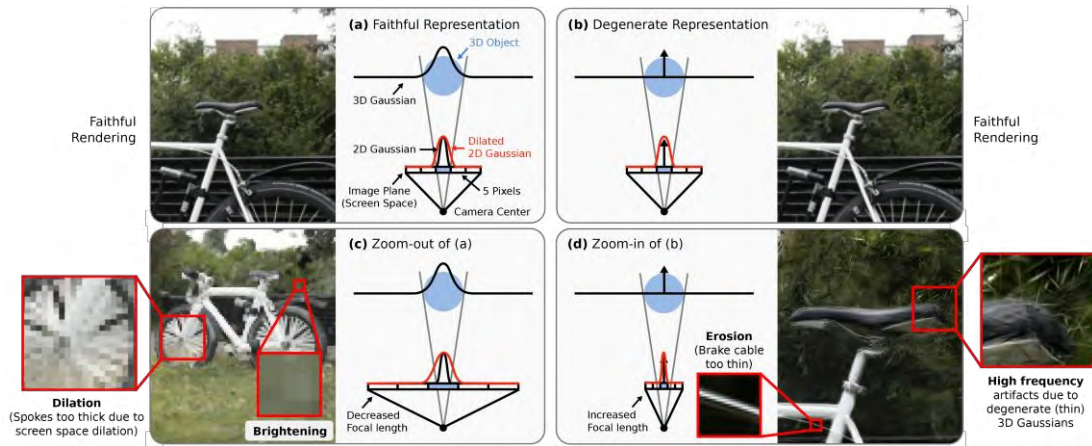


图6 三维高斯溅射技术的混叠现象^[35]

Fig. 6 Aliasing in 3D Gaussian Splatting^[35]

Mip-Splatting^[35] 提出了一种 3D 平滑滤波器，通过输入视角对应的最大采样频率限制三维高斯点的大小，从而消除了放大时的高频伪影；还采用了二维 Mip 滤波器替换原本 3DGS 中对二维高斯的简单膨胀操作，有效地减轻了混叠现象。Analytic-Splatting^[36] 指出，3D 图形渲染中出现的混叠现象主要源于将每个像素视为独立的单点，而非一个整体区域。为了克服这一问题，Analytic-Splatting 通过计算像素内的高斯分布积分，而非直接采样中间像素，提升了抗混叠能力。具体而言，它采用条件逻辑斯蒂函数 (conditioned logistic function) 作为一维高斯信号累积分布函数 (Cumulative Distribution Function, CDF) 的近似，通过两个 CDF 的差值计算出高斯积分，并将这种近似进一步扩展到二维高斯分布。MS-3DGS^[37] 提出了一种多尺度的三维高斯溅射技术来减轻影像放大时候的混叠现象。其核心思想是在渲染不同分辨率的影像时，使用不同尺度的三维高斯点。具体而言，当渲染高分辨率影像时，MS-3DGS 采用更多的小尺寸高斯点来保留细节信息；而在渲染低分辨率影像时，则采用较少的大尺寸高斯点以达到抗锯齿的效果。然而，这种方案需要多尺度的影像共同参与训练。SA-GS^[38] 在渲染过程中对每个高斯点施加 2D 尺度适应滤波，可以在不重新训练的情况下，赋予现有预训练的三维高斯场景以抗混叠的能力。

4 几何与材质提取

4.1 表面格网提取

表面格网准确地描述了物体的几何形状和表面特征，在各种下游任务中都发挥了重要的作用，特别是物理模拟动画方面。格网表面为物理模拟提供了更真实、更精确的基础，使得

动画更加的生动和真实。

SuGar^[39] 是首个从三维高斯场景中提取表面网格的工作。它通过引入正则化项鼓励三维高斯点靠近表面,并通过泊松重建从三维高斯点云中提取网格。最后, SuGar 将三维高斯点与表面网格进行绑定,并使用 3DGS 技术进行进一步优化。Mesh-GS^[40] 采用一种基于网格的三维高斯表达方式,实现交互式形变。具体而言, Mesh-GS 将三维高斯定义在显式网格上,并相互绑定。通过 3DGS 引导网格面的分裂,达到自适应网格细化的目的。显式网格约束抑制了质量较差的高斯(例如未对齐的高斯、长条形高斯),从而增强了视觉质量并避免了变形过程中的伪影。GaMes^[41] 同样将三维高斯绑定到显式格网上,通过网格面的顶点对高斯点进行参数化。因此,通过三维高斯点在格网中的位置就可以定义其所有参数。这种方法使得在编辑过程中,高斯点能够自动根据位置进行适应性的调整,从而实现可编辑三维高斯的实时渲染。一些工作则致力于从三维高斯表达中提取更高质量的表面格网。3DGSR^[42] 将隐式的有向距离场 (Signed Distance Field, SDF) 和三维高斯表达结合起来,通过两者的对齐和共同优化,从而实现隐式表面的重建。3DGSR 使用了一个可微分的钟形 (bell-shaped) 函数将 SDF 值转换为三维高斯点的不透明度,从而为三维高斯点增加表面约束。另一方面,三维高斯优化过程为 SDF 提供了监督信号。然而,三维高斯点的稀疏性导致 SDF 的监督信号不足。为了解决这个问题,3DGSR 分别使用三维高斯渲染和 SDF 的体渲染两种方式获得深度图和法线图,并对这两种方式生成的深度图和法线图施加一致性约束。NeuSG^[43] 将三维高斯溅射作为一种隐式表达重建的引导,以实现更富细节的表面。NeuSG 使用尺度上的正则化将三维高斯点拉扁,从而使得其中心点靠近表面,并利用由神经隐式模型预测的表面法线作为先验来优化 3D 高斯投影生成的点云。DN-Splatter^[44] 在三维高斯溅射技术中引入了几何先验,以应对复杂的室内场景。具体而言, DN-Splatter 使用普通手持设备上自带的传感器获取的深度信息,或者从现有单目深度和法线估计网络模型预测得到的几何先验,来对 3DGS 进行正则化,从而实现更优质的几何表面网格重建。2DGS^[45] 认为三维高斯难以准确表达几何表面的原因在于其天生的多视角不一致性。为了解决这个问题,2DGS 采用了圆盘状的二维高斯来替代三维高斯进行场景表达。它引入了一种透视精确的溅射技术,取代了透视近似的 EWA 溅射,通过“光线-圆盘”相交和光栅化实现可微分渲染。此外,2DGS 还加入了深度畸变和法线一致性项,进一步提高了几何表面重建的质量。相比于原始的 3DGS 方法和早期的 SuGar 方法,2DGS 获得了相当出色的格网提取结果,如图 7 所示。

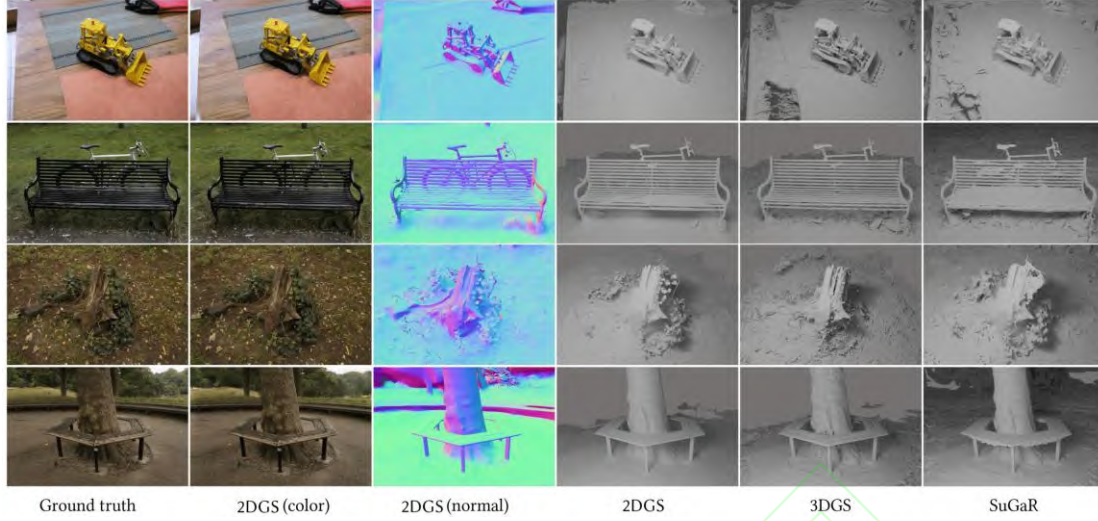


图7 2DGS与3DGS以及SuGaR在表面格网提取任务上的对比^[45]

Fig. 7 Comparison of 2DGS with 3DGS and SuGaR on the surface mesh extraction task^[45]

4.2 材质解耦

基于物理的逆向渲染旨在根据对场景的多视角观测来恢复场景的属性，如材质和照明，以便于进行后续的材质编辑或重打光应用。将三维高斯喷射技术与逆向渲染结合存在两个主要挑战：一是如何计算高斯点云表达的法线；二是如何表示场景的照明和材质以进行基于物理的渲染。针对这两个问题，一些研究者已经取得了一些进展。

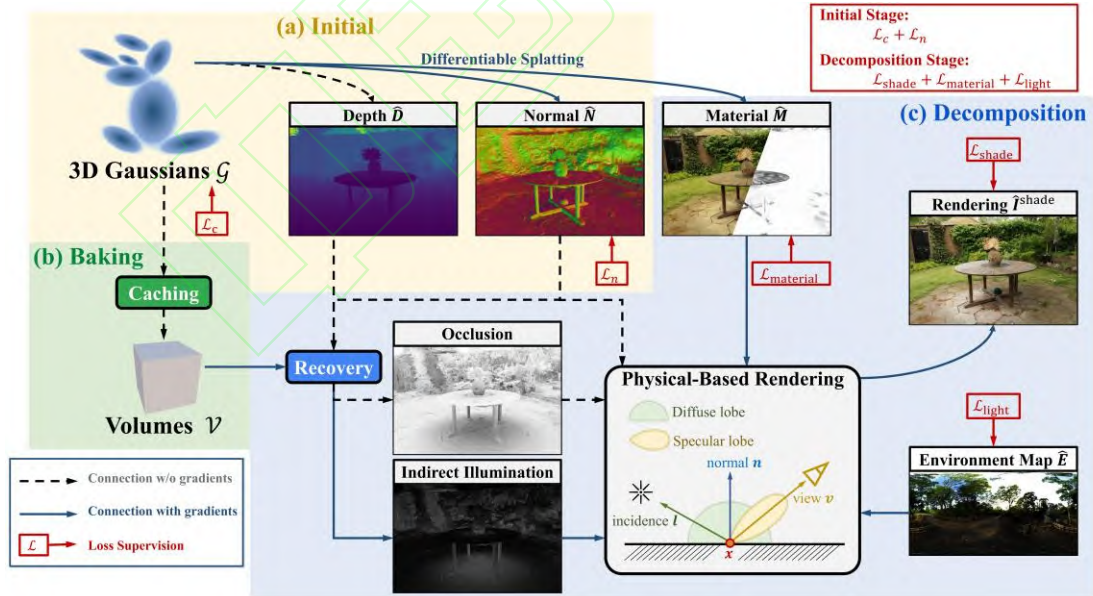


图8 GS-IR技术路线^[46]

Fig. 8 Roadmap for GS-IR^[46]

GaussianShader^[47] 旨在提升 3DGS 在反射表面上的性能。对于法线估计，它将三维高斯的最短轴方向作为基础，并为每个高斯点引入可学习的法线残差以提高准确性。在

GaussianShader 中, 渲染方程被简化为一个着色函数, 而光照则表示为一个可学习的、分辨率为 $6 \times 64 \times 64$ 的立方体光照贴图。GIR^[48] 同样采用三维高斯的最短轴作为法线估计值, 并基于法线与视线方向之间的一致性生成可见掩码, 在渲染过程中辅助剔除背向视角的高斯点。在渲染方程上, GIR 采用微表面反射模型 (Microfacet Model), 通过全卷积网络 (Fully Convolutional Network, FCN) 学习场景的环境光, 并使用球谐系数和视角相关的遮挡来近似对间接光照建模。GS-IR^[46] 采用深度图对法线进行监督学习, 其中每个像素的深度是光线经过所有高斯点的深度加权和。在几何结构优化完成后, GS-IR 将场景进行均匀格网的划分, 并在每一个格网位置上渲染出立方体深度贴图, 使用手动设置的距离阈值获取遮挡信息。接着, GS-IR 将遮挡信息烘焙到每个格网的球谐系数中。除了遮挡格网外, GS-IR 还维护了一个光照格网, 以便对间接光照进行建模。GS-IR 的技术路线如图 8 所示。Relightable 3D Gaussian^[49] 提出了一种基于高斯点云的真实渲染管线, 支持场景重光照、场景编辑和光线追踪。它利用渲染深度图派生的法线图作为每个高斯法线估计的直接监督; 对于复杂的现实世界场景, 还使用多视图立体匹配 (Multi-view Stereo, MVS) 的深度和法线进行额外的监督。针对离散的三维高斯, 设计了基于层次包围盒 (Bounding Volume Hierarchy, BVH) 的光线追踪方法, 并以此为监督优化每个高斯点的可视性。其光照建模包括一个场景共享的全局环境光和每个高斯点上的局部光场, 两者通过可视性组合为高斯点的入射光场。基于这一管线, 多个对象可以被组合到一个新的场景, 并且可以在新的环境光下渲染出真实的阴影效果, 如图 9。

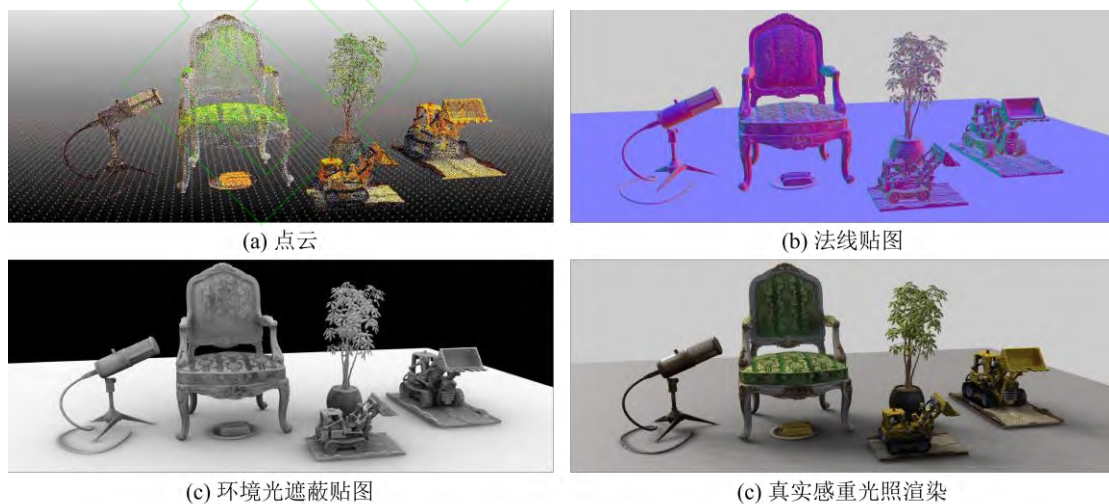


图9 基于物理的真实感点云渲染管线^[49]

Fig. 9 Physical based realistic point cloud rendering pipeline^[49]

5 复杂场景重建

5.1 动态场景重建

在动态场景重建任务中，基于 NeRF 的重建方案已经证明了使用隐式表示重建高质量动态场景的可行性^[50-56]。然而，动态 NeRF 依赖于计算密集型神经网络，导致训练和推理时间延长，不适合虚拟现实等实时应用。而相比 NeRF，3DGS 的训练速度和渲染速度都得到了巨大的提升，因而有研究者尝试将 3DGS 技术引入到动态场景的重建中。

一些方法尝试通过多层感知机 (Multilayer Perceptron, MLP) 来对场景中三维高斯点的运动变化进行建模。4DGS^[57] 和 Control4D^[58] 分别使用 Tri-Plane^[59] 和 HexPlane^[55] 来编码场景的空间特征，并采用 MLP 进行解码，从而获得每个高斯点的运动变化。D3DGS^[60] 和 GauFre^[61] 直接利用一个大规模的 MLP 来预测每个高斯点的运动。其中，GauFre 还采用静态和动态高斯的分离，以区分场景的静态和动态部分。SC-GS^[62] 在场景中引入一组控制点，利用 MLP 预测每个控制点的变化场。之后，通过 K-近邻 (K Nearest Neighbor, KNN) 算法寻找每个高斯点的临近控制点，并通过对临近控制点的变化场插值出高斯点的运动变化。DynMF^[63] 将三维高斯点的运动建模为多项式模型，并利用一个小型的 MLP 来查询每一时刻的多项式系数，从而获得每个高斯随时间的运动变化。NPG^[64] 致力于从单目视频中重建动态物体，采用两阶段方法来进行优化。在第一阶段中，训练一个 MLP 来生成某一时刻的粗略点云模型；接着，在第二阶段，利用粗糙点云模型定义局部体，并在局部体中进行三维高斯点的优化。CoGS^[65] 旨在解决动态场景的操控问题。它采用基于 MLP 的运动变化场建模来实现动态场景的重建，并通过 MLP 将操控信号引入到系统中，从而实现动态场景的实时控制。

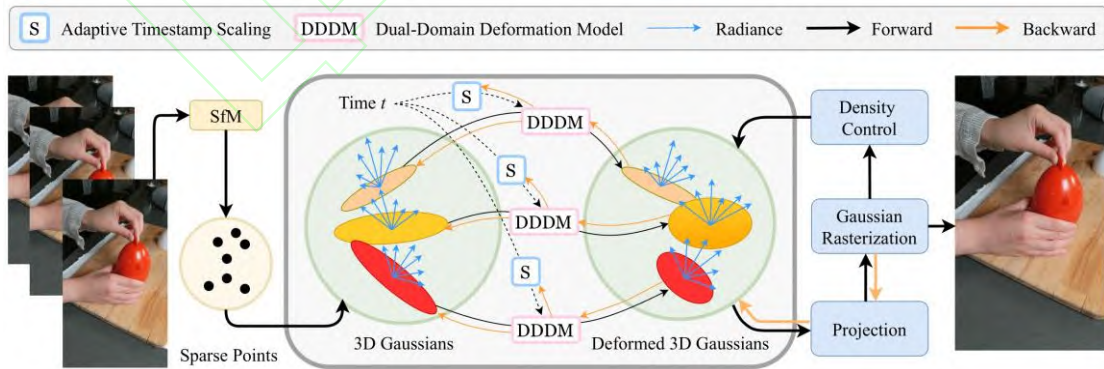


图10 Gaussian-Flow技术流程图^[66]

Fig. 10 Flowchart of Gaussian-Flow^[66]

尽管基于 3DGS 的方案相较于基于 NeRF 的动态重建方案已经取得了显著的效率提升，但引入 MLP 仍需进行一定程度的密集计算。相比之下，另一些方案完全采用显式的运

动场表达。这些显式建模运动的方法只需在 3DGS 的基础上引入少量的计算成本，更适合实时应用。Dy3DGS^[67] 假设高斯点的运动仅包括位置和旋转的变化，而颜色、不透明度和大小不随时间变化。该方法首先对第一帧进行静态重建，然后在剩余帧中进行长期密集六自由度跟踪。然而，它无法重建进入场景的新对象，且仅适用于多摄像机系统。Spacetime Gaussian^[68] 采用多项式来预测高斯变形，并引入时变权重来控制每个高斯的不透明度。GaussianFlow^[66] 采用多项式和傅里叶级数的联合建模来预测每个高斯的运动，同时捕捉时域和频域中的位置和旋转变化，其技术路线如图 10 所示。E3DGS^[69] 分别使用傅立叶级数和线性模型对高斯点位置和旋转变化进行近似，并利用光流进行额外的监督。4DG^[70] 通过在三维高斯表达的基础上加入了新的时间维度，构成了场景的四维高斯表达。在进行渲染时，将每个四维高斯点分解为三维高斯再采用 3DGS 技术进行渲染。

表 2 动态重建方法在 D-NeRF 合成数据集上的定量对比结果

Table 2 Quantitative Comparison Results of Dynamic Reconstruction Methods on D-NeRF Synthetic Datasets

| | 4DGS | D3DGS | CoGS | E3DGS | 4DG | NPG | SCGS | DynMF | GauFRe |
|--------|-------|--------------|--------------|------------|------------|-------|-------|-------|--------|
| PSNR↑ | 34.05 | 39.51 | <u>37.90</u> | 32.07 | 34.09 | 33.86 | 43.31 | 36.80 | 34.80 |
| SSIM↑ | 0.980 | 0.990 | <u>0.983</u> | 0.960 | 0.980 | 0.974 | 0.998 | 0.983 | 0.982 |
| LPIPS↓ | 0.020 | 0.012 | <u>0.017</u> | - | 0.020 | 0.042 | 0.008 | 0.020 | 0.020 |
| FPS↑ | 82 | 30 | - | 150 | <u>114</u> | 20 | - | > 300 | 50 |
| Size | 800 | 800 | 800 | 800 | 800 | 800 | 400 | 400 | 400 |

本文对上述部分方法在 D-NeRF 数据集上^[71] 的评估结果整理为表 2，以供参考。由于大多数方法采用原始 800×800 像素分辨率进行评估，而少量方法采用二倍降采样，本文只对采用原始分辨率的方法在各个指标上选出最优和次优，并分别用粗体和下划线标出。

5.2 大场景重建

大场景重建面临两个关键问题：首先，由于场景庞大，直接整体训练对资源要求过高。分块处理是一种解决方案，但如何确定分块的边界以及每个块所需的点和相机是一个挑战。其次，在大场景拍摄中，曝光和时间等因素可能导致不同区域的外观差异，甚至是季节性差异。对于这些问题，一些研究工作者尝试提出解决方案，探索将 3DGS 技术应用于大场景重建任务之中。

Fred3DGS^[72] 将 3DGS 与联邦学习结合起来，实现了可扩展的三维重建框架，并引入了外观建模来处理如季节性差异等场景变化，达到了与集中式方法相当的渲染图像质量。CityGaussian^[73] 采用分治的训练策略，首先使用 3DGS 对整个场景进行优化，获得场景的全局先验，再根据这种先验进行场景的划分，并提出使用基于块的层次细节选择策略来实现

多级层次细节 (Level-of-Detail, LoD) 渲染。VastGaussian^[74] 采用渐进式分割策略将整个大场景分割为多个单元,并设计了一种可视性准则来确保每一个单元内相机和点云是合理分布的。此外,VastGaussian 意识到大场景区域中的影像常常出现曝光的不一致,并对影像拍摄时的外观变化进行建模,减少了由此导致的伪影。

本文对大场景重建方案的结果进行了摘录和整理,如表 3 所示。表中,每个指标的最优和次优分别用粗体和下划线标出。除了基于 3DGS 的技术方案外,表中还额外摘录了三种基于 NeRF 的大场景重建方案(即 MegaNeRF^[75],SwitchNeRF^[76]和 GridNeRF^[77]),以供参考。

表 3 大场景重建结果定量比较
Table 3 Quantitative Comparison of Large Scene Reconstruction Results

| | Residence | | | Rubble | | | Building | | | Campus | | |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | PSNR↑ | SSIM↑ | LPIPS↓ | PSNR↑ | SSIM↑ | LPIPS↓ | PSNR↑ | SSIM↑ | LPIPS↓ | PSNR↑ | SSIM↑ | LPIPS↓ |
| MegaNeRF | 22.08 | 0.628 | 0.489 | 24.06 | 0.553 | 0.516 | 20.93 | 0.547 | 0.504 | 23.42 | 0.537 | 0.618 |
| SwitchNeRF | 22.57 | 0.654 | 0.457 | 24.31 | 0.562 | 0.496 | 21.54 | 0.579 | 0.474 | 23.62 | 0.541 | 0.609 |
| GridNeRF | <u>23.77</u> | 0.802 | <u>0.137</u> | 24.13 | 0.767 | <u>0.207</u> | - | - | - | <u>24.90</u> | <u>0.757</u> | <u>0.162</u> |
| Fed3DGS | 20.00 | 0.665 | 0.344 | 20.62 | 0.588 | 0.437 | 18.66 | 0.602 | 0.362 | 21.64 | 0.635 | 0.436 |
| CityGaussian | 22.00 | <u>0.813</u> | 0.211 | <u>25.77</u> | <u>0.813</u> | 0.228 | <u>21.55</u> | <u>0.778</u> | <u>0.246</u> | - | - | - |
| VastGaussian | 24.25 | 0.852 | 0.124 | 26.92 | 0.823 | 0.132 | 23.50 | 0.804 | 0.130 | 26.00 | 0.816 | 0.151 |

5.3 稀疏视角重建

3DGS 需要足够数量的输入图片来实现对静态场景的重建与新视角渲染。因此,在输入图片数目受限的情况下,3DGS 会在训练过程中过拟合,并在新视角渲染中出现大量空中伪影和背景塌陷等问题。一系列工作^[78-82] 致力于缓解 3DGS 对图片数量的依赖: SparseGS^[78] 通过引入深度监督和生成网络的约束,并提出一种去除空中浮点伪影的方法,来达到使用少量图片训练即可生成高质量新视角渲染结果的目的。FSGS^[79] 同样引入了深度监督,并提出了一种高斯上采样的方法提升新视角泛化能力。CoherentGS^[80] 提出了一种一致性 3DGS 优化方法,并引入了深度和光流约束。DNGaussian^[81] 改善了上述方法中的深度监督方式,即引入软深度,硬深度,和深度局部归一化,并将 3DGS 点中的球谐系数替换为神经网络,缓解优化过程中的过拟合问题。CoR-GS^[82] 通过训练两个高斯辐射场,并提出了协同修剪点云与伪视图协同正则化的方法来识别和抑制两个辐射场之间的不一致。InstantSplat^[83] 结合了 Dust3r 的几何先验与位姿先验,来为 3D 高斯点云提供初始化点云和位姿。同时,一些研究^[84-88]专注于稀疏视角下的推理式的新视角渲染,即直接通过卷积网络推理而非优化得到 3DGS 表征的系数。Pixelsplat^[87] 与 MVsplat^[84] 通过卷积神经网络推理得到特征图,随即

使用另一个神经网络来推理每个像素对应高斯点云的参数，进而实现推理式高斯表征；基于 Pixelsplat 与 MVSpLat 这种回归式的方法，LatentSplat^[85] 结合了生成式的方法，使得渲染结果保留了更多的高频细节。Gamba^[86] 则注重于单视角图像的新视角渲染与重建。

本文摘录了部分稀疏视角重建方法在 LLFF^[89] 数据集上的定量结果，如表 4 所示，覆盖的视角数包含 3、6、9 视角，供读者参考。表中，每个指标的最优和次优分别用粗体和下划线标出。

表 4 稀疏视角重建结果在 LLFF 数据集上定量比较

| Table 4 Quantitative comparison of sparse view reconstruction results on the LLFF dataset | | | | | | | | | |
|---|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | 3 views | | | 6 views | | | 9 views | | |
| | PSNR↑ | SSIM↑ | LPIPS↓ | PSNR↑ | SSIM↑ | LPIPS↓ | PSNR↑ | SSIM↑ | LPIPS↓ |
| DNGaussian | 19.12 | 0.591 | 0.294 | 22.18 | 0.755 | 0.198 | 23.17 | 0.788 | 0.180 |
| CoherentGS | 20.33 | 0.725 | 0.180 | - | - | - | - | - | - |
| FSGS | <u>20.43</u> | <u>0.682</u> | 0.248 | <u>24.09</u> | <u>0.823</u> | <u>0.145</u> | <u>25.31</u> | <u>0.860</u> | <u>0.122</u> |
| CoR-GS | 20.45 | <u>0.712</u> | <u>0.196</u> | 24.49 | 0.837 | 0.115 | 26.06 | 0.874 | 0.089 |

6 总 结

本文回顾了三维高斯溅射技术在可微分渲染领域的最新研究成果，展现了其在该领域的前沿应用与潜力。然而，由于篇幅所限，本文未能详尽覆盖所有相关话题和领域。例如，三维高斯场景表达的压缩^[90-92]，从运动模糊影像中进行的重建^[93, 94]，以及在重建过程中对相机位姿进行优化^[95, 96]等话题。此外，将三维高斯溅射技术应用于即时定位与地图构建 (Simultaneous Localization And Mapping, SLAM)^[97-99]、三维生成^[100-102]、三维语义分割^[103-105]和编辑^[106, 107]等任务中，也展现出广阔的应用前景，值得进一步的探索与研究。

尽管三维高斯溅射技术在重建效率、渲染质量及速度方面取得了显著成就，但它仍存在一定的局限性。首先，由于三维高斯点云采用的是离散表达方式，辐射场的范围被限制在特定的离散空间内，这使得该技术对于未被直接观测到的区域无法进行有效的外推。其次，三维高斯溅射技术仍然依赖于如 SfM 等技术来获取相机的位姿信息。再者，虽然三维高斯点采用了类似点云的展示形式，但其在本质上并不等同于传统意义上的表面点云，更与图形渲染管线中常用的三角格网存在显著差异，因此需要从业者进行更深入的理解与掌握。

随着生成式扩散模型^[108, 109]的日益成熟和深化，本文预见其与三维高斯溅射技术将实现深度的融合。这种融合将显著提升复杂场景的重建质量，并大幅增强其外推能力。未来，三维高斯溅射技术还极有可能将 SfM 等技术无缝集成至可微分重建管道中，直接从无序影像甚至是网络图片集合中精准地重建出三维辐射场。此外，该技术若能与传统的渲染管线相融

合，并实现多平台的广泛部署，必将为三维实景、数字文物以及元宇宙等前沿领域提供坚实而强大的技术支撑。

参考文献

- [1] 韩开, 徐娟. 3D 场景渲染技术—神经辐射场的研究综述[J]. 计算机应用研究, 2024, 41(8).
KAI HAN, JUAN XU. Comprehensive review of 3d scene rendering technique-neural radiance fields[J]. Application Research of Computers, 2024, 41(8).
- [2] 何高湘, 朱斌, 解博, 等. 基于神经辐射场的新视角合成研究进展[J]. 激光与光电子学进展, 2024, 61(12):1200005.
GAOXIANG HE, BIN ZHU, BO XIE, et al. Progress in novel view synthesis using neural radiance fields[J]. Laser & Optoelectronics Progress, 2024, 61(12):1200005.
- [3] 李吉洋, 程乐超, 何靖璇, 等. 神经辐射场的研究现状与展望[J]. 计算机辅助设计与图形学学报, 2024.
JIYANG LI, LECHAO CHENG, JINGXUAN HE, et al. Current status and prospects of research on neural radiance fields[J]. Journal of Computer-Aided Design & Computer Graphics, 2024.
- [4] 陈琪, 秦芝宝, 蔡晓誉, 等. 面向机器人手术模拟器的神经辐射场软组织动态三维重建[J]. 光学学报, 2024, 44(07):0733001.
QI CHEN, ZHIBAO QIN, XIAOYU CAI, et al. Dynamic three-dimensional reconstruction of soft tissue in neural radiation field for robotic surgery simulators[J]. Acta Optica Sinica, 2024, 44(07):0733001.
- [5] 侯耀斐, 黄海松, 范青松, 等. 基于改进多层感知机的神经辐射场三维重建方法[J]. 激光与光电子学进展, 2024, 61(04):0415004.
YAOFEI HOU, HAI SONG HUANG, QINGSONG FAN, et al. 3D reconstruction of neural radiation field based on improved multiple layer perceptron[J]. Laser & Optoelectronics Progress, 2024, 61(04):0415004.
- [6] 苗源, 刘畅, 邱钧. 基于神经辐射场的光场角度域超分辨[J]. 光学学报, 2023, 43(14):1411001.
YUAN MIAO, CHANG LIU, JUN QIU. Neural radiance field-based light field super-resolution in angular domain[J]. Acta Optica Sinica, 2023, 43(14):1411001.
- [7] MILDENHALL B, SRINIVASAN P P, TANCIK M, et al. Nerf: Representing scenes as neural radiance fields for view synthesis [J]. Communications of the ACM, 2021,65(1): 99-106.
- [8] KERBL B, KOPANAS G, LEIMKÜHLER T, et al. 3D gaussian splatting for real-time radiance field rendering [J]. ACM Transactions on Graphics, 2023,42(4): 1-14.
- [9] AGARWAL S, FURUKAWA Y, SNAVELY N, et al. Building rome in a day [J]. Communications of the ACM, 2011,54(10): 105-12.
- [10] CRANDALL D, OWENS A, SNAVELY N, et al. Discrete-continuous optimization for large-scale structure from motion [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2011: 3001-8.
- [11] MOULON P, MONASSE P, MARLET R. Global fusion of relative motions for robust, accurate and scalable structure from motion [C]//Proceedings of the IEEE international conference on

- computer vision. 2013: 3248-55.
- [12] SNAVELY N, SEITZ S M, SZELISKI R. Modeling the world from internet photo collections [J]. International journal of computer vision, 2008,80:189-210.
 - [13] SCHONBERGER J L, FRAHM J-M. Structure-from-motion revisited [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 4104-13.
 - [14] WU C. Towards linear-time incremental structure from motion [C]//International Conference on 3D Vision. 2013: 127-34.
 - [15] GU X, FAN Z, ZHU S, et al. Cascade cost volume for high-resolution multi-view stereo and stereo matching [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 2495-504.
 - [16] GALLIANI S, LASINGER K, SCHINDLER K. Massively parallel multiview stereopsis by surface normal diffusion [C]//Proceedings of the IEEE international conference on computer vision. 2015: 873-81.
 - [17] YAO Y, LUO Z, LI S, et al. Mvsnet: Depth inference for unstructured multi-view stereo [C]//Proceedings of the European conference on computer vision. 2018: 767-83.
 - [18] SCHÖNBERGER J L, ZHENG E, POLLEFEYS M, et al. Pixelwise view selection for unstructured multi-view stereo [C]//European Conference on Computer Vision (ECCV). 2016: 501-18.
 - [19] YAO Y, LUO Z, LI S, et al. Recurrent mvsnet for high-resolution multi-view stereo depth inference [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 5525-34.
 - [20] ZHANG J, LI S, LUO Z, et al. Vis-mvsnet: Visibility-aware multi-view stereo network [J]. International Journal of Computer Vision, 2023,131(1): 199-214.
 - [21] LIU Y, WANG P, LIN C, et al. Nero: Neural geometry and brdf reconstruction of reflective objects from multiview images [J]. ACM Transactions on Graphics, 2023,42(4): 1-22.
 - [22] YAO Y, ZHANG J, LIU J, et al. Neirlf: Neural incident light field for physically-based material estimation [C]//Proceedings of the European conference on computer vision. 2022: 700-16.
 - [23] ZHANG J, YAO Y, LI S, et al. Neirlf++: Inter-reflectable light fields for geometry and material estimation [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 3601-10.
 - [24] ZWICKER M, PFISTER H, VAN BAAR J, et al. Ewa splatting [J]. IEEE Transactions on Visualization and Computer Graphics, 2002,8(3): 223-38.
 - [25] HUANG L-T, BAI J-X, GUO J, et al. On the error analysis of 3D gaussian splatting and an optimal projection strategy [EB/OL]. (2024-02-01)[2024-05-27] <https://arxiv.org/abs/2402.00752>.
 - [26] CHENG K, LONG X, YANG K, et al. Gaussianpro: 3D gaussian splatting with progressive propagation [EB/OL]. (2024-02-22)[2024-05-27] <https://arxiv.org/abs/2402.14650>.
 - [27] ZHANG Z, HU W, LAO Y, et al. Pixel-gs: Density control with pixel-aware gradient for 3D gaussian splatting [EB/OL]. (2024-03-22)[2024-05-27] <https://arxiv.org/abs/2403.15530>.
 - [28] FANG G, WANG B. Mini-splatting: Representing scenes with a constrained number of gaussians [EB/OL]. (2024-03-21)[2024-05-27] <https://arxiv.org/abs/2403.14166>.
 - [29] ZHANG J, ZHAN F, XU M, et al. Fregs: 3D gaussian splatting with progressive frequency regularization [EB/OL]. (2024-03-11)[2024-05-27] <https://arxiv.org/abs/2403.06908>.
 - [30] LU T, YU M, XU L, et al. Scaffold-gs: Structured 3d gaussians for view-adaptive rendering

- [EB/OL]. (2023-11-30)[2024-05-27] <https://arxiv.org/abs/2312.00109>.
- [31] HAMDI A, MELAS-KYRIAZI L, QIAN G, et al. Ges: Generalized exponential splatting for efficient radiance field rendering [EB/OL]. (2024-02-15)[2024-05-27] <https://arxiv.org/abs/2402.10128>.
- [32] YANG Z, GAO X, SUN Y, et al. Spec-gaussian: Anisotropic view-dependent appearance for 3D gaussian splatting [EB/OL]. (2024-02-24)[2024-05-27] <https://arxiv.org/abs/2402.15870>.
- [33] BARRON J T, MILDENHALL B, VERBIN D, et al. Mip-nerf 360: Unbounded anti-aliased neural radiance fields [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 5470-9.
- [34] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 586-95.
- [35] YU Z, CHEN A, HUANG B, et al. Mip-splatting: Alias-free 3d gaussian splatting [EB/OL]. (2023-11-27)[2024-05-27] <https://arxiv.org/abs/2311.16493>.
- [36] LIANG Z, ZHANG Q, HU W, et al. Analytic-splatting: Anti-aliased 3D gaussian splatting via analytic integration [EB/OL]. (2024-03-17)[2024-05-27] <https://arxiv.org/abs/2403.11056>.
- [37] YAN Z, LOW W F, CHEN Y, et al. Multi-scale 3d gaussian splatting for anti-aliased rendering [EB/OL]. (2023-11-28)[2024-05-27] <https://arxiv.org/abs/2311.17089>.
- [38] SONG X, ZHENG J, YUAN S, et al. Sa-gs: Scale-adaptive gaussian splatting for training-free anti-aliasing [EB/OL]. (2024-03-28)[2024-05-27] <https://arxiv.org/abs/2403.19615>.
- [39] GUÉDON A, LEPETIT V. Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering [EB/OL]. (2023-11-21)[2024-05-27] <https://arxiv.org/abs/2311.12775>.
- [40] GAO L, YANG J, ZHANG B-T, et al. Mesh-based gaussian splatting for real-time large-scale deformation [EB/OL]. (2024-02-07)[2024-05-27] <https://arxiv.org/abs/2402.04796>.
- [41] WACZYŃSKA J, BORYCKI P, TADEJA S, et al. Games: Mesh-based adapting and modification of gaussian splatting [EB/OL]. (2024-02-02)[2024-05-27] <https://arxiv.org/abs/2402.01459>.
- [42] LYU X, SUN Y-T, HUANG Y-H, et al. 3dgsr: Implicit surface reconstruction with 3d gaussian splatting [EB/OL]. (2024-03-30)[2024-05-27] <https://arxiv.org/abs/2404.00409>.
- [43] CHEN H, LI C, LEE G H. Neusg: Neural implicit surface reconstruction with 3d gaussian splatting guidance [EB/OL]. (2023-12-01)[2024-05-27] <https://arxiv.org/abs/2312.00846>.
- [44] TURKULAINEN M, REN X, MELEKHOV I, et al. Dn-splatter: Depth and normal priors for gaussian splatting and meshing [EB/OL]. (2024-03-26)[2024-05-27] <https://arxiv.org/abs/2403.17822>.
- [45] HUANG B, YU Z, CHEN A, et al. 2d gaussian splatting for geometrically accurate radiance fields [EB/OL]. (2024-03-26)[2024-05-27] <https://arxiv.org/abs/2403.17888>.
- [46] LIANG Z, ZHANG Q, FENG Y, et al. Gs-ir: 3d gaussian splatting for inverse rendering [EB/OL]. (2023-11-26)[2024-05-27] <https://arxiv.org/abs/2311.16473>.
- [47] JIANG Y, TU J, LIU Y, et al. Gaussianshader: 3D gaussian splatting with shading functions for reflective surfaces [EB/OL]. (2023-11-29)[2024-05-27] <https://arxiv.org/abs/2311.17977>.
- [48] SHI Y, WU Y, WU C, et al. Gir: 3d gaussian inverse rendering for relightable scene factorization [EB/OL]. (2023-12-08)[2024-05-27] <https://arxiv.org/abs/2312.05133>.
- [49] GAO J, GU C, LIN Y, et al. Relightable 3d gaussian: Real-time point cloud relighting with brdf

- decomposition and ray tracing [EB/OL]. (2023-11-27)[2024-05-27] <https://arxiv.org/abs/2311.16043>.
- [50] DU Y, ZHANG Y, YU H-X, et al. Neural radiance flow for 4d view synthesis and video processing [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 14304-14.
 - [51] GAO C, SARAF A, KOPF J, et al. Dynamic view synthesis from dynamic monocular video [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 5712-21.
 - [52] SONG L, CHEN A, LI Z, et al. Nerfplayer: A streamable dynamic scene representation with decomposed neural radiance fields [J]. IEEE Transactions on Visualization and Computer Graphics, 2022,29(5): 2732-42.
 - [53] SHAO R, ZHENG Z, TU H, et al. Tensor4d: Efficient neural 4d decomposition for high-fidelity dynamic reconstruction and rendering [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 16632-42.
 - [54] FRIDOVICH-KEIL S, MEANTI G, WARBURG F, et al. K-planes: Explicit radiance fields in space, time, and appearance [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023: 12479-88.
 - [55] CAO A, JOHNSON J. Hexplane: A fast representation for dynamic scenes [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 130-41.
 - [56] JANG H, KIM D. D-tensorf: Tensorial radiance fields for dynamic scenes [EB/OL]. (2022-12-05)[2024-05-27] <https://arxiv.org/abs/2212.02375>.
 - [57] WU G, YI T, FANG J, et al. 4d gaussian splatting for real-time dynamic scene rendering [EB/OL]. (2023-12-07)[2024-05-27] <https://arxiv.org/abs/2310.08528>.
 - [58] SHAO R, SUN J, PENG C, et al. Control4d: Efficient 4d portrait editing with text [EB/OL]. (2023-03-31)[2024-05-27] <https://arxiv.org/abs/2305.20082>.
 - [59] CHEN A, XU Z, GEIGER A, et al. Tensorf: Tensorial radiance fields [C]//Proceedings of the European Conference on Computer Vision. 2022: 333-50.
 - [60] YANG Z, GAO X, ZHOU W, et al. Deformable 3d gaussians for high-fidelity monocular dynamic scene reconstruction [EB/OL]. (2023-09-22)[2024-05-27] <https://arxiv.org/abs/2309.13101>.
 - [61] LIANG Y, KHAN N, LI Z, et al. Gaufré: Gaussian deformation fields for real-time dynamic novel view synthesis [EB/OL]. (2023-12-18)[2024-05-27] <https://arxiv.org/abs/2312.11458>.
 - [62] HUANG Y-H, SUN Y-T, YANG Z, et al. Sc-gs: Sparse-controlled gaussian splatting for editable dynamic scenes [EB/OL]. (2023-12-04)[2024-05-27]
 - [63] KRATIMENOS A, LEI J, DANIILIDIS K. Dynmf: Neural motion factorization for real-time dynamic view synthesis with 3d gaussian splatting [EB/OL]. (2023-11-30)[2024-05-27] <https://arxiv.org/abs/2312.00112>.
 - [64] DAS D, WEWER C, YUNUS R, et al. Neural parametric gaussians for monocular non-rigid object reconstruction [EB/OL]. (2023-12-02)[2024-05-27] <https://arxiv.org/abs/2312.01196>.
 - [65] YU H, JULIN J, MILACSKI Z Á, et al. Cogs: Controllable gaussian splatting [EB/OL]. (2023-12-09)[2024-05-27] <https://arxiv.org/abs/2312.05664>.
 - [66] LIN Y, DAI Z, ZHU S, et al. Gaussian-flow: 4d reconstruction with dynamic 3d gaussian particle [EB/OL]. (2023-12-06)[2024-05-27] <https://arxiv.org/abs/2312.03431>.
 - [67] LUITEN J, KOPANAS G, LEIBE B, et al. Dynamic 3d gaussians: Tracking by persistent

- dynamic view synthesis [EB/OL]. (2023-08-18)[2024-05-27] <https://arxiv.org/abs/2308.09713>.
- [68] LI Z, CHEN Z, LI Z, et al. Spacetime gaussian feature splatting for real-time dynamic view synthesis [EB/OL]. (2023-12-28)[2024-05-27] <https://arxiv.org/abs/2312.16812>.
- [69] KATSUMATA K, VO D M, NAKAYAMA H. An efficient 3d gaussian representation for monocular/multi-view dynamic scenes [EB/OL]. (2023-11-21)[2024-05-27] <https://arxiv.org/abs/2311.12897>.
- [70] YANG Z, YANG H, PAN Z, et al. Real-time photorealistic dynamic scene representation and rendering with 4d gaussian splatting [EB/OL]. (2023-10-16)[2024-05-27] <https://arxiv.org/abs/2310.10642>.
- [71] PUMAROLA A, CORONA E, PONS-MOLL G, et al. D-nerf: Neural radiance fields for dynamic scenes [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10318-27.
- [72] SUZUKI T. Fed3dgs: Scalable 3D gaussian splatting with federated learning [EB/OL]. (2024-03-18)[2024-05-27] <https://arxiv.org/abs/2403.11460>.
- [73] LIU Y, GUAN H, LUO C, et al. Citygaussian: Real-time high-quality large-scale scene rendering with gaussians [EB/OL]. (2024-08-01)[2024-05-27] <https://arxiv.org/abs/2404.01133>.
- [74] LIN J, LI Z, TANG X, et al. Vastgaussian: Vast 3D gaussians for large scene reconstruction [EB/OL]. (2024-02-27)[2024-05-27] <https://arxiv.org/abs/2402.17427>.
- [75] TURKI H, RAMANAN D, SATYANARAYANAN M. Mega-nerf: Scalable construction of large-scale nerfs for virtual fly-throughs [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 12922-31.
- [76] ZHENXING M I, XU D. Switch-nerf: Learning scene decomposition with mixture of experts for large-scale neural radiance fields [C]//The Eleventh International Conference on Learning Representations. 2022
- [77] XU L, XIANGLI Y, PENG S, et al. Grid-guided neural radiance fields for large urban scenes [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 8296-306.
- [78] XIONG H, MUTTUKURU S, UPADHYAY R, et al. Sparsegts: Real-time 360° sparse view synthesis using gaussian splatting [EB/OL]. (2023-11-30)[2024-05-27] <https://arxiv.org/abs/2312.00206>.
- [79] ZHU Z, FAN Z, JIANG Y, et al. Fsgs: Real-time few-shot view synthesis using gaussian splatting [EB/OL]. (2023-12-01)[2024-05-27] <https://arxiv.org/abs/2312.00451>.
- [80] PALIWAL A, YE W, XIONG J, et al. Coherentgts: Sparse novel view synthesis with coherent 3D gaussians [EB/OL]. (2024-03-28)[2024-05-27] <https://arxiv.org/abs/2403.19495>.
- [81] LI J, ZHANG J, BAI X, et al. Dngaussian: Optimizing sparse-view 3D gaussian radiance fields with global-local depth normalization [EB/OL]. (2024-03-11)[2024-05-27] <https://arxiv.org/abs/2403.06912>.
- [82] ZHANG J, LI J, YU X, et al. Cor-gs: Sparse-view 3D gaussian splatting via co-regularization [EB/OL]. (2024-05-20)[2024-05-27] <https://arxiv.org/pdf/2405.12110>.
- [83] FAN Z, CONG W, WEN K, et al. Instantsplat: Unbounded sparse-view pose-free gaussian splatting in 40 seconds [EB/OL]. (2024-03-29)[2024-05-27] <https://arxiv.org/abs/2403.20309>.
- [84] CHEN Y, XU H, ZHENG C, et al. Mvsplat: Efficient 3D gaussian splatting from sparse multi-view images [EB/OL]. (2024-03-21)[2024-05-27] <https://arxiv.org/abs/2403.14627>.
- [85] WEWER C, RAJ K, ILG E, et al. Latentsplat: Autoencoding variational gaussians for fast

- generalizable 3D reconstruction [EB/OL]. (2024-03-24)[2024-05-27] <https://arxiv.org/abs/2403.16292>.
- [86] SHEN Q, YI X, WU Z, et al. Gamba: Marry gaussian splatting with mamba for single view 3D reconstruction [EB/OL]. (2024-03-27)[2024-05-27] <https://arxiv.org/abs/2403.18795>.
- [87] CHARATAN D, LI S, TAGLIASACCHI A, et al. Pixelsplat: 3d gaussian splats from image pairs for scalable generalizable 3d reconstruction [EB/OL]. (2023-12-19)[2024-05-27] <https://arxiv.org/abs/2312.12337>.
- [88] SZYMANOWICZ S, RUPPRECHT C, VEDALDI A. Splatter image: Ultra-fast single-view 3d reconstruction [EB/OL]. (2023-12-20)[2024-05-27] <https://arxiv.org/abs/2312.13150>.
- [89] MILDENHALL B, SRINIVASAN P P, ORTIZ-CAYON R, et al. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines [J]. ACM Transactions on Graphics, 2019,38(4): 1-14.
- [90] NIEDERMAYR S, STUMPFEGGER J, WESTERMANN R. Compressed 3d gaussian splatting for accelerated novel view synthesis [EB/OL]. (2023-11-17)[2024-05-27] <https://arxiv.org/abs/2401.02436>.
- [91] FAN Z, WANG K, WEN K, et al. Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200+ fps [EB/OL]. (2023-11-28)[2024-05-27] <https://arxiv.org/abs/2311.17245>.
- [92] LEE J C, RHO D, SUN X, et al. Compact 3d gaussian representation for radiance field [EB/OL]. (2023-11-22)[2024-05-27] <https://arxiv.org/abs/2311.13681>.
- [93] ZHAO L, WANG P, LIU P. Bad-gaussians: Bundle adjusted deblur gaussian splatting [EB/OL]. (2024-03-18)[2024-05-27] <https://arxiv.org/abs/2403.11831>.
- [94] LEE B, LEE H, SUN X, et al. Deblurring 3D gaussian splatting [EB/OL]. (2024-06-01)[2024-05-27] <https://arxiv.org/abs/2401.00834>.
- [95] FU Y, LIU S, KULKARNI A, et al. Colmap-free 3d gaussian splatting [EB/OL]. (2023-12-12)[2024-05-27] <https://arxiv.org/pdf/2312.07504>.
- [96] DARMON F, PORZI L, ROTA-BULÒ S, et al. Robust gaussian splatting [EB/OL]. (2024-04-05)[2024-05-27]
- [97] LI M, LIU S, ZHOU H. Sgs-slam: Semantic gaussian splatting for neural dense slam [EB/OL]. (2024-02-05)[2024-05-27] <https://arxiv.org/abs/2402.03246>.
- [98] YAN C, QU D, WANG D, et al. Gs-slam: Dense visual slam with 3d gaussian splatting [EB/OL]. (2023-11-20)[2024-05-27] <https://arxiv.org/abs/2311.11700>.
- [99] MATSUKI H, MURAI R, KELLY P H J, et al. Gaussian splatting slam [EB/OL]. (2023-12-11)[2024-05-27] <https://arxiv.org/abs/2312.06741>.
- [100] CHEN Z, WANG F, LIU H. Text-to-3d using gaussian splatting [EB/OL]. (2023-09-28)[2024-05-27] <https://arxiv.org/abs/2309.16585>.
- [101] TANG J, REN J, ZHOU H, et al. Dreamgaussian: Generative gaussian splatting for efficient 3d content creation [EB/OL]. (2023-09-28)[2024-05-27] <https://arxiv.org/abs/2309.16653>.
- [102] LING H, KIM S W, TORRALBA A, et al. Align your gaussians: Text-to-4d with dynamic 3d gaussians and composed diffusion models [EB/OL]. (2023-12-21)[2024-05-27] <https://arxiv.org/abs/2312.13763>.
- [103] LAN K, LI H, SHI H, et al. 2d-guided 3d gaussian segmentation [EB/OL]. (2023-12-26)[2024-05-27] <https://arxiv.org/abs/2312.16047>.
- [104] HU X, WANG Y, FAN L, et al. Sagd: Boundary-enhanced segment anything in 3D gaussian via gaussian decomposition [EB/OL]. (2024-06-31)[2024-05-27] <https://arxiv.org/abs/2401.17857>.

- [105] CEN J, FANG J, YANG C, et al. Segment any 3d gaussians [EB/OL]. (2023-12-01)[2024-05-27] <https://arxiv.org/abs/2312.00860>.
- [106] CHEN Y, CHEN Z, ZHANG C, et al. Gaussianeditor: Swift and controllable 3d editing with gaussian splatting [EB/OL]. (2023-11-24)[2024-05-27] <https://arxiv.org/abs/2311.14521>.
- [107] FANG J, WANG J, ZHANG X, et al. Gaussianeditor: Editing 3d gaussians delicately with text instructions [EB/OL]. (2023-11-27)[2024-05-27] <https://arxiv.org/abs/2311.16037>.
- [108] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 10684-95.
- [109] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models [J]. Advances in neural information processing systems, 2020,336840-51.



网络首发:

标题: 基于三维高斯溅射技术的可微分渲染研究进展

作者: 高建, 陈林卓, 沈秋, 曹汛, 姚遥

收稿日期: 2024-05-27

录用日期: 2024-06-27

DOI: 10.3788/LOP241369

引用格式:

高建, 陈林卓, 沈秋, 曹汛, 姚遥. 基于三维高斯溅射技术的可微分渲染研究进展[J]. 激光与光电子学进展, 2024, 61(16): 1611010.

网络首发文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

光学相干层析成像在生物医学中的研究进展

颜佳涛 唐云 袁锐 邹鹏辉 廖萍 李菁锋 彭旭翔

湖南科技大学物理与电子科学学院智能传感器与新型传感材料湖南省重点实验室, 湖南 湘潭 411201

激光与光电子学进展, 2024, 61(12): 1200004

基于神经辐射场的新视角合成研究进展

何高湘 朱斌 解博 陈熠

国防科技大学电子对抗学院脉冲功率激光技术国家重点实验室, 安徽 合肥 230037

激光与光电子学进展, 2024, 61(12): 1200005

用于全息三维显示的数据获取方法进展

佟卓键 桂进斌 胡磊 胡先飞

昆明理工大学理学院, 云南 昆明 650500

激光与光电子学进展, 2024, 61(10): 1000007

基于改进多层感知机的神经辐射场三维重建方法

侯耀斐 黄海松 范青松 肖婧 韩正功

贵州大学现代制造技术教育部重点实验室, 贵州 贵阳 550025

激光与光电子学进展, 2024, 61(4): 0415004

基于相移条纹分析的相位误差补偿技术发展综述 (特邀)

李乐阳 吴周杰 张启灿

四川大学电子信息学院, 四川 成都 610065

激光与光电子学进展, 2024, 61(2): 0211008