3D Gaussian Splatting: Survey, Technologies, Challenges, and Opportunities

3D 高斯点绘 (3D Gaussian Splatting): 综述、技术、挑战与机遇

Yanqi Bao, Tianyu Ding, Jing Huo*, Yaoli Liu, Yuxin Li, Wenbin Li, Yang Gao, Member, IEEE, and Jiebo Luo, Fellow, IEEE,

鲍彦琦，丁天宇，霍晶 *，刘耀利，李宇新，李文斌，高扬，IEEE 会员，罗杰博，IEEE 会士，

Abstract-3D Gaussian Splatting (3DGS) has emerged as a prominent technique with the potential to become a mainstream method for 3D representations. It can effectively transform multiview images into explicit 3D Gaussian through efficient training, and achieve real-time rendering of novel views. This survey aims to analyze existing 3DGS-related works from multiple intersecting perspectives, including related tasks, technologies, challenges, and opportunities. The primary objective is to provide newcomers with a rapid understanding of the field and to assist researchers in methodically organizing existing technologies and challenges. Specifically, we delve into the optimization, application, and extension of 3DGS, categorizing them based on their focuses or motivations. Additionally, we summarize and classify nine types of technical modules and corresponding improvements identified in existing works. Based on these analyses, we further examine the common challenges and technologies across various tasks, proposing potential research opportunities.

摘要-3D 高斯点绘 (3DGS) 作为一种具有潜力成为主流三维表示方法的突出技术，能够通过高效训练将多视角图像转换为显式的三维高斯表示，并实现新视角的实时渲染。本文综述旨在从相关任务、技术、挑战与机遇等多个交叉视角分析现有 3DGS 相关工作。主要目标是帮助新入门者快速理解该领域，并协助研究者系统梳理现有技术与挑战。具体而言，我们深入探讨了 3DGS 的优化、应用与扩展，并根据其关注点或动机进行分类。此外，我们总结并归纳了现有工作中识别出的九类技术模块及相应改进。基于这些分析，我们进一步考察了各类任务中的共性挑战与技术，提出潜在的研究机会。

Index Terms-3D Representations, Rendering, 3D Gaussian Splatting.

关键词-三维表示，渲染，3D 高斯点绘。

I. INTRODUCTION

一、引言

TI HE advent of Neural Radiance Fields (NeRF) [1] has ignited considerable interest in the pursuit of photorealistic 3D content. Despite substantial recent advancements that have markedly enhanced NeRF's potential for practical applications, its inherent efficiency challenges have remained unresolved. The introduction of 3D Gaussian Splatting (3DGS) has decisively addressed this bottleneck, enabling high-quality real-time ($\geq 30\text{fps}$) novel view synthesis at 1080p resolution. This rapid development has quickly attracted researchers and led to a proliferation of related works.

神经辐射场 (Neural Radiance Fields，NeRF)[1] 的出现激发了对逼真三维内容的广泛关注。尽管近期取得了显著进展，大幅提升了 NeRF 在实际应用中的潜力，但其固有的效率瓶颈仍未解决。3D 高斯点绘 (3DGS) 的引入有效突破了这一瓶颈，实现了高质量实时 ($\geq$ 30fps) 新视角合成，分辨率达到 1080p。这一快速发展迅速吸引了研究者，催生了大量相关工作。

Owing to the efficiency and controllable explicit representation of 3DGS, its applications extend across a diverse array of fields. These include enhancing immersive environments in Virtual Reality (VR) and Augmented Reality (AR), improved spatial awareness in robotics and autonomous systems, and urban planning and architecture, etc.

由于 3DGS 具有效率高且显式表示可控的优势，其应用涵盖了广泛领域，包括增强虚拟现实 (VR) 和增强现实 (AR) 中的沉浸式环境、机器人与自主系统中的空间感知提升，以及城市规划与建筑等。

To assist readers in quickly grasping the progress in 3DGS research, we provide a comprehensive survey of 3DGS and its derivative works. This survey systematically compiles the most important and recent literature on the subject, offering detailed classifications and discussions of their tasks and techniques. In examining the technological commonalities across numerous 3DGS variants, we present a structured analysis of technical improvements in fundamental components of vanilla 3DGS, such as initialization, attribute configurations and regularization, etc. Based on this summary of techniques, we aim to help readers synthesize the connections among different improved techniques and provide approaches to enhance various components of vanilla 3DGS to meet their customized tasks. Moreover, we conduct a systematic investigation into the relationships between downstream tasks and their enabling technologies in 3DGS, identifying and analyzing four fundamental challenges. Through careful examination of these challenges, we propose promising research directions to advance this rapidly evolving field, providing a roadmap for future innovations in 3DGS.

为帮助读者快速把握 3DGS 研究进展，我们提供了对 3DGS 及其衍生工作的全面综述。该综述系统整理了该领域最重要且最新的文献，详细分类并讨论了其任务与技术。在考察众多 3DGS 变体的技术共性时，我们对基础 3DGS 的关键组成部分如初始化、属性配置与正则化等技术改进进行了结构化分析。基于技术总结，我们旨在帮助读者综合不同改进技术之间的联系，并提供提升基础 3DGS 各组件以满足定制任务的途径。此外，我们系统探讨了下游任务与其支撑技术之间的关系，识别并分析了四大核心挑战。通过对这些挑战的深入审视，我们提出了推动该快速发展领域的有前景研究方向，为 3DGS 未来创新提供路线图。
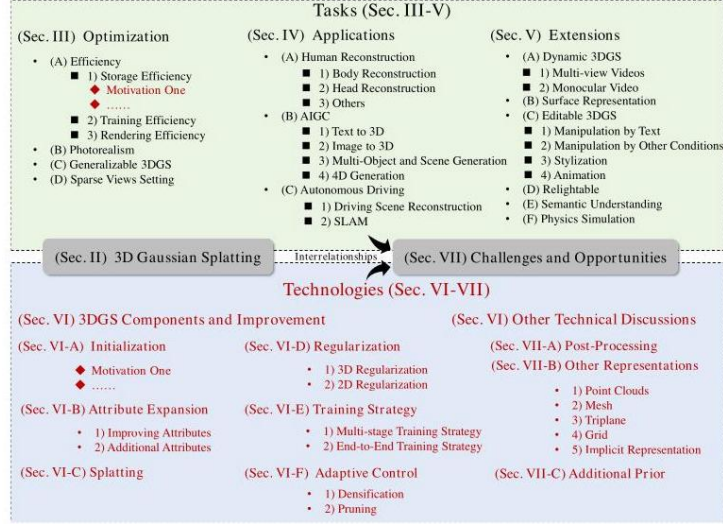
Fig. 1: The introduction of this survey, with the RED parts indicating the new content compared to existing reviews.

图 1: 本综述介绍，红色部分表示相较于现有综述的新内容。

Although some comprehensive reviews have documented recent advances in 3DGS [2]-[4], they focus on categorizing and discussing existing works by downstream tasks, overlooking technical connections between different tasks, which leads to redundant discussions. Our work is distinctive in providing discussions at two levels: Tasks and Techniques. Specifically, we categorize and discuss existing downstream tasks according to their different motivations or focuses, rather than reviewing all works sequentially. More significantly, we present a thorough examination of technical improvements implemented across various modules of the vanilla 3DGS by existing variants, enabling readers to establish clear relationships between different research fields that share similar methodological foundations. Building upon these, we further investigate the their underlying commonalities and delineate core challenges and opportunities, as shown in Fig. 1. Through this approach, we aim to synthesize recent technical breakthroughs in the 3DGS field and direct researchers' attention to the core unique challenges facing 3DGS. The primary contributions of this survey can be summarized as follows:

尽管已有一些综合性综述记录了 3DGS 的最新进展 [2]-[4]，但它们主要从下游任务角度分类讨论现有工作，忽视了不同任务间的技术联系，导致讨论重复。我们的工作独特之处在于从任务与技术两个层面展开讨论。具体而言，我们根据不同动机或关注点对现有下游任务进行分类讨论，而非按顺序回顾所有工作。更重要的是，我们深入审视了现有变体在基础 3DGS 各模块中实施的技术改进，使读者能够明确不同研究领域间基于相似方法论的联系。在此基础上，我们进一步探讨其潜在共性，勾勒核心挑战与机遇，如图 1 所示。通过此方法，我们旨在整合 3DGS 领域的最新技术突破，引导研究者关注 3DGS 面临的核心独特挑战。本综述的主要贡献可总结如下：

1) This survey discusses 3DGS and its various derivative tasks, including the optimization, application, and extension of 3DGS. We provide a classification and discussion based on focuses or motivations, which enables readers to gain a more comprehensive understanding of tasks and establish research directions.

Yanqi Bao, Jing Huo, Yaoli Liu, Yuxin Li, Wenbin Li and Yang Gao are with the State Key Laboratory for Novel Software Technology, Nanjing University, China, 210023 (e-mail: {yq_bao, yaoliliu, liyuxin16}@smail.nju.edu.cn; {huojing, liwenbin, gaoy}@nju.edu.cn).

1) 本综述讨论了三维高斯散点 (3DGS) 及其各种衍生任务，包括 3DGS 的优化、应用和扩展。我们基于关注点或动机对其进行了分类和讨论，使读者能够更全面地理解相关任务并确立研究方向。

2) More importantly, we comprehensively analyze the enhancements in various techniques within 3DGS across different works, offering detailed classifications and in-depth discussions. This enables readers to discern the interrelationships among various improved techniques, thereby assisting in their application to customized tasks.

2) 更重要的是，我们全面分析了不同工作中 3DGS 各项技术的改进，提供了详细的分类和深入的讨论，帮助读者辨别各种改进技术之间的内在联系，从而辅助其在定制任务中的应用。

3) Based on the analysis of existing works and techniques, we identify the commonalities and associations among 3DGS-related tasks and summarize the core challenges.

3) 基于对现有工作和技术的分析，我们识别了 3DGS 相关任务的共性与关联，并总结了核心挑战。

4) In addressing common challenges, this survey elucidates potential opportunities and provides insightful analyses.

4) 针对常见挑战，本综述阐明了潜在机遇并提供了富有洞见的分析。

5) We have published an open-source project on GitHub that compiles 3DGS-related articles, and will continue to add new works and technologies into this project. https://github.com/ qqqqqy0227/awesome-3DGS. We hope that researchers can use it to access the latest research information.

5) 我们在 GitHub 上发布了一个开源项目，汇编了与 3DGS 相关的文献，并将持续添加新的工作和技术。https://github.com/qqqqqy0227/awesome-3DGS。希望研究人员能借此获取最新的研究信息。

II. PRELIMINARIES

二、预备知识

鲍彦琦，霍晶，刘耀利，李宇新，李文斌，高扬均隶属于中国南京大学新型软件技术国家重点实验室，邮编 210023(电子邮箱:{yq_bao, yaoliliu, liyuxin16}@smail.nju.edu.cn；{huojing, liwenbin, gaoy}@nju.edu.cn)。

Tianyu Ding is with the Applied Sciences Group, Microsoft Corporation, Redmond, USA (e-mail: tianyuding@microsoft.com).

丁天宇隶属于美国雷德蒙微软公司应用科学组 (电子邮箱:tianyuding@microsoft.com)。

Jiebo Luo is with the Department of Computer Science, University of Rochester, America (e-mail: jluo@cs.rochester.edu).

Jiebo Luo 任职于美国罗切斯特大学计算机科学系 (电子邮件:jluo@cs.rochester.edu)。

*Corresponding authors: Jing Huo.

* 通讯作者: 霍晶。

In this section, we primarily introduce 3DGS to the readers. More background, including Neural Implicit Fields [5] and Point-based Rendering [6], is described in Supplementary-A.

本节主要向读者介绍 3DGS。更多背景内容，包括神经隐式场 (Neural Implicit Fields)[5] 和基于点的渲染 (Point-based Rendering)[6]，详见补充材料 A。

3DGS [7] combines the advantages of neural implicit field and point-based rendering methods, achieving the high-fidelity rendering quality of the former while maintaining the real-time rendering capability of the latter. Specifically, as shown in Fig. 2, 3DGS defines points in the point cloud as 3D Gaussian primitives with volumetric density:

3DGS [7] 结合了神经隐式场和基于点的渲染方法的优势，实现了前者的高保真渲染质量，同时保持了后者的实时渲染能力。具体而言，如图 2 所示，3DGS 将点云中的点定义为具有体积密度的三维高斯基元:

$$G\left(\mathbf{x}\right) = \exp\left(-\frac{1}{2}(\mathbf{x})^T \sum{}^{-1}(\mathbf{x})\right), \tag{1}$$

where $\sum$ is the 3D covariance matrix and $\mathbf{x}$ is the position from the point (Gaussian mean) μ. To ensure the semi-positive definiteness of the covariance matrix, 3DGS reparameterizes the covariance matrix as a combination of a rotation matrix $\mathbf{R}$ and a scaling matrix $\mathbf{S}$:

其中 $\sum$ 为三维协方差矩阵，$\mathbf{x}$ 为点的位置 (高斯均值) μ。为确保协方差矩阵的半正定性，3DGS 将协方差矩阵重新参数化为旋转矩阵 $\mathbf{R}$ 与缩放矩阵 $\mathbf{S}$ 的组合:

$$\sum = \mathbf{R}\mathbf{S}\mathbf{S}^T\mathbf{R}^T, \tag{2}$$

where the 3D scaling matrix $\mathbf{S}$ can be represented by a 3D vector s, and the rotation matrix $\mathbf{R}$ is obtained through a learnable quaternion q, resulting in a total of 7 learnable parameters. Compared to the commonly employed Cholesky decomposition, which guarantees the semi-positive-definiteness of matrices, the reparameterization method utilized by 3DGS, albeit introducing an additional learnable parameter, facilitates the imposition of geometric constraints on Gaussian primitives (e.g., constraining the scaling vector to give Gaussian primitives a flattened characteristic). In addition to geometric attributes, each Gaussian primitive also stores an opacity α and a set of learnable Spherical Harmonic (SH) parameters to represent view-dependent appearance. Thus, the collection of all primitives can be regarded as a discretized representation that only stores the non-empty parts of the neural field.

其中三维缩放矩阵 $\mathbf{S}$ 可由三维向量 s 表示，旋转矩阵 $\mathbf{R}$ 通过可学习的四元数 q 获得，共计 7 个可学习参数。相比常用的保证矩阵半正定性的 Cholesky 分解，3DGS 采用的重新参数化方法虽引入了额外的可学习参数，但便于对高斯基元施加几何约束 (例如限制缩放向量使高斯基元呈扁平特性)。除几何属性外，每个高斯基元还存储一个不透明度 α 及一组可学习的球谐函数 (Spherical Harmonic, SH) 参数以表示视角相关的外观。因此，所有基元的集合可视为仅存储神经场非空部分的离散表示。

At the beginning of training, the initial Gaussian primitives are either initialized from a sparse point cloud provided by Structure-from-Motion or randomly initialized. The initial number of Gaussian primitives may be

insufficient for high-quality scene reconstruction; hence, 3DGS offers a method for adaptively controlling Gaussian primitives. This method evaluates whether a primitive is "under-reconstructed" or "over-reconstructed" by observing the gradient of each Gaussian primitive's position attributes in view space. Based on this evaluation, the method increases the number of Gaussian primitives by cloning or splitting the primitives to enhance scene representation capability. Additionally, the opacity of all Gaussian primitives is periodically reset to zero to mitigate the presence of artifacts during the optimization process. This adaptive process allows 3DGS to start optimization with a smaller initial set of Gaussians, thus alleviating the dependency on dense point clouds that previous point-based differentiable rendering methods required.

> 训练初期，初始高斯基元要么由结构光束法 (Structure-from-Motion) 提供的稀疏点云初始化，要么随机初始化。初始高斯基元数量可能不足以实现高质量场景重建，因此 3DGS 提供了一种自适应控制高斯基元数量的方法。该方法通过观察每个高斯基元在视图空间中位置属性的梯度，评估其是否"重建不足"或"重建过度"。基于此评估，通过克隆或拆分基元增加高斯基元数量，以增强场景表示能力。此外，为减轻优化过程中的伪影，所有高斯基元的不透明度会定期重置为零。该自适应过程使 3DGS 能够以较少的初始高斯基元开始优化，缓解了以往基于点的可微渲染方法对稠密点云的依赖。

During rendering, 3DGS projects 3D Gaussian primitives onto the 2D imaging plane using the EWA splatting method [8] and employs α blending to compute the final pixel color. This process is described in detail in the Supplementary-A.

> 渲染过程中，3DGS 采用 EWA 散点法 (EWA splatting)[8] 将三维高斯基元投影到二维成像平面，并利用 α 混合计算最终像素颜色。该过程详见补充材料 A。

III. OPTIMIZATION OF 3D GAUSSIAN SPLATTING

> ## 三、三维高斯散点的优化

A. Efficiency

> ### A. 效率

Efficiency is one of the core metrics for evaluating 3D reconstruction [9]. In this section, we describe it from three perspectives: storage, training, and rendering efficiency.

> 效率是评估 3D 重建 [9] 的核心指标之一。本节将从存储、训练和渲染效率三个方面进行描述。

1) Storage Efficiency: 3DGS requires millions of different Gaussian primitives to fit the geometry and appearance in a scene, leading to high storage overhead: a typical reconstruction of an outdoor scene often requires several hundred megabytes to multiple gigabytes of explicit storage space. Given that the geometric and appearance attributes of different Gaussian primitives may be highly similar, storing attributes for each primitive individually can lead to potential redundancy. Some quantitative reconstruction results are reported in Table 1.

1) 存储效率:3D 高斯斑点 (3DGS) 需要数百万个不同的高斯基元来拟合场景中的几何形状和外观，导致存储开销较大: 一个典型的户外场景重建通常需要数百兆字节到数吉字节的显式存储空间。鉴于不同高斯基元的几何和外观属性可能高度相似，单独存储每个基元的属性可能导致潜在的冗余。部分定量重建结果见表 1。

Existing works [10]-[12] primarily focus on applying Vector Quantization [13] (VQ) techniques to compress 3DGS.

现有工作 [10]-[12] 主要集中于应用向量量化 (Vector Quantization， VQ)[13] 技术来压缩 3DGS。

Among them, Compact3D [11] applies VQ to compress different attributes into four corresponding codebooks and stores the index of each Gaussian in these codebooks to reduce the storage overhead. After establishing the codebooks, the training gradients are copied and backpropagated to the original non-quantized Gaussian parameters via the codebooks, updating both the quantized and non-quantized parameters, and discarding the non-quantized parameters when the training is done. Despite achieving efficient 3DGS compression, these methods inevitably encounter quantization errors following discretization and remain sensitive to hyperparameter configurations.

其中，Compact3D[11] 将 VQ 应用于将不同属性压缩到四个对应的码本中，并存储每个高斯在这些码本中的索引以减少存储开销。建立码本后，训练梯度通过码本复制并反向传播到原始非量化的高斯参数，更新量化和非量化参数，训练完成后丢弃非量化参数。尽管实现了高效的 3DGS 压缩，这些方法不可避免地在离散化后遇到量化误差，并且对超参数配置仍然敏感。
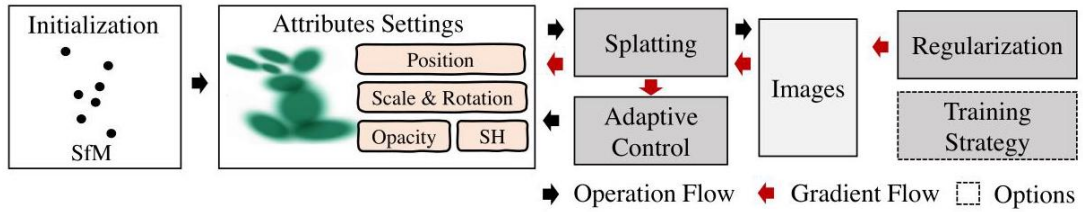

Fig. 2: Pipeline and Technologies of 3D Gaussian Splatting.

图 2:3D 高斯斑点的流程和技术。

TABLE I: Comparison of compression on MipNeRF360 [20]

表 I:MipNeRF360[20] 上的压缩比较

Method	PSNR↑	SSIM↑	LPIPS↓	Size (MB) ↓
3DGS	27.49	0.813	0.222	744.7
Scaffold-GS [16]	27.50	0.806	0.252	253.9
HAC [19]	27.53	0.807	0.238	15.26
Compact-3DGS [21]	27.08	0.798	0.247	48.80
EAGLES	10|	27.15	0.808	0.238	68.89
LightGaussian [15]	27.00	0.799	0.249	44.54
Gaussian-SLAM [18]	26.01	0.772	0.259	23.90
Compact3d [11]	27.16	0.808	0.228	50.30

Method	峰值信噪比 ↑	结构相似性指数 ↑	感知相似性指标 ↓	大小 (MB) ↓
3DGS	27.49	0.813	0.222	744.7
Scaffold-GS [16]	27.50	0.806	0.252	253.9
HAC [19]	27.53	0.807	0.238	15.26
Compact-3DGS [21]	27.08	0.798	0.247	48.80
EAGLES	10|	27.15	0.808	0.238	68.89
LightGaussian [15]	27.00	0.799	0.249	44.54
Gaussian-SLAM [18]	26.01	0.772	0.259	23.90
Compact3d [11]	27.16	0.808	0.228	50.30

Furthermore, some works [14], [15] aim at developing efficient pruning strategies.

此外，一些工作 [14]，[15] 旨在开发高效的剪枝策略。

LightGaussian [15] introduces a Gaussian pruning strategy based on the global significance score and a distillation strategy for high-degree spherical harmonic parameters. Similarly, the work by Lee et al. [14] introduces a learnable mask to reduce the number of original Gaussians. Such methods heavily rely on determining which primitives are non-essential.

LightGaussian[15] 引入了一种基于全局重要性评分的高斯剪枝策略以及针对高阶球谐参数的蒸馏策略。类似地，Lee 等人 [14] 的工作引入了可学习的掩码以减少原始高斯基元的数量。这些方法在很大程度上依赖于确定哪些基元是非必要的。

Furthermore, there are works [16]-[19] focused on improving efficient Gaussian representations or attributes.

此外，还有一些工作 [16]-[19] 专注于提升高效高斯表示或属性。

Scaffold-GS [16] designs anchors and additional attributes for efficient representation, which have the capability to convert to 3DGS. Based on this representation, Scaffold-GS proposes a set of strategies for the growth and pruning of anchors on multi-resolution voxel grids. Despite their widespread implementation, these approaches face inherent compression limitations due to their unstructured characteristics, which HAC [19] later mitigated through the incorporation of structured hash grid.

Scaffold-GS[16] 设计了锚点和附加属性以实现高效表示，具备转换为三维高斯混合模型 (3DGS) 的能力。基于此表示，Scaffold-GS 提出了一套在多分辨率体素网格上锚点增长与剪枝的策略。尽管这些方法被广泛应用，但由于其无结构特性，存在固有的压缩限制，HAC[19] 通过引入结构化哈希网格缓解了这一问题。

2) Training Efficiency: Improving training efficiency is also important for 3DGS. DISTWAR [22] introduces an advanced technique aimed at accelerating atomic operations in raster-based differentiable rendering applications, which typically encounter significant bottlenecks during gradient computation due to the high volume of atomic updates. By leveraging intra-warp locality in atomic updates and addressing the variability in atomic traffic among warps, DISTWAR implements warp-level reduction of threads at the SM sub-cores using registers. These strategies enable DISTWAR to achieve an average $2.44\times$ acceleration in performance.

2) 训练效率: 提升训练效率对于 3DGS 同样重要。DISTWAR[22] 引入了一种先进技术，旨在加速基于光栅的可微渲染应用中的原子操作，这类应用在梯度计算过程中通常因大量原子更新而成为瓶颈。通过利用原子更新中的线程组内局部性并解决线程组间原子流量的差异，DISTWAR 在 SM 子核中使用寄存器实现了线程的线程组级归约。这些策略使 DISTWAR 实现了平均 $2.44\times$ 倍的性能加速。

3) Rendering Efficiency: Real-time rendering is one of the core advantages of Gaussian-based methods. Some works that improve storage efficiency can simultaneously enhance rendering performance, for example, by reducing the number of Gaussian primitives. Several studies [23], [24] have specifically addressed this issue. After training the 3DGS, the work by [23] involves pre-identifying and excluding unnecessary Gaussian primitives through offline clustering based on their spatial proximity and potential impact on the final rendered 2D image. Furthermore, this work introduces a specialized hardware architecture designed to support this technique, achieving a speedup of $10.7\times$ compared to a GPU.

3) 渲染效率: 实时渲染是基于高斯方法的核心优势之一。一些提升存储效率的工作也能同时增强渲染性能，例如通过减少高斯基元数量。若干研究 [23]，[24] 专门针对该问题进行了探讨。训练完成 3DGS 后，[23] 的工作通过基于空间邻近性和对最终渲染 2D 图像潜在影响的离线聚类，预先识别并剔除不必要的高斯基元。此外，该工作还引入了一种专用硬件架构以支持该技术，实现了相较 GPU $10.7\times$ 的加速。

B. Photorealism

B. 真实感

Photorealism is also a topic worth attention [25]. 3DGS is expected to achieve realistic rendering in various scenarios. Some [26]-[28] focus on optimizing under vanilla settings. Among them, GaussianPro [26] introduces an innovative paradigm for joint 2D-3D training. Building upon the 3D plane definition and patch matching technology, it proposes a progressive Gaussian propagation strategy, which harnesses the consistency of 3D views and projection relationships to refine the rendered 2D depth and normal maps. Although these methods demonstrate superior capabilities in handling artifacts and 3D inconsistencies compared to 3DGS, further exploration is still needed for complex geometric reconstruction. To further mitigate this issue, the work [29] introduces a scalable and efficient N-dimensional Gaussian Mixture Model for fast, accurate high-dimensional modeling without domain-specific heuristics or sacrificing computational efficiency but is limited by its reliance on dense data and conservative culling.

真实感也是一个值得关注的话题 [25]。3DGS 期望在多种场景中实现逼真渲染。一些工作 [26]-[28] 聚焦于在基础设置下的优化。其中，GaussianPro[26] 提出了一种创新的二维-三维联合训练范式。基于三维平面定义和补丁匹配技术，提出了渐进式高斯传播策略，利用视角一致性和投影关系来细化渲染的 2D 深度和法线图。尽管这些方法在处理伪影和三维不一致性方面优于 3DGS，但在复杂几何重建方面仍需进一步探索。为进一步缓解该问题，工作 [29] 引入了一种可扩展且高效的 N 维高斯混合模型，用于快速、准确的高维建模，无需领域特定启发式且不牺牲计算效率，但受限于对密集数据的依赖和保守的剔除策略。

The sharp decline in multi-scale rendering performance is also a topic worth attention [30]-[33]. Among them,

Mip-splatting [31], addressing the issue from the perspective of the sampling rate, introduces a Gaussian low-pass filter and 2D Mip filter based on Nyquist's theorem to constrain the frequency of the Gaussians according to the maximal sampling rate across all observed samples. Then, to address the over-smoothing issue in Mip-splatting, Analytic-Splatting [33] proposes a novel method that analytically approximates the integral of 2D Gaussian signals within pixel window areas through a conditioned logistic function. While it enhances detail fidelity, it comes at the expense of efficiency.

多尺度渲染性能的急剧下降也是一个值得关注的问题 [30]-[33]。其中，Mip-splatting[31] 从采样率角度出发，基于奈奎斯特定理引入了高斯低通滤波器和二维 Mip 滤波器，以根据所有观测样本的最大采样率约束高斯的频率。随后，为解决 Mip-splatting 中过度平滑的问题，Analytic-Splatting[33] 提出了一种新方法，通过条件逻辑函数解析近似二维高斯信号在像素窗口区域内的积分。该方法提升了细节保真度，但以效率为代价。

Other works attempt to reconstruct challenging scenes, such as reflective surfaces [34]-[39] and Deblurring [40]-[43].

其他工作尝试重建具有挑战性的场景，如反射表面 [34]-[39] 和去模糊 [40]-[43]。

GaussianShader [34] reconstructs reflective surfaces by employing a hybrid color representation and integrating the specular GGX [44] and normal estimation module, which encompasses diffuse color, direct specular reflection, and a residual color component that accounts for phenomena such as scattering and indirect light reflections. The work [38] introduces a local Gaussian density mixture representation that combines geometry-guided ray interpolation with neural blending and end-to-end optimization to model high-frequency view-dependent reflections and refractions. Although these methods effectively manage complex reflective surfaces, they unavoidably sacrifice computational efficiency relative to vanilla 3DGS. Recently, the work [37] proposes a deferred shading method for 3DGS, which overcomes the challenge of normal estimation for environment map reflection and demonstrates enhanced efficiency by propagating accurate normals across neighboring Gaussians and per-pixel shading in screen space.

GaussianShader[34] 通过采用混合颜色表示并整合镜面 GGX[44] 和法线估计模块重建反射表面，该模块涵盖漫反射色、直接镜面反射及考虑散射和间接光反射等现象的残余色。工作 [38] 引入了一种局部高斯密度混合表示，结合几何引导的光线插值、神经融合及端到端优化，以建模高频视角依赖的反射和折射。尽管这些方法有效处理复杂反射表面，但相较于基础 3DGS 不可避免地牺牲了计算效率。近期，工作 [37] 提出了一种针对 3DGS 的延迟着色方法，克服了环境贴图反射法线估计的难题，并通过在邻近高斯间传播准确法线及屏幕空间的逐像素着色，展示了提升的效率。

Existing deblurring approaches predominantly focus on motion blur or lens defocusing [40]-[42], introducing blur process modeling to enable sharp reconstruction supervised by blurred images. To address a wider range of scenarios, BAGS [43] introduces a Blur Agnostic robust modeling by a Blur Proposal Network and coarse-to-fine optimization scheme.

现有的去模糊方法主要集中于运动模糊或镜头失焦 [40]-[42]，通过引入模糊过程建模，实现以模糊图像为监督的清晰重建。为应对更广泛的场景，BAGS[43] 提出了一种通过模糊提议网络和粗到细优化方案实现的模糊无关鲁棒建模。

C. Generalizable 3DGS

C. 可泛化的三维高斯散射 (3DGS)

The objective of existing generalizable 3D reconstruction or novel view synthesis tasks is to leverage extensive auxiliary datasets to learn scene-agnostic representations.

现有可泛化三维重建或新视角合成任务的目标是利用大量辅助数据集学习场景无关的表示。

The 3DGS's explicit representation has led to a substantial body of works focused on using reference images to directly infer corresponding Gaussian primitives on a per-pixel basis, which are subsequently employed to render images from target views [45], [46]. To achieve this, early works such as Splatter Image [45] propose a novel paradigm for converting images into Gaussian attribute images. MVSplat [47] proposes representing the cost volume using plane sweeps in 3D space and predicting the depths in sparse reference inputs, precisely locating the centers of Gaussian primitives. While demonstrating generalization ability, this technique's application is limited by its restricted synthesis range and the emergence of distractor-data. Subsequent works, FreeSplat [48] and DGGS [49], address these limitations through Pixel-wise Triplet Fusion strategy and distractor-free training and inference paradigms, respectively. Most recently, G3R [50] extends generalizable 3DGS to dynamic scenes, achieving generalized dynamic scene reconstruction through extra LiDAR data.

3DGS 的显式表示催生了大量工作，专注于利用参考图像逐像素直接推断对应的高斯基元 (Gaussian primitives)，随后用于从目标视角渲染图像 [45], [46]。为实现这一点，早期工作如 Splatter Image[45] 提出了一种将图像转换为高斯属性图的新范式。MVSplat[47] 提出使用三维空间中的平面扫描表示代价体积，并在稀疏参考输入中预测深度，精确定位高斯基元的中心。尽管展示了泛化能力，该技术的应用受限于合成范围有限及干扰数据的出现。后续工作 FreeSplat[48] 和 DGGS[49] 分别通过像素级三元融合策略和无干扰训练及推理范式解决了这些限制。最近，G3R[50] 将可泛化 3DGS 扩展到动态场景，通过额外的激光雷达 (LiDAR) 数据实现了动态场景的泛化重建。

Furthermore, some studies [51], [52] focus on introducing triplane to achieve generalization capabilities, which infers Gaussian attributes by querying triplane features. Recent works [53], [54] seek to extend similar paradigms to large-scale 3D datasets, utilizing transformer-based architectures to enable direct 3D asset inference from sparse image inputs.

此外，一些研究 [51], [52] 聚焦于引入三平面 (triplane) 以实现泛化能力，通过查询三平面特征推断高斯属性。近期工作 [53], [54] 尝试将类似范式扩展到大规模三维数据集，利用基于变换器 (transformer) 的架构实现从稀疏图像输入直接推断三维资产。

D. Sparse Views Setting

D. 稀疏视角设置

Reconstructing from sparse inputs presents significant challenges, wherein the methodology of 3DGS is fundamentally analogous to that of NeRF [55], which aim to develop novel regularization strategies and integrate supplementary information, such as depth data or Diffusion model [56]-[63].

从稀疏输入重建面临重大挑战，其中 3DGS 的方法论在本质上与 NeRF[55] 类似，旨在开发新颖的正则化策略并整合补充信息，如深度数据或扩散模型 (Diffusion model)[56]-[63]。

Specifically, the incorporation of depth data is crucial for 3DGS as an explicit representation, alleviating the model's demand for spatial comprehension under sparse inputs. And multi-view trained diffusion models also can provide important prior knowledge for expanding sparse views into refined dense information. Nevertheless, these methods typically depend on the diffusion models' ability to preserve 3D consistency.

具体而言，深度数据的引入对作为显式表示的 3DGS 至关重要，缓解了模型在稀疏输入下对空间理解的需求。多视角训练的扩散模型也能为将稀疏视角扩展为精细密集信息提供重要先验知识。然而，这些方法通常依赖于扩散模型保持三维一致性的能力。

TABLE II: Comparison to works for Body Reconstruction.

表 II: 人体重建相关工作的比较。

Pose-dependent Deformation	Novel Pose Animation	Fast Training	Rendering > 60FPS	Monocular Input	Super-resolution	Generalization	
✓	✓	✗	✗	✓	✗	✗	HuGS [65]
-	-	-	-	-	✓	✓	GPS-Gaussian [66]
✗	✓	✓	✓	✗	✗	✗	HUGS [67]
✓	✓	✓	✓	✗	✗	✗	GaussianAvatar [68]
✗	✗	✓	✓	✓	✗	✗	GauHuman [69]
✓	✓	✓	✗	✓	✗	✗	3DGS-Avatar [70]
✓	✓	-	✗	✓	✗	✗	ASH [71]
✓	✓	✗	✗	✗	✗	✗	Animatable Guassian [72]

Additional regularization is commonly achieved through incorporating explicit spatial distribution rules or pseudo-views supervision. In addition to these common methods, some studies have focused on the initialization and optimization strategy. GaussianObject [64] introduces an initialization strategy based on Visual Hull and an optimization method using distance statistical data to eliminate floaters.

额外的正则化通常通过引入显式空间分布规则或伪视角监督实现。除这些常见方法外，一些研究关注初始化和优化策略。GaussianObject[64] 引入了基于视觉包络 (Visual Hull) 的初始化策略和利用距离统计数据消除漂浮物的优化方法。

IV. APPLICATIONS OF 3D GAUSSIAN SPLATTING

IV. 三维高斯散射的应用

A. Human Reconstruction

A. 人体重建

1) Body Reconstruction: Body reconstruction mainly focuses on reconstructing deformable human avatars from multiview or monocular videos [73], as well as providing real-time rendering. We list comparisons of recent works in Tab. III

> 1) 身体重建: 身体重建主要关注从多视角或单目视频重建可变形人体模型 [73]，并提供实时渲染。我们在表 III 中列出了近期工作的比较。

Most works [65], [67]-[70] prefer to use well-preconstructed human models like SMPL [74] or SMPL-X [75] as strong prior knowledge. Nevertheless, SMPL is limited to introducing prior knowledge about the human body itself, thus posing challenges for the reconstruction and deformation of outward features such as garments and hair. For the reconstruction of outward appearance, HUGS [67] utilizes SMPL and LBS only at the initial stage, allowing Gaussian primitives to deviate from the initial mesh to accurately represent garments and hair. Additionally, Animatable Gaussians [72] uses a template that can fit outward appearance as guidance and leverages StyleGAN to learn the pose-dependent Gaussian maps, enhancing the capability for modeling detailed dynamic appearances.

> 大多数工作 [65]，[67]-[70] 倾向于使用预先构建良好的人体模型如 SMPL[74] 或 SMPL-X[75] 作为强先验知识。然而，SMPL 仅限于引入人体本身的先验知识，因此在服装和头发等外部特征的重建与变形方面存在挑战。针对外观重建，HUGS[67] 仅在初始阶段使用 SMPL 和线性混合骨骼 (LBS)，允许高斯基元偏离初始网格以准确表示服装和头发。此外，Animatable Gaussians[72] 使用可拟合外观的模板作为指导，并利用 StyleGAN 学习与姿态相关的高斯图，增强了对细节动态外观的建模能力。

Some studies project the problem space from 3D to 2D, thereby reducing complexity and introducing well-established 2D networks for parameter learning [71], [72]. Among them, ASH [71] generates a motion-related template mesh via a deformation network and predicts Gaussian parameters through a 2D network using a motion-related texture map derived from this mesh. These methods are often limited by the difficulty of isolating clothing from the reconstructed 3DGS.

> 一些研究将问题空间从三维投影到二维，从而降低复杂度并引入成熟的二维网络进行参数学习 [71]，[72]。其中，ASH[71] 通过变形网络生成与运动相关的模板网格，并利用该网格生成的运动相关纹理图通过二维网络预测高斯参数。这些方法通常受限于难以将服装从重建的 3DGS 中分离。

Although existing methods achieve high-fidelity reconstruction, their applications are limited by scene-specific training requirements, prompting research efforts toward generalizable human reconstruction [66].

> 尽管现有方法实现了高保真重建，但其应用受限于场景特定的训练需求，促使研究向可泛化人体重建方向发展 [66]。

2) Head Reconstruction: In the domain of human head reconstruction, the works [76]-[80], following the prevalent approach of utilizing SMPL as a strong prior, incorporates FLAME [81] meshes to provide geometric guidance or coarse reconstruction for 3DGS, achieving superior rendering quality. However, Gaussian Head Avatar [82] challenges the conventional use of FLAME meshes and LBS for facial deformation, arguing that these simplified linear operations are inadequate for capturing intricate facial expressions. As an alternative, it introduces an MLP-based approach that directly predicts Gaussian displacements during the transition from neutral to target expressions, enabling high-resolution head rendering at up to 2 K resolution. And then, the works [83],[84] re-

place FLAME-based initialization with geometric guidance derived from signed distance field and DMTet. Several works [85], [86] focus on rendering efficiency, achieving frame rates exceeding 300 FPS. For downstream tasks, various studies [87]-[89] seek to integrate audio for controlling dynamic head reconstruction, thereby achieving audio-visual synchronization.

2) 头部重建: 在人头重建领域，文献 [76]-[80] 遵循利用 SMPL 作为强先验的主流方法，结合 FLAME[81] 网格为 3DGS 提供几何引导或粗略重建，实现了优异的渲染质量。然而，Gaussian Head Avatar[82] 挑战了传统使用 FLAME 网格和线性混合骨骼 (LBS) 进行面部变形的做法，认为这些简化的线性操作不足以捕捉复杂的面部表情。作为替代，它引入了一种基于多层感知机 (MLP) 的方法，直接预测从中性表情到目标表情的高斯位移，实现了最高 2 K 分辨率的高精度头部渲染。随后，文献 [83],[84] 用基于符号距离场和 DMTet 的几何引导替代了基于 FLAME 的初始化。若干研究 [85],[86] 聚焦于渲染效率，实现了超过 300 帧每秒的帧率。针对下游任务，多个研究 [87]-[89] 尝试融合音频以控制动态头部重建，从而实现视听同步。

(a) Text to 3D Objects [95].

(b) 4D Generation [96].

(c) Image to 3D Object [97].

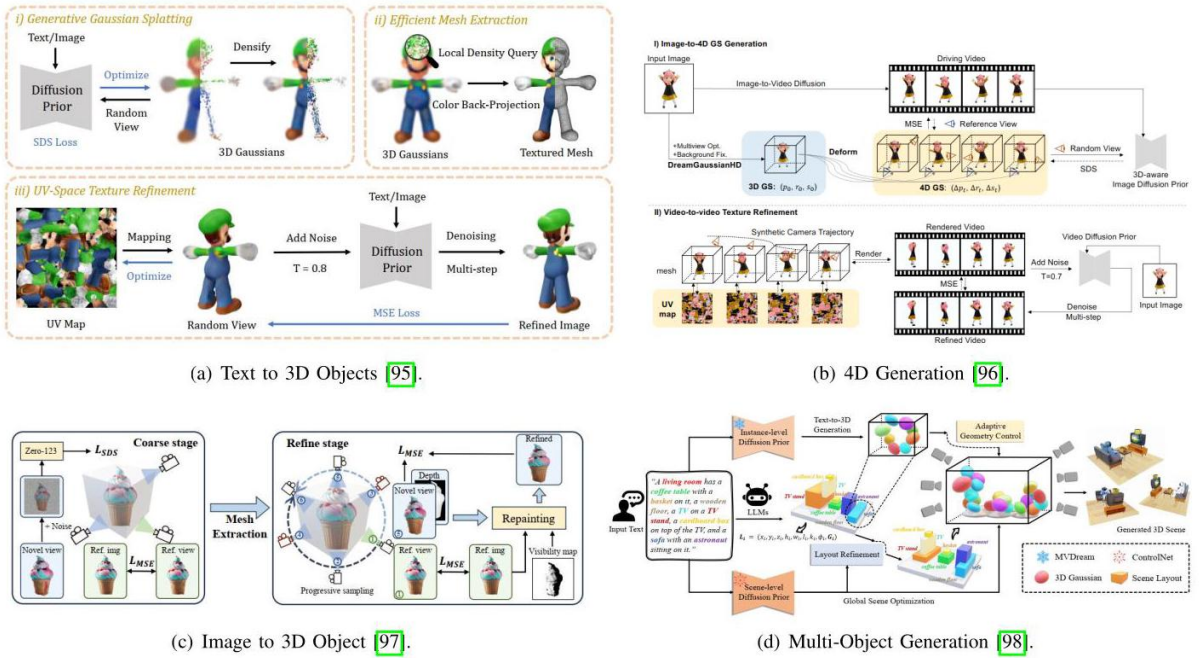(d) Multi-Object Generation [98].

Fig. 3: Four typical tasks of 3DGC in AIGC Applications.

图 3:AIGC 应用中 3DGC 的四个典型任务。

3) Others: 3DGS has introduced innovative solutions in other human-related areas [90], [91]. GaussianHair [90] focuses on the reconstruction of human hair, using linked cylindrical Gaussian modeling for the strands. Additionally, some research [92]-[94] have explored the integration of 3DGS with generative models.

3) 其他:3DGS 在其他与人体相关领域引入了创新解决方案 [90],[91]。GaussianHair[90] 专注于人类头发的重建，采用链式圆柱高斯模型对发丝进行建模。此外，一些研究 [92]-[94] 探索了 3DGS 与生成模型的融合。

B. Artificial Intelligence-Generated Content (AIGC)

B. 人工智能生成内容 (AIGC)

Artificial Intelligence Generated Content (AIGC) leverages artificial intelligence technologies to autonomously produce content. In this sector, we systematically classify contemporary algorithms based on the types of prompts and the objects they generate. The categories include Image-to-3D Object Generation, Text-to-3D Object Generation, Multi-Object and Scene Generation, and 4D Generation.

人工智能生成内容 (AIGC) 利用人工智能技术自主生成内容。在该领域，我们基于提示类型及生成对象系统地分类了当代算法，类别包括图像到 3D 对象生成、文本到 3D 对象生成、多对象及场景生成，以及 4D 生成。

1) Text to 3D Objects: Extensive existing research endeavors to utilize the superior generative capabilities of 2D generative models to achieve coherent 3D content creation. Benefiting from reduced dependency on extensive 3D training data, Score Distillation Sampling-based paradigms garnered significant attention in early research [99]. Some works [95], [100]-[102] focus on improving the framework to apply score distillation loss to 3DGS. Building upon Score Distillation Sampling (SDS), DreamGaussian [95] ensures the geometric consistency of the generated models by extracting explicit Mesh representations from the 3DGS and refines texture in the UV space to enhance the quality of the renderings. However, the mode-seeking paradigm of score distillation frequently leads to oversaturation, excessive smoothing, and lack-detail in the generated outcomes [103]-[107]. Among them, LucidDreamer [104] addresses the challenges of over-smoothing and insufficient sampling steps inherent in traditional SDS. By introducing deterministic diffusion trajectories (DDIM [108]) and interval-based score matching mechanisms Eq. 9, it achieves superior generation quality and efficiency.

1) 文本到 3D 对象: 大量现有研究致力于利用 2D 生成模型的卓越生成能力，实现连贯的 3D 内容创作。得益于对大规模 3D 训练数据依赖的降低，基于评分蒸馏采样 (Score Distillation Sampling, SDS) 的范式在早期研究中受到广泛关注 [99]。部分工作 [95],[100]-[102] 聚焦于改进框架，将评分蒸馏损失应用于 3DGS。在 SDS 基础上，DreamGaussian[95] 通过从 3DGS 中提取显式网格表示确保生成模型的几何一致性，并在 UV 空间细化纹理以提升渲染质量。然而，评分蒸馏的模式寻求范式常导致生成结果过度饱和、过度平滑及细节缺失 [103]-[107]。其中，LucidDreamer[104] 针对传统 SDS 中存在的过度平滑和采样步数不足问题，通过引入确定性扩散轨迹 (DDIM[108]) 和基于区间的评分匹配机制 (公式 9)，实现了更优的生成质量和效率。

To further mitigate the inherent limitations of SDS, several works [109]-[113] seek to leverage video or multi-view generative models to obtain more data for reconstruction. Although these approaches introduce direct prior guidance for 3D generation, the inherent lack of guaranteed 3D consistency in both multi-view and video generation still leads to instability in reconstruction. To enhance the efficiency of 3D asset generation, works [114]-[116] aim to generate using only feed-forward networks without the need for scene-specific training. BrightDreamer [114] predicts positional offsets following fixed initialization and employs a text-guided triplane generator to process extracted textual features for predicting additional 3DGS attributes, achieving text-to-3D model conversion in 77 ms , albeit at the cost of some reconstruction quality. For enhanced geometric detail control, SketchDream [117] proposes a framework for sketch-based text-to-3D generation and editing, which integrates hand-drawn sketches and text prompts to achieve fine-grained geometry and appearance control, enabling high-quality 3D content creation and local editing through a two-stage coarse-to-fine approach.

为进一步缓解 SDS 固有的局限性，若干研究 [109]-[113] 尝试利用视频或多视角生成模型获取更多重建数据。尽管这些方法为 3D 生成引入了直接的先验引导，但多视角和视频生成本质上缺乏 3D 一致性保障，仍导致重建不稳定。为提升 3D 资产生成效率，文献 [114]-[116] 旨在仅通过前馈网络生成，无需场景特定训练。BrightDreamer[114] 在固定初始化后预测位置偏移，采用文本引导的三平面生成器处理提取的文本特征以预测额外的 3DGS 属性，实现了文本到 3D 模型的转换，尽管牺牲了一定的重建质量。为增强几何细节控制，SketchDream[117] 提出了基于草图的文本到 3D 生成与编辑框架，融合手绘草图与文本提示，实现细粒度的几何与外观控制，通过粗到细的两阶段方法支持高质量 3D 内容创作及局部编辑。

More directly, several works [118]-[121] aim to incorporate 3DGS representations into 3D generative models. Among these, L3DG [120] proposes a latent diffusion framework for compressed 3D Gaussian representation, achieving superior visual quality and real-time rendering. However, these approaches are inherently limited by the availability of 3D data, which impacts their reconstruction capability for complex targets.

更直接地，若干工作 [118]-[121] 致力于将 3DGS 表示融入 3D 生成模型。其中，L3DG[120] 提出了压缩 3D 高斯表示的潜在扩散框架，实现了卓越的视觉质量和实时渲染。然而，这些方法本质上受限于 3D 数据的可用性，影响其对复杂目标的重建能力。

Some works [92], [93], [122] also attempt to apply this generative paradigm to areas such as digital human generation. HumanGaussian [93] combines RGB and depth rendering to improve the SDS, thereby jointly supervising the optimization of the structural perception of human appearance and geometry.

一些工作 [92], [93], [122] 也尝试将这一生成范式应用于数字人类生成等领域。HumanGaussian[93] 结合了 RGB 和深度渲染以提升 SDS，从而共同监督人体外观和几何结构感知的优化。

2) Image to 3D Object: Similar to works on NeRF, recent studies [97], [123], [124] have also focused on generating entire 3DGS from a single image. These approaches share fundamental similarities with text-to-3D object methods. As an example, following a process similar to DreamGaussian [95], Repaint123 [97] employs Zero-123 [125] and SDS for coarse 3DGS, followed by a fine stage where mesh representation is extracted and refined using depth-guided and visibility-aware repainting on novel views for consistent 3DGS fine-tuning.

2) 图像到 3D 对象: 类似于 NeRF 相关工作，近期研究 [97], [123], [124] 也聚焦于从单张图像生成完整的 3DGS。这些方法与文本到 3D 对象的方法在基本原理上相似。例如，继承 DreamGaussian[95] 的流程，Repaint123[97] 采用 Zero-123[125] 和 SDS 进行粗略 3DGS 生成，随后通过深度引导和可见性感知的重绘，在新视角上提取并细化网格表示，实现一致的 3DGS 微调。

3) Multi-Object and Scene Generation: In addition to single-object generation, multi-object and scene generation is more crucial in most application scenarios.

3) 多对象与场景生成: 除了单对象生成，多对象及场景生成在大多数应用场景中更为关键。

Multi-Object Generation: Several studies [98], [126]-[128] have explored the generation of multiple composite objects, which not only concentrate on the individual objects but also aim to investigate the interactions between multiple objects. For predicting the interactions between multiple objects, CG3D [126] leverages SDS and probabilistic graph models extracted from text to predict the relative relationships between objects and incorporates

priors such as gravity and contact relationships between objects, CG3D achieves models with realistic physical interactions.

> 多对象生成: 若干研究 [98]，[126]-[128] 探讨了多复合对象的生成，不仅关注单个对象，还致力于研究多对象间的交互。为预测多对象间的交互，CG3D[126] 利用 SDS 和从文本中提取的概率图模型预测对象间的相对关系，并结合重力和接触关系等先验，实现具有真实物理交互的模型。

Scene Generation: Unlike object-centric generation, scene generation typically requires the incorporation of additional information, such as pre-trained depth estimation models [129], [130] or Large Language Models [127], [131], where the former provides spatial understanding for projecting images into 3D space, while the latter enhances the quality of text prompts. LucidDreamer [2] [129] employs a two-stage approach: first initializing point clouds using text-to-image and depth estimation models with inpainting [132] for consistency, then converting to 3DGS with extended image supervision.

> 场景生成: 与以对象为中心的生成不同，场景生成通常需要整合额外信息，如预训练的深度估计模型 [129]，[130] 或大型语言模型 (Large Language Models)[127]，[131]，前者提供将图像投射到 3D 空间的空间理解，后者提升文本提示的质量。LucidDreamer [2] [129] 采用两阶段方法: 首先利用文本到图像和深度估计模型结合修复 (inpainting)[132] 初始化点云以保证一致性，然后通过扩展的图像监督转换为 3DGS。

4) 4D Generation: Analogous to static scene generation using text-to-image SDS, it is natural to consider that text-to-video SDS could potentially generate dynamic scenes [96], [133]-[137]. These works primarily focus on designing video-based SDS losses or exploring hybrid supervision with T2I (Text to Image) and T2V (Text to Video) models. As an example, Align Your Gaussians [133] adopts a two-stage approach: first reconstructing static 3DGS using MVDream [138] and text-to-image supervision, then extending to 4DGS with text-to-video guidance and simplified score distillation loss. Although these methods are effective, the inherent limitations of the aforementioned SDS-based paradigm persist.

> 4)4D 生成: 类似于使用文本到图像 SDS 进行静态场景生成，自然考虑文本到视频 SDS 可能生成动态场景 [96]，[133]-[137]。这些工作主要关注设计基于视频的 SDS 损失或探索 T2I(文本到图像) 与 T2V(文本到视频) 模型的混合监督。例如，Align Your Gaussians[133] 采用两阶段方法: 先利用 MVDream[138] 和文本到图像监督重建静态 3DGS，再通过文本到视频引导和简化的分数蒸馏损失扩展到 4DGS。尽管这些方法有效，但上述基于 SDS 范式的固有限制依然存在。

To mitigate this issue, subsequent works focus on generating pseudo-labeled images from additional views to facilitate dynamic 3DGS reconstruction [139]-[142]. Among them, 4DGen [139] generates multi-view pseudo-labels per frame, while employing Hexplane's [143] multi-scale features to maintain temporal consistency in 4DGS generation.

> 为缓解该问题，后续工作聚焦于从额外视角生成伪标签图像以促进动态 3DGS 重建 [139]-[142]。其中，4DGen[139] 为每帧生成多视角伪标签，同时利用 Hexplane[143] 的多尺度特征保持 4DGS 生成的时间一致性。

Furthermore, some studies focus on animating static canonical 3DGS [144], [145] to achieve better control. Among them, BAGS [144] introduces neural bones and skinning weights to describe the spatial deformation based on canonical space. Using diffusion model priors and rigid body constraints, BAGS can be manually manipulated

to achieve novel pose rendering.

此外, 一些研究致力于对静态规范 3DGS[144], [145] 进行动画化以实现更好控制。其中, BAGS[144] 引入神经骨骼和蒙皮权重描述基于规范空间的空间变形。借助扩散模型先验和刚体约束, BAGS 可通过手动操作实现新姿态渲染。

C. Autonomous Driving

C. 自动驾驶

1) Autonomous Driving Scene Reconstruction: Reconstruction driving scenes is a challenging task, involving multiple technical domains such as large-scale scene reconstruction, dynamic object reconstruction, static object reconstruction, and Gaussian mixture reconstruction. Existing works [146]- [148] partitions the entire process into static background and dynamic target reconstruction. DrivingGaussian [146] reconstructs large-scale driving scenes by combining depth-binned static 3DGS for background and dynamic Gaussian graphs for multiple targets, utilizing multi-sensor data. Then StreetGaussians [147] extends this approach by incorporating semantic attributes and employing Fourier transforms for efficient SH temporal modeling in dynamic 3DGS. Subsequent works [149], aiming to further improve reconstruction efficiency, attempt to directly reconstruct entire scenes by Tightly Coupled LiDAR-Camera Gaussian Splatting.

1) 自动驾驶场景重建: 重建驾驶场景是一项挑战性任务, 涉及大规模场景重建、动态对象重建、静态对象重建及高斯混合重建等多个技术领域。现有工作 [146]-[148] 将整个过程划分为静态背景和动态目标重建。DrivingGaussian[146] 结合深度分箱的静态 3DGS 作为背景和多目标动态高斯图, 利用多传感器数据重建大规模驾驶场景。随后, StreetGaussians[147] 通过引入语义属性并采用傅里叶变换实现动态 3DGS 中高效的球谐 (SH) 时域建模, 扩展了该方法。后续工作 [149] 为进一步提升重建效率, 尝试通过紧耦合激光雷达-摄像头高斯点云渲染直接重建整个场景。

Moreover, 3DGS have been applied to multimodal spatiotemporal calibration tasks [150]. By leveraging the LiDAR point cloud as a reference for the Gaussians' positions, 3DGS-Calib [150] constructs a continuous scene representation and enforces both geometrical and photometric consistency across all sensors, achieving accurate and robust calibration with improved efficiency compared to NeRF-based works.

此外, 3DGS 已应用于多模态时空校准任务 [150]。3DGS-Calib[150] 利用激光雷达点云作为高斯位置参考, 构建连续场景表示, 并在所有传感器间强制几何和光度一致性, 实现比基于 NeRF 的工作更准确、鲁棒且高效的校准。

2) Simultaneous Localization and Mapping (SLAM): SLAM is a key problem in robotics and computer vision, where a device builds a map of an unknown environment while locating itself within it. Some studies [151]-[156] have retained the traditional inputs and approached this from two perspectives: online tracking and incremental mapping. In early work, GS-SLAM [151] utilizes 3DGS for SLAM with adaptive primitive expansion and employs a coarse-to-fine optimization strategy: first optimizing camera poses using sparse pixels, then refining them through selective re-rendering of reliable Gaussians. In parallel, Photo-SLAM [152] combines ORB features [157] and Gaussian attributes in a Hyper Primitives Map, utilizing LM optimization [158] and loop closure [157] for photorealistic SLAM reconstruction. While these approaches achieve higher efficiency than NeRF-based methods, further optimization of computational performance is essential for real-world deployment. Therefore, CG-SLAM

[155] leverages an uncertainty-aware 3D Gaussian field and a GPU-accelerated framework, achieving superior efficiency. Then, RGBD GS-ICP SLAM [156] enhances efficiency by integrating G-ICP [159] with shared co-variances and scale alignment techniques for faster convergence. However, these methods remain susceptible to sensor noise in practical applications. Some quantitative reconstruction results are reported in Table III.

2) 同时定位与地图构建 (SLAM):SLAM 是机器人学和计算机视觉中的一个关键问题，设备在构建未知环境地图的同时定位自身。一些研究 [151]-[156] 保留了传统输入，从在线跟踪和增量建图两个角度进行探讨。在早期工作中，GS-SLAM[151] 利用 3D 高斯 (3DGS) 进行 SLAM，采用自适应基元扩展和由粗到细的优化策略: 先用稀疏像素优化相机位姿，再通过选择性重渲染可靠高斯进行细化。与此同时，Photo-SLAM[152] 结合 ORB 特征 [157] 和高斯属性构建超基元地图，利用 LM 优化 [158] 和回环检测 [157] 实现光度真实感 SLAM 重建。尽管这些方法效率高于基于 NeRF 的方法，但为了实际应用，计算性能的进一步优化仍然必要。因此，CG-SLAM[155] 利用不确定性感知的 3D 高斯场和 GPU 加速框架，实现了更优的效率。随后，RGBD GS-ICP SLAM[156] 通过结合共享协方差和尺度对齐技术的 G-ICP[159] 提升效率，加快收敛速度。然而，这些方法在实际应用中仍易受传感器噪声影响。部分定量重建结果见表 III。

TABLE III: Comparison of 3DGS-SLAM on Replica [165

表 III:Replica 数据集上 3DGS-SLAM 的比较 [165]

Method		SSIM↑	LPIPS↓	Tracking RMSE↓
GS-SLAM [166]	34.27	0.98	0.08	0.50
Photo-SLAM 152	34.96	0.94	0.06	0.60
SplaTAM [153]	34.11	0.97	0.10	0.36
GS-ICP SLAM [156]	38.83	0.98	0.04	0.16
MotionGS [167	39.60	0.98	0.04	0.49
LoopSplat [168]	36.63	0.99	0.11	0.26

Method		结构相似性指数 (SSIM)↑	感知图像补丁相似性 (LPIPS)↓	跟踪均方根误差 (RMSE)↓
GS-SLAM [166]	34.27	0.98	0.08	0.50
Photo-SLAM 152	34.96	0.94	0.06	0.60
SplaTAM [153]	34.11	0.97	0.10	0.36
GS-ICP SLAM [156]	38.83	0.98	0.04	0.16
MotionGS [167	39.60	0.98	0.04	0.49
LoopSplat [168]	36.63	0.99	0.11	0.26

Incorporating scene understanding capabilities is equally crucial in SLAM tasks [160]-[162], thus prompting several works to integrate semantic information. Among them, SGS-SLAM [160] employs multi-channel geometric, appearance, and semantic features for rendering and optimization and proposes a keyframe selection strategy based on geometric and semantic constraints to enhance performance and efficiency.

在 SLAM(同步定位与地图构建) 任务中，融入场景理解能力同样至关重要 [160]-[162]，因此促使多项工作将语义信息整合其中。其中，SGS-SLAM[160] 采用多通道的几何、外观和语义特征进行渲染与优化，并提出基于几何和语义约束的关键帧选择策略，以提升性能和效率。

Additionally, there are several works focusing on related issues such as localization [163] and navigation [164]. Specifically, 3DGS-ReLoc [163] leverages LiDAR initialization and 2D voxelized submaps with KD-tree for

efficient memory usage, while achieving precise localization through feature-based PnP optimization. In the context of indoor navigation, GaussNav [164] focuses on the instance image navigation task. Based on reconstructed 3DGS maps, GaussNav proposes an image target navigation algorithm, achieving impressive performance through classification, matching, and path planning.

此外，还有若干工作关注定位 [163] 和导航 [164] 等相关问题。具体而言，3DGS-ReLoc[163] 利用激光雷达初始化和带有 KD 树的二维体素子图以实现高效内存使用，同时通过基于特征的 PnP 优化实现精确定位。在室内导航方面，GaussNav[164] 聚焦于实例图像导航任务。基于重建的 3DGS 地图，GaussNav 提出了一种图像目标导航算法，通过分类、匹配和路径规划实现了出色的性能。

V. EXTENSIONS OF 3D GAUSSIAN SPLATTING

五、三维高斯点扩展

A. Dynamic 3D Gaussian Splatting

A. 动态三维高斯点扩展

The study of dynamic 3DGS has recently garnered significant attention. Here, we categorize them into Multi-view Videos and Monocular Video based on different inputs.

动态三维高斯点扩展 (3DGS) 的研究近年来受到广泛关注。这里，我们根据不同输入将其分为多视角视频和单目视频两类。

1) Multi-view Videos: Some works [169], [170] attempted to directly construct dynamic 3DGS frame by frame. An early work [169] extends 3DGS to dynamic scenes by enabling temporal Gaussian motion while preserving static attributes. It employs online temporal reconstruction with previous-frame initialization and incorporates physical priors, including local rigidity, local rotational-similarity, and long-term local-isometry, for motion regularization, as shown in Eq. 4.6 Similarly, 3DGStream [170] employs a two-stage training with Neural Transformation Cache and I-NGP [171] for dynamic reconstruction, followed by gradient-based adaptive Gaussian densification. Despite promising results, these methods are limited to reconstructing only scene elements visible in the initial frame. To address this limitation, other works [172], [173] aim to achieve such performance by predicting deformations. SWAGS [172] proposes a window-based 4DGS with flow-guided adaptive window division and dynamic MLP optimization, employing inter-window consistency loss for seamless temporal reconstruction.

1) 多视角视频: 部分工作 [169], [170] 尝试逐帧直接构建动态 3DGS。早期工作 [169] 通过引入时间高斯运动同时保持静态属性，将 3DGS 扩展至动态场景。其采用基于前一帧初始化的在线时间重建，并结合物理先验，包括局部刚性、局部旋转相似性和长期局部等距性，用于运动正则化，如公式 4.6 所示。类似地，3DGStream[170] 采用两阶段训练，结合神经变换缓存和 I-NGP[171] 进行动态重建，随后进行基于梯度的自适应高斯密化。尽管结果令人鼓舞，这些方法仅限于重建初始帧中可见的场景元素。为解决此限制，其他工作 [172], [173] 通过预测形变以实现此类性能。SWAGS[172] 提出基于窗口的四维高斯点扩展，采用流引导的自适应窗口划分和动态多层感知机优化，并利用窗口间一致性损失实现无缝的时间重建。

TABLE IV: Comparison of Dynamic 3DGS on D-NeRF [186]

Method		SSIM↑	LPIPS↓		FPS↑
K-Planes [NeRF-based]	31.07	0.97	0.02	54 min	1.20
Deformable3DGS [174]	39.31	0.99	0.01	26 min	85.45
CoGS [187]	37.90	0.983	0.027	-	-
SC-GS [184]	43.31	0.997	0.0063	-	-
GauFRe [175]	34.5	0.98	0.02	13mins	112
Deformable4DGS | 176	32.99	0.97	0.05	13 min	104.00
RealTime4DGS [188]	32.71	0.97	0.03	10 min	289.07
4DRotorGS [189	34.26	0.97	0.03	5 min	1257.63

Method		结构相似性指数 (SSIM)↑	感知图像补丁相似性 (LPIPS)↓		帧率 (FPS)↑
K-Planes [基于 NeRF(神经辐射场)]	31.07	0.97	0.02	54 分钟	1.20
Deformable3DGS [174]	39.31	0.99	0.01	26 分钟	85.45
CoGS [187]	37.90	0.983	0.027	-	-
SC-GS [184]	43.31	0.997	0.0063	-	-
GauFRe [175]	34.5	0.98	0.02	13 分钟	112
Deformable4DGS | 176	32.99	0.97	0.05	13 分钟	104.00
RealTime4DGS [188]	32.71	0.97	0.03	10 分钟	289.07
4DRotorGS [189	34.26	0.97	0.03	5 分钟	1257.63

2) Monocular Video or Multi-view Images: Some works [174]-[180] tend to divide into two stages: canonical reconstruction and deformation prediction. The studies [174], [175] reconstruct static 3DGS and predicts temporal deformation in terms of positions, rotations and scales through position-time encoding. Similarly, 4D-GS [176] introduces the multi-scale HexPlane [143] as the foundational representation to encode temporal and spatial information. To further decouple motion and shape parameters in 4D-GS, ST-4DGS [179] introduces a spatial-temporally consistent 4DGS framework that incorporates motion-aware shape regularization and spatial-temporal density control to learn compact 4D representations. Although these methods achieve stable performance, they struggle to handle abrupt motions and sudden object appearances.

2) 单目视频或多视角图像: 一些工作 [174]-[180] 倾向于分为两个阶段: 规范重建和变形预测。研究 [174]、[175] 重建静态三维高斯表面 (3DGS),并通过位置-时间编码预测位置、旋转和尺度的时间变形。类似地, 4D-GS[176] 引入多尺度 HexPlane[143] 作为基础表示, 用于编码时空信息。为了进一步解耦 4D-GS 中的运动和形状参数, ST-4DGS[179] 提出了一个时空一致的 4DGS 框架,结合运动感知的形状正则化和时空密度控制,以学习紧凑的 4D 表示。尽管这些方法实现了稳定的性能,但在处理突发运动和突然出现的物体时仍存在困难。

Instead of discrete offsets, exploring temporally continuous motion can promote smoothness in the time dimension [181]- [184]. Gaussian-Flow [183] aims to develop a representation capable of fitting variable motion by analyzing the advantages and disadvantages of polynomial [182], [185] and Fourier series fitting [181]. It then proposes a dual-domain deformation model with adaptive time-step scaling and temporal-rigid constraints for stable and continuous motion prediction.

与离散偏移不同，探索时间连续的运动有助于提升时间维度的平滑性 [181]-[184]。Gaussian-Flow[183] 旨在通过分析多项式 [182]、[185] 和傅里叶级数拟合 [181] 的优缺点，开发一种能够拟合可变运动的表示。随后提出了一个具有自适应时间步长缩放和时间刚性约束的双域变形模型，以实现稳定且连续的运动预测。

Recent works aim to extend 3DGS to 4D space for dynamic 3D scenes representation. Among them, the work [188] achieves end-to-end 4D training by jointly modeling spatial-temporal variables with 4D Gaussian primitives, incorporating 4D rotation, scaling, and temporal-aware spherical harmonics for color variation. Similarly, the work [189] introduces a rotor-based 4DGS representation with eight-component rotation decomposition, enabling temporal slicing for dynamic objects and enforcing 4D consistency through a dedicated loss. Although these approaches demonstrate robustness in complex scene reconstruction, the compressibility of their representations remains a notable consideration. Quantitative reconstruction results are reported in Table IV.

近期工作旨在将 3DGS 扩展到 4D 空间以表示动态三维场景。其中，研究 [188] 通过联合建模时空变量与 4D 高斯基元，实现了端到端的 4D 训练，结合 4D 旋转、缩放及时间感知的球谐函数以表现颜色变化。类似地，研究 [189] 引入基于旋转子的 4DGS 表示，采用八分量旋转分解，实现动态物体的时间切片，并通过专门的损失函数强制 4D 一致性。尽管这些方法在复杂场景重建中表现出鲁棒性，但其表示的可压缩性仍是一个重要考量。定量重建结果见表 IV。

Dynamic 3DGS reconstruction demonstrate broad applicability across diverse domains, extending beyond previously mentioned human reconstruction IV-A, AIGC IV-B, and autonomous driving IV-C to include robotic manipulation [190], physical simulation [191], and various downstream tasks.

动态 3DGS 重建在多个领域展现出广泛适用性，除了前述的人体重建 IV-A、人工智能生成内容 (AIGC)IV-B 和自动驾驶 IV-C 外，还涵盖机器人操作 [190]、物理仿真 [191] 及各种下游任务。

B. Surface Representation

B. 表面表示

Although 3DGS enables highly realistic rendering, extracting surface representations remains challenging. In this line of works, Signed Distance Functions are an indispensable topic [192]-[195].In early work, SuGaR [192] proposes an idealized SDF to constrain Gaussian surface alignment, enabling efficient mesh extraction through Poisson reconstruction and optional mesh-guided Gaussian refinement for high-quality results. Similarly, 3DGSR [194] integrates neural implicit SDF with 3DGS through a differentiable SDF-to-opacity transformation, maintaining consistency between volumetric and 3DGS-derived depth properties. Another line of research [193], [195] focuses on jointly optimizing NeuS [196] and 3DGS for surfaces. However, these methods exhibit limitations in handling unbounded scenes and computational overhead.

尽管 3DGS 实现了高度逼真的渲染，提取表面表示仍然具有挑战性。在这类工作中，符号距离函数 (Signed Distance Functions，SDF) 是不可或缺的主题 [192]-[195]。早期工作 SuGaR[192] 提出理想化的 SDF 以约束高斯表面对齐，通过泊松重建实现高效网格提取，并可选用网格引导的高斯细化以获得高质量结果。类似地，3DGSR[194] 通过可微分的 SDF 到不透明度转换，将神经隐式 SDF 与 3DGS 结合，保持体积和 3DGS 导出深度属性的一致性。另一类研究 [193]、[195] 聚焦于联合优化 NeuS[196] 与 3DGS 以实现表面表示，但这些方法在处理无界场景和计算开销方面存在局限。

Other studies [197]-[200] aim to address this issue by improving 3DGS representation. The work [197] proposes Gaussian Surfels with depth-normal consistency loss and volumetric cutting for improved surface reconstruction, followed by screened Poisson mesh generation. Similarly, 2D Gaussian Splatting [199] (2DGS) replaces 3DGS with planar disks to represent surfaces, which are defined within the local tangent plane. Then, Gaussian Opacity Fields (GOF) [198] are developed based on 3DGS, wherein 3DGS is normalized along the ray to form a 1DGS for volume rendering. Although these methods achieve precise geometry reconstruction, they inevitably compromise rendering fidelity and face challenges in handling semi-transparent surfaces.

其他研究 [197]-[200] 致力于通过改进 3DGS 表示来解决该问题。研究 [197] 提出具有深度-法线一致性损失和体积切割的高斯点云 (Gaussian Surfels)，以提升表面重建质量，随后进行筛选泊松网格生成。类似地，二维高斯点云 (2D Gaussian Splatting，2DGS)[199] 用局部切平面内定义的平面圆盘替代 3DGS 表示表面。随后，基于 3DGS 开发了高斯不透明度场 (Gaussian Opacity Fields，GOF)[198]，其中 3DGS 沿射线归一化形成一维高斯点云 (1DGS) 用于体积渲染。尽管这些方法实现了精确的几何重建，但不可避免地牺牲了渲染保真度，并在处理半透明表面时面临挑战。

C. Editable 3D Gaussian Splatting

C. 可编辑的三维高斯点云

3DGS, with its advantages of explicit representation, has naturally attracted considerable attention from researchers focusing on 3DGS editing. Unfortunately, current editable 3DGS works often lack precise training supervision, which poses a significant challenge for editing. In this section, we categorize existing works according to different tasks.

3DGS 凭借其显式表示的优势，自然吸引了大量关注 3DGS 编辑的研究者。不幸的是，当前的可编辑 3DGS 工作往往缺乏精确的训练监督，这对编辑构成了重大挑战。本节根据不同任务对现有工作进行分类。

1) Manipulation by Text: To address this challenge, the existing works can be classified into two distinct categories. The first type introduces the score distillation loss as in Eq. 8 These methods require editing prompts as additional conditions to guide the editing process [201], [202]. GaussianEditor [201] enables semantic-controlled 3DGS editing through SDS, utilizing hierarchical 3DGS and anchor loss for stability, while incorporating 2D inpainting guidance for object manipulation. Following Dreamgaussian, GSEdit [202] uses the pre-trained Instruct-Pix2Pix [203] model instead of the image generation model for SDS. However, such methods remain constrained by pre-trained diffusion models, particularly when handling complex editing prompts.

1) 基于文本的操作: 为应对该挑战，现有工作可分为两类。第一类引入了如公式 8 所示的评分蒸馏损失 (score distillation loss)，这些方法需要编辑提示作为额外条件以指导编辑过程 [201]、[202]。GaussianEditor[201] 通过 SDS 实现语义控制的 3DGS 编辑，利用分层 3DGS 和锚点损失保证稳定性，同时结合二维修补引导进行物体操作。继 Dreamgaussian 之后，GSEdit[202] 使用预训练的 Instruct-Pix2Pix[203] 模型替代图像生成模型进行 SDS。然而，此类方法仍受限于预训练扩散模型，尤其在处理复杂编辑提示时表现有限。

The second type focuses on editing multi-view images before reconstructing. GaussianEditor [2] [204] employs multi-modal, language, and segmentation models to locate editable regions from text inputs, then optimizes targeted Gaussians based on 2D-edited images. However, this paradigm introduces an intuitive problem: how to ensure consistency in multi-view editing [205], [206]. GaussCtrl [205] introduces a depth-guided image editing network, ControlNet [207], utilizing its ability to perceive geometry and maintain multi-view consistency in the editing network. It also introduces a latent code alignment strategy in the attention layers, ensuring that the edited multiview images remain consistent with the references.

第二类方法侧重于在重建前编辑多视图图像。GaussianEditor [2] [204] 利用多模态、语言和分割模型从文本输入中定位可编辑区域，然后基于二维编辑图像优化目标高斯。然而，该范式引入了一个直观问题: 如何确保多视图编辑的一致性 [205]，[206]。GaussCtrl [205] 引入了基于深度引导的图像编辑网络 ControlNet [207]，利用其感知几何结构和保持多视图一致性的能力。同时在注意力层中引入潜代码对齐策略，确保编辑后的多视图图像与参考图像保持一致。

Unlike editing methods for 3DGS, recent discussions have increasingly focused on editing 4DGS. Recent work, Control4D [208] leverages 4D GaussianPlanes to structurally decompose four-dimensional space, ensuring spatiotemporal consistency through Tensor4D representation, while incorporating a super-resolution GAN-based 4D generator [209] that learns from diffusion-generated edited images. However, it remains challenged on non-rigid movements.

与 3D 高斯场 (3DGS) 的编辑方法不同，近期讨论越来越多地聚焦于 4DGS 的编辑。最新工作 Control4D [208] 利用 4D 高斯平面结构化分解四维空间，通过 Tensor4D 表示确保时空一致性，同时结合基于超分辨率 GAN 的 4D 生成器 [209]，该生成器从扩散生成的编辑图像中学习。然而，该方法在处理非刚性运动方面仍面临挑战。

2) Manipulation by Other Conditions: In addition to text-controlled editing, existing works have explored 3DGS editing methods under various conditions, such as mixed conditions [210] and points [211]. TIP-Editor [210] enables fine-grained 3DGS control through text, reference image, and location inputs, utilizing stepwise 2D personalization and coarse-fine editing strategies to support diverse tasks like object insertion and stylization. And Point'n Move [211] enables object-level editing through point annotations, utilizing a dual-stage process of segmentation, inpainting, and recomposition, which demonstrates improved control capability. Recent research [212] has introduced a training-free 3DGS splitting paradigm that achieves editing plane control by formulating it as a constrained minimization problem, preserving visual fidelity through moment conservation while mitigating Gaussian overflow via an analytically derived closed-form solution.

2) 基于其他条件的操作: 除了文本控制编辑, 现有工作还探索了在多种条件下的 3DGS 编辑方法, 如混合条件 [210] 和点标注 [211]。TIP-Editor [210] 通过文本、参考图像和位置输入实现细粒度 3DGS 控制, 采用逐步二维个性化和粗细编辑策略, 支持对象插入和风格化等多样任务。Point'n Move [211] 通过点注释实现对象级编辑, 利用分割、修补和重组的双阶段流程, 展现了更强的控制能力。近期研究 [212] 引入了一种无训练的 3DGS 分割范式, 将编辑平面控制形式化为约束最小化问题, 通过矩保持视觉保真, 同时通过解析闭式解缓解高斯溢出。

3) Stylization: In the realm of style transfer for 3DGS, early explorations have been made by [213]. Similar to traditional style transfer works [214], this work designs a 2D stylization module on the rendered images and a 3D color module on the 3DGS. By aligning the stylized 2D results of both modules, this approach achieves multi-view consistent 3DGS stylization without altering the geometry.

3) 风格化: 在 3DGS 风格迁移领域, 早期探索见于 [213]。类似传统风格迁移工作 [214], 该工作设计了基于渲染图像的二维风格化模块和基于 3DGS 的三维颜色模块。通过对齐两个模块的风格化二维结果, 实现了多视图一致的 3DGS 风格化, 且不改变几何结构。

4) Animation: As described in Sec. V-A, some dynamic 3DGS works, such as SC-GS [184], can achieve animation effects by animating sparse control points. AIGC-related works, such as BAGS [144], aim to utilize video input and generation models to animate existing 3DGS. Similar research has also been mentioned in the context of Human Reconstruction. Additionally, CoGS [187] discusses how to control animation. Based on dynamic representations [169], [174], it uses a small MLP to extract relevant control signals and align the deformation of each Gaussian primitives.

4) 动画: 如第五节 A 部分所述, 一些动态 3DGS 工作, 如 SC-GS [184], 通过动画稀疏控制点实现动画效果。与 AIGC 相关的工作, 如 BAGS [144], 旨在利用视频输入和生成模型为现有 3DGS 赋予动画。类似研究也在人体重建背景中提及。此外, CoGS [187] 讨论了动画控制方法。基于动态表示 [169], [174], 其使用小型多层感知机 (MLP) 提取相关控制信号, 并对每个高斯基元的变形进行对齐。

D. Relightable

D. 可重光照

Relightable 3DGS has also emerged as one of the recent challenges gaining significant attention. Decoupling texture and lighting represents a common approach in relighting tasks. In early work, Relightable 3D Gaussian [215] and GS-IR [216] represent scenes using points with normal, BRDF, and decomposed lighting attributes for relighting. However, these methods face challenges in handling reflective scenes. Therefore, followup work [34] introduce accurate normal estimation and residual color terms to effectively model view-dependent reflections and complex lighting interactions. To address limitations in complex materials like semi-transparent volumes and furs, OLAT-GS [217] decomposes observed color into the attenuated light intensity, received incident illumination and scattering value. And then, GS^3 [218] combines spatial and angular Gaussians with a triple splatting process to model geometry and reflectance properties, incorporating neural networks for self-shadowing and global illumination. Despite their capability in handling challenging geometry and appearance, further exploration is needed for transparent materials and indirect lighting.

可重光照 3DGS 作为近期备受关注的挑战之一也逐渐兴起。纹理与光照的解耦是重光照任务中的常见方法。早期工作中，Relightable 3D Gaussian [215] 和 GS-IR [216] 使用带有法线、双向反射分布函数 (BRDF) 和分解光照属性的点表示场景以实现重光照。然而，这些方法在处理反射场景时存在挑战。因此，后续工作 [34] 引入了精确的法线估计和残余颜色项，有效建模视角依赖反射和复杂光照交互。为解决半透明体积和毛发等复杂材质的限制，OLAT-GS [217] 将观测颜色分解为衰减光强、接收的入射光照和散射值。随后，GS^3 [218] 结合空间和角度高斯，通过三重溅射过程建模几何和反射属性，并引入神经网络处理自阴影和全局光照。尽管其在处理复杂几何和外观方面表现出色，但对透明材质和间接光照的探索仍需深入。

E. Semantic Understanding

E. 语义理解

Endowing 3DGS with semantic understanding capabilities allows for the extension of 2D semantic models into 3D space, thereby enhancing the model's comprehension in 3D environments. This can be applied to various tasks such as 3D detection, segmentation, and editing. Many works attempt to leverage pre-trained 2D semantic-aware models for extra supervision on semantic attributes [219]-[222]. Feature 3DGS [220] leverages pre-trained 2D models to create joint 3DGS and feature fields, enabling spatial understanding through feature rasterization and regularization for downstream promptable tasks. However, these approaches remain constrained by multi-view consistency and open-world perception challenges. Subsequent research [223], 224 aims to incorporate contrastive learning losses as auxiliary supervision to achieve interactive 3D segmentation. To mitigate the dependence on 2D pre-trained models in the aforementioned works, the work [225] attempts to leverage reconstructed Featured 3DGS to fine-tune the 2D pre-trained encoder.

赋予 3DGS 语义理解能力，使得 2D 语义模型能够扩展到 3D 空间，从而增强模型在三维环境中的理解力。这可以应用于 3D 检测、分割和编辑等多种任务。许多工作尝试利用预训练的二维语义感知模型对语义属性进行额外监督 [219]-[222]。Feature 3DGS [220] 利用预训练的二维模型创建联合的 3DGS 和特征场，通过特征光栅化和正则化实现空间理解，以支持下游的可提示任务。然而，这些方法仍受限于多视图一致性和开放世界感知的挑战。后续研究 [223]、224 旨在引入对比学习损失作为辅助监督，实现交互式 3D 分割。为减少上述工作对 2D 预训练模型的依赖，工作 [225] 尝试利用重建的 Feature 3DGS 对二维预训练编码器进行微调。

Others focus on incorporating text-visual alignment features for open-world understanding [226]-[228]. A significant challenge is the high dimensionality of CLIP features, which makes direct training and storage difficult compared to original Gaussian attributes. The work [226] introduces corresponding continuous semantic vectors into the 3DGS by extracting and discretizing dense features from CLIP [229] and DINO [230], which are used to predict semantic indices m in Discrete Feature Space by MLPs as in VQ-VAE [231]. Subsequently, FMGS [228] mitigates the issue of large CLIP feature dimensions by introducing multi-resolution hash encoders [171].

其他研究则侧重于引入文本-视觉对齐特征以实现开放世界理解 [226]-[228]。一个重要挑战是 CLIP 特征的高维度，使得直接训练和存储相比原始高斯属性更为困难。工作 [226] 通过提取并离散化 CLIP [229] 和 DINO [230] 的密集特征，将相应的连续语义向量引入 3DGS，利用 MLP 在离散特征空间中预测语义索引 m，类似于 VQ-VAE [231]。随后，FMGS [228] 通过引入多分辨率哈希编码器 [171] 缓解了 CLIP 特征维度过大的问题。

F. Physics Simulation

F. 物理仿真

Recent efforts aim to extend 3DGS to simulation. Based on the "what you see is what you simulate" philosophy, Phys-Gaussian [191] reconstructs a static 3DGS as the discretization of the scene to be simulated, and then incorporates continuum mechanics theory along with the Material Point Method [232] solver to endow 3DGS with physical properties. Similar approaches have also been explored in PhysDreamer [233]. Nevertheless, these physics-based simulations typically entail significant computational overhead. VR-GS [234] design an efficient interactive simulation system for VR, offering users a rich platform for 3D content manipulation.

近期工作致力于将 3DGS 扩展至仿真领域。基于"所见即所仿"的理念，Phys-Gaussian [191] 将静态 3DGS 重建为待仿真场景的离散化表示，随后结合连续介质力学理论和材料点法 (Material Point Method)[232] 求解器，赋予 3DGS 物理属性。类似方法也在 PhysDreamer [233] 中有所探索。然而，这些基于物理的仿真通常计算开销较大。VR-GS [234] 设计了一个高效的交互式虚拟现实仿真系统，为用户提供丰富的 3D 内容操作平台。

VI. 3DGS COMPONENTS AND IMPROVEMENTS

VI. 3DGS 组件与改进

3DGS can generally be divided into the following stages, as illustrated in Fig. 2: initialization, attribute optimization, splatting, regularization, training strategy, and adaptive control. In follow-up studies, researchers have enhanced and refined the fundamental 3DGS components. These technical improvements not only enhance the rendering performance of the original 3DGS but also address specific tasks in derivative works. Consequently, this section delves into the technological improvements in 3DGS, with the aim of providing valuable insights for researchers in related fields.

3DGS 通常可分为以下几个阶段，如图 2 所示: 初始化、属性优化、光斑渲染、正则化、训练策略和自适应控制。在后续研究中，学者们对基础 3DGS 组件进行了增强和改进。这些技术提升不仅提高了原始 3DGS 的渲染性能，还针对衍生任务进行了优化。因此，本节深入探讨 3DGS 的技术改进，旨在为相关领域的研究者提供有价值的参考。
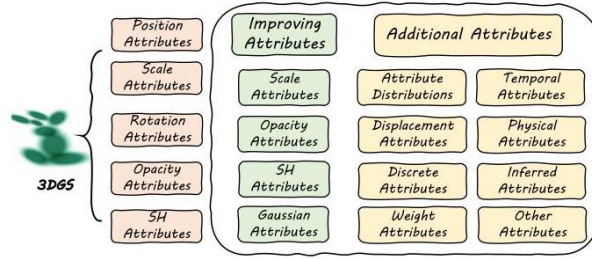
Fig. 4: Overview of Attribute Expansion Strategies.

图 4: 属性扩展策略概览。

A. Initialization

A. 初始化

Proper initialization has been shown to be crucial, as it directly affects the optimization process [235]. The initialization of 3DGS is typically performed using sparse points derived from Structure-from-Motion (SfM) or through random generation. However, these methods are often unreliable, especially under weak supervision signals.

适当的初始化被证明至关重要，因为它直接影响优化过程 [235]。3DGS 的初始化通常采用来自结构光束法 (Structure-from-Motion，SfM) 的稀疏点或随机生成。然而，这些方法在弱监督信号下往往不够可靠。

Combining pre-trained models is an optional approach. Pre-training a 3D model on a limited number of 3D samples and using it as an initialization prior is a viable strategy [52]. This approach can enhance the performance of initialization to some extent, although its effectiveness is contingent upon the data used. To address this limitation, pretrained 3D generative models [100], [101], [105] or monocular depth estimation models [62], [129], [130] are also introduced for initialization. Based on these, the work [100] also introduces some perturbation points to achieve a more comprehensive geometric representation.

结合预训练模型是一种可选方案。在有限的 3D 样本上预训练 3D 模型并将其作为初始化先验是一种可行策略 [52]。该方法在一定程度上提升了初始化性能，但其效果依赖于所用数据。为解决此限制，还引入了预训练的 3D 生成模型 [100]、[101]、[105] 或单目深度估计模型 [62]、[129]、[130] 用于初始化。在此基础上，工作 [100] 还引入了一些扰动点，以实现更全面的几何表示。

Improving initialization strategies is also important. Based on the analysis of the role of SfM in capturing low-frequency signals within the spectrum, Sparse-large-variance initialization is designed to effectively focus on the low-frequency distribution identified by SfM [235].

改进初始化策略同样重要。基于对 SfM 在捕捉频谱低频信号中作用的分析，设计了稀疏大方差初始化方法，有效聚焦于 SfM 识别的低频分布 [235]。

Utilizing other representations can also enhance initialization capabilities. By determining the Local Volumes from a coarse model, a small number of Gaussians are initialized within each volume, thereby avoiding excessive assumptions [236]. Similarly, some initialization based on Visual Hull [64], Flame [78] or Mesh [80], [217], [237] are proposed, enabling the acquisition of structural geometric priors.

> 利用其他表示形式也能增强初始化能力。通过从粗糙模型确定局部体积，在每个体积内初始化少量高斯，避免过度假设 [236]。类似地，还提出了基于视觉包络 (Visual Hull)[64]、Flame [78] 或网格 (Mesh)[80]、[217]、[237] 的初始化方法，实现结构几何先验的获取。

Discussion: The incorporation of pre-trained or structured information during the initialization of 3DGS is crucial, as high-fidelity initialization can mitigate training instabilities, particularly in under-determined scenarios. It is feasible to determine initialization protocols according to the available prior knowledge.

> 讨论: 在 3D 高斯表示 (3DGS) 初始化过程中引入预训练或结构化信息至关重要，因为高保真度的初始化能够缓解训练中的不稳定性，尤其是在欠定场景中。根据可用的先验知识确定初始化方案是可行的。

B. Attribute Expansion

> ## B. 属性扩展

The original attributes of 3DGS include the position, scale, rotation, Spherical Harmonic (SH) coefficients, and opacity value. Some works have extended these attributes to make them more suitable for downstream tasks. It can be categorized into improvements of existing attributes or the introduction of novel attributes, as shown in Fig. 4.

> 3DGS 的原始属性包括位置、尺度、旋转、球谐函数 (Spherical Harmonic, SH) 系数和不透明度值。一些工作对这些属性进行了扩展，使其更适合下游任务。可分为对现有属性的改进或引入新属性，如图 4 所示。

1) Improving Attributes: Certain attributes of vanilla Gaussian can be custom-tailored, thereby making 3DGS suitable for a wider range of tasks.

> 1) 属性改进: 可以定制部分基础高斯属性，从而使 3DGS 适用于更广泛的任务。

Scale: By collapsing the z-scale to zero and incorporating additional supervision on depth, normal, or shell maps, the works [26], [92], [192], [197], [199] aim to improve Gaussian primitives to make them flatter and more suitable for surface reconstruction, where z direction can be approximated as the normal direction. Conversely, a scale constraint, which limits the ratio of the major axis length to the minor axis length [150], [191], [238], ensures that the Gaussian primitives remain spherical to mitigate the issue of unexpected plush artifacts caused by overly skinny kernels.

SH: By combining hash grids and MLP, the corresponding color attributes are encoded, effectively addressing the storage issues caused by a large number of SH parameters [14].

Opacity: By constraining the transparency to approach either 0 or 1, thereby minimizing the number of semi-transparent Gaussian primitives, the works [34], [192] achieve clearer Gaussian surfaces, effectively alleviating artifacts.

Gaussian: By introducing shape parameters, an attempt is made to replace the original Gaussians with a Generalized Exponential Family (GEF) mixture [17]. Traditional 3DGS can be viewed as a special case of the GEF mixture $(\beta = 2)$, enhancing the representational efficiency of Gaussians,

$$\widehat{G}(\mathbf{x}) = \exp\left\{-\frac{1}{2}\left((\mathbf{x} - \mu)^{\top}\sum{}^{-1}(\mathbf{x} - \mu)\right)^{\frac{\beta}{2}}\right\}. \tag{3}$$

2) Additional Attributes: By adding new attributes and corresponding supervisions, the original representation capabilities of 3DGS can be augmented.

Semantic Attributes: By introducing them and corresponding supervision, works such as [147], [148], [160]-[162], [220], [223]-[225] are endowed with enhanced spatial semantic awareness, which is crucial for tasks such as SLAM and editing. After the semantic attributes' splatting, the 3DGS's semantic attributes are supervised using 2D semantic segmentation maps. Additionally, methods to improve the extraction of semantic information [211] and introducing high-dimensional semantic-text, such as CLIP and DINO features [226]-[228], have been employed to address a wider range of downstream tasks. Similar to semantic attributes, the identity encoding attributes can group 3DGS that belong to the same instance or stuff [219], which is more effective for multi-object scenes.

语义属性: 通过引入语义属性及对应监督, 文献 [147]、[148]、[160]-[162]、[220]、[223]-[225] 赋予 3DGS 更强的空间语义感知能力, 这对 SLAM 和编辑等任务至关重要。语义属性的投影后, 3DGS 的语义属性通过二维语义分割图进行监督。此外, 改进语义信息提取的方法 [211] 及引入高维语义文本特征, 如 CLIP 和 DINO 特征 [226]-[228], 被用于应对更广泛的下游任务。类似于语义属性, 身份编码属性可将属于同一实例或物体的 3DGS 分组 [219], 对多目标场景更为有效。

Attribute Distributions: Learning position distributions with reparameterization techniques instead of a fixed value is an effective approach to prevent local minima [103] and mitigate its reliance on Adaptive Control of 3DGS [46]. In addition to these works focusing on the distribution prediction of position attributes, the distribution of the scale has also been incorporated [103]. By sampling the predicted distributions, Gaussian primitives can be obtained.

属性分布: 采用重参数化技术学习位置分布而非固定值, 是防止陷入局部最优 [103] 及减轻对 3DGS 自适应控制依赖 [46] 的有效方法。除关注位置属性分布预测的工作外, 尺度分布也被纳入考虑 [103]。通过采样预测分布, 可获得高斯基元。

Temporal Attributes: Replacing the original static attributes with temporal attributes is key to animating the 3DGS [140], [147], [188], [189]. For 4D attributes, including rotation, scale, and position, existing works render 3DGS on timestep t by either taking time slices [189] or decoupling the t dimension from 4D attributes [140], [188]. Also, the introduction of 4D SH is crucial for time-varying color attributes. For this, the Fourier series is typically used as the adopted basis functions to endow SH with temporal capabilities [147], [188]. Note that due to involving different timesteps, these attributes often require video-based training. This regularization primarily aims to improve attributes in Gaussian primitives [26], [80], [98], [191], [238], as in Sec. VI-B.

时间属性: 用时间属性替代原始静态属性是实现 3DGS 动画的关键 [140]、[147]、[188]、[189]。对于包括旋转、尺度和位置的 4D 属性, 现有工作通过时间切片 [189] 或将 t 维度从 4D 属性中解耦 [140]、[188] 来渲染 3DGS。此外, 引入 4D SH 对时变颜色属性至关重要。通常采用傅里叶级数作为基函数, 赋予 SH 时间能力 [147]、[188]。注意, 由于涉及不同时间步, 这些属性通常需要基于视频的训练。此正则化主要旨在提升高斯基元中的属性 [26]、[80]、[98]、[191]、[238], 如第 VI-B 节所述。

Displacement Attributes: They can describe the relationship between the final and initial attributes in Gaussian primitives and be classified based on their condition. Condition-independent displacement attributes are often used to refine coarse attributes, which can be directly optimized in the same manner as other attributes [45]. Condition-dependent displacement attributes can describe the changes of static 3DGS, thereby achieving dynamic representations and controllable representations. This approach often involves introducing a small MLP to predict displacement based on timestep t [174]- [176], expression and other control signals [41], [70], [76], [79], [83]–[86], [187].

位移属性: 描述高斯基元最终属性与初始属性之间的关系, 可根据条件分类。无条件位移属性常用于细化粗糙属性, 可像其他属性一样直接优化 [45]。有条件位移属性描述静态 3DGS 的变化, 实现动态和可控表示。此方法通常引入小型 MLP, 根据时间步 t [174]-[176]、表情及其他控制信号 [41]、[70]、[76]、[79]、[83]-[86]、[187] 预测位移。

Physical Attributes: They encompass a broad range of properties describing the physical laws governing Gaus-

sian primitives, thus endowing 3DGS with more realistic representation. For instance, shading-related attributes like diffuse color, direct specular reflection, residual color, shadow, and anisotropic spherical Gaussian can be used for specular reconstruction and relighting [34]-[36], [177], [217], [218]. Additionally, the velocity attributes can represent the transient information of Gaussian, essential for describing dynamic objects [178]. These attributes are typically optimized by considering the influence of physical laws at specific rendering positions [34], [36], [177] or by incorporating supplementary information, such as flow maps [178].

> 物理属性: 它们涵盖描述高斯基元 (Gaussian primitives) 物理规律的广泛属性, 从而赋予 3D 高斯表示 (3DGS) 更逼真的表现。例如, 与着色相关的属性如漫反射颜色、直接镜面反射、残余颜色、阴影和各向异性球面高斯可用于镜面反射重建和重光照 [34]-[36], [177], [217], [218]。此外, 速度属性能够表示高斯的瞬态信息, 对于描述动态物体至关重要 [178]。这些属性通常通过考虑特定渲染位置的物理规律影响 [34], [36], [177] 或结合辅助信息如流场图 [178] 进行优化。

Discrete Attributes: Utilizing discrete attributes in place of continuous ones is an effective method for compressing high-dimensional representations and representing complex motion. This is often achieved by storing the index values of the VQ codebook [10]-[12], [14], [120] or the motion coefficient for motion basis [182] as the discrete attributes in Gaussian primitives. However, discrete attributes inevitable lead to performance degradation; combining them with compressed continuous attributes may be a potential solution [226].

> 离散属性: 使用离散属性替代连续属性是一种有效的高维表示压缩和复杂运动表达方法。通常通过存储向量量化 (VQ) 码本的索引值 [10]-[12], [14], [120] 或运动基的运动系数 [182] 作为高斯基元中的离散属性实现。然而, 离散属性不可避免地导致性能下降；将其与压缩的连续属性结合可能是一种潜在的解决方案 [226]。

Inferred Attributes: These attributes do not require optimization; they are inferred from other attributes. The Parameter-Sensitivity attributes reflects the impact of parameter changes on reconstruction performance and are represented by the gradient of the parameter, guiding compression clustering [12]. The Pixel-Coverage attributes determines the relative size of Gaussian primitives at the current resolution. It is related to the horizontal or vertical size of the Gaussian primitives and guides their scale to meet sampling requirements in multi-scale rendering [30].

> 推断属性: 这些属性无需优化, 而是从其他属性推断得出。参数敏感性属性反映参数变化对重建性能的影响, 通过参数的梯度表示, 指导压缩聚类 [12]。像素覆盖属性决定当前分辨率下高斯基元的相对大小, 关联高斯基元的水平或垂直尺寸, 指导其尺度以满足多尺度渲染中的采样需求 [30]。

Weight Attributes: They rely on structured representations, such as Local Volumes [236], Gaussian-kernel RBF [184], Mesh [237], and SMPL [239], to determine the attributes of query points by calculating the weights of structured points.

> 权重属性: 它们依赖于结构化表示, 如局部体积 (Local Volumes)[236]、高斯核径向基函数 (Gaussian-kernel RBF)[184]、网格 (Mesh)[237] 和 SMPL[239], 通过计算结构点的权重来确定查询点的属性。

Other Attributes: The Uncertainty Attributes can help maintain training stability by reducing the loss weight in areas with high uncertainty [58], [226]. The ORB-Features Attributes, extracted from image frames [157], play a crucial role in establishing 2D-to-2D and 2D-to-3D correspondences [152].

Discussion: The modification of Gaussian attributes facilitates the execution of a wider range of downstream tasks, offering an efficient approach as it obviates the need for additional structural elements. Moreover, the integration of new attributes with supplementary information constraints also has the potential to significantly enhance the representational efficacy of the original 3DGS. For instance, semantic attributes can, in certain scenarios, yield more precise object boundaries.

讨论: 高斯属性的修改促进了更广泛下游任务的执行，提供了一种高效方法，因为无需额外结构元素。此外，将新属性与辅助信息约束结合，也有望显著提升原始 3DGS 的表现能力。例如，语义属性在某些场景下能够提供更精确的物体边界。

C. Splatting

C. 点溅 (Splatting)

The role of Splatting is to efficiently transform 3D Gaussian data into high-quality 2D images, ensuring smooth, continuous projections and significantly improving rendering efficiency. As a core technology in traditional computer graphics, there are also efforts aimed at improving it from the perspectives of efficiency and performance.

点溅的作用是高效地将 3D 高斯数据转换为高质量的二维图像，确保平滑连续的投影，并显著提升渲染效率。作为传统计算机图形学的核心技术，也有从效率和性能角度进行改进的努力。

Several studies focus on enhancing splatting mechanisms. By analyzing the residual errors from the first-order Taylor expansion, the work [240] establishes a correlation between these errors and the Gaussian mean position and an unified projection plane to mitigate the projection errors through the Unit Sphere Based Rasterizer. Analytic-Splatting [33] advances pixel-center splatting through the introduction of pixel-window-based approximation, archiving anti-aliasing. To enhance 3DGS performance in complex scenarios, a GPU-accelerated ray tracing algorithm is introduced [241] for semi-transparent particle-based representations, achieving real-time rendering with support for complex effects such as secondary rays, depth of field, and distorted cameras. Additional improvements focus on optimizing splat ordering during the blending process. StopThePop [28] proposes a hierarchical per-pixel sorting strategy, which eliminates popping artifacts and ensures view-consistent real-time rendering by accurately computing the depth of Gaussians along individual rays.

多项研究聚焦于增强点溅机制。通过分析一阶泰勒展开的残差误差，文献 [240] 建立了这些误差与高斯均值位置及统一投影平面之间的关联，利用基于单位球的光栅化器减轻投影误差。解析点溅 (Analytic-Splatting)[33] 通过引入基于像素窗口的近似，推进了像素中心点溅，实现了抗锯齿效果。为提升 3DGS 在复杂场景中的性能，提出了基于 GPU 加速的光线追踪算法 [241]，用于半透明粒子表示，实现了支持次级光线、景深和畸变相机等复杂效果的实时渲染。其他改进聚焦于混合过程中的点溅排序优化。StopThePop[28] 提出了分层的逐像素排序策略，通过精确计算沿单条光线的高斯深度，消除弹跳伪影，确保视图一致的实时渲染。

D. Regularization

D. 正则化

Regularization is crucial for 3D reconstruction. We categorize the regularization terms into 2D and 3D regularization, as shown in Fig. 5 The 3D regularization directly constrains 3DGS, while the 2D regularization imposes constraints on the rendered images, thus influencing attribute optimization.

正则化对于 3D 重建至关重要。我们将正则化项分为二维和三维正则化，如图 5 所示。三维正则化直接约束 3DGS，而二维正则化则对渲染图像施加约束，从而影响属性优化。

1) 3D Regularization: The 3D regularization has garnered significant attention due to its intuitive constraint capabilities. These efforts categorized based on their targeted objectives into individual Gaussian primitive, local, and global regularization.

1) 三维正则化: 三维正则化因其直观的约束能力而备受关注。这些工作根据目标对象分为单个高斯基元、本地和全局正则化。

Individual Gaussian Primitive Regularization: This regularization primarily aims to improve attributes in Gaussian primitives [26], [80], [98], [191], [238], as Sec. VI-B.

单个高斯基元正则化: 该正则化主要旨在提升高斯基元中的属性 [26], [80], [98], [191], [238]，详见第 VI-B 节。

Local Regularization: Owing to the explicit representation of 3DGS, it is meaningful to impose constraints on Gaussian primitives within local regions. Such constraints can ensure the continuity and feasibility of Gaussian primitives in the local space. Physics-related Regularization is often used to ensure the local rigidity of deformable targets, which includes short-term local-rigidity loss, local rotation similarity loss, and long-term local isometry loss. Short-term local rigidity implies that nearby Gaussians should move following a rigid body transformation between time steps,

局部正则化: 由于 3D 高斯场 (3DGS) 的显式表示，在局部区域对高斯基元施加约束是有意义的。这些约束可以确保局部空间中高斯基元的连续性和可行性。与物理相关的正则化通常用于保证可变形目标的局部刚性，包括短期局部刚性损失、局部旋转相似性损失和长期局部等距损失。短期局部刚性意味着相邻的高斯应当遵循时间步之间的刚体变换移动，

$$\mathcal{L}_{i,j}^{\text{rigid}} = w_{i,j} \left\| (\mu_{j,t-1} - \mu_{i,t-1}) - \mathbf{R}_{i,t-1}\mathbf{R}_{i,t}^{-1} (\mu_{j,t} - \mu_{i,t}) \right\|_2,$$

(4)

where μ is Gaussian mean position, i and j are indexing of neighboring points, t is timestep, and $\mathbf{R}$ reperents rotation; Local rotation similarity enforces that adjacent Gaussian primitives have the same rotation over time steps,

> 其中 μ 是高斯均值位置，i 和 j 是邻近点的索引，t 是时间步，$\mathbf{R}$ 表示旋转；局部旋转相似性强制相邻高斯基元在时间步之间具有相同的旋转，

$$\mathcal{L}_{i,j}^{\text{rot}} = w_{i,j} \left\| \widehat{\mathbf{q}}_{j,t}\widehat{\mathbf{q}}_{j,t-1}^{-1} - \widehat{\mathbf{q}}_{i,t}\widehat{\mathbf{q}}_{i,t-1}^{-1} \right\|_2,$$

(5)

where $\widehat{\mathbf{q}}$ is the normalized quaternion representation of each Gaussian's rotation; Long-term local isometry loss prevents elements of the scene from drifting apart,

> 其中 $\widehat{\mathbf{q}}$ 是每个高斯旋转的归一化四元数表示；长期局部等距损失防止场景元素相互分离，

$$\mathcal{L}_{i,j}^{\text{iso}} = w_{i,j} \left| \left\| \mu_{j,0} - \mu_{i,0} \right\|_2 - \left\| \mu_{j,t} - \mu_{i,t} \right\|_2 \right|,$$

(6)

[133], [134], [169], [177], [182], [184], [187], [236]. Subsequently, some works have also adopted similar paradigms to constrain local rigidity [144], [179], [183].

> , [134], [169], [177], [182], [184], [187], [236]。随后，一些工作也采用了类似范式来约束局部刚性 [144], [179], [183]。

In addition to rigidity loss, Conservation of Momentum Regularization can also be used as a constraint in dynamic scene reconstruction. It encourages a constant velocity vector and applies a low-pass filter effect to the 3D trajectories, thereby smoothing out trajectories with sudden changes [177]. In addition, there are some Local Consistency Regularization terms that also aim to constrain the Gaussian primitives within local regions to maintain similar attributes, such as semantic attributes [219], [221], [226], position [126], [139], time [183], frame [172], normal [242], weight [38] and depth [56], [62].

> 除了刚性损失，动量守恒正则化也可作为动态场景重建中的约束。它鼓励速度向量保持恒定，并对 3D 轨迹施加低通滤波效果，从而平滑突变轨迹 [177]。此外，还有一些局部一致性正则项，旨在约束局部区域内的高斯基元保持相似属性，如语义属性 [219], [221], [226]，位置 [126], [139]，时间 [183]，帧 [172]，法线 [242]，权重 [38] 和深度 [56], [62]。

Global Regularization: Unlike the local regularization within neighboring regions, global regularization aims to constrain the overall 3DGS. Physics-related Regularization introduces real-world physical laws to constrain the state of 3DGS, including gravity loss and contact loss, among others. Gravity loss is used to constrain the relationship between the object and the floor, while contact loss targets the relationships among multiple objects [126].

> 全局正则化: 不同于邻域内的局部正则化，全局正则化旨在约束整体 3DGS。与物理相关的正则化引入现实物理规律来约束 3DGS 的状态，包括重力损失和接触损失等。重力损失用于约束物体与地面的关系，而接触损失针对多个物体之间的关系 [126]。

Benefiting from the explicit representation, depth and normal attributes can be directly calculated and constrained during training, particularly for surface reconstruction tasks. Depth-Normal Regularization achieves depth-normal consistency by comparing the normal computed from depth values with the predicted normal [34], [194], [197], [198]. This method enforces constraints on both normal and depth simultaneously. Additionally, directly constraining either the normal or the depth is also feasible. Normal Regularization often adopts a self-supervised paradigm due to the lack of direct supervision signals, which can be implemented by designing pseudo-labels from gradients [193], the shortest axis direction of Gaussian primitives [34], or SDF [194], [195]. Similarly, Depth Regularization adopts a similar approach; however, it not only aims for accurate depth values but also seeks to ensure clear surfaces in 3DGS. Depth Distortion loss [20] aggregates Gaussian primitives along the ray,

得益于显式表示，深度和法线属性可以在训练过程中直接计算和约束，尤其适用于表面重建任务。深度-法线正则化通过比较由深度值计算的法线与预测法线，实现深度-法线一致性 [34], [194], [197], [198]。该方法同时对法线和深度施加约束。此外，单独约束法线或深度也是可行的。法线正则化通常采用自监督范式，因缺乏直接监督信号，可通过设计基于梯度的伪标签 [193]、高斯基元最短轴方向 [34] 或符号距离函数 (SDF)[194], [195] 实现。类似地，深度正则化采用类似方法；但其不仅追求深度值的准确性，还力求确保 3DGS 表面的清晰度。深度畸变损失 [20] 沿射线聚合高斯基元，

$$\mathcal{L}_d = \sum_{i,j} \omega_i \omega_j \left| z_i - z_j \right|, \tag{7}$$

where z is the intersection depth of Gaussian [198],[199]. In addition to self-supervised methods, incorporating additional pre-trained models to estimate normal [197] and depth [35], [57], [59], [243] has proven to be more effective in Normal Regularization and Depth Regularization. Derived works introduce hard depth and soft depth regularization to address the geometry degradation and achieve more complete surfaces [59]. Similarly, SDF Regularization is also a constraint strategy for surface reconstruction. It achieves the desired surface by constraining the SDF that corresponds to 3DGS to an ideal distribution [192]-[195], [244].

其中 z 是高斯的交点深度 [198],[199]。除了自监督方法，结合额外的预训练模型来估计法线 [197] 和深度 [35], [57], [59], [243] 已被证明在法线正则化和深度正则化中更为有效。衍生工作引入硬深度和软深度正则化，以解决几何退化并实现更完整的表面 [59]。类似地，SDF 正则化也是表面重建的一种约束策略，通过将对应于 3DGS 的 SDF 约束为理想分布，实现期望的表面 [192]-[195], [244]。
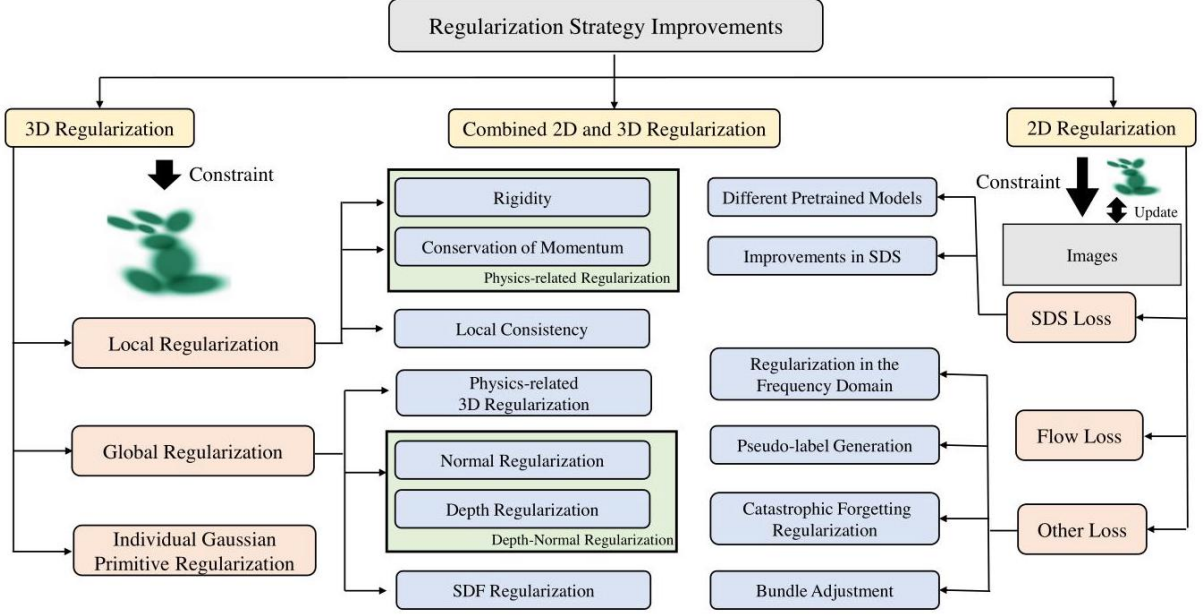
Fig. 5: Overview of Regularization Strategy Improvement in existing works.

图 5: 现有工作中正则化策略改进的概览。

2) 2D Regularization: Unlike the intuitive constraints in 3D, 2D regularization is often used to address under-constrained situations where original loss functions alone are insufficient.

2) 二维正则化: 不同于 3D 中的直观约束，二维正则化常用于解决原始损失函数不足以约束的问题。

SDS loss: An important example is the SDS loss, as shown in Eq 8, which uses a pre-trained 2D diffusion model to supervise 3DGS training via distillation paradigms [95], [201].

SDS 损失: 一个重要示例是 SDS 损失，如公式 8 所示，利用预训练的二维扩散模型通过蒸馏范式监督 3DGS 训练 [95], [201]。

$$\nabla_\theta = \mathbb{E}\left[w_t\left(\epsilon_\phi\left(\mathbf{x_t}, t, y\right) - \epsilon\right)\frac{\partial \mathbf{x}}{\partial \theta}\right]. \tag{8}$$

This approach extends to distill pre-trained 3D diffusion models [245], multi-view diffusion models [246], image editing models [203], and video diffusion models. Introducing 3D [100], [247] and multi-view diffusion models [96]-[98], [117], [133], [134], [139] enhances geometry and multi-view consistency. Image editing models [202] enable controllable edits, while video diffusion models [133] support dynamic temporal scene generation. Additionally, distillation on multimodal images, like RGB-Depth [93], also holds potential, providing more constraints from pre-trained diffusion models.

该方法扩展至蒸馏预训练的 3D 扩散模型 [245]、多视角扩散模型 [246]、图像编辑模型 [203] 和视频扩散模型。引入 3D[100], [247] 和多视角扩散模型 [96]-[98], [117], [133], [134], [139] 增强了几何和多视角一致性。图像编辑模型 [202] 实现了可控编辑，而视频扩散模型 [133] 支持动态时序场景生成。此外，对多模态图像如 RGB-深度 [93] 的蒸馏也具有潜力，为预训练扩散模型提供更多约束。

Some improvements specifically target inherent issues in SDS [104], [106]. Interval Score Matching is proposed to address issues of randomness and single-step sampling.

一些改进专门针对 SDS(Score-based Diffusion Models)[104]，[106] 中的固有问题。提出了区间分数匹配 (Interval Score Matching) 以解决随机性和单步采样的问题。

$$\nabla_\theta \mathcal{L}_{\text{ISM}}(\theta) := \mathbb{E}_{t,c}\left[\omega(t)\left(\epsilon_\phi(\mathbf{x_t}, t, y) - \epsilon_\phi(\mathbf{x_s}, s, \varnothing)\right)\frac{\partial g(\theta, c)}{\partial \theta}\right]$$

(9)where $s = t - \delta_T$ and δ_T denotes a small step size [104]. Introducing Negative Prompts [248] is a method [93], [104], [247] to mitigate the impact of random noise ϵ and enhance stability by replacing random noise with negative prompts $[\epsilon_\phi(\mathbf{x}_t; y_{\text{neg}})]$. And, LODS incorporates LoRA terms [249] $\left[\frac{(1-w)\epsilon_\psi(\mathbf{x}_t; \varnothing, t)}{w} - \frac{\epsilon}{w}\right]$ to replace traditional random noise ϵ, thereby alleviating the impact of out-of-distribution [106].

其中 $s = t - \delta_T$ 和 δ_T 表示一个小步长 [104]。引入负面提示 (Negative Prompts)[248] 是一种方法 [93]，[104]，[247]，通过用负面提示替代随机噪声 ϵ 来减轻随机噪声的影响并增强稳定性 $[\epsilon_\phi(\mathbf{x}_t; y_{\text{neg}})]$。此外，LODS 结合了 LoRA 项 [249] $\left[\frac{(1-w)\epsilon_\psi(\mathbf{x}_t; \varnothing, t)}{w} - \frac{\epsilon}{w}\right]$ 以替代传统的随机噪声 ϵ，从而缓解了分布外 (out-of-distribution) 问题 [106]。

Flow loss: It is a commonly used regularization term for dynamic 3DGS and uses the output of a pre-trained 2D optical flow estimation model as ground truth. Predicted flow is rendered by calculating the displacement of Gaussian primitives over a unit time and splatting these 3D displacements onto a 2D plane [134], [148], [179], [181]. However, this approach has a significant gap since optical flow is a 2D attribute and susceptible to noise. Selecting Gaussians with correct depth and introducing uncertainty through KL divergence to constrain optical flow is a potentially feasible method [178].

光流损失: 这是动态 3D 高斯场 (3DGS) 中常用的正则化项，使用预训练的二维光流估计模型的输出作为真实值。预测的光流通过计算高斯基元在单位时间内的位移并将这些 3D 位移投影到二维平面上进行渲染 [134]，[148]，[179]，[181]。然而，该方法存在显著差距，因为光流是二维属性且易受噪声影响。选择具有正确深度的高斯基元并通过 KL 散度引入不确定性以约束光流是一种潜在可行的方法 [178]。

Other loss: There are also some 2D regularization terms worth discussing. For example, constraining the differences in amplitude and phase between the rendered image and the ground truth in the frequency domain can serve as a loss function to aid training, thereby alleviating overfitting issues [27]. Introducing pseudo-labels for hypothetical viewpoints through noise perturbation can assist training in sparse-view settings [57]. In large-scale scene mapping, constraining the changes in attributes before and after optimization can prevent catastrophic forgetting in 3DGS [250]. Additionally, bundle adjustment is usually an important constraint in pose estimation problems [151], [152], [161].

其他损失: 还有一些值得讨论的二维正则化项。例如，在频域中约束渲染图像与真实图像在幅度和相位上的差异，可以作为损失函数辅助训练，从而缓解过拟合问题 [27]。通过噪声扰动引入假设视角的伪标签，可以辅助稀疏视角设置下的训练 [57]。在大规模场景映射中，约束优化前后的属性变化可以防止 3DGS 中的灾难性遗忘 [250]。此外，束调整 (bundle adjustment) 通常是位姿估计问题中的重要约束 [151]，[152]，[161]。

Noted that, whether 2D or 3D regularization is used, overall updating is sometimes suboptimal due to the large number of primitives. Some primitives often have an uncontrollable impact on the results. Therefore, it is necessary to guide the optimization by selecting important primitives using methods such as visibility [153], [160], [162], [164].

> 需要注意的是，无论使用 2D 还是 3D 正则化，由于基元数量庞大，整体更新有时并不理想。一些基元往往对结果产生不可控的影响。因此，有必要通过可见性等方法 [153]，[160]，[162]，[164] 选择重要基元来指导优化。

Discussion: The incorporation of regularization terms serves as an effective approach to enhance the reconstruction performance of 3DGS. These regularization terms can impose constraints on various attributes of 3DGS, including geometry and spatial distribution, etc., in accordance with specific task requirements. Furthermore, in under-determined scenarios, the optimization process can be further constrained by introducing additional prior information. For a particular task, it is feasible to incorporate multiple distinct constraints simultaneously.

> 讨论: 引入正则化项是提升 3DGS 重建性能的有效手段。这些正则化项可以根据具体任务需求，对 3DGS 的几何形状、空间分布等多种属性施加约束。此外，在欠定场景中，可以通过引入额外的先验信息进一步约束优化过程。对于特定任务，完全可以同时引入多种不同的约束。

E. Training Strategy

> ### E. 训练策略

Training strategy is also an important topic. In this section, we divide it into multi-stage training strategy and end-to-end training strategy, which can be applied to different tasks.

> 训练策略也是一个重要话题。本节将其划分为多阶段训练策略和端到端训练策略，分别适用于不同任务。

1) Multi-stage Training Strategy: Multi-stage training strategy is a common training paradigm, often involving coarse-to-fine reconstruction. It is widely used for under-determined tasks, such as AIGC, SLAM, etc..

> 1) 多阶段训练策略: 多阶段训练策略是一种常见的训练范式，通常涉及由粗到细的重建。它广泛应用于欠定任务，如 AIGC、SLAM 等。

Using different 3D representations in different training stages is a typical multi-stage training paradigm. 3DGS → Mesh (training 3DGS first, converting to Mesh, then optimizing Mesh) [95]-[97], [115], [202], [247] ensures geometric consistency in the generated 3D model. Additionally, generating multi-view images [109], [123], [140], [204]-[206], [247] in the first stage to aid reconstruction in the second stage can alleviate optimization difficulties.

在不同训练阶段使用不同的三维表示是典型的多阶段训练范式。3DGS → 网格 (先训练 3DGS，转换为网格，再优化网格)[95]-[97]，[115]，[202]，[247] 确保生成模型的几何一致性。此外，在第一阶段生成多视角图像 [109]，[123]，[140]，[204]-[206]，[247] 以辅助第二阶段的重建，可以缓解优化难题。

Two-stage reconstruction for static and dynamic reconstruction is also important in dynamic 3DGS. This type of work typically involves training a time-independent static 3DGS in the first stage, and then training a time-dependent deformation field in the second stage to characterize dynamic Gaussians [174]-[177], [181], [208]. Additionally, incremental reconstruction of dynamic scenes frame by frame is also a focus in some works, often relying on the performance of previous reconstructions [169], [170].

静态与动态重建的两阶段重建在动态 3DGS 中也很重要。这类工作通常在第一阶段训练时间无关的静态 3DGS，第二阶段训练时间相关的变形场以表征动态高斯 [174]-[177]，[181]，[208]。此外，逐帧增量重建动态场景也是部分工作的重点，通常依赖于之前重建的性能 [169]，[170]。

In multi-objective optimization tasks, multi-stage training paradigms can enhance stability and performance. For example, the coarse-to-fine camera tracking strategy first obtains a coarse camera pose from a sparse pixel set, then refines it based on optimized rendering results [151], [163].

在多目标优化任务中，多阶段训练范式可以提升稳定性和性能。例如，粗到细的相机跟踪策略先从稀疏像素集获得粗略的相机位姿，然后基于优化的渲染结果进行细化 [151]，[163]。

Additionally, some works aim to refine the 3DGS trained in the first stage [52], [64], [100], [105], [117], [210], [221], [236] or endow them with additional capabilities, such as semantics [164], [222] and stylization [213]. There are many such training strategies, which are also effective in maintaining training stability and avoiding local optima [15]. Furthermore, iterative optimization of the final result to enhance performance is also feasible [61], [206].

此外，一些工作旨在细化第一阶段训练的三维高斯场 (3DGS)[52]，[64]，[100]，[105]，[117]，[210]，[221]，[236]，或赋予其额外能力，如语义 [164]，[222] 和风格化 [213]。此类训练策略众多，且在保持训练稳定性和避免局部最优方面同样有效 [15]。此外，对最终结果进行迭代优化以提升性能也是可行的 [61]，[206]。

2) End-to-End Training Strategy: These strategies are often more efficient and can be applied to a wider range of downstream tasks. Some typical works are described in Fig. 6

2) 端到端训练策略: 这些策略通常更高效，且可应用于更广泛的下游任务。部分典型工作见图 6。

Progressive Optimization Strategy: This commonly used strategy helps 3DGS prioritize learning global representations before locally optimizing details. In the frequency domain, this can be viewed as progressively learning from low-frequency to high-frequency components. It is often implemented by gradually increasing the proportion of high-frequency signals [27], [235] or introducing progressively larger image/feature sizes for supervision [10], [36], [152], which can also improve efficiency [43], [150]. In generative tasks, progressively selecting the camera pose is also an easy-to-difficult training strategy, optimizing from positions close to the initial viewpoint to those further away [97], [130].

渐进优化策略: 这一常用策略帮助三维高斯场优先学习全局表示, 再局部优化细节。在频域中, 可视为从低频到高频成分的渐进学习。通常通过逐步增加高频信号比例 [27], [235] 或引入逐渐增大的图像/特征尺寸进行监督 [10], [36], [152] 实现, 也能提升效率 [43], [150]。在生成任务中, 渐进选择相机位姿也是一种由易到难的训练策略, 从接近初始视点的位置优化到更远的位置 [97], [130]。
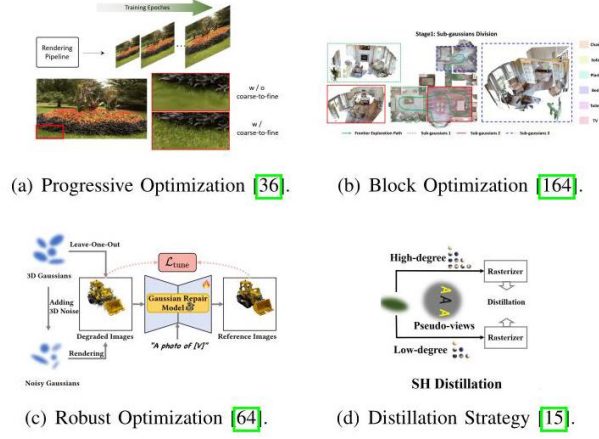

(a) Progressive Optimization [36]. (b) Block Optimization [164].

(c) Robust Optimization [64]. (d) Distillation Strategy [15].

Fig. 6: Four typical End-to-End Training Strategies.

图 6: 四种典型的端到端训练策略。

Block Optimization Strategy: This strategy is often used in large-scale scene reconstruction to improve efficiency and alleviate catastrophic forgetting [154], [163], [164]. However, such paradigms are often influenced by block partitioning and training data selection. Consequently, several studies have proposed designing Primitives and Data Division strategies to mitigate workload imbalances caused by numerous empty blocks, while enhancing detail reconstruction capabilities [251]. To improve efficiency, introducing Level of Detail and hierarchical reconstruction prove effective, especially in large-scale scene processing [251], [252]. It can also achieve reconstruction by partitioning the scene into static backgrounds and dynamic objects [146]-[148], [175]. Additionally, this approach is applied in AIGC and Semantic Understanding, where refining submap reconstruction quality enhances overall performance [105], [227]. Unlike submaps divided by spatial regions, Gaussians can be categorized into different generations during their densification process, allowing for the application of distinct regularization strategies to each generation, effectively regulating their fluidity [201]. Categorizing Gaussians into those on smooth surfaces and independent points is also feasible for geometric representation. By designing distinct initialization and densification strategies, better representation can be achieved [242]. Additionally, some works design keyframe (or window) selection strategies based on inter-frame covisibility or geometric overlap ratio in temporal data for reconstructions [151], [153], [160], [172], [238], [250].

Robust Optimization Strategy: Introducing noise perturbations is a common method to enhance the robustness of training [49], [64], [115], [174]. Such perturbations can target camera poses, timesteps, and images, and can be regarded as a form of data augmentation to prevent overfitting. Additionally, some strategies mitigate catastrophic forgetting by avoiding continuous training from a single viewpoint [154], [156].

Distillation-based Strategy: To compress model parameters, some distillation strategies use the original 3DGS as the teacher model and a low-dimensional SH 3DGS as the student model, introducing more pseudo views to enhance the performance of the low-dimensional SH [15].

Discussion: Improving training strategies is an efficient way to optimize the training process of 3DGS and can enhance performance in many tasks. End-to-end training strategies, in particular, can improve performance while ensuring efficiency.

F. Adaptive Control

Adaptive Control of 3DGS is an important process for regulating the number of Gaussian primitives, including cloning, splitting, and pruning. In the following sections, we will summarize existing techniques from the perspectives of densification (cloning and splitting) and pruning.

1) Densification: Densification is crucial, especially for detail reconstruction. we will analyze it from the perspectives of "Where to densify" and "How to densify". Additionally, we will discuss how to avoid excessive densification.

> 1) 致密化: 致密化尤为关键，特别是对细节重建。我们将从"在哪里致密化"和"如何致密化"两个方面进行分析，并讨论如何避免过度致密化。

Where to Densification: Densification techniques focus on identifying positions requiring densification, governed by gradients in the original 3DGS and extendable to dynamic scene reconstruction [170]. Regions with low opacity, silhouette, or high depth-rendered error, considered unreliable, guide densification to fill holes or improve 3D inconsistencies [26], [151], [164], [185], [197], [250]. Some approaches improve based on gradients by weighting the number of pixels covered by each Gaussian in different views to dynamically average view gradients, enhancing point cloud growth conditions [253]. Additionally, SDF value, motion masks and neighbor distance are important criteria, with locations closer to the surface, motion regions and lower compactness being more prone to densification [100], [179], [192], [195].

> 密集化的位置: 密集化技术侧重于识别需要密集化的位置，这些位置由原始三维高斯场 (3DGS) 的梯度决定，并可扩展到动态场景重建 [170]。低不透明度、轮廓或高深度渲染误差的区域被视为不可靠，指导密集化以填补空洞或改善三维不一致性 [26], [151], [164], [185], [197], [250]。一些方法基于梯度改进，通过加权不同视图中每个高斯覆盖的像素数，动态平均视图梯度，增强点云增长条件 [253]。此外，SDF 值、运动掩码和邻居距离是重要标准，靠近表面、运动区域和较低紧凑度的位置更易于密集化 [100], [179], [192], [195]。

How to Densification: Numerous works have improved densification methods. Graph structures explore Gaussian relationships and define new Gaussians at edge centers based on proximity scores, mitigating sparse viewpoint impacts [57]. To prevent excessive Gaussian growth, the Candidate Pool Strategy stores pruned Gaussians for densification [118]. Additionally, work [254] introduces three conservation rules for visual consistency and employs integral tensor equations to model densification.

> 密集化的方法: 大量工作改进了密集化方法。图结构探索高斯关系，并基于接近度分数在边缘中心定义新的高斯，缓解稀疏视点的影响 [57]。为防止高斯过度增长，候选池策略存储被剪枝的高斯以供密集化 [118]。此外，文献 [254] 引入了三条视觉一致性守恒规则，并采用积分张量方程对密集化进行建模。

Excessive densification is also unnecessary, as it directly impacts the efficiency of 3DGS. In cases where two Gaussian functions are in close proximity, limiting their densification is a straightforward idea, where the distance between Gaussians can be measured by Gaussian Divergent Significance [123] (GDS) or Kullback-Leibler divergence [69], where $\mu_1, \sum_1, \mu_2, \sum_2$ belongs to two adjacent Gaussians.

> 过度密集化也是不必要的，因为它直接影响 3DGS 的效率。在两个高斯函数距离较近的情况下，限制它们的密集化是一种直接的思路，高斯间距离可通过高斯发散显著性 (Gaussian Divergent Significance, GDS)[123] 或 Kullback-Leibler 散度 [69] 来衡量，其中 $\mu_1, \sum_1, \mu_2, \sum_2$ 属于两个相邻高斯。

$$\text{GDS}: \left\| \mu_1 - \mu_2 \right\|^2 + \text{tr}\left(\sum_1 + \sum_2 - 2\left(\sum_1^{-1} \sum_2 \sum_1^{-1} \right)^{1/2} \right),$$

$$\text{KL}: \frac{1}{2}\left(\text{tr}\left({\textstyle\sum}_2^{-1}{\textstyle\sum}_1 \right) + \ln\frac{\det {\textstyle\sum}_1}{\det {\textstyle\sum}_2} + (\mu_2 - \mu_1)^T {\textstyle\sum}_2^{-1}(\mu_2 - \mu_1) - \right.$$

(10)

And DeblurGS [40] incorporates a Gaussian Densification Annealing strategy to prevent the densification of inaccurate Gaussians during the early training stages at imprecise camera motion estimation. Furthermore, in some downstream tasks, densification is sometimes abandoned to prevent 3DGS from overfitting to each image, which could lead to incorrect geometric shapes [150], [151], [153], [238].

DeblurGS[40] 引入了高斯密集退火策略，以防止在相机运动估计不准确的早期训练阶段对不准确高斯的密集化。此外，在某些下游任务中，有时会放弃密集化，以防止 3DGS 对每张图像过拟合，避免产生错误的几何形状 [150], [151], [153], [238]。

2) Pruning: Removing unimportant Gaussian primitives can ensure efficient representation. In the initial 3DGS framework, opacity was employed as the criterion for determining the significance of a Gaussian. Subsequent research has explored the incorporation of scale as a guiding factor or distractor masks for pruning [49], [93]. However, these approaches primarily focus on individual Gaussian primitives, lacking a comprehensive consideration of the global representation. Therefore, subsequent derivative techniques are discussed.

2) 剪枝: 移除不重要的高斯基元可以确保高效表示。在初始 3DGS 框架中，不透明度被用作判断高斯重要性的标准。后续研究探索了将尺度作为指导因素或干扰掩码用于剪枝 [49], [93]。然而，这些方法主要关注单个高斯基元，缺乏对全局表示的综合考虑。因此，后续衍生技术被讨论。

Importance scores: The volume and hit count on training views, along with opacity, can be used to jointly determine the global significance score of a Gaussian primitive [15].

重要性评分: 训练视图中的体积和命中次数，以及不透明度，可联合用于确定高斯基元 (Gaussian primitive) 的全局重要性评分 [15]。

$$\text{GS}_j = \sum_{i=1}^{MHW} \mathbf{1}\left(G\left(\mathbf{X}_j \right), r_i \right) \cdot \sigma_j \cdot \gamma\left({\textstyle\sum}_j \right),$$

(11)

where $\gamma\left(\sum_j \right)$ and $\mathbf{1}\left(G\left(\mathbf{X}_j \right), r_i \right)$ are volume and hit count, and M represents the number of training views. Subsequently, Gaussians are ranked according to their global scores, and the ones with the lowest scores are pruned. Similar importance scores were improved in other works [255], [256].

其中 $\gamma\left(\sum_j \right)$ 和 $\mathbf{1}\left(G\left(\mathbf{X}_j \right), r_i \right)$ 分别是体积和命中次数，M 表示训练视图的数量。随后，根据全局评分对高斯进行排序，得分最低的将被剪枝。类似的重要性评分在其他工作中得到了改进 [255], [256]。

Multi-view consistency: Multi-view consistency is a key criterion for determining whether Gaussians need to be pruned. For example, [238] prunes newly added Gaussians that are not observed by three keyframes within a local keyframe window, while [162] prunes Gaussians that are invisible in all virtual views but visible in real views.

多视图一致性: 多视图一致性是判断高斯是否需要剪枝的关键标准。例如，[238] 剪枝在局部关键帧窗口内未被三个关键帧观测到的新添加高斯，而 [162] 剪枝在所有虚拟视图中不可见但在真实视图中可见的高斯。

Distance Metric: Surface-aware methods often use distance to the surface [151] and SDF values [195] to prune Gaussian primitives far from the surface. The distance between Gaussians is also a key metric [179]. GauHuman [69] aims to merge Gaussians with small scale and low KL divergence, as mentioned in Eq. 10.

距离度量: 基于表面的算法通常使用到表面距离 [151] 和符号距离函数 (SDF) 值 [195] 来剪枝远离表面的高斯基元。高斯之间的距离也是关键度量 [179]。GauHuman[69] 旨在合并尺度小且 KL 散度低的高斯，如公式 10 所示。

Learnable control parameter: Introducing a learnable mask based on scale and opacity to determine whether Gaussian primitives should be removed effectively prevents 3DGS from becoming overly dense [14].

可学习控制参数: 引入基于尺度和不透明度的可学习掩码来决定是否移除高斯基元，有效防止 3D 高斯点云 (3DGS) 过度密集 [14]。

Others: CoR-GS [63] aims to leverage mismatched regions between two 3DGS models, trained in parallel under identical conditions, as guidance for pruning.

其他:CoR-GS[63] 旨在利用在相同条件下并行训练的两个 3DGS 模型之间的不匹配区域，作为剪枝的指导。

Discussion: Adaptive Control strategies play a pivotal role in enhancing rendering fidelity and computational efficiency. However, excessive densification or pruning can adversely affect both the efficiency and performance of 3D Gaussian Splatting. Therefore, it is crucial to examine and establish an optimal balance between these two strategic approaches.

讨论: 自适应控制策略在提升渲染质量和计算效率方面起着关键作用。然而，过度密集或剪枝均会对 3D 高斯点云的效率和性能产生负面影响。因此，审视并建立这两种策略之间的最佳平衡至关重要。

VII. OTHER TECHNICAL DISCUSSIONS

七、其他技术讨论

Several existing works have explored the augmentation of the original 3DGS pipeline through the integration of supplementary components, including post-processing modules and additional information as well as representations. These architectural enhancements typically yield improved performance metrics and facilitate optimization under ill-posed conditions.

已有多项工作通过集成后处理模块、附加信息及表示等补充组件，探索了对原始 3DGS 流程的增强。这些架构改进通常带来性能提升，并有助于在病态条件下的优化。

A. Post-Processing

Post-processing strategies for pre-trained Gaussians are important, as they can improve the original efficiency and performance. Common post-processing often enhances Gaussian representations through various optimization strategies. This type of work has been discussed in Sec. VI-E

对预训练高斯的后处理策略十分重要，因为它们能提升原始效率和性能。常见的后处理通常通过多种优化策略增强高斯表示。此类工作已在第六章 E 节讨论。

Representation Conversion: Pre-trained 3DGS can be converted to Mesh using Poisson reconstruction [257] on sampled 3D points [192], [197]. Similarly, GOF [198] uses 3D bounding boxes to convert 3DGS to a Tetrahedral Grid, then extracts meshes using Binary Search of Level Set. Additionally, LGM [115] converts Pre-trained 3DGS to NeRF, then uses NeRF2Mesh [258] for Mesh conversion.

表示转换: 预训练的 3DGS 可通过对采样的三维点 [192], [197] 进行泊松重建 (Poisson reconstruction)[257] 转换为网格。同样，GOF[198] 利用三维包围盒将 3DGS 转换为四面体网格，再通过水平集的二分搜索提取网格。此外，LGM[115] 将预训练 3DGS 转换为 NeRF，再使用 NeRF2Mesh[258] 进行网格转换。

Performance and Efficiency: Some works enhance 3DGS performance in specific tasks through post-processing, such as multi-scale rendering. SA-GS [32] introduces a 2D scale-adaptive filter to dynamically adjust scales based on rendering frequency, enhancing anti-aliasing when zooming out. For efficiency, removing redundant Gaussian primitives from pre-trained 3DGS [23] or introducing a Gaussian caching mechanism [259] can improve rendering efficiency.

性能与效率: 部分工作通过后处理提升 3DGS 在特定任务中的表现，如多尺度渲染。SA-GS[32] 引入二维尺度自适应滤波器，根据渲染频率动态调整尺度，增强缩小时的抗锯齿效果。为提升效率，移除预训练 3DGS 中的冗余高斯基元 [23] 或引入高斯缓存机制 [259] 均可提高渲染效率。

B. Integration with Other Representations

The convertible nature of 3D representations facilitates the integration of 3DGS with other representations, leveraging their advantages to improve the original 3DGS.

3D 表示的可转换特性促进了 3DGS 与其他表示的集成，借助它们的优势改进原始 3DGS。

1) Point Clouds: Point clouds, as a 3D representation related to 3DGS, are often used to initialize positions. Converting point clouds to 3DGS can effectively fill holes [129], [130] or improve reconstruction details [119], typically after high-precision reconstruction. Conversely, 3DGS can be converted into point clouds, voxelized into 3D voxels, and projected onto 2D BEV grids, which guide navigation tasks [164]. Additionally, anchor points in space can assist 3DGS. These methods use voxel centers as anchor points to represent the scene. Each anchor point comprises a local context feature, a scaling factor, and multiple learnable offsets. By decoding other attributes based

on these offsets and features, the anchors transform into local neural Gaussians, which helps mitigate redundant expansion [16], [36], [195].

> 1) 点云: 点云作为与 3DGS 相关的三维表示, 常用于初始化位置。将点云转换为 3DGS 可有效填补空洞 [129], [130] 或提升重建细节 [119], 通常在高精度重建后进行。反之, 3DGS 可转换为点云, 体素化为三维体素, 并投影到二维鸟瞰图 (BEV) 网格, 用于导航任务 [164]。此外, 空间中的锚点可辅助 3DGS。这些方法使用体素中心作为锚点表示场景。每个锚点包含局部上下文特征、缩放因子及多个可学习偏移。通过基于这些偏移和特征解码其他属性, 锚点转化为局部神经高斯, 有助于缓解冗余扩展 [16], [36], [195]。

2) Mesh: Meshes have better geometric representation capabilities and can, to some extent, alleviate artifacts or blurry pixels caused by 3DGS [173]. They are still the most widely used 3D representation in downstream tasks [115]. Much work has discussed converting 3DGS to Mesh, as mentioned in Sec. V-B. Once converted, they can be optimized for better geometry and appearance [71], [95], [96], [202]. Jointly optimizing 3DGS and Mesh is also an optional strategy. 3DGS is suitable for constructing complex geometric structures, while Mesh can be used to reconstruct detailed color appearances on smooth surfaces. Combining the two can enhance reconstruction performance [173] and large-scale deformation control [237].

> 2) 网格 (Mesh): 网格具有更好的几何表示能力, 在一定程度上可以缓解由 3D 高斯场 (3DGS)[173] 引起的伪影或模糊像素问题。它们仍然是下游任务中最广泛使用的三维表示形式 [115]。大量工作讨论了将 3DGS 转换为网格的方法, 如第 V-B 节所述。转换后, 可以对其进行优化以获得更好的几何形状和外观 [71], [95], [96], [202]。联合优化 3DGS 和网格也是一种可选策略。3DGS 适合构建复杂的几何结构, 而网格可用于在光滑表面上重建细节丰富的颜色外观。两者结合可以提升重建性能 [173] 和大规模变形控制 [237]。

3) Triplane: Triplane, known for its compactness and efficient expressiveness [51], is often used in generalization tasks. It consists of three orthogonal feature planes: $X-Y, Y$-Z, and $X-Z$. Features can be obtained by querying positions in the space, and subsequently decoded to predict Gaussian attributes [51], [52], [67], [114]. Recent works [139], [176], [179], [208] extend triplane to 4D space ($XYZ-T$) using multi-scale HexPlanes [143] or 4D GaussianPlanes [208] to enhance 4DGS continuity in the spatiotemporal dimension.

> 3) 三平面 (Triplane): 三平面以其紧凑性和高效表达能力著称 [51], 常用于泛化任务。它由三个正交的特征平面组成: $X-Y, Y$、Z 和 $X-Z$。通过查询空间中的位置可以获得特征, 随后解码以预测高斯属性 [51], [52], [67], [114]。近期工作 [139], [176], [179], [208] 利用多尺度 HexPlanes[143] 或 4D 高斯平面 (4D GaussianPlanes)[208] 将三平面扩展到四维空间 ($XYZ-T$), 以增强时空维度上的 4D 高斯场 (4DGS) 连续性。

4) Grid: Grid is also an efficient representation, as it can access grid corners and interpolate to obtain features or attributes at specific positions. Hash grid [171], a representative method, can compress scenes and achieve a more compact and efficient 3DGS [14], [19], [70], [118], [120], [150], [213], [228]. Furthermore, Self-Organizing Gaussian [18] maps unstructured 3D Gaussians onto a 2D grid to preserve local spatial relationships, where adjacent Gaussians have similar attribute values, reducing memory storage and maintaining continuity in 3D space.

4) 网格 (Grid): 网格也是一种高效的表示方法，因为它可以访问网格顶点并通过插值获得特定位置的特征或属性。哈希网格 (Hash grid)[171] 作为代表性方法，能够压缩场景，实现更紧凑高效的 3DGS[14], [19], [70], [118], [120], [150], [213], [228]。此外，自组织高斯 (Self-Organizing Gaussian)[18] 将无结构的 3D 高斯映射到二维网格上，以保持局部空间关系，其中相邻高斯具有相似的属性值，减少内存存储并维持 3D 空间的连续性。

Particularly, GaussianVolumes are also used for generalizable representations [118], where a volume is composed of a fixed number of 3DGS. This maintains the efficiency of 3DGS and offers greater manipulability compared to triplane.

特别地，高斯体积 (GaussianVolumes) 也被用于泛化表示 [118]，其中一个体积由固定数量的 3DGS 组成。这保持了 3DGS 的效率，并相比三平面提供了更大的可操作性。

5) Implicit Representation: Implicit representations, benefiting from their representational capability, can be used to mitigate the condition difficulty and surface artifacts of 3DGS [91]. Specifically, introducing NeRF to encode color and opacity can significantly enhance the representation's adjustability [260]. Moreover, by designing an SDF-to-opacity transformation function [194] or employing mutual geometry supervision [195] to jointly optimize 3DGS and SDF representations, the surface reconstruction performance of 3DGS can be improved.

5) 隐式表示 (Implicit Representation): 隐式表示因其强大的表达能力，可用于缓解 3DGS 的条件难题和表面伪影 [91]。具体而言，引入神经辐射场 (NeRF) 编码颜色和不透明度，能显著增强表示的可调节性 [260]。此外，通过设计从有符号距离函数 (SDF) 到不透明度的转换函数 [194] 或采用互相几何监督 [195] 联合优化 3DGS 和 SDF 表示，可以提升 3DGS 的表面重建性能。

Discussion: Given the inherently unstructured characteristics of 3DGS, the incorporation of structured representations emerges as a viable prior, particularly advantageous for tasks such as human body reconstruction, facilitating both cross-representation transformations and geometric reconstruction, thereby enabling enhanced performance in downstream applications. However, due to certain limitations in representational characteristics, this approach may be accompanied by some degradation in rendering performance.

讨论: 鉴于 3DGS 固有的无结构特性，结合结构化表示作为先验成为一种可行策略，特别适用于人体重建等任务，促进跨表示转换和几何重建，从而提升下游应用的性能。然而，由于表示特性的某些限制，该方法可能伴随渲染性能的部分下降。

C. Guidance by Additional Prior

C. 额外先验的引导

When dealing with under-determined problems, such as sparse view settings III-C introducing additional priors is a straightforward method to improve 3DGS performance.

在处理欠定问题时，如稀疏视角设置 III-C，引入额外先验是一种直接有效的提升 3DGS 性能的方法。

Pre-trained Models: Introducing pre-trained models is an effective paradigm that can guide the optimization through the model's knowledge. Pre-trained monocular depth models and point cloud prediction models are a common type of priors, where the predicted depth values and positions can be used for the initialization and regularization [56]-[58], [62], [129], [130], [162]. Pre-trained 2D image (or 3D and video) generative models are also important in some AIGC-related tasks. They can be used not only for optimization in combination with SDS Loss [100], [133], [247] but also for directly generating (or editing) images for training [61], [109], [129], [130], [140]. Similarly, some works introduce pre-trained image inpainting networks to alleviate difficulties caused by occlusion as well as overlap [97], [129], [130], [201], [211] or super-resolution models for a high level of detail [130], [208] during the generation process. Additionally, pre-trained ControlNet [207] or Large Language Models can also be used to guide 3D generation. The former can enhance geometric consistency under depth guidance [97], [98], [205], while the latter can predict layout maps to guide spatial relationships in multi-object 3D generation scenarios [98]. Notably, certain pre-trained models can endow 3DGS with additional capabilities, such as semantic understanding models, as discussed in Sec. V-E and spatial understanding models [162].

> 预训练模型: 引入预训练模型是一种有效范式, 可以通过模型知识指导优化。预训练的单目深度模型和点云预测模型是常见的先验类型, 预测的深度值和位置可用于初始化和正则化 [56]-[58], [62], [129], [130], [162]。预训练的二维图像 (或三维及视频) 生成模型在某些 AIGC 相关任务中也非常重要。它们不仅可结合 SDS 损失 [100], [133], [247] 用于优化, 还可直接用于生成 (或编辑) 训练图像 [61], [109], [129], [130], [140]。类似地, 一些工作引入预训练的图像修复网络以缓解遮挡和重叠带来的困难 [97], [129], [130], [201], [211], 或在生成过程中使用超分辨率模型以获得高细节 [130], [208]。此外, 预训练的 ControlNet[207] 或大型语言模型也可用于指导三维生成。前者可在深度引导下增强几何一致性 [97], [98], [205], 后者可预测布局图以指导多目标三维生成场景中的空间关系 [98]。值得注意的是, 某些预训练模型还能赋予 3DGS 额外能力, 如语义理解模型 (详见第 V-E 节) 和空间理解模型 [162]。

More Sensors: Due to the 3D-agnostic nature of 2D images, reconstructing 3DGS can be challenging, especially in large-scale reconstructions such as SLAM and autonomous driving. Therefore, incorporating additional sensors for 3D depth information, including depth sensors [154], [156], [161], [162], [250], audio [87]-[89], LiDAR [50], [146], [149], [150], [163], and optical tactile sensors [58], has the potential to alleviate this issue.

> 更多传感器: 由于二维图像的三维无关性, 重建三维几何结构 (3DGS) 具有挑战性, 尤其是在大规模重建如同步定位与地图构建 (SLAM) 和自动驾驶中。因此, 结合额外的三维深度信息传感器, 包括深度传感器 [154]、[156]、[161]、[162]、[250], 音频 [87]-[89], 激光雷达 (LiDAR)[50]、[146]、[149]、[150]、[163], 以及光学触觉传感器 [58], 有望缓解这一问题。

Task-specific Priors: Some reconstruction tasks, such as human reconstruction, target subjects with certain common characteristics. These characteristics, such as template models and Linear Blend Skinning, can be extracted as priors to guide the reconstruction of similar targets. In the reconstruction, animation, and generation of non-rigid objects, many works utilize SMPL [74] and SMAL [261] to provide strong priors for representing the motion and deformation of non-rigid objects like humans [65], [67], [69], [70], [93] and animals [144], [239]. Subsequently, based on the SMPL template, Shell Maps [262] and template meshes are also introduced in combination with 3DGS to address issues of low efficiency in 3DGAN [92], [94] and unclear geometry [71], [72]. Similarly, in head and face reconstruction and animation tasks, some works [77], [82] also use the FLAME model [81] as a prior. Linear Blend Skinning [263] is also employed as prior knowledge to assist in the prediction of 3DGS motion [82], [184]. Additionally, in 3D urban scene reconstruction tasks, HUGS [148] introduces the Unicycle Model to

model the motion of vehicles, thereby making the motion modeling of moving objects smoother.

> 任务特定先验: 某些重建任务, 如人体重建, 针对具有某些共性特征的目标。这些特征, 如模板模型和线性混合蒙皮 (Linear Blend Skinning), 可作为先验提取, 用以指导相似目标的重建。在非刚性物体的重建、动画和生成中, 许多工作利用 SMPL[74] 和 SMAL[261] 提供强有力的先验, 用于表示人体 [65]、[67]、[69]、[70]、[93] 及动物 [144]、[239] 等非刚性物体的运动和变形。随后, 基于 SMPL 模板, 结合三维几何结构 (3DGS) 引入了 Shell Maps[262] 和模板网格, 以解决 3DGAN[92]、[94] 中效率低下及几何形状不清晰的问题 [71]、[72]。类似地, 在头部和面部重建及动画任务中, 一些工作 [77]、[82] 也采用 FLAME 模型 [81] 作为先验。线性混合蒙皮 [263] 亦被用作先验知识, 辅助预测 3DGS 运动 [82]、[184]。此外, 在三维城市场景重建任务中, HUGS[148] 引入了单轮车模型 (Unicycle Model) 来建模车辆运动, 从而使移动物体的运动建模更加平滑。

Discussion: Accessible auxiliary information has the capability to enhance the performance of 3DGS across numerous tasks, serving as prior knowledge to facilitate spatial comprehension, particularly in inherently ill-posed problems. Although certain priors or sensors may lead to increased computational overhead and costs, they have the capability to significantly enhance the representational capacity of 3DGS .

> 讨论: 可获取的辅助信息能够提升三维几何结构 (3DGS) 在众多任务中的表现, 作为先验知识促进空间理解, 尤其是在本质上病态的问题中。尽管某些先验或传感器可能导致计算开销和成本增加, 但它们能够显著增强 3DGS 的表达能力。

VIII. CHALLENGES AND OPPORTUNITIES

八、挑战与机遇

The preceding discussion indicates that various 3DGS-related tasks share similar technical approaches, which stems from common challenges across different tasks. To provide readers with a deeper understanding of this phenomenon, this section examines the commonalities among different tasks, summarizes four core challenges as well as their corresponding technical solutions, and outlines future opportunities.

> 前述讨论表明, 各种与三维几何结构 (3DGS) 相关的任务共享相似的技术方法, 这源于不同任务间的共性挑战。为使读者更深入理解这一现象, 本节探讨不同任务的共性, 总结四大核心挑战及其对应的技术解决方案, 并展望未来机遇。

A. Interrelationships

A. 任务间关系

We have extensively discussed various 3DGS-related tasks in Sec. III, Sec. IV, and Sec. V, revealing common challenges and techniques across these tasks. As illustrated in Tab V, we categorize existing tasks according to four core challenges, demonstrating that solutions from different tasks can be mutually instructive. Furthermore, there are some interrelationships between different tasks that have not been mentioned. For instance, Surface Reconstruction techniques (Sec V-B) are often referenced in the context of Editable 3DGS (Sec V-C), etc. We anticipate that this analysis will offer valuable insights for future research endeavors in related tasks.

我们在第三、四、五节中广泛讨论了多种与三维几何结构 (3DGS) 相关的任务，揭示了这些任务中的共性挑战与技术。如表 V 所示，我们根据四大核心挑战对现有任务进行分类，展示不同任务的解决方案可以相互借鉴。此外，不同任务间还存在一些未被提及的关联。例如，表面重建技术 (第五节 B 部分) 常被用于可编辑三维几何结构 (第五节 C 部分) 等。我们期望此分析为相关任务的未来研究提供有价值的见解。

B. Suboptimal Data

B. 数据不足

Challenges. In real-world scenarios, collecting large volumes of high-quality training data is often impractical. Without access to 3D data and sufficient multi-view images, relying on limited 2D image supervision is insufficient for accurate 3DGS reconstruction. For example, inferring the back appearance from only a frontal image is highly challenging. Additionally, data quality is critical, as accurate poses and clear images directly influence reconstruction performance.

挑战: 在现实场景中，收集大量高质量训练数据往往不切实际。缺乏三维数据和足够的多视角图像时，仅依赖有限的二维图像监督难以实现准确的三维几何结构 (3DGS) 重建。例如，仅凭正面图像推断背面外观极具挑战性。此外，数据质量至关重要，准确的姿态和清晰的图像直接影响重建效果。

Opportunities. An ideal 3DGS training process requires sufficient high-quality data, but this is often excessively challenging in practical applications. Although introducing priors can mitigate this problem to some extent, optimizing a large number of Gaussians under underconstrained conditions remains difficult. A potential solution is to reduce the number of Gaussian primitives based on their uncertainty while enhancing the representational capacity of individual primitives [17]. This involves finding a trade-off between the number of Gaussians and rendering performance, thereby improving the efficiency of utilizing sparse samples. Then, poor-quality data should also be taken into consideration. Unconstrained in-the-wild images are a typical case, encompassing transient occlusions and dynamic appearance changes, such as varying sky, weather, and lighting, which have been extensively discussed in NeRF [266]-[268]. To enhance efficiency, existing works have addressed this issue in the context of 3DGS [269], [270], attempting to model appearance changes and handle transient objects. However, their performance struggles, especially in scenes with complex lighting changes and frequent occlusions. Thanks to the explicit representation characteristics of 3DGS, decoupling geometric representations and introducing geometric consistency constraints across different scenes is a promising approach to mitigate instability during the training process.

C. Generalization

Challenges. Despite the improved training efficiency compared to NeRF, the scene-specific training paradigm remains a major bottleneck for the application of 3DGS. It is hard to imagine having to train for each target or scene individually, especially in multi-target and scene reconstruction (generation).

Opportunities. Although existing generalization-related works can directly obtain scene representations through forward inference, their performance is often unsatisfactory and limited by the type of scene [45], [47], [51], [114]. We hypothesize that this is due to the difficulty of feedforward networks in performing the adaptive control of 3DGS, as also mentioned in [46]. In future research, designing a reference-feature-based feedforward adaptive control strategy is a potential solution, which can predict the positions requiring adaptive control through reference features and be plug-and-play into existing generalization-related works. Additionally, existing generalization-related works rely on accurate poses, which are often difficult to obtain in practical applications [265], [271], [272]. Therefore, discussing generalizable 3DGS under pose-missing conditions is also promising [259].

TABLE V: The Relationships among Challenges, Tasks, and Technological Improvement, where the first column represents the core challenges, while the second and third columns denote the related downstream tasks and the corresponding technological advancements associated with these challenges, respectively

Challenges	Major Tasks	Major Technological Improvements
Suboptimal Data Chal- lenges VIII-B	Limited Number: Sparse Views III-B Autonomous Driving Reconstruction IV-C1 Dynamic 3DGS V-A (Monocular Video Part V-A2), AIGC IV-B and Editable 3DGS V-C Limited Quality: Slam IV-C2 and Reconstruction under blurred images (Sec III-B) or without poses [264], [265]	Limited Number: Initialization VI-A Regulariza- tion VI-D Adaptive Control VI-F Training Strate- gies V-6 and Guidance by Additional Prior VII-C Limited Quality: Training Strategies VI-6 and Inte- gration with Other Representations (Sec. VI-G2)
Generalization Chal- lenges VIII-C	Generalization III-C and Generalization-related tasks in the Human Reconstruc- tion IV-A and AIGC IV-B	Initialization VI-A Adaptive Control VI-F and Inte- gration with Other Representations VI-G2
Physical Chal- lenges VIII-D	Physical Motion: Dynamic 3DGS V-A Physics Simulation V-F Anime- tion V-C4 Dynamic Human IV-A and Autonomous Driving Reconstruction IV-C1 Physical Rendering: Photorealism III-B and Physics Simulation V-F	Physical Motion and Rendering: Attribute Ex- pansion VI-B-partsen VI-B Regularization Strategy VI-D and Guidance by Additional Prior VII-C
Realness and Efficiency Challenges VIII-E	Realness: Photorealism III-B Surface Reconstruction V-B Semantic Under- standing V-E some AIGC-related IV-B and Autonomous Driving IV-C works. Efficiency III-A some works in Autonomous Driving IV-C and Semantic Understanding V-E	Realness: Most of the technologies in VI Efficiency: Attribute Expansion VI-B Post-Processing VI-E2 Adaptive Control VI-F and Splatting VI-C

挑战	主要任务	主要技术改进
次优渲染质量 VIII-B	数据有限：稀疏观测 III-D 自动驾驶视角重建 IV-C1 动态三维几何建模 (3DGS) 项目视频部分 V-A2, 人工智能生成内容 (AIGC) IV-B 和可编辑三维几何建模 V-C 质量有限，同步定位与地图构建 (SLAM) IV-C2 以及模糊图像 (第 III-B 节) 或无位姿估计下的重建 [264], [265]	数据增强：初始化 VI-A 正则化 VI-D 自适应控制 VI-F 训练策略 VI-E 及渲染先验引导 VII-C 质量增强：训练策略 VI-E 及其他表示方法的整合 (第 VI-F2 节)
泛化挑战 VIII-C	泛化：III-C 及人体建模 IV-A 和人工智能生成内容 (AIGC) IV-B 中泛化相关任务	初始化 VI-A 自适应控制 VI-F 及其他表示方法的整合 VI-F2
物理属性 VIII-D	物理运动：动态三维几何建模 (3DGS) V-A 物理防真 V-F 动画 V-C4 动态 IV-A 及自动驾驶重建 IV-C 物理渲染：照片真实感 III-B 及物理的真实 V-f	物理运动的渲染：属性扩展 V-B 及正则化建模 VI-D 支撑特化分引导 VII-C
真实感与效率难题 VIII-E	真实感：照片真实感 III-B 表面建模 V-B 语义理解 V-E 及部分人工智能生成内容 (AIGC) IV-B 相关工作，效率：效率 III-A 及自动驾驶 IV-C 和语义理解 V-E 中的部分工作	真实感：大部分 V-E 中的技术改进 属性扩展 V-B 后处理 VI-F2 自适应控制 VI-F 及点云溅射 (Splatting) V-C

D. Physics Reconstruction and Rendering

D. 物理重建与渲染

    Challenges. Traditional 3DGS only considers static rendering and neglects the laws of physical motion, which are important in simulations [191]. Additionally, Physically-based rendering is a significant step towards applying 3DGS to simulate the physical world and achieve more realistic effects.

> 挑战。传统的三维高斯点渲染 (3DGS) 仅考虑静态渲染，忽视了物理运动规律，而这些规律在仿真中至关重要 [191]。此外，基于物理的渲染是将 3DGS 应用于模拟物理世界并实现更逼真效果的重要一步。

    Opportunities. Ensuring that the 3DGS's motion adheres to physical laws is essential for unifying simulation and rendering [191]. Although rigidity-related regularization have been introduced, as described in Sec. VI-D1, most existing works focus on animating 3DGS while neglecting the physical attributes of the Gaussian primitives themselves (Sec. V-A). Some pioneering works attempt to introduce velocity attributes [178] and Newtonian dynamics rules [191], but this is not sufficient to fully describe the physical motion of 3DGS in space. A potential solution is to introduce more physical attributes in Gaussian primitives, such as material [216], acceleration, and force distribution, which can be regularized by priors from certain simulation tools and physics knowledge. Physically-based rendering is also a direction worth attention, as it enables 3DGS to handle relighting and material editing, producing outstanding inverse rendering results [215]. Future works can explore decoupling geometry and appearance in 3DGS, conducting research from the perspectives of normal reconstruction and the modeling of illumination and materials [91], [216], [273].

> 机遇。确保 3DGS 的运动遵循物理定律对于统一仿真与渲染至关重要 [191]。尽管如第 VI-D1 节所述，已有刚性相关的正则化方法被引入，但大多数现有工作侧重于 3DGS 的动画制作，而忽视了高斯基元本身的物理属性 (见第 V-A 节)。一些开创性工作尝试引入速度属性 [178] 和牛顿动力学规则 [191]，但这仍不足以全面描述 3DGS 在空间中的物理运动。一个潜在的解决方案是为高斯基元引入更多物理属性，如材料 [216]、加速度和力分布，这些属性可以通过某些仿真工具和物理知识的先验进行正则化。基于物理的渲染也是一个值得关注的方向，因为它使 3DGS 能够处理重新光照和材质编辑，产生出色的逆向渲染结果 [215]。未来的工作可以探索在 3DGS 中解耦几何与外观，从法线重建以及光照和材质建模的角度开展研究 [91], [216], [273]。

E. Realness and Efficiency

E. 真实感与效率

    Challenges. Realness and efficiency challenges are fundamental issues. They are investigated in various works and have been discussed in Sec. III. In this part, we discuss downstream tasks and techniques optimized for performance and efficiency.

Opportunities. The difficulty in reconstructing clear surfaces has always been a significant challenge affecting rendering realism. As discussed in Sec. V-B, some works have addressed it by attempting to represent surfaces with planar Gaussians. However, this can result in a decline in rendering performance, possibly due to the reduced representational capacity of planar Gaussian primitives or the training instability. Therefore, designing Gaussian primitives better suited for surface representation and introducing a multi-stage training paradigm along with regularization are potential solutions. Storage efficiency is one of the critical bottlenecks of 3DGS. Existing works focus on introducing VQ techniques and compressing SH parameters, as discussed in Sec. III-A1. However, such approaches inevitably affect performance. Therefore, exploring how to design more efficient representations based on 3DGS is a potential way to enhance efficiency [16], [17] while maintaining performance.

机遇。重建清晰表面一直是影响渲染真实感的重要挑战。如第 V-B 节所述，一些工作尝试用平面高斯表示表面，但这可能导致渲染性能下降，原因可能是平面高斯基元的表达能力降低或训练不稳定。因此，设计更适合表面表示的高斯基元，并引入多阶段训练范式及正则化，是潜在的解决方案。存储效率是 3DGS 的关键瓶颈之一。现有工作集中于引入向量量化 (VQ) 技术和压缩球谐函数 (SH) 参数，如第 III-A1 节所述，但这些方法不可避免地影响性能。因此，探索基于 3DGS 设计更高效的表示方法，是提升效率 [16]，[17] 同时保持性能的潜在途径。

F. Other Technical Challenges and Opportunities

F. 其他技术挑战与机遇

An increasing number of research and engineering projects have found that initialization is important in 3DGS. Traditional SfM is not suitable for constrained scenarios, such as sparse view settings, AIGC, and low-light reconstruction. Therefore, more robust initialization methods should be designed to replace original initialization in these scenarios. And Splatting also plays an important role in 3DGS, but it is rarely mentioned in existing works [240], [274]. Designing efficient parallel splatting strategies on pretrained 3DGS has the potential to impact rendering performance and efficiency.

越来越多的研究和工程项目发现，初始化在 3DGS 中非常重要。传统的结构光束法 (SfM) 不适用于受限场景，如稀疏视角设置、人工智能生成内容 (AIGC) 和低光重建。因此，应设计更鲁棒的初始化方法以替代这些场景中的原始初始化。同时，点溅 (Splatting) 在 3DGS 中也扮演重要角色，但现有工作中很少提及 [240]，[274]。在预训练 3DGS 上设计高效的并行点溅策略，有望提升渲染性能和效率。

IX. Conclusion

IX. 结论

The burgeoning interest in the field of 3D Gaussian Splatting (3DGS) has precipitated the emergence of a myriad of related downstream tasks and technologies, thereby contributing to increasing complexity and confu-

sion within the domain. This complexity manifests in various forms, including similar motivations across different works, the incorporation of similar technologies across disparate tasks, and the nuanced differences and interconnections between various technologies. This survey systematically categorizes existing works based on their motivations and critically discusses associated technologies. Our objective is to elucidate common challenges across tasks and technologies, providing a coherent framework for understanding this evolving field. Additionally, we highlight prospective research avenues to inspire continued innovation in 3DGS. REFERENCES

对三维高斯点渲染 (3DGS) 领域日益增长的兴趣催生了大量相关的下游任务和技术，导致该领域的复杂性和混乱性不断增加。这种复杂性表现为不同工作间相似的动机、不同任务中相似技术的融合，以及各种技术间细微的差异和联系。本文系统地根据动机对现有工作进行分类，并对相关技术进行批判性讨论。我们的目标是阐明任务和技术中的共性挑战，提供理解这一不断发展的领域的连贯框架。此外，我们还强调了未来的研究方向，以激发 3DGS 领域的持续创新。参考文献

[1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R.Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," Commun. ACM, vol. 65, no. 1, pp. 99-106, 2021.

B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R.Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," Commun. ACM, vol. 65, no. 1, pp. 99-106, 2021.

[2] G. Chen and W. Wang, "A survey on 3d gaussian splatting," arXiv:2401.03890, 2024.

G. Chen and W. Wang, "A survey on 3d gaussian splatting," arXiv:2401.03890, 2024.

[3] T. Wu, Y.-J. Yuan, L.-X. Zhang, J. Yang, Y.-P. Cao, L.-Q. Yan, and L. Gao, "Recent advances in 3d gaussian splatting," Comput. Vis. Media, pp. 1-30, 2024.

T. Wu, Y.-J. Yuan, L.-X. Zhang, J. Yang, Y.-P. Cao, L.-Q. Yan, and L. Gao, "Recent advances in 3d gaussian splatting," Comput. Vis. Media, pp. 1-30, 2024.

[4] B. Fei, J. Xu, R. Zhang, Q. Zhou, W. Yang, and Y. He, "3d gaussian splatting as new era: A survey," IEEE Trans. Vis. Comput. Graph., 2024.

B. Fei, J. Xu, R. Zhang, Q. Zhou, W. Yang, 和 Y. He, "3D 高斯点渲染新时代: 综述," IEEE 可视化与计算机图形学汇刊, 2024.

[5] G. Liao, K. Zhou, Z. Bao, K. Liu, and Q. Li, "Ov-nerf: Open-vocabulary neural radiance fields with vision and language foundation models for 3d semantic understanding," IEEE Trans. Circuits Syst. Video Technol., 2024.

G. Liao, K. Zhou, Z. Bao, K. Liu, 和 Q. Li, "Ov-nerf: 基于视觉与语言基础模型的开放词汇神经辐射场用于 3D 语义理解," IEEE 电路系统视频技术汇刊, 2024.

[6] P. Zhang, X. Wang, L. Ma, S. Wang, S. Kwong, and J. Jiang, "Progressive point cloud upsampling via differentiable rendering," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 12, pp. 4673-4685, 2021.

P. Zhang, X. Wang, L. Ma, S. Wang, S. Kwong, 和 J. Jiang, " 通过可微渲染的渐进式点云上采样," IEEE 电路系统视频技术汇刊, 第 31 卷, 第 12 期, 页 4673-4685, 2021.

[7] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3d gaussian splatting for real-time radiance field rendering," ACM Trans. Graph., vol. 42, no. 4, pp. 1-14, 2023.

> B. Kerbl, G. Kopanas, T. Leimkühler, 和 G. Drettakis, " 用于实时辐射场渲染的 3D 高斯点渲染," ACM 图形汇刊, 第 42 卷, 第 4 期, 页 1-14, 2023.

[8] M. Zwicker, H. Pfister, J. Van Baar, and M. Gross, "Ewa volume splatting," in Proc. IEEE Vis. IEEE, 2001, pp. 29-538.

> M. Zwicker, H. Pfister, J. Van Baar, 和 M. Gross, "EWA 体积点渲染," 载 IEEE 可视化会议论文集, 2001, 页 29-538.

[9] J. Ding, Y. He, B. Yuan, Z. Yuan, P. Zhou, J. Yu, and X. Lou, "Ray reordering for hardware-accelerated neural volume rendering," IEEE Trans. Circuits Syst. Video Technol., 2024.

> J. Ding, Y. He, B. Yuan, Z. Yuan, P. Zhou, J. Yu, 和 X. Lou, " 用于硬件加速神经体渲染的光线重排序," IEEE 电路系统视频技术汇刊, 2024.

[10] S. Girish, K. Gupta, and A. Shrivastava, "Eagles: Efficient accelerated 3d gaussians with lightweight encodings," arXiv:2312.04564, 2023.

> S. Girish, K. Gupta, 和 A. Shrivastava, "Eagles: 具有轻量编码的高效加速 3D 高斯点," arXiv:2312.04564, 2023.

[11] K. Navaneet, K. P. Meibodi, S. A. Koohpayegani, and H. Pirsiavash, "Compact3d: Compressing gaussian splat radiance field models with vector quantization," arXiv:2311.18159, 2023.

> K. Navaneet, K. P. Meibodi, S. A. Koohpayegani, 和 H. Pirsiavash, "Compact3d: 使用矢量量化压缩高斯点辐射场模型," arXiv:2311.18159, 2023.

[12] S. Niedermayr, J. Stumpfegger, and R. Westermann, "Compressed 3d gaussian splatting for accelerated novel view synthesis," arXiv:2401.02436, 2023.

> S. Niedermayr, J. Stumpfegger, 和 R. Westermann, " 用于加速新视角合成的压缩 3D 高斯点渲染," arXiv:2401.02436, 2023.

[13] W. H. Equitz, "A new vector quantization clustering algorithm," IEEE Trans. Acoust. Speech Signal Process., vol. 37, no. 10, pp. 1568-1575, 1989.

> W. H. Equitz, " 一种新的矢量量化聚类算法," IEEE 声学语音信号处理汇刊, 第 37 卷, 第 10 期, 页 1568-1575, 1989.

[14] J. C. Lee, D. Rho, X. Sun, J. H. Ko, and E. Park, "Compact 3d gaussian representation for radiance field," arXiv:2311.13681, 2023.

J. C. Lee, D. Rho, X. Sun, J. H. Ko, 和 E. Park, ” 用于辐射场的紧凑 3D 高斯表示,” arXiv:2311.13681, 2023.

[15] Z. Fan, K. Wang, K. Wen, Z. Zhu, D. Xu, and Z. Wang, ”Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200 + fps ,” arXiv:2311.17245, 2023.

Z. Fan, K. Wang, K. Wen, Z. Zhu, D. Xu, 和 Z. Wang, ”Lightgaussian: 具有 15x 降维和 200 + fps 的无界 3D 高斯压缩,” arXiv:2311.17245, 2023.

[16] T. Lu, M. Yu, L. Xu, Y. Xiangli, L. Wang, D. Lin, and B. Dai, ”Scaffold-gs: Structured 3d gaussians for view-adaptive rendering,” arXiv:2312.00109, 2023.

T. Lu, M. Yu, L. Xu, Y. Xiangli, L. Wang, D. Lin, 和 B. Dai, ”Scaffold-gs: 用于视角自适应渲染的结构化 3D 高斯点,” arXiv:2312.00109, 2023.

[17] A. Hamdi, L. Melas-Kyriazi, J. Mai, G. Qian, R. Liu, C. Vondrick, B. Ghanem, and A. Vedaldi, ”Ges: Generalized exponential splatting for efficient radiance field rendering,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 19812-19822.

A. Hamdi, L. Melas-Kyriazi, J. Mai, G. Qian, R. Liu, C. Vondrick, B. Ghanem, 和 A. Vedaldi, ”GES: 用于高效辐射场渲染的广义指数点渲染,” 载 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 页 19812-19822.

[18] W. Morgenstern, F. Barthel, A. Hilsmann, and P. Eisert, ”Compact 3d scene representation via self-organizing gaussian grids,” arXiv:2312.13299, 2023.

W. Morgenstern, F. Barthel, A. Hilsmann, 和 P. Eisert, ” 通过自组织高斯网格实现紧凑 3D 场景表示,” arXiv:2312.13299, 2023.

[19] Y. Chen, Q. Wu, W. Lin, M. Harandi, and J. Cai, ”Hac: Hash-grid assisted context for 3 d gaussian splatting compression,” in Eur. Conf. Comput. Vis. Springer, 2025, pp. 422-438.

Y. Chen, Q. Wu, W. Lin, M. Harandi, 和 J. Cai, ”Hac: 基于哈希网格辅助的高斯点云压缩,” 载于欧洲计算机视觉会议, Springer, 2025, 页 422-438。

[20] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, ”Mip-nerf 360: Unbounded anti-aliased neural radiance fields,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 5470-5479.

J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, 和 P. Hedman, ”Mip-nerf 360: 无界抗锯齿神经辐射场,” 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2022, 页 5470-5479。

[21] J. C. Lee, D. Rho, X. Sun, J. H. Ko, and E. Park, ”Compact 3d gaussian representation for radiance field,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 21719-21728.

J. C. Lee, D. Rho, X. Sun, J. H. Ko, 和 E. Park, ” 用于辐射场的紧凑三维高斯表示,” 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 页 21719-21728。

[22] S. Durvasula, A. Zhao, F. Chen, R. Liang, P. K. Sanjaya, and N. Vijaykumar, "Distwar: Fast differentiable rendering on raster-based rendering pipelines," arXiv:2401.05345, 2023.

S. Durvasula, A. Zhao, F. Chen, R. Liang, P. K. Sanjaya, 和 N. Vijaykumar, "Distwar: 基于光栅渲染管线的快速可微分渲染," arXiv:2401.05345, 2023。

[23] J. Jo, H. Kim, and J. Park, "Identifying unnecessary 3d gaussians using clustering for fast rendering of 3 d gaussian splatting," arXiv:2402.13827, 2024.

J. Jo, H. Kim, 和 J. Park, " 利用聚类识别不必要的三维高斯以加速高斯点云渲染," arXiv:2402.13827, 2024。

[24] J. Lee, S. Lee, J. Lee, J. Park, and J. Sim, "Gscore: Efficient radiance field rendering via architectural support for 3 d gaussian splatting," in Proc. 29th ACM Int. Conf. Archit. Support Program. Lang. Oper. Syst., Volume 3, 2024, pp. 497-511.

J. Lee, S. Lee, J. Lee, J. Park, 和 J. Sim, "Gscore: 通过架构支持实现高效的辐射场渲染," 载于第 29 届 ACM 国际架构支持程序语言与操作系统会议, 第 3 卷, 2024, 页 497-511。

[25] D. Verbin, P. Hedman, B. Mildenhall, T. Zickler, J. T. Barron, and P. P. Srinivasan, "Ref-nerf: Structured view-dependent appearance for neural radiance fields," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. IEEE, 2022, pp. 5481-5490.

D. Verbin, P. Hedman, B. Mildenhall, T. Zickler, J. T. Barron, 和 P. P. Srinivasan, "Ref-nerf: 结构化视角依赖外观的神经辐射场," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, IEEE, 2022, 页 5481-5490。

[26] K. Cheng, X. Long, K. Yang, Y. Yao, W. Yin, Y. Ma, W. Wang, and X. Chen, "Gaussianpro: 3d gaussian splatting with progressive propagation," arXiv:2402.14650, 2024.

K. Cheng, X. Long, K. Yang, Y. Yao, W. Yin, Y. Ma, W. Wang, 和 X. Chen, "Gaussianpro: 具有渐进传播的三维高斯点云," arXiv:2402.14650, 2024。

[27] J. Zhang, F. Zhan, M. Xu, S. Lu, and E. Xing, "Fregs: 3d gaussian splatting with progressive frequency regularization," arXiv:2403.06908, 2024.

J. Zhang, F. Zhan, M. Xu, S. Lu, 和 E. Xing, "Fregs: 具有渐进频率正则化的三维高斯点云," arXiv:2403.06908, 2024。

[28] L. Radl, M. Steiner, M. Parger, A. Weinrauch, B. Kerbl, and M. Steinberger, "Stopthepop: Sorted gaussian splatting for view-consistent real-time rendering," ACM Trans. Graph., vol. 43, no. 4, pp. 1-17, 2024.

L. Radl, M. Steiner, M. Parger, A. Weinrauch, B. Kerbl, 和 M. Steinberger, "Stopthepop: 用于视图一致实时渲染的排序高斯点云," ACM 图形学汇刊, 第 43 卷, 第 4 期, 页 1-17, 2024。

[29] S. Diolatzis, T. Zirr, A. Kuznetsov, G. Kopanas, and A. Kaplanyan, "N-dimensional gaussians for fitting of high dimensional functions," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-11.

S. Diolatzis, T. Zirr, A. Kuznetsov, G. Kopanas, 和 A. Kaplanyan, " 用于高维函数拟合的 N 维高斯," 载于 ACM SIGGRAPH 会议论文, 2024, 页 1-11。

[30] Z. Yan, W. F. Low, Y. Chen, and G. H. Lee, "Multi-scale 3d gaussian splatting for anti-aliased rendering," arXiv:2311.17089, 2023.

Z. Yan, W. F. Low, Y. Chen, 和 G. H. Lee, " 多尺度三维高斯点云用于抗锯齿渲染," arXiv:2311.17089, 2023。

[31] Z. Yu, A. Chen, B. Huang, T. Sattler, and A. Geiger, "Mip-splatting: Alias-free 3d gaussian splatting," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 19447-19456.

Z. Yu, A. Chen, B. Huang, T. Sattler, 和 A. Geiger, "Mip-splatting: 无别名三维高斯点云," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 页 19447-19456。

[32] X. Song, J. Zheng, S. Yuan, H.-a. Gao, J. Zhao, X. He, W. Gu, and H. Zhao, "Sa-gs: Scale-adaptive gaussian splatting for training-free anti-aliasing," arXiv:2403.19615, 2024.

X. Song, J. Zheng, S. Yuan, H.-a. Gao, J. Zhao, X. He, W. Gu, 和 H. Zhao, "Sa-gs: 无需训练的尺度自适应高斯点云抗锯齿," arXiv:2403.19615, 2024。

[33] Z. Liang, Q. Zhang, W. Hu, L. Zhu, Y. Feng, and K. Jia, "Analytic-splatting: Anti-aliased 3d gaussian splatting via analytic integration," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 281-297.

Z. Liang, Q. Zhang, W. Hu, L. Zhu, Y. Feng, 和 K. Jia, "解析散点法: 通过解析积分实现的抗锯齿三维高斯散点," 载于欧洲计算机视觉会议, Springer 出版社, 2025 年, 第 281-297 页。

[34] Y. Jiang, J. Tu, Y. Liu, X. Gao, X. Long, W. Wang, and Y. Ma, "Gaussianshader: 3 d gaussian splatting with shading functions for reflective surfaces," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 5322-5332.

Y. Jiang, J. Tu, Y. Liu, X. Gao, X. Long, W. Wang, 和 Y. Ma, "GaussianShader: 用于反射表面的带阴影函数的高斯散点," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024 年, 第 5322-5332 页。

[35] J. Meng, H. Li, Y. Wu, Q. Gao, S. Yang, J. Zhang, and S. Ma, "Mirror-3dgs: Incorporating mirror reflections into 3d gaussian splatting," arXiv:2404.01168, 2024.

J. Meng, H. Li, Y. Wu, Q. Gao, S. Yang, J. Zhang, 和 S. Ma, "Mirror-3dgs: 将镜面反射纳入三维高斯散点," arXiv:2404.01168, 2024 年。

[36] Z. Yang, X. Gao, Y. Sun, Y. Huang, X. Lyu, W. Zhou, S. Jiao, X. Qi, and X. Jin, "Spec-gaussian: Anisotropic view-dependent appearance for 3d gaussian splatting," arXiv:2402.15870, 2024.

Z. Yang, X. Gao, Y. Sun, Y. Huang, X. Lyu, W. Zhou, S. Jiao, X. Qi, 和 X. Jin, "Spec-Gaussian: 用于三维高斯散点的各向异性视角依赖外观," arXiv:2402.15870, 2024 年。

[37] K. Ye, Q. Hou, and K. Zhou, "3d gaussian splatting with deferred reflection," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-10.

K. Ye, Q. Hou, 和 K. Zhou,"带延迟反射的三维高斯散点,"载于 ACM SIGGRAPH 会议论文集,2024 年,第 1-10 页。

[38] X. Wu, J. Xu, C. Wang, Y. Peng, Q. Huang, J. Tompkin, and W. Xu, "Local gaussian density mixtures for unstructured lumigraph rendering," in ACM SIGGRAPH Conf. Papers, ser. SA '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: https://doi.org/10.1145/3680528.3687659

X. Wu, J. Xu, C. Wang, Y. Peng, Q. Huang, J. Tompkin, 和 W. Xu,"用于无结构光场渲染的局部高斯密度混合模型,"载于 ACM SIGGRAPH 会议论文集,系列 SA '24。纽约,纽约州,美国: 计算机协会,2024 年。[在线]. 可访问:https://doi.org/10.1145/3680528.3687659

[39] J. Liu, X. Tang, F. Cheng, R. Yang, Z. Li, J. Liu, Y. Huang, J. Lin, S. Liu, X. Wu et al., "Mirrorgaussian: Reflecting 3d gaussians for reconstructing mirror reflections," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 377-393.

J. Liu, X. Tang, F. Cheng, R. Yang, Z. Li, J. Liu, Y. Huang, J. Lin, S. Liu, X. Wu 等,"MirrorGaussian: 用于重建镜面反射的反射三维高斯,"载于欧洲计算机视觉会议,Springer 出版社,2025 年,第 377-393 页。

[40] J. Oh, J. Chung, D. Lee, and K. M. Lee, "Deblurgs: Gaussian splatting for camera motion blur," arXiv:2404.11358, 2024.

J. Oh, J. Chung, D. Lee, 和 K. M. Lee,"DeblurGS: 用于相机运动模糊的高斯散点,"arXiv:2404.11358,2024 年。

[41] B. Lee, H. Lee, X. Sun, U. Ali, and E. Park, "Deblurring 3d gaussian splatting," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 127-143.

B. Lee, H. Lee, X. Sun, U. Ali, 和 E. Park,"三维高斯散点去模糊,"载于欧洲计算机视觉会议,Springer 出版社,2025 年,第 127-143 页。

[42] O. Seiskari, J. Ylilammi, V. Kaatrasalo, P. Rantalankila, M. Turkulainen, J. Kannala, E. Rahtu, and A. Solin, "Gaussian splatting on the move: Blur and rolling shutter compensation for natural camera motion," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 160-177.

O. Seiskari, J. Ylilammi, V. Kaatrasalo, P. Rantalankila, M. Turkulainen, J. Kannala, E. Rahtu, 和 A. Solin,"动态高斯散点: 自然相机运动的模糊与滚动快门补偿,"载于欧洲计算机视觉会议,Springer 出版社,2025 年,第 160-177 页。

[43] C. Peng, Y. Tang, Y. Zhou, N. Wang, X. Liu, D. Li, and R. Chellappa, "Bags: Blur agnostic gaussian splatting through multi-scale kernel modeling," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 293-310.

C. Peng, Y. Tang, Y. Zhou, N. Wang, X. Liu, D. Li, 和 R. Chellappa, "BAGS: 通过多尺度核建模实现的模糊无关高斯散点," 载于欧洲计算机视觉会议, Springer 出版社, 2025 年, 第 293-310 页。

[44] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance, "Microfacet models for refraction through rough surfaces," in Proc. 18th Eurographics Conf. Rendering Tech., 2007, pp. 195-206.

B. Walter, S. R. Marschner, H. Li, 和 K. E. Torrance, "粗糙表面折射的微面模型," 载于第 18 届欧洲图形学会议渲染技术, 2007 年, 第 195-206 页。

[45] S. Szymanowicz, C. Rupprecht, and A. Vedaldi, "Splatter image: Ultrafast single-view 3d reconstruction," arXiv:2312.13150, 2023.

S. Szymanowicz, C. Rupprecht, 和 A. Vedaldi, "Splatter Image: 超快速单视图三维重建," arXiv:2312.13150, 2023 年。

[46] D. Charatan, S. Li, A. Tagliasacchi, and V. Sitzmann, "Pixelsplat: 3d gaussian splats from image pairs for scalable generalizable 3 d reconstruction," arXiv:2312.12337, 2023.

D. Charatan, S. Li, A. Tagliasacchi, 和 V. Sitzmann, "PixelSplat: 基于图像对的三维高斯散点用于可扩展通用重建," arXiv:2312.12337, 2023 年。

[47] Y. Chen, H. Xu, C. Zheng, B. Zhuang, M. Pollefeys, A. Geiger, T.-J. Cham, and J. Cai, "Mvsplat: Efficient 3d gaussian splatting from sparse multi-view images," arXiv:2403.14627, 2024.

Y. Chen, H. Xu, C. Zheng, B. Zhuang, M. Pollefeys, A. Geiger, T.-J. Cham, 和 J. Cai, "Mvsplat: 基于稀疏多视图图像的高效三维高斯点绘制", arXiv:2403.14627, 2024。

[48] Y. Wang, T. Huang, H. Chen, and G. H. Lee, "Freesplat: Generalizable 3d gaussian splatting towards free-view synthesis of indoor scenes," arXiv:2405.17958, 2024.

Y. Wang, T. Huang, H. Chen, 和 G. H. Lee, "Freesplat: 面向室内场景自由视角合成的通用三维高斯点绘制", arXiv:2405.17958, 2024。

[49] Y. Bao, J. Liao, J. Huo, and Y. Gao, "Distractor-free generalizable 3d gaussian splatting," arXiv:2411.17605, 2024.

Y. Bao, J. Liao, J. Huo, 和 Y. Gao, "无干扰通用三维高斯点绘制", arXiv:2411.17605, 2024。

[50] Y. Chen, J. Wang, Z. Yang, S. Manivasagam, and R. Urtasun, "G3r: Gradient guided generalizable reconstruction," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 305-323.

Y. Chen, J. Wang, Z. Yang, S. Manivasagam, 和 R. Urtasun, "G3r: 梯度引导的通用重建", 发表于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer, 2025, 页 305-323。

[51] Z.-X. Zou, Z. Yu, Y.-C. Guo, Y. Li, D. Liang, Y.-P. Cao, and S.-H. Zhang, "Triplane meets gaussian splatting: Fast and generalizable single-view 3d reconstruction with transformers," arXiv:2312.09147, 2023.

Z.-X. Zou, Z. Yu, Y.-C. Guo, Y. Li, D. Liang, Y.-P. Cao, 和 S.-H. Zhang, "三平面遇见高斯点绘制: 基于变换器的快速通用单视图三维重建", arXiv:2312.09147, 2023。

[52] D. Xu, Y. Yuan, M. Mardani, S. Liu, J. Song, Z. Wang, and A. Vahdat, "Agg: Amortized generative 3 d gaussians for single image to 3 d ," arXiv:2401.04099, 2024.

D. Xu, Y. Yuan, M. Mardani, S. Liu, J. Song, Z. Wang, 和 A. Vahdat, "AGG: 用于单图像到 3 d 的摊销生成高斯模型", arXiv:2401.04099, 2024。

[53] K. Zhang, S. Bi, H. Tan, Y. Xiangli, N. Zhao, K. Sunkavalli, and Z. Xu, "Gs-lrm: Large reconstruction model for 3d gaussian splatting," Eur. Conf. Comput. Vis., 2024.

K. Zhang, S. Bi, H. Tan, Y. Xiangli, N. Zhao, K. Sunkavalli, 和 Z. Xu, "GS-LRM: 用于三维高斯点绘制的大型重建模型", 欧洲计算机视觉会议, 2024。

[54] Y. Xu, Z. Shi, W. Yifan, H. Chen, C. Yang, S. Peng, Y. Shen, and G. Wetzstein, "Grm: Large gaussian reconstruction model for efficient 3d reconstruction and generation," arXiv:2403.14621, 2024.

Y. Xu, Z. Shi, W. Yifan, H. Chen, C. Yang, S. Peng, Y. Shen, 和 G. Wetzstein, "GRM: 高效三维重建与生成的大型高斯重建模型", arXiv:2403.14621, 2024。

[55] S. Guo, Q. Wang, Y. Gao, R. Xie, L. Li, F. Zhu, and L. Song, "Depth-guided robust point cloud fusion nerf for sparse input views," IEEE Trans. Circuits Syst. Video Technol., 2024.

S. Guo, Q. Wang, Y. Gao, R. Xie, L. Li, F. Zhu, 和 L. Song, "基于深度引导的鲁棒点云融合 NeRF, 用于稀疏输入视图", IEEE 电路与系统视频技术汇刊, 2024。

[56] J. Chung, J. Oh, and K. M. Lee, "Depth-regularized optimization for 3d gaussian splatting in few-shot images," arXiv:2311.13398, 2023.

J. Chung, J. Oh, 和 K. M. Lee, "基于深度正则化的少样本图像三维高斯点绘制优化", arXiv:2311.13398, 2023。

[57] Z. Zhu, Z. Fan, Y. Jiang, and Z. Wang, "Fsgs: Real-time few-shot view synthesis using gaussian splatting," arXiv:2312.00451, 2023.

Z. Zhu, Z. Fan, Y. Jiang, 和 Z. Wang, "FSGS: 基于高斯点绘制的实时少样本视图合成", arXiv:2312.00451, 2023。

[58] A. Swann, M. Strong, W. K. Do, G. S. Camps, M. Schwager, and M. Kennedy III, "Touch-gs: Visual-tactile supervised 3d gaussian splatting," arXiv:2403.09875, 2024.

A. Swann, M. Strong, W. K. Do, G. S. Camps, M. Schwager, 和 M. Kennedy III, "Touch-GS: 视觉-触觉监督的三维高斯点绘制", arXiv:2403.09875, 2024。

[59] J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, and L. Gu, "Dngaussian: Optimizing sparse-view 3d gaussian radiance fields with global-local depth normalization," arXiv:2403.06912, 2024.

> J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, 和 L. Gu, "DNGaussian: 基于全局-局部深度归一化的稀疏视图三维高斯辐射场优化", arXiv:2403.06912, 2024。

[60] Z. Liu, J. Su, G. Cai, Y. Chen, B. Zeng, and Z. Wang, "Georgs: Geometric regularization for real-time novel view synthesis from sparse inputs," IEEE Trans. Circuits Syst. Video Technol., 2024.

> Z. Liu, J. Su, G. Cai, Y. Chen, B. Zeng, 和 Z. Wang, "GEORGS: 用于稀疏输入的实时新视角合成的几何正则化", IEEE 电路与系统视频技术汇刊, 2024。

[61] X. Liu, J. Chen, S.-H. Kao, Y.-W. Tai, and C.-K. Tang, "Deceptive-nerf/3dgs: Diffusion-generated pseudo-observations for high-quality sparse-view reconstruction," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 337-355.

> X. Liu, J. Chen, S.-H. Kao, Y.-W. Tai, 和 C.-K. Tang, "Deceptive-NeRF/3DGS: 基于扩散生成的伪观测用于高质量稀疏视图重建", 发表于欧洲计算机视觉会议, Springer, 2025, 页 337-355。

[62] A. Paliwal, W. Ye, J. Xiong, D. Kotovenko, R. Ranjan, V. Chandra, and N. K. Kalantari, "Coherentgs: Sparse novel view synthesis with coherent 3d gaussians," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 19-37.

> A. Paliwal, W. Ye, J. Xiong, D. Kotovenko, R. Ranjan, V. Chandra, 和 N. K. Kalantari, "Coherentgs: 基于相干三维高斯 (3D Gaussians) 的稀疏新视角合成," 发表于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer 出版社, 2025 年, 第 19-37 页。

[63] J. Zhang, J. Li, X. Yu, L. Huang, L. Gu, J. Zheng, and X. Bai, "Cor-gs: sparse-view 3d gaussian splatting via co-regularization," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 335-352.

> J. Zhang, J. Li, X. Yu, L. Huang, L. Gu, J. Zheng, 和 X. Bai, "Cor-gs: 通过协同正则化实现的稀疏视角三维高斯点喷溅 (3D Gaussian Splatting)," 发表于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer 出版社, 2025 年, 第 335-352 页。

[64] C. Yang, S. Li, J. Fang, R. Liang, L. Xie, X. Zhang, W. Shen, and Q. Tian, "Gaussianobject: Just taking four images to get a high-quality 3d object with gaussian splatting," arXiv:2402.10259, 2024.

> C. Yang, S. Li, J. Fang, R. Liang, L. Xie, X. Zhang, W. Shen, 和 Q. Tian, "Gaussianobject: 仅用四张图像即可通过高斯点喷溅获得高质量三维物体," arXiv:2402.10259, 2024 年。

[65] A. Moreau, J. Song, H. Dhamo, R. Shaw, Y. Zhou, and E. Pérez-Pellitero, "Human gaussian splatting: Real-time rendering of animatable avatars," arXiv:2311.17113, 2023.

> A. Moreau, J. Song, H. Dhamo, R. Shaw, Y. Zhou, 和 E. Pérez-Pellitero, "人体高斯点喷溅: 可动画头像的实时渲染," arXiv:2311.17113, 2023 年。

[66] S. Zheng, B. Zhou, R. Shao, B. Liu, S. Zhang, L. Nie, and Y. Liu, "GPS-Gaussian: Generalizable Pixel-wise 3D Gaussian Splatting for Real-time Human Novel View Synthesis," 2024.

> S. Zheng, B. Zhou, R. Shao, B. Liu, S. Zhang, L. Nie, 和 Y. Liu, "GPS-Gaussian: 用于实时人体新视角合成的可泛化像素级三维高斯点喷溅,"2024 年。

[67] M. Kocabas, J.-H. R. Chang, J. Gabriel, O. Tuzel, and A. Ranjan, "Hugs: Human gaussian splats," arXiv:2311.17910, 2023.

> M. Kocabas, J.-H. R. Chang, J. Gabriel, O. Tuzel, 和 A. Ranjan, "Hugs: 人体高斯点喷溅, " arXiv:2311.17910, 2023 年。

[68] L. Hu, H. Zhang, Y. Zhang, B. Zhou, B. Liu, S. Zhang, and L. Nie, "Gaussianavatar: Towards realistic human avatar modeling from a single video via animatable 3d gaussians," arXiv:2312.02134, 2023.

> L. Hu, H. Zhang, Y. Zhang, B. Zhou, B. Liu, S. Zhang, 和 L. Nie, "Gaussianavatar: 通过可动画三维高斯实现单视频逼真人体头像建模,"arXiv:2312.02134, 2023 年。

[69] S. Hu and Z. Liu, "Gauhuman: Articulated gaussian splatting from monocular human videos," arXiv:2312.02973, 2023.

> S. Hu 和 Z. Liu, "Gauhuman: 基于单目人体视频的关节高斯点喷溅,"arXiv:2312.02973, 2023 年。

[70] Z. Qian, S. Wang, M. Mihajlovic, A. Geiger, and S. Tang, "3dgs-avatar: Animatable avatars via deformable 3d gaussian splatting," arXiv:2312.09228, 2023.

> Z. Qian, S. Wang, M. Mihajlovic, A. Geiger, 和 S. Tang, "3dgs-avatar: 通过可变形三维高斯点喷溅实现的可动画头像,"arXiv:2312.09228, 2023 年。

[71] H. Pang, H. Zhu, A. Kortylewski, C. Theobalt, and M. Habermann, "Ash: Animatable gaussian splats for efficient and photoreal human rendering," arXiv:2312.05941, 2023.

> H. Pang, H. Zhu, A. Kortylewski, C. Theobalt, 和 M. Habermann, "Ash: 用于高效且逼真人体渲染的可动画高斯点喷溅,"arXiv:2312.05941, 2023 年。

[72] Z. Li, Z. Zheng, L. Wang, and Y. Liu, "Animatable gaussians: Learning pose-dependent gaussian maps for high-fidelity human avatar modeling," arXiv:2311.16096, 2023.

> Z. Li, Z. Zheng, L. Wang, 和 Y. Liu, "Animatable gaussians: 学习基于姿态的高斯映射以实现高保真人体头像建模,"arXiv:2311.16096, 2023 年。

[73] Z. Sheng, F. Liu, M. Liu, F. Zheng, and L. Nie, "Open-set synthesis for free-viewpoint human body reenactment of novel poses," IEEE Trans. Circuits Syst. Video Technol., 2024.

> Z. Sheng, F. Liu, M. Liu, F. Zheng, 和 L. Nie, "开放集合成用于新姿态的人体自由视角重演,"IEEE 电路与系统视频技术汇刊 (IEEE Trans. Circuits Syst. Video Technol.), 2024 年。

[74] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: A skinned multi-person linear model," ACM Trans. Graph. (Proc. SIGGRAPH Asia), vol. 34, no. 6, pp. 248:1-248:16, oct 2015.

M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, 和 M. J. Black, "SMPL: 一种蒙皮多人体线性模型," ACM 图形学汇刊 (ACM Trans. Graph., SIGGRAPH Asia 会议论文集), 第 34 卷第 6 期, 第 248:1-248:16 页, 2015 年 10 月。

[75] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, and M. J. Black, "Expressive body capture: 3D hands, face, and body from a single image," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 10975-10985.

G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, 和 M. J. Black, "表现力强的身体捕捉: 基于单张图像的三维手部、面部和身体," 发表于 IEEE 计算机视觉与模式识别会议 (Proc. IEEE Conf. Comput. Vis. Pattern Recognit.), 2019 年, 第 10975-10985 页。

[76] Y. Chen, L. Wang, Q. Li, H. Xiao, S. Zhang, H. Yao, and Y. Liu, "Monogaussianavatar: Monocular gaussian point-based head avatar," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-9.

Y. Chen, L. Wang, Q. Li, H. Xiao, S. Zhang, H. Yao, 和 Y. Liu, "Monogaussianavatar: 基于单目高斯点的头部头像," 发表于 ACM SIGGRAPH 会议论文集, 2024 年, 第 1-9 页。

[77] S. Qian, T. Kirschstein, L. Schoneveld, D. Davoli, S. Giebenhain, and M. Nießner, "Gaussianavatars: Photorealistic head avatars with rigged 3d gaussians," arXiv:2312.02069, 2023.

S. Qian, T. Kirschstein, L. Schoneveld, D. Davoli, S. Giebenhain, 和 M. Nießner, "Gaussianavatars: 具有绑定三维高斯 (3D Gaussians) 模型的照片级真实感头部头像," arXiv:2312.02069, 2023.

[78] Z. Zhao, Z. Bao, Q. Li, G. Qiu, and K. Liu, "Psavatar: A point-based morphable shape model for real-time head avatar creation with 3 d gaussian splatting," arXiv:2401.12900, 2024.

Z. Zhao, Z. Bao, Q. Li, G. Qiu, 和 K. Liu, "Psavatar: 一种基于点的可变形形状模型, 用于实时头部头像创建, 结合 3 d 高斯溅射 (Gaussian Splatting)," arXiv:2401.12900, 2024.

[79] H. Dhamo, Y. Nie, A. Moreau, J. Song, R. Shaw, Y. Zhou, and E. Pérez-Pellitero, "Headgas: Real-time animatable head avatars via 3d gaussian splatting," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 459-476.

H. Dhamo, Y. Nie, A. Moreau, J. Song, R. Shaw, Y. Zhou, 和 E. Pérez-Pellitero, "Headgas: 通过三维高斯溅射实现的实时可动画头部头像," 载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer, 2025, 页 459-476.

[80] K. Teotia, H. Kim, P. Garrido, M. Habermann, M. Elgharib, and C. Theobalt, "Gaussianheads: End-to-end learning of drivable gaussian head avatars from coarse-to-fine representations," ACM Trans. Graph., vol. 43, no. 6, pp. 1-12, 2024.

K. Teotia, H. Kim, P. Garrido, M. Habermann, M. Elgharib, 和 C. Theobalt,"Gaussianheads: 从粗到细表示的端到端可驱动高斯头部头像学习,"ACM 图形学汇刊 (ACM Trans. Graph.), 第 43 卷第 6 期, 页 1-12, 2024.

[81] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero, "Learning a model of facial shape and expression from 4D scans," ACM Trans. Graph. (Proc. SIGGRAPH Asia), vol. 36, no. 6, pp. 194:1-194:17, 2017. [Online]. Available: https://doi.org/10.1145/3130800.3130813

T. Li, T. Bolkart, M. J. Black, H. Li, 和 J. Romero, "从 4D 扫描学习面部形状与表情模型,"ACM 图形学汇刊 (ACM Trans. Graph.)(SIGGRAPH Asia 会议论文集), 第 36 卷第 6 期, 页 194:1-194:17, 2017. [在线]. 可访问: https://doi.org/10.1145/3130800.3130813

[82] Y. Xu, B. Chen, Z. Li, H. Zhang, L. Wang, Z. Zheng, and Y. Liu, "Gaussian head avatar: Ultra high-fidelity head avatar via dynamic gaussians," arXiv:2312.03029, 2023.

Y. Xu, B. Chen, Z. Li, H. Zhang, L. Wang, Z. Zheng, 和 Y. Liu, "Gaussian 头部头像: 通过动态高斯实现的超高保真头部头像,"arXiv:2312.03029, 2023.

[83] ——, "Gaussian head avatar: Ultra high-fidelity head avatar via dynamic gaussians," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 1931-1941.

——, "Gaussian 头部头像: 通过动态高斯实现的超高保真头部头像,"载于 IEEE/CVF 计算机视觉与模式识别会议 (Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.), 2024, 页 1931-1941.

[84] Y. Xu, L. Wang, Z. Zheng, Z. Su, and Y. Liu, "3d gaussian parametric head model," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 129-147.

Y. Xu, L. Wang, Z. Zheng, Z. Su, 和 Y. Liu, "三维高斯参数化头部模型,"载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.),Springer, 2025, 页 129-147.

[85] S. Ma, Y. Weng, T. Shao, and K. Zhou, "3d gaussian blendshapes for head avatar animation," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-10 .

S. Ma, Y. Weng, T. Shao, 和 K. Zhou, "用于头部头像动画的三维高斯混合形状 (Blendshapes),"载于 ACM SIGGRAPH 会议论文集, 2024, 页 1-10.

[86] J. Xiang, X. Gao, Y. Guo, and J. Zhang, "Flashavatar: High-fidelity head avatar with efficient gaussian embedding," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 1802-1812.

J. Xiang, X. Gao, Y. Guo, 和 J. Zhang, "Flashavatar: 具有高效高斯嵌入的高保真头部头像,"载于 IEEE/CVF 计算机视觉与模式识别会议 (Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.), 2024, 页 1802-1812.

[87] J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, and L. Gu, "Talkinggaussian: Structure-persistent 3d talking head synthesis via gaussian splatting," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 127-145.

J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, 和 L. Gu, "Talkinggaussian: 通过高斯溅射实现结构持久的三维会说话头部合成," 载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer 出版社, 2025 年, 第 127-145 页。

[88] K. Cho, J. Lee, H. Yoon, Y. Hong, J. Ko, S. Ahn, and S. Kim, "Gaussiantalker: Real-time talking head synthesis with 3 d gaussian splatting," in Proc. ACM Int. Conf. Multimedia, 2024, pp. 10985- 10994.

K. Cho, J. Lee, H. Yoon, Y. Hong, J. Ko, S. Ahn, 和 S. Kim, "Gaussiantalker: 基于 3 d 高斯溅射的实时会说话头部合成," 载于 ACM 国际多媒体会议论文集 (Proc. ACM Int. Conf. Multimedia), 2024 年, 第 10985-10994 页。

[89] H. Yu, Z. Qu, Q. Yu, J. Chen, Z. Jiang, Z. Chen, S. Zhang, J. Xu, F. Wu, C. Lv et al., "Gaussiantalker: Speaker-specific talking head synthesis via 3d gaussian splatting," in Proc. ACM Int. Conf. Multimedia, 2024, pp. 3548-3557.

H. Yu, Z. Qu, Q. Yu, J. Chen, Z. Jiang, Z. Chen, S. Zhang, J. Xu, F. Wu, C. Lv 等, "Gaussiantalker: 通过三维高斯溅射实现说话人特定的会说话头部合成," 载于 ACM 国际多媒体会议论文集 (Proc. ACM Int. Conf. Multimedia), 2024 年, 第 3548-3557 页。

[90] H. Luo, M. Ouyang, Z. Zhao, S. Jiang, L. Zhang, Q. Zhang, W. Yang, L. Xu, and J. Yu, "Gaussianhair: Hair modeling and rendering with light-aware gaussians," arXiv:2402.10483, 2024.

H. Luo, M. Ouyang, Z. Zhao, S. Jiang, L. Zhang, Q. Zhang, W. Yang, L. Xu, 和 J. Yu, "Gaussianhair: 基于光感知高斯的头发建模与渲染," arXiv:2402.10483, 2024 年。

[91] L. Bolanos, S.-Y. Su, and H. Rhodin, "Gaussian Shadow Casting for Neural Characters," 2024.

L. Bolanos, S.-Y. Su, 和 H. Rhodin, "神经角色的高斯阴影投射," 2024 年。

[92] R. Abdal, W. Yifan, Z. Shi, Y. Xu, R. Po, Z. Kuang, Q. Chen, D.-Y. Yeung, and G. Wetzstein, "Gaussian shell maps for efficient 3d human generation," arXiv:2311.17857, 2023.

R. Abdal, W. Yifan, Z. Shi, Y. Xu, R. Po, Z. Kuang, Q. Chen, D.-Y. Yeung, 和 G. Wetzstein, "高斯壳映射用于高效三维人体生成," arXiv:2311.17857, 2023 年。

[93] X. Liu, X. Zhan, J. Tang, Y. Shan, G. Zeng, D. Lin, X. Liu, and Z. Liu, "Humangaussian: Text-driven 3d human generation with gaussian splatting," arXiv:2311.17061, 2023.

X. Liu, X. Zhan, J. Tang, Y. Shan, G. Zeng, D. Lin, X. Liu, 和 Z. Liu, "Humangaussian: 基于文本驱动的三维人体生成与高斯溅射," arXiv:2311.17061, 2023 年。

[94] T. Kirschstein, S. Giebenhain, J. Tang, M. Georgopoulos, and M. Nießner, "Gghead: Fast and generalizable 3d gaussian heads," in ACM SIGGRAPH Asia Conf. Papers, 2024, pp. 1-11.

T. Kirschstein, S. Giebenhain, J. Tang, M. Georgopoulos, 和 M. Nießner, "Gghead: 快速且通用的三维高斯头部," 载于 ACM SIGGRAPH Asia 会议论文集, 2024 年, 第 1-11 页。

[95] J. Tang, J. Ren, H. Zhou, Z. Liu, and G. Zeng, "Dreamgaus-sian: Generative gaussian splatting for efficient 3d content creation," arXiv:2309.16653, 2023.

J. Tang, J. Ren, H. Zhou, Z. Liu, 和 G. Zeng, "Dreamgaussian: 用于高效三维内容创作的生成式高斯溅射," arXiv:2309.16653, 2023 年。

[96] J. Ren, L. Pan, J. Tang, C. Zhang, A. Cao, G. Zeng, and Z. Liu, "Dreamgaussian4d: Generative 4d gaussian splatting," arXiv:2312.17142, 2023.

J. Ren, L. Pan, J. Tang, C. Zhang, A. Cao, G. Zeng, 和 Z. Liu, "Dreamgaussian4d: 生成式四维高斯溅射," arXiv:2312.17142, 2023 年。

[97] J. Zhang, Z. Tang, Y. Pang, X. Cheng, P. Jin, Y. Wei, W. Yu, M. Ning, and L. Yuan, "Repaint123: Fast and high-quality one image to 3d generation with progressive controllable 2d repainting," arXiv:2312.13271, 2023.

J. Zhang, Z. Tang, Y. Pang, X. Cheng, P. Jin, Y. Wei, W. Yu, M. Ning, 和 L. Yuan, "Repaint123: 通过渐进可控的二维重绘实现快速高质量单图到三维生成," arXiv:2312.13271, 2023 年。

[98] X. Zhou, X. Ran, Y. Xiong, J. He, Z. Lin, Y. Wang, D. Sun, and M.-H. Yang, "Gala3d: Towards text-to-3d complex scene generation via layout-guided generative gaussian splatting," arXiv:2402.07207, 2024.

X. Zhou, X. Ran, Y. Xiong, J. He, Z. Lin, Y. Wang, D. Sun, 和 M.-H. Yang, "Gala3d: 通过布局引导的生成式高斯溅射实现文本到三维复杂场景生成," arXiv:2402.07207, 2024 年。

[99] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "Dreamfusion: Text-to-3d using 2d diffusion," arXiv:2209.14988, 2022.

B. Poole, A. Jain, J. T. Barron, 和 B. Mildenhall, "Dreamfusion: 利用二维扩散实现文本到三维," arXiv:2209.14988, 2022 年。

[100] Z. Chen, F. Wang, and H. Liu, "Text-to-3d using gaussian splatting," arXiv:2309.16585, 2023.

Z. Chen, F. Wang, 和 H. Liu, "基于高斯溅射的文本到三维," arXiv:2309.16585, 2023 年。

[101] T. Yi, J. Fang, J. Wang, G. Wu, L. Xie, X. Zhang, W. Liu, Q. Tian, and X. Wang, "Gaussiandreamer: Fast generation from text to 3d gaussians by bridging 2 d and 3 d diffusion models," arXiv preprint arXiv, vol. 2310, 2023.

T. Yi, J. Fang, J. Wang, G. Wu, L. Xie, X. Zhang, W. Liu, Q. Tian, 和 X. Wang, "Gaussiandreamer: 通过桥接 2 d 和 3 d 扩散模型实现从文本到 3D 高斯的快速生成," arXiv 预印本 arXiv, 卷 2310, 2023.

[102] Q. Shen, X. Yang, M. B. Mi, and X. Wang, "Vista3d: Unravel the 3d darkside of a single image," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 405-421.

Q. Shen, X. Yang, M. B. Mi, 和 X. Wang, "Vista3d: 揭示单幅图像的 3D 暗面," 载于欧洲计算机视觉会议，Springer, 2025, 页 405-421.

[103] X. Li, H. Wang, and K.-K. Tseng, "Gaussiandiffusion: 3d gaussian splatting for denoising diffusion probabilistic models with structured noise," arXiv:2311.11221, 2023.

X. Li, H. Wang, 和 K.-K. Tseng, "Gaussiandiffusion: 用于带结构噪声的去噪扩散概率模型的 3D 高斯点渲染," arXiv:2311.11221, 2023.

[104] Y. Liang, X. Yang, J. Lin, H. Li, X. Xu, and Y. Chen, "Luciddreamer: Towards high-fidelity text-to-3d generation via interval score matching," arXiv:2311.11284, 2023.

Y. Liang, X. Yang, J. Lin, H. Li, X. Xu, 和 Y. Chen, "Luciddreamer: 通过区间评分匹配实现高保真文本到 3D 生成," arXiv:2311.11284, 2023.

[105] D. Di, J. Yang, C. Luo, Z. Xue, W. Chen, X. Yang, and Y. Gao, "Hyper- 3dg: Text-to-3d gaussian generation via hypergraph," arXiv:2403.09236, 2024.

D. Di, J. Yang, C. Luo, Z. Xue, W. Chen, X. Yang, 和 Y. Gao, "Hyper-3dg: 通过超图实现文本到 3D 高斯生成," arXiv:2403.09236, 2024.

[106] X. Yang, Y. Chen, C. Chen, C. Zhang, Y. Xu, X. Yang, F. Liu, and G. Lin,"Learn to optimize de-noising scores for 3 d generation: A unified and improved diffusion prior on nerf and 3d gaussian splatting," arXiv:2312.04820, 2023.

X. Yang, Y. Chen, C. Chen, C. Zhang, Y. Xu, X. Yang, F. Liu, 和 G. Lin, " 学习优化去噪评分以生成 3 d :NeRF 和 3D 高斯点渲染上的统一且改进的扩散先验," arXiv:2312.04820, 2023.

[107] W. Zhuo, F. Ma, H. Fan, and Y. Yang, "Vividdreamer: Invariant score distillation for hyper-realistic text-to-3d generation," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 122-139.

W. Zhuo, F. Ma, H. Fan, 和 Y. Yang, "Vividdreamer: 用于超真实文本到 3D 生成的不变评分蒸馏," 载于欧洲计算机视觉会议，Springer, 2025, 页 122-139.

[108] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," arXiv:2010.02502, 2020.

J. Song, C. Meng, 和 S. Ermon, " 去噪扩散隐式模型," arXiv:2010.02502, 2020.

[109] L. Melas-Kyriazi, I. Laina, C. Rupprecht, N. Neverova, A. Vedaldi, O. Gafni, and F. Kokkinos, "Im-3d: Iterative multiview diffusion and reconstruction for high-quality 3 d generation," arXiv:2402.08682,2024.

L. Melas-Kyriazi, I. Laina, C. Rupprecht, N. Neverova, A. Vedaldi, O. Gafni, 和 F. Kokkinos, "Im-3d: 用于高质量 3 d 生成的迭代多视角扩散与重建," arXiv:2402.08682, 2024.

[110] V. Voleti, C.-H. Yao, M. Boss, A. Letts, D. Pankratz, D. Tochilkin, C. Laforte, R. Rombach, and V. Jampani, "Sv3d: Novel multi-view synthesis and 3 d generation from a single image using latent video diffusion," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 439-457.

V. Voleti, C.-H. Yao, M. Boss, A. Letts, D. Pankratz, D. Tochilkin, C. Laforte, R. Rombach, 和 V. Jampani, "Sv3d: 利用潜在视频扩散从单幅图像实现新颖多视角合成和 3 d 生成," 载于欧洲计算机视觉会议, Springer, 2025, 页 439-457.

[111] R. Gao, A. Holynski, P. Henzler, A. Brussee, R. Martin-Brualla, P. Srinivasan, J. T. Barron, and B. Poole, "Cat3d: Create anything in 3d with multi-view diffusion models," arXiv:2405.10314, 2024.

R. Gao, A. Holynski, P. Henzler, A. Brussee, R. Martin-Brualla, P. Srinivasan, J. T. Barron, 和 B. Poole, "Cat3d: 使用多视角扩散模型创建任意 3D 内容," arXiv:2405.10314, 2024.

[112] J. Han, F. Kokkinos, and P. Torr, "Vfusion3d: Learning scalable 3d generative models from video diffusion models," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 333-350.

J. Han, F. Kokkinos, 和 P. Torr, "Vfusion3d: 从视频扩散模型学习可扩展的 3D 生成模型," 载于欧洲计算机视觉会议, Springer, 2025, 页 333-350.

[113] H. Yang, Y. Chen, Y. Pan, T. Yao, Z. Chen, C.-W. Ngo, and T. Mei, "Hi3d: Pursuing high-resolution image-to-3d generation with video diffusion models," in Proc. ACM Int. Conf. Multimedia, 2024, pp. $6870-6879$.

H. Yang, Y. Chen, Y. Pan, T. Yao, Z. Chen, C.-W. Ngo, 和 T. Mei, "Hi3d: 利用视频扩散模型追求高分辨率图像到 3D 生成," 载于 ACM 国际多媒体会议论文集, 2024, 页 $6870-6879$.

[114] L. Jiang and L. Wang, "Brightdreamer: Generic 3d gaussian generative framework for fast text-to-3d synthesis," arXiv:2403.11273, 2024.

L. Jiang 和 L. Wang, "Brightdreamer: 用于快速文本到三维合成的通用三维高斯生成框架," arXiv:2403.11273, 2024.

[115] J. Tang, Z. Chen, X. Chen, T. Wang, G. Zeng, and Z. Liu, "Lgm: Large multi-view gaussian model for high-resolution 3d content creation," arXiv:2402.05054, 2024.

J. Tang, Z. Chen, X. Chen, T. Wang, G. Zeng 和 Z. Liu, "Lgm: 用于高分辨率三维内容创作的大型多视角高斯模型," arXiv:2402.05054, 2024.

[116] B. Zhang, Y. Cheng, J. Yang, C. Wang, F. Zhao, Y. Tang, D. Chen, and B. Guo, "Gaussiancube: Structuring gaussian splatting using optimal transport for 3d generative modeling," arXiv:2403.19655, 2024.

B. Zhang, Y. Cheng, J. Yang, C. Wang, F. Zhao, Y. Tang, D. Chen 和 B. Guo, "Gaussiancube: 利用最优传输结构化高斯散点用于三维生成建模," arXiv:2403.19655, 2024.

[117] F.-L. Liu, H. Fu, Y.-K. Lai, and L. Gao, "Sketchdream: Sketch-based text-to-3d generation and editing," ACM Trans. Graph., vol. 43, no. 4, pp. 1-13, 2024.

F.-L. Liu, H. Fu, Y.-K. Lai 和 L. Gao, "Sketchdream: 基于草图的文本到三维生成与编辑," ACM Trans. Graph., 第 43 卷, 第 4 期, 第 1-13 页, 2024.

[118] X. He, J. Chen, S. Peng, D. Huang, Y. Li, X. Huang, C. Yuan, W. Ouyang, and T. He, "Gvgen: Text-to-3d generation with volumetric representation," arXiv:2403.12957, 2024.

X. He, J. Chen, S. Peng, D. Huang, Y. Li, X. Huang, C. Yuan, W. Ouyang 和 T. He, "Gvgen: 基于体积表示的文本到三维生成," arXiv:2403.12957, 2024.

[119] L. Lu, H. Gao, T. Dai, Y. Zha, Z. Hou, J. Wu, and S.-T. Xia, "Large point-to-gaussian model for image-to-3d generation," in Proc. ACM Int. Conf. Multimedia, 2024, pp. 10 843-10 852.

L. Lu, H. Gao, T. Dai, Y. Zha, Z. Hou, J. Wu 和 S.-T. Xia, "大型点到高斯模型用于图像到三维生成," 见 ACM 国际多媒体会议论文集, 2024, 第 10843-10852 页.

[120] B. Roessle, N. Müller, L. Porzi, S. Rota Bulò, P. Kontschieder, A. Dai, and M. Nießner, "L3dg: Latent 3d gaussian diffusion," in ACM SIGGRAPH Asia Conf. Papers, 2024, pp. 1-11.

B. Roessle, N. Müller, L. Porzi, S. Rota Bulò, P. Kontschieder, A. Dai 和 M. Nießner, "L3dg: 潜在三维高斯扩散," 见 ACM SIGGRAPH Asia 会议论文集, 2024, 第 1-11 页.

[121] H. Yu, W. Gong, J. Chen, and H. Ma, "Get3dgs: Generate 3d gaussians based on points deformation fields," IEEE Trans. Circuits Syst. Video Technol., 2024.

H. Yu, W. Gong, J. Chen 和 H. Ma, "Get3dgs: 基于点变形场生成三维高斯," IEEE Trans. Circuits Syst. Video Technol., 2024.

[122] Y. Yuan, X. Li, Y. Huang, S. De Mello, K. Nagano, J. Kautz, and U. Iqbal, "Gavatar: Animatable 3d gaussian avatars with implicit mesh learning," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 896-905.

Y. Yuan, X. Li, Y. Huang, S. De Mello, K. Nagano, J. Kautz 和 U. Iqbal, "Gavatar: 具有隐式网格学习的可动画三维高斯头像," 见 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 第 896-905 页.

[123] Q. Feng, Z. Xing, Z. Wu, and Y.-G. Jiang, "Fdgaussian: Fast gaussian splatting from single image via geometric-aware diffusion model," arXiv:2403.10242, 2024.

Q. Feng, Z. Xing, Z. Wu 和 Y.-G. Jiang, "Fdgaussian: 基于几何感知扩散模型的单图像快速高斯散点," arXiv:2403.10242, 2024.

[124] Y. Chen, J. Fang, Y. Huang, T. Yi, X. Zhang, L. Xie, X. Wang, W. Dai, H. Xiong, and Q. Tian, "Cascade-zero123: One image to highly consistent 3 d with self-prompted nearby views," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 311-330.

Y. Chen, J. Fang, Y. Huang, T. Yi, X. Zhang, L. Xie, X. Wang, W. Dai, H. Xiong 和 Q. Tian, "Cascade-zero123: 单图像到高度一致的 3 d，结合自提示的邻近视图," 见欧洲计算机视觉会议, Springer, 2025, 第 311-330 页.

[125] R. Liu, R. Wu, B. Van Hoorick, P. Tokmakov, S. Zakharov, and C. Vondrick, "Zero-1-to-3: Zero-shot one image to 3d object," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2023, pp. 9298-9309.

R. Liu, R. Wu, B. Van Hoorick, P. Tokmakov, S. Zakharov 和 C. Vondrick，"Zero-1-to-3: 零样本单图像到三维对象，"见 IEEE/CVF 国际计算机视觉会议论文集, 2023, 第 9298-9309 页.

[126] A. Vilesov, P. Chari, and A. Kadambi, "Cg3d: Compositional generation for text-to-3d via gaussian splatting," arXiv:2311.17907, 2023.

A. Vilesov, P. Chari 和 A. Kadambi， "Cg3d: 通过高斯散点实现文本到三维的组合生成，" arXiv:2311.17907, 2023.

[127] H. Li, H. Shi, W. Zhang, W. Wu, Y. Liao, L. Wang, L.-h. Lee, and P. Y. Zhou, "Dreamscene: 3d gaussian-based text-to-3d scene generation via formation pattern sampling," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 214-230.

H. Li, H. Shi, W. Zhang, W. Wu, Y. Liao, L. Wang, L.-h. Lee 和 P. Y. Zhou， "Dreamscene: 基于三维高斯的文本到三维场景生成，通过形态模式采样，"见欧洲计算机视觉会议，Springer, 2025, 第 214-230 页.

[128] Y. Chen, T. Wang, T. Wu, X. Pan, K. Jia, and Z. Liu, "Comboverse: Compositional 3d assets creation using spatially-aware diffusion guidance," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 128-146.

Y. Chen, T. Wang, T. Wu, X. Pan, K. Jia, 和 Z. Liu, "Comboverse: 使用空间感知扩散引导的组合式三维资产创建，"载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.)，Springer 出版社，2025 年，第 128-146 页。

[129] J. Chung, S. Lee, H. Nam, J. Lee, and K. M. Lee, "Lucid-dreamer: Domain-free generation of 3 d gaussian splatting scenes," arXiv:2311.13384, 2023.

J. Chung, S. Lee, H. Nam, J. Lee, 和 K. M. Lee, "Lucid-dreamer: 无领域限制的 3 d 高斯点扩散场景生成，" arXiv:2311.13384, 2023 年。

[130] H. Ouyang, K. Heal, S. Lombardi, and T. Sun, "Text2immersion: Generative immersive scene with 3d gaussians," arXiv:2312.09242, 2023.

H. Ouyang, K. Heal, S. Lombardi, 和 T. Sun, "Text2immersion: 基于三维高斯体的生成式沉浸式场景，" arXiv:2312.09242, 2023 年。

[131] S. Zhou, Z. Fan, D. Xu, H. Chang, P. Chari, T. Bharadwaj, S. You, Z. Wang, and A. Kadambi, "Dream-scene360: Unconstrained text-to- 3d scene generation with panoramic gaussian splatting," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 324-342.

S. Zhou, Z. Fan, D. Xu, H. Chang, P. Chari, T. Bharadwaj, S. You, Z. Wang, 和 A. Kadambi, "Dream-scene360: 使用全景高斯点扩散的无约束文本到三维场景生成，" 载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.)，Springer 出版社，2025 年，第 324-342 页。

[132] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, "Repaint: Inpainting using denoising diffusion probabilistic models," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp.

11461-11471.

> A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, 和 L. Van Gool, "Repaint: 利用去噪扩散概率模型进行图像修复," 载于 IEEE/CVF 计算机视觉与模式识别会议 (Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.)，2022 年，第 11461-11471 页。

[133] H. Ling, S. W. Kim, A. Torralba, S. Fidler, and K. Kreis, "Align your gaussians: Text-to-4d with dynamic 3d gaussians and composed diffusion models," arXiv:2312.13763, 2023.

> H. Ling, S. W. Kim, A. Torralba, S. Fidler, 和 K. Kreis, " 对齐你的高斯体: 结合动态三维高斯体与组合扩散模型的文本到四维生成," arXiv:2312.13763, 2023 年。

[134] Q. Gao, Q. Xu, Z. Cao, B. Mildenhall, W. Ma, L. Chen, D. Tang, and U. Neumann, "Gaussianflow: Splatting gaussian dynamics for 4d content creation," arXiv:2403.12365, 2024.

> Q. Gao, Q. Xu, Z. Cao, B. Mildenhall, W. Ma, L. Chen, D. Tang, 和 U. Neumann, "Gaussianflow: 用于四维内容创作的高斯动力学点扩散," arXiv:2403.12365, 2024 年。

[135] S. Bahmani, I. Skorokhodov, V. Rong, G. Wetzstein, L. Guibas, P. Wonka, S. Tulyakov, J. J. Park, A. Tagliasacchi, and D. B. Lindell, "4d-fy: Text-to-4d generation using hybrid score distillation sampling," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 7996-8006.

> S. Bahmani, I. Skorokhodov, V. Rong, G. Wetzstein, L. Guibas, P. Wonka, S. Tulyakov, J. J. Park, A. Tagliasacchi, 和 D. B. Lindell, "4d-fy: 使用混合评分蒸馏采样的文本到四维生成," 载于 IEEE/CVF 计算机视觉与模式识别会议 (Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.)，2024 年，第 7996-8006 页。

[136] S. Bahmani, X. Liu, W. Yifan, I. Skorokhodov, V. Rong, Z. Liu, X. Liu, J. J. Park, S. Tulyakov, G. Wetzstein et al., "Tc4d: Trajectory-conditioned text-to-4d generation," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 53-72.

> S. Bahmani, X. Liu, W. Yifan, I. Skorokhodov, V. Rong, Z. Liu, X. Liu, J. J. Park, S. Tulyakov, G. Wetzstein 等, "Tc4d: 基于轨迹条件的文本到四维生成," 载于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer 出版社，2025 年，第 53-72 页。

[137] Y. Zheng, X. Li, K. Nagano, S. Liu, O. Hilliges, and S. De Mello, "A unified approach for text-and image-guided 4d scene generation," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 7300 − 7309.

> Y. Zheng, X. Li, K. Nagano, S. Liu, O. Hilliges, 和 S. De Mello, " 文本与图像引导的四维场景生成统一方法," 载于 IEEE/CVF 计算机视觉与模式识别会议 (Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.)，2024 年，第 7300 − 7309 页。

[138] Y. Shi, P. Wang, J. Ye, M. Long, K. Li, and X. Yang, "Mvdream: Multi-view diffusion for 3d generation," arXiv:2308.16512, 2023.

Y. Shi, P. Wang, J. Ye, M. Long, K. Li, 和 X. Yang, "Mvdream: 用于三维生成的多视角扩散," arXiv:2308.16512, 2023 年。

[139] Y. Yin, D. Xu, Z. Wang, Y. Zhao, and Y. Wei, "4dgen: Grounded 4d content generation with spatial-temporal consistency," arXiv:2312.17225, 2023.

Y. Yin, D. Xu, Z. Wang, Y. Zhao, 和 Y. Wei, "4dgen: 具有时空一致性的基于地面真实数据的四维内容生成," arXiv:2312.17225, 2023 年。

[140] Z. Pan, Z. Yang, X. Zhu, and L. Zhang, "Fast dynamic 3d object generation from a single-view video," arXiv:2401.08742, 2024.

Z. Pan, Z. Yang, X. Zhu, 和 L. Zhang, " 基于单视角视频的快速动态三维物体生成," arXiv:2401.08742, 2024 年。

[141] Q. Sun, Z. Guo, Z. Wan, J. N. Yan, S. Yin, W. Zhou, J. Liao, and H. Li, "Eg4d: Explicit generation of 4d object without score distillation," arXiv:2405.18132, 2024.

Q. Sun, Z. Guo, Z. Wan, J. N. Yan, S. Yin, W. Zhou, J. Liao, and H. Li, "Eg4d: 无需分数蒸馏的显式 4D 对象生成," arXiv:2405.18132, 2024.

[142] Y. Zeng, Y. Jiang, S. Zhu, Y. Lu, Y. Lin, H. Zhu, W. Hu, X. Cao, and Y. Yao, "Stag4d: Spatial-temporal anchored generative 4d gaussians," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 163-179.

Y. Zeng, Y. Jiang, S. Zhu, Y. Lu, Y. Lin, H. Zhu, W. Hu, X. Cao, and Y. Yao, "Stag4d: 时空锚定生成 4D 高斯模型," 载于欧洲计算机视觉会议 (Springer), 2025, 页 163-179.

[143] A. Cao and J. Johnson, "Hexplane: A fast representation for dynamic scenes," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2023, pp. 130-141.

A. Cao and J. Johnson, "Hexplane: 动态场景的快速表示方法," 载于 IEEE/CVF 计算机视觉与模式识别会议, 2023, 页 130-141.

[144] T. Zhang, Q. Gao, W. Li, L. Liu, and B. Chen, "Bags: Building animatable gaussian splatting from a monocular video with diffusion priors," arXiv:2403.11427, 2024.

T. Zhang, Q. Gao, W. Li, L. Liu, and B. Chen, "Bags: 基于扩散先验从单目视频构建可动画高斯散点," arXiv:2403.11427, 2024.

[145] Z. Wu, C. Yu, Y. Jiang, C. Cao, F. Wang, and X. Bai, "Sc4d: Sparse-controlled video-to-4d generation and motion transfer," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 361-379.

Z. Wu, C. Yu, Y. Jiang, C. Cao, F. Wang, and X. Bai, "Sc4d: 稀疏控制的视频到 4D 生成及运动迁移," 载于欧洲计算机视觉会议 (Springer), 2025, 页 361-379.

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

> X. Zhou, Z. Lin, X. Shan, Y. Wang, D. Sun, and M.-H. Yang, "Drivinggaussian: 用于周边动态自动驾驶场景的复合高斯散点," arXiv:2312.07920, 2023.

[147] Y. Yan, H. Lin, C. Zhou, W. Wang, H. Sun, K. Zhan, X. Lang, X. Zhou, and S. Peng, "Street gaussians for modeling dynamic urban scenes," arXiv:2401.01339, 2024.

> Y. Yan, H. Lin, C. Zhou, W. Wang, H. Sun, K. Zhan, X. Lang, X. Zhou, and S. Peng, "Street gaussians: 用于动态城市场景建模的街道高斯模型," arXiv:2401.01339, 2024.

[148] H. Zhou, J. Shao, L. Xu, D. Bai, W. Qiu, B. Liu, Y. Wang, A. Geiger, and Y. Liao, "Hugs: Holistic urban 3d scene understanding via gaussian splatting," arXiv:2403.12722, 2024.

> H. Zhou, J. Shao, L. Xu, D. Bai, W. Qiu, B. Liu, Y. Wang, A. Geiger, and Y. Liao, "Hugs: 通过高斯散点实现的整体城市 3D 场景理解," arXiv:2403.12722, 2024.

[149] C. Zhao, S. Sun, R. Wang, Y. Guo, J.-J. Wan, Z. Huang, X. Huang, Y. V. Chen, and L. Ren, "Telc-gs: Tightly coupled lidar-camera gaussian splatting for surrounding autonomous driving scenes," arXiv:2404.02410, 2024.

> C. Zhao, S. Sun, R. Wang, Y. Guo, J.-J. Wan, Z. Huang, X. Huang, Y. V. Chen, and L. Ren, "Telc-gs: 紧耦合激光雷达-摄像头高斯散点用于周边自动驾驶场景," arXiv:2404.02410, 2024.

[150] Q. Herau, M. Bennehar, A. Moreau, N. Piasco, L. Roldao, D. Tsishkou, C. Migniot, P. Vasseur, and C. Demonceaux, "3dgs-calib: 3d gaussian splatting for multimodal spatiotemporal calibration," arXiv:2403.11577, 2024.

> Q. Herau, M. Bennehar, A. Moreau, N. Piasco, L. Roldao, D. Tsishkou, C. Migniot, P. Vasseur, and C. Demonceaux, "3dgs-calib: 用于多模态时空校准的 3D 高斯散点," arXiv:2403.11577, 2024.

[151] C. Yan, D. Qu, D. Wang, D. Xu, Z. Wang, B. Zhao, and X. Li, "Gs-slam: Dense visual slam with 3d gaussian splatting," arXiv:2311.11700, 2023.

> C. Yan, D. Qu, D. Wang, D. Xu, Z. Wang, B. Zhao, and X. Li, "Gs-slam: 基于 3D 高斯散点的密集视觉 SLAM," arXiv:2311.11700, 2023.

[152] H. Huang, L. Li, H. Cheng, and S.-K. Yeung, "Photo-slam: Real-time simultaneous localization and photorealistic mapping for monocular, stereo, and rgb-d cameras," arXiv:2311.16728, 2023.

> H. Huang, L. Li, H. Cheng, and S.-K. Yeung, "Photo-slam: 支持单目、立体及 RGB-D 摄像头的实时定位与光照真实感建图," arXiv:2311.16728, 2023.

[153] N. Keetha, J. Karhade, K. M. Jatavallabhula, G. Yang, S. Scherer, D. Ramanan, and J. Luiten, "Splatam: Splat, track & map 3d gaussians for dense rgb-d slam," arXiv:2312.02126, 2023.

N. Keetha, J. Karhade, K. M. Jatavallabhula, G. Yang, S. Scherer, D. Ramanan, and J. Luiten, "Splatam: 用于密集 RGB-D SLAM 的高斯散点跟踪与建图," arXiv:2312.02126, 2023.

[154] V. Yugay, Y. Li, T. Gevers, and M. R. Oswald, "Gaussian-slam: Photo-realistic dense slam with gaussian splatting," arXiv:2312.10070, 2023.

V. Yugay, Y. Li, T. Gevers, and M. R. Oswald, "Gaussian-slam: 基于高斯散点的光照真实感密集 SLAM," arXiv:2312.10070, 2023.

[155] J. Hu, X. Chen, B. Feng, G. Li, L. Yang, H. Bao, G. Zhang, and Z. Cui, "Cg-slam: Efficient dense rgb-d slam in a consistent uncertainty-aware 3d gaussian field," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 93-112.

J. Hu, X. Chen, B. Feng, G. Li, L. Yang, H. Bao, G. Zhang, 和 Z. Cui,"Cg-slam: 一种在一致性不确定性感知的三维高斯场中高效的密集 RGB-D SLAM,"发表于欧洲计算机视觉会议 (Eur. Conf. Comput. Vis.), Springer 出版社, 2025 年, 第 93-112 页。

[156] S. Ha, J. Yeon, and H. Yu, "Rgbd gs-icp slam," arXiv:2403.12550, 2024.

S. Ha, J. Yeon, 和 H. Yu,"RGBD GS-ICP SLAM,"arXiv:2403.12550, 2024 年。

[157] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in Proc. Int. Conf. Comput. Vis. (ICCV). IEEE, 2011, pp. 2564-2571.

E. Rublee, V. Rabaud, K. Konolige, 和 G. Bradski,"ORB: 一种高效替代 SIFT 或 SURF 的方法,"发表于国际计算机视觉会议 (ICCV) 论文集, IEEE, 2011 年, 第 2564-2571 页。

[158] J. J. Moré, "The levenberg-marquardt algorithm: implementation and theory," in Proc. Dundee Biennial Conf. Numer. Anal. Springer, 2006, pp. 105-116.

J. J. Moré,"Levenberg-Marquardt 算法: 实现与理论,"发表于邓迪双年数值分析会议论文集, Springer 出版社, 2006 年, 第 105-116 页。

[159] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp," in Robotics: Sci. Syst. (RSS), vol. 2, no. 4. Seattle, WA, 2009, p. 435.

A. Segal, D. Haehnel, 和 S. Thrun,"广义 ICP(Generalized-ICP),"发表于机器人科学系统会议 (RSS), 第 2 卷第 4 期, 西雅图, 华盛顿, 2009 年, 第 435 页。

[160] M. Li, S. Liu, and H. Zhou, "Sgs-slam: Semantic gaussian splatting for neural dense slam," arXiv:2402.03246, 2024.

M. Li, S. Liu, 和 H. Zhou,"SGS-SLAM: 用于神经密集 SLAM 的语义高斯点溅射 (Semantic Gaussian Splatting),"arXiv:2402.03246, 2024 年。

[161] S. Zhu, R. Qin, G. Wang, J. Liu, and H. Wang, "Semgauss-slam: Dense semantic gaussian splatting slam," arXiv:2403.07494, 2024.

S. Zhu, R. Qin, G. Wang, J. Liu, 和 H. Wang, "SemGauss-SLAM: 密集语义高斯点溅射 SLAM," arXiv:2403.07494, 2024 年。

[162] Y. Ji, Y. Liu, G. Xie, B. Ma, and Z. Xie, "Neds-slam: A novel neural explicit dense semantic slam framework using 3 d gaussian splatting," arXiv:2403.11679, 2024.

Y. Ji, Y. Liu, G. Xie, B. Ma, 和 Z. Xie, "NEDS-SLAM: 一种基于 3 d 高斯点溅射的神经显式密集语义 SLAM 新框架," arXiv:2403.11679, 2024 年。

[163] P. Jiang, G. Pandey, and S. Saripalli, "3dgs-reloc: 3d gaussian splatting for map representation and visual relocalization," arXiv:2403.11367, 2024.

P. Jiang, G. Pandey, 和 S. Saripalli, "3DGS-RELOC: 用于地图表示和视觉重定位的三维高斯点溅射," arXiv:2403.11367, 2024 年。

[164] X. Lei, M. Wang, W. Zhou, and H. Li, "Gaussnav: Gaussian splatting for visual navigation," arXiv:2403.11625, 2024.

X. Lei, M. Wang, W. Zhou, 和 H. Li, "GaussNav: 用于视觉导航的高斯点溅射," arXiv:2403.11625, 2024 年。

[165] J. Straub, T. Whelan, L. Ma, Y. Chen, E. Wijmans, S. Green, J. J. Engel, R. Mur-Artal, C. Ren, S. Verma et al., "The replica dataset: A digital replica of indoor spaces," arXiv:1906.05797, 2019.

J. Straub, T. Whelan, L. Ma, Y. Chen, E. Wijmans, S. Green, J. J. Engel, R. Mur-Artal, C. Ren, S. Verma 等, "Replica 数据集: 室内空间的数字复制品," arXiv:1906.05797, 2019 年。

[166] C. Yan, D. Qu, D. Xu, B. Zhao, Z. Wang, D. Wang, and X. Li, "Gs-slam: Dense visual slam with 3d gaussian splatting," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 19595-19604.

C. Yan, D. Qu, D. Xu, B. Zhao, Z. Wang, D. Wang, 和 X. Li, "GS-SLAM: 基于三维高斯点溅射的密集视觉 SLAM," 发表于 IEEE/CVF 计算机视觉与模式识别会议 (CVPR), 2024 年, 第 19595-19604 页。

[167] X. Guo, W. Zhang, R. Liu, P. Han, and H. Chen, "Motiongs: Compact gaussian splatting slam by motion filter," arXiv:2405.11129, 2024.

X. Guo, W. Zhang, R. Liu, P. Han, 和 H. Chen, "MotionGS: 通过运动滤波实现的紧凑型高斯点溅射 SLAM," arXiv:2405.11129, 2024 年。

[168] L. Zhu, Y. Li, E. Sandström, S. Huang, K. Schindler, and I. Ar-meni, "Loopsplat: Loop closure by registering 3 d gaussian splats," arXiv:2408.10154, 2024.

L. Zhu, Y. Li, E. Sandström, S. Huang, K. Schindler, 和 I. Armeni, "LoopSplat: 通过注册 3 d 高斯点实现的回环闭合," arXiv:2408.10154, 2024 年。

[169] J. Luiten, G. Kopanas, B. Leibe, and D. Ramanan, "Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis," arXiv:2308.09713, 2023.

J. Luiten, G. Kopanas, B. Leibe, 和 D. Ramanan, "动态三维高斯: 通过持久动态视图合成进行跟踪," arXiv:2308.09713, 2023 年。

[170] J. Sun, H. Jiao, G. Li, Z. Zhang, L. Zhao, and W. Xing, "3dgstream: On-the-fly training of 3 d gaussians for efficient streaming of photo-realistic free-viewpoint videos," arXiv:2403.01444, 2024.

J. Sun, H. Jiao, G. Li, Z. Zhang, L. Zhao, 和 W. Xing, "3DGStream: 用于高效流式传输逼真自由视点视频的 3 d 高斯在线训练," arXiv:2403.01444, 2024 年。

[171] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," ACM Trans. Graph., vol. 41, no. 4, pp. 1-15, 2022.

T. Müller, A. Evans, C. Schied, 和 A. Keller, "具有多分辨率哈希编码的即时神经图形基元," ACM 图形学汇刊, 第 41 卷, 第 4 期, 页 1-15, 2022 年。

[172] R. Shaw, J. Song, A. Moreau, M. Nazarczuk, S. Catley-Chandar, H. Dhamo, and E. Perez-Pellitero, "Swags: Sampling windows adaptively for dynamic 3d gaussian splatting," arXiv:2312.13308, 2023.

R. Shaw, J. Song, A. Moreau, M. Nazarczuk, S. Catley-Chandar, H. Dhamo, 和 E. Perez-Pellitero, "Swags: 用于动态三维高斯点撒布的自适应采样窗口," arXiv:2312.13308, 2023 年。

[173] Y. Xiao, X. Wang, J. Li, H. Cai, Y. Fan, N. Xue, M. Yang, Y. Shen, and S. Gao, "Bridging 3d gaussian and mesh for freeview video rendering," arXiv:2403.11453, 2024.

Y. Xiao, X. Wang, J. Li, H. Cai, Y. Fan, N. Xue, M. Yang, Y. Shen, 和 S. Gao, "连接三维高斯与网格的自由视角视频渲染," arXiv:2403.11453, 2024 年。

[174] Z. Yang, X. Gao, W. Zhou, S. Jiao, Y. Zhang, and X. Jin, "Deformable 3d gaussians for high-fidelity monocular dynamic scene reconstruction," arXiv:2309.13101, 2023.

Z. Yang, X. Gao, W. Zhou, S. Jiao, Y. Zhang, 和 X. Jin, "用于高保真单目动态场景重建的可变形三维高斯," arXiv:2309.13101, 2023 年。

[175] Y. Liang, N. Khan, Z. Li, T. Nguyen-Phuoc, D. Lanman, J. Tompkin, and L. Xiao, "Gaufre: Gaussian deformation fields for real-time dynamic novel view synthesis," arXiv:2312.11458, 2023.

Y. Liang, N. Khan, Z. Li, T. Nguyen-Phuoc, D. Lanman, J. Tompkin, 和 L. Xiao, "Gaufre: 用于实时动态新视角合成的高斯变形场," arXiv:2312.11458, 2023 年。

[176] G. Wu, T. Yi, J. Fang, L. Xie, X. Zhang, W. Wei, W. Liu, Q. Tian, and X. Wang, "4d gaussian splatting for real-time dynamic scene rendering," arXiv:2310.08528, 2023.

G. Wu, T. Yi, J. Fang, L. Xie, X. Zhang, W. Wei, W. Liu, Q. Tian, 和 X. Wang, "用于实时动态场景渲染的四维高斯点撒布," arXiv:2310.08528, 2023 年。

[177] B. P. Duisterhof, Z. Mandi, Y. Yao, J.-W. Liu, M. Z. Shou, S. Song, and J. Ichnowski, "Md-splatting: Learning metric deformation from 4d gaussians in highly deformable scenes," arXiv:2312.00583, 2023.

B. P. Duisterhof, Z. Mandi, Y. Yao, J.-W. Liu, M. Z. Shou, S. Song, 和 J. Ichnowski, "Md-splatting: 从四维高斯中学习高度可变形场景的度量变形," arXiv:2312.00583, 2023 年。

[178] Z. Guo, W. Zhou, L. Li, M. Wang, and H. Li, "Motion-aware 3d gaussian splatting for efficient dynamic scene reconstruction," IEEE Trans. Circuits Syst. Video Technol., 2024.

Z. Guo, W. Zhou, L. Li, M. Wang, 和 H. Li, "运动感知三维高斯点撒布用于高效动态场景重建," IEEE 电路系统视频技术汇刊, 2024 年。

[179] D. Li, S.-S. Huang, Z. Lu, X. Duan, and H. Huang, "St-4dgs: Spatial-temporally consistent 4 d gaussian splatting for efficient dynamic scene rendering," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-11.

D. Li, S.-S. Huang, Z. Lu, X. Duan, 和 H. Huang, "St-4dgs: 空间-时间一致的 4 d 高斯点撒布用于高效动态场景渲染," ACM SIGGRAPH 会议论文, 2024 年, 页 1-11。

[180] Z. Lu, X. Guo, L. Hui, T. Chen, M. Yang, X. Tang, F. Zhu, and Y. Dai, "3d geometry-aware deformable gaussian splatting for dynamic view synthesis," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 8900-8910.

Z. Lu, X. Guo, L. Hui, T. Chen, M. Yang, X. Tang, F. Zhu, 和 Y. Dai, "面向动态视图合成的三维几何感知可变形高斯点撒布," IEEE/CVF 计算机视觉与模式识别会议论文集, 2024 年, 页 8900-8910。

[181] K. Katsumata, D. M. Vo, and H. Nakayama, "An efficient 3d gaussian representation for monocular/multi-view dynamic scenes," arXiv:2311.12897, 2023.

K. Katsumata, D. M. Vo, 和 H. Nakayama, "一种高效的三维高斯表示用于单目/多视角动态场景," arXiv:2311.12897, 2023 年。

[182] A. Kratimenos, J. Lei, and K. Daniilidis, "Dynmf: Neural motion factorization for real-time dynamic view synthesis with 3 d gaussian splatting," arXiv:2312.00112, 2023.

A. Kratimenos, J. Lei, 和 K. Daniilidis, "Dynmf: 用于实时动态视图合成的神经运动分解与 3 d 高斯点撒布," arXiv:2312.00112, 2023 年。

[183] Y. Lin, Z. Dai, S. Zhu, and Y. Yao, "Gaussian-flow: 4d reconstruction with dynamic 3d gaussian particles," arXiv:2312.03431, 2023.

Y. Lin, Z. Dai, S. Zhu, 和 Y. Yao, "Gaussian-flow: 基于动态三维高斯粒子的四维重建," arXiv:2312.03431, 2023.

[184] Y.-H. Huang, Y.-T. Sun, Z. Yang, X. Lyu, Y.-P. Cao, and X. Qi, "Sc-gs: Sparse-controlled gaussian splatting for editable dynamic scenes," arXiv:2312.14937, 2023.

Y.-H. Huang, Y.-T. Sun, Z. Yang, X. Lyu, Y.-P. Cao, 和 X. Qi, "Sc-gs: 用于可编辑动态场景的稀疏控制高斯散点," arXiv:2312.14937, 2023.

[185] Z. Li, Z. Chen, Z. Li, and Y. Xu, "Spacetime gaussian feature splatting for real-time dynamic view synthesis," arXiv:2312.16812, 2023.

Z. Li, Z. Chen, Z. Li, 和 Y. Xu, "时空高斯特征散点用于实时动态视图合成," arXiv:2312.16812, 2023.

[186] A. Pumarola, E. Corona, G. Pons-Moll, and F. Moreno-Noguer, "D-nerf: Neural radiance fields for dynamic scenes," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2021, pp. 10318-10327.

A. Pumarola, E. Corona, G. Pons-Moll, 和 F. Moreno-Noguer, "D-nerf: 动态场景的神经辐射场," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2021, 页 10318-10327.

[187] H. Yu, J. Julin, Z. Á. Milacski, K. Niinuma, and L. A. Jeni, "Cogs: Controllable gaussian splatting," arXiv:2312.05664, 2023.

H. Yu, J. Julin, Z. Á. Milacski, K. Niinuma, 和 L. A. Jeni, "Cogs: 可控高斯散点," arXiv:2312.05664, 2023.

[188] Z. Yang, H. Yang, Z. Pan, X. Zhu, and L. Zhang, "Real-time photorealistic dynamic scene representation and rendering with 4 d gaussian splatting," arXiv:2310.10642, 2023.

Z. Yang, H. Yang, Z. Pan, X. Zhu, 和 L. Zhang, "基于 4 d 高斯散点的实时逼真动态场景表示与渲染," arXiv:2310.10642, 2023.

[189] Y. Duan, F. Wei, Q. Dai, Y. He, W. Chen, and B. Chen, "4d-rotor gaussian splatting: towards efficient novel view synthesis for dynamic scenes," in ACM SIGGRAPH Conf. Papers, 2024, pp. 1-11.

Y. Duan, F. Wei, Q. Dai, Y. He, W. Chen, 和 B. Chen, "4d-rotor 高斯散点: 面向动态场景高效新视角合成," 载于 ACM SIGGRAPH 会议论文, 2024, 页 1-11.

[190] G. Lu, S. Zhang, Z. Wang, C. Liu, J. Lu, and Y. Tang, "Manigaussian: Dynamic gaussian splatting for multi-task robotic manipulation," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 349-366.

G. Lu, S. Zhang, Z. Wang, C. Liu, J. Lu, 和 Y. Tang, "Manigaussian: 用于多任务机器人操作的动态高斯散点," 载于欧洲计算机视觉会议, Springer, 2025, 页 349-366.

[191] T. Xie, Z. Zong, Y. Qiu, X. Li, Y. Feng, Y. Yang, and C. Jiang, "Physgaussian: Physics-integrated 3d gaussians for generative dynamics," arXiv:2311.12198, 2023.

T. Xie, Z. Zong, Y. Qiu, X. Li, Y. Feng, Y. Yang, 和 C. Jiang, "Physgaussian: 集成物理的三维高斯生成动力学," arXiv:2311.12198, 2023.

[192] A. Guédon and V. Lepetit, "Sugar: Surface-aligned gaussian splatting for efficient 3 d mesh reconstruction and high-quality mesh rendering," arXiv:2311.12775, 2023.

A. Guédon 和 V. Lepetit, "Sugar: 面向高效 3 d 网格重建与高质量网格渲染的表面对齐高斯散点," arXiv:2311.12775, 2023.

[193] H. Chen, C. Li, and G. H. Lee, "Neusg: Neural implicit surface reconstruction with 3d gaussian splatting guidance," arXiv:2312.00846, 2023.

H. Chen, C. Li, 和 G. H. Lee, "Neusg: 基于三维高斯散点引导的神经隐式表面重建,"arXiv:2312.00846, 2023.

[194] X. Lyu, Y.-T. Sun, Y.-H. Huang, X. Wu, Z. Yang, Y. Chen, J. Pang, and X. Qi, "3dgsr: Implicit surface reconstruction with 3d gaussian splatting," arXiv:2404.00409, 2024.

X. Lyu, Y.-T. Sun, Y.-H. Huang, X. Wu, Z. Yang, Y. Chen, J. Pang, 和 X. Qi, "3dgsr: 基于三维高斯散点的隐式表面重建," arXiv:2404.00409, 2024.

[195] M. Yu, T. Lu, L. Xu, L. Jiang, Y. Xiangli, and B. Dai, "Gsdf: 3dgs meets sdf for improved rendering and reconstruction," arXiv:2403.16964, 2024.

M. Yu, T. Lu, L. Xu, L. Jiang, Y. Xiangli, 和 B. Dai, "Gsdf: 三维高斯散点与有符号距离函数 (SDF) 的结合以提升渲染与重建,"arXiv:2403.16964, 2024.

[196] P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura, and W. Wang, "Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction," arXiv:2106.10689, 2021.

P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura, 和 W. Wang, "Neus: 通过体渲染学习神经隐式表面以实现多视角重建,"arXiv:2106.10689, 2021.

[197] P. Dai, J. Xu, W. Xie, X. Liu, H. Wang, and W. Xu, "High-quality surface reconstruction using gaussian surfels," arXiv:2404.17774, 2024.

P. Dai, J. Xu, W. Xie, X. Liu, H. Wang, 和 W. Xu, "基于高斯表面元素的高质量表面重建," arXiv:2404.17774, 2024.

[198] Z. Yu, T. Sattler, and A. Geiger, "Gaussian opacity fields: Efficient and compact surface reconstruction in unbounded scenes," arXiv:2404.10772, 2024.

Z. Yu, T. Sattler, 和 A. Geiger, "高斯不透明度场: 用于无界场景的高效紧凑表面重建," arXiv:2404.10772, 2024.

[199] B. Huang, Z. Yu, A. Chen, A. Geiger, and S. Gao, "2d gaussian splatting for geometrically accurate radiance fields," arXiv:2403.17888, 2024.

B. Huang, Z. Yu, A. Chen, A. Geiger, 和 S. Gao,"用于几何精确辐射场的二维高斯散点法 (2D Gaussian Splatting),"arXiv:2403.17888, 2024.

[200] C. Reiser, S. Garbin, P. Srinivasan, D. Verbin, R. Szeliski, B. Mildenhall, J. Barron, P. Hedman, and A. Geiger, "Binary opacity grids: Capturing fine geometric detail for mesh-based view synthesis," ACM Trans. Graph., vol. 43, no. 4, pp. 1-14, 2024.

C. Reiser, S. Garbin, P. Srinivasan, D. Verbin, R. Szeliski, B. Mildenhall, J. Barron, P. Hedman, 和 A. Geiger, "二进制不透明度网格: 捕捉基于网格视图合成的细微几何细节,"ACM Trans. Graph., 第 43 卷, 第 4 期, 第 1-14 页, 2024.

[201] Y. Chen, Z. Chen, C. Zhang, F. Wang, X. Yang, Y. Wang, Z. Cai, L. Yang, H. Liu, and G. Lin, "Gaussianeditor: Swift and controllable 3d editing with gaussian splatting," arXiv:2311.14521, 2023.

Y. Chen, Z. Chen, C. Zhang, F. Wang, X. Yang, Y. Wang, Z. Cai, L. Yang, H. Liu, 和 G. Lin,"Gaussianeditor: 基于高斯散点法的快速且可控的三维编辑,"arXiv:2311.14521, 2023.

[202] F. Palandra, A. Sanchietti, D. Baieri, and E. Rodolà, "Gsedit: Efficient text-guided editing of 3 d objects via gaussian splatting," arXiv:2403.05154, 2024.

F. Palandra, A. Sanchietti, D. Baieri, 和 E. Rodolà, "Gsedit: 通过高斯散点法实现的高效文本引导对象编辑,"arXiv:2403.05154, 2024.

[203] T. Brooks, A. Holynski, and A. A. Efros, "Instructpix2pix: Learning to follow image editing instructions," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2023, pp. 18392-18402.

T. Brooks, A. Holynski, 和 A. A. Efros,"Instructpix2pix: 学习遵循图像编辑指令,"载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2023, 第 18392-18402 页.

[204] J. Fang, J. Wang, X. Zhang, L. Xie, and Q. Tian, "Gaussianeditor: Editing 3d gaussians delicately with text instructions," arXiv:2311.16037, 2023.

J. Fang, J. Wang, X. Zhang, L. Xie, 和 Q. Tian, "Gaussianeditor: 基于文本指令的精细三维高斯编辑,"arXiv:2311.16037, 2023.

[205] J. Wu, J.-W. Bian, X. Li, G. Wang, I. Reid, P. Torr, and V. A. Prisacariu, "Gaussctrl: Multi-view consistent text-driven 3d gaussian splatting editing," arXiv:2403.08733, 2024.

J. Wu, J.-W. Bian, X. Li, G. Wang, I. Reid, P. Torr, 和 V. A. Prisacariu, "Gaussctrl: 多视角一致的文本驱动三维高斯散点编辑,"arXiv:2403.08733, 2024.

[206] Y. Wang, X. Yi, Z. Wu, N. Zhao, L. Chen, and H. Zhang, "View-consistent 3 d editing with gaussian splatting," arXiv:2403.11868, 2024.

Y. Wang, X. Yi, Z. Wu, N. Zhao, L. Chen, 和 H. Zhang, "基于高斯散点法的视角一致对象编辑,"arXiv:2403.11868, 2024.

[207] L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2023, pp. 3836-3847.

> L. Zhang, A. Rao, 和 M. Agrawala, "为文本到图像扩散模型添加条件控制," 载于 IEEE/CVF 国际计算机视觉会议论文集, 2023, 第 3836-3847 页.

[208] R. Shao, J. Sun, C. Peng, Z. Zheng, B. Zhou, H. Zhang, and Y. Liu, "Control4d: Dynamic portrait editing by learning 4 d gan from 2 d diffusion-based editor," arXiv:2305.20082, vol. 2, no. 6, p. 16, 2023.

> R. Shao, J. Sun, C. Peng, Z. Zheng, B. Zhou, H. Zhang, 和 Y. Liu, "Control4d: 通过学习基于扩散的编辑器的生成对抗网络实现动态肖像编辑," arXiv:2305.20082, 第 2 卷, 第 6 期, 第 16 页, 2023.

[209] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," Commun. ACM, vol. 63, no. 11, pp. 139-144, 2020.

> I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, 和 Y. Bengio, "生成对抗网络 (Generative Adversarial Networks)," Commun. ACM, 第 63 卷, 第 11 期, 第 139-144 页, 2020.

[210] J. Zhuang, D. Kang, Y.-P. Cao, G. Li, L. Lin, and Y. Shan, "Tip-editor: An accurate 3d editor following both text-prompts and image-prompts," arXiv:2401.14828, 2024.

> J. Zhuang, D. Kang, Y.-P. Cao, G. Li, L. Lin, 和 Y. Shan, "Tip-editor: 精准遵循文本和图像提示的三维编辑器," arXiv:2401.14828, 2024.

[211] J. Huang and H. Yu, "Point'n move: Interactive scene object manipulation on gaussian splatting radiance fields," arXiv:2311.16737, 2023.

> J. Huang 和 H. Yu, "Point'n move: 基于高斯散点辐射场的交互式场景对象操作," arXiv:2311.16737, 2023.

[212] Q.-Y. Feng, G.-C. Cao, H.-X. Chen, Q.-C. Xu, T.-J. Mu, R. Martin, and S.-M. Hu, "Evsplitting: An efficient and visually consistent splitting algorithm for 3d gaussian splatting," in SIGGRAPH Asia 2024 Conf. Papers, 2024, pp. 1-11.

> Q.-Y. Feng, G.-C. Cao, H.-X. Chen, Q.-C. Xu, T.-J. Mu, R. Martin, 和 S.-M. Hu, "Evsplitting: 一种高效且视觉一致的三维高斯散点分割算法," 载于 SIGGRAPH Asia 2024 会议论文集, 2024, 第 1-11 页.

[213] A. Saroha, M. Gladkova, C. Curreli, T. Yenamandra, and D. Cremers, "Gaussian splatting in style," arXiv:2403.08498, 2024.

> A. Saroha, M. Gladkova, C. Curreli, T. Yenamandra, 和 D. Cremers, "风格中的高斯点渲染 (Gaussian splatting)," arXiv:2403.08498, 2024.

[214] Y.-H. Huang, Y. He, Y.-J. Yuan, Y.-K. Lai, and L. Gao, "Stylizednerf: consistent 3d scene stylization as stylized nerf via 2d-3d mutual learning," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 18342-18352.

Y.-H. Huang, Y. He, Y.-J. Yuan, Y.-K. Lai, 和 L. Gao, "StylizedNeRF: 通过二维-三维互学习实现一致的三维场景风格化作为风格化 NeRF," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2022, 页 18342-18352.

[215] J. Gao, C. Gu, Y. Lin, H. Zhu, X. Cao, L. Zhang, and Y. Yao, "Relightable 3d gaussian: Real-time point cloud relighting with brdf decomposition and ray tracing," arXiv:2311.16043, 2023.

J. Gao, C. Gu, Y. Lin, H. Zhu, X. Cao, L. Zhang, 和 Y. Yao, "可重光照三维高斯: 基于 BRDF 分解和光线追踪的实时点云重光照," arXiv:2311.16043, 2023.

[216] Z. Liang, Q. Zhang, Y. Feng, Y. Shan, and K. Jia, "Gs-ir: 3d gaussian splatting for inverse rendering," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 21 644-21 653.

Z. Liang, Q. Zhang, Y. Feng, Y. Shan, 和 K. Jia, "GS-IR: 用于逆向渲染的三维高斯点渲染," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 页 21644-21653.

[217] Z. Kuang, Y. Yang, S. Dong, J. Ma, H. Fu, and Y. Zheng, "Olat gaussians for generic relightable appearance acquisition," in ACM SIGGRAPH Asia Conf. Papers, 2024, pp. 1-11.

Z. Kuang, Y. Yang, S. Dong, J. Ma, H. Fu, 和 Y. Zheng, "OLAT 高斯用于通用可重光照外观获取," 载于 ACM SIGGRAPH Asia 会议论文, 2024, 页 1-11.

[218] Z. Bi, Y. Zeng, C. Zeng, F. Pei, X. Feng, K. Zhou, and H. Wu, "Gs3: Efficient relighting with triple gaussian splatting," in ACM SIGGRAPH Asia Conf. Papers, 2024, pp. 1-12.

Z. Bi, Y. Zeng, C. Zeng, F. Pei, X. Feng, K. Zhou, 和 H. Wu, "GS3: 基于三重高斯点渲染的高效重光照," 载于 ACM SIGGRAPH Asia 会议论文, 2024, 页 1-12.

[219] M. Ye, M. Danelljan, F. Yu, and L. Ke, "Gaussian grouping: Segment and edit anything in 3d scenes," arXiv:2312.00732, 2023.

M. Ye, M. Danelljan, F. Yu, 和 L. Ke, "高斯分组: 三维场景中的任意分割与编辑," arXiv:2312.00732, 2023.

[220] S. Zhou, H. Chang, S. Jiang, Z. Fan, Z. Zhu, D. Xu, P. Chari, S. You, Z. Wang, and A. Kadambi, "Feature 3dgs: Supercharging 3d gaussian splatting to enable distilled feature fields," arXiv:2312.03203, 2023.

S. Zhou, H. Chang, S. Jiang, Z. Fan, Z. Zhu, D. Xu, P. Chari, S. You, Z. Wang, 和 A. Kadambi, "Feature 3DGS: 增强三维高斯点渲染以实现蒸馏特征场," arXiv:2312.03203, 2023.

[221] K. Lan, H. Li, H. Shi, W. Wu, Y. Liao, L. Wang, and P. Zhou, "2d-guided 3d gaussian segmentation," arXiv:2312.16047, 2023.

K. Lan, H. Li, H. Shi, W. Wu, Y. Liao, L. Wang, 和 P. Zhou, "二维引导的三维高斯分割," arXiv:2312.16047, 2023.

[222] B. Dou, T. Zhang, Y. Ma, Z. Wang, and Z. Yuan, "Cosseggaussians: Compact and swift scene segmenting 3 d gaussians," arXiv:2401.05925, 2024.

B. Dou, T. Zhang, Y. Ma, Z. Wang, 和 Z. Yuan, "CossegGaussians: 紧凑且快速的场景分割高斯," arXiv:2401.05925, 2024.

[223] Q. Gu, Z. Lv, D. Frost, S. Green, J. Straub, and C. Sweeney, "Egolifter: Open-world 3d segmentation for egocentric perception," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 382-400.

Q. Gu, Z. Lv, D. Frost, S. Green, J. Straub, 和 C. Sweeney, "EgoLifter: 面向自我中心感知的开放世界三维分割,"载于欧洲计算机视觉会议, Springer, 2025, 页 382-400.

[224] S. Choi, H. Song, J. Kim, T. Kim, and H. Do, "Click-gaussian: Interactive segmentation to any 3d gaussians," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 289-305.

S. Choi, H. Song, J. Kim, T. Kim, 和 H. Do, "Click-Gaussian: 面向任意三维高斯的交互式分割,"载于欧洲计算机视觉会议, Springer, 2025, 页 289-305.

[225] Y. Yue, A. Das, F. Engelmann, S. Tang, and J. E. Lenssen, "Improving 2d feature representations by 3d-aware fine-tuning," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 57-74.

Y. Yue, A. Das, F. Engelmann, S. Tang, 和 J. E. Lenssen, "通过三维感知微调提升二维特征表示,"载于欧洲计算机视觉会议, Springer, 2025, 页 57-74.

[226] J.-C. Shi, M. Wang, H.-B. Duan, and S.-H. Guan, "Language embedded 3 d gaussians for open-vocabulary scene understanding," arXiv:2311.18482, 2023.

J.-C. Shi, M. Wang, H.-B. Duan, 和 S.-H. Guan, "嵌入语言的 3 d 高斯用于开放词汇场景理解," arXiv:2311.18482, 2023.

[227] M. Qin, W. Li, J. Zhou, H. Wang, and H. Pfister, "Langsplat: 3d language gaussian splatting," arXiv:2312.16084, 2023.

M. Qin, W. Li, J. Zhou, H. Wang, 和 H. Pfister, "LangSplat: 三维语言高斯点渲染," arXiv:2312.16084, 2023.

[228] X. Zuo, P. Samangouei, Y. Zhou, Y. Di, and M. Li, "Fmgs: Foundation model embedded 3d gaussian splatting for holistic 3 d scene understanding," arXiv:2401.01970, 2024.

Zuo X., Samangouei P., Zhou Y., Di Y., Li M., "Fmgs: 嵌入基础模型的三维高斯点绘制用于整体场景理解," arXiv:2401.01970, 2024.

[229] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark et al., "Learning transferable visual models from natural language supervision," in Proc. Int. Conf. Mach. Learn. PMLR, 2021, pp. 8748-8763.

Radford A., Kim J. W., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J. 等，"从自然语言监督中学习可迁移的视觉模型，"载于国际机器学习大会论文集 PMLR, 2021, 页 8748-8763.

[230] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin, "Emerging properties in self-supervised vision transformers," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2021, pp. 9650-9660.

Caron M., Touvron H., Misra I., Jégou H., Mairal J., Bojanowski P., Joulin A., "自监督视觉变换器中的新兴特性，"载于 IEEE/CVF 国际计算机视觉大会论文集, 2021, 页 9650-9660.

[231] A. Van Den Oord, O. Vinyals et al., "Neural discrete representation learning," Adv. Neural Inf. Process. Syst., vol. 30, 2017.

Van Den Oord A., Vinyals O. 等，"神经离散表示学习，"神经信息处理系统进展, 第 30 卷, 2017.

[232] A. De Vaucorbeil, V. P. Nguyen, S. Sinaie, and J. Y. Wu, "Material point method after 25 years: Theory, implementation, and applications," Adv. Appl. Mech., vol. 53, pp. 185-398, 2020.

De Vaucorbeil A., Nguyen V. P., Sinaie S., Wu J. Y., "材料点方法 25 年回顾: 理论、实现与应用，"应用力学进展, 第 53 卷, 页 185-398, 2020.

[233] T. Zhang, H.-X. Yu, R. Wu, B. Y. Feng, C. Zheng, N. Snavely, J. Wu, and W. T. Freeman, "Physdreamer: Physics-based interaction with 3d objects via video generation," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 388-406.

Zhang T., Yu H.-X., Wu R., Feng B. Y., Zheng C., Snavely N., Wu J., Freeman W. T., "Physdreamer: 基于物理的三维物体交互视频生成，"欧洲计算机视觉大会, Springer, 2025, 页 388-406.

[234] Y. Jiang, C. Yu, T. Xie, X. Li, Y. Feng, H. Wang, M. Li, H. Lau, F. Gao, Y. Yang et al., "Vr-gs: A physical dynamics-aware interactive gaussian splatting system in virtual reality," in ACM SIGGRAPH 2024 Conf. Papers, 2024, pp. 1-1.

Jiang Y., Yu C., Xie T., Li X., Feng Y., Wang H., Li M., Lau H., Gao F., Yang Y. 等，"Vr-gs: 虚拟现实中具备物理动力学感知的交互式高斯点绘制系统，"ACM SIGGRAPH 2024 会议论文, 2024, 页 1-1.

[235] J. Jung, J. Han, H. An, J. Kang, S. Park, and S. Kim, "Relaxing accurate initialization constraint for 3 d gaussian splatting," arXiv:2403.09413, 2024.

Jung J., Han J., An H., Kang J., Park S., Kim S., "放宽高斯点绘制的精确初始化约束，"arXiv:2403.09413, 2024.

[236] D. Das, C. Wewer, R. Yunus, E. Ilg, and J. E. Lenssen, "Neural parametric gaussians for monocular non-rigid object reconstruction," arXiv:2312.01196, 2023.

Das D., Wewer C., Yunus R., Ilg E., Lenssen J. E., "用于单目非刚性物体重建的神经参数高斯模型，"arXiv:2312.01196, 2023.

[237] L. Gao, J. Yang, B.-T. Zhang, J.-M. Sun, Y.-J. Yuan, H. Fu, and Y.-K. Lai, "Mesh-based gaussian splatting for real-time large-scale deformation," arXiv:2402.04796, 2024.

Gao L., Yang J., Zhang B.-T., Sun J.-M., Yuan Y.-J., Fu H., Lai Y.-K., "基于网格的高斯点绘制实现实时大规模变形," arXiv:2402.04796, 2024.

[238] H. Matsuki, R. Murai, P. H. Kelly, and A. J. Davison, "Gaussian splatting slam," arXiv:2312.06741, 2023.

Matsuki H., Murai R., Kelly P. H., Davison A. J., "高斯点绘制 SLAM," arXiv:2312.06741, 2023.

[239] J. Lei, Y. Wang, G. Pavlakos, L. Liu, and K. Daniilidis, "Gart: Gaussian articulated template models," arXiv:2311.16099, 2023.

Lei J., Wang Y., Pavlakos G., Liu L., Daniilidis K., "Gart: 高斯关节模板模型," arXiv:2311.16099, 2023.

[240] L. Huang, J. Bai, J. Guo, and Y. Guo, "Gs++: Error analyzing and optimal gaussian splatting," arXiv:2402.00752, 2024.

Huang L., Bai J., Guo J., Guo Y., "Gs++: 误差分析与最优高斯点绘制," arXiv:2402.00752, 2024.

[241] N. Moenne-Loccoz, A. Mirzaei, O. Perel, R. de Lutio, J. Martinez Esturo, G. State, S. Fidler, N. Sharp, and Z. Gojcic, "3d gaussian ray tracing: Fast tracing of particle scenes," ACM Trans. Graph., vol. 43, no. 6, pp. 1-19, 2024.

Moenne-Loccoz N., Mirzaei A., Perel O., de Lutio R., Martinez Esturo J., State G., Fidler S., Sharp N., Gojcic Z., "三维高斯光线追踪: 快速粒子场景追踪," ACM 图形学汇刊, 第 43 卷第 6 期, 页 1-19, 2024.

[242] Y. Li, C. Lyu, Y. Di, G. Zhai, G. H. Lee, and F. Tombari, "Ge-ogaussian: Geometry-aware gaussian splatting for scene rendering," arXiv:2403.11324, 2024.

Y. Li, C. Lyu, Y. Di, G. Zhai, G. H. Lee, 和 F. Tombari, "Ge-ogaussian: 面向场景渲染的几何感知高斯点溅射 (Geometry-aware Gaussian Splatting)," arXiv:2403.11324, 2024.

[243] H. Xiong, S. Muttukuru, R. Upadhyay, P. Chari, and A. Kadambi, "Sparsegs: Real-time 360°sparse view synthesis using gaussian splatting," arXiv:2312.00206, 2023.

H. Xiong, S. Muttukuru, R. Upadhyay, P. Chari, 和 A. Kadambi, "Sparsegs: 基于高斯点溅射的实时 360," arXiv:2312.00206, 2023.

[244] L. Bolanos, S.-Y. Su, and H. Rhodin, "Gaussian shadow casting for neural characters," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 20997-21006.

L. Bolanos, S.-Y. Su, 和 H. Rhodin, "用于神经角色的高斯阴影投射," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 页 20997-21006.

[245] A. Nichol, H. Jun, P. Dhariwal, P. Mishkin, and M. Chen, "Point-e: A system for generating 3 d point clouds from complex prompts," arXiv:2212.08751, 2022.

> A. Nichol, H. Jun, P. Dhariwal, P. Mishkin, 和 M. Chen, "Point-e: 一个基于复杂提示生成点云的系统," arXiv:2212.08751, 2022.

[246] Y. Liu, C. Lin, Z. Zeng, X. Long, L. Liu, T. Komura, and W. Wang, "Syncdreamer: Generating multiview-consistent images from a single-view image," arXiv:2309.03453, 2023.

> Y. Liu, C. Lin, Z. Zeng, X. Long, L. Liu, T. Komura, 和 W. Wang, "Syncdreamer: 从单视图图像生成多视图一致图像," arXiv:2309.03453, 2023.

[247] Z. Li, Y. Chen, L. Zhao, and P. Liu, "Controllable text-to-3d generation via surface-aligned gaussian splatting," arXiv:2403.09981, 2024.

> Z. Li, Y. Chen, L. Zhao, 和 P. Liu, "通过表面对齐高斯点溅射实现可控文本到三维生成," arXiv:2403.09981, 2024.

[248] M. Armandpour, A. Sadeghian, H. Zheng, A. Sadeghian, and M. Zhou, "Re-imagine the negative prompt algorithm: Transform 2d diffusion into 3d, alleviate janus problem and beyond," arXiv:2304.04968, 2023.

> M. Armandpour, A. Sadeghian, H. Zheng, A. Sadeghian, 和 M. Zhou, "重新构想负面提示算法: 将二维扩散转换为三维, 缓解双面神问题及更多," arXiv:2304.04968, 2023.

[249] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "Lora: Low-rank adaptation of large language models," arXiv:2106.09685, 2021.

> E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, 和 W. Chen, "LoRA: 大规模语言模型的低秩适配," arXiv:2106.09685, 2021.

[250] S. Sun, M. Mielle, A. J. Lilienthal, and M. Magnusson, "High-fidelity slam using gaussian splatting with rendering-guided densification and regularized optimization," arXiv:2403.12535, 2024.

> S. Sun, M. Mielle, A. J. Lilienthal, 和 M. Magnusson, "基于高斯点溅射的高保真 SLAM, 结合渲染引导的密集化和正则化优化," arXiv:2403.12535, 2024.

[251] Y. Liu, C. Luo, L. Fan, N. Wang, J. Peng, and Z. Zhang, "Citygaussian: Real-time high-quality large-scale scene rendering with gaussians," in Eur. Conf. Comput. Vis. Springer, 2025, pp. 265-282.

> Y. Liu, C. Luo, L. Fan, N. Wang, J. Peng, 和 Z. Zhang, "Citygaussian: 基于高斯点的实时高质量大规模场景渲染," 载于欧洲计算机视觉会议论文集, Springer, 2025, 页 265-282.

[252] B. Kerbl, A. Meuleman, G. Kopanas, M. Wimmer, A. Lanvin, and G. Drettakis, "A hierarchical 3d gaussian representation for real-time rendering of very large datasets," ACM Trans. Graph., vol. 43, no. 4, pp. 1-15, 2024.

B. Kerbl, A. Meuleman, G. Kopanas, M. Wimmer, A. Lanvin, 和 G. Drettakis, "一种用于超大规模数据集实时渲染的分层三维高斯表示," ACM 计算机图形学汇刊, 第 43 卷第 4 期, 页 1-15, 2024.

[253] Z. Zhang, W. Hu, Y. Lao, T. He, and H. Zhao, "Pixel-gs: Density control with pixel-aware gradient for 3 d gaussian splatting," arXiv:2403.15530, 2024.

Z. Zhang, W. Hu, Y. Lao, T. He, 和 H. Zhao, "Pixel-gs: 基于像素感知梯度的高斯点溅射密度控制," arXiv:2403.15530, 2024.

[254] Q. Feng, G. Cao, H. Chen, T.-J. Mu, R. R. Martin, and S.-M. Hu, "A new split algorithm for 3d gaussian splatting," arXiv:2403.09143, 2024.

Q. Feng, G. Cao, H. Chen, T.-J. Mu, R. R. Martin, 和 S.-M. Hu, "一种新的三维高斯点溅射分割算法," arXiv:2403.09143, 2024.

[255] M. Niemeyer, F. Manhardt, M.-J. Rakotosaona, M. Oechsle, D. Duck-worth, R. Gosula, K. Tateno, J. Bates, D. Kaeser, and F. Tombari, "Radsplat: Radiance field-informed gaussian splatting for robust real-time rendering with 900+ fps," arXiv:2403.13806, 2024.

M. Niemeyer, F. Manhardt, M.-J. Rakotosaona, M. Oechsle, D. Duckworth, R. Gosula, K. Tateno, J. Bates, D. Kaeser, 和 F. Tombari, "Radsplat: 基于辐射场信息的高斯点溅射，实现 900+ 帧每秒的鲁棒实时渲染," arXiv:2403.13806, 2024.

[256] G. Fang and B. Wang, "Mini-splatting: Representing scenes with a constrained number of gaussians," arXiv:2403.14166, 2024.

G. Fang 和 B. Wang, "Mini-splatting: 用有限数量高斯点表示场景," arXiv:2403.14166, 2024.

[257] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in Proc. Eurographics Symp. Geom. Process. (SGP), vol. 7, no. 4, 2006.

M. Kazhdan, M. Bolitho, 和 H. Hoppe, "泊松表面重建 (Poisson surface reconstruction)," 载于 Eurographics 几何处理研讨会 (SGP) 论文集，第 7 卷，第 4 期，2006 年。

[258] J. Tang, H. Zhou, X. Chen, T. Hu, E. Ding, J. Wang, and G. Zeng, "Delicate textured mesh recovery from nerf via adaptive surface refinement," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2023, pp. 17739-17749.

J. Tang, H. Zhou, X. Chen, T. Hu, E. Ding, J. Wang, 和 G. Zeng, "通过自适应表面细化从 NeRF 恢复精细纹理网格," 载于 IEEE/CVF 国际计算机视觉会议论文集，2023 年，第 17739-17749 页。

[259] H. Li, Y. Gao, D. Zhang, C. Wu, Y. Dai, C. Zhao, H. Feng, E. Ding, J. Wang, and J. Han, "Ggrt: Towards generalizable 3d gaussians without pose priors in real-time," arXiv:2403.10147, 2024.

H. Li, Y. Gao, D. Zhang, C. Wu, Y. Dai, C. Zhao, H. Feng, E. Ding, J. Wang, 和 J. Han, "GGRt: 面向实时无姿态先验的可泛化三维高斯 (3D Gaussians)," arXiv:2403.10147, 2024 年。

[260] D. Malarz, W. Smolak, J. Tabor, S. Tadeja, and P. Spurek, "Gaussian splatting with nerf-based color and opacity."

D. Malarz, W. Smolak, J. Tabor, S. Tadeja, 和 P. Spurek, "基于 NeRF 的颜色和不透明度的高斯点绘制 (Gaussian splatting)。"

[261] S. Zuffi, A. Kanazawa, D. W. Jacobs, and M. J. Black, "3d menagerie: Modeling the 3d shape and pose of animals," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 6365-6373.

S. Zuffi, A. Kanazawa, D. W. Jacobs, 和 M. J. Black, "3D 动物模型库: 动物三维形状与姿态建模," 载于 IEEE 计算机视觉与模式识别会议论文集，2017 年，第 6365-6373 页。

[262] S. D. Porumbescu, B. Budge, L. Feng, and K. I. Joy, "Shell maps," ACM Trans. Graph., vol. 24, no. 3, pp. 626-633, 2005.

S. D. Porumbescu, B. Budge, L. Feng, 和 K. I. Joy, "壳映射 (Shell maps)," ACM 计算机图形学汇刊，第 24 卷，第 3 期，第 626-633 页，2005 年。

[263] R. W. Sumner, J. Schmid, and M. Pauly, "Embedded deformation for shape manipulation," in ACM SIGGRAPH 2007 Papers, 2007, pp. 80-es.

R. W. Sumner, J. Schmid, 和 M. Pauly, "嵌入式变形用于形状操控," 载于 ACM SIGGRAPH 2007 论文集，2007 年，第 80-es 页。

[264] Z. Fan, W. Cong, K. Wen, K. Wang, J. Zhang, X. Ding, D. Xu, B. Ivanovic, M. Pavone, G. Pavlakos et al., "Instantsplat: Unbounded sparse-view pose-free gaussian splatting in 40 seconds," IEEE Trans. Circuits Syst. Video Technol., 2024.

Z. Fan, W. Cong, K. Wen, K. Wang, J. Zhang, X. Ding, D. Xu, B. Ivanovic, M. Pavone, G. Pavlakos 等, "Instantsplat:40 秒内实现无姿态限制的无限稀疏视角高斯点绘制," IEEE 电路系统与视频技术汇刊，2024 年。

[265] Y. Sun, X. Wang, Y. Zhang, J. Zhang, C. Jiang, Y. Guo, and F. Wang, "icomma: Inverting 3 d gaussians splatting for camera pose estimation via comparing and matching," arXiv:2312.09031, 2023.

Y. Sun, X. Wang, Y. Zhang, J. Zhang, C. Jiang, Y. Guo, 和 F. Wang, "iComma: 通过比较匹配反演高斯点绘制以估计相机姿态," arXiv:2312.09031, 2023 年。

[266] R. Martin-Brualla, N. Radwan, M. S. Sajjadi, J. T. Barron, A. Dosovit-skiy, and D. Duckworth, "Nerf in the wild: Neural radiance fields for unconstrained photo collections," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2021, pp. 7210-7219.

R. Martin-Brualla, N. Radwan, M. S. Sajjadi, J. T. Barron, A. Dosovitskiy, 和 D. Duckworth, "野外 NeRF: 用于无约束照片集合的神经辐射场," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集，2021 年，第 7210-7219 页。

[267] X. Chen, Q. Zhang, X. Li, Y. Chen, Y. Feng, X. Wang, and J. Wang, "Hallucinated neural radiance fields in the wild," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 12943-12952.

> X. Chen, Q. Zhang, X. Li, Y. Chen, Y. Feng, X. Wang, 和 J. Wang, "野外幻觉神经辐射场," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2022 年, 第 12943-12952 页。

[268] Y. Yang, S. Zhang, Z. Huang, Y. Zhang, and M. Tan, "Cross-ray neural radiance fields for novel-view synthesis from unconstrained image collections," in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2023, pp. 15901-15911.

> Y. Yang, S. Zhang, Z. Huang, Y. Zhang, 和 M. Tan, "跨射线神经辐射场用于无约束图像集合的新视角合成," 载于 IEEE/CVF 国际计算机视觉会议 (ICCV) 论文集, 2023 年, 第 15901-15911 页。

[269] H. Dahmani, M. Bennehar, N. Piasco, L. Roldao, and D. Tsishkou, "Swag: Splatting in the wild images with appearance-conditioned gaussians," arXiv:2403.10427, 2024.

> H. Dahmani, M. Bennehar, N. Piasco, L. Roldao, 和 D. Tsishkou, "SWAG: 基于外观条件高斯的野外图像点绘制," arXiv:2403.10427, 2024 年。

[270] D. Zhang, C. Wang, W. Wang, P. Li, M. Qin, and H. Wang, "Gaussian in the wild: 3d gaussian splatting for unconstrained image collections," arXiv:2403.15704, 2024.

> D. Zhang, C. Wang, W. Wang, P. Li, M. Qin, 和 H. Wang, "野外高斯: 用于无约束图像集合的三维高斯点绘制," arXiv:2403.15704, 2024 年。

[271] Y. Fu, S. Liu, A. Kulkarni, J. Kautz, A. A. Efros, and X. Wang, "Colmap-free 3d gaussian splatting," arXiv:2312.07504, 2023.

> Y. Fu, S. Liu, A. Kulkarni, J. Kautz, A. A. Efros, 和 X. Wang, "无 Colmap 的三维高斯点绘制 (Colmap-free 3D Gaussian Splatting)," arXiv:2312.07504, 2023.

[272] D. Cai, J. Heikkilä, and E. Rahtu, "Gs-pose: Cascaded framework for generalizable segmentation-based 6d object pose estimation," arXiv:2403.10683, 2024.

> D. Cai, J. Heikkilä, 和 E. Rahtu, "GS-Pose: 基于分割的通用 6D 物体姿态估计的级联框架," arXiv:2403.10683, 2024.

[273] S. Saito, G. Schwartz, T. Simon, J. Li, and G. Nam, "Relightable gaussian codec avatars," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2024, pp. 130-141.

> S. Saito, G. Schwartz, T. Simon, J. Li, 和 G. Nam, "可重光照的高斯编码头像 (Relightable Gaussian Codec Avatars)," 载于 IEEE/CVF 计算机视觉与模式识别会议论文集, 2024, 第 130-141 页.

[274] L. Franke, D. Rückert, L. Fink, and M. Stamminger, "Trips: Trilinear point splatting for real-time radiance field rendering," in Comput. Graph. Forum. Wiley Online Library, 2024, p. e15012.

L. Franke, D. Rückert, L. Fink, 和 M. Stamminger, "TRIPS: 用于实时辐射场渲染的三线性点绘制 (Trilinear Point Splatting)," 载于计算机图形论坛 (Comput. Graph. Forum)，Wiley 在线图书馆, 2024, 页码 e15012.