



研究与开发

基于神经表达的动态场景重建综述

陆平¹, 孙俊杰², 李婉³, 林家欣³, 刘枫⁴, 冯大权^{1,3,5}, 施文哲¹

(1. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518057;

2. 国家无线电监测中心检测中心, 北京 100041;

3. 广东省数字创意技术工程实验室, 广东 深圳 518060;

4. 中国信息通信研究院泰尔终端实验室, 北京 100083;

5. 广东省哲学社会科学重点实验室 (文化数字化与文化创新发展), 广东 深圳 518060)

摘要: 动态三维场景重建在计算机视觉和虚拟现实领域中具有重要的研究价值。近年来, 得益于神经表达技术的快速发展, 动态场景重建技术也取得了显著突破。过去4年中, 基于神经辐射场和3D高斯泼溅的动态场景重建方法被相继提出, 并取得了令人瞩目的效果。然而, 由于相关文献的数量庞大, 研究者很难全面跟踪所有相关工作。为了应对这一挑战, 基于神经表达技术在动态场景重建中的发展历程, 系统总结了该领域的代表性研究, 总结归纳为基于神经辐射场和3D高斯泼溅两大类。此外, 还介绍了当前动态三维场景重建的代表性数据集, 并总结了常用的算法评估标准。最后, 分析了当前动态三维场景重建技术面临的主要挑战, 并展望了未来的发展趋势。

关键词: 动态重建; 神经表达; 神经辐射场; 3D高斯泼溅

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025152

A review of dynamic scene reconstruction based on neural representation

LU Ping¹, SUN Junjie², LI Wan³, LIN Jiaxin³, LIU Hong⁴, FENG Daquan^{1,3,5}, SHI Wenzhe¹

1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518057, China

2. The State Radio Monitoring_center Testing Center, Beijing 100041, China

3. Guangdong Province Engineering Laboratory for Digital Creative Technology, Shenzhen 518060, China

4. China Telecommunication Technology Labs, Beijing 100083, China

5. Guangdong Provincial Key Laboratory of Philosophy and Social Sciences, Shenzhen 518060, China

Abstract: Dynamic scene reconstruction holds significant research value in the fields of computer vision and virtual reality. Recent advancements in neural representation technologies have facilitated rapid progress in this task. Over

收稿日期: 2025-02-25; 修回日期: 2025-05-29

通信作者: 李婉, liwan2022@email.szu.edu.cn

基金项目: 中兴通讯研究基金资助项目 (No. 2023ZTE03-04)

Foundation Item: ZTE Industry-University-Institute Cooperation Funds (No. 2023ZTE03-04)

the past four years, methods based on neural radiance fields and 3D Gaussian splatting have been proposed, achieving remarkable results. However, the large number of literature presents a challenge for individuals to comprehensively track comprehensive relevant works. To address this issue, typical work for dynamic scene reconstruction based on neural representation was summarized, categorizing them into methods based on neural radiance fields and 3D Gaussian splatting. Furthermore, representative datasets were highlighted and common evaluation metrics for algorithms were summarized. Finally, the persistent challenges in current methodologies were discussed and potential directions for future development trends were proposed.

Key words: dynamic reconstruction, neural representation, neural radiance field, 3D Gaussian splatting

0 引言

三维动态场景重建是计算机视觉和机器人技术中的一个重点研究课题,在自动驾驶、虚拟现实、增强现实及数字媒体等领域具有重要应用价值。该技术通过场景的实时感知与特征建模,为自动驾驶系统的行车跟踪、增强现实(AR)/混合现实(MR)的虚实交互以及影视的特效融合提供了重要技术支撑。尽管现有方法已经取得了显著进展,但在动态环境中实现高精度和实时性仍然面临诸多挑战。

传统的动态场景重建方法主要基于显式几何表征(如点云、网格和体素)的运动跟踪与预测。这类方法在处理刚性物体的变化时具有显著优势,尤其适用于拓扑结构保持不变的场景元素,如车辆、机械等。通过对物体的点、边、面或体素建立精确的几何运动模型,能够有效地描述和预测刚性物体在时空维度上的连续位移变化。然而,当面对复杂动态场景,如火焰、云朵等非刚性物体时,其缺乏固定的几何形状且运动模式具有高度非线性和随机性,基于显式几何表征的方法无法适应拓扑结构的剧烈变化,且难以准确捕捉其复杂的形变过程,从而导致重建失败。因此,如何高效地重建这类物体的动态变化,是传统方法所面临的一大挑战。近年来,随着神经表达学习的快速发展,神经辐射场(neural radiance field, NeRF)^[1]等隐式表征技术为动态场景重建提供了新的解决思路。这类技术通过

连续的神经网络参数化场景表示,能够自然地建模复杂的光场变化和几何形变,为解决传统方法面临的挑战提供了潜在的技术路径。NeRF利用深度神经网络预测均匀采样点的颜色、密度和变形,并通过体积渲染技术重建整个场景。相较于显式方法,隐式方法能够更好地捕捉细腻的静态细节和动态变化,显著提升了重建的视觉质量。然而,基于全局采样的NeRF及其变体在提高视觉质量的同时,也带来了高昂的计算成本和较长的处理时间。尽管已有一些加速技术得到应用^[2-5],动态NeRF在实时应用中的局限性仍然存在,特别是在需要维持每秒60帧的最低可接受交互帧率场景,如虚拟现实系统、实时三维视频会议、移动端增强现实等,难以同时满足实时性和高保真度的双重要求。为克服这一限制,近期的研究开始探索结合显式和隐式表征方法的新途径。例如,基于体素^[6]、低秩张量^[7]等显式表征形式为几何计算提供了坚实的基础,而深度神经网络强大的建模能力则使得这些显式结构能够精准地模拟复杂的动态变化。

在此基础上,3D高斯泼溅(3D Gaussian splatting, 3DGS)^[8]技术因其展示的巨大潜力而引起广泛关注。该技术利用一组三维高斯函数来表达和渲染动态场景,并与高度定制化的渲染器结合使用,提供了一种高效的场景表示方法。最新研究进展已将3DGS应用于包含复杂形变的动态场景建模领域,充分展现了其在几何快速处理和动态特征捕捉方面的独特优势。其核心技术在



于通过高斯函数精确模拟物体表面在不同光照条件下的反射特性，并在迭代优化过程中基于场景几何分布动态调整高斯参数，从而在保证渲染质量的同时显著提升计算效率。值得关注的是，3DGS技术自2023年被提出以来仍处于快速发展阶段，现有动态3DGS方案在保持时序一致性方面仍存在明显局限，表现为轨迹漂移现象、快速运动重建失败等^[9-12]。尽管现有监督机制显著提升了重建可靠性，但在动态场景的建模精度和泛化能力方面仍有较大的优化空间。

当前，国内外关于隐式重建的综述研究呈现出明显的两极分化态势：文献[13-14]等综述，或宽泛地覆盖整个隐式重建领域，或仅聚焦于NeRF或3DGS其中之一。值得关注的是，针对动态场景隐式重建这一重要分支的系统性总结与评述仍近乎空白。因此，本文系统性地梳理并对比分析了传统显示动态重建、神经辐射场隐式重建、3D高斯泼溅显-隐重建3类技术的核心原理与优劣势，动态三维重建技术发展历程总述见表1。后续章节重点围绕神经辐射场和3D高斯泼溅两大前沿技术展开了深入探讨，以时间轴为主线详细阐述了动态NeRF的技术演进历程，并基于现有技术路径对动态3DGS技术的发展脉络进行了剖析。此外，本文还系统调研了动态场景重建领域的常用基准数据集，并对当前主流重建算法的

量化评估体系进行了全面总结。最后，本文针对复杂场景的动态三维重建技术提出了未来可能的研究方向与发展趋势。

1 基于神经辐射场的动态场景重建方法

1.1 神经辐射场的基本原理

Mildenhall等^[1]提出了一种新型的隐式几何表征方法—NeRF。该方法仅使用RGB图像作为监督信息，能精确拟合高分辨率的几何形状隐函数，并实现复杂场景的高质量视点合成。与传统的显式几何基元不同，NeRF将场景信息以参数化形式进行编码，并利用多层感知机（multilayer perceptron, MLP）拟合表示参数，神经辐射场表达方法如图1所示。

NeRF将三维场景建模为一个连续且可微分的辐射场 F_θ ：

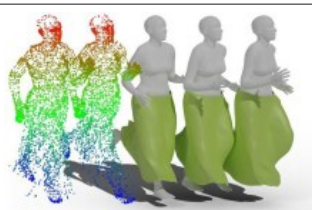
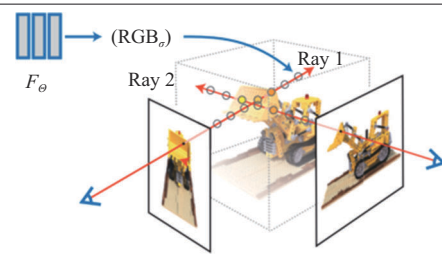
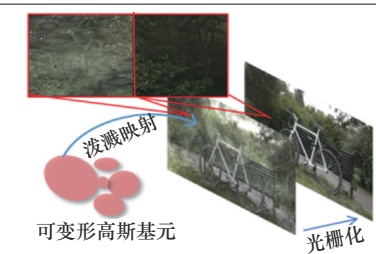
$$F_\theta(\mathbf{x}, \mathbf{d}) = [\sigma, c] \quad (1)$$

其中， $\mathbf{x}$ 为采样点的空间位置信息 (x, y, z) ， $\mathbf{d}$ 表示观测方向 (d_x, d_y, d_z) ， θ 表示为一组可学习的场景信息参数， σ 表示为光线到达采样点 $\mathbf{x}$ 处的概率， c 表示为采样点的颜色。因此，从相机位置 $\mathbf{o} \in \mathbb{R}^3$ 出发，光线 $\mathbf{l}$ 与成像平面上的像素点相连得到观测方向 $\mathbf{d} \in \mathbb{R}^3$ ，该光线可表示为：

$$\mathbf{l}(t) = \mathbf{o} + t\mathbf{d}, t \in (-\infty, +\infty) \quad (2)$$

像素点颜色由该光线上每个采样点的颜色值

表1 动态三维重建技术发展历程总述

传统显示动态重建（1990年—）	神经辐射场隐式重建（2020年—）	3D高斯泼溅显-隐重建（2023年—）
原理  通过跟踪几何体的运动状态重建动态物体	 利用神经辐射场预测采样点运动形变，结合体渲染技术进行重建	 采用可变形的高斯函数，通过光栅化技术融合映射图层
优势 刚性重建稳定；理论基础坚实	重建对象不受限；模型鲁棒性强	内存利用率高；渲染速度快
劣势 具有透明度的流体运动重建效果差	训练时间长、渲染速度慢	理论发展阶段、精度不足

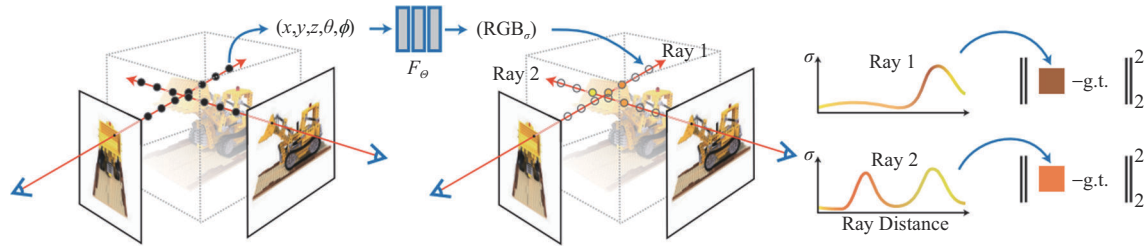


图1 神经辐射场表达方法

通过体渲染进行加权累积，成像颜色 C 表示为：

$$C = \int_{t_n}^{t_f} W(t) \sigma(l(t)) c(l(t), d) dt \quad (3)$$

其中， $W(t)$ 表示为光线累积量：

$$W(t) = \exp\left(-\int_{t_n}^t \sigma(l(t)) dt\right) \quad (4)$$

可微分的体渲染过程可以与神经网络的训练流程结合，使训练仅依赖 RGB 作为监督。此外，为了避免合成图像中高频信息的丢失，NeRF 在输入变量上引入了位置编码技术：

$$x \rightarrow [\sin(Mx), \cos(Mx)] \quad (5)$$

其中，

$$M = [2^0 \pi \ 2^1 \pi \ 2^2 \pi \ \dots \ 2^{p-1} \pi]^T \quad (6)$$

其中， p 代表一个控制使用频率的超参数，其值取决于目标信号的分辨率。能够为神经网络提供结构化特征表示，从而增强其对输入数据高频分量的捕捉能力。

1.2 动态神经辐射场方法

NeRF 最初被应用于静态场景，自 2021 年一些研究将其拓展到动态领域。动态 NeRF 的关键是如何充分利用时间信息。现有处理方法主要可

分为两类：第一类采用直接建模，将时间变量作为 NeRF 的显式输入，通过神经网络直接计算目标时刻的运动状态，这类方法虽然直观高效，但通常需要引入复杂的正则化约束来维持时空一致性；第二类则采用解耦建模，通过构建独立的变形场将时间因素与 NeRF 分离，间接计算随时间变化的形变量，这种架构设计不仅实现了刚性物体运动与拓扑变化的独立操作，还显著提升了模型的可解释性和可控性。本文对 2021 年至 2024 年间动态 NeRF 的代表性研究进行了总结，动态 NeRF 代表性方法的发展脉络如图 2 所示，直观地反映了其在不同发展阶段的研究热度。

2021 年是动态 NeRF 的奠基之年，多项开创性工作为后续研究奠定了基础。其中，NSFF (Neural Scene Flow Fields)^[15] 率先提出了神经场景流场，该方法将时空数据作为直接输入，利用时间变化的连续函数来预测场景的几何和运动属性。D-NeRF^[16] 则将时间视为额外输入，将时变场景运动统一编码至规范空间，通过可学习变形

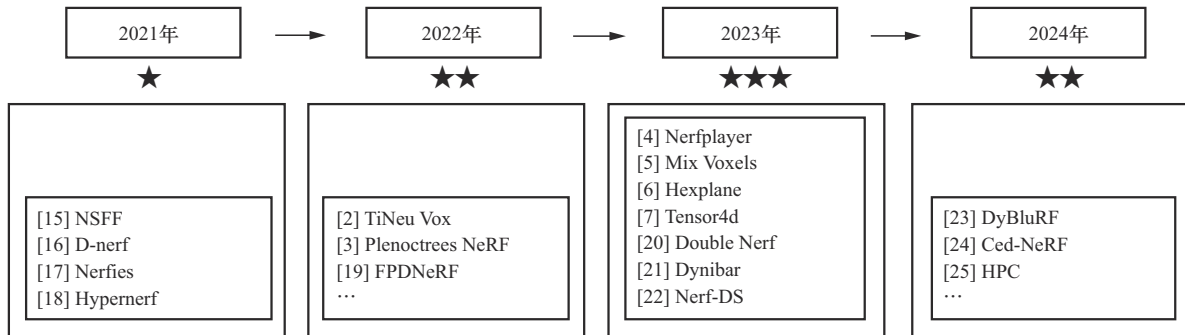


图2 动态NeRF代表性方法的发展脉络



场将规范表示映射到各时间。在此基础上, Nerfies^[17]提出在变形场中将每个观察点扭曲为规范的5D-NeRF, 并使用弹性正则化提升重建精度。HyperNeRF^[18]则将每个单独的输入图像对应的5D辐射场表示为“超空间”的表面切片, 有效解决拓扑变化建模的局限性。这一开创性工作因其出色的性能和通用性, 已成为当前该领域最具代表性的基准方法之一。

2022年是动态NeRF研究进入快速发展的关键阶段, 研究重心明显转向性能优化与效率提升。这一时期的创新工作主要围绕速度瓶颈展开。TiNeuVox^[2]提出了一种基于时间感知体素特征的辐射场框架, 通过一个小型坐标变形网络来模拟粗略的运动轨迹, 并通过多距离插值方法来模拟不同范围的运动, 其训练时间仅需8 min, 且仅占用8 MB存储空间。这一成果为后续研究树立了新的效率标杆。Plenotrees NeRF^[3]通过将NeRF转换为PlenOctree并通过封闭形式的球形基函数分解外观, 实现了每秒超过150帧的渲染速度。FPDNeRF^[19]采用空间混合生成树结构, 将运动表示为具有时变密度和颜色属性的傅里叶系数, 从而使训练速度比原始NeRF快了3 000倍。

2023年, 动态神经重建的研究迅速发展, 涌现出许多创新的动态NeRF方法。在时效提升方面, NeRFPlayer^[4]将场景划分为静态、动态和未见区域, 并为每个区域单独设计神经场表达和规范化方法, 实现每帧10 s的重建和交互式渲染。MixVoxels^[5]使用轻量级模型处理静态区域, 并使用内积时间查询方法高效地查询动态体素, 训练时间仅需15 min。Hexplane^[6]通过融合来自6个时空平面的向量来计算时空中点的特征, 并结合小型MLP进行颜色预测, 训练时间缩短超百倍。在重建精度方面, Tensor4d^[7]将动态场景表达为四维时空张量, 并将四维时空张量投影到时间感知的体积中, 采用分层分解方法, 实现了单目高质量动态重建。Double DANOs^[20]提出了一

种可微分的物理模拟管道, 通过参数化NeRF的密度场, 计算由碰撞产生的正向和摩擦力, 显著提升了动态场景的物理一致性。Dynibar^[21]采用基于体积的图像渲染框架, 通过以场景运动感知的方式从附近视图聚合特征合成新视点, 实现了在高动态和野外场景高质量渲染。NeRF-DS^[22]针对镜面物体, 计算NeRF在观测空间中的表面位置和方向, 从而在映射到标准空间时保持不同反射颜色。

NeRF可以有效提高新视图合成结果的质量, 但在动态高分辨率(>1 K)下, 依然存在训练时间长(>10 h)、渲染速率慢(<1 FPS)的问题, 难以在消费级GPU上实现长序列的快速训练(1~2 h)和高渲染速率(>30 FPS)。随着3DGS技术在差分光栅器的快速收敛上展现出的优势, 2024年学术界逐渐将重点从基于全局采样的NeRF方法转向基于稀疏点云的3DGS方法。然而, 仍有许多具有创新性的NeRF工作值得关注。例如, DyBluRF^[23]提出了针对运动模糊的解决方案, 通过捕获相机轨迹和场景物体的离散余弦变换轨迹, 采用全局跨时间的渲染方法, 确保了场景在时间维度上的连贯性。Ced-NeRF^[24]通过网格表达的编码特征正则化损失, 解决了动态场景中混合表达引发的过拟合问题。HPC^[25]提出了具有多分辨率残差辐射场的分层表达, 减少长序列中的时间冗余。

动态神经辐射场技术经过4年的快速发展, 已从最初的静态场景拓展到复杂的动态场景重建领域。本节梳理了2021—2024年该技术的演进历程: 2021年的奠基性工作确立了时空建模的基本框架; 2022年的研究重点转向效率优化, 通过数据交叉表征和特征压缩显著提升了计算性能; 2023年则涌现出多模态融合、物理模拟等创新方向; 2024年虽受3DGS技术冲击, 但在运动模糊处理、稀疏视图重建等细分领域仍取得突破。未来研究应着力解决3个关键问题: (1) 如何构建

更高效的时空混合表达以平衡质量与速度；(2) 如何增强模型在复杂光照和极端运动条件下的鲁棒性；(3) 如何实现长序列动态场景的一致性建模。这些突破将推动动态 NeRF 技术在数字孪生、虚拟制作等领域的深度应用，并为三维视觉提供核心技术支持。

2 基于 3DGS 的动态场景重建方法

2.1 3DGS 的基本原理

3DGS 技术^[8]保留了神经渲染的优点，采用一种以点云为初始化的球面谐波空间加速结构，能够在分钟级时间内，产生与基于 NeRF 的高级框架相当的结果。

3DGS 表达框架如图 3 所示，3DGS 技术将场景表达为一组各向异性高斯函数，仅针对输入图像即可实现高质量、快速合成。具体来说，场景表示为高斯函数集合，每个高斯函数包含均值 $\mathbf{x}_i$ 、协方差矩阵 Σ_i 、不透明度 α_i 以及和视角相关的颜色 c_i 属性。其中，协方差矩阵 Σ_i 形式为：

$$\Sigma_i = R_i \text{diag}(s_i) \text{diag}(s_i)^T R_i^T \quad (7)$$

参数 q_i 、 s_i 、 R_i 通过可微分光栅化的梯度下

降法，在给定一组已知相机内参 $\mathbf{K}$ 和外参 $\mathbf{V}$ 的图像输入下直接计算得到。图像渲染通过对像素 (u, v) 的相机射线按深度排序的可见高斯分布 $\{G_j\}_{j=1}^N$ 的贡献进行 α 混合，因此像素颜色表示为：

$$\mathbf{G}(u, v) = \sum_{j \in N} c_j \alpha_j p_j(u, v) \prod_{k=1}^{j-1} (1 - \alpha_k p_k(u, v)) \quad (8)$$

其中， $p_j(u, v)$ 表示像素 (u, v) 在第 j 个高斯的概率密度。在相机固有 $\mathbf{K}$ 和外参 $\mathbf{V}$ 的条件下，将 3D 高斯 $G_j = N(\mathbf{x}_j, \Sigma_j)$ 通过透视变换转化为 2D 高斯 $G'_j = N(\mathbf{x}'_j, \Sigma'_j)$ ，其均值和协方差分别表示为：

$$\mathbf{x}' = \mathbf{K} \mathbf{V} \mathbf{x}, \Sigma' = \mathbf{J} \mathbf{V} \Sigma \mathbf{V}^T \mathbf{J}^T \quad (9)$$

其中， $\mathbf{J}$ 为透视投影的雅可比矩阵。为了平衡渲染质量和计算效率，3DGS 使用自适应致密化，通过定期执行高斯函数的克隆和修剪适配整个场景。

2.2 动态 3D 高斯泼溅方法

将 3DGS 扩展到动态场景的关键是如何表达时间 T 。大多数基于 3DGS 的动态场景重建方法侧重于高斯的运动与随时间 T 的形变预测。现有基于动态 3DGS 的方法概述见表 2，总结了现有基于动态 3DGS 运动时间编码的类型，主要包括

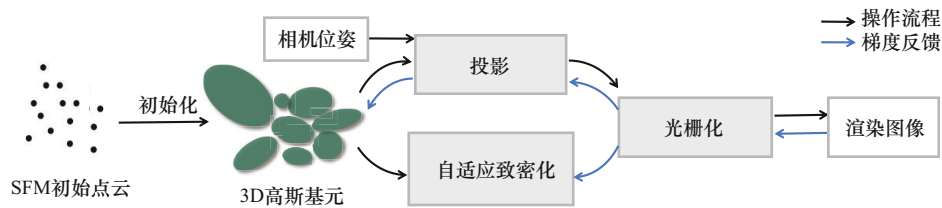


图3 3DGS 表达框架

表2 现有基于动态 3DGS 的方法概述

方法	侧重点	类型	方法	侧重点	类型
Dynamic 3D Gaussians ^[26]	Quality	Iterative	MoDGS ^[31]	Quality	3D+motion
3DGStream ^[27]	Quality	Iterative	SpaceTime-GS ^[32]	Quality	3D+motion
TC3DGS ^[39]	Memory	Iterative	GaGS ^[33]	Quality	3D+motion
QUEEN ^[40]	Memory	Iterative	D3DGS ^[34]	Quality	3D+motion
4DGS ^[9]	Quality	4D	GauFre ^[10]	Speed	3D+motion
RTGS ^[29]	Quality	4D	EffGS ^[35]	Speed	3D+motion
4D-Rotor GS ^[30]	Quality	4D	Gaussian-flow ^[36]	Speed	3D+motion
Deformable-GS ^[28]	Quality	3D+motion	BTimer ^[37]	Speed	3D+motion



Iterative、4D 和 3D+motion。Iterative 依赖基于时间迭代的形变预测，代表性方法包括 Dynamic 3D Gaussians^[26] 和 3DGStream^[27]。Dynamic 3D Gaussians^[26] 使用基于密集场景元素的六自由度（6-DOF）追踪，3DGStream^[27] 则采用了一个紧凑的神经变换缓存模拟平移和旋转。然而，这种迭代式的运动预测方法，在长序列下会产生积累误差，影响模型稳定性。

目前使用最多的方法是额外的变形场模拟运动。Deformable-GS^[28] 是最早基于 3DGS，使用 3D+motion 方式进行重建的方法，Deformable-GS 渲染管线如图 4 所示。该方法提出了一种可变形 3DGS 框架，将高斯函数映射到规范空间中并学习一个变形场来模拟单目场景运动。同时，引入了一个无额外开销的退火平滑训练机制，可以减轻时间插值任务的不准确姿态的影响。与此同时，4DGS^[9] 创新性地提出了使用神经体素编码高斯特征，随后一个轻量级 MLP 用于预测高斯形变。这两项开创性工作为动态 3DGS 奠定了重要基础，不仅成为该领域的基准方法，更为后续复杂动态场景建模、实时性能优化等维度的深入探索提供了新范式。

基于上述基础性工作，该领域已发展出两大主要技术分支。一方面，以 4DGS^[9] 为基础的技术路线，RTGS^[29] 通过优化一组具有显式几何和外观建模的几何体来近似场景的时空 4D 体素。4D-RotorGS^[30] 通过对各向异性的高斯函数进行时

间切片来模拟每个时间的运动状态。另一方面，基于 Deformable-GS^[28] 框架的 MoDGS^[31] 采用了单视图深度估计作为额外监督，并提出了一种基于 3D 感知初始化方法学习有效变形场，能够从单目视频中生成稳健的合成视角图像。为进一步提升重建质量，部分研究在场景表达层面进行融合创新。例如，SpaceTime-GS^[32] 融入了时空特征，通过嵌入时间不透明度和参数化的形变量，分别捕捉静态、动态以及瞬变内容。GaGS^[33] 将几何信息融入高斯渲染，通过在变形过程中强制执行场景几何约束，提升重建几何一致性。此外，另一些方法通过优化框架设计来提升精度。例如，D3DGS^[34] 改变了单一的基于坐标的变形场设计，采用高斯与时间的联合嵌入重新定义变形，并对粗细粒度运动进行层次建模。

一些方法针对速度进行了优化。EffGS^[35] 将位置和旋转定义为时间函数，同时保持静态高斯的其他属性时不变，实现了 1 352 dpi×1 014 dpi 分辨率下每秒 118 帧的最高渲染速度。Gaussian-Flow^[36] 将每个属性的时间依赖残差通过时间域中的多项式和频率域中的傅里叶级数拟合，使训练速度提高了 5 倍。不同于 4DGS^[9] 后向变形场，GauFR^[10] 采用时间依赖的前向预测场，能在 20 min 内完成训练。BTimer^[37] 则提出一个实时的运动感知前馈模型，汇总来自所有上下文帧的信息，在 150 ms 内实现给定目标时间的全景重建。

基于 3DGS 的方法通常使用由结构光或运动

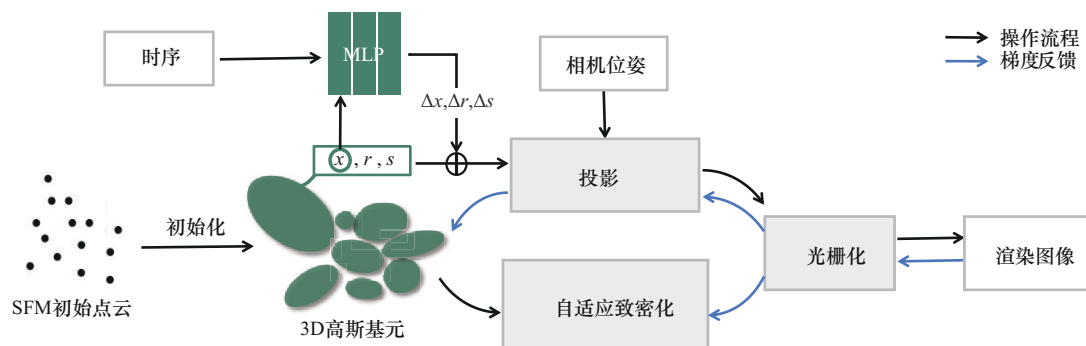


图4 Deformable-GS 渲染管线

恢复结构 (structure from motion, SfM) 算法计算得出的稀疏点云作为初始输入。近年来,研究者通过融合多源传感器数据,成功将 3DGS 拓展到更复杂的应用场景。以 LiHi-GS^[38] 为例,该方法创新性地融合 LiDAR 提供的稠密点云初始化高斯,不仅实现了高保真合成,还支持场景的实时交互编辑。这种融合路径使得 3DGS 技术能够应用于室内导航、增强现实 (AR)/虚拟现实 (VR) 场景构建等需要实时稠密建图的场景,其生成的高精度地图还可用于服务机器人路径规划等下游任务。此外,模型的紧凑性优化也是该领域的一个重要研究方向。TC3DGS^[39] 根据高斯的时间相关性有选择地修剪高斯,并采用梯度感知混合精度量化,最终实现了 67 倍的模型压缩率。QUEEN^[40] 直接学习连续帧之间的高斯属性残差,并通过可学习的门控模块来稀疏化位置残差,将模型的存储大小压缩至每帧仅 0.7 MB。

动态 3DGS 技术通过高效的点云表达和可微光栅化,显著提升了动态场景重建的速度与质量。现有方法主要围绕高斯运动与形变建模展开,形成了 3 种主流技术路线,其中 3D+motion 凭借其灵活性与性能优势成为当前主流,能够通过显式建模运动场实现高质量动态重建,同时保持实时渲染能力。然而,现有方法仍面临长序列误差累积、复杂运动建模不足以及初始点云依赖性强等挑战。未来研究可从以下方向

突破:(1) 运动建模优化,结合物理模拟或光流约束提升复杂动态(如流体、弹性体)的建模精度;(2) 跨模态初始化,利用 LiDAR、事件相机等多源数据增强初始点云质量,减少对 SfM 算法的依赖;(3) 轻量化与泛化,通过动态高斯剪枝和增量学习降低存储开销,适配动态大规模长序列场景;(4) 任务驱动扩展,探索动态 3DGS 在 SLAM、机器人交互等领域的应用潜力。随着进一步的创新,3DGS 有望成为动态场景重建的核心技术,为实时三维感知与生成提供高效解决方案。

3 动态场景数据集

神经辐射场 (NeRF) 与 3D 高斯泼溅 (3DGS) 技术在动态场景重建中的性能,高度依赖训练数据。单目数据集与多目数据集的根本差异在于时空信息获取方式:单目序列通过时间连续性推断几何变化,而多目阵列通过空间视角覆盖直接捕捉立体信息。然而,现有数据集在运动幅度(如大位移物体、小范围变形)、纹理复杂度(如高光、半透明物体)和采集配置(如相机数量、同步性)上的显著差异,导致不同方法泛化能力受限,当前主流动态场景重建数据集的典型训练样本可视化如图 5 所示。本节系统分析了单目数据集和多目数据集的特性与局限,为动态重建任务的数据选择与算法设计提供依据。



图5 当前主流动态场景重建数据集的典型训练样本可视化



3.1 单目数据集

对于单目数据集，序列变化对数据集内部和数据集之间的性能影响显著。与场景运动较大、摄像机运动较小的序列相比，场景运动较小或摄像机运动相对于物体运动大的序列重建效果更佳。同时，高纹理场景通常需要更多采样点。当前被 NeRF 和 3DGS 广泛使用的单目动态场景重建常用数据集见表 3。

D-NeRF^[16]：该数据集包含由 360° 轨道内向相机捕捉的合成物体运动，共有 8 个场景。背景为纯白色，场景突出表现复杂几何形状和具有非 Lambertian 纹理的物体缓慢运动，是动态场景三维重建中最常用且最经典的数据集之一。

Nerfies^[17]：该数据集使用 2 个刚性约束的摄像机，收集时间同步数据，涵盖真实世界非刚性变形场景，用于衡量在未见视角下场景重建的可信度。数据集包括两点。（1）由前置摄像头拍摄的毫秒级同步自拍序列，共计 5 个人类主体；（2）由后置摄像头录制的双序列同步视频，并使用 COLMAP 配准，包含人脸、动物和物体共计 4 个场景。训练集和验证集通常交替分配左、右视角，以避免未见区域对重建结果的影响。

HyperNeRF^[18]：该数据集是单目动态场景重建中使用最广泛的数据集之一，包含 3 个子集。（1）misc 子集，由单目摄像头采集的包含手部弯曲等来自真实世界的 7 个长序列视频；（2）interp 子集，由单目摄像头采集的包括厨房运动等非刚体变形共计 6 个场景，用于验证场景的时序一致性；（3）vrig 子集，包含由两个同步相机拍摄的

4 个场景，与 Nerfies^[5]一致，采用左、右相机交替重建验证新视角合成能力。该数据集存在大幅度的运动、复杂的光照条件和细小的物体结构，被认为是该领域极具挑战性的数据集之一。

NeRF-DS^[22]：该数据集是针对动态镜面反射重建的专用数据集，由 8 个日常环境中的场景组成，涵盖物体刚性移动和变形的高光物体。每个场景包含双序列视频，由两台同步摄像头拍摄。

DyCheck^[41]：该数据集使用 3 部同步 iPhone 拍摄，共计 7 个动态场景。场景内物体运动幅度较大，摄像机运动相对较小。在实验设置中，通常选取来自 1 部 iPhone 的图像训练，其他相机帧用于评估。

3.2 多目数据集

多目数据集是由多个摄像头从不同视角同时捕获的图像数据组成，通过空间分布的多视角固定相机采集系统，可构建具有立体信息的稠密观测数据。目前被 NeRF 和 3DGS 广泛使用的多目动态场景重建常用数据集见表 4。

Plenoptic Video^[11]：该数据集通过一个移动多视角捕捉系统，使用 21 台 GoPro Black Hero 7 摄像机进行视频捕捉，数据分辨率 2 028 dpi×2 704 dpi，帧率为 30 FPS。场景包含高光泽物体、半透明物体，具有变化拓扑的场景变化和运动，具有强视角依赖效应的纠缠运动物体，以及各种光照条件（如白天、夜晚、侧面聚光灯）。该数据集是多目动态场景重建领域中使用最广泛的数据集之一。

NVIDIA^[42]：该数据集通过一个相机阵列

表 3 单目动态场景重建常用数据集

数据集	场景数量	运动特性	采集设备	纹理/光照	挑战性
D-NeRF ^[16]	8	物体缓慢运动，复杂几何	合成数据	纯白背景	非刚性变形建模
Nerfies ^[17]	9	非刚性变形	双同步相机	自然光照	拓扑变化
HyperNeRF ^[18]	17	刚体/非刚体，大幅运动	双同步相机	自然光照	大幅度运动
NeRF-DS ^[22]	8	镜面反射物体运动	双同步相机	高光材质	镜面反射重建
DyCheck ^[41]	7	大物体运动，小相机运动	3 部同步 iPhone	自然光照	视角稀疏

表4 多目动态场景重建常用数据集

数据集	场景数量	运动特性	采集设备	纹理/光照	挑战性
Plenoptic ^[11]	21	拓扑变化, 高光/半透明物体	21台移动GoPro阵列	多光照条件	拓扑变化
NVIDIA ^[42]	8	人体/物体互动, 刚体运动	12台GoPro静态阵列	自然光照	多目标运动
Google Immersive ^[43]	11	动态街景	46鱼眼相机阵列	自然光照	大规模场景
Light-Field ^[44]	7	面部、中景动画	4×4相机阵列	剧场光照	窄基线视角合成
PanopticSports ^[45]	6	快速体育动作	31静态相机	室外运动场	高速运动与遮挡

(包含12台GoPro Black Edition摄像机)采集具有深度信息和掩码信息的真实场景,包括人的运动、交互运动、道路场景以及玩偶在不同位置逐帧拍摄的定格动画。数据集原始分辨率为1 920 dpi × 924 dpi,每个视角采集192帧的时间序列。

Google Immersive^[43]:包含由时间同步的46鱼眼摄像机装置捕获的各种室内和室外环境的光场视频,包括焊工、火焰、卡车、展品、面部彩绘、洞穴等。

Light-Field Video^[44]:该数据集提供了一组同步光场视频序列,使用4×4摄像机阵列捕捉,帧率为30 FPS,分辨率为2 048 dpi×1 088 dpi。数据集包括面部表情特写序列、中景角度的场景(如画家、生日场景)和来自人偶、剧院、火车的动画场景。

PanopticSports^[45]:该数据集使用31个静态摄像机收集,一般使用0、10、15和30号位相机作为测试数据。该数据集包含6个动态运动场景(杂耍、拳击、垒球、网球、足球和篮球)。每个场景的分辨率为640 dpi×360 dpi,共计150帧,帧率为30FPS。

动态场景重建的性能表现与数据集的时空特性密切相关。单目数据集通过时间序列的连续性捕捉动态变化,为评估算法在物体运动主导场景下的性能提供了重要基准;而多目数据集则凭借密集的空间采样显著丰富了数据特征,在动态拓扑建模和高速运动重建方面展现出独特优势,但其对设备同步精度和标定质量提出了更高要求。值得注意的是,两类数据集在运动幅度、纹理复杂度以及采集配置等方面的显著差异,直接决定

了算法模型的泛化能力。通过深入理解不同数据集的特性优势,并结合合理且有效的评估手段,将有效推动该研究领域的持续发展。

4 动态场景评估方法

4.1 评估方法

在场景重建和视图合成的研究中,常见的任务类型包括模型拟合、时间插值和新视角合成。模型拟合任务旨在评估模型在精确捕捉和再现场景细节特征方面的能力,通常通过将训练数据集用作验证集来反映模型对训练场景的学习效果,从而直接评估其在细节重建上的表现。时间插值任务则聚焦于评估模型保持场景结构一致性的能力。该任务要求从相机捕获的图像序列中生成视觉连贯的中间视图,以检验模型对场景深度、几何结构的理解及其在时间轴上的插值准确性。而新视角合成任务则涉及使用已有视角的图像或视频数据,在任意时刻和视角下生成高质量的图像。该任务的挑战在于,模型不仅需精确重建三维场景的空间结构,还需在时间维度上保持逻辑一致性和连贯性。因此,新视角合成任务被认为是目前最常见的评估方法之一,被广泛应用于动态场景和视角重建的评估中。

4.2 量化指标

基于NeRF和3DGS的方法中,场景的表达采用参数形式,因此其量化评估一般在二维像素空间中进行。常用指标主要包括峰值信噪比(peak signal-to-noise ratio, PSNR)、结构相似性指数(structural similarity index, SSIM)和学习



感知度量 (learned perceptual image patch similarity, LPIPS)。其中 PSNR 和 SSIM 反映了像素级误差, 表达为:

$$\text{PSNR} = 10 \cdot \lg\left(\frac{\text{MAX}^2}{\text{MSE}}\right) \quad (10)$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\gamma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\gamma_x^2 + \gamma_y^2 + C_2)} \quad (11)$$

其中, μ 和 γ 分别表示均值和标准差。PSNR 和 SSIM 值越高, 表明渲染图像的像素质量越好。LPIPS 则用于衡量渲染图像与真值图像之间的人眼感知误差, 通常基于 VGG 网络训练的特征来计算, 数值越低表明图像的主观感知质量越高。值得注意的是, PSNR 和 SSIM 指标倾向于偏好模糊图像, 因此可能存在 PSNR 和 SSIM 指标较高而 LPIPS 指标不佳的情况。因此, LPIPS 值应被视为更具决定性的图像质量评价指标。

在性能评估方面, 重建速度和渲染速度是衡量效率的关键指标, 分别通过模型训练所需时间及渲染帧率 (frames per second, FPS) 来量化。此外, 运行内存的使用情况反映了模型的紧凑性。对于基于 3DGS 的方法, 通过统计所需的高斯基元数量, 可以评估模型的潜在冗余性。这些综合性指标为模型在实际应用中的可行性和效能提供了重要参考依据。

5 结束语

本文系统梳理了动态场景重建领域的研究进展, 重点分析了基于神经辐射场与 3D 高斯泼溅技术的代表性方法。与基于神经辐射场的全参表达方法相比, 3DGS 通过显式表征和可微分渲染管线, 在训练效率 (提升 10~100 倍) 和渲染速度 (达到实时级) 方面实现了突破性进展。尽管在过去一年中取得了令人瞩目的成果, 但实现高精度实时的复杂场景动态重建仍面临以下挑战。

(1) 大规模数据处理: 特征密集场景, 或来自多摄像头、多传感器的海量数据, 对存储、传输和一致性管理提出了高要求。(2) 高速场景的捕获和重建: 车辆快速移动等高速场景易出现纹理模糊和建模错误, 且连续帧的显著位置变化增加了重建难度。(3) 实时性能要求。(4) 遮挡和视角限制: 随着应用场景复杂化, 未来还将面临多模态数据时空误差随数据量级呈指数增长, 极端光照条件下的材质估计偏差导致表面散射等非朗伯体效应, 长序列拓扑结构变化的累计误差超过可接受阈值等诸多新挑战。

针对当前及未来隐式动态场景重建技术的关键挑战, 本文研究从以下维度探讨潜在突破路径。(1) 设计高效混合表征网络, 探索超越传统点基元的高质量显式表征与快速隐式计算的融合方法, 以协同优化计算效率与几何细节表达能力。(2) 基于 Transformer、2D/3D 光流、物理引擎等, 设计时空自适应建模模块, 增强动态场景下基元对极端运动与复杂光照的鲁棒表征能力。(3) 优化资源管理和计算效率, 开发高效算法或结合云计算资源, 满足实时动态重建的需求。

希望本文探讨能为未来研究提供有益参考, 助力动态场景重建技术向更高水平发展。

参考文献:

- [1] MILDENHALL B, SRINIVASAN P P, TANCIK M, et al. NeRF: representing scenes as neural radiance fields for view synthesis[C]//Proceedings of the European Conference on Computer Vision (ECCV) 2020. Cham: Springer, 2020: 405-421.
- [2] FANG J, YI T, WANG X, et al. Fast dynamic radiance fields with time-aware neural voxels[C]//Proceedings of the SIGGRAPH Asia 2022 Conference Papers. New York: ACM Press, 2022: 1-9.
- [3] YU A, LI R, TANCIK M, et al. Plenotrees for real-time rendering of neural radiance fields[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 5752-5761.
- [4] SONG L C, CHEN A P, LI Z, et al. Nerfplayer: a streamable dynamic scene representation with decomposed neural radiance fields[J]. IEEE Transactions on Visualization and Computer Graphics, 2023, 29(5): 2732-2742.

- [5] WANG F, TAN S N, LI X H, et al. Mixed neural voxels for fast multi-view video synthesis[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2023: 19706-19716.
- [6] CAO A, JOHNSON J. Hexplane: a fast representation for dynamic scenes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 130-141.
- [7] SHAO R Z, ZHENG Z R, TU H Z, et al. Tensor4d: efficient neural 4D decomposition for high-fidelity dynamic reconstruction and rendering[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 16632-16642.
- [8] KERBL B, KOPANAS G, LEIMKÜHLER T, et al. 3D Gaussian splatting for real-time radiance field rendering[J]. ACM Transactions on Graphics, 2023, 42(4).
- [9] WU G J, YI T R, FANG J M, et al. 4D Gaussian splatting for real-time dynamic scene rendering[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 20310-20320.
- [10] LIANG Y Q, KHAN N, LI Z Q, et al. GauFR: Gaussian deformation fields for real-time dynamic novel view synthesis[C]//Proceedings of the 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE Press, 2025: 2642-2652.
- [11] LI T, SLAVCHEVA M, ZOLLHOEFER M, et al. Neural 3D video synthesis from multi-view video[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 5521-5531.
- [12] LI W, PAN X, LIN J X, et al. FRPGS: fast, robust, and photorealistic monocular dynamic scene reconstruction with deformable 3D Gaussians[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025.
- [13] 李吉洋, 程乐超, 何靖璇, 等. 神经辐射场的研究现状与展望[J]. 计算机辅助设计与图形学学报, 2024, 36(7): 995-1013.
LI J Y, CHENG L C, HE J X, et al. Current status and prospects of research on neural radiance fields[J]. Journal of Computer-Aided Design & Computer Graphics, 2024, 36(7): 995-1013.
- [14] WU T, YUAN Y J, ZHANG L X, et al. Recent advances in 3D Gaussian splatting[J]. Computational Visual Media, 2024, 10(4): 613-642.
- [15] LI Z, NIKLAUS S, SNAVELY N, et al. Neural scene flow fields for space-time view synthesis of dynamic scenes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 6498 - 6508.
- [16] PUMAROLA A, CORONA E, PONS-MOLL G, et al. D-NeRF: Neural radiance fields for dynamic scenes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 10318-10327.
- [17] PARK K, SINHA U, BARRON J T, et al. Nerfies: deformable neural radiance fields[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 5865-5874.
- [18] PARK K, SINHA U, HEDMAN P, et al. HyperNeRF: a higher-dimensional representation for topologically varying neural radiance fields[J]. ACM Transactions on Graphics, 2021, 40(6): 238.
- [19] WANG L, ZHANG J, LIU X, et al. Fourier plenotrees for dynamic radiance field rendering in real-time[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 13524-13534.
- [20] LE CLEAC'H S, YU H X, GUO M, et al. Differentiable physics simulation of dynamics-augmented neural objects[J]. IEEE Robotics and Automation Letters, 2023, 8(5): 2780-2787.
- [21] LI Z Q, WANG Q Q, COLE F, et al. Dynibar: neural dynamic image-based rendering[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 4273-4284.
- [22] YAN Z W, LI C, LEE G H. NeRF-DS: neural radiance fields for dynamic specular objects[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 8285-8295.
- [23] BUI M Q, PARK J, OH J, et al. DyBluRF: dynamic neural radiance fields from blurry monocular video[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 7517-7527.
- [24] LIN Y. Ced-NeRF: a compact and efficient method for dynamic neural radiance fields[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Los Altos: AAAI, 2024, 38(4): 3504-3512.
- [25] ZHENG Z H, ZHONG H Q, HU Q, et al. HPC: hierarchical progressive coding framework for volumetric video[C]//Proceedings of the 32nd ACM International Conference on Multimedia. New York: ACM Press, 2024: 7937-7946.
- [26] LUITEN J, KOPANAS G, LEIBE B, et al. Dynamic 3D gaussians: tracking by persistent dynamic view synthesis[C]//2024 International Conference on 3D Vision (3DV). Piscataway: IEEE Press, 2024: 800-809.
- [27] SUN J K, JIAO H, LI G Y, et al. 3DGStream: on-the-fly training of 3D Gaussians for efficient streaming of photo-realistic



- free-viewpoint videos[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 20675-20685.
- [28] YANG Z Y, GAO X Y, ZHOU W, et al. Deformable 3D Gaussians for high-fidelity monocular dynamic scene reconstruction[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 20331-20341.
- [29] YANG Z Y, YANG H Y, PAN Z J, et al. Real-time photorealistic dynamic scene representation and rendering with 4D Gaussian splatting[EB]. 2024.
- [30] DUAN Y X, WEI F Y, DAI Q Y, et al. 4D-Rotor Gaussian splatting: towards efficient novel view synthesis for dynamic scenes[C]//Proceedings of the ACM Press SIGGRAPH 2024 Conference Papers. Denver: ACM Press, 2024: 1-11.
- [31] LIU Q M, LIU Y, WANG J P, et al. MoDGS: dynamic Gaussian splatting from casually-captured monocular videos[EB]. 2024.
- [32] LI Z, CHEN Z, LI Z, et al. Spacetime Gaussian feature splatting for real-time dynamic view synthesis[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 8508-8520.
- [33] LU Z C, GUO X, HUI L, et al. 3D geometry-aware deformable Gaussian splatting for dynamic view synthesis[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 8900-8910.
- [34] BAE J, KIM S, YUN Y, et al. Per-Gaussian embedding-based deformation for deformable 3D Gaussian splatting[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 321-335.
- [35] KATSUMATA K, VO D M, NAKAYAMA H. A compact dynamic 3D Gaussian representation for real-time dynamic view synthesis[EB]. 2023.
- [36] LIN Y T, DAI Z Z, ZHU S Y, et al. Gaussian-flow: 4D reconstruction with dynamic 3D gaussian particle[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 21136-21145.
- [37] LIANG H X, REN J W, MIRZAEI A, et al. Feed-forward bullet-time reconstruction of dynamic scenes from monocular videos[EB]. 2024.
- [38] KUNG P C, ZHANG X, SKINNER K A, et al. LiHi-GS: LiDAR-supervised gaussian splatting for highway driving scene reconstruction[EB]. 2024.
- [39] JAVED S, KHAN A J, DUMERY C, et al. Temporally compressed 3D Gaussian splatting for dynamic scenes[EB]. 2024.
- [40] GIRISH S, LI T Y, MAZUMDAR A, et al. QUEEN: quantized efficient encoding of dynamic Gaussians for streaming free-viewpoint videos[EB]. 2024.
- [41] GAO H, LI R, TULSIANI S, et al. Monocular dynamic view synthesis: a reality check[J]. Advances in Neural Information Processing Systems, 2022, 35: 33768-33780.
- [42] YOON J S, KIM K, GALLO O, et al. Novel view synthesis of dynamic scenes with globally coherent depths from a monocular camera[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 5336-5345.
- [43] BROXTON M, FLYNN J, OVERBECK R, et al. Immersive light field video with a layered mesh representation[J]. ACM Transactions on Graphics (TOG), 2020, 39(4): 1-86.
- [44] SABATER N, BOISSON G, VANDAME B, et al. Dataset and pipeline for multi-view light-field video[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2017: 30-40.
- [45] JOO H, LIU H, TAN L, et al. Panoptic studio: a massively multiview system for social motion capture[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 3334-3342.

[作者简介]



陆平 (1971-), 男, 博士, 移动网络和移动多媒体技术国家重点实验室副主任、正高级工程师, 东南大学、深圳大学博士生导师, 中兴通讯股份有限公司副总裁, 主要研究方向为云计算、增强现实、媒体大数据。



孙俊杰 (1992-), 男, 博士, 国家无线电监测中心检测中心工程师, 主要研究方向为卫星互联网和数字媒体技术。



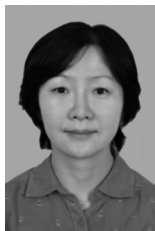
李婉 (1995-), 女, 深圳大学电子与信息工程学院博士生, 主要研究方向为计算机视觉、神经渲染和生成人工智能。



林家欣 (1996-), 男, 深圳大学电子与信息工程学院博士生, 主要研究方向为隐式表面表示和重建、几何计算和衣着人体。



冯大权 (1986-), 男, 博士, 深圳大学电子与信息工程学院副教授、博士生导师, 广东省哲学社会科学重点实验室(文化数字化与文化创新发展)副主任, 主要研究方向为生成式人工智能、沉浸式通信、VR/AR。



刘飒 (1973-), 女, 现就职于中国信息通信研究院泰尔终端实验室, 主要研究方向为人工智能、区块链、网络优化等。



施文哲 (1989-), 男, 移动网络和移动多媒体技术国家重点实验室工程师, 中兴通讯 XRExplore 产品规划经理, 主要研究方向为室内视觉 AR 导航、三维重建、实时云渲染。