

中图法分类号: 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-14

论文引用格式: Qiu Wenan, Qian Wenhua, Liu Peng, Li Huaguang. Symplectic Dual-Path Decoupling 3D Gaussian Splatting for Motion Deblurring[J/OL]. Journal of Image and Graphics, XXXX:1-14. DOI: 10.11834/jig.260388. (丘文安, 钱文华, 刘朋, 李华光. 面向运动去模糊的辛几何双路解耦三维高斯泼溅[J/OL]. 中国图象图形学报, XXXX:1-14. DOI: 10.11834/jig.260388.) [DOI: 10.11834/jig.260388]

面向运动去模糊的辛几何双路解耦三维高斯泼溅

丘文安, 钱文华, 刘朋, 李华光

云南大学信息学院, 昆明 650500

摘要: 目的 运动模糊是图像采集过程中常见的退化现象, 严重影响着三维场景重建与新视角合成的质量。现有基于三维高斯泼溅的去模糊方法存在两个问题: 一是相机轨迹建模缺乏物理一致性保证, 二是采用单一变换难以精确描述复杂非匀速运动。方法 针对上述问题, 提出一种面向运动去模糊的辛几何双路解耦三维高斯泼溅方法。该方法引入哈密顿动力学框架, 将相机运动建模为相空间中的位置-动量对偶演化, 再利用辛几何微分方程求解器确保相流满足辛保持性质, 从物理本质上保证相机运动轨迹的连续性; 在此基础上, 提出一种物理先验与残差精修协同的双路自适应运动解耦框架, 将相机运动分解为提供全局位姿基线的刚体运动分支和补偿细微运动偏差与积分近似误差的残差精修分支, 两条支路通过自适应协同组合实现复杂非匀速运动的稳定表达; 结果 在 Deblur-NeRF 合成/真实数据集和 ExBluRF 极端模糊数据集上进行验证实验, 与主流方法相比, 在真实数据集和 ExBluRF 数据集上指标均达到最优, 在合成数据集上 PSNR 和 LPIPS 达到最优, 其中 PSNR 在三个数据集上较次优方法提升分别达 0.12 dB、1.75 dB、1.87 dB, 在复杂非匀速运动场景下优势尤为明显。结论 实验结果表明, 所提方法通过辛几何建模与双路解耦策略, 有效提升了运动模糊场景下的三维重建质量。

关键词: 三维高斯泼溅; 运动模糊; 辛几何; 运动解耦; 新视角合成

Symplectic Dual-Path Decoupling 3D Gaussian Splatting for Motion Deblurring

Qiu Wenan, Qian Wenhua, Liu Peng, Li Huaguang

School of Information Science and Engineering, Yunnan University, Kunming 650500, China

Abstract: **Objective** Motion blur is a common degradation in image acquisition that severely degrades three-dimensional scene reconstruction and novel view synthesis. In practice, motion blur arises from prolonged camera exposure or relative motion, causing loss of high-frequency details and geometric degradation. Existing deblurring methods based on neural radiance fields or 3D Gaussian Splatting have made notable progress. However, two critical limitations remain. First, camera trajectory modeling relies on discrete interpolation or standard neural ordinary differential equations, lacking physical consistency guarantees. Second, a single global transformation is adopted, which is insufficient to capture complex non-uniform motions such as curved trajectories and acceleration phases. This study aims to develop a motion deblurring 3DGS framework that ensures physically consistent trajectory modeling and enables precise representation of complex non-uniform motions. **Method** The proposed framework introduces two core technical contributions. First, a Hamiltonian dynamics framework based on symplectic geometry is established for continuous camera trajectory modeling. The camera state is represented as a position-momentum dual vector in phase space, coupled through the canonical symplectic structure. A symplectic neural derivative network parameterizes the Hamiltonian, and a symplectic Euler integrator performs numerical inte-

收稿日期: 2026-06-30; 修回日期: 2026-08-27

基金项目: 云南省基础研究面上项目(202501AT070219)

Supported by: the Basic Research Program-General Project of the Science and Technology Department of Yunnan Province(202501AT070219)

© 中国图象图形学报版权所有

gration by alternately updating position and momentum components. This alternating update preserves the symplectic structure at the discrete level, thereby guaranteeing trajectory continuity and long-term numerical stability. The camera pose is decoded from the position component via the Rodrigues formula and $SE(3)$ exponential map. Second, a dual-path adaptive motion decoupling framework is designed. This framework decomposes motion into two independent branches: a rigid motion branch that inherits the symplectic trajectory estimation to ensure physically consistent global camera motion, and a continuous motion refinement (CMR) branch that learns residual adjustments in an unconstrained space to compensate for discretized integration errors. An orthogonality regularization is applied to the CMR rotation matrix to maintain geometric plausibility, and the two branches are combined via matrix multiplication to produce final virtual camera poses. For deblurring rendering, a 3D frequency constraint filter is applied to Gaussian primitives to suppress aliasing artifacts. Multiple virtual cameras are sampled along the estimated trajectory, each rendering a sharp image via 3DGS. A pixel-level weighted fusion module and a blur-region mask network are introduced to adaptively synthesize the deblurred output. The experiments are conducted on the Deblur-NeRF synthetic dataset (5 Blender-rendered scenes with blur simulated in linear RGB space), the Deblur-NeRF real-world dataset (10 scenes captured with a Canon EOS RP camera), and the ExBluRF extreme blur dataset (8 scenes with challenging camera motions). Camera poses and initial point clouds are recovered using COLMAP. Three evaluation metrics are adopted: PSNR, SSIM, and LPIPS. Ablation experiments are designed to independently verify the contributions of the symplectic geometry module and the dual-path decoupling framework. **Result** The experiments are conducted on the Deblur-NeRF synthetic dataset, Deblur-NeRF real-world dataset, and ExBluRF extreme blur dataset. On the synthetic dataset, the proposed method achieves 30.74 dB in PSNR, 0.901 in SSIM, and 0.073 in LPIPS, outperforming all compared methods in PSNR and LPIPS. The PSNR improvement over the second-best method SMURF reaches 0.12 dB, while the LPIPS improvement over the second-best method Deblurring 3DGS is 0.032. On the real-world dataset, the proposed method achieves optimal performance across all three metrics: 28.45 dB in PSNR, 0.853 in SSIM, and 0.084 in LPIPS, surpassing all compared methods. On the ExBluRF extreme blur dataset, the proposed method maintains robust performance with 30.74 dB in PSNR, 0.752 in SSIM, and 0.307 in LPIPS, where the PSNR advantage over the second-best method Deblur-NeRF (28.87 dB) reaches 1.87 dB. The experimental results demonstrate that the proposed method achieves optimal performance on all evaluation metrics across the real-world and ExBluRF datasets, and optimal PSNR and LPIPS on the synthetic dataset, with particularly pronounced advantages in complex non-uniform motion scenarios. **Conclusion** The proposed framework successfully addresses the stated objectives by integrating symplectic geometry-based Hamiltonian dynamics with a dual-path adaptive motion decoupling strategy. The symplectic neural ordinary differential equation ensures physical consistency and numerical stability in trajectory estimation, while the dual-path decomposition enables accurate modeling of complex non-uniform motions. The method demonstrates superior reconstruction quality on standard and extreme blur benchmarks. The findings suggest that enforcing physical consistency in camera trajectory modeling substantially benefits motion deblurring in 3D scene reconstruction. However, the current method is designed for static scenes and has limited handling capability for dynamic foreground objects. Additionally, the high-dimensional symplectic neural derivative network introduces non-negligible computational overhead. Future work will focus on reducing computational complexity while preserving trajectory modeling accuracy, as well as extending the framework to dynamic scenes.

Key words: 3D Gaussian Splatting; motion blur; symplectic geometry; motion decoupling; novel view synthesis

论文引用格式:[DOI:10.11834/jig.260388]

0 引言

新视角合成(novel view synthesis, NVS)旨在根据已知视角的观测图像重建三维场景,并生成任意视角下的逼真图像,是计算机视觉与计算机图形学

领域的研究热点,在虚拟现实、增强现实以及三维内容生成等应用中具有重要价值。近年来,神经辐射场(neural radiance field, NeRF)(Mildenhall等,2020)的提出显著推动了新视角合成技术的发展,其通过隐式体积表示建模场景,在渲染质量上取得了突破性进展。然而,NeRF依赖体渲染过程,训练与渲染效率较低,为解决这一问题,Instant-NGP(Müller等,

© 中国图象图形学报版权所有

2022)通过多分辨率哈希编码将训练时间从数小时缩短至数秒,但仍难以满足实时应用的需求。随后,针对这一问题,三维高斯泼溅(3D Gaussian Splatting, 3DGS)(Kerbl 等, 2023)通过显式高斯表示与可微分光栅化渲染,在保持高质量渲染的同时实现了实时性能,成为神经渲染领域的重要发展方向。在实际场景采集过程中,由于相机曝光时间较长或拍摄时存在相对运动,图像往往会发生运动模糊。运动模糊会导致图像高频细节的丢失与几何结构的退化,严重制约了基于多视角输入的三维重建与新视角合成的质量(Hu 等, 2024)。传统图像去模糊方法主要在二维图像空间中进行建模,如基于卷积神经网络的方法 SRN-Deblur(Tao 等, 2018)或基于生成对抗网络的方法 DeblurGAN(Kupyn 等, 2018),尽管在单幅图像恢复方面取得了一定进展,但难以充分利用多视角数据中的几何一致性信息,也无法直接适配三维场景表示。

针对神经渲染中的运动模糊问题,已有研究尝试在渲染管线中显式建模模糊形成过程。Deblur-NeRF(Ma 等, 2022)首先将去模糊机制引入 NeRF 框架,通过学习模糊核来模拟曝光期间的图像退化。该方法假设模糊由相机运动引起的空间卷积过程产生,通过在训练时引入模糊核并在渲染时排除该核来生成清晰图像。然而,该方法仅关注模糊的最终效果,未对相机运动轨迹本身进行建模,限制了其在复杂运动场景下的表达能力。图像模糊可以归纳为非均匀模糊与复合模糊等多种类型,真实场景下去模糊性能提升的核心瓶颈在于对运动轨迹的精确建模(Wang 等, 2024)。为更好地捕捉相机运动特性,后续研究将相机运动显式建模为三维刚体运动。DP-NeRF(Lee 等, 2023)通过预测曝光开始与结束时的相机位姿,并假设匀速运动来插值中间位姿,以保持场景的几何一致性。尽管刚体建模能够在一定程度上刻画相机运动,但匀速假设在实际场景中往往难以成立。此外,束调整方法 BAD-NeRF(Wang 等, 2023)、BAD-Gaussians(Zhao 等, 2024)采用更为灵活的轨迹表示,通过联合优化相机轨迹与场景表示提升建模精度。这类方法通常使用 B 样条或线性插值来预测连续位姿,而在面对大幅度或非匀速运动时,离散插值容易产生位姿跳变,导致轨迹不平滑。为解决轨迹不连续的问题,近期研究尝试引入更连续的轨迹建模策略。SMURF(Lee 等, 2024)通过神经

常微分方程建模连续的动态过程,实现了比离散插值方法更平滑的轨迹估计。尽管如此,将物理守恒律显式引入相机轨迹建模的研究仍相对匮乏,在神经渲染框架下探索物理一致的轨迹建模是值得深入的方向(Xing 等, 2024)。

综上所述,这些方法仍面临两方面挑战。首先是连续轨迹的建模缺乏物理一致性保证,上述方法在建模相机轨迹时,或采用离散插值策略,或采用标准神经 ODE,都是依赖数据驱动的轨迹参数化,但均未从物理本质上对相机运动施加约束。相机轨迹应满足连续、平滑、几何一致的物理约束,即相机轨迹应在时间维度上连续无跳变、相机轨迹的速度与加速度符合物理规律、相机轨迹变换应保持场景的几何结构。现有建模框架未能结合物理规律,缺乏对上述约束的显式保证,导致复杂运动场景下轨迹估计的稳定性和物理一致性难以保证。其次是对复杂运动建模的能力有限,现有方法通常将运动模糊归因于相机刚体运动,采用单一变换进行全局建模,然而真实相机运动往往包含非匀速、曲线等复杂特性,单一刚体变换难以精确描述细微运动偏差,并且对离散数值积分带来的近似误差也缺乏有效补偿机制。这种粗粒度的建模方式将不同性质的运动混叠于单一变换中,容易导致运动表示的不准确,制约模型在复杂运动场景下的表达能力。

受上述分析启发,本文从连续轨迹建模与结构化解耦表示两个角度出发,提出一种面向运动去模糊的辛几何双路解耦三维高斯泼溅方法。在连续轨迹建模方面,本文引入基于辛几何结构的哈密顿动力学框架,将相机运动状态建模为相空间中的位置-动量对偶分量,并设计辛神经导数网络学习状态演化规律,确保相流满足辛保持性质,从本质上保证了相机轨迹的连续性与物理一致性。在复杂运动处理层面,本文设计双路自适应运动解耦框架,将相机运动分解为对应物理先验的刚性运动分支与对应残差精修的连续运动精调分支两部分独立建模。刚性运动分支继承辛几何轨迹估计并采用螺旋轴参数化(Eade, 2017)刻画曝光期间的整体位姿变换,为后续精修提供结构稳定且数学正当的全局位姿基线;连续运动精调分支通过无约束变换矩阵(Whyte 等, 2010)捕捉刚性运动无法描述的细微运动偏差与积分近似误差,并施加正交性正则约束以维持几何合理性。两分支由粗到细、先稳定后细化,通过自适应

协同组合实现对复杂相机运动的结构化分解与精确建模。

本文方法在 Deblur-NeRF 合成/真实模糊数据集 (Ma 等, 2022) 和 ExBluRF 极端模糊数据集 (Lee 等, 2023) 上进行了全面实验验证。定量与定性结果表明, 与现有方法相比, 本文方法在图像重建质量、细节恢复能力以及复杂运动场景适应性方面均取得了明显提升。

1 三维高斯泼溅与运动模糊成像模型

1.1 三维高斯泼溅

三维高斯泼溅是一种基于光栅化渲染的神经渲染方法, 其核心思想是将三维场景显式表示为一组三维高斯分布的集合。与神经辐射场的隐式体积表示不同, 3DGS 中的每个三维高斯由以下参数定义

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

式中, $x \in \mathbf{R}^3$ 为空间任意点, $\mu \in \mathbf{R}^3$ 为均值向量, 即高斯分布的中心位置, $\Sigma \in \mathbf{R}^{3 \times 3}$ 为协方差矩阵, 控制高斯分布的形状与尺度。协方差矩阵 Σ 由可学习的缩放向量 $S \in \mathbf{R}^3$ 和旋转四元数 $q \in \mathbf{R}^4$ 通过 $\Sigma = RSS^T R^T$ 计算得到, 其中 S 为由 s 构建的对角缩放矩阵, R 为由四元数 q 转换得到的旋转矩阵。此外, 每个三维高斯还包含球谐函数系数 (SH), 用于表示视角相关的颜色信息; 以及不透明度 α , 控制该高斯的可见性。

在渲染阶段, 3DGS 将三维空间中的高斯投影到二维相机平面, 并使用 α 混合计算最终像素颜色。对于按深度从近到远排序后的第 i 个高斯, 其颜色 c_i 和不透明度 α_i 对像素颜色的贡献为该高斯前方所有高斯的不透明度乘积与自身颜色的乘积。遍历所有 N 个有序高斯求和即为最终像素颜色, 具体为

$$I_c = \sum_{i=1}^N c_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j) \quad (2)$$

式中求和与乘积遍历从近到远排序的高斯集合, c_i 和 α_i 分别为第 i 个高斯的颜色和不透明度。

1.2 运动模糊成像模型

运动模糊是相机在曝光期间发生运动时, 场景辐射在成像平面上积分的结果。当相机沿某一轨迹从位置 τ_0 移动到 τ_N 时, 模糊图像的获取过程可以建模为曝光时间窗口内多个清晰帧的加权积分, 具

体为

$$I_{blur} = \frac{w_0 I_{\tau_0} + w_1 I_{\tau_1} + \cdots + w_N I_{\tau_N}}{w_0 + w_1 + \cdots + w_N} \quad (3)$$

式中 I_{τ_i} 为曝光时间 τ_i 时刻对应的清晰图像帧, w_i 为对应的积分权重, 且满足 $\sum w_i = 1$ 。该式揭示了运动模糊建模的核心问题, 即如何准确估计曝光时间内相机沿连续轨迹运动所对应的一系列虚拟相机位姿。需要特别指出的是, 上述求和本质上是对连续积分的数值近似, 而相机在实际曝光过程中的运动轨迹 $T(\tau)$ 是连续的, 因此最终场景重建的质量在很大程度上取决于相机轨迹建模的质量。

2 方法

本研究的目标是以多视角运动模糊图像及对应的相机标定参数作为输入, 重建精确的三维场景并渲染出清晰的新视角图像。本研究以物理先验为原则, 设计了面向运动去模糊的辛几何双路解耦三维高斯泼溅框架, 包含辛几何驱动连续相机轨迹建模、双路运动解耦建模、面向去模糊的 3DGS 渲染管线以及多损失联合优化策略。本文通过 COLMAP (Schönberger 和 Frahm, 2016) 获得每幅图像的相机姿势和初始点云。

整体流程如图 1 所示, 给定相机标定位姿 $T_{\tau_u} = [R_{\tau_u} | t_{\tau_u}]$, 首先编码为辛几何相空间状态, 再利用辛几何微分方程求解器分为刚性运动与连续运动精调两个独立分支对连续相机轨迹进行建模得到虚拟相机位姿序列; 随后对 N 个虚拟相机位姿独立渲染得到 N 张清晰图像; 最后, 通过像素级加权融合与模糊区域掩码生成最终的运动模糊预测图像, 并将其与输入模糊图像进行监督训练。在端到端优化过程中, 三维高斯参数与运动建模参数同步更新, 从而实现精确的场景重建与运动轨迹估计。

2.1 连续时间相机轨迹建模

真实相机运动受拍摄者意图、抖动与场景约束等多重因素影响, 难以用单一显式方程精确刻画, 但其连续轨迹仍呈现出可观察的运动学共性, 即时间维度的连续性、相邻帧位姿变化的有界性以及长程运动的整体稳定性。上述共性与哈密顿系统在数学上所具有的辛保持性、相流体积守恒性及长时间数值稳定性在形式上高度一致 (Hairer 等, 2006)。因

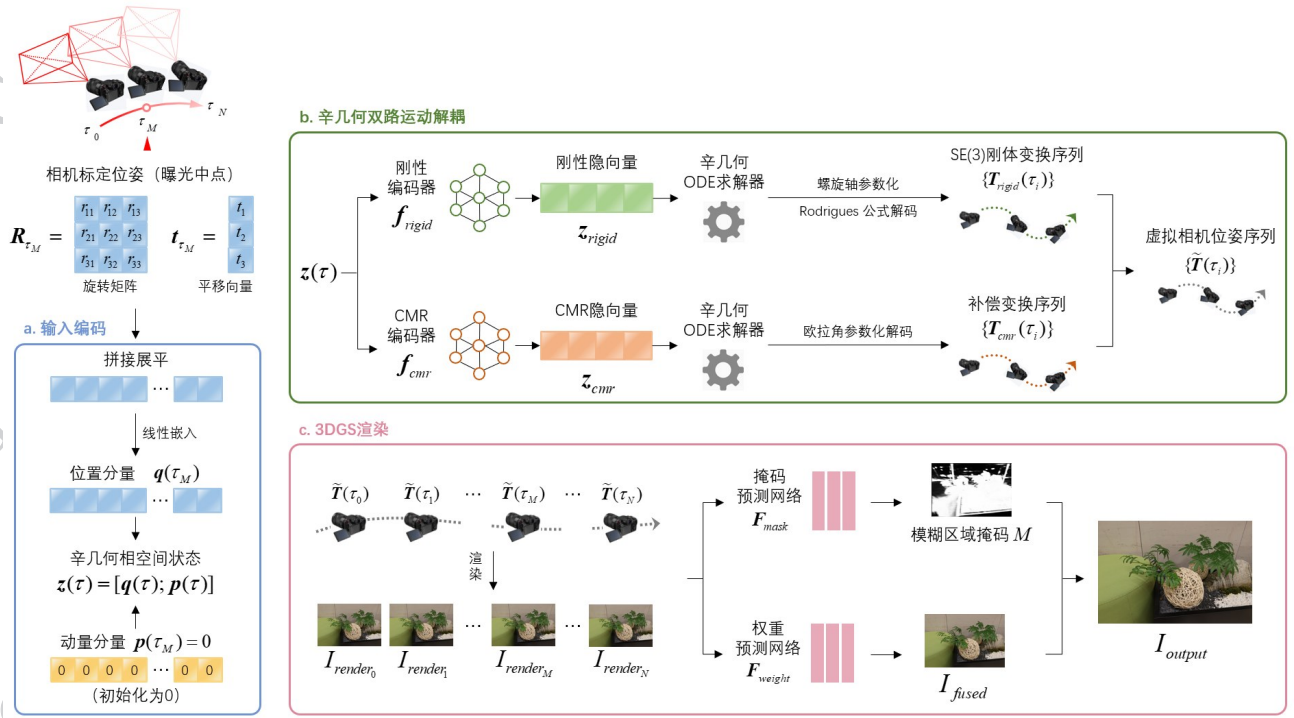


图1 整体方法流程图

Fig. 1 Overall framework of the proposed method

此,本文将哈密顿动力学视为一种正则化先验,通过位置-动量对偶结构与辛保持积分器,将"轨迹连续、演化结构稳定"这一运动学直觉显式编码为归纳偏置,从而约束神经网络的解空间。该先验并不声称真实相机运动严格满足某一守恒律,而是为长程外推提供数值上更稳定的参数化形式(Greydanus 等, 2019; Matsubara 等, 2021)。

2.1.1 状态空间表示

相机轨迹在曝光时间 $[\tau_0, \tau_N]$ 内演化,其中 τ_0 和 τ_N 分别对应曝光的起始和终止时刻。由于仅有输入模糊图像的采集时刻 τ_M ,即曝光中点具有标定相机位姿 T_{τ_M} ,因此将 τ_M 作为初始状态。将相机在曝光时间内的运动建模为连续时间演化过程,核心在于在相空间中定义状态表示(Greydanus 等, 2019)。将相空间状态向量 $z(\tau_M)$ 分为位置 q 和动量 p 两个互补分量,具体为

$$z(\tau) = [q(\tau); p(\tau)] \quad (4)$$

式中 $q \in \mathbb{R}^{\text{view_dim}}$ 和 $p \in \mathbb{R}^{\text{view_dim}}$ 为相机状态的隐位置和隐动量分量,两者通过辛几何结构耦合,共同刻画相机运动的完整动力学特征。

输入图像的标定相机位姿 $T_{\tau_M} = [R_{\tau_M} | t_{\tau_M}]$,将标定旋转矩阵 R_{τ_M} 与平移向量 t_{τ_M} 拼接为12维向量后经

线性嵌入层映射到 view_dim 维空间,得到初始位置分量 $q(\tau_M)$;动量分量 $p(\tau_M)$ 则初始化为零向量,在训练过程中通过反向传播逐步优化。随后将初始位置分量 $q(\tau_M)$ 和动量分量 $p(\tau_M)$ 拼接,就得到了输入标定相机位姿 T_{τ_M} 对应的相空间初始状态 $z(\tau_M)$ 。此外,每个视图具有独立的可学习视图嵌入向量 $v_i \in \mathbb{R}^{\text{view_dim}}$ (view_dim 默认64),用于编码该视角的模糊特征,使模型能够为不同视角的模糊图像生成对应的运动轨迹。

2.1.2 辛几何神经导数

相机运动的连续时间演化由一阶常微分方程(ODE)系统描述(Barfoot, 2017),具体为

$$\frac{dz}{d\tau} = f_{\Phi}(z(\tau), \tau; \Phi) \quad (5)$$

式中 $f_{\Phi}(\cdot)$ 为神经网络参数化的神经导数函数, Φ 为可学习参数。不同于标准神经 ODE 直接回归状态导数(Chen 等, 2018),本文遵循辛几何的理论框架(Matsubara 等, 2021),将相机状态演化建模为哈密顿系统。定义哈密顿量

$$H(z) = T(p) + V(q) \quad (6)$$

式中 $T(p)$ 为动能项(动量的函数), $V(q)$ 为势能项。哈密顿动力学给出标准的一阶辛方程组为

$$\begin{aligned}\frac{dq}{d\tau} &= \frac{\partial H}{\partial p} \\ \frac{dp}{d\tau} &= -\frac{\partial H}{\partial q}\end{aligned}\quad (7)$$

该方程组可统一写为

$$\frac{dz}{d\tau} = J^{-1} \cdot \nabla H(z) \quad (8)$$

式中 $J = [[0, I]; [-I, 0]]$ 为标准辛矩阵。辛矩阵 J 满足 $J^2 = -I$ 和 $J^T = -J$, 这一代数结构在几何上对应于相流的体积保持性质, 即哈密顿系统的相流在时间演化过程中始终保持相空间的几何结构不变。在本文实现中, 神经导数函数 f_Φ 被分解为两个并行的分支, 均为单层线性层后接 ReLU 非线性激活, 分别建模动能梯度 f_p 和势能梯度 f_q , 具体为

$$\begin{aligned}f_\Phi(z, \tau) &= [f_p(z, \tau; \Phi_p); f_q(z, \tau; \Phi_q)] \\ (f_p &= \frac{\partial H}{\partial p}, f_q = -\frac{\partial H}{\partial q})\end{aligned}\quad (9)$$

为使模型能够感知相机运动的时间位置, 本文为曝光窗口内的 N_{interp} 个采样时间点分别引入了可学习的时间嵌入向量 $e_i \in \mathbf{R}^{\text{time_dim}}$ (time_dim 默认 8), 并将其与状态向量的对应分量拼接后输入神经导数网络, 从而实现时间感知的状态演化。

2.1.3 辛几何微分方程求解器

状态 ODE 通过辛几何微分方程求解器进行数值积分。与标准 Runge-Kutta 方法 (Butcher, 2016) 不同, 辛积分器在每一步迭代中交替更新位置和动量分量, 从而在离散化层面保持辛结构。给定初始状态 $z(\tau_m)$, 求解器分别在 $\tau_m \rightarrow \tau_N$ 和 $\tau_0 \leftarrow \tau_m$ 两个方向上独立积分, 最后将两部分轨迹拼接, 得到完整的 N 个连续相机状态 $\{z(\tau_i)\}$ 。本文采用辛欧拉法作为积分器, 其更新规则 (以正向为例) 为

$$p_{k+1} = p_k + \Delta\tau \cdot f_p(q_k, p_k) \quad (10)$$

$$q_{k+1} = q_k + \Delta\tau \cdot f_q(q_{k+1}, p_{k+1}) \quad (11)$$

式中 $\Delta\tau$ 为时间步长 (固定为 0.1)。该交替更新策略在每一步均保持辛结构, 确保长时间积分的数值稳定性。

2.1.4 位姿解码

为了让运动轨迹能被后续的三维重建直接使用, 本文将相机状态 z 中的位置分量 q 通过轻量级 MLP 解码器解码为 SE(3) 刚体变换矩阵 (Barfoot, 2017)。首先, 将 q 均分为 q_w 和 q_v 两部分, 再分别解码旋转向量 w 旋转角度 θ 和平移向量 v , 具体为

$$w = \exp_{\text{map}}(\text{decoder}_w(q_w)) \quad (12)$$

$$\theta = \text{decoder}_\theta(q_w) \quad (13)$$

$$v = \text{decoder}_v(q_v) \quad (14)$$

式中 $\exp_{\text{map}}(\cdot)$ 为对数映射后的归一化操作, decoder_w 、 decoder_θ 和 decoder_v 为独立的头网络。旋转向量 w 经斜对称矩阵变换后, 通过 Rodrigues 公式转换为旋转矩阵 R_{exp} ; 平移向量 v 经 SE(3) 指数映射后得到平移量 $t = G(\theta) \cdot v$ 。最终构建 SE(3) 变换矩阵为

$$T_{\text{rigid}} = [[R_{\text{exp}}, t]; [0, 0, 0, 1]] \quad (15)$$

该推导过程遵循刚体运动的几何学原理, 确保了旋转矩阵的正当性 ($\det(R_{\text{exp}}) = 1$) 和平移向量的连续性。Rodrigues 公式和 SE(3) 指数映射中用于处理平移分量的矩阵 $G(\theta)$ 的具体形式为

$$\begin{aligned}R_{\text{exp}} &= I + \sin(\theta) \cdot [w_\wedge] + (1 - \cos(\theta)) \cdot [w_\wedge]^2 \\ G(\theta) &= I\theta + (1 - \cos(\theta)) \cdot [w_\wedge] + (\theta - \sin(\theta)) \cdot [w_\wedge]^2\end{aligned}\quad (16)$$

式中 $[w_\wedge]$ 为 w 对应的 3×3 斜对称矩阵。设旋转向量为 $w = [w_x, w_y, w_z]^T$, 则 $[w_\wedge]$ 定义为

$$[w_\wedge] = \begin{bmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{bmatrix} \quad (18)$$

该矩阵满足叉积性质, 即 $[w_\wedge] \cdot v = w \cdot v$ 。

2.1.5 辛几何轨迹建模方法分析

本文所提出的辛几何轨迹建模方法在数学与数值层面具有以下性质: 1) 连续性。辛几何常微分方程系统在给定初始条件下具有唯一解, 其解作为时间的连续函数在积分区间内无跳变。2) 平滑性。辛结构的近能量守恒特性使相机状态在相空间中沿光滑相轨道演化, 速度与加速度在数值层面具有连续性, 缓解了纯数据驱动方法中常见的非物理抖动现象。3) 长程数值稳定性。辛积分器在每一步迭代中维持辛二形式不变, 从而在长时间外推与多步反向传播中表现出优于一般 Runge-Kutta 方法 (Butcher, 2016) 的数值稳定性。需要再次强调, 上述性质是模型所施加的正则先验在数学与数值层面所具备的优良特性, 而非对真实相机运动严格满足物理守恒律的强假设; 其作用在于为学习过程提供稳定的归纳偏置, 而非完整刻画物理过程本身。

尽管辛几何框架在轨迹连续性与数值稳定性方面具有理论优势, 但其并非在所有场景下均优于显式轨迹平滑约束, 仍存在适用边界: 1) 极端非线性运

动。当相机运动呈现高度非线性特征(如急转弯、急停等)时,哈密顿动力学的平滑相轨道假设可能与实际运动产生偏差,显式平滑约束在此类场景下可能更为稳定。2)长曝光大幅运动。辛积分器在长轨迹累积位移较大时仍会累积数值误差。如3.3节所示,当时间采样点数过多时性能出现回落,表明本文方法对长轨迹建模存在一定局限。3)计算开销。本文方法需训练辛神经导数网络,计算开销高于简单的显式插值或平滑方法,在实时应用中需权衡精度与效率。

综上,本文方法的核心优势在于物理一致的连续轨迹建模与长程数值稳定性,其局限性主要体现在极端非线性运动与计算开销方面,实际应用时应根据场景特性与资源约束选择合适的轨迹建模策略。

2.2 双路运动解耦

本文提出一种双路自适应运动解耦框架,将模糊运动投影到相机刚性运动分支和连续运动精调(CMR)分支两个独立的演化空间中分别建模。

2.1节建立的辛几何轨迹建模框架为相机轨迹提供了连续性的结构保障,然而实际的相机运动模糊包含了两类性质迥异的因素:一类是由拍摄者主动控制的、占主导地位的整体刚体运动导致的模糊,另一类则是由相机抖动、加速度变化以及离散辛积分截断所引入的细微残差导致的模糊。单一全局变换难以同时刻画复杂运动中导致模糊的两类因素,所以本文还设计了双路运动解耦,兼顾两类成分的建模需求。刚性运动分支承担物理先验的角色,为后续精修提供全局位姿基线,而连续运动精调(CMR)分支承担残差精修的角色,对先验未能覆盖的偏差进行数据驱动的学习式补偿。

2.2.1 刚性运动分支

本分支承担物理先验通路的角色,完全继承2.1节所描述的辛几何轨迹估计方法,使用辛几何神经ODE在相空间中演化相机状态,从初始位姿 T_{τ_u} 出发,生成一系列连续的刚体变换 $\{T_{\text{rigid}}(\tau_i)\}$ 。本分支的设计意图并非完整刻画真实相机运动的全部分量,而是为后续CMR分支提供一个结构稳定、数值鲁棒、且与连续运动直觉一致的全局位姿基线。该基线在数学上由SE(3)上的螺旋轴参数化(Eade, 2017)保证刚体变换的正当性,在数值上由辛欧拉积

分器保证长程外推的稳定性。

2.2.2 连续运动精调分支

本分支承担残差精修通路的角色,其设计动机源于对刚性先验固有局限的考量:1)辛欧拉积分器在固定步长下不可避免地引入离散化截断误差;2)刚体SE(3)参数化在数学上排除了非刚体形变的可能;3)真实相机运动中往往存在非匀速、抖动与轻微非线性等无法被单一刚体变换完全吸收的成分。为补偿上述局限,CMR分支在非约束的欧氏空间中直接学习一个轻量级变换矩阵,不显式遵循SE(3)约束,仅通过正交正则维持旋转部分的几何合理性,从而获得更高的自由度来拟合复杂的连续运动。CMR分支使用独立的视图嵌入和编码器,在独立的隐状态空间中演化,随后通过独立的解码器映射到欧氏空间参数(Goodfellow等, 2016),并直接构建为变换矩阵。为防止训练初期CMR分支过度干扰刚性运动,CMR的平移和旋转偏移量初始化为接近零的值。CMR变换矩阵同样施加正交性正则约束,防止产生过度非刚性形变,具体为

$$L_{\text{ortho}} = \|R_{\text{cmr}}^T \cdot R_{\text{cmr}} - I\|_F^2 \quad (19)$$

式中 R_{cmr} 为 T_{cmr} 中的旋转分量。该正交约束使CMR的旋转部分在训练过程中逐渐接近有效旋转矩阵,从而在灵活性与几何合理性之间取得平衡。

2.2.3 自适应融合策略

双路运动分支独立演化后,通过矩阵乘法组合为总体变换,具体为

$$T_{\text{transform}}(\tau_i) = T_{\text{rigid}}(\tau_i) \times T_{\text{cmr}}(\tau_i) \quad (20)$$

T_{rigid} 为刚性运动分支的变换矩阵,建模全局相机运动,即主导运动模糊的主要来源, T_{cmr} 为连续运动精调分支的变换矩阵,对相机运动轨迹叠加局部精调修正。这两个变换矩阵皆接近单位矩阵,则相乘后得到的组合变换场也接近单位矩阵,确保了组合变换的数学性质。最终,将组合变换作用于标定相机位姿 T_{τ_u} ,得到 N 个虚拟相机位姿 $\tilde{T}_i$ 为

$$\tilde{T}_i = T_{\tau_u} \times T_{\text{transform}}(\tau_i), i = 0, 1, \dots, N-1 \quad (21)$$

2.3 面向去模糊的3DGS渲染模型

为消除高斯空间频率伪影,本文参考Mip-Splatting(Yu等, 2024)的做法,在三维高斯上施加三维频率约束滤波器,根据各三维高斯到相机的深度和相机焦距计算频率约束项,将其叠加到高斯的协方差矩阵和透明度上,从而约束高斯在空间频率上

满足奈奎斯特采样条件,消除渲染过程中的锯齿与高频伪影。沿 2.2 节所得到的虚拟相机位姿序列 $\{\tilde{T}_i\}$, 对每个虚拟相机独立执行三维高斯泼溅渲染管线, 分别渲染出 N 张清晰图像 $\{I_{render_i}\}$ 。随后将 N 张渲染图像的通道拼接输入轻量级卷积神经网络 F_{weight} 预测权重图 $\{w_i\}$, 再采用像素级加权融合策略按照归一化像素权重 $w_i \in R^{H \times W \times 3}$ 将 N 张渲染图像合成为加权融合结果 I_{fused} 。此外, 为区分图像中的模糊区域与清晰区域, 本文引入模糊区域掩码 $M \in [0, 1]^{H \times W \times 3}$, 使模型能够自适应地为模糊区域输出多帧加权融合结果 I_{fused} , 同时为清晰区域保留曝光中点帧的高置信度渲染 I_{render_u} 。掩码 M 由另一网络 F_{mask} 预测, 其输入为 F_{weight} 中间特征的通道拼接, 输出为 $[0, 1]$ 之间的掩码值。最终输出图像为

$$I_{output} = (1 - M) \odot I_{render_u} + M \odot I_{fused} \quad (22)$$

式中 I_{render_u} 为曝光中点相机位姿的渲染结果。

2.4 损失函数

本文方法的整体优化目标由重建损失、结构感知损失与运动正则损失组成。重建损失衡量模型生成的模糊预测图像与真实输入模糊图像之间的像素级差异, 本文采用 L_1 损失 $L_{L1} = |I_{output} - I_{blur_gt}|_1$, 其中 $|\cdot|_1$ 为平均绝对误差损失, I_{blur_gt} 为输入的真实模糊图像。感知损失衡量两张图像的结构相似性, 能够有效捕捉模糊图像中的结构退化, 本文采用结构性判别器相似度损失 $L_{DSSIM} = 1 - DSSIM(I_{output}, I_{blur_gt})$ 。运动正则损失包括防止 CMR 分支产生过度非刚性形变的正交正则 $L_{ortho} = |R_{cmr}^T \cdot R_{cmr} - I|_F^2$ 和使模型倾向于依赖中点清晰帧渲染结果的掩码稀疏正则 $L_{mask} = \text{mean}(M)$ 。总损失函数为上述各损失的加权组合, 具体为

$$L_{total} = (1 - \lambda_{dssim}) \cdot L_{L1} + \lambda_{dssim} \cdot L_{DSSIM} + \lambda_{mask} \cdot L_{mask} + \lambda_{ortho} \cdot L_{ortho} \quad (23)$$

式中 $\lambda_{dssim} = 0.3$, $\lambda_{mask} = 10^{-3}$, $\lambda_{ortho} = 10^{-4}$ 。 L_1 与 L_{DSSIM} 的组合确保了像素级精度与结构感知能力的平衡。

2.5 优化策略

本文方法的待优化参数分为两组: 1) 三维高斯参数。包括高斯的位置、缩放、旋转、不透明度、球谐函数系数以及三维频率约束滤波器参数; 2) 运动建模参数。包括辛几何神经导数网络权重、位姿解码器权重、CMR 变换参数以及像素级加权融合网络权

重。两组参数采用独立的 Adam 优化器 (Kingma 和 Ba, 2015), 其中高斯参数使用 3DGS 原版的指数衰减学习率调度策略, 运动建模参数使用固定学习率 10^{-4} , 且在训练过程中线性衰减至 10^{-6} 。

为保证训练稳定性, 本文采用渐进式激活机制: 前 k_1 次迭代 ($k_1 = 1000$) 仅训练三维高斯参数, 运动建模模块保持冻结; k_1 至 k_2 次迭代 ($k_2 = 3000$) 激活双路运动解耦模块, 但仍冻结像素级加权网络和模糊掩码网络, 此时 N 张渲染图像直接取均值合成模糊预测; k_2 次迭代之后全量激活所有模块, 包括像素级加权网络和掩码网络。本文还沿用 3DGS 的自适应密集化策略, 即在迭代次数低于 20000 时, 根据梯度阈值对高斯进行克隆和分裂操作, 增加场景表示的精细度, 而达到 20000 后仅进行剪枝操作以移除不透明度低于阈值的高斯。在每次密集化操作后, 重新计算三维频率约束滤波器参数。

3 实验结果

本文使用的数据集包括: Deblur-Nerf 合成数据集、Deblur-Nerf 真实世界数据集和 ExBluRF 极端模糊数据集。Deblur-Nerf 合成数据集由 Blender (Blender Foundation, 2018) 生成的 5 个场景组成, 其中图像在线性 RGB 空间中组合, 以创建最终的模糊图像。Deblur-Nerf 真实世界数据集包括用佳能 EOS RP 相机捕捉的 10 个场景, 通过手动设置曝光时间以产生模糊图像。ExBluRF 真实世界数据集由 8 个场景组成, 是具有挑战性的极端相机运动模糊场景。对于定量结果, 使用三个指标来评估: 峰值信噪比 (PSNR)、结构相似性指数 (SSIM) 和学习感知图像块相似性 (LPIPS)。

3.1 对比实验

为验证本文方法的有效性, 在 Deblur-NeRF 合成数据集、Deblur-NeRF 真实数据集和 ExBluRF 极端模糊数据集上与 7 种主流方法进行对比, 包括基于神经辐射场的 Deblur-NeRF 和 ExBluRF, 基于三维高斯泼溅的 Deblurring 3DGS (Lee 等, 2024)、Mip-Splatting 和 BAGS, 束调整方法 BAD-Gaussians 以及连续动力学方法 SMURF。定量结果如表 1 所示, 在合成数据集上, 本文方法在 PSNR、LPIPS 两项指标上达到最优, 尽管 SMURF 在 SSIM 上略高于本文方

表 1 对比实验结果
Table 1 Rendering results

方法	Synthetic Scene			Real-World Scene			ExBluRF Real-World Scene		
	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
ExBluRF	23.54	0.662	0.388	23.41	0.675	0.219	23.45	0.550	0.381
Mip-Splatting	23.28	0.677	0.282	21.87	0.627	0.307	24.52	0.566	0.600
Deblur-NeRF	28.40	0.866	0.118	24.75	0.806	0.155	<u>28.87</u>	<u>0.709</u>	0.402
Deblurring 3DGS	28.24	0.858	<u>0.105</u>	25.65	0.822	0.114	27.36	0.680	0.399
BAD-Gaussians	29.43	0.793	0.248	19.66	0.635	0.353	26.83	0.663	<u>0.322</u>
BAGS	27.34	0.835	0.112	<u>26.70</u>	<u>0.824</u>	<u>0.096</u>	24.70	0.584	0.528
SMURF	<u>30.62</u>	0.912	0.117	26.33	0.790	0.188	16.52	0.331	0.669
ours	30.74	<u>0.901</u>	0.073	28.45	0.853	0.084	30.74	0.752	0.307

注:加粗、下划线字体分别表示各列最优、次优结果。

法 1.2%,但在 LPIPS 上落后于本文方法 37.6%,表明 SMURF 的实际场景重建效果并不如本文方法。在真实数据集上,本文方法的 PSNR 达到 28.45 dB,领先次优方法 BAGS 达 1.75 dB,SSIM 和 LPIPS 均优于所有对比方法。在 ExBluRF 极端模糊数据集上,本文方法的 PSNR 为 30.74 dB,领先次优方法 Deblur-NeRF 达 1.87 dB,SSIM 和 LPIPS 同样均达到最优。

定性对比如图 2 所示,本文方法在细节恢复方面优势明显。在绿色盆栽场景中,本文方法能够更清晰地还原叶片的纹理细节和边缘,而对比方法均存在不同程度的模糊或伪影;在两个包含沙发的场景中,本文方法渲染的沙发表面纹理更加清晰,对比方法则出现了明显的细节丢失;在池塘场景中,本文方法能够更准确地呈现水和木质架子的边界细节,对比方法则存在模糊和伪影问题。上述定性结果进一步验证了本文方法在复杂运动模糊场景下的有效性和优越性。

3.2 消融实验

为了验证辛几何模块和双路运动解耦的贡献,本节在 Deblur-NeRF 合成数据集、Deblur-NeRF 真实数据集和 ExBluRF 极端模糊数据集上分别对这两个模块进行消融实验。所有实验均在搭载 NVIDIA GeForce RTX 4090 GPU(单卡 24 GB GDDR6X,驱动版本 550.163.01)、Intel Xeon Platinum 8269CY CPU(@ 2.50 GHz)的服务器上完成,深度学习框架采用 PyTorch 2.6.0(CUDA 12.4),不同消融配置除目标

模块外共享所有超参数与随机种子,以排除随机性干扰。实验包括:1)标准神经 ODE。使用 torchdiffeq (Chen 等,2018)的 dopri5 求解器替代本文的辛欧拉积分器;2)仅刚性解耦。移除 CMR 分支,仅保留辛几何刚性运动轨迹估计;3)辛几何双路解耦。即本文的方法,使用辛欧拉积分器保持哈密顿动力学约束,并使用刚性分支与 CMR 分支的加权组合。

定量实验结果如表 2 所示,从实验 1)和实验 3)的对比可以看出,辛几何建模在所有指标上均优于标准神经 ODE,验证了本文的理论分析,辛几何约束确保了相流的结构保持性质,使得长时间积分的数值稳定性更好,在三维重建这种需要多步积分的运动模糊建模任务中能取得更好性能。实验 2)和实验 3)的对比结果显示,仅使用刚性分支时,模型难以补偿离散积分带来的近似误差,在大幅度运动场景下性能受限,而双路解耦框架利用刚性分支提供稳定的轨迹基础和 CMR 分支在局部进行的精细修正,取得更好性能。不同方法的渲染结果如图 3 所示,从图像中可以看出,辛几何双路解耦的方案不仅改善了图像中的结构模糊现象,还能渲染出更清晰的纹理。如第一个场景沙发的部分,辛几何双路解耦能还原出大部分沙发材质的纹理,相较其他方法有明显提升,再如第二个场景中的盆栽,辛几何双路解耦相较其他方法能更好缓解运动带来的结构模糊现象。

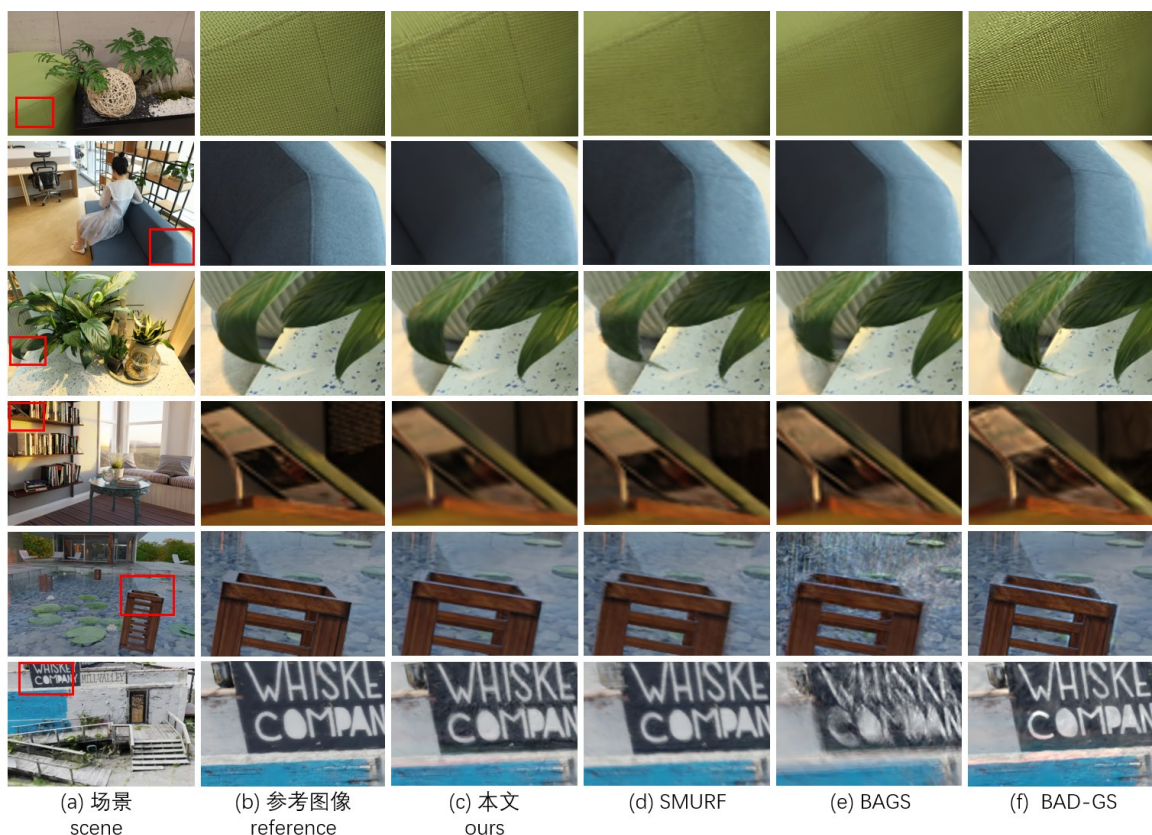


图2 对比实验结果

Fig. 2 Comparison of deblurring results

除渲染质量与轨迹精度外,计算开销亦是衡量神经渲染方法实用性的重要维度。表3报告了在Deblur-NeRF合成数据集上不同消融配置的模型计算效率对比结果。从结果可以看出,辛几何双路解耦的模型参数数量和最低的峰值显存都达到了最低,表明辛几何先验与双路解耦结构虽然带来额外的计算冗余,但由于本文采用自适应密集化策略(2.5节),高斯参数数量会随训练动态调整,而辛几何双路解耦能够更准确地建模相机运动轨迹,使密集化阶段可更早收敛于一组稀疏且充分的高斯表示,最终保留的高斯数显著少于仅刚性解耦和标准神经ODE,以更高效率的参数化方式压缩了整体表示开销,验证了所提方法在精度与参数利用率上的优势。在训练阶段,辛几何双路解耦单场景平均训练耗时约16736秒(4.65小时),高于仅刚性解耦的13059秒(3.63小时)。这一额外开销来源于辛欧拉积分器在每一步迭代中均需执行位置与动量分量的交替更新,以及对应的反向传播链。考虑到完整模型相较仅刚性解耦在合成数据集上PSNR提升1.83 dB,上述训练成本的增加是合理的,符合精度收益驱动的

工程权衡。在推理阶段,完整模型单帧推理时间约342ms(2.93 FPS),与仅刚性解耦的335ms处于同一量级,未引入显著延迟。需要指出的是,由于所提方法在曝光时间内对多个虚拟相机独立渲染并通过像素级加权融合生成最终输出,推理耗时的增加主要源自运动模糊的多帧建模机制本身,而非所提算法结构。在实际离线重建场景下,所提方法的训练与推理代价仍处于可接受范围。

3.3 时间采样点数分析

运动模糊图像的形成过程中,相机在曝光时间 $[\tau_0, \tau_N]$ 内沿连续轨迹运动,捕获的像素值是沿轨迹各采样点接收到光照的积分累积,而时间采样点数 N_{warp} 决定了沿轨迹采样的虚拟相机数量,直接影响模糊核的建模精度与渲染计算开销。为了分析时间采样点数 N_{warp} 带来的影响,本节在Deblur-NeRF真实世界数据集上进行实验。如表4,实验结果表明,当 $N_{warp} = 9$ 时性能达到最优,继续增加采样点数反而导致性能轻微下降。当 $N_{warp} = 3$ 时,采样点过于稀疏,仅能捕获轨迹的粗略轮廓,无法准确描述模糊的形成过程,导致重建图像存在明显的残余模糊;当

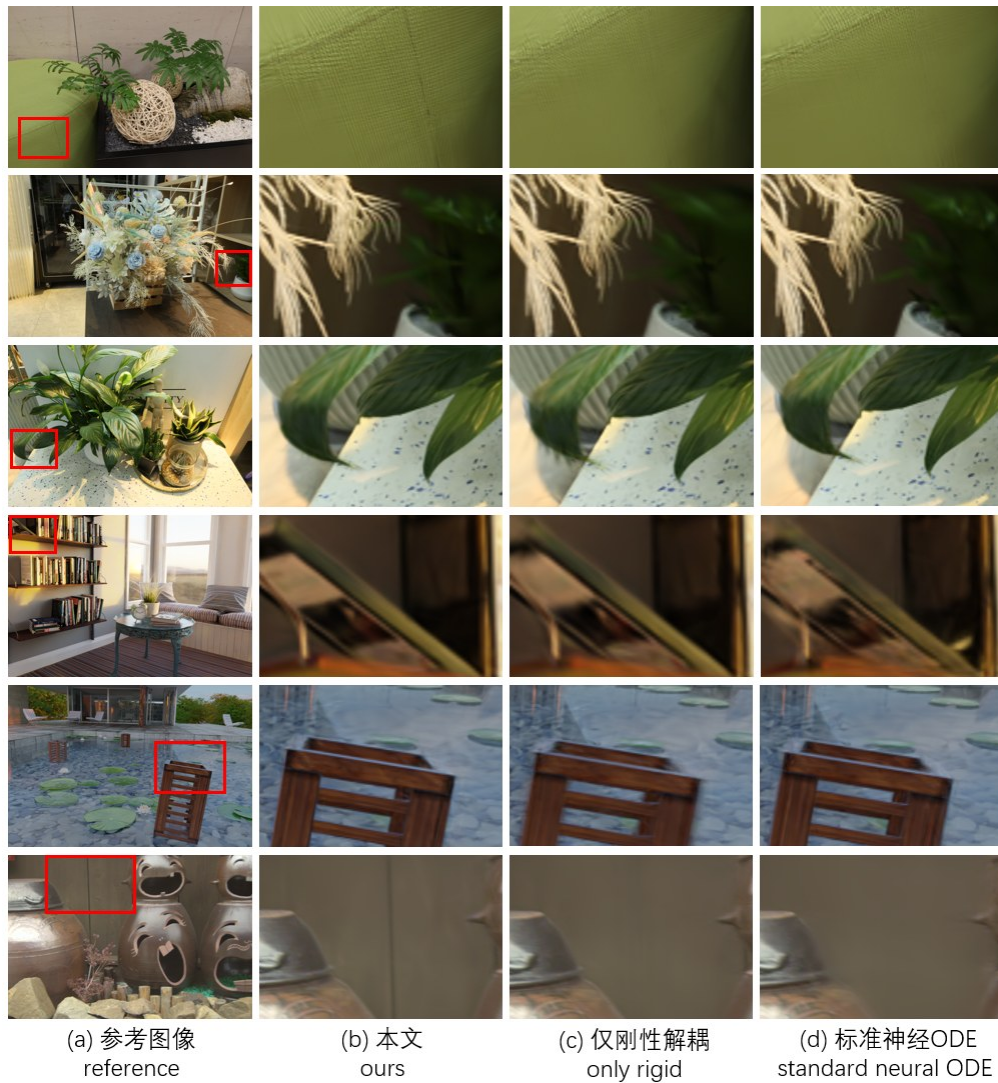


图3 消融实验结果

Fig. 3 Ablation study results

表2 消融实验结果

Table 2 Ablation study results

方法	Real-World Scene			Synthetic Scene			ExBluRF Real-World Scene		
	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
标准神经ODE	28.12	0.840	0.089	30.42	0.893	0.074	30.49	0.743	0.317
仅刚性解耦	27.85	0.834	0.096	28.91	0.869	0.112	30.73	0.751	0.308
辛几何双路解耦	28.45	0.853	0.084	30.74	0.901	0.073	30.74	0.752	0.307

注:加粗字体表示各列最优结果。

$N_{\text{warp}} = 5$ 时,性能稳步提升,表明更多的采样点能够更精细地近似模糊积分过程;当 $N_{\text{warp}} = 9$ 时达到性能峰值,三项指标均达到最优水平;当 $N_{\text{warp}} = 15$ 时,性能出现轻微回落,这可能是由于过密的轨迹近似引入了额外的数值误差。

从物理角度分析,采样点数的选择应与模糊程

度相匹配,对于轻度模糊场景,较少采样点数即可充分建模,而对于重度模糊场景则需要更多采样点以捕获长轨迹的运动细节。然而,过密的采样会引入额外的渲染开销和梯度反向传播负担,可能导致训练不稳定。本文综合考虑精度与效率的平衡,选择 $N_{\text{warp}} = 9$ 作为默认配置。

表 3 计算效率对比结果

Table 3 Computational efficiency comparison

方法	训练时间(s)	推理时间(ms/img)	FPS	高斯点数量	参数量(M)	峰值显存(GB)
标准神经ODE	9241	334.57	2.99	72804	5.35	2.67
仅刚性解耦	13059	334.69	2.99	69764	5.20	2.64
辛几何双路解耦	16736	342.18	2.93	55677	4.37	2.54

表 4 时间采样点数对比结果

Table 4 Comparison of time sampling points

采样点数	PSNR ↑	SSIM ↑	LPIPS ↓
3	26.65	0.807	0.142
5	27.72	0.836	0.108
9	28.45	0.853	0.084
15	28.15	0.844	0.085

注:加粗字体表示各列最优结果。

3.4 轨迹物理一致性分析

为验证辛几何双路解耦建模相机轨迹的实际效果,本节可视化了学习得到的相机运动轨迹。图 4 展示了训练视角的相机分布,图中使用彩色锥体表示各训练视角的相机位置,颜色从青色渐变至黄色,表示视角的采集顺序。以图 4 紫框中的视角为例,图 5 展示了该视角下沿运动轨迹采样的虚拟相机位姿,使用彩色锥体表示沿估计的运动轨迹在曝光时间内采样到的多个虚拟相机位置,颜色从红色渐变至浅色表示轨迹的时间演化方向。从图中可以观察到,相机锥体沿运动方向均匀分布,颜色过渡平滑,表明速度和加速度连续,符合相机运动的物理规律。

表 5 轨迹物理一致性分析定量结果

Table 5 Quantitative results of physical consistency analysis of trajectories

方法	ATE ↓	ATEP ↓
标准神经ODE	1.966	2.254
仅刚性解耦	2.461	3.042
辛几何双路解耦	1.501	1.774

注:加粗字体表示各列最优结果。

为进一步验证预测相机轨迹的质量,本节还使用绝对轨迹误差(ATE)及其 95 百分位(ATEP)定量评估相机轨迹的重建精度。ATE 通过将估计轨迹与真值轨迹进行 Sim(3) 对齐后计算均方根误差

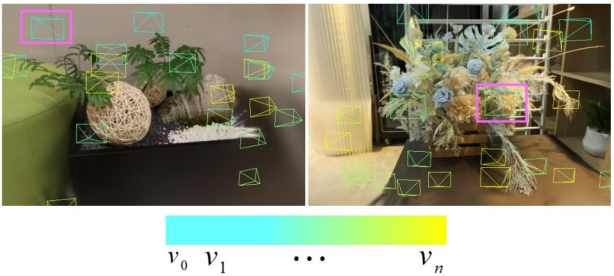


图 4 训练数据视角覆盖情况

Fig. 4 Training view coverage

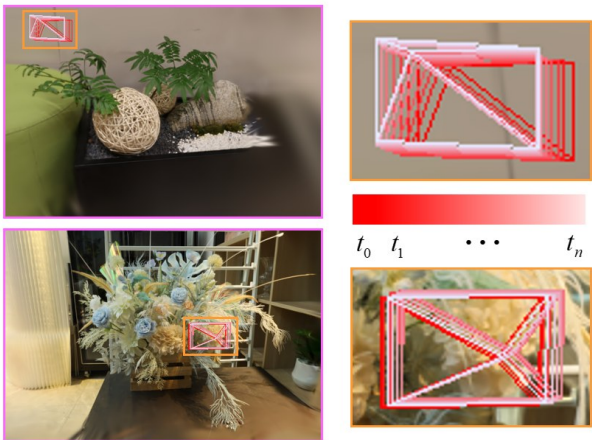


图 5 单视角运动轨迹

Fig. 5 Single-view motion trajectory

(Sturm 等, 2012),ATEP 则在 ATE 基础上取 95 分位数以反映误差的尾部分布(Cadena 等, 2016)。表 5 展示了在 Deblur-Nerf 合成数据集下使用三种消融配置的相机轨迹物理一致性分析定量结果。可以看出,辛几何双路解耦预测的相机运动轨迹在两项定量指标上均取得最优性能,ATE 由仅刚性解耦配置下的 2.46 m 下降至 1.50 m,相对降幅达 39.0%;同时 ATEP 由 3.04 m 下降至 1.77 m,相对降幅达 41.6%。上述结果表明,辛几何双路解耦显著提升了相机轨迹的重建精度,且在误差的尾部分布上同样取得改善,验证了本文方法建模出的相机运动轨迹的物理一致性。

4 结 论

针对引言部分提出的相机轨迹建模缺乏物理一致性保证、单一变换难以描述复杂非匀速运动这两个核心问题, 本文开展系统研究, 建立了基于辛几何的相机轨迹物理建模框架。不同于传统方法采用标准神经 ODE 或离散插值导致轨迹不稳定的问题, 本文引入哈密顿动力学框架, 将相机运动建模为相空间中的位置-动量对偶演化, 通过辛几何微分方程求解器确保相流满足辛保持性质, 从物理本质上保证了轨迹连续性, 并提升了长时间积分过程的数值稳定性。在此基础上, 本文还提出双路自适应运动解耦策略, 将相机运动分解为对应物理先验的刚性运动与对应残差精修的连续运动精调两条独立分支, 刚性分支继承辛几何轨迹估计提供结构稳定的全局位姿基线, 连续运动精调分支在非约束空间中学习轻量级补偿以吸收积分截断误差与刚体参数化的固有局限, 两支路通过自适应协同组合, 有效解决了相机复杂运动场景下的建模难题。为验证所提方法的有效性, 本文在三维高斯泼溅框架下构建了面向去模糊的渲染管线, 引入三维频率约束滤波与模糊区域掩码机制, 实验结果表明所提方法在 Deblur-NeRF 合成/真实数据集和 ExBluRF 极端模糊数据集上均取得了优于主流方法的重建质量。然而当前方法主要针对静态场景设计, 对动态前景目标的处理能力有限, 且高维辛神经导数网络存在较大计算开销, 如何在保持轨迹建模精度的前提下降低计算复杂度并扩展至动态场景, 是未来研究的重点方向。

参考文献 (References)

- Barfoot T D. 2017. *State Estimation for Robotics*. Cambridge: Cambridge University Press.
- Blender Foundation. 2018. *Blender: A 3D Modelling and Rendering Package* [EB/OL]. [2026-06-23].
<https://www.blender.org>
- Butcher J C. 2016. *Numerical Methods for Ordinary Differential Equations*. 3rd ed. Chichester: John Wiley & Sons.
- Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, et al. 2016. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, 32(6): 1309-1332. [DOI: 10.1109/TRO.2016.

2624754]

Chen R T Q. 2018. torchdiffeq [EB/OL]. [2026-06-23].

<https://github.com/rtqichen/torchdiffeq>Chen R T Q, Rubanova Y, Bettencourt J and Duvenaud D. 2018. Neural Ordinary Differential Equations // *Advances in Neural Information Processing Systems*. Montréal, Canada: NeurIPS Foundation: 6571-6583.Eade E. 2017. *Lie Groups for 2D and 3D Transformations* [EB/OL]. [2026-06-23].<https://ethaneade.com/lie.pdf>Goodfellow I, Bengio Y and Courville A. 2016. *Deep Learning*. Cambridge: MIT Press.Greydanus S, Dzamba M and Yosinski J. 2019. Hamiltonian Neural Networks // *Advances in Neural Information Processing Systems*. Vancouver, Canada: NeurIPS Foundation: 15379-15389.Hairer E, Lubich C and Wanner G. 2006. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. 2nd ed. Berlin: Springer.Hu Z Y, Zhou Q, Chen M J, Cui J C, Wu X F and Zheng B Y. 2024. Survey of Image Deblurring. *Journal of Image and Graphics*, 29(4): 841-861. [DOI: 10.11834/jig.230555]Kerbl B, Kopanas G, Leimkühler T and Drettakis G. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, 42(4): 1-14. [DOI: 10.1145/3592433]Kingma D P and Ba J. 2015. Adam: A Method for Stochastic Optimization // *International Conference on Learning Representations*. San Diego, USA: ICLR.Kupyn O, Budzan V, Mykhailych M, Mishkin D and Matas J. 2018. DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 8183-8192. [DOI: 10.1109/CVPR.2018.00854]Lee B, Lee H, Sun X, Ali U and Park E. 2024. Deblurring 3D Gaussian Splatting [EB/OL]. [2026-06-23].
<https://arxiv.org/abs/2401.00834>Lee D, Lee M, Shin C and Lee S. 2023. DP-NeRF: Deblurred Neural Radiance Field with Physical Scene Priors // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada: IEEE: 12386-12395. [DOI: 10.1109/CVPR52729.2023.01192]Lee D, Oh J, Rim J, Cho S and Lee K M. 2023. ExBluRF: Efficient Radiance Fields for Extreme Motion Blurred Images // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Paris, France: IEEE: 17639-17648.

Lee J, Lee D, Lee M, Kim D and Lee S. 2024. SMURF: Continuous Dynamics for Motion-Deblurring Radiance Fields [EB/OL]. [2026-06-23].

<https://arxiv.org/abs/2403.07547>

Ma L, Li X, Liao J, Zhang Q, Wang X, Wang J, et al. 2022. Deblur-

- NeRF: Neural Radiance Fields from Blurry Images//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 12861-12870.
- Matsubara T, Miyatake Y and Yaguchi T. 2021. Symplectic Adjoint Method: Efficient Gradient Computation for Hamiltonian Neural Networks//Advances in Neural Information Processing Systems. 34: 28486-28498.
- Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis//Computer Vision - ECCV 2020. Cham: Springer: 405-421. [DOI:10.1007/978-3-030-58452-8_24]
- Müller T, Evans A, Schied C and Keller A. 2022. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. ACM Transactions on Graphics, 41 (4) : 1-15. [DOI: 10.1145/3528223.3530127]
- Schönberger J L and Frahm J M. 2016. Structure-from-Motion Revisited//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 4104-4113. [DOI: 10.1109/CVPR.2016.445]
- Sturm J, Engelhard N, Endres F, Burgard W and Cremers D. 2012. A Benchmark for the Evaluation of RGB-D SLAM Systems//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal: IEEE: 573-580. [DOI: 10.1109/IROS.2012.6385773]
- Tao X, Gao H, Shen X, Wang J and Jia J. 2018. Scale-Recurrent Network for Deep Image Deblurring//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8174-8182. [DOI:10.1109/CVPR.2018.00853]
- Wang P, Zhao L, Ma R and Liu P. 2023. BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 4170-4179. [DOI: 10.1109/CVPR52729.2023.00406]
- Wang P, Zhang H, Li M and Zhao Q. 2024. Real-World Image Deblurring: Challenges and Prospects. Journal of Image and Graphics. [DOI:10.11834/jig.230802]
- Whyte O, Sivic J, Zisserman A and Ponce J. 2010. Non-Uniform Deblurring for Shaken Images. International Journal of Computer Vision, 98(2): 168-186. [DOI:10.1007/s11263-011-0502-7]
- Xing J K and Xu K. 2024. Survey of Physics-Based Differentiable Rendering. Journal of Image and Graphics, 29 (10) : 2912-2925. [DOI:10.11834/jig.230715]
- Yu Z, Chen A, Huang B, Sattler T and Geiger A. 2024. Mip-Splatting: Alias-Free 3D Gaussian Splatting//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 19447-19456.
- Zhao L, Wang P and Liu P. 2024. BAD-Gaussians: Bundle Adjusted Deblur Gaussian Splatting//Computer Vision - ECCV 2024. Cham: Springer: 233-250. [DOI: 10.1007/978-3-031-72698-9_14]

作者简介

- 丘文安,男,硕士研究生,主要研究方向为计算机视觉和三维重建。E-mail: qiuwenan@stu.ynu.edu.cn
- 李华光,通信作者,男,讲师,主要研究方向为计算机视觉和三维重建。E-mail: huagl@ynu.edu.cn
- 钱文华,男,教授,主要研究方向为计算机视觉、图形图像处理、文化计算。E-mail: qwhua003@sina.com
- 刘朋,男,讲师,主要研究方向为计算机视觉和轻量化模型设计。E-mail: liupeng@ynu.edu.cn