

复现论文题目：3D Gaussian Splatting for Real-Time Radiance Field Rendering

1. 背景介绍

网格和点云是最常见的 3D 场景表示，因为它们是显示的，非常适合基于 GPU/CUDA 的快速光栅化。相反，最近的 NeRF 方法建立在连续场景表示的基础上，通常使用体积射线来优化 MLP，用户捕获场景的新视图合成。虽然这些方法的连续性有助于优化，但渲染所需的随机采样成本高昂，并且可能导致噪声。本文引入了一种新的方法，提出的 3D 高斯表达能够以 SOTA 的视觉质量和快速的训练时间进行优化，而本文基于 Splatting 的渲染方案能够以 1080p 分辨率进行 SOTA 质量实时渲染。

我们的目标是允许对多张照片所捕捉的场景进行实时渲染，并创建一种场景表达，使得在经典真实场景下该种场景表达的优化速度和过去最高效的方法一样快。最近一些工作能够达到很快的训练速度，但同时难以达到目前 NeRF 方法的 SOTA 质量，这种训练速度快但质量较低的辐射场方法可以根据场景实现交互式渲染，但无法实现高分辨率场景的实时渲染。

本文的解决方案主要有三个组成模块。首先选择以 3D 高斯作为场景表达元素，输入数据为 SfM 稀疏点云和相机位姿，而不像大多数基于点云的方法需要 MVS 作为输入。第二个关键点在于对 3D 高斯属性的优化，根据高斯点对当前所处场景的表达能力对其属性进行调整。第三点在于本文受基于 tile 的渲染方式启发而开发的实时渲染方案。

2. 主体方法框架

本文方法的输入是一组静态场景的图像，以及由 SfM 校准的相应相机位姿，利用 SfM 可以产生稀疏点云。根据这些点云可以创建一组 3D 高斯，其由坐标、协方差矩阵和不透明度来定义。本文方法通过 3D 高斯参数的一系列优化步骤来创建辐射场表示。本文算法的效率凸出之处关键在于基于瓦片的光栅化器，它允许各向异性的 α 混合。Pipeline 如下图 1 所示

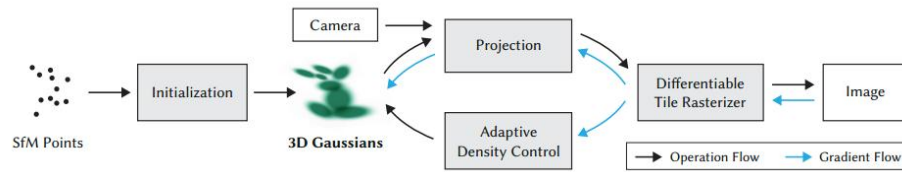


图 1 3D Gaussian Splatting Pipeline

Initialization 将 SfM 生成的稀疏点云初始化为 3D 高斯表达，如下公式 1 和公式 2 所示

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

$$\Sigma = RSS^T R^T \quad (\text{公式 2})$$

Adaptive Density Control 本文提出了针对 3D 高斯的优化方法，并同时进行自适应密度控制，存储在点中的高斯参数会在反向传播时进行优化更新。本文根据 3D 高斯位置参数的梯度来自适应调整点云分布，对于过大的 3D 高斯，无法很好地拟合该处的细节，就对该点进行分割，用分割后的两个点来表达；若某处的 3D 高斯太密集，则将多个 3D 高斯合并成为一个新的 3D 高斯，如下图 2 所示

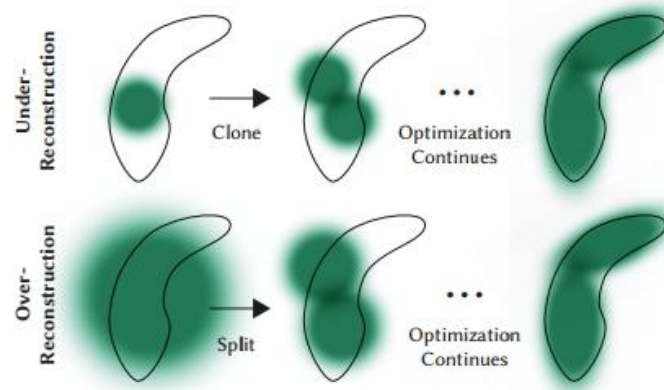


图 2 自适应密度控制

Tile-based Rasterizer 直接采用 Splatting 方法显然无法实现高帧率的渲染效率，本文的渲染目标是快速实现整体渲染和排序，从而实现近似 α -blending 并且不再限制能够接收梯度的 splats 数量，为了达到目的，对 Gaussian splats 进行分块处理，主要步骤如下

1. 将整个图像划分为 16×16 个 tiles，每个 tile 视锥内挑选可视的 3D Gaussian。
2. 每个视锥内只取置信度大于 99% 的高斯，并将 3D 高斯实例化，按高斯对象中包含的深度信息进行排序。
3. 并行地在每个 tile 进行 splat。
4. 有像素的不透明度达到饱和就停止对应线程。
5. 反向传播误差时按 tile 对高斯进行索引。

3. 实验结果

对论文中所给的数据集进行测试，如下所示测试了 playroom 和 train 两个

数据集以及一个自己建立的 Hisense 数据集。

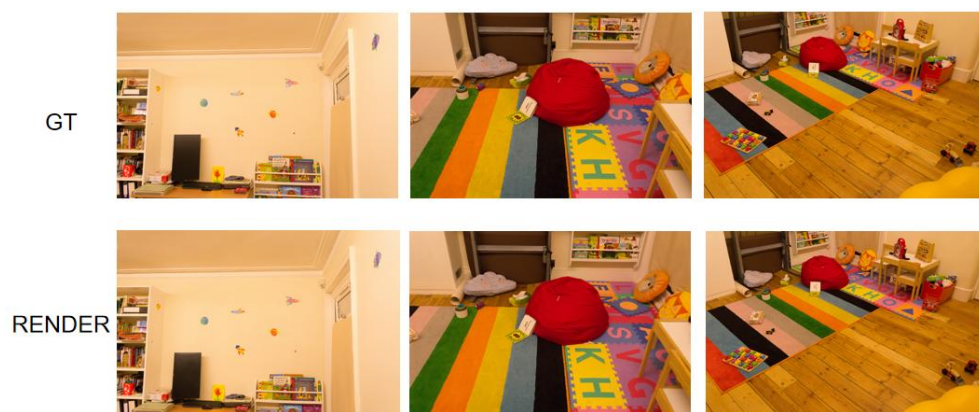


图 3 playroom 数据集测试结果



图 4 train 数据集测试结果

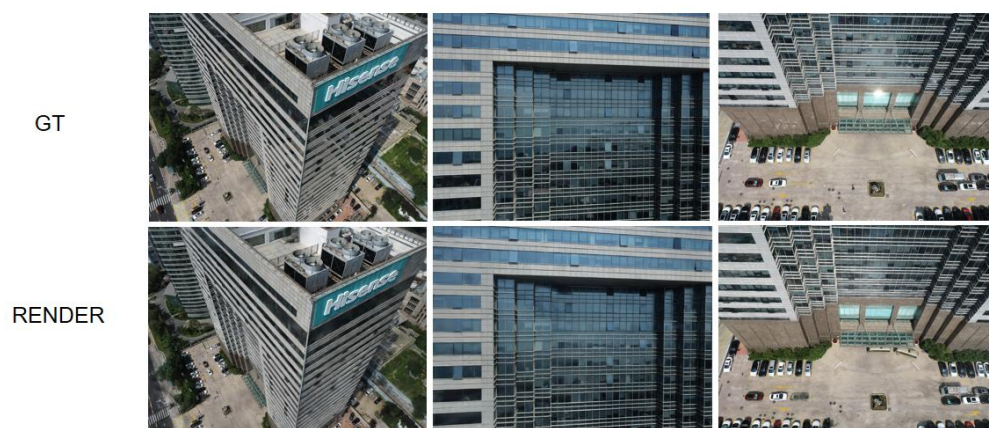


图 5 Hisense 数据集测试结果

4. 结论以及改进思路

3D Gaussian Splatting 是一种用于实时渲染的光栅化技术，同 NeRF 一样，可以对从小图像样本集中学习到的逼真场景进行极其快速的渲染。但 3D Gaussian Splatting 本质与 NeRF 还是有比较大的区别，前者更偏向于光栅化的渲染，后者则有点类似于 ray tracing 的渲染方式。并且由于 3D Gaussian Splatting 的生成结果是点云的显式表达，因此在有遮挡物体建模、生成静态场景的动画等情境下更占优势。

本文选择的 3D 高斯基元保留了体渲染的特性以进行优化，同时直接允许 splatting 的快速光栅化，该工作表明，与广泛接受的观点相反，场景的连续表达并不是实现快速高质量辐射场训练的严格必要条件。

本文证实了基于实时渲染原理、基于 GPU 的强大功能以及软件光栅化流水线速度构建架构的重要性。这些设计选择是训练和实时渲染性能的关键，与以前的体积射线相比，它们在性能上具有竞争力。

该工作也同样有局限性，对于输入图像没有很好捕获到的场景区域以及反光区域，重建结果中可能会出现伪影。同时，大量的 3D 高斯会消耗巨大的显存，对硬件设备的要求较高。

参考文献：

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).