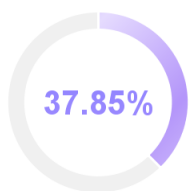


NO. 53ee24e364bf979b | 2025-04-27 20:18:31

- 题目：刘杨毕设初稿
- 作者：刘杨
- 检测所属单位： -

📄 论文字符数：29171    📖 论文页数：46    📊 表格数量：2    🖼️ 图片数量：12

## 检测结果



**37.85%**

全文疑似AIGC生成

**62.15%**

全文人写概率

## 结果分布

序号	章节	AI生成文字/章节总字数	AI生成章节占比	人工占比
1	摘要	0/789	0.0%	100.0%
2	绪论前言	4473/12887	34.0%	66.0%
3	一、球谐函数概述	84/289	29.0%	71.0%
4	二、球谐函数在3DGS中的应用	110/310	35.0%	65.0%
5	三、SH建模在渲染中的实际应用	2332/4455	52.0%	48.0%

\*注:格式规范的情况下可准确识别章节,若论文中无章节,可能会识别有误。

片段分布



文字标注

■ 自写片段    ■ 疑似AI生成

刘杨毕设初稿



本科毕业论文（设计）

题目： 基于高斯泼溅的三维重建系统设计与实现

学 院： 信息科学与工程学院

专 业： 人工智能

学 生 姓 名： 刘杨

学 号： 632102090923

指 导 教 师： 杨庆

评 阅 教 师：

完 成 时 间：

重庆交通大学

CHONGQING JIAOTONG UNIVERSITY

本科毕业论文（设计）原创性声明

本人郑重声明：所提交的毕业论文（设计），是本人在导师指导下，独立进行研究工作所取得的成果。除文中已注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文研究做出过重要贡献的个人和集体，均已在文中以明确方式标明。

本人完全意识到本声明的法律后果由本人承担。

作者签名（亲笔）： 年 月 日

## 本科毕业论文（设计）版权使用授权书

本毕业论文（设计）作者完全了解学校有关保留、使用学位论文的规定，本科生在校攻读期间毕业论文（设计）工作的知识产权单位属重庆交通大学，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权重庆交通大学可以将毕业论文（设计）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编毕业设计（论文）。

作者签名（亲笔）： 年 月 日

导师签名（亲笔）： 年 月 日

## 摘 要

随着计算机视觉以及三维重建技术持续取得发展，从多视图图像当中自动去还原出高精度的三维场景，已经成为了研究领域的一个热点所在。传统的那些用于三维重建的方法，往往是要依靠对图像特征来做精确的匹配，并且要构建出稠密的点云，之后再据此去推断场景的几何结构。然而，此类技术在复杂光照环境或反射性材质等非朗伯面情况下表现不佳，且难以实现自由视角下的真实感渲染。基于神经网络的重建方法，又面临训练时间长，渲染消耗大，难以落地。在2023年的时候，Kerbl等人提出了一种三维高斯泼溅方法，也就是3DGS，这个方法有效地对NeRF的效率短板进行了弥补，从而为三维场景建模以及实时渲染开辟出了新的可能性。

此篇论文着重围绕稀疏点云的构建以及优化方面展开探讨，同时也针对以高斯泼溅为基础的高质量三维重建的相关方法来深入开展研究工作，其意图在于搭建起一个能够实现高效化、真实感颇为突出、且具备可微优化特性的三维场景重建框架体系。

我使用COLMAP算法以及开源软件来做完图像数据的预处理工作，包括特征提取，特征匹配优化，增强点云的结构一致性与空间分布的完整性，最后生成稀疏点云数据。在此基础上，构建高斯点云表示，利用各点的空间位置、各向异性尺度、透明度和颜色等参数进行建模。

本文详细探讨了基于 3DGS 的可微渲染方法，引入球谐函数（Spherical Harmonics）对点的反射属性进行编码，实现方向感知的真实光照建模。该方法不仅具有良好的可视化效果，同时也便于在训练中通过反向传播优化高斯点的参数。最后，论文还探讨了高斯点的初始化策略、训练管线设计、优化与压缩策略，以实现轻量化、高精度、高效率的三维重建。

本文所构建起来的那个依托于3DGS的三维重建系统，成功达成了快速渲染这一目标，并且还能够实现可交互式展示。如此一来，便给游戏开发领域、数字孪生方面以及虚拟现实等诸多应用开拓出了全新的技术路径。

关键词为三维重建、多视图几何、稀疏点云、高斯泼溅、可微渲染以及球谐函数。

The Subject of Undergraduate Graduation Project (Thesis) of DUT

## Abstract

With the continuous development of computer vision and 3D reconstruction technology, automatically restoring high-precision 3D scenes from multi-view images has become a hot topic in the research field. Traditional 3D reconstruction methods often rely on precise matching of image

features and the construction of dense point clouds to infer the geometric structure of the scene. However, such techniques perform poorly in complex lighting environments or with reflective materials (non-Lambertian surfaces) and are difficult to achieve realistic rendering from arbitrary viewpoints. Neural network-based reconstruction methods, on the other hand, face challenges such as long training times, high rendering costs, and difficulty in practical application. In 2023, Kerbl et al. proposed a 3D Gaussian Splatting (3DGS) method, which effectively addresses the efficiency shortcomings of NeRF, opening up new possibilities for 3D scene modeling and real-time rendering.

This paper focuses on the construction and optimization of sparse point clouds and conducts in-depth research on high-quality 3D reconstruction methods based on Gaussian splatting. The aim is to establish a 3D scene reconstruction framework that is efficient, visually realistic, and supports differentiable optimization.

I used the COLMAP algorithm and open-source software to preprocess the image data, including feature extraction, feature matching optimization, enhancing the structural consistency and spatial distribution integrity of the point cloud, and finally generating sparse point cloud data. On this basis, a Gaussian point cloud representation was constructed, using parameters such as the spatial position, anisotropic scale, transparency, and color of each point for modeling.

This paper elaborates on the differentiable rendering method based on 3DGS, introducing spherical harmonics to encode the reflection properties of points, achieving direction-aware realistic lighting modeling. This method not only has good visualization effects but also facilitates the optimization of Gaussian point parameters through backpropagation during training. Finally, the paper also discusses the initialization strategy of Gaussian points, training pipeline design, optimization, and compression strategies to achieve lightweight, high-precision, and high-efficiency 3D reconstruction.

The 3D reconstruction system based on 3DGS constructed in this paper successfully achieves the goal of fast rendering and can also support interactive display. This opens up new technical paths for applications such as game development, digital twins, and virtual reality.

The key words are 3D reconstruction, multi-view geometry, sparse point cloud, Gaussian splatting, differentiable rendering, and spherical harmonics.

目 录

摘 要II

AbstractIII

1 绪论1

1.1 研究背景与意义1

1.2 国内外研究现状2

1.2.1 三维重建技术研究现状	2
1.2.2 3DGS研究现状	2
1.3 研究目标与内容	3
1.4 章节安排	4
2 三维重建相关技术综述	5
2.1 三维重建基础理论	5
2.1.1 多视图几何	5
2.1.2 相机模型与标定	7
2.1.3 常见的三维重建表达方式	9
2.2 基于神经辐射场的体素渲染	10
2.2.1. 核心思想与数学原理	10
2.2.2. NeRF的训练过程:	12
2.2.3. NeRF 的优缺点	12
3 图像数据预处理与稀疏点云重建	13
3.1 基于COLMAP的相机位姿估计	13
3.1.1 相机模型与投影几何	13
3.1.2 特征提取与匹配	13
3.1.3 增量式 SfM 重建流程	14
3.1.4 输出格式与可视化	14
3.2 稀疏点云生成	14
3.2.1 三角测量与深度恢复	14
3.2.2 稀疏点云的特性与局限性	14
3.3 基于点云神经网络的系数点云优化	15
3.3.1 点云处理挑战	15
3.3.2 PointNet架构原理	15
3.3.3 在稀疏重建中的应用	15
3.3.4 训练与部署流程	16
4 基于高斯泼溅的三维重建方法	17
4.1 高斯泼溅表示	17
4.1.1 高斯分布的数学模型	17
4.1.2 向量化与效率优化	17
4.2 可微渲染过程	18
4.2.1 渲染原理	18
4.2.2 可微渲染计算与梯度传播	19

4.2.3 引入球谐函数进行光照建模	19
4.3 参数初始化与训练策略	21
4.3.1 初始化策略	22
4.3.2 训练损失与优化目标	22
4.3.3 多阶段训练策略	23
4.4 优化与高斯压缩策略	23
4.4.1 点密度增长策略 (Density-Guided Splitting)	23
4.4.2 点合并与压缩策略 (Gaussian Pruning)	24
4.4.3 性能优化与加速策略	24
4.4.4 压缩率与质量的权衡	24
5 总结与展望	25
5.1 全文总结	25
5.2 展望	26
致 谢	27
参 考 文 献	28

## 1 绪论

### 1.1 研究背景与意义

三维重建技术在计算机视觉以及图形学领域属于核心的研究方向，主要的任务就是从二维图像或者传感器数据当中把三维场景的几何结构以及外观信息给恢复出来。传统的一些方法呢，主要是运用显式几何表示的方式，像是点云、网格、体素这些，或者也会采用隐式神经表示，就比如神经辐射场NeRF，但各自存在显著局限性[1-6]。

显式几何表示固然可以对场景结构予以直接表达，在针对复杂材质以及光照变化展开建模工作的时候，它是存在着诸多不足之处的。它还对多视图间的精确配准有着高度的依赖性，就使得整体的流程变得颇为复杂，计算效率也呈现出低下的状况。而隐式神经表示是借助多层感知机来对辐射场实施隐式编码的，通过这样的方式能够生成高保真的渲染效果。但它的训练过程是极为耗时的，渲染速度也相当缓慢，还难以对场景编辑以及实时交互提供有力的支持，这一系列的情况也就限制了它在工业场景当中的实际应用程度。

在这背景之下，3DGS技术被提出来，这已然成为了可微渲染领域当中极为重要的一个突破点。3DGS它是利用显式高斯点云的方式来针对场景展开建模工作的，并且还和可微分光栅化管线相互结合起来。它有能与NeRF相媲美的渲染质量，而且还把训练时间大幅缩减，缩减到了分钟这样的级别，与此同时，其渲染帧率也得到了极大提升，能够提升至数百FPS，这样计算效率以及场景可控性都有了明显的提高。靠着快速渲染以及轻量化这两大优势，该技术已经快速地在多个领域当中得到了应用，有力地推动了像移动增强现实（AR、VR）、数字文化遗产等这些实际应用不断向前发展。

随着元宇宙、数字孪生以及智能机器人这些前沿领域不断快速发展，实时高保真三维重建的需求正变得日益迫切起来。就比如说，工厂场景的三维重建工作得在秒级时间内完成，并且还要能够支持交互式编辑操作；在医疗领



域当中，为了能够还原器官的解剖结构，就要求达到亚毫米级的重建精度；而对于移动应用而言，则需要那种具备低功耗特点的轻量化算法，像3D Moment App就已经实现了用户级别的三维内容创作。

3DGS所构建的三维重建系统，一方面有着重要的理论层面的研究价值，另一方面也实实在在地给实际应用给出了高效的解决办法。它的不断发展，是很有希望促使下一代三维感知、交互以及生成技术取得突破的。所以说，去展开针对3DGS技术的研究以及完成系统的实现工作，这在学术方面有着重要意义，在产业方面也有着广阔的发展前景[7]。

## 1.2 国内外研究现状

### 1.2.1 三维重建技术研究现状

三维重建属于计算机视觉跟图形学之间极为重要的交叉领域方向，一直以来都是运用结构光、双目视觉以及激光雷达等多种多样的方法去达成对场景三维信息的重建工作。在这之中，运动恢复结构（Structure from Motion, SfM）还有多视图立体（Multi-View Stereo, MVS）这两项技术实际上构成了传统三维重建流程里面的核心部分[8]。这类方法大体上是要依靠图像特征的精准匹配操作，进而构建出稠密的点云，再依据这些来推断出场景具体的几何结构。要是处于复杂的光照环境之下，或者是面对具有反射性、非朗伯特材质的情况时，这些技术所展现出来的效果往往就不是那么理想了，并且想要在自由视角下实现那种具有真实感的渲染也是比较困难的事情[9]。

当可微渲染这一概念被提出来之后，研究者们便着手借助神经网络，采用端到端的方式去对三维结构和图像的一致性加以优化。在这当中，神经辐射场（Neural Radiance Fields, NeRF）的诞生可以说是一个极为关键的转折点。NeRF利用一个多层感知机（MLP）网络，能够把空间里的任意一个点映射成为相应的颜色以及密度值，并且与体积渲染过程相结合[10]，进而从静态图像当中完成对连续场景表示的重建工作。这种方法在新视角图像合成这个方面已然取得了颇为显著的进展，在建筑建模、虚拟现实、文化遗产数字化等等诸多领域都得到了广泛的应用[11]。

不过，NeRF模型明显存在着两个局限之处：其一，该模型所采用的隐式表示方式在可解释性以及可编辑性方面是有所欠缺的；其二，它需要依靠大量的光线采样以及前向网络计算，如此一来便使得训练周期被拉长了，而且渲染效率也比较低，很难去满足实时渲染以及交互式应用所提出的要求。针对这些存在的问题，学术界开始一步步地尝试引入一些更为高效的场景表示形式，像是体素网格、有符号距离函数这类显式结构。尽管有所突破，但在表达细节与效率之间仍面临取舍。

### 1.2.2 3DGS研究现状

在2023年的时候，Kerbl等人提出了三维高斯泼溅方法。这一方法成功弥补了NeRF在效率方面存在的不足，给三维场景建模以及实时渲染开拓出了新的发展方向。该方法是把场景通过一种方式表示出来，具体而言，是将其表示为一组带有高斯分布特性的点，并且运用了可微光栅化渲染管线，如此一来，便极大地提高了训练的速度，在推理效率方面也有明显提升，还很好地保留了高质量的渲染效果。和NeRF不一样的是，NeRF需要针对整条光线展开密集的采样操作，而且还要多次调用MLP来进行推理，而3DGS则是直接把高斯点投影到图像平面之上，然后再对其进行加权融合，通过这样的操作，有效地削减了计算方面的开销[12]。

3DGS的显式表示一方面增强了模型的可控性，另一方面还便于和传统图形学管线集成起来，进而拓展了它在实际应用当中的可行性，像城市级别场景、动态重建以及SLAM系统等方面都是如此。就拿VastGaussian来说，它引入了依据可见性判断的空间划分策略，从而实现了大规模城市场景的高效重建。而4D-GS则是把HexPlane编码与轻量级

MLP结合到一起，达成了对动态场景的时空建模。GS-SLAM通过高斯点的自适应更新机制，在实时定位与地图构建方面，跟NeRF-SLAM相比，展现出了更为出色的精度与速度表现[13]。

虽说当下3DGS技术已经有了颇为明显的发展进程，可依旧存在着一些有待攻克挑战。就好比说，在那种大规模的场景状况之下，针对高斯点展开的存储工作以及相关管理事宜，是很有可能引发较为明显的内存方面的耗费情况的；而处于像复杂材质、镜面反射这类并非十分理想的场景当中的时候，它在针对光照物理展开建模的这一能力层面上，还是有进一步提升的必要的；除此之外，当下其在语义理解方面所具备的能力还是比较有限的，如此一来，便极有可能致使在开展虚拟世界建模这项任务的过程当中，在几何与语义两者之间出现不相符、不一致的情况。为了能妥善处理这些存在的问题，相关的研究者们正在努力尝试着从像是模型压缩、物理建模以及多模态语义融合等这样的一些不同角度出发，进而提出一些能够起到改进作用的方案。比如LightGaussian，它借助八叉树结构来对高斯点集进行压缩处理，以此达到降低内存消耗的目的，而DrivingGaussian则是通过独立开展建模工作并且确保时间上的一致性，从而实现了动态对象进行精确渲染的效果，进一步拓宽了3DGS技术在实际应用当中的范围边界[14]。

从整体情况来讲，3DGS技术已然变成了在三维重建以及新视角渲染领域里的研究热点所在。它在重建精度这块，还有效率方面以及可控性等方面所具备的优势，实实在在地为构建下一代真实虚拟环境打下了牢固的根基。

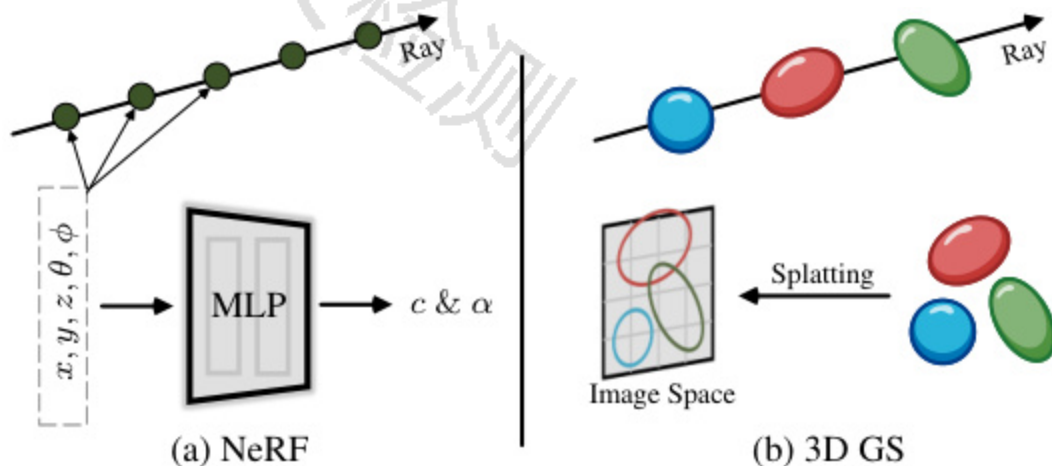


图 1.1 Nerf和3DGS原理示意图

### 1.3 研究目标与内容

本文对三维高斯泼溅（3DGS）技术的三维重建系统展开研究，关键目标是要处理现有方法在复杂场景当中所存在的效率、可扩展性以及适应性方面的诸多问题，达成更高精度的三维重建渲染效果，同时让系统的计算性能得以优化[15]。

我着重对3DGS技术的核心算法展开研究，同时剖析它在不同场景当中的具体表现，进而提出与之对应的优化方案。会细致探讨高斯点生成、存储、渲染以及合成的整个过程，还会摸索怎样借助算法的优化来提高重建以及渲染的效率。设计出一套完整的三维重建流程，涵盖从数据采集、做预处理，再到高斯点生成以及可微渲染等环节，以此来达成一个高效的三维重建框架。研究在大规模场景里面计算成本和重建精度二者间的平衡状态，确保系统能够高效地运转起来[16]。



本研究还打算探讨怎样把深度信息、纹理信息以及光照模型等多模态数据引入到3DGS当中，以此强化系统针对复杂场景予以理解以及表达的能力。研究把语义分割、光照模型还有材质估计技术相互结合起来，促使三维重建的结果朝着更高精度以及更高真实感的方向去发展。除此之外，还提出一种依据分层存储与空间划分数据结构所制定的优化方案，该方案是用来优化大规模场景重建过程里的内存管理以及渲染性能的，并且要保证实时渲染的性能不会受到任何影响。本研究还会去探索点云优化、并行计算以及GPU加速等技术在场景重建方面的具体应用[17]。

## 1.4 章节安排

本文围绕基于高斯泼溅（3D Gaussian Splatting, 3DGS）的三维重建系统展开研究，结合 COLMAP 工具获取的稀疏点云数据，通过对高斯渲染模型的解析与优化，实现三维场景的可视化重建与Web端实时展示。全文主要分为六章，具体安排如下：

第一章，绪论。本章介绍本研究的背景与意义，阐述当前国内外在三维重建和3DGS领域的研究现状，明确本文的研究目标与内容，并概述各章节的结构安排[18]。

第二章，三维重建相关技术综述。本章系统回顾了与本课题密切相关的三维重建基础理论，包括多视图几何、相机模型与标定方法等，并简要介绍了基于体素的神经渲染方法（如 NeRF）及其发展趋势，为后文的技术选型提供理论支持。

第三章，图像数据预处理与稀疏点云重建。本章聚焦于前期数据处理过程，详细介绍如何基于 COLMAP 工具完成图像的特征提取、相机姿态估计与稀疏点云生成，并进一步探讨了点云优化方法及其在稀疏重建中的应用[19]。

第四章，基于高斯泼溅的三维重建方法。本章作为本文的核心部分，系统阐述了3DGS的建模与训练方法，包括高斯分布的数学建模、可微渲染原理、球谐光照建模、训练损失设计与多阶段优化策略，并探讨了点云压缩与性能优化技术[20]。

第五章，三维重建系统设计与实现。本章介绍了整个三维重建系统的实现过程，主要从系统功能需求出发，设计前后端架构，并详细说明了后端文件管理与任务调度逻辑、前端基于 React + Three.js 的场景渲染流程、高斯点可视化以及用户交互界面设计。

第六章 总结与展望。本章总结了本文的研究工作，分析系统的实际效果与不足之处，并对未来在高质量重建、编辑功能扩展以及系统部署等方面的发展方向进行了展望。

## 2 三维重建相关技术综述

三维重建技术是从多维度信息中推导出三维几何形态的过程。近年来，随着计算机视觉、图形学和深度学习的快速发展，三维重建方法不断得到优化和拓展。三维重建不仅在虚拟现实（VR）、增强现实（AR）、机器人导航和自动驾驶等领域具有重要应用，也在数字孪生、医学影像、文化遗产保护等领域发挥着重要作用。为了实现高质量的三维重建，需要掌握从图像信息到三维模型的转换方法，涉及多个学科的交叉，包括多视图几何、相机模型与标定、点云与网格表示等基础理论[21]。

### 2.1 三维重建基础理论

三维重建的基础理论为从多视图图像中恢复三维场景的几何结构提供了坚实的理论框架。本节将介绍三维重建中的三大核心理论：多视图几何、相机模型与标定、以及常见的三维重建表达方式。

#### 2.1.1 多视图几何

多视图几何 (Multi-view Geometry) 是计算机视觉中的一个重要研究方向,旨在通过从不同视角拍摄的多张图像,恢复出真实世界中场景的三维结构与摄像机的运动轨迹。这一技术是现代三维重建、增强现实、无人驾驶导航、文物数字化等应用的核心基础[22]。

从本质上来说,多视图几何是在二维图像的基础上,利用图像之间的几何对应关系和投影模型,反推出三维空间中物体的形状、相对位置与相机参数。它建立了图像空间与三维空间之间的映射关系,是三维计算机视觉中的桥梁[23]。

多视图几何主要包括以下几个核心技术模块:

1. 特征提取与匹配:

多视图几何的第一步是从图像中提取稳定、具有描述性的局部特征点。常用的特征提取算法包括 SIFT (尺度不变特征变换)、ORB、SURF 等。提取后的特征会被用于图像之间的匹配,通过匹配可以建立图像间的点对应关系。这些对应点是后续几何推理与三维重建的基础。

2. 立体匹配与视差估计:

立体视觉 (Stereo Vision) 是多视图几何中的一个基本模型。它通常基于两张或多张从不同角度拍摄的图像,通过特征点或像素之间的匹配,计算图像间的视差 (Disparity)。由于图像中同一点在不同视角下的投影位置不同,这种视差可以用来反推出该点的深度信息。视差越大,点距离相机越近。视差图的质量直接决定了后续三维重建的精度。

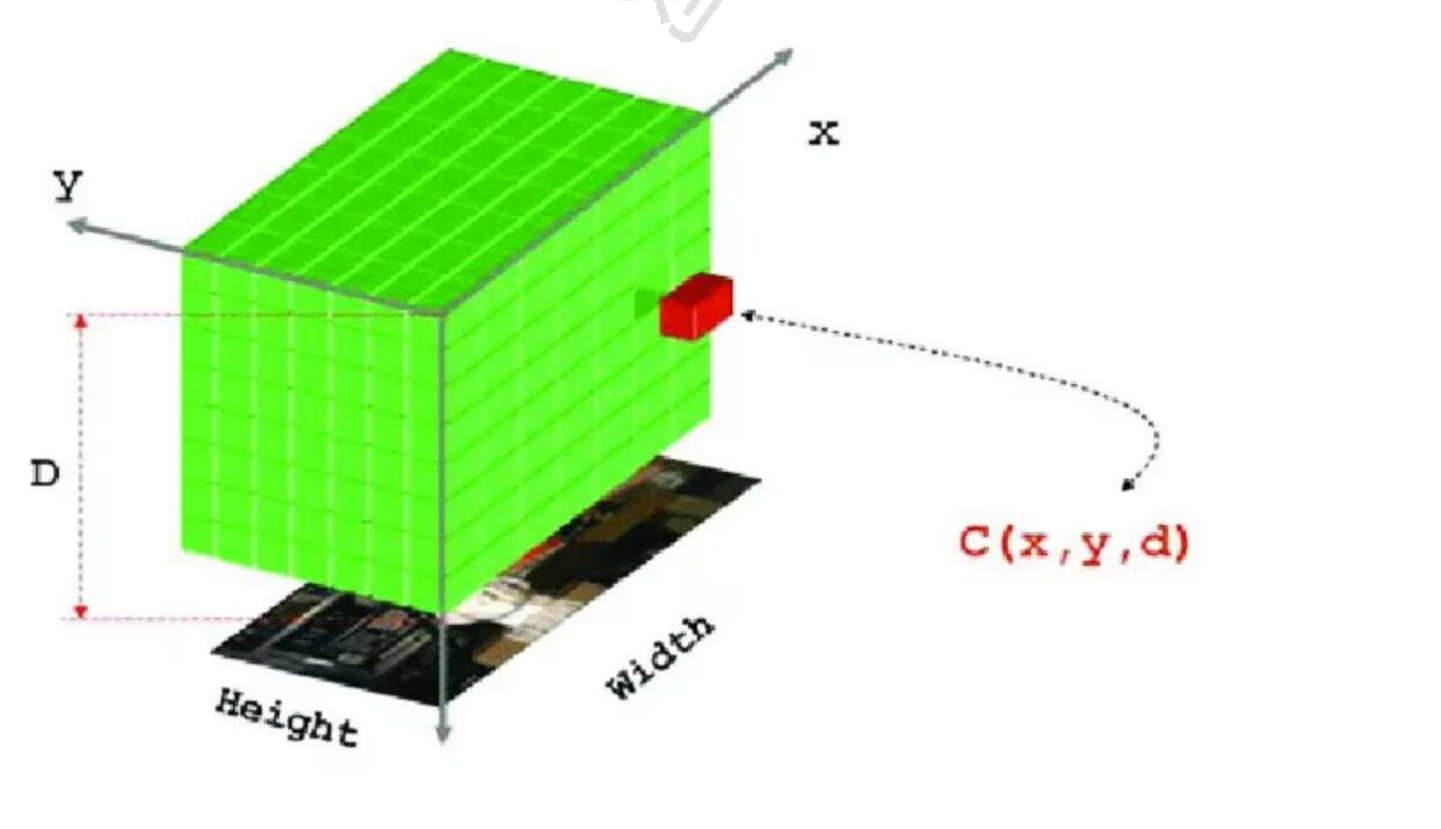


图2.1 视差空间示意图

3. 视图几何约束关系:

图像之间的几何关系可用一系列数学工具来刻画，其中最典型的包括：基础矩：描述两个视图对应点之间的约束，满足极线几何关系；本质矩阵：在已知相机内参情况下，描述相机之间的相对运动；单应性矩阵：在同一平面上的点之间的映射关系，常用于平面场景的配准；极几何：描述一对图像之间的几何对应关系，是进行匹配和约束搜索的重要依据。这些几何约束关系可通过 RANSAC 等稳健估计方法从匹配点对中计算得出，进一步用于三角测量和相机姿态估计。

4. 三角测量与三维点恢复：

一旦建立了图像间的匹配关系并获得相机的相对姿态，就可以通过三角测量方法将二维图像中的匹配点反推出三维空间中的位置。三角测量的基本原理是利用两个视图中同一三维点的投影，通过几何交会求解其空间坐标。数学上，这可以通过求解一组线性方程组或最小化重投影误差来实现。最终获得的三维点集构成了场景的稀疏点云，为后续表面重建或三维模型优化提供了基础数据[24]。

5. 运动恢复结构 (Structure from Motion, SfM)：

SfM 是多视图几何中的经典方法，目标是在无需已知相机位姿的前提下，从多张图像中联合恢复三维场景结构 (Structure) 与相机的运动轨迹 (Motion)。其核心流程包括：图像特征提取与全局匹配；增量式或全局式的相机位姿估计；三角测量生成稀疏点云；捆绑调整进行全局非线性优化，提高重建精度。SfM 通常输出稀疏三维点云、每张图像对应的外参矩阵 (位姿) 和内参参数[25]，为后续稠密重建 (如 MVS) 和建模提供基础。本系统使用的是 COLMAP 来获取稀疏点云数据，它是是一个基于 Structure-from-Motion (SfM) 和 Multi-View Stereo (MVS) 的开源三维重建工具箱。

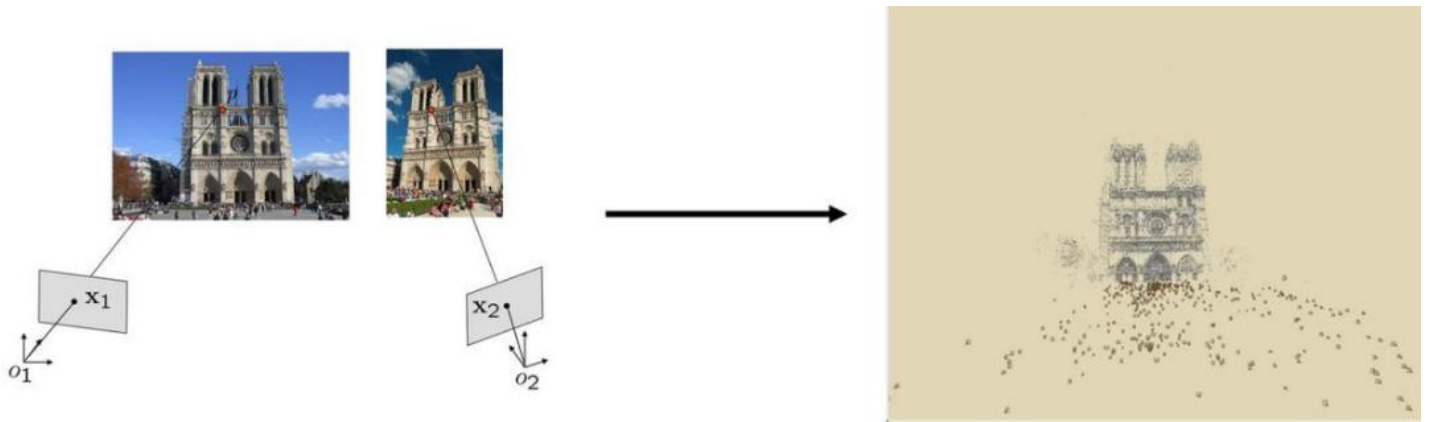


图2.2 运动恢复结构 (SfM) 示意

6. 多视图几何的实际应用：

多视图几何作为三维视觉的基础，在多个实际领域中有着广泛的应用，包括但不限于：摄影测量与地图制作：利用航拍图像重建大范围地形；机器人视觉与 SLAM：实现视觉里程计与三维定位；增强现实 (AR) 与虚拟现实 (VR)：提供精确的场景建模与相机追踪；文化遗产数字化：对文物、古迹等物理对象进行高精度三维采集；自动驾驶与环境感知：从多个摄像头中获取车辆周围环境的三维结构[26]。

2.1.2 相机模型与标定

相机模型与标定是三维重建系统中的基础性环节，其核心目的是准确描述三维世界中的点如何通过摄像机投影

到二维图像上。只有在相机模型准确、标定结果精确的前提下，才能保证后续的几何恢复和三维重建的精度。因此，相机模型和标定方法不仅是图像投影计算的数学依据，也是三维重建系统误差控制的关键所在。

(一) 相机模型的基本概念

相机模型用于模拟现实相机的成像过程，即描述三维空间中的点如何经过透视投影映射到二维图像平面上。在计算机视觉中，最常用的模型是针孔相机模型（Pinhole Camera Model）。

(1) 针孔相机模型

针孔模型是一种理想化的简化模型，忽略了镜头的复杂光学结构与畸变效应，仅保留了投影变换的核心部分。该模型假设成像过程为线性透视投影，三维点  $[X\ Y\ Z]$  会被投影到二维图像点  $[u\ v]$  上，其关系可由以下投影矩阵  $P$  表示：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim K \cdot [R|t] \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (0.1)$$

其中：

- $K$  为相机内参矩阵，包含焦距、主点、像素比例等；
- $[R|t]$  为相机的外参矩阵，描述从世界坐标系到相机坐标系的变换；
- $\sim$  表示齐次坐标的等价变换（即忽略比例因子）。

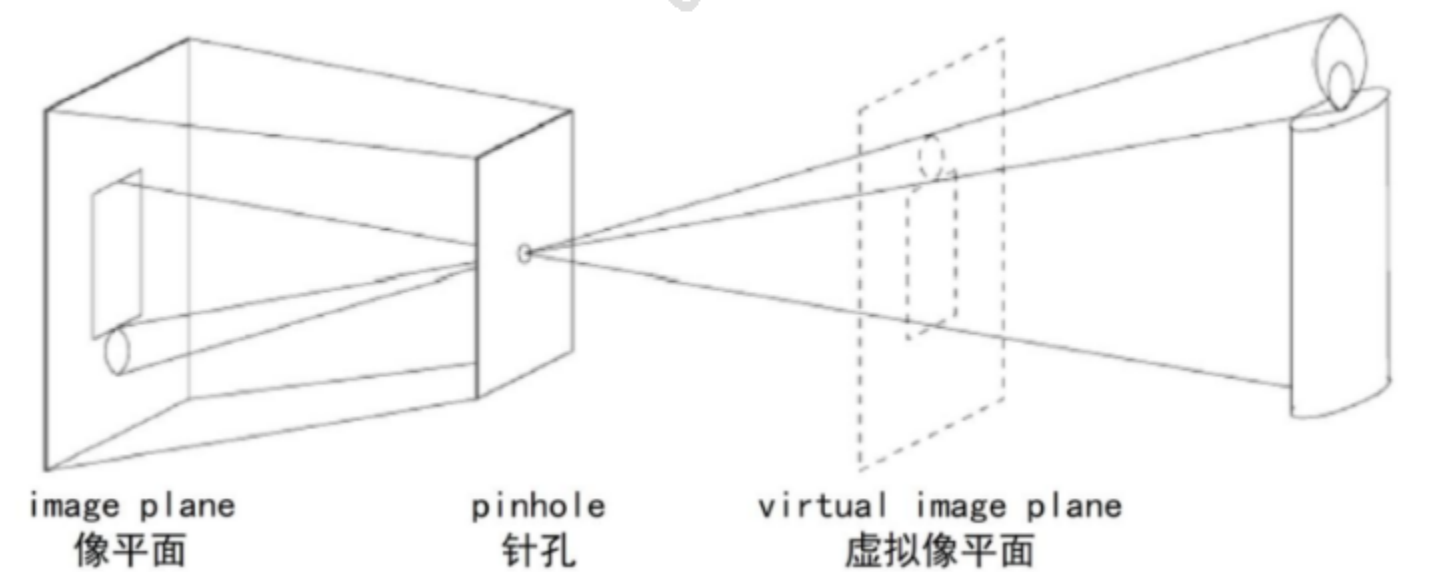


图2.3针孔照相机模型示意

(2) 畸变模型与广角镜头建模

现实世界中的相机镜头存在多种类型的成像畸变，这在使用广角镜头、鱼眼镜头等非理想成像系统时尤为明显。常见的畸变类型主要包括：

- 径向畸变：图像边缘位置会被拉伸或压缩，形成桶形或枕形失真；
- 切向畸变：由于镜头与图像平面不完全对齐导致的非对称畸变。

为了解决这类畸变带来的误差，通常会在相机模型中引入畸变系数（如  $[k_1, k_2, p_1, p_2]$  等），并在图像处理阶段进行图像去畸变操作，从而还原真实图像的几何结构[27]。

## （二）相机模型中的坐标系

相机模型存在4个坐标系：世界坐标系、摄像机坐标系、图像物理坐标系和图像像素坐标系。光心：图像的中心。光轴：穿过图像的光心，与图像平面垂直的轴。

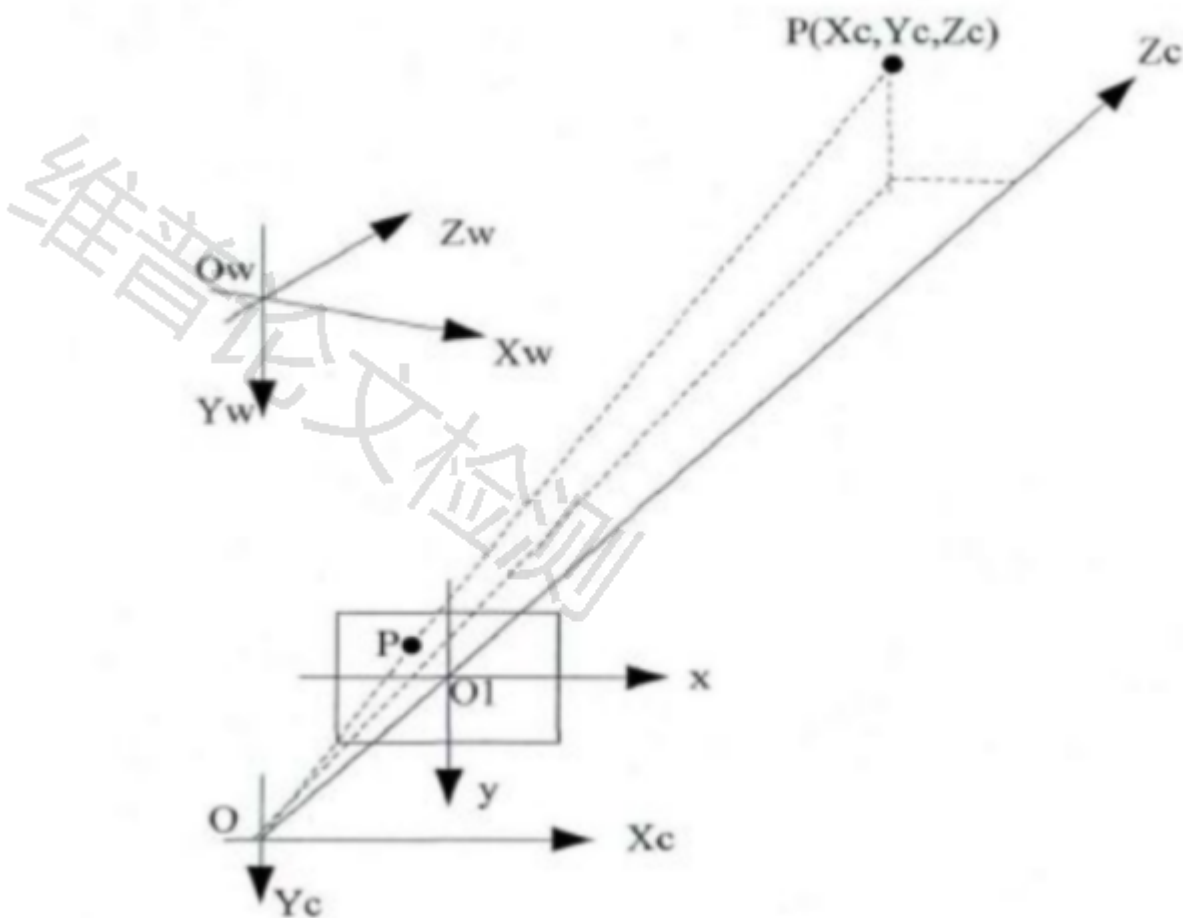


图2.4 相机模型中的坐标系

世界坐标系：是客观三维世界的绝对坐标系，也称客观坐标系，就是物体在真实世界中的坐标，世界坐标系是随着物体的大小和位置变化的，单位是长度单位

相机坐标系：以相机的光心为坐标原点，以平行于图像的x和y方向为x轴和y轴，z轴和光轴平行，x，y，z互相垂直，单位是长度单位。

图像物理坐标系：以主光轴Zc和图像平面交点为坐标原点，x和y方向如图所示，单位是长度单位。

图像像素坐标系：以图像的上/下/左/右顶点为坐标原点，u和v方向平行于x' 和y' 方向，单位是像素。

相机模型与标定是三维重建系统的理论基石，准确的模型不仅能够提供高质量的图像几何校正，更能为后续 SfM、MVS、3DGS 等重建算法提供精确的几何输入。在实际系统中，良好的相机标定往往意味着更高的重建精度、更少的误差积累，也为三维高斯泼溅等神经渲染技术提供可靠的初始化条件与视角一致性保证。



### 2.1.3 常见的三维重建表达方式

根据不同的重建方法和数据特征，三维重建结果可以采用多种不同的数据结构与表达形式。常见的三维重建表达方式包括：点云（Point Cloud）、网格模型（Mesh）、体素（Voxel Grid）、深度图（Depth Map）、神经场（Neural Field）和三维高斯泼溅（3D Gaussian Splatting）等。

#### 1、点云（Point Cloud）

点云是一种最基础、最通用的三维表示方式，由大量三维空间中的离散点组成，每个点记录其空间位置（ $x, y, z$ ），并可附带颜色、法线或语义等属性。

结构简单、易于从多视图几何算法中直接获取（如SfM、MVS等）；适合大规模场景的稀疏表示。缺乏拓扑结构，不能表示连续的表面信息；在渲染和碰撞检测中较为低效。大多应用于激光雷达扫描、SfM初始输出、3D重建初步展示等。

#### 2、三角网格（Mesh）

三角网格通过将点云中的点按照一定拓扑连接形成面片（通常为三角形）来构成表面。一个网格通常由顶点集、边集与面集组成，能较好地表达物体的几何形状。表达能力强，可用于精细建模、纹理映射与物理仿真；主流图形引擎对Mesh有良好支持。网格构建复杂，重建过程对噪声敏感；不适合表示体积数据。大多应用于游戏建模、三维打印、仿真建模、工业逆向工程等。

#### 3、体素网格（Voxel Grid）

体素（Voxel）是三维空间中均匀划分的小立方体单元，可类比于二维图像中的像素。体素网格常用于重建实体物体或体积密度场。结构规则、易于处理和分析；适合表示体积对象，如医学图像、构造块模型等。内存开销大，分辨率受限；在表示细节丰富场景时效率较低。大多应用于医学CT/MRI建模、机器人路径规划、体绘制（Volume Rendering）等。

#### 4、深度图（Depth Map）

深度图是二维图像的一种扩展，其每个像素值表示对应点到相机的深度值。通过对多张深度图进行融合，可重建出三维模型。与图像结构一致，易于采集与处理；适合多视图融合与立体视觉计算。不具备空间一致性，融合难度大；噪声敏感。大多应用于RGB-D相机采集、AR场景定位、SLAM系统、MVS重建输入等。

#### 5、神经场（Neural Fields）

神经场通过使用神经网络对空间进行隐式建模，例如神经辐射场（NeRF）等技术，通过坐标-特征映射函数建模整个三维场景。可以实现连续、无网格的高质量重建；渲染效果真实，支持新视角合成。训练开销大，渲染效率低；难以实时交互或导出显式几何。

#### 6、三维高斯泼溅（3D Gaussian Splatting）

三维高斯泼溅是一种新兴的神经渲染表达形式，通过在三维空间中分布高斯核（每个高斯拥有位置、方向、颜色、透明度等属性），并进行光栅化渲染，达到近似真实的视图合成效果。重建速度快，可实现高质量实时渲染；支持显式空间结构，便于交互。精度受限于输入图像质量；在结构稀疏区域可能存在冗余或模糊。大多应用于快速三维扫描、人脸建模、轻量级数字孪生、移动端AR渲染等[28]。

### 2.2 基于神经辐射场的体素渲染



### 2.2.1. 核心思想与数学原理

神经辐射场 (Neural Radiance Fields, NeRF) 是一种使用神经网络建模三维场景的技术, 最早由 Mildenhall 等人在 2020 年提出。NeRF 的核心思想是: 通过一个多层感知器 (MLP) 学习一个函数, 将三维空间中的位置和视角映射到颜色和体密度, 从而实现对真实场景的建模与新视角图像的合成。NeRF 不直接输出显式几何 (如点云、网格), 而是以“体积渲染”的方式间接表达三维结构, 是神经渲染领域的重要突破。NeRF 的核心是一个神经网络函数:

$$F_{\theta}(\mathbf{x}, \mathbf{d}) = (\mathbf{c}, \sigma) \quad (2.2)$$

其中:

$\mathbf{x} \in \mathbb{R}^3$ : 表示空间中的三维位置;

$\mathbf{d} \in \mathbb{R}^3$ : 表示光线的方向;

$\mathbf{c} \in \mathbb{R}^3$ : RGB 颜色;

$\sigma \in \mathbb{R}$ : 体密度, 用于体积渲染。

神经网络的任务是: 对于空间中的任意一点和视角, 输出该点的颜色和密度。

NeRF 不直接渲染像素, 而是模拟光线穿过体积时的累积效果。其颜色计算公式为:

$$\hat{\mathbf{c}}(\mathbf{r}) = \int_{t_n}^{t_f} T(t) \cdot \sigma(\mathbf{r}(t)) \cdot \mathbf{c}(\mathbf{r}(t), \mathbf{d}) dt$$

N: 采样点数量;

$\delta_i$ : 第 i 个采样点的采样间隔;

$\sigma_i$ : 第 i 个采样点的密度;

$\mathbf{c}_i$ : 第 i 个采样点的颜色;

$T_i$ : 第 i 个点未被遮挡的概率。

所谓体渲染, 直观地说, 我们知道相机的焦点, 焦点和像素的连线可以连出来一条射线, 我们可以对这条射线上所有的点的颜色做某种求和就可以得到这个像素的颜色值。理论上, 可以对这条射线经过空间上的每个点的密度 (只和空间坐标相关) 和颜色 (同时依赖空间坐标和入射角) 进行某种积分就可以得到每个像素的颜色。当每个像素的颜色都计算出来, 那么这个视角下的图像就被渲染出来了[29]。

### 2.2.2. NeRF的训练过程:

NeRF 的训练数据通常是多张具有已知相机位姿的图像。

训练步骤包括:

1. 相机参数估计: 使用 COLMAP 等工具获取相机内参与位姿;
2. 射线生成: 每个像素反投影为一条光线;
3. 采样与渲染: 在每条射线中采样若干点, 使用神经网络计算每个点的颜色和密度;
4. 损失函数: 计算合成图像与真实图像之间的均方误差 (MSE);

## 5. 反向传播训练：使用梯度下降算法优化网络参数。

### 2.2.3. NeRF 的优缺点

优点	说明
高质量合成	能够实现逼真的新视角图像生成，细节丰富
隐式建模	不需要显式构建网格，表示灵活
可扩展性	可以结合其他模块如语义、深度、时序等进行拓展

缺点	说明
训练慢	原始 NeRF 训练时间长（数小时以上）
渲染效率低	推理时需大量采样，难以实时渲染
无显式几何	不方便与传统几何系统对接，如物理引擎或网格操作

## 3 图像数据预处理与稀疏点云重建

在基于三维高斯泼溅（3DGS）的三维重建系统中，输入图像数据的预处理与点云的初步构建是整个系统流程中的关键前置步骤。本章将详细介绍如何利用 COLMAP 实现相机位姿估计与稀疏点云生成，并进一步探索结合点云神经网络 PointNet 对初始点云质量进行结构性优化的方法。

### 3.1 基于COLMAP的相机位姿估计

#### 3.1.1 相机模型与投影几何

COLMAP 使用透视相机模型对真实世界进行建模。一个三维点  $X = (X, Y, Z)^T \in \mathbb{R}^3$  通过相机内参矩阵  $K$  与外参矩阵  $[R|t]$  被投影到图像平面上：

$$x = \pi(K[R|t]X) = \frac{1}{Z_c} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [R|t]X$$

其中  $(f_x, f_y)$  表示焦距， $(c_x, c_y)$  是主点位置， $R$  为旋转矩阵， $t$  为平移向量， $Z_c$  是点在相机坐标系下的深度。

### 3.1.2 特征提取与匹配

系统首先对所有输入图像进行特征点提取，默认采用 SIFT（尺度不变特征变换）算法提取局部纹理关键点。这些特征点具有尺度、旋转和光照不变性，适合多视角图像的匹配。

特征匹配分为两个阶段：

**粗匹配：**基于特征向量间的欧氏距离使用最近邻匹配，结合比值测试（如 Lowe's ratio test）剔除歧义匹配；

**精匹配：**通过几何验证，如 RANSAC 估计基础矩阵  $F$ ，筛选出满足极线约束的匹配点对。

### 3.1.3 增量式 SfM 重建流程

COLMAP 使用增量式结构恢复方法，从图像对开始，逐步添加新视角，更新已有三维点坐标与新相机的外参。重建过程中包含以下关键步骤：

**三角测量 (Triangulation)：**由两个视角的匹配点反推出三维点位置；

**位姿估计 (PnP)：**根据已知三维点和二维图像点求解新相机的位姿；

**全局优化 (Bundle Adjustment)：**最小化整个系统的重投影误差：

$$\min_{\{R_j, t_j, X_i\}} \sum_{i,j} \| \pi(K_j(R_j X_i + t_j)) - x_{ij} \|^2$$

该优化通过稀疏 Levenberg - Marquardt 算法高效实现。

### 3.1.4 输出格式与可视化

COLMAP 最终输出包括：每张图像的相机内外参数  $(K, R, t)$ ；三维稀疏点集  $X_i$ ；每个点的颜色（通过图像平均采样）与可视视角索引；匹配图像对的信息。

此外，还可输出可视化结果，如图像之间的匹配连线图、稀疏点云视图等，辅助后续模型质量评估。

## 3.2 稀疏点云生成

### 3.2.1 三角测量与深度恢复

给定图像间的相对位姿与像素点匹配，可以使用线性或非线性三角测量方法求出匹配点的三维位置。以最小化重投影误差为目标，最常用的是线性三角测量：

$$AX = 0 \Rightarrow X = \arg \min_{\|X\|=1} \|AX\|$$

其中  $A$  由两个或多个相机投影矩阵与像素坐标构造。得到的点云  $X_i$  存储在 COLMAP 数据库中，初步构建出场景的三维几何结构。

### 3.2.2 稀疏点云的特性与局限性

COLMAP 输出的点云具有以下特点：

点的数量级通常为  $10^3 - 10^5$ ，取决于图像数量和视角重叠程度；

点主要集中在图像重合区域，非纹理区域（如天空、纯色墙面）较少；

精度较高，但分布不均、存在遮挡与缺失区域；

部分点由于匹配误差或误差传播可能存在偏离或漂浮现象。

这类稀疏点云适合作为结构引导，但不适用于直接渲染，因此需要进一步的优化与补全。

### 3.3 基于点云神经网络的系数点云优化

#### 3.3.1 点云处理挑战

在传统三维重建系统中，点云的噪声、高维特征提取、局部语义理解等处理问题往往难以通过几何方法解决。特别是在稀疏点云中，由于点密度低、分布不均匀，传统滤波或插值方法难以获得良好的效果。

点云神经网络（Point-based Deep Network）提供了更强大的结构感知能力。PointNet 是最早且广泛使用的点云处理架构之一，能够从原始点集直接学习特征，无需构建网格或体素。

#### 3.3.2 PointNet架构原理

PointNet 的核心思想在于对无序点集的“对称处理”：每个点  $\mathbf{x}_i \in \mathbb{R}^3$  被单独送入一个共享的多层感知器（MLP），提取局部特征  $h(\mathbf{x}_i)$ ；所有特征通过一个对称函数（如最大池化）汇聚成全局特征：

$$f(X) = \gamma(\text{MAX}h(\mathbf{x}_1), h(\mathbf{x}_2), \dots, h(\mathbf{x}_n))$$

全局特征可用于分类、回归、分割等任务，也可反馈优化每个点的表示。该结构具有以下优点：兼容任意顺序的点集输入；能够处理不规则、稀疏、非均匀采样的点云；学习能力强，适应性广泛。

#### 3.3.3 在稀疏重建中的应用

在 3DGS 系统中，我们可将 PointNet 引入稀疏点云后处理阶段，用于完成以下任务：

##### （1）点云去噪与剔除

训练一个分类器网络识别异常点，包括：配准失败的漂浮点；噪声图像特征产生的伪点；少视角支持的低可信度点。通过引入已知场景的先验数据，可将这些点识别为异常并剔除，提高点云的一致性与可控性。

##### （2）点云补全与插值

使用 Point Completion 网络（如 Point Completion Network、PCN）从已有点云中学习完整结构，插入缺失区域的点，有效提升稠密度。

##### （3）语义聚类与分类

PointNet 可对点云进行空间聚类，为高斯模型赋予局部结构标签，如“边缘点”、“曲面点”、“角点”等。这一信息可在高斯参数初始化时用作先验。

##### （4）多尺度结构理解

通过扩展结构（如 PointNet++），可以在多尺度邻域中聚合特征，对不同尺度下的结构保持敏感，适应建筑、自然场景等多种不同类型数据。

#### 3.3.4 训练与部署流程

部署 PointNet 模块需进行如下步骤：

数据准备：从 COLMAP 导出点云，标注或合成带标签的训练样本；

训练阶段：使用 PyTorch 或 TensorFlow 实现 PointNet 模型，对任务（如二分类、补全）进行监督训练；  
推理与后处理：将处理结果应用于原始稀疏点云，实现剔除、插值或聚类，生成优化后的点云结构。

## 4 基于高斯泼溅的三维重建方法

## 4.1 高斯泼溅表示

在传统的三维重建方法中，如三角网格（Mesh）、体素（Voxel）或点云（Point Cloud）表示方法，通常存在着建模灵活性有限、表示精度不足或者数据冗余等问题。而高斯泼溅（Gaussian Splatting，简称3DGS）提供了一种介于稀疏点云与连续体之间的中间表示形式，它通过各向异性的3D高斯函数来近似建模三维空间中表面的密度与颜色分布，能够以连续、可微且紧凑的方式表达复杂场景的细节和结构。

在3DGS中，一个三维物体由多个“高斯点”组成，每个点表示为一个椭球形的三维高斯核。这些高斯核不仅记录了位置和颜色，还携带有方向、尺度、透明度等参数，用于在渲染过程中通过图像空间投影完成融合[30]。

### 4.1.1 高斯分布的数学模型

我们用  $G_i$  表示第  $i$  个高斯点，其核心参数包括：

位置：  $\mathbf{x}_i \in \mathbb{R}^3$

尺度（协方差矩阵）：  $\Sigma_i \in \mathbb{R}^{3 \times 3}$ ，用于控制高斯在各方向上的扩展；

颜色：  $\mathbf{c}_i \in \mathbb{R}^3$ （RGB 格式）；

不透明度（density/opacity）：  $\alpha_i \in [0, 1]$ ；

朝向（由视锥约束的缩放系数）：用于描述点的可见性与方向性。

在空间中，高斯点  $G_i$  对任意采样位置  $\mathbf{x}$  的贡献可由如下公式表示：

$$G_i(\mathbf{x}) = \alpha_i \cdot \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_i)^T \Sigma_i^{-1}(\mathbf{x}-\mathbf{x}_i)\right) \cdot \mathbf{c}_i$$

这个函数在空间中形成一个带有颜色的“模糊球”，其影响范围由协方差矩阵  $\Sigma_i$  控制。通过调节协方差可以让高斯核表现出各向异性的形状，从而更灵活地贴合场景中的不同几何细节。

### 4.1.2 向量化与效率优化

由于重建场景常常包含数万个高斯点，直接计算其三维卷积代价极高，因此3DGS方法往往采用图像空间的二维渲染策略。具体做法是：

先将所有3D高斯投影到相机视平面；在屏幕空间按像素聚合其影响；最终通过透明度混合、球谐光照等方法合成颜色。这种做法兼顾了三维信息的准确性与二维渲染的高效性，并天然支持GPU并行化计算。

## 4.2 可微渲染过程

高斯泼溅重建方法的重要特性之一是：其整个渲染过程是端到端可微的。这意味着渲染结果可以作为神经网络的输出参与损失函数计算，从而通过反向传播优化每个高斯点的参数（如位置、颜色、尺度等）。这种可微特性为无监督学习和图像到几何的端到端建模提供了基础。

### 4.2.1 渲染原理

在渲染过程中，我们首先将每个高斯点  $G_i$  投影到图像平面上，并根据其空间位置、尺寸和视角确定其在屏幕上形成的覆盖区域（即2D椭圆）。每个像素的颜色最终由多个高斯点在该像素位置的贡献加权融合而成。

具体渲染过程可表述如下：

$$I(\mathbf{u}) = \sum_{i=1}^N \alpha_i(\mathbf{u}) \cdot \mathbf{c}_i \cdot T_i(\mathbf{u})$$

$$\sum_{i=1}^N$$

其中：

$\mathbf{u} \in \mathbb{R}^2$ ：像素坐标；

$N$ ：所有参与该像素渲染的高斯点数目；

$\alpha_i(\mathbf{u})$ ：第 $i$ 个高斯点在像素  $\mathbf{u}$  的覆盖透明度（依据其投影与协方差矩阵计算）；

$\mathbf{c}_i$ ：该点的颜色；

$T_i(\mathbf{u})$ ：该像素未被前方高斯点遮挡的概率（光线穿透率），按如下方式近似建模：

此公式表示从前往后叠加的透明度，体现了前方点的遮挡作用，类似体积渲染中的光线累积透明度计算。

#### 4.2.2 可微渲染计算与梯度传播

由于上述渲染公式中每一部分（高斯投影、遮挡概率、透明度混合）均为连续可导函数，故可以对损失函数

$\mathcal{L} = ||I(\mathbf{u}) - I_{gt}(\mathbf{u})||^2$  求偏导，得到高斯点参数的梯度：

$$\frac{\partial \mathcal{L}}{\partial \mathbf{x}_i}, \quad \frac{\partial \mathcal{L}}{\partial \Sigma_i}, \quad \frac{\partial \mathcal{L}}{\partial \mathbf{c}_i}, \quad \frac{\partial \mathcal{L}}{\partial \alpha_i}$$

这些梯度可以通过反向传播机制送回神经网络或优化器，用于更新模型参数。高斯点的位置调整将推动重建表面更贴近真实几何，而颜色和透明度的优化则提升了渲染图像的视觉一致性。

#### 4.2.3 引入球谐函数进行光照建模

在高斯泼溅方法中，为了进一步增强渲染过程中对真实光照变化的模拟，尤其是方向性反射的建模能力，研究者引入了球谐函数（Spherical Harmonics, SH）来表达每个高斯点的颜色响应与观察方向之间的关系。该方法能够在保持较低计算开销的同时有效地捕捉复杂的光照变化，成为高斯泼溅渲染中的重要组成部分。

##### 一、球谐函数概述

球谐函数是定义在单位球面上的一组正交函数系，可看作是傅里叶级数在球面坐标系上的推广。它们广泛用于计算机图形学、物理学、量子力学中的球面对称问题中。

球谐函数的标准形式为：

$$Y_l^m(\theta, \phi) = N_l^m \cdot P_l^m(\cos \theta) \cdot e^{im\phi}$$

其中：

$l \in \mathbb{N}_0$ ：阶数，表示频率层次；

$m \in [-l, l]$ ：次数；

$\theta \in [0, \pi]$ 、 $\phi \in [0, 2\pi]$ ：分别为球坐标系中的极角（zenith）与方位角（azimuth）；

$P_l^m$ ：关联勒让德多项式；

$N_l^m$ ：归一化系数，确保函数正交性。

球谐函数具有如下重要性质：

完备性：任意在球面上的连续函数  $f(\theta, \phi)$  都可以展开为球谐函数的加权和；



方向建模能力：能够表示随方向变化的函数（如反射、光照）；

低频有效性：在渲染中常用  $l=0 \sim 2$  的低阶球谐来逼近光照与反射函数，计算高效且表达力强。

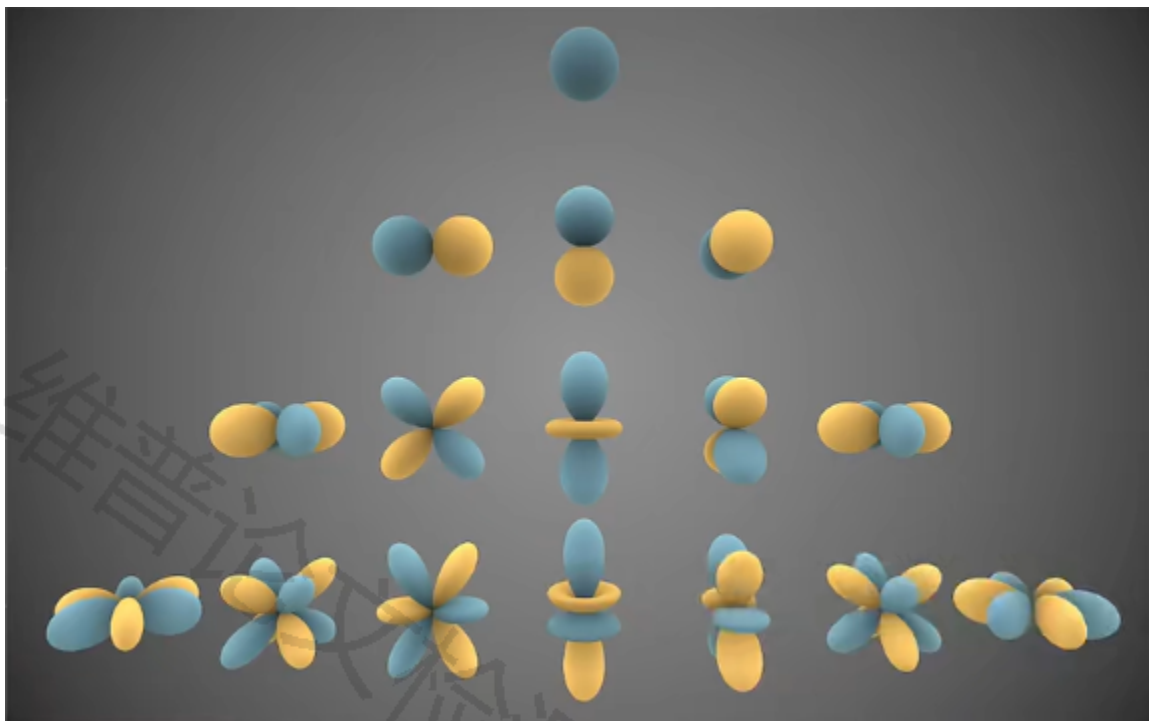


图4.1 球谐函数示意图

## 二、球谐函数在3DGS中的应用

在高斯泼溅的可微渲染中，每个高斯点不仅携带一个静态颜色  $\mathbf{c}_i \in \mathbb{R}^3$ ，而是拥有一组用于描述方向性反射的球谐系数  $\mathbf{c}_i^{SH} = \{c_{lm}^i\}$ 。渲染时，针对当前像素的视角方向  $\mathbf{v}$ ，我们用其在单位球面上的方向向量  $(\theta, \phi)$  作为自变量，计算该点在该方向下的颜色响应：

具体做法是引入球谐函数（Spherical Harmonics, SH）作为点的颜色表示的一部分：

$$\mathbf{c}_i(\mathbf{n}) = \sum_{l=0}^L \sum_{m=-l}^l \mathbf{c}_{lm}^i Y_{lm}(\mathbf{n})$$

其中：

$L$  是球谐展开的最大阶数（一般为2或3）；

$\mathbf{c}_{lm}^i \in \mathbb{R}^3$ ：每个频率分量上的RGB系数；

$Y_{lm}(\theta, \phi)$ ：球谐基函数，输入为当前相机与点之间的方向角。

通过这种方法，每个点的颜色可以动态地随着光照或视角发生变化，实现更真实的BRDF（双向反射分布函数）效果。这一机制在3DGS中用于拟合更复杂的光照环境，增强点云渲染的真实性。

## 三、SH建模在渲染中的实际应用

引入球谐光照建模后，3DGS在渲染过程中表现出更真实的视觉特性，特别适用于下列场景：

自然光照环境（IBL）：球谐函数非常适合用来拟合自然图像中的低频全局光照；

细节可控：通过调节球谐展开的阶数 $0020L$ ，可以平衡渲染真实感与计算效率；

与神经网络兼容：SH系数可以作为高斯点的一部分，在网络训练中直接回传梯度进行优化，从而实现基于图像监督的三维重建。

图形学中，使用三阶球谐（ $L = 2$ ）通常就能获得较好的方向感知效果，只需9个SH基函数即可表示每个点的光照响应曲线，相比传统BRDF或微面模型，成本极低。

通过引入球谐函数，高斯泼溅方法不仅能表示更丰富的几何结构，同时具备更强的光照建模能力，显著提升了重建结果在视觉感知上的真实度与一致性，为高质量、实时可视化的三维重建提供了强有力的支持。

4.3 参数初始化与训练策略

高质量的三维重建不仅依赖于强大的模型结构，还依赖于良好的参数初始化和合理的训练策略。在高斯泼溅表示中，每一个高斯点都携带多种参数，包括位置、尺度、方向、透明度、颜色和可能的球谐系数（SH），这些参数的初始值对于收敛速度和最终效果都有重要影响。

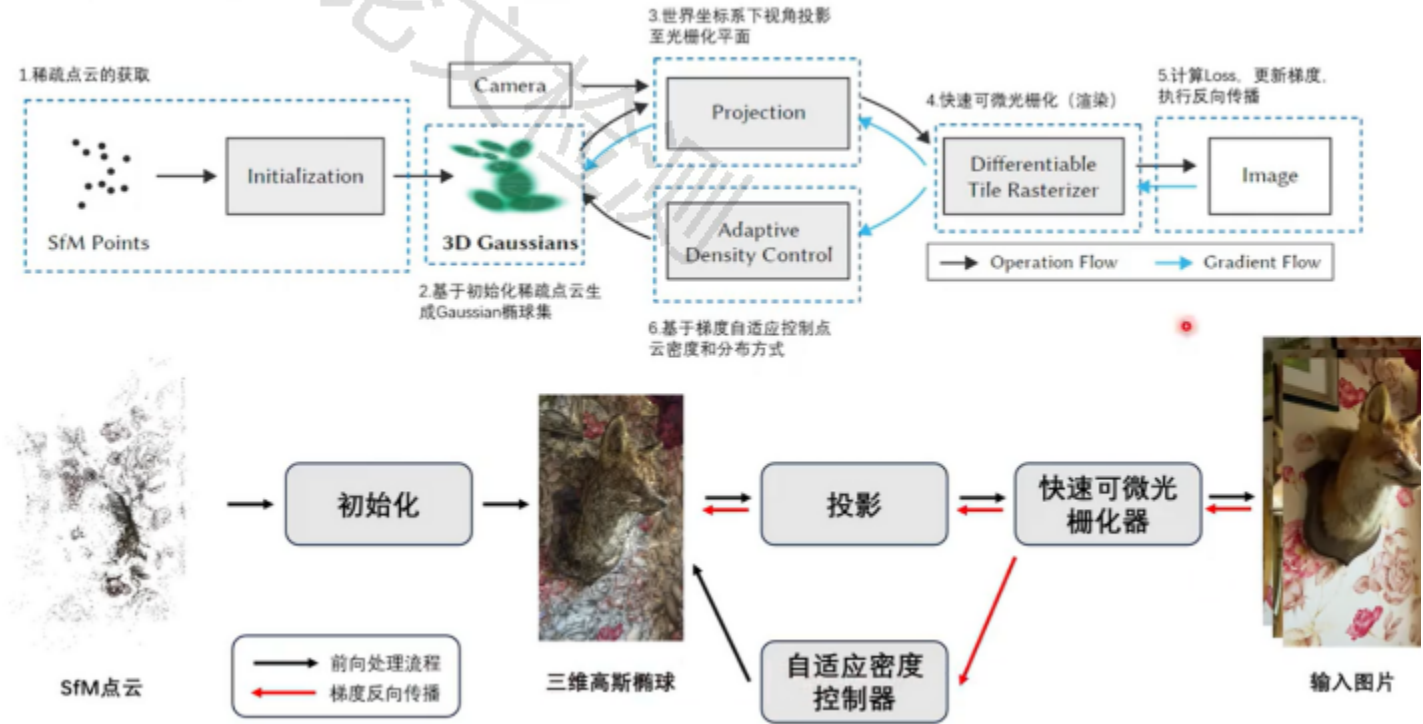


图4.2 高斯泼溅核心算法流程

4.3.1 初始化策略

在训练开始之前，通常需要通过稀疏点云（例如来自 COLMAP 的 SfM 结果）来初始化高斯点的位置。这些初始化点是三维空间中稀疏且较为准确的特征点，具有较强的几何结构信息。

每个点的初始参数如下：

位置  $\underline{x}_i \in \mathbb{R}^3$ ：由 COLMAP 稀疏点云直接提供；

尺度（协方差） $\underline{\Sigma}_i$ ：通常初始化为各向同性小球，即  $\underline{\Sigma}_i = \sigma_0^2 \mathbf{I}$ ，其中  $\sigma_0$  是一个小的常数；

颜色  $\mathbf{c}_i \in \mathbb{R}^3$ ：可通过多个视角图像投影并平均采样得到；

方向（用于椭球分布）：初始设为单位向量或者从 PCA 中估计；

透明度或密度参数  $\sigma_i$ ：初始化为常数或与深度成反比；

球谐系数：常设为均值颜色乘以一个经验常数或零。这种初始化策略可以保证模型在一开始就拥有基本的几何和颜色一致性，从而有效减少训练初期的不稳定性。

#### 4.3.2 训练损失与优化目标

高斯点的训练目标是最小化渲染图像与真实图像之间的差异。常见的损失函数包括：

像素级重建损失（MSE）：

$$L_{\text{pixel}} = \frac{1}{HW} \sum_{u,v} (I_{\text{render}}(u,v) - I_{\text{gt}}(u,v))^2$$

感知损失（Perceptual / LPIPS）：利用预训练网络提取图像特征，衡量图像相似性；

深度一致性损失：可选，在有深度图 supervision 时使用；

正则项：限制点的尺度、密度、方向，避免训练时的参数发散：

$$L_{\text{reg}} = \lambda_1 \sum_i |\Sigma_i|^2 + \lambda_2 \sum_i |c_i|^2$$

#### 4.3.3 多阶段训练策略

为提升训练稳定性与最终质量，3DGS 采用逐步细化的训练流程：

初始阶段：仅训练位置和颜色；

中期阶段：引入方向和尺度参数优化；

后期阶段：加入球谐系数、密度参数的联合优化；

分辨率提升：逐步引入高分辨率图像和更复杂视角；

每一步训练都可以使用不同的学习率和调度策略。常用优化器为 Adam，搭配 Cosine Annealing 或 Exponential Decay 学习率下降曲线。

#### 4.4 优化与高斯压缩策略

在训练过程中，高斯点的数量和精度会不断提升。若无优化和压缩机制，模型将面临渲染速度慢、显存占用高、训练不稳定等问题。因此，动态点数管理成为 3DGS 成功的关键之一。

##### 4.4.1 点密度增长策略（Density-Guided Splitting）

随着训练深入，需要在图像细节丰富或重建误差大的区域增加更多点。通常的点增长方式包括：

图像误差驱动：在重建误差大的区域插入新点；

梯度驱动法：使用点的梯度大小判断其重要性；

光线采样密度控制：增加点来覆盖疏采样区域；

椭球拉伸复制：将一个尺度较大的点拉伸方向复制多个点（split），细化边缘结构。

每次新增点后，其初始化可通过邻域插值、权重传播或使用图像特征嵌入网络生成。

#### 4.4.2 点合并与压缩策略 (Gaussian Pruning)

为了压缩模型体积，提高渲染效率，需要周期性压缩或合并冗余点：

空间聚类合并 (Spatial Clustering)：将空间上邻近、颜色和尺度接近的点合并为一个点：

$$\mathbf{x}_{\text{new}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i, \quad c_{\text{new}} = \frac{1}{n} \sum_{i=1}^n c_i$$

不可见点删除 (Visibility Culling)：统计每个点在训练图像中可见帧数，若低于阈值则剔除；

权重贡献剔除：根据每个点在渲染图像中所占权重，对低权重点执行惩罚或删除；

SH 活跃性检测：若球谐系数长期未更新，可认定该点贡献小而移除。

#### 4.4.3 性能优化与加速策略

为了平衡质量与效率，可进一步引入以下技术：

分块渲染 (Tile-based Splats Rendering)：将图像划分为 tile，按需加载高斯点；

视锥裁剪 (Frustum Culling)：仅加载当前视角下可见高斯点；

KD-tree / Octree 加速结构：快速查询局部点邻域，提高渲染速度；

混合显存与磁盘缓存：将不活跃点转存磁盘，释放 GPU 显存；

#### 4.4.4 压缩率与质量的权衡

在部署阶段，模型大小和渲染速度成为关键性能指标：

高压缩率意味着少量高质量的点；

高质量需求则要求保留更多局部细节点；

通常使用 PSNR、LPIPS 与FPS等多指标综合评估最终效果。

### 5 三维重建系统设计与实现

本系统旨在实现一个基于 Web 的三维高斯泼溅 (3D Gaussian Splatting, 3DGS) 可视化与编辑平台，支持从 COLMAP 导出的稀疏点云与位姿信息出发，通过加载训练好的 3DGS 模型，在浏览器中实现实时渲染与交互，并为研究者和用户提供可视化调试、参数调节与数据分析的工具。

#### 5.1 系统需求与分析

##### 5.1.1 功能需求

为了满足三维重建的教学、科研和应用展示等需求，系统应具备以下功能模块：

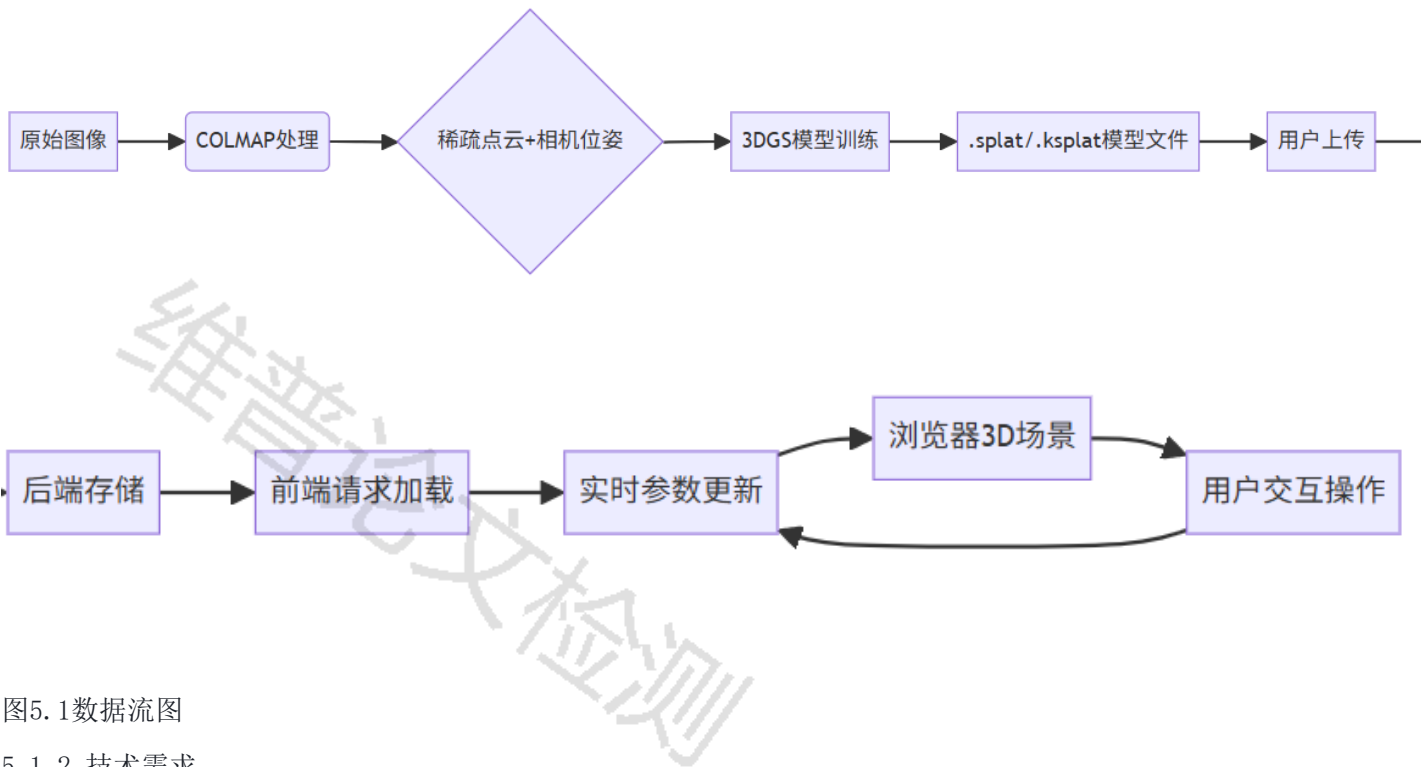
1. 相机姿态与稀疏点云获取：借助 COLMAP 工具实现稀疏点云重建，并解析获取相机位姿，用于初始化场景相机；

2. 3DGS 模型加载与渲染：支持用户本地上传训练好的 .splat 或 .ksplat 格式高斯点数据，系统能解析并还原三维场景；

3. 三维可视化与漫游：提供 Web 端交互式三维场景浏览，用户可通过鼠标旋转、缩放、拖动等方式进行漫游；

4. 高斯泼溅参数调节：支持对高斯点的尺寸、颜色、方向、透明度等属性进行可视化调整；

- 5. 光照控制与 SH 拟合展示：允许用户调节环境光方向、颜色，展示球谐光照模拟效果；
- 6. 界面友好，体验流畅：系统整体界面应简洁直观，响应迅速，具有良好的用户交互体验；



5.1.2 技术需求

前端采用 React + Three.js 实现实时 Web 端 3D 渲染；  
后端使用 Node.js 和 Express 框架搭建，负责文件管理、任务调度；  
模型训练基于原作者的算法模型，对点云数据进行处理，并生成3dgs格式的模型

5.2 后端模块设计

系统后端主要基于 Node.js 平台实现，并使用SQLite保存用户注册的数据。

5.3 前端渲染模块设计

前端采用 React 构建组件化页面结构，结合 Three.js 实现 3D 可视化功能，核心模块包括：

5.3.1 场景构建与数据加载

基于 THREE.Scene、THREE.Camera 等类构建基础渲染场景；支持 GLTF 模型、ksplat格式点云的解析与转换。

5.3.2 高斯泼溅可视化实现

模拟高斯分布形状（如球体、椭球体）并动态调整其半径、透明度与颜色；使用基于 ShaderMaterial 的 GLSL 编写渲染材质，支持对 SH 光照效果的简单模拟；提供参数调节面板，允许用户实时调节高斯点属性（如尺度、方向、颜色等）。能够编辑加载和处理的GLTF模型和动画。

5.3.3 用户交互与界面设计

提供鼠标操作支持，包括旋转、缩放、拖拽等；引入调试面板提升交互体验。





图5.2系统用户登录界面

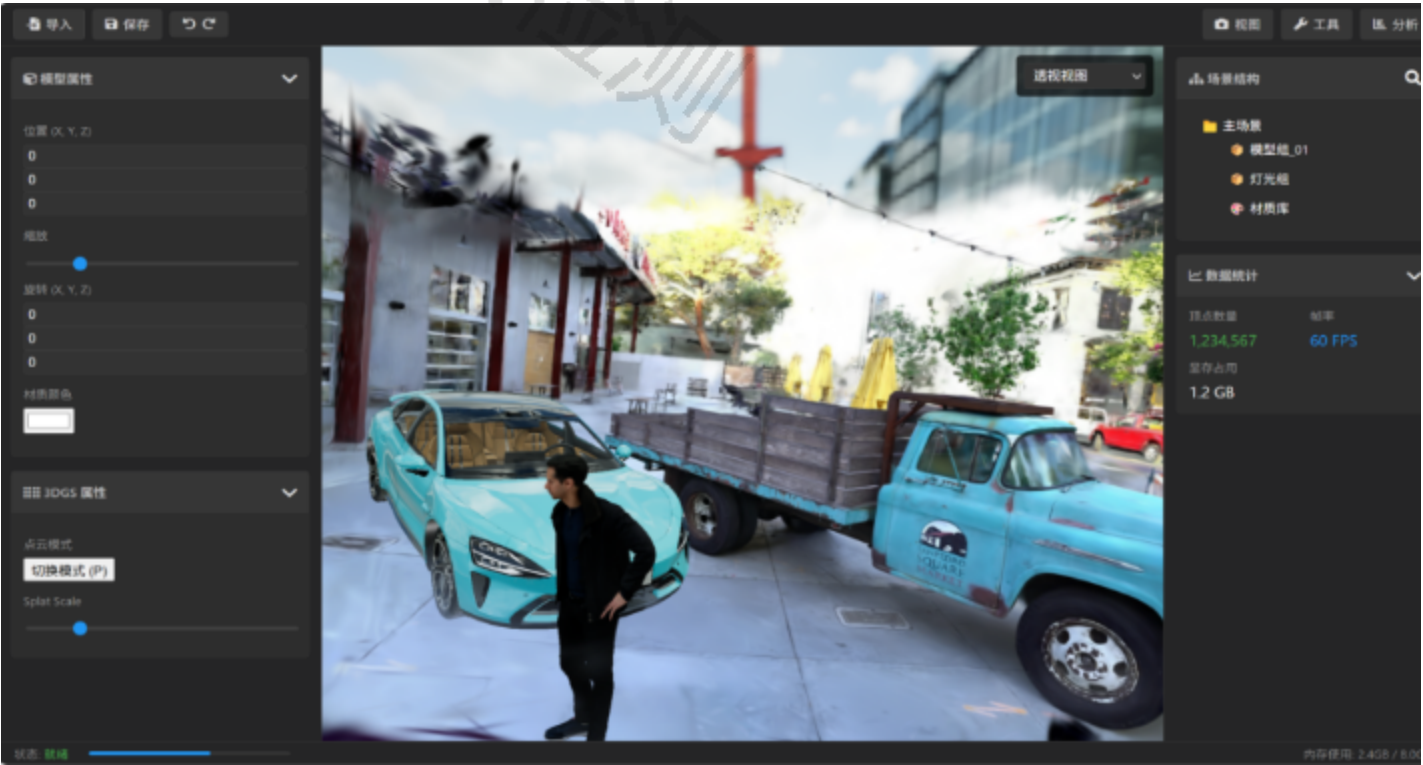


图5.3系统渲染操作界面

## 6 总结与展望

### 6.1 全文总结

本文围绕基于高斯泼溅的三维重建系统展开研究，重点探讨了从图像输入到高质量点云渲染输出的整个技术流程。全文的研究工作可分为以下几个方面：



首先，在图像预处理与稀疏点云重建方面，本文基于 COLMAP 工具对多视角图像进行特征匹配和结构光束法束束优化，得到稀疏点云和相机位姿，为后续三维建模提供了可靠的初始结构。为了提升点云质量，引入了 PointNet 网络对稀疏点云进行感知优化，在保证全局结构的同时提高了局部几何表达的准确性。

其次，在高斯泼溅的三维表示方法方面，本文系统介绍了 3D Gaussian Splatting (3DGS) 的表示原理、渲染机制与参数建模。与传统的体素和网格表示不同，3DGS 通过大量具有空间分布和方向性的高斯点实现了高效且可微的三维建模与渲染过程，为无网格重建技术开辟了新方向。

在渲染部分，本文重点分析了可微渲染的数学原理，包括基于密度和透明度的融合机制、投影机制、遮挡建模，以及梯度传播方式。此外，还引入球谐函数建模点的方向性反射特征，进一步提升了渲染光照的真实性和一致性，为高斯泼溅模型在复杂光照环境下的表现提供了理论基础。

接着，本文详细讨论了参数初始化与训练策略，通过多阶段、多目标的优化方法，使得高斯点能够从稀疏粗糙的初始状态逐步优化为稠密、细节丰富的三维表示。为了进一步优化模型性能和存储开销，本文提出了一套高斯点的压缩与裁剪机制，动态调整高斯点数量，保证模型在推理时具备良好的效率与可用性。

最后，本文的研究成果不仅可直接应用于无网格三维建模、全景重建、AR/VR 等前沿领域，也为点云渲染、可微图形学等方向的研究提供了可推广的思路与方法。

综上所述，本文从稀疏点云的生成与优化出发，系统构建了一套以 3DGS 为核心的高质量三维重建流程，验证了其在精度、速度、扩展性等方面的优势，展示了基于高斯泼溅的三维建模技术的巨大潜力。

## 6.2 展望

尽管本文在 3D 高斯泼溅建模领域取得了阶段性成果，但当前系统仍有诸多可扩展和优化的方向，未来的研究可以从以下几个方面深入探索：

### (1) 稠密点云与几何细节建模

目前系统主要依赖稀疏点云和单一图像信息进行重建，尽管通过可微优化与高斯点增强提高了质量，但在一些纹理重复或无纹理区域仍难以还原细节。未来可以融合多源信息（如深度图、法线图、激光点云）以获得更高质量的几何形态。

### (2) 实时渲染与轻量化模型压缩

高斯泼溅的实时渲染能力虽然较强，但在高点数场景中仍可能面临带宽与显存瓶颈。后续可进一步探索基于硬件特性的图形加速（如 WebGPU、CUDA 光线追踪）与压缩技术（如点的分层 LOD、稀疏索引结构）来提升系统响应效率，拓展移动端和 Web 应用的适配能力。

### (3) 动态场景建模与时序一致性优化

当前系统主要面向静态场景。如何将高斯泼溅拓展至可动对象与动态场景是一个值得探索的方向。可以考虑引入时间维度上的高斯点轨迹建模、基于物体跟踪的动态姿态预测机制，结合连续帧之间的可微约束，实现对动态环境的连续重建。

### (4) 融合深度学习的高斯生成器

当前点的初始化与优化仍然依赖于显式手工步骤，未来可探索使用生成式模型（如 Diffusion 模型、VAE 或 NeRF 结构）预测高斯点的位置与属性，甚至通过文本或语义描述直接生成三维高斯模型，实现语义驱动的三维重

建。

#### (5) 交互式编辑与应用拓展

3DGS 不仅适合用于重建与可视化, also具有良好的可编辑性与物理可解释性。未来可以将其集成到三维编辑工具中, 实现高斯点级别的交互修改、分割、重定向等操作, 推动其在游戏制作、虚拟现实、数字文物保护等领域的实际落地。

#### 致 谢

时光荏苒, 毕业论文的写作即将画上句号, 在此我怀着无比感激的心情向在我学习和研究过程中给予指导与帮助的老师、同学和亲友致以诚挚的谢意。

首先, 我要衷心感谢我的导师杨庆老师, 在本论文的选题、研究、撰写等各个阶段, 给予了我细致入微的指导和无私的帮助。导师严谨的治学态度、深厚的专业素养和踏实认真的工作作风, 对我影响深远, 使我在毕设工作上受益良多。

其次, 感谢所在课题组的各位同学与前辈们, 大家在项目实施和学术交流过程中给予了我极大的启发和鼓励。特别是在实验搭建、代码实现和论文讨论中, 大家的建议与协作让我少走了许多弯路, 也让我体会到了团队合作的力量。

感谢信息科学与工程学院大数据与人工智能系梁宗保教授、刘新龙老师、王笛老师, 杨翔立老师等在学业方面给予的指导和帮助; 感谢辅导员张秋玉老师、金其奇老师在本科学习生活中给予的照顾和关怀。

感谢我的同学李财华、甘孝平等, 前专业的袁倩, 吴智博, 四年本科时光同你们度过留下许多美好的回忆, 大家在相互帮助、鼓励和学习中一起度过了这段最美好的时光, 你们是我求学探索道路上的最重要的力量, 祝你们未来一切顺利。

同时, 也要感谢我在实习期间接触到的同校学长朱悦凯和导师谢传金, 在我人生中第一次参加工作的时候给予了非常多的帮助, 无论是生活上, 技术上, 还是个人发展上, 对我影响深远。

此外, 我还要感谢我的家人, 是他们一如既往的支持与鼓励, 给予我坚实的后盾, 使我在遇到困难与挑战时依旧充满勇气和动力。

最后, 感谢母校为我提供了优良的学习与研究环境, 使我有机会在图形学与人工智能交叉领域深入探索, 为将来迈向更高的学术台阶打下了坚实基础。

谨以此文, 献给所有关心和支持过我的人。

#### 参 考 文 献

[1] 高建, 陈林卓, 沈秋, 等. 基于三维高斯溅射技术的可微分渲染研究进展[J/OL]. 激光与光电子学进展, 1-23[2025-04-13].

[2] 朱东林, 陈淼, 毛宇岩, 等. 三维高斯泼溅技术在场景重建中的现状与挑战[J/OL]. 集成技术, 1-17[2025-04-13].

[3] 韩开, 徐娟. 3D 场景渲染技术—神经辐射场的研究综述[J]. 计算机应用研究, 2024, 41(8)..

[4] 侯耀斐, 黄海松, 范青松, 等. 基于改进多层感知机的神经辐射场三维重建方法[J]. 激光与光电子学进展, 2024, 61(04):0415004. 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.

- [5]王锋, 银莹, 王佳炎, 等. 基于高斯泼溅的轻量级重建场景分割方法[J/OL]. 计算机学报, 1-12[2025-04-22].
- [6]何天琪, 宋佳洁, 程景春, 等. 基于三维高斯泼溅技术的航拍场景建模[J]. 火箭军工程大学学报, 2025, 39(02): 1-12+21.
- [7]宋明清, 郭尧, 李晓峰, 等. 基于光照探针的三维高斯辐射场压缩算法[J/OL]. 计算机应用研究, 1-8[2025-04-22].
- [8]张月婷, 李文劼, 郭嘉逸, 等. 基于高斯基的SAR图像重建技术研究[J]. 激光与红外, 2024, 54(12): 1936-1940.
- [9]Kerbl B, Kopanas G, Leimkühler T, et al. 3d gaussian splatting for real-time radiance field rendering[J]. ACM Transactions on Graphics, 2023, 42(4): 1-14.
- [10]Xin YY, Zuo XX, Lu DY, et al. SimpleMapping: Real-Time Visual-Inertial Dense Mapping with Deep Multi-View Stereo[C] // 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2023: 273-282.
- [11]Guedon A and Lepetit V. SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering[C] // 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 5354-5363.
- [12]Liang Z, Zhang Q, Feng Y, et al. GS-IR: 3D Gaussian Splatting for Inverse Rendering," 2024 IEEE /CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C] // 2024: 21644-21653.
- [13]Fei B, Xu J, Zhang R, et al. 3D Gaussian Splatting as New Era: A Survey[J]. IEEE Transactions on Visualization and Computer Graphics , Early Access, 2024.
- [14]Zhu Z, Peng SY, Larsson V, et al. NICE-SLAM: Neural Implicit Scalable Encoding for SLAM[C] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 12776-12786.
- [15]Zhu DL, Wang ZL, Lu TF, et al. PMF-SLAM: Pose-Guided and Multiscale Feature Interaction-Based Semantic SLAM for Autonomous Wheel Loader[J]. IEEE Sensors Journal, 2024, 24(7): 11625-11638.
- [16]Fridovich-Keil, S.; Meanti, G.; Warburg, F. R.; Recht, B.; Kanazawa, A. K-planes: Explicit radiance fields in space, time, and appearance. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 12479 - 12488, 2023.
- [17]Jiang, Y.; Yu, C.; Xie, T.; Li, X.; Feng, Y.; Wang, H.; Li, M.; Lau, H.; Gao, F.; Yang, Y.; et al. VRGS: A physical dynamics-aware interactive Gaussian splatting system in virtual reality. arXiv preprint arXiv:2401.16663, 2024.
- [18]MILDENHALL B, SRINIVASAN P P, ORTIZ-CAYON R, et al. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines [J]. ACM Transactions on Graphics, 2019, 38(4): 1-14.
- [19]TANG J, REN J, ZHOU H, et al. Dreamgaussian: Generative gaussian splatting for efficient 3d content creation [EB/OL]. (2023-09-28)[2024-05-27].
- [20]Rusinkiewicz, S.; Levoy, M. QSplat: A multiresolution point rendering system for large meshes. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, 343 - 352, 2000.

- [21]Botsch, M.; Kobbelt, L. High-quality point-based rendering on modern GPUs. In: Proceedings of the 11th Pacific Conference on Computer Graphics and Applications, 335–343, 2003.
- [22]Ge, W.; Hu, T.; Zhao, H.; Liu, S.; Chen, Y. C. Ref-NeuS: Ambiguity-reduced neural implicit surface learning for multi-view reconstruction with reflection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 4228–4237, 2023.
- [23]Bonet, J.; Wood, R. D. Nonlinear Continuum Mechanics for Finite Element Analysis. Cambridge, UK: Cambridge University Press, 2008.
- [24]Brooks, T.; Holynski, A.; Efros, A. A. InstructPix2Pix: Learning to follow image editing instructions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 18392–18402, 2023.
- [25]Seiskari, O.; Ylilammi, J.; Kaatrasalo, V.; Rantalankila, P.; Turkulainen, M.; Kannala, J.; Rahtu, E.; Solin, A. Gaussian splatting on the move: Blur and rolling shutter compensation for natural camera motion. arXiv preprint arXiv:2403.13327, 2024.
- [26]Xiong, B.; Li, Z.; Li, Z. GauU-scene: A scene reconstruction benchmark on large scale 3D reconstruction dataset using Gaussian splatting. arXiv preprint arXiv:2401.14032, 2024.
- [27]HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models [J]. Advances in neural information processing systems, 2020, 33:6840–51.
- [28]Gao, L.; Yang, J.; Wu, T.; Yuan, Y.; Fu, H.; Lai, Y.; Zhang, H. SDM-NET: Deep generative network for structured deformable mesh. ACM Transactions on Graphics Vol. 38, No. 6, Article No. 243, 2019.
- [29]FANG J, WANG J, ZHANG X, et al. Gaussianeditor: Editing 3d gaussians delicately with text instructions [EB/OL]. (2023-11-27)[2024-05-27].
- [30]ZHENXING M I, XU D. Switch-nerf: Learning scene decomposition with mixture of experts for large-scale neural radiance fields [C]//The Eleventh International Conference on Learning Representations. 2022.
- [31]SONG L, CHEN A, LI Z, et al. Nerfplayer: A streamable dynamic scene representation with decomposed neural radiance fields [J]. IEEE Transactions on Visualization and Computer Graphics, 2022, 29(5): 2732–42.
- [32]GAO C, SARAF A, KOPF J, et al. Dynamic view synthesis from dynamic monocular video [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 5712–21.
- [33]BARRON J T, MILDENHALL B, VERBIN D, et al. Mip-nerf 360: Unbounded anti-aliased neural radiance fields [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 5470–9.

## 须知：

- 报告编号系送检论文检测报告在本系统中的唯一编号
- 本报告为维普论文检测系统算法自动生成，仅对您所选择比对资源范围内检验结果负责，仅供参考。



微信公众号

---

唯一官网：<https://vpcs.fanyu.com> | 客服邮箱：[vpcs@fanyu.com](mailto:vpcs@fanyu.com) | 客服热线：400-607-5550 | 客服QQ：4006075550

维普论文检测