

# 基于神经辐射场的 SAR 图像高效三维重建方法

汪章怀<sup>1,2</sup>, 薛丰利<sup>1</sup>, 刘开雨<sup>1</sup>

(1. 中国科学院空天信息创新研究院 航天微波遥感系统部, 北京 100190;

2. 中国科学院大学 电子电气与通信工程学院, 北京 100049)

**摘要:** 为了解决合成孔径雷达图像三维重构中数据集稀缺、重构精度不足等问题, 针对 SAR 图像处理, 提出了一种基于神经辐射场模型的三维重构方法; 采用双向解析射线追踪方法生成 SAR 仿真数据集, 并利用 ColMap 工具获取相机姿态和稀疏重构数据; 在数据处理方面, 使用图像增强、散斑噪声抑制和旁瓣抑制等技术, 确保输入数据的高质量; 通过优化神经辐射场模型编码方式和网络结构, 引入哈希编码技术, 训练 NeRF 模型实现建筑物三维重构; 实验测试实现了对建筑物主要轮廓和结构特征的较好恢复, 尤其在柱体部分重建精度较高, 光学仿真模型与重构模型尺寸比例误差在 6% 以内; 研究表明该方法为 SAR 建筑物图像三维重构提供了新的思路与技术支持, 尽管地基部分重建精度有待提升, 但整体具有应用潜力。

**关键词:** 合成孔径雷达; 三维重构; 神经辐射场; 图像增强

## Three-dimensional Efficient Reconstruction Method for SAR Images Based on Neural Radiation Field

WANG Zhanghuai<sup>1,2</sup>, XUE Fengli<sup>1</sup>, LIU Kaiyu<sup>1</sup>

(1. Institute of Aerospace Information Innovation, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** In order to solve the problems of scarce data sets and insufficient reconstruction accuracy in the three-dimensional reconstruction of synthetic aperture radar (SAR) images, a three-dimensional reconstruction method based on neural radiation field model is proposed for SAR image processing. A bidirectional analytical ray tracing method is used to generate the SAR simulation data set, and the ColMap tool is used to obtain the attitude and sparse reconstruction data of the camera. In data processing, the image enhancement, speckle noise suppression and sidelobe suppression are used to ensure the high quality of input data. By optimizing the coding mode and network structure of the neural radiation field model, and introducing the Hush coding technology, the neural radiance field (NeRF) model is trained to realize the three-dimensional reconstruction of buildings. Experimental results show that this method effectively restores main outline and structural characteristics of the building, especially in column part, with a high reconstruction accuracy, and the size ratio error between the optical simulation model and the reconstruction model is within 6%. The research shows that this method provides a new idea and technical support for the three-dimensional reconstruction of SAR building images. Although the accuracy of foundation reconstruction needs to be improved, it has application potential as a whole.

**Keywords:** SAR; three-dimensional reconstruction; neural radiance field; image enhancement

## 0 引言

合成孔径雷达 (SAR, synthetic aperture radar) 在地球遥感领域发挥着重要作用, 其全天候、全天时的观测能力使其成为获取地表信息的关键技术之一。然而, 尽

管 SAR 图像在三维重建中的应用潜力巨大, 但该领域仍面临诸多挑战。目前, SAR 图像的三维重建主要依赖传统的几何重建方法, 如层析成像<sup>[1]</sup>。该方法通过多视角 SAR 影像的配准与相位信息处理, 能够重建目标的三维结构<sup>[2]</sup>。然而, 层析成像对目标表面纹理信息的依赖性

收稿日期: 2025-03-25; 修回日期: 2025-04-23。

作者简介: 汪章怀 (2000-), 男, 硕士研究生。

引用格式: 汪章怀, 薛丰利, 刘开雨. 基于神经辐射场的 SAR 图像高效三维重建方法[J]. 计算机测量与控制, 2025, 33(9): 271-282.

较强,在复杂场景中难以捕捉精细结构,且对噪声高度敏感,导致重建精度受限;此外,SAR图像对观测配置的高度敏感性导致在不同条件下获取的图像之间存在显著差异,相较于光学图像,SAR图像因成像机制特殊,表面纹理信息贫乏且噪声干扰严重,导致特征匹配点不足,尤其在多路径效应影响下,SFM常出现匹配错误或发散问题,这进一步增加了三维重建的难度<sup>[3]</sup>。

近年来,基于二维图像的三维场景合成技术已成为研究热点,在多个领域展现出重要的应用价值。许多经典方法,如运动结构恢复(SFM, structure from motion)<sup>[4]</sup>和开放多视图几何(openMVG, open multiple view geometry)<sup>[5]</sup>,已在特定场景下取得了一定进展。三维场景的表示方式主要分为显式表示和隐式表示两类。显式表示方法包括体素网络(3D-R2N2, 3D recurrent reconstruction neural network)<sup>[6]</sup>、点云网络(PointNet, point network)<sup>[7]</sup>和网格网络(Pixel2mesh, pixel to mesh)<sup>[8]</sup>等,通常通过点云、体素等形式进行描述,具有便捷的编辑性、高计算效率和广泛的软件支持等优点。这些方法通常依赖神经网络提取特征,并进一步建模特征与已知深度之间的关系。然而,在当前研究场景下,该方法存在明显的局限性。一方面,由于图像表面纹理信息较为贫乏,难以提取足够的特征匹配点,从而限制了高精度图像关联的实现。另一方面,卫星与建筑物之间的距离往往超出了传感器的有效测量范围,获取准确深度图像面临巨大挑战,这对显式方法的应用造成了较大的限制<sup>[9]</sup>。因此,显式方法在SAR领域的适用性受到质疑。

与显式表示方法不同,三维场景的隐式表示通过一个函数来描述物体的几何形状,仅需提供图像及其对应的相机姿态,便能有效避免显式表达方式中因表面纹理信息不足而引发的问题。然而,该方法在精确获取相机内外参数时存在显著困难,尤其是在将隐式表达式与真实虚拟视图相结合时。在引入神经辐射场(NeRF, neural radiance fields)<sup>[10]</sup>之前,这种结合几乎是不可能的。与符号距离函数<sup>[11]</sup>和占用场<sup>[12]</sup>等传统方法相比,NeRF作为一种基于隐式表示的三维场景建模方法,其核心理念是结合光学相机的光线投射机制,通过神经网络对三维体素的辐射度和密度进行建模。NeRF模型巧妙地将光学3D体素的渲染物理模型与深度学习技术深度结合,能够从多视图图像中实现高精度的三维密集重建,为三维场景合成技术的发展开辟了新的道路。

现有研究尝试将NeRF技术迁移到SAR领域,以SAR建筑物幅度图像重构三维模型,但在实践中面临诸多挑战。首先,由于真实SAR图像获取难度大且涉及隐私保护,现有数据资源极为有限。其次,SAR建筑物图像与光学图像在成像机制上存在显著差异:SAR

建筑物图像的成像结果受到雷达波与目标相互作用的影响,其纹理、亮度和结构与光学图像存在显著差异<sup>[13]</sup>。其次,SAR图像普遍存在较强的散斑噪声和多路径效应,这些噪声和干扰严重影响了三维重建的精度和稳定性<sup>[14]</sup>,并导致NeRF模型在SAR建筑物图像中的应用变得极为困难。此外,由于NeRF的优化基于每个像素的渲染损失,这使得NeRF在SAR建筑物图像上的耗时会尤为严重<sup>[15]</sup>。

为解决上述问题,本文提出了一种创新的SAR建筑物三维重建方法,结合了NeRF模型与SAR建筑物图像的特性,从数据生成、图像处理、相机姿态获取及模型优化等方面进行了多层次改进:首先,在数据生成方面,本文采用双向解析射线追踪(BART, bidirectional analytical ray tracing)<sup>[16-17]</sup>技术生成高质量仿真数据集,模拟真实成像特性,解决数据稀缺问题。图像处理上,采用图像增强、散斑噪声抑制和旁瓣抑制等技术,提升图像质量。相机姿态获取环节,提出时序匹配算法,利用图像时间序列信息,提高匹配精度和效率。针对NeRF模型,对其编码方式和网络结构进行优化,引入哈希编码技术,提升计算效率和泛化能力,降低训练复杂度,减少对输入图像质量的依赖。优化后的模型显著提升了计算效率和泛化能力,有效降低了训练复杂度,同时确保了高质量的三维重建效果。

## 1 研究方法

本研究以SFM与多视图立体重建(MVS, multi-view stereo)三维重构技术为基础,结合NeRF及其变种模型,自主构建了SAR建筑模型的仿真环境,并生成了适配的SAR三维重构数据集。本文研究方法的流程架构如图1所示。

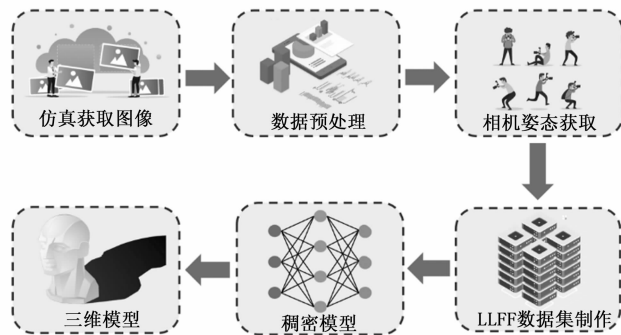


图1 技术路线

### 1.1 仿真算法

在本研究中,所采用的仿真算法为BART算法。BART方法是一种高效计算三维电磁大型复杂目标在随机粗糙表面上的复合散射的先进算法,其关键特点如下:

1) 双向追踪机制: BART方法在入射和散射方向上进行光线追踪,能够精准捕捉目标和表面之间的高阶

散射相互作用, 从而为复杂场景下的电磁散射分析提供了更为全面和精确的解决方案。

2) 多边形光束的分析跟踪: 该方法利用多边形光束的分析跟踪技术, 精确计算目标面元的照明、反射和阴影情况。这一技术的优势在于避免了传统方法中对细网格的依赖, 显著提高了计算效率, 同时保持了较高的精度。

该技术的核心公式用来计算多阶散射路径的散射场矢量, 公式如下:

$$E_{\text{scat}} = \int_A \mathbf{R}_{\text{GO}} dS e^{-jk \cdot r} \mathbf{P}_{\text{PO}} \tag{1}$$

其中:  $\mathbf{R}_{\text{GO}}$  表示几何光学反射矩阵,  $dS$  表示目标上微小面元的面积,  $k$  是入射和散射波矢量,  $r$  是参考点的传播距离,  $\mathbf{P}_{\text{PO}}$  是物理光学散射矩阵。

借助双向追踪机制, BART 能够在多次反射过程中精准构建完整的散射路径, 进而实现对复合散射现象的高精度捕捉与刻画。

1.2 图像预处理

在 SAR 图像的三维重建过程中, 仿真得到的初始图像常面临全局像素值过高、散斑噪声过大、旁瓣强度过高等问题。这些问题导致 SAR 图像质量低下, 进而严重影响后续的相机姿态获取以及三维重构工作的准确性和有效性。因此, 对 SAR 图像进行预处理是确保三维重建质量的关键步骤。

本文针对仿真结果图像出现的问题, 采用了对比度增强、散斑噪声抑制、旁瓣抑制和 SAR 图像边缘增强的方法。

1.2.1 对比度增强

针对全局像素值过高的挑战, 本文首先采用直方图均衡化 (HE, histogram equalization) 方法。通过调整图像的直方图, 降低全局的像素值, 使图像的动态范围扩展, 从而增强图像中建筑物的可视性。然而, 这种方法可能会导致局部噪声的强度过高。为了克服这一问题, 本文进一步使用局部对比度增强技术, 如局部直方图均衡化 (CLAHE, contrast limited adaptive histogram equalization)<sup>[18]</sup>。CLAHE 能够在不过度增强背景噪声的同时, 提升建筑物等重要特征的对比如。

假设我们有一个图像, 其像素值范围为  $[0, L-1]$ , 局部直方图均衡化的基本公式为:

$$s = T(r) = \frac{\sum_{i=0}^r h(i)}{M \times N} \tag{2}$$

其中:  $s$  是输出图像的像素值,  $r$  是输入图像的像素值,  $h(i)$  是局部直方图中第  $i$  个灰度级的频数,  $M$  和  $N$  分别是局部区域的行数和列数, 即局部区域的面积。

1.2.2 散斑噪声抑制

在 SAR 图像处理中, 散斑噪声的存在是一个普遍

且显著的问题, 它严重影响了图像的质量和后续分析的准确性。为了在不影响图像质量的前提下有效抑制 SAR 图像中的散斑噪声, 同时保留目标建筑的边缘和结构信息, 本文选择了 Frost 滤波器<sup>[19]</sup>, Frost 滤波器的选用是基于其能够根据局部统计特性自适应调整滤波器权重的特点, 这使其在保留目标建筑的边缘和结构信息方面表现优异。相比之下, 同样基于局部统计特性的 Lee 滤波器虽然在处理包含大量的图像时表现出色, 但其在弱化表面细节信息的同时, 可能无法充分保留目标的结构和边缘信息, 而这正是 SAR 三维重构所重点关注的。因此, Frost 滤波器更符合本文的研究需求。

Frost 滤波器根据局部统计特性自适应调整滤波器的权重, 具体公式如下:

$$w(i, j) = e^{-\alpha \cdot d(i, j)} \tag{3}$$

$$\hat{I}(i, j) = \frac{\sum_{(k, l) \in W} w(k, l) \cdot I(k, l)}{\sum_{(k, l) \in W} w(k, l)} \tag{4}$$

其中:  $w(i, j)$  表示滤波器在位置  $(i, j)$  的权重,  $\alpha$  是一个控制权重变化的参数, 用于调节滤波器的平滑程度,  $d(i, j)$  是位置  $(i, j)$  与中心像素的距离,  $I(k, l)$  是输入图像在位置  $(k, l)$  的像素值,  $w$  是滤波器的窗口, 包含了中心像素及其邻域内的像素,  $\hat{I}(i, j)$  是滤波后的输出图像在位置  $(i, j)$  的像素值。

1.2.3 旁瓣抑制

由于单一的旁瓣抑制方法可能对图像的质量造成较大的影响, 本文采用带阻滤波器来抑制特定频率范围内的旁瓣信号。带阻滤波器的频率响应函数可以用如下公式表示:

$$H(f) = \begin{cases} 1, & \text{if } |f| < f_L \text{ or } |f| > f_H \\ 0, & \text{if } f_L \leq |f| \leq f_H \end{cases} \tag{5}$$

带阻滤波器通常通过设计一个高通滤波器和一个低通滤波器的组合来实现, 即将某个频段的频率信号进行衰减, 具体的滤波效果也可以通过无限脉冲响应 (IIR, infinite impulse response) 或有限脉冲响应 (FIR, finite impulse response) 滤波器来实现。其差分方程形式为:

$$y[n]x[n] + \sum_{i=1}^N a_i \cdot x[n-i] + \sum_{j=1}^M b_j \cdot y[n-j] \tag{6}$$

为了进一步优化图像质量, 本文采用了最小均方 (LMS, Least Mean Square) 自适应滤波器。LMS 自适应滤波器通过调整图像中每个像素点的权重, 根据其周围像素的强度和分布情况来降低旁瓣的影响。这种方法需要精确地估计旁瓣的贡献, 并相应地调整权重, 以减少旁瓣对图像质量的影响, 同时尽量保留图像细节和质量。

LMS 自适应滤波器的权重更新公式如下:

$$w[n+1] = w[n] + 2\mu e[n]x[n] \tag{7}$$

终的滤波输出为:

$$y[n] = \sum_{i=0}^{N-1} w[i]x[n-i] \quad (8)$$

通过上述方法, LMS 自适应滤波器能够有效地调整权重, 以适应不同的图像特征, 从而在降低旁瓣影响的同时, 保留图像的细节和质量, 为后续的图像处理和提供高质量的图像基础。

#### 1.2.4 SAR 图像边缘增强

考虑到三维重构重点关注建筑的边缘和结构信息, 而这些信息主要保存在图像的边缘特征中, 本文采用 Sobel 算子进行边缘检测。Sobel 算子<sup>[20]</sup>通过两个卷积核分别计算水平方向 ( $G_x$ ) 和垂直方向 ( $G_y$ ) 的梯度:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (9)$$

对于图像中的每个像素, 计算出水平方向和垂直方向的梯度后, 组合成总体的梯度幅度:

$$G = \sqrt{G_x^2 + G_y^2} \quad (10)$$

为了进一步增强这些边缘特征, 本文应用拉普拉斯滤波器。拉普拉斯算子是各向同性的, 因此它对所有方向的边缘都同样敏感。拉普拉斯滤波器常用的卷积核为:

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (11)$$

对于每个像素, 拉普拉斯滤波器计算的是该像素与其周围像素之间的差异, 即:

$$\text{Laplacian}(x, y) = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \quad (12)$$

通过上述方法, Sobel 算子和拉普拉斯滤波器的结合不仅能够有效地检测和增强图像中的边缘特征, 还能在保留图像细节的同时, 减少旁瓣对图像质量的影响。这为后续的三维重构提供了更清晰、更准确的边缘信息, 从而可以提高三维重建的精度和质量。

#### 1.3 相机姿态获取

本文选取 ColMap 来获取相机姿态, 其算法流程如图 2 所示。该流程从输入的一系列不同视角的图像开始, 这些图像用于全面覆盖待重建场景的各个方面。首先, 算法执行特征提取、特征匹配和几何验证, 以识别输入图像中的场景重叠区域, 并定位重叠图像中

相同点的投影位置。该步骤的输出为经过几何验证的图像对以及每个点的图像投影数据。

在增量重建阶段, 算法采用迭代方式逐步添加图像并进行处理。初始化阶段作为增量重建的起点, 通过匹配点生成三维点, 并估算初始图像对的相机姿态。随后进行图像配准, 将新的图像纳入系统, 并通过匹配当前三维模型中的点来计算新增图像的相机姿态。在此基础上, 算法进行三角化, 生成新的三维点。

由于后续的相机姿态估计是以特征匹配得到的匹配点为基础数据, 因此特征匹配对后续工作的影响很大, 当特征匹配的数量不足时, 会导致后续增量重建算法的失败。具体来说, 匹配点数量过少或匹配质量差, 会导致三角化生成的三维点不足, 从而使得初始场景重建不稳定。在增量重建中, 每次加入新的相机都需要对其与已有场景的特征点匹配进行验证。如果匹配点不足, 新的相机无法正确对齐, 模型会逐步失去稳定性甚至崩溃。SAR 图像相对于光学图像更加复杂且难以解读, 更容易出现上述的情况。此外, 由于采用了仿真的方式, 图像间时序性更强, 这是值得利用的信息。因此, 本文摒弃了原来的枚举匹配算法, 采用了时序性更强的时序匹配算法<sup>[21]</sup>。重点关注相邻图像间的匹配信息, 通过传递完成所有图像之间的信息匹配。具体算法描述如下。

##### 1) 时序图像预处理:

将连续采集的 SAR 图像按时间顺序划分为若干组, 每组包含相邻的张图像。每组内图像共享时序关联性, 确保匹配连续性。对每组图像采用 SIFT 特征检测, 提取关键点及其描述子。为适应 SAR 图像特性, 引入

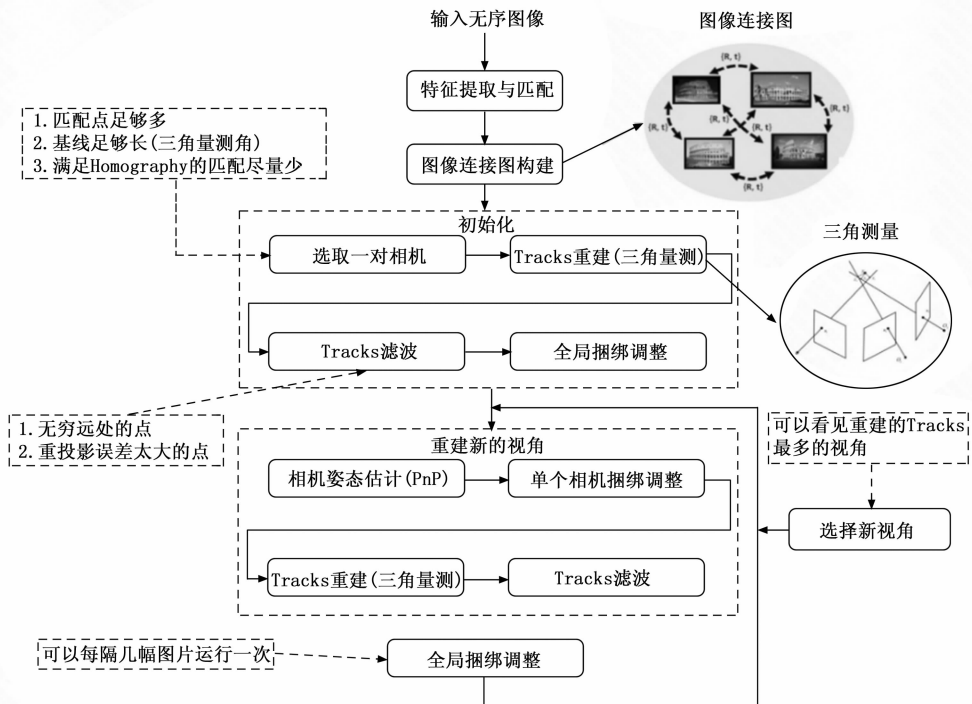


图 2 增量运动结构管道

Frost 滤波预处理,抑制散斑噪声对特征稳定性的影响。

2) 定义绝对误差:

$$\epsilon(i, j, s, t) = |S_{i,j}(s, t) - \bar{S}_{i,j} - T(s, t) + \bar{T}| \quad (13)$$

其中:带有上划线的分别表示子图、模板的均值:

$$\bar{S}_{i,j} = E(S_{i,j}) = \frac{1}{M \times N} \sum_{s=1}^M \sum_{t=1}^N S_{i,j}(s, t) \quad (14)$$

$$\bar{T} = E(T) = \frac{1}{M \times N} \sum_{s=1}^M \sum_{t=1}^N T(s, t) \quad (15)$$

绝对误差即子图与模板图各自去掉其均值后,对应位置之差的绝对值。

3) 设定阈值  $T_h$ :

手动设定固定阈值  $T_h$ ,通过阈值控制匹配误差。当计算得到的绝对误差超过该阈值时,认为子图与模板图在该位置不匹配,提前终止计算,减少不必要的计算量。从图像序列中,选择相邻的图像作为模板图像和叠加子图像。优先选择时间上相邻且视角变化较小的图像对,因为这些图像之间的特征相关性较高,更有利于匹配。

4) 误差累积计算:

在随机选择模板图像与叠加子图像的未重复点时,计算模板图像与叠加子图像在对应位置的误差  $\epsilon$ 。当  $\epsilon < T_h$  时,继续选择下一个点并进行误差计算,同时累积误差;当  $\epsilon > T_h$  时,重新比较累积误差  $\sum \epsilon$  与  $T_h$ ,并记录累积次数  $H$ 。定义序贯相似性检测算法(SSDA, sequential similarity detection algorithm)的检测面为:

$$R(i, j) = \{H \min_{1 \leq H \leq M \times N} [\sum_{h=1}^H \epsilon(i, j, s, t) \geq T_h]\} \quad (16)$$

取  $R(i, j)$  的最大值位置  $(i, j)$  作为匹配点,该点通过在阈值  $T_h$  上对匹配子图与模板图的迭代累加所得到的最佳匹配位置。误差累积情况如图3所示。图中, A 和 B 的误差累积比 C 快,且比 C 早到达阈值。C 累积次数最多,可能代表正确的匹配位置。

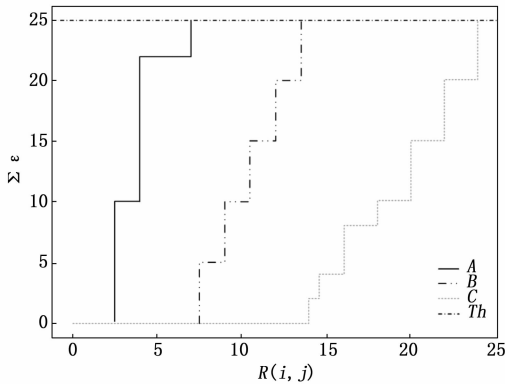


图3 时序匹配算法误差的累积曲线

5) 匹配点传递与补全:

若相邻图像间匹配点数量不足,利用前一帧的匹配结果作为先验,通过光流法预测当前帧特征点位置。基于亮度恒定假设,光流方程可表示为:

$$\nabla I \cdot v + I_t = 0 \quad (17)$$

其中:  $v$  为光流向量,  $\nabla I$  为图像梯度,  $I_t$  为时间导数。通过求解光流方程,可以预测出当前帧中特征点可能出现的位置,实现匹配点从前一帧到当前帧的传递。

由于光流法预测存在误差,会引入一些误匹配的,这些点会影响后续姿态估计的准确性。因此,采用双向一致性检验来判断预测点的准确性。利 RANSAC 算法进一步剔除离群点, RANSAC 能够通过迭代随机采样最小样本集,统计内点数量,最终选择内点最多的模型。公式为:

$$\|y_i - f(x_i; \theta)\| < \tau \quad (18)$$

其中:  $y_i$  为观测值,  $f(x_i; \theta)$  为模型预测值,  $\tau$  为距离阈值。经过多次迭代后,最终选择内点最多的模型,将不符合该模型的离群点剔除,从而提高匹配点的质量。采用双线性插值的方法补充匹配点。双线性插值是一种常用的图像插值方法,它根据周围4个已知像素点的灰度值,通过线性插值计算出目标位置的像素值:

$$I(x, y) = (1 - \alpha)(1 - \beta) I_{00} + \alpha(1 - \beta) I_{10} + (1 - \alpha)\beta I_{01} + \alpha\beta I_{11} \quad (19)$$

其中:  $\alpha$  和  $\beta$  为归一化坐标偏移量。  $I_{00}$ ,  $I_{10}$ ,  $I_{01}$ ,  $I_{11}$  为插值权重的相邻4个像素。在匹配点补全中,利用周围已有的匹配点信息,通过双线性插值计算出稀疏区域的匹配点,确保图像中各个区域都有足够的匹配点用于姿态估计,进而保证姿态估计的连续性和准确性。

#### 1.4 NeRF

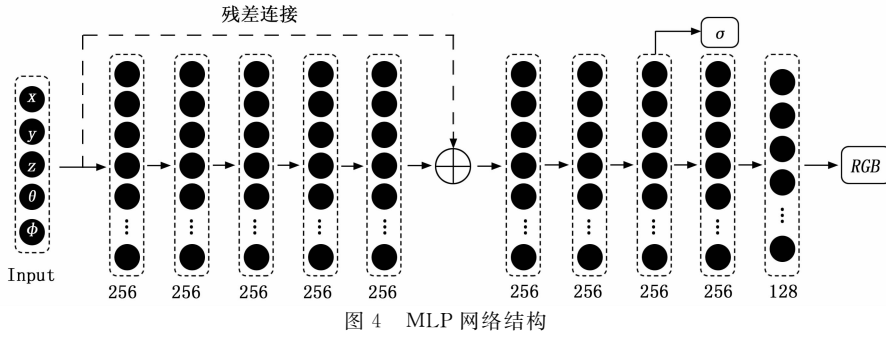
NeRF 的核心思想是训练一个多层感知机 (MLP, multi-layer perceptron)<sup>[22]</sup> 来拟合场景的辐射场,并通过少量的2D图像及其相机参数作为输入,生成任意视角下的高质量图像。与传统方法相比,NeRF能够在处理复杂几何形状和精细纹理方面展现出更强的表现力。

##### 1.4.1 NeRF 模型

NeRF 模型通过神经网络表示三维场景的辐射场。辐射场描述了在各个视角下,场景中每一点的颜色和密度。具体地,辐射场的数学表达式为:  $F(x, \theta, \varphi) \rightarrow (c, \sigma)$ 。其中,  $x$  表示三维空间中的坐标点,  $(\theta, \varphi)$  表示射线的水平角和俯仰角,  $c$  表示颜色,  $\sigma$  表示体积网格密度。该函数由一个多层感知机,记为  $F_\theta$ , 进行近似。 $(\theta, \varphi)$  也可以表达为三维单位方向向量  $d$ 。

该表达需满足多视图一致性,即:场景的内容  $(\sigma)$  与视角无关,而颜色  $(c)$  则与坐标和视角均有关。通常,NeRF 模型采用两个 MLP 网络,第一个 MLP 以  $x$  为输入,输出体积密度  $\sigma$  以及一个高维特征向量  $h$ :  $\sigma, h = \sigma_{\text{MLP}}(x)$ 。然后特征向量  $h$  和视线方向  $d$  拼接后输入第二个 MLP,输出最终的颜色  $c$ :  $c = c_{\text{MLP}}[h; d]$ 。

其中 NeRF 近似使用的 MLP 的具体网络结构如图4所示。



#### 1.4.2 新视图合成

NeRF 模型通过沿相机光线采样 5D 坐标（包括位置和观察方向）来合成图像，如图 5 中（a）所示。每个采样点包括一个三维空间位置  $(x, y, z)$  和视角方向  $(\theta, \varphi)$ ，这两者共同构成了一个 5D 输入向量。首先，这些输入会经过三角函数编码处理，将低频信息映射到高频，从而更好地适应神经网络学习。（b）展示了光度追踪算法得到的仿真光线 Ray1、Ray2，并将这些位置输入 MLP 以生成颜色 RGB 和体积密度  $\sigma$ 。（c）展示了使用体积渲染技术将这些得到的参数合成为图像，通过沿射线路径在多个点上采样体积密度和颜色，使用积分渲染公式来计算最终的颜色输出。（d）展示了模型训练过程，训练目标是通过最小化渲染损失来优化网络参数，该损失是基于射线预测颜色和实际颜色（g. t., ground truth）之间的差异，通常使用 L2 范数<sup>[23]</sup>（即欧氏距离）来衡量损失。

#### 1.4.3 体积网格渲染

设相机射线为  $r(t) = o + td$ ，其中  $o$  为相机位置， $d$  为射线方向。射线上任意位置的颜色  $C(r)$  可通过以下公式计算：

$$C(r) = \int_{t_1}^{t_2} T(t) \cdot \sigma[r(t)] \cdot c[r(t), d] dt \quad (20)$$

其中： $\sigma[r(t)]$  和  $c[r(t), d]$  分别为射线  $r(t)$  处的体积网格密度和颜色， $T(t)$  为累积透明度，表示射线从  $t_1$  到  $t$  之间未被遮挡的概率：

$$T(t) = \exp\left(-\int_{t_1}^t \sigma[r(u)] du\right) \quad (21)$$

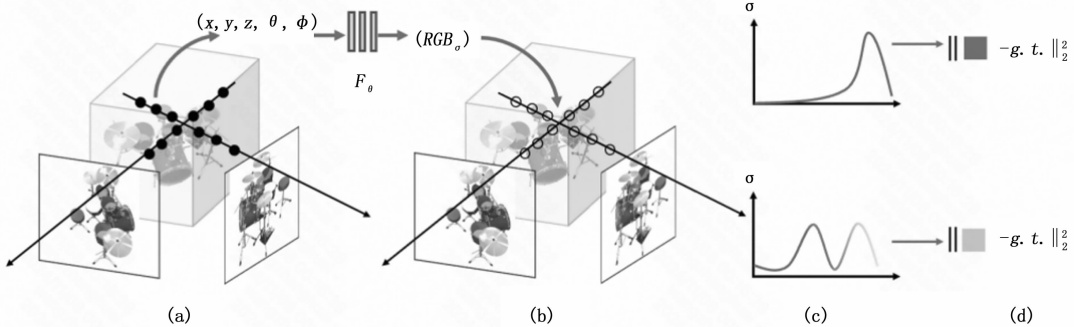


图 5 神经辐射场场景表示和可微渲染过程

通过追踪待合成图像的像素对应的相机射线，可以计算积分。但通常会通过分层抽样法来计算其近似值。具体来说，将射线分为等长的段，然后在各段内均匀抽样一个点，并用求和来近似上述积分：

$$\hat{C}(r) = \sum_{i=1}^N \alpha_i T_i c_i; T_i = \exp\left(-\sum_{j=1}^{i-1} \sigma_j \delta_j\right) \quad (22)$$

其中： $\delta_i$  表示采样点  $i$  与  $i+1$  的距离； $(\sigma_i, c_i)$  表示射线上采样点的密度和颜色（通过 NeRF 的 MLP 计算得到）。 $\alpha_i$  表示采样点  $i$  处的合成透明度/不透明度： $\alpha_i = 1 - \exp(-\sigma_i \delta_i)$ 。

#### 1.4.4 期望深度计算

期望的深度可以通过沿射线计算积累透明度得到，其计算公式为：

$$d(r) = \int_{t_1}^{t_2} T(t) \cdot \sigma[r(t)] \cdot t \cdot dt \quad (23)$$

上式同样可以通过求和近似：

$$\hat{D}(r) = \sum_{i=1}^N \alpha_i t_i T_i \quad (24)$$

#### 1.4.5 损失函数

为优化 MLP 参数，本文考虑了每个像素的平方误差光度损失。具体而言，总体损失函数定义如下：

$$L = \sum_{r \in R} \|\hat{C}(r) - C_{gt}(r)\|^2 \quad (25)$$

其中： $C_{gt}(r)$  是与射线  $r$  关联像素的真实颜色， $R$  是待合成图像的相机射线集合。

#### 1.4.6 位置编码

NeRF 模型通常使用位置编码来提高图像渲染的细节。最初的编码方式是将位置  $x$  的坐标（归一化到  $[-1, 1]$  区间）以及射线单位方向向量  $d$  的每个分量  $v$  进行如下变换：

$$\gamma(v) = \sin(2^0 \pi v), \cos(2^0 \pi v), \dots, \sin(2^{N-1} \pi v), \cos(2^{N-1} \pi v) \quad (26)$$

对 3D 位置， $\gamma(v)$  设置  $N=10$ ；对于视角信息， $\gamma(v)$  设置  $N=4$ 。

然而，由于最原始的 NeRF 模型通常需要大量具有

小基线且高质量的输入图像来生成精确的 3D 网格, 并且每条光线上的点都要经过 MLP 查询, 同时 NeRF 的位置编码方式需要网络自适应选择某些维度来生成体密度和颜色值, 这些问题导致网络规模过大, 计算效率低, 且增加了模型过拟合的风险。为此, 本文借鉴了 Instant-ngp<sup>[24]</sup> 方法的思路, 解决了 NeRF 在全连接神经网络参数化时的效率问题。该方法提出一种新的编码方式, 算法细节如图 6 所示, 使得可以使用较小规模的网络实现 NeRF, 同时不损失精度。

具体而言, 这种编码方法包括以下几点:

1) 多层次结构

编码由  $L$  层组成, 每层包含一个哈希表, 用于存储特征向量。每层的分辨率逐级递增, 使得可以捕获不同尺度的细节。

2) 网格分辨率的几何进展

每层网格的分辨率  $N_i$  按几何级数分布, 最粗和最细分辨率之间的计算公式为:  $N_i = \lfloor N_{min} \cdot b^i \rfloor$ 。

其中,  $b$  是增长因子, 定义为:

$$b = \exp\left(\frac{\ln N_{max} - \ln N_{min}}{L - 1}\right)$$

(27)

3) 空间哈希函数

在更细的分辨率层中, 使用空间哈希函数  $h(x)$  将坐标映射到特征向量数组中, 公式为:

$$h(x) = (\bigoplus_{i=1}^d x_i \pi_i) \bmod T$$

(28)

其中:  $\pi_i$  是不同的质数, 确保哈希操作中的维度分散。  $T$  是哈希表的大小, 决定了存储特征向量的容量。

4) 线性插值

为了减少哈希冲突带来的误差, 每层的特征向量在其  $d$  维网格中进行线性插值。具体来说, 对于每个特征向量  $x$ , 我们在  $d$  维网格中的相邻两个点之间进行插值。插值权重定义为:  $w_i = x_i - \lfloor x_i \rfloor$ , 其中  $\lfloor x_i \rfloor$

表示  $x_i$  的向下取整, 而  $x_i - \lfloor x_i \rfloor$  表示  $x_i$  距离下一个整数点的距离。通过这种方式, 线性插值可以更平滑地处理特征向量的变化, 从而减少因哈希冲突导致的误差。这样可以根据输入位置的相对位置平滑地插入特征值。

这些编码后的特征向量被串联形成最终的神经网络输入, 用于提升模型在多种任务中的表现和计算效率。该网络通过多分辨率哈希表来增强特征向量, 并基于随机梯度下降 (SGD, stochastic gradient descent)<sup>[25]</sup> 进行优化。多分辨率结构有助于图形处理单元 (GPU, graphics processing unit) 并行, 并通过消除哈希冲突减少计算量。该方法的最大亮点在于将 NeRF 从小时级大幅缩减至秒级, 同时显著降低了模型对输入图像质量的敏感性, 展现了对低质量图像的更强鲁棒性, 从而提升了其在 SAR 领域中的泛化能力。

2 实验与结果

2.1 SAR 建筑物仿真图像

本文选取复旦大学的光华楼作为仿真模型, 该建筑由中间两幢 31 层的主楼及东西两侧各一幢 8 层的辅楼组成, 整个建筑东西总长约 180 米, 南北宽约 50 米, 高约 140 米。如图 7 所示。光华楼复杂结构, 层次差异显著, 是评估 SAR 图像处理算法性能的理想对象。

在本次仿真工作中, 本文采用的仿真参数设置如表 1 所示。

表 1 仿真参数设置

| 参数名称   | 参数值            | 具体值    |
|--------|----------------|--------|
| 中心频率   | $f_c$          | 10 GHz |
| 入射角    | $\theta_i$     | 60°    |
| 方位向分辨率 | $R_a$          | 0.1 m  |
| 距离向分辨率 | $R_r$          | 0.1 m  |
| 成像间隔   | $\Delta\theta$ | 5°     |

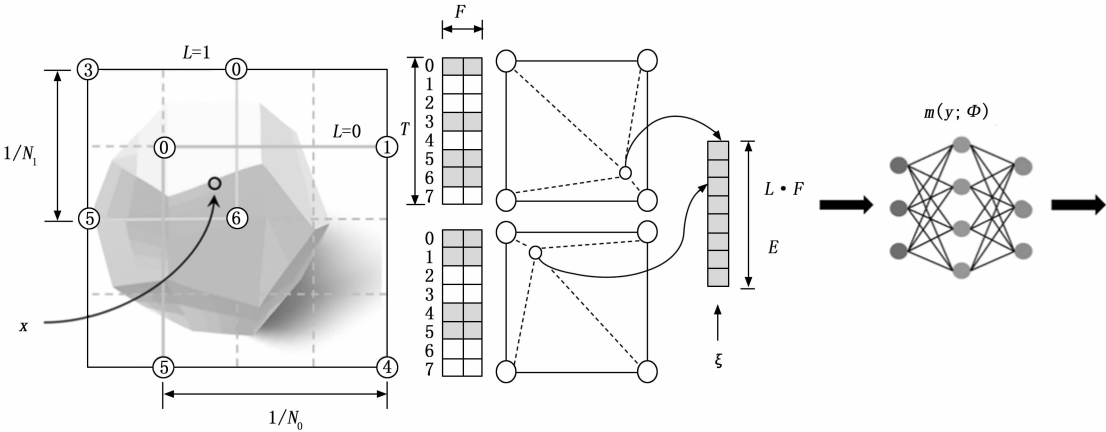


图 6 多尺度 hash 编码

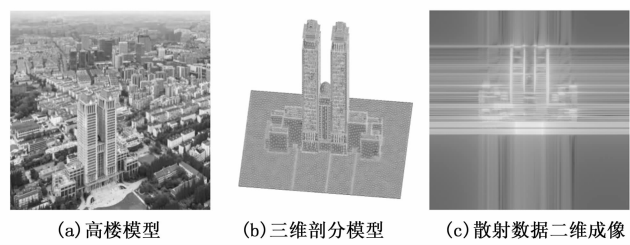


图 7 建筑结构与散射成像分析框架

中心频率  $f_c=10\text{ GHz}$ ，这一高频选择旨在提高图像的分辨率和细节表现力。入射角  $\theta_i=60^\circ$ ，该角度平衡了信号的反射效率和覆盖范围，确保了对建筑结构的有效成像。方位向和距离向分辨率为  $0.1\text{ m}$ ，这一高分辨率配置有助于捕捉建筑的精细特征。成像方式采用类似圆迹的方式，间隔角度设定为  $5^\circ$ ，这种设置能够在保证数据采集效率的同时，提供足够的视角多样性，以支持后续的三维重建和分析工作。

经过精细的参数配置与仿真处理，本研究最终生成了 72 张 SAR 建筑物仿真图像，如图 8 所示，这些图像涵盖了从不同角度对光华楼进行观测的丰富数据，将用于后续的图像分析与三维重建。

2.2 图像预处理

依据 1.2 节所述的图像增强策略，本文对 SAR 建筑物仿真图像实施了一系列数据预处理步骤，涵盖对比度增强、散斑噪声抑制、旁瓣抑制、边缘增强以及掩膜

处理。经此处理，SAR 图像质量实现显著跃升。具体而言，对比度增强有效提升了目标与背景之间的亮度差异，令目标特征轮廓愈发鲜明，便于后续识别与分析；散斑噪声抑制环节精准削弱了雷达成像过程中不可避免的随机噪声干扰，显著提升了图像的信噪比，使图像细节得以清晰呈现；旁瓣抑制技术精准削减了主波束邻近区域的能量溢出，优化了图像分辨率，增强了细节的可辨识度；边缘增强技术进一步凸显了目标的边缘轮廓线条，赋予结构特征更加锐利的视觉效果。如图 9 和表 2 所示。综合上述处理手段，通过 Frost 滤波器和 LMS 自适应滤波器，散斑噪声显著降低，SNR 从 18.6 dB 提升至 24.3 dB。通过直方图均衡化（HE）与局部对比度受限自适应直方图均衡化（CLAHE）技术，图像标准差从预处理前的 28.6 提升至 41.5，对比度提升率达 45.1%。这表明预处理后图像的全局与局部对比度均显著增强，目标建筑物与背景的区分度更加明显，细节纹理得到有效凸显。SAR 图像的视觉品质与结构细节保真度均获大幅提升，为目标检测与三维重建作业奠定了数据基础。

表 2 图像预处理定量评估结果

| 评估指标   | 预处理前 | 预处理后 | 提升率/% |
|--------|------|------|-------|
| 信噪比/dB | 18.6 | 24.3 | 30.6  |
| 等效视数   | 4.2  | 11.5 | 173.8 |
| 对比度标准差 | 28.6 | 41.5 | 45.1  |

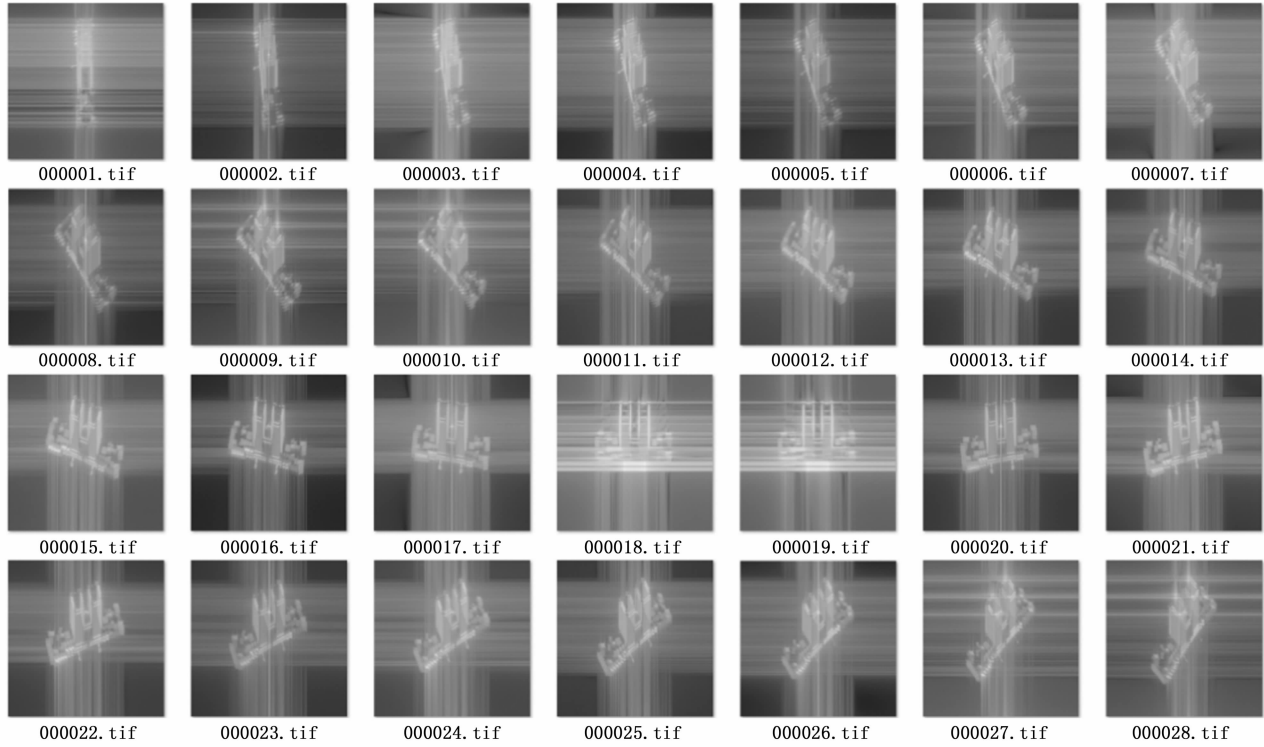


图 8 SAR 建筑物仿真图像

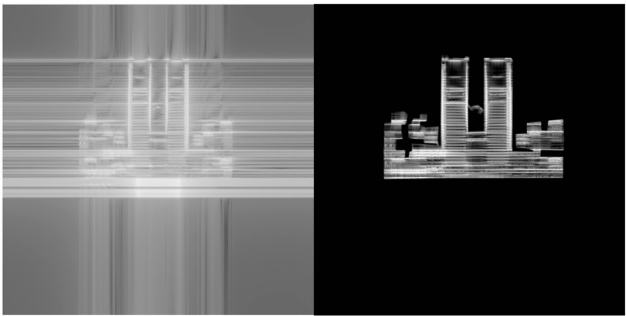
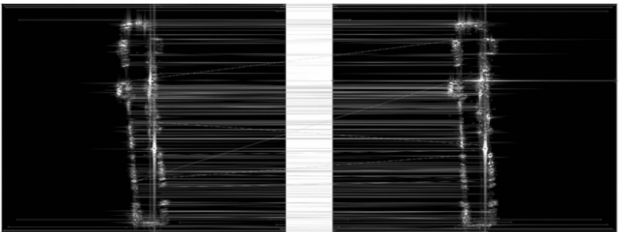


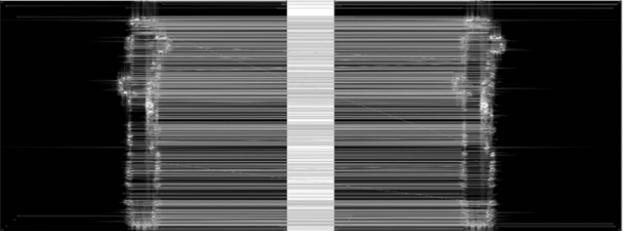
图 9 图像预处理

2.3 ColMap 相机姿态获取

在本文中, 对枚举匹配算法和时序匹配算法的匹配效果进行了深入评估。鉴于本文仿真的数据集模型为建筑物, 枚举匹配算法在处理此类数据时表现出明显的不足。具体而言, 当目标的 4 个面结构信息较为接近时, 如本文所使用的光华楼数据集, 枚举匹配算法不仅耗时较长, 而且匹配效果也不尽如人意。这种情况下, 算法往往只能匹配到物体一个面的信息, 导致其余 3 个面的信息丢失, 进而引发相机姿态的错误, 使得后续重构的三维模型出现大量发散点, 甚至无法进行有效的后续重构。相比之下, 时序匹配算法通过选择相邻的  $x$  张图像进行匹配, 充分利用了图像间的时序信息。这种方法不仅避免了枚举匹配算法丢失三维信息的问题, 还能够获取更为稠密的匹配点信息, 从而显著提高了匹配的准确性和效率。如图 10 所示, 时序匹配算法在处理具有相似结构信息的多面体时, 能够更全面地捕捉到目标的各个面的信息, 为后续的三维重建提供了更可靠的数据支持。



(a) 枚举匹配算法



(b) 时序匹配算法

图 10 匹配结果对比

在匹配算法替换后, 本文利用 ColMap 软件包来获

取 SAR 建筑物仿真图像的相机姿态和稀疏重构模型, 如图 11 所示, 通过这一处理流程, 我们成功获得了稀疏重构后的三维模型以及对应的图像的相机姿态参数。这些参数详细描述了相机在拍摄每张图像时的位置和方向, 为后续的三维重建工作提供了关键的几何信息。



图 11 相机姿态与稀疏重构模型

我们同时获得了图像的深度图信息。深度图中的单个像素值表示空间中某个点到垂直于左红外镜头光轴并通过镜头光心 (深度相机光学零点) 平面的垂直距离。如图 12 所示。

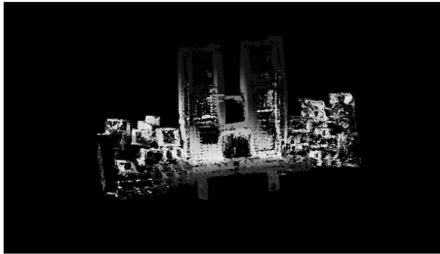


图 12 深度图

深度图是一个由深度相机视野内所有点的深度数据构成的矩阵。该深度图为 16 bit 位深的单通道矩阵, 单位为毫米。对于没有深度信息的点, 其值为 0。为了直观地表示不同的距离数值, 输出的深度图均被映射到了 RGB 色彩空间, 因此显示的结果为 RGB 三通道、8 bit 位深的位图。需要注意的是, 深度图的数据没有外参信息和畸变修正, 仅提供用于转换点云数据的内参。因此, 深度相机的每一帧的深度图像相当于在相机局部三维坐标系中的点云模型。

2.4 NeRF 三维重建

在本文中, NeRF 模型的训练在单张 Nvidia 3090 GPU 上训练, 采用 RMSProp 优化器, 初始学习率设为  $5 \times 10^{-5}$ 。训练过程中, 批处理大小设置为 1 024 条射线, 每条射线被离散化为 8 192 个均匀分布的三维点。模型在  $2 \times 10^6$  次迭代 (共 17 个 epoch) 下完成训练。训练损失使用均方误差 (MSE, mean squared error) 定义如下:

$$MSE(I,G) = \frac{1}{n} \sum_{i=1}^n (G_i - I_i)^2 \tag{29}$$

其中： $I$  表示新视角生成的图像， $G$  表示真实值。

本文使用前文得到的 72 张 SAR 图像仿真数据集进行三维重构，模型所需的相机姿态以及稀疏重构的信息都通过 2.3 节获取，得到 3 个文件，分别为 camera.txt、image.txt、point3D.txt。这些文件中分别存储着相机的内参、外参以及二维特征点与稀疏三维模型点的映射关系。

为了将这些文件整合为 LLFF 格式，我们进行了以下步骤：

1) 通过 LLFF 官方格式转换步骤，将 camera.txt、image.txt、point3D.txt 这 3 个 txt 文件整合为一个 poses\_bounds.npy 文件。LLFF 格式需要 3 个图像文件夹，分别是像素值为  $4\ 032 \times 3\ 024$  的 images 文件夹，以及下采样 4 倍、8 倍的 images\_4 和 images\_8 文件夹。

2) 本文对 2.2 节预处理得到的图像进行像素扩展以及下采样，以满足 LLFF 格式的要求。通过上述步骤，NeRF 训练所需要的自建数据集才算完成。

通过上述实验可以观察到，重构的三维模型成功地恢复了预期的结构轮廓，尤其是建筑的柱体结构及楼梯中间的圆盘状特征得到了较为清晰的展示。然而，模型表面细节的呈现仍有改进空间，特别是建筑底层结构存在一定程度的模糊，而边缘部分则呈现出雾状的重构效果。这些问题可能源于相机姿态的误差，导致网络在该空间区域内难以收敛，或者网络错误地将该区域识别为包含物体结构信息的区域。总体而言，模型重构结果基本符合预期，成功恢复了目标建筑的三维信息。实验结果如图 13 所示。

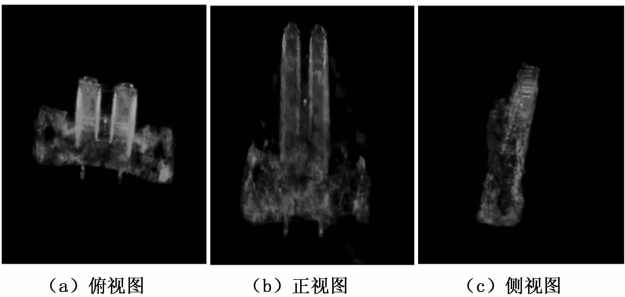


图 13 SAR 建筑物图像渲染结果

神经辐射场生成的三维模型采用隐式表示，其中目标的三维结构特征被编码在神经网络的权重和参数中，而非以显式的几何形式（例如点云或网格）直接呈现。这种隐式表示方法赋予了模型极大的灵活性，使其能够高效地捕捉复杂的细节特征。然而，对于许多实际应用而言，隐式表示的三维模型缺乏直观性或

易操作性。因此，为了将 NeRF 生成的隐式模型转化为更常用的显式形式，本文选择将其转换为点云模型。这一转化过程包括以下几个关键步骤：空间采样、辐射值查询以及点云生成。最终，通过这一系列操作，NeRF 模型的隐式表示被成功转化为点云形式，转化后的结果如图 14 所示。

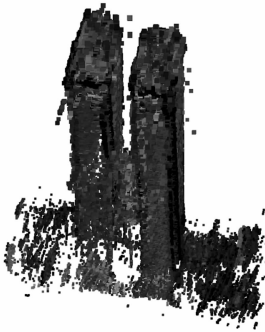


图 14 SAR 建筑物图像三维重构结果

为了全面评估所提出算法的重构性能，本文对最终得到的三维模型与真实模型的几何比例进行了量化对比。这一评估方法通过比较模型的高度、长度和宽度比例，能够有效反映重构结果的准确性。具体评估结果如表 3 所示。

表 3 三维重建算法性能与模型尺寸综合评估表

| 性能/模型            | 指标/模型尺寸       | 柱体区域 | 地基区域 | 全局   |
|------------------|---------------|------|------|------|
| 性能评估             | RMSE/m        | 0.22 | 0.68 | 0.45 |
|                  | Hausdorff/m   | 0.45 | 1.89 | 1.25 |
|                  | 完整度/%         | 95   | 62   | 82   |
|                  | 重叠率/%         | 92   | 58   | 78   |
| 光学仿真模型(高度:长度:宽度) | 7.3:4.55:1    | —    | —    | —    |
| 重构模型(高度:长度:宽度)   | 7.874:5.058:1 | —    | —    | —    |

在对光华楼的三维重建和图像处理算法进行性能评估时，我们发现算法在柱体区域的表现十分出色。具体来说，柱体区域的均方根误差（RMSE，root mean squared error）仅为 0.22 米，这一数值显著低于地基区域的 0.68 米，显示出算法在主体区域的预测精度较高。全局 RMSE 为 0.45 米，表明算法在整体上保持了良好的精度水平；在衡量点集匹配度的 Hausdorff 距离指标上，柱体部分的距离为 0.45 米，而地基部分的距离为 1.89 米，这表明算法在地基部分的匹配度有待提高。尽管全局 Hausdorff 距离为 1.25 米，略高于柱体部分，但整体上仍然显示出算法在点集匹配方面具有较好的性能；在完整度方面，柱体部分达到了 95%，这一高比例表明算法在柱体部分能够高效地重建目标的完整结构。地基部分的完整度为 62%，虽然相对较低，但这

也为算法在地基部分的进一步优化提供了明确的方向。全局完整度为 82%，这一数值虽然略低于柱体部分，但仍然显示出算法在大多数情况下能够保持目标的完整性；重叠率指标也显示了算法的整体一致性。柱体部分的重叠率为 92%，这表明算法在柱体部分的重建结果与真实情况具有很高的 consistency。地基部分的重叠率为 58%，虽然这一比例较低，但它指出了算法在地基部分一致性方面需要改进的地方。全局重叠率为 78%，这一数值虽然略低于柱体部分，但仍然表明算法在整体上具有较好的一致性。通过对比光学仿真模型和重构模型的尺寸比例，我们发现光学仿真模型的比例为高：长：宽=7.3：4.55：1，而重构模型的比例为高：长：宽=7.874：5.058：1。两者之间的误差控制在 6% 以内，这一结果表明，所提出的算法在保持模型几何特征方面表现出色，能够以较高的精度实现三维模型的重构。这种量化评估不仅为算法的性能提供了直观的证据，也为未来在更复杂场景下的应用提供了信心和参考。总体而言，算法在柱体区域的性能令人满意，而在地基区域的性能虽然相对较低，但也显示出了改进的潜力。未来的工作将集中在对算法进行特别的调整或优化，以提高其在地基区域的重建精度和一致性，从而进一步提升算法的整体性能。

为验证本文提出的基于 NeRF 的 SAR 图像三维重建方法的有效性，本节从重建精度、计算效率和资源消耗 3 个维度对模型全局进行对比实验，选用传统层析成像和运动结构恢复作为基线方法。RMSE 量化重建模型与真实模型的空间几何误差，反映整体重建精度。Hausdorff 距离衡量两模型表面间的最大局部偏差，表征最劣情况下的重建误差。总耗时为包括数据预处理、特征匹配、模型训练/优化的全流程时间。内存/显存占用反映算法对计算资源的依赖程度。3 种方法在光华楼数据集上的性能对比如表 4 所示。

| 表 4 3 种方法在光华楼数据集上的性能对比 |        |             |       |                |
|------------------------|--------|-------------|-------|----------------|
| 方法                     | RMSE/m | Hausdorff/m | 总耗时/h | 内存/显存<br>占用/GB |
| 传统层析成像                 | 1.32   | 3.45        | 4.5   | 12.8           |
| 运动结构恢复                 | 0.89   | 2.76        | 6.2   | 9.6            |
| 本文方法                   | 0.45   | 1.25        | 2.8   | 5.2            |

如表 4 所示，本文方法的 RMSE 较传统层析成像和 SFM 分别降低 66% 和 49%，表明 NeRF 通过隐式辐射场建模可更精准捕捉复杂几何特征。Hausdorff 距离优于传统方法，证明算法在最大局部偏差控制上表现稳健，这是因为时序匹配算法，通过利用 SAR 图像的时序信息生成更稠密的匹配点，减少了特征稀疏导致的局部误差。总耗时最少，主要归因于通过相邻图像信息传

递，减少了特征匹配时间。显存占用仅 5.2 GB，验证了模型轻量化设计的有效性；传统方法因依赖 CPU 并行计算，内存占用高达 12.8 GB。实验表明，本文方法在精度、效率与资源占用上均显著优于传统方法，为 SAR 图像三维重建提供了高效可靠的解决方案。

### 3 结束语

本文致力于将神经辐射场模型创新性地应用于合成孔径雷达图像的三维重建，旨在开发一种高效且精准的建筑物三维重建技术。鉴于当前可用于 NeRF 研究的 SAR 公开数据集极为有限，本研究巧妙地运用双向解析射线追踪方法，成功生成了高质量的仿真 SAR 数据集。同时，借助 ColMap 工具，精确获取了相机的姿态信息和稀疏重构数据，为 NeRF 模型的训练提供了坚实可靠的数据基础。此外，本研究针对 SAR 图像的独特特性，精心设计了一套综合性的图像预处理流程，涵盖图像增强、散斑噪声抑制、旁瓣抑制等关键技术环节，显著提升了 SAR 图像的整体质量，进而优化了重建模型的性能表现。

实验结果充分彰显了本文提出方法在建筑物三维重建领域的卓越效能。在建筑物的整体轮廓和结构特征重建方面，该方法呈现出令人瞩目的精准度，尤其在复杂形态的柱体结构和楼梯中间的圆盘状特征的重现上，达到了较高的清晰度。然而，由于 SAR 图像固有的特殊成像机制，部分细节在重建结果中仍呈现出一定程度的模糊现象，尤其是建筑物地基区域，这一问题尤为突出。经深入分析，主要存在以下几方面原因：

1) 相机姿态误差影响：在获取相机姿态信息时，尽管采用了 ColMap 工具与时序匹配算法，但 SAR 图像成像环境复杂，相机与地基部分的相对位置和角度多变，易受周围环境干扰，导致相机姿态估计存在误差。这使得地基部分点云数据定位偏差，直接影响了重建精度。

2) 散斑噪声干扰：SAR 图像中的散斑噪声是影响重建精度的重要因素。即便经过散斑噪声抑制处理，仍有部分噪声残留。地基部分纹理信息相对匮乏，对噪声更为敏感，噪声干扰了图像特征的提取与匹配，使得在相机姿态获取和三维重建过程中，对地基结构的判断出现偏差。

3) 模型泛化能力受限：NeRF 模型原本针对光学图像设计，而 SAR 图像与光学图像成像机制差异巨大。对于地基这种特殊区域，模型难以适应其复杂结构和成像特点，在学习和重建其几何形状与细节时存在困难，导致泛化能力不足，重建精度受限。

综上所述，本研究不仅有力地验证了 NeRF 模型在

SAR 图像三维重构中的可行性, 更为 SAR 图像的三维重建提供了一种全新的技术路径。尽管目前算法在柱体区域表现优异, 为了进一步提升算法的重建精度和一致性, 未来的工作将聚焦于以下几个关键方向:

探索多传感器融合的创新方法, 融合激光雷达等设备获取的高精度深度信息, 辅助相机姿态的精确估计。同时, 持续改进时序匹配算法, 引入更先进的特征提取和匹配策略, 增强算法对复杂环境的适应性, 减少匹配错误, 从而提高相机姿态估计的准确性。此外引入自适应学习机制, 使模型能够根据 SAR 图像的特点自动调整学习策略; 通过数据增强技术, 增加数据的多样性, 让模型学习到更丰富的特征, 从而提高模型的鲁棒性和重建精度。同时, 探索将不同类型的数据集进行融合, 如结合光学图像和 SAR 图像的优势, 为模型训练提供更全面的信息。

#### 参考文献:

- [1] 李 念, 刘 杰, 于君明, 等. 基于多域联合直达波估计的建筑布局层析成像方法 [J]. 雷达学报 (中英文), 2025, 14 (2): 309–321.
- [2] 李耀东. 分布式 SAR 三维成像与点云去噪方法研究 [D]. 成都: 电子科技大学, 2024.
- [3] 李雨轩, 刘 峥, 冉 磊. 改进三维最大类间方差的 SAR 图像海陆分割算法 [J]. 雷达科学与技术, 2024, 22 (4): 416–426.
- [4] 黄泽贤, 吴凡路, 傅 瑶, 等. 基于深度学习的遥感图像舰船目标检测算法综述 [J]. 光学精密工程, 2023, 31 (15): 2295–2318.
- [5] 李春升, 王伟杰, 王鹏波, 等. 星载 SAR 技术的现状与发展趋势 [J]. 电子与信息学报, 2016, 38 (1): 229–240.
- [6] 贾赫成, 蒲欣洋, 王燕妮, 等. 融合可微分渲染的 SAR 多视角样本增广 [J]. 雷达学报, 2024, 13 (2): 457–470.
- [7] 陈健堃, 彭凌霄, 仇晓兰, 等. 基于深度神经网络的 SAR 建筑目标三维重建方法 [J]. 中国科学: 信息科学, 2019, 49 (12): 1606–1625.
- [8] WEN C, ZHANG Y, FU X Y. Pixel2Mesh++: 3D mesh generation and refinement from Multi-View images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (2): 2166–2180.
- [9] 钟承来. 结合几何先验与显式辐射场的高效三维人头重建与应用 [D]. 合肥: 中国科学技术大学, 2024.
- [10] 陈 涛, 杨启亮, 陈 寅. 神经辐射场技术及应用综述 [J]. 计算机辅助设计与图形学学报, 2025, 37 (1): 51–74.
- [11] 余 瑛, 万贻前, 洪 瀚, 等. 基于符号距离函数的体渲染表面重建研究综述 [J]. 计算机应用研究, 2024, 41 (12): 3533–3542.
- [12] 曹梦晴. 基于深度学习的 3D 点云生成和表面重建算法研究 [D]. 西安: 西安电子科技大学, 2023.
- [13] HAGHIABI Z, MOKARI N, ABBASI ARAND B, et al. Building detection in SAR images based on fusion of classic and deep learning features [J]. International Journal of Remote Sensing, 2024, 45 (11): 3577–3596.
- [14] GE X, LIANG X, LI H, et al. An efficient synthetic aperture radar interference suppression method based on image domain regularization [J]. Electronics, 2025, 14 (5): 1054–1054.
- [15] JIN Z, XU Z, FENG H, et al. Reliable image dehazing by NeRF [J]. Optics Express, 2024, 32 (3): 3528–3550.
- [16] ZHOU Y, WANG H P, CHEN S Z. SAR automatic target recognition based on numerical scattering simulation and model-based matching [J]. Journal of Radars, 2015, 4 (6): 666–673.
- [17] DAI J W, JIN Y Q, XU F. Scattering simulation and reconstruction of a 3D complex target using Downward-Looking Step-Frequency radar [J]. Science China Information Sciences, 2011, 41 (8): 912–928.
- [18] 余燕娟. 基于图像熵确定 CLAHE 参数算法研究 [J]. 兰州文理学院学报 (自然科学版), 2024, 38 (6): 44–50.
- [19] 曲长文, 李 智, 周 强, 等. 基于改进 Frost 滤波的 SAR 图像斑噪抑制算法 [J]. 火力与指挥控制, 2018, 43 (11): 98–101.
- [20] 张志勇, 潘宁慧, 赵廷玉. 基于标准差加权高斯滤波函数及多方向 Sobel 算子的感兴趣区域中心子区图像期望清晰度值计算方法 [J]. 激光与光电子学进展, 2024, 61 (18): 395–403.
- [21] LI F, CHEN Y, SHI Q, et al. Improved Low-Light image feature matching algorithm based on the superglue net model [J]. Remote Sensing, 2025, 17 (5): 905–905.
- [22] 陈存宇, 陈爱喜, 戚晓秋, 等. 基于 MLP 神经网络优化改进的 BW 模型 [J]. 物理学报, 2025, 74 (1): 36–46.
- [23] 路仲伟, 陈 勇, 莫 云, 等. 基于 TRCSP 和 L2 范数的脑电通道选择方法 [J]. 电子测量技术, 2023, 46 (7): 94–102.
- [24] 邹梓越, 王永丽, 殷文艺, 等. 基于 Instant-NGP 平台的 NeRF 复现研究 [J]. 现代计算机, 2024, 30 (16): 44–50.
- [25] 王杰昌, 刘玉岭, 张 平, 等. 基于小批量随机梯度下降法的 SVM 训练隐私保护方案 [J]. 信息安全研究, 2024, 10 (10): 967–974.