文章编号:1674-2869(2024)03-0317-08

基于结构光相机与隐式表面重建的室内建模

王一川[1,2],王家奎[3],熊 伦[1,2],余子洋*[1,2]
1. 光学信息与模式识别湖北省重点实验室,湖北 武汉 430205;
2. 武汉工程大学光电信息与能源工程学院,湖北 武汉 430205;
3. 武汉唯理科技有限公司,湖北 武汉 430223

摘 要:深度相机在室内建模领域中以其快速、精准、鲁棒的特性而备受青睐。然而,在实际应用中,深度相机在获取精准深度信息时面临诸多挑战,如环境光干扰、点云噪声众多以及深度缺值。针对上述问题,提出了一种新型的结构光神经隐式表面(SL-NeuS)重建方法。通过结合广泛的三维重建技术,包括多目立体视觉(MVS)、神经网络、神经辐射场和神经隐式表面(NeuS),克服了现有技术在室内建模中的局限性。特别地,利用基于可微分循环优化启发设计的同步定位和建图(DROID-SLAM)算法精确获取相机外参,并结合单目深度估计和单目法向估计先验信息,有效地实现了精确的室内三维建模。实验结果显示,SL-NeuS重建方法在不同深度相机的性能分析中表现优越,成功降低了环境光影响和光学失真对建模精度的干扰。应用该方法对各种室内场景进行建模,误差可控制在极小范围内,实现了高精度的室内三维重建。本方法不仅提高了室内建模的准确性,还大幅减少了训练时间,为数字建筑和数字导航等领域提供了重要的技术支持。

关键词:室内建模;结构光;神经隐式表面;三维重建
中图分类号:TP391 **文献标识码**:A **DOI**:10.19843/j.cnki.CN42-1779/TQ.202311014

Indoor modeling based on structured-light camera and implicit surface reconstruction

***WANG Yichuan*[1,2], *WANG Jiakui*[3], *XIONG Lun*[1,2], *YU Ziyang**[1,2]**
1. Hubei Key Laboratory of Optical Information and Pattern Recognition, Wuhan 430205, China;
2. School of Optical Information and Energy Engineering, Wuhan Institute of Technology, Wuhan 430205, China;
3. Wuhan Veily Technology Co., Ltd, Wuhan 430223, China

Abstract: Depth cameras are highly valued in indoor modeling for their rapidness, precision, and robustness. Yet, they encounter challenges such as environmental light interference, point cloud noise, and incomplete data depth. To overcome these hurdles, an innovative structured light neural implicit surface (SL-NeuS) reconstruction method integrates a broad spectrum of 3D reconstruction technologies, including multi-view stereo (MVS), neural networks, neural radiance fields, and neural implicit surfaces (NeuS). This method leverages the differentiable recurrent optimization-inspired design simultaneous localization and mapping (DROID-SLAM) algorithm to accurately determine camera extrinsics. By incorporating monocular depth estimation and monocular normal estimation of prior information, it achieves precise indoor 3D modeling. Experimental findings demonstrate that the SL-NeuS reconstruction method excels in performance analysis across diverse depth cameras, effectively minimizing the impact of environmental light and optical distortions

收稿日期:2023-11-23
基金项目:国家自然科学基金(91963207)
作者简介:王一川,硕士研究生。Email:18955215829@163.com
***通信作者**:余子洋,博士,讲师。 Email:tommyu91@163.com
引文格式:王一川,王家奎,熊伦,等. 基于结构光相机与隐式表面重建的室内建模[J]. 武汉工程大学学报,2024,46(3):317-324.

on modeling accuracy. Applying this method to model various indoor environments keeps errors within a narrow range, ensuring high-precision indoor 3D reconstruction. In addition to enhancing indoor modeling precision, it significantly reduces training time, offering essential technical support for digital architecture and digital navigation fields.

Keywords: indoor modeling; structured light; neural implicit surfaces; three dimension reconstruction

在三维重建领域中,室内建模扮演着至关重要的角色。这一领域挑战重重,涉及的对象包括各种形态和结构的室内环境,条件瞬息万变,且在建模前无法完全确知详细信息。室内建模技术利用先进的计算机视觉和图形学技术,旨在立体、直观、准确地映射室内场景为三维模型。这不仅是数字建筑和虚拟导航领域的重要研究课题,也是实现虚拟现实、增强现实和室内导航等多领域应用的关键。

当前市场上深度相机[1]种类繁多,如何选择最适合室内建模的相机成为一个复杂且变化多端的问题。双目相机估算深度不准确,结构光相机和飞行时间法相机[2]测量深度准确,但是量程短,点云噪声比较严重,部分区域会有深度缺值,而神经网络可以弥补深度相机的缺陷。传统三维重建方法[3-4],如多目立体视觉(multiview stereo,MVS)[5],虽然在某些方面取得了进展,但在处理复杂室内环境时往往存在局限性。现有的多视图三维模型重建方法,包括基于几何匹配和神经网络的MVS,都面临着各自的技术瓶颈。例如,基于几何匹配的方法在弱纹理区域效果不佳,而基于神经网络的方法虽然可以处理这些区域,但在训练时受限于深度真值数据。

为了解决这些问题,提出了一种新型的室内建模方法——结构光神经隐式表面(structured light neural implicit surfaces,SL-NeuS)重建方法。SL-NeuS重建方法将多种技术整合在一起,包括深度相机数据获取、基于可微分循环优化启发设计的同步定位和建图(differentiable recurrent optimization-inspired design simultaneous localization and mapping,DROID-SLAM)算法[6]的相机外参获取以及先验信息的添加。不仅聚焦于深度相机的性能分析,还包括了SL-NeuS重建方法在室内建模中的应用。实验结果表明,SL-NeuS重建方法在室内建模中的表现优于现有方法,特别是在结合双目结构光相机时,能够在更短的时间内实现更高质量的重建效果,为未来利用深度相机进行室内建模的学者提供了有价值的参考方案。通过结构光相机与神经隐式表面(neural implicit surfaces,NeuS)重建的融合,提高了整个系统在复杂室内环境中的鲁棒性和精确度,使室内建模技术在数字建筑和数字导航等领域的应用更加可靠和高效。

1 深度相机性能分析与室内重建方法

选用市面上较为先进的3种深度相机,对其获取深度信息的精准度与抗环境光干扰能力等性能做出具体分析。NeuS重建方法[7]已成为多视图三维重建领域的研究热点。与传统的多视图立体方法[8-9]相比,NeuS重建方法倾向于产生更平滑、更完整的重建结果,这归因于神经网络的归纳平滑性和建模中使用的法向约束损失。由于深度相机最重要的是其获取精准深度信息的能力,所以研究将围绕着三原色深度(red green blue depth,RGB-D)数据集展开,为此选用在RGB-D数据集上表现较为出色的基于深度测量监督的神经三原色深度(neural red green blue depth,Neural RGB-D)方法[10],以其重建效果对深度相机的性能做出评估。此外还选用了基于单目深度估计先验和单目法向估计先验的单目符号距离场(monocular signed distance field,MonoSDF)方法[11],该方法添加了几何约束,大幅提高了重建效果。最终,提出SL-NeuS重建方法,该方法以NeuS[12]作为基线,使用RGB-D数据集与单目先验信息作为输入,旨在融合结构光相机与NeuS重建,并且在SL-NeuS重建方法中添加几何约束等多种先验信息,提高重建效果。

1.1 深度相机性能分析

双目结构光相机利用内置的结构光投射器向目标场景投射编码的红外光斑,2个红外相机同时捕捉结构光照射形成的红外图像,通过匹配对应点和计算位移,实现对目标表面深度信息的高精度获取。单目结构光相机通过投射特定的红外结构光斑模式至被测物体表面,利用单个红外摄像机捕捉反射回的光斑图像,识别光斑的编号再进行三角化计算来获取三维深度信息。双目结构光相机的本质是利用结构光投射器来丰富物体表面的纹理,提升双目立体匹配的精度。单目结构光

相机则依赖于图像识别，如果它不能精准识别光斑对应的编号，就不能精准地获得深度。一般来说，双目结构光相机具有更高的测量精度，单目结构光相机具有一定的成本优势。双目相机是一种模拟人眼立体视觉的系统，由两个三原色（red green blue，RGB）相机组成，它们之间的平行距离称为基线。通过捕捉同一场景的两个不同视角的图像，双目相机可以利用点到点匹配计算物体的深度信息，从而实现三维视觉。

本文引入的双目结构光相机（RealSense D455）、单目结构光相机（Orbbec Astra+）和双目相机（ZED 2i）的硬件信息如表1所示，其中，H表示水平方向视角，V表示垂直方向视角，D表示对角线视角。

1.2　SL-NeuS重建方法

首先通过3种不同的深度相机获取RGB-D数据，通过相机各自的软件开发工具包读取相机的内参，然后通过DROID-SLAM算法获取相机外参，并在此基础上添加图像的单目先验信息[13]，包含单目法向先验信息与单目深度先验信息，此处的单目先验信息由Omnidata神经网络模型[14]提供。SL-NeuS将场景几何图形表示为符号距离场（signed distance field，SDF），在给定场景姿态图像的情况下，SL-NeuS的目标是重建场景表面，该表面由一个神经隐式SDF的零水平集S表示。

$$S = \left\{ x \in R_3 | f(x) = 0 \right\} \tag{1}$$

其中，x是空间中任意一个3D点，R_3代表三维空间，$f(x)$是SDF函数。

为学习神经网络的权值，SL-NeuS用隐式SDF进行图像渲染，并最小化渲染图像与输入图像之间的差异。通过引入概率密度函数$\varphi_s(x)$，实现通过体积渲染[15]来训练SDF网络。$\varphi_s(x)$的形式如下：

$$\phi_s(x) = s\exp(-sx)/[1+\exp(-sx)]^2 \tag{2}$$

其中，$\varphi_s(x)$是sigmoid函数$F_s(x)=[1+\exp(-sx)]^{-1}$的导数，$s$为一个可学习的参数。

给定一个像素，以该像素为原点发射一条光线，将其表示为：

$$\left\{ p(t) = o + t\boldsymbol{v} | t \geqslant 0 \right\} \tag{3}$$

其中o是相机中心，$\boldsymbol{v}$是该像素对应的光线的单位方向向量，t为该像素深度值。

沿光线积累的颜色定义为：

$$C(o,\boldsymbol{v}) = \int_0^{+\infty} w(t)c\left[p(t),\boldsymbol{v}\right]\mathrm{d}t \tag{4}$$

其中，$C(o,\boldsymbol{v})$是该像素输出的颜色，$p(t)$是像素点，$w(t)$是点$p(t)$的积分权重，$c[p(t),\boldsymbol{v}]$是沿着光线方向的点$p(t)$处的颜色。

SL-NeuS引用了NeuS中定义$w(t)$的方式，定义不透明的密度函数$\rho(t)$，该函数对应于标准体渲染中的体积密度$\sigma(t)$。

$$\sigma(t) = \varphi_s\left(f\left(p(t)\right)\right) \tag{5}$$

$$\rho(t) = \frac{-\dfrac{\mathrm{d}F_s}{\mathrm{d}t}\left(f\left(p(t)\right)\right)}{F_s\left(f\left(p(t)\right)\right)} \tag{6}$$

通过下述方程式计算新的权重函数$w(t)$和沿光线的累计投射率$T(t)$：

$$w(t) = T(t)\rho(t),\ T(t) = \exp\left(-\int_0^t \rho(u)\mathrm{d}u\right) \tag{7}$$

SL-NeuS使用RGB损失函数L_{RGB}，用于优化场景表示：

$$L_{\mathrm{RGB}} = \sum_{r \in R} ||\hat{C}(r) - C(r)|| \tag{8}$$

其中，$\hat{C}(r)$是神经网络预测颜色，$C(r)$是相机拍摄的颜色，r表示当前光线，R表示光线集合。

此外，SL-NeuS在采样点上添加程函损失函数L_{eikl}，用于正则化三维空间中的SDF值：

$$L_{\mathrm{eikl}} = \sum_{x \in X} (||\nabla f(x)||_2 - 1)^2 \tag{9}$$

其中，X是一组均匀采样点和近表面点。

由于添加了几何约束，所以SL-NeuS引入基

表1　深度相机硬件信息表

Tab. 1　Depth camera hardware information table

相机品牌	RealSense D455	Orbbec Astra+	ZED 2i
深度视场角	87°(H)×58°(V)	57°(H)×45.5°(V)×69°(D)	无
彩色视场角	90°(H)×65°(V)	69.7°(H)×42.8°(V)×77.2°(D)	110°(H)×70°(V)×120°(D)
拍摄范围	0.6~6.0 m	0.6~8.0 m	0.2~20.0 m
拍摄精度	(4.000±0.008) m	(1.000±0.003) m	(3.000±0.001) m
彩色最大分辨率	1 280×800	1 920×1 080	2 208×1 242
深度最大分辨率	1 280×720	1 280×1 024	2 208×1 242
使用场景	室内	室内	室内，室外
软件复杂度	低	中	高

于单目深度估计先验和单目法向估计先验的MonoSDF方法中的单目深度一致性损失与单目法向一致性损失。使用预训练的单目深度网络预测的深度与法向作为每个图像的附加约束。对于深度损失,在估算深度损失函数$L_{\text{estd-d}}$基础上添加测量深度损失函数$L_{\text{mead-d}}$:

$$L_{\text{estd-d}}=\sum_{r\in R}||(w\hat{D}(r)+q)-\bar{D}(r)||^2 \quad (10)$$

$$L_{\text{mead-d}}=\sum_{r\in R}||\hat{D}(r)-D(r)||^2 \quad (11)$$

其中$\hat{D}$是神经网络预测深度,$\bar{D}$是单目深度估计深度,D是利用深度相机测量的深度,w、q分别是用于对齐$\hat{D}$和$\bar{D}$的比例和位移。

使用两种深度损失函数是由于测量的深度在平均意义上比较准确,但是存在比较严重的噪声和缺值问题,和单目深度估计有一定的互补性。单目深度估计虽然总体上不准确,但是其细节丰富度和完整性远强于深度相机测量的深度,可以补足测量深度作为监督信号的不足。在未来还可以探索一些自适应的方案,例如有些区域可以以测量深度为主,有些区域可以以估算深度为主。

采用单目法向一致性损失函数L_{n}将体积渲染出的法向$\hat{N}$和单目估计法向$\bar{N}$转换为相同角度的坐标系。

$$L_{\text{n}}=\sum_{r\in R}||\hat{N}(r)-\bar{N}(r)||+||1-\hat{N}(r)^{\text{T}}\bar{N}(r)|| \quad (12)$$

SL-NeuS重建方法使用体积渲染[15]来训练SDF网络,并通过密度函数进行优化。在此过程中,将输入图像用作监督信号,并结合单目先验线索作为约束以引导和优化重建过程。之后将如下损失函数最小化:

$$L=L_{\text{RGB}}+0.1L_{\text{eikl}}+0.1L_{\text{estd-d}}+0.3L_{\text{mead-d}}+0.05L_{\text{n}} \quad (13)$$

使得最终得到的SDF零水平集可以准确地代表重建表面的几何信息。SL-NeuS重建方法融合了双目结构光相机、体积渲染、深度学习网络以及先验信息,以实现对三维场景的精确重建。

2 实验分析

为比较基于高精度结构光相机与神经辐射场的室内建模性能,在Ubuntu20.04系统上使用pytorch框架构建实验,采用NVIDIA RTX 3090建模,对象为工作建筑内的室内场景,并使用3种不同的深度相机进行数据采集。同时,采用两种不同的方法获取相机外参构建数据集。再将这些数据集分别应用于3种方法中,得到重建模型并进行比较分析。

2.1 数据采集

实验将使用RealSense D455相机、Orbbec Astra+相机和ZED 2i相机进行初始数据采集。RealSense D455相机能够提供高分辨率的深度图像,并且具有宽视场角,可以捕捉更广阔的场景,适用于室内场景应用,且该相机采用USB接口,易于与各种计算机和嵌入式系统集成。Orbbec Astra+相机能够提供高分辨率的深度图像,捕捉场景中物体的精确距离信息,相比于RealSense D455相机,Orbbec Astra+相机获取的深度图的噪声更少,并且Orbbec Astra+相机同样适用于各种计算机系统且具有用户友好的接口和丰富的开发工具。此外,Orbbec Astra+相机价格更加实惠。ZED 2i相机采用立体视觉技术,具有双目摄像头,能够捕捉场景中的深度信息,并且可以提供高分辨率的深度图像,捕捉物体的精确距离和轮廓信息,适用于室内场景和室外场景使用。表1中的相机性能参数均来源于官方文档,实测中ZED 2i相机的拍摄精度远不如RealSense D455相机。

相机外参是数据集中至关重要的组成部分。为获取准确的相机外参,引入了COLMAP算法[16]与DROID-SLAM算法。COLMAP算法在处理RGB数据集时表现良好,但是在处理RGB-D数据集时存在缺陷,其中最突出的问题是COLMAP算法的输入仅限于RGB图像,导致其估计的相机外参与数据集的深度图之间存在坐标系尺寸差异,使COLMAP算法输出的基于RGB-D数据集的相机外参不准确,为此引入DROID-SLAM算法。DROID-SLAM算法用于单目、双目和RGB-D相机的深度视觉同步定位与建图(simultaneous localization and mapping,SLAM)系统,该系统获取数据集相机外参的流程如下:首先,将数据集和相机内参作为输入;其次,SLAM系统初始化确定相机的初始位置和姿态;接着,系统从输入图像中提取特征点,并在不同的图像帧中进行匹配以实现位置跟踪,通过跟踪特征点,SLAM系统估计相机在空间中的轨迹,在此过程中,SLAM系统构建三维地图,包括特征点的位置、地标或密集的点云数据,以及相机的轨迹信息(如图1所示)。SLAM系统使用光束法平差(bundle adjustment,BA)算法[17]来提高地图的精度,并优化相机的轨迹;最终,SLAM系统会提供估计的相机外参。与COLMAP算法不同,DROID-SLAM算法使用真实的相机内参,因此其估计的相机外参更加准确。此外,DROID-SLAM算法采用稀疏重建方法,可

以更好地控制重建的精度。在运行时间方面，DROID-SLAM算法相较于COLMAP算法，在实验中运行时间更短，大约节省了8倍的时间。

为了获取单目先验信息，引入了Omnidata神经网络模型。为此，配置Omnidata神经网络模型的虚拟环境，并将对应的原始数据放入Omnidata神经网络模型中，以估计深度值与法向值。

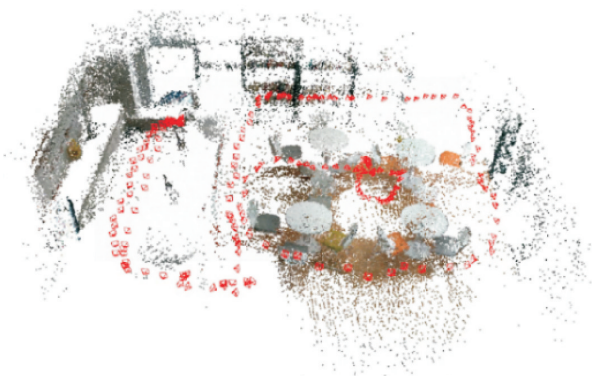

图1　DROID-SLAM三维地图

Fig. 1　DROID-SLAM three dimension map

2.2　数据结果分析

2.2.1　重建效果分析　基于深度测量监督的Neural RGB-D方法：基于pytorch框架，使用自适应矩估计算法（adaptive moment estimation，Adam）优化器[18]，学习率为5×10^{-4}，在1×10^{5}次迭代中，指数学习速率衰减为10^{-1}，每个批次在每个训练迭代中使用的光线数为2 048。根据场景大小，在基于单个NVIDIA RTX 3090的硬件条件下，训练时间为3~5 h。

图2（a-c）分别代表3个相机拍摄的数据集基于深度测量监督的Neural RGB-D方法进行训练后的重建结果。如图2（a）所示，由于RealSense D455相机获取的深度图效果最佳，其深度图与彩色图之间的对齐效果最好，且该相机对环境光适应能力较强，因此RealSense D455相机拍摄的数据集重建效果最佳。Orbbec Astra+相机获取的深度图数据虽然较为精确，但彩色图与深度图之间存在偏差，所以由图2（b）可以看出其重建效果一般。ZED 2i相机并非结构光相机，因此获取的深度信息存在较大波动，并且在使用过程中ZED 2i相机的抗环境光干扰能力最差，所以图2（c）重建效果最差。由本组实验可以看出双目结构光相机在室内建模方法中表现最为优秀。

基于单目深度估计先验和单目法向估计先验的MonoSDF方法同样使用pytorch框架和Adam优化器，学习率为5×10^{-4}，迭代次数为2×10^{5}，每个批次在每个训练迭代中使用的光线数为2 048。根据场景大小，在基于单个NVIDIA RTX 3090的硬件条件下，训练时间为13~15 h。

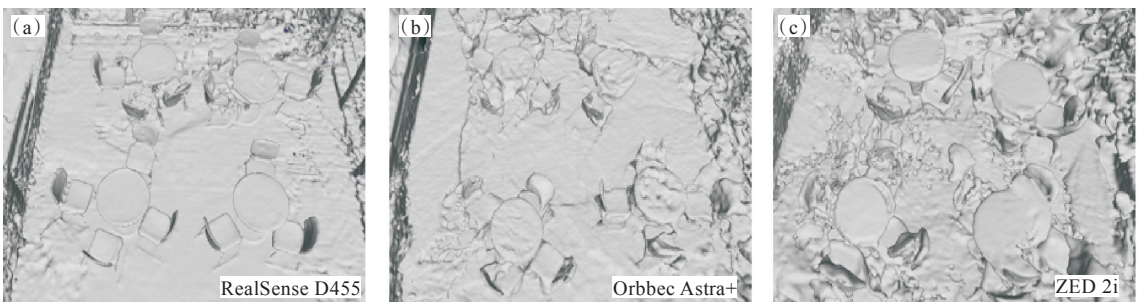

图2　不同相机数据集下基于深度测量监督的Neural RGB-D方法的重建结果

Fig. 2　Reconstruction results of the Neural RGB-D method based on depth measurement supervision under different camera datasets

图3（a-c）分别代表3个相机拍摄的数据集基于单目深度估计先验和单目法向估计先验的MonoSDF方法进行训练后的重建结果。通过对比图3与图2中的6幅重建结果图，可以发现，加入单目先验信息后，有效解决了噪点众多问题，显著提升了重建效果。这是因为光度一致性线索表面重建方法和单目先验网络使用的识别线索是互补的：在弱纹理区域（如墙壁中），表面法向能够可靠地预测结构化的三维场景布局，而光度一致性线索允许在纹理区域建立全局精确的三维几何图形，法向和深度线索只提供局部几何信息。然而实现高精度重建也带来了更高的时间成本，与基于测量深度监督的Neural RGB-D方法相比，增加了3倍训练时间，这是由于基于单目深度估计先验和单目法向估计先验的MonoSDF方法具有更加复杂的神经网络结构，并且其训练迭代的次数也比Neural RGB-D方法多1倍。

SL-NeuS重建方法同样使用pytorch框架和Adam优化器，学习率为5×10^{-4}，迭代次数为1×10^{5}，每个批次在每个训练迭代中使用的光线数为

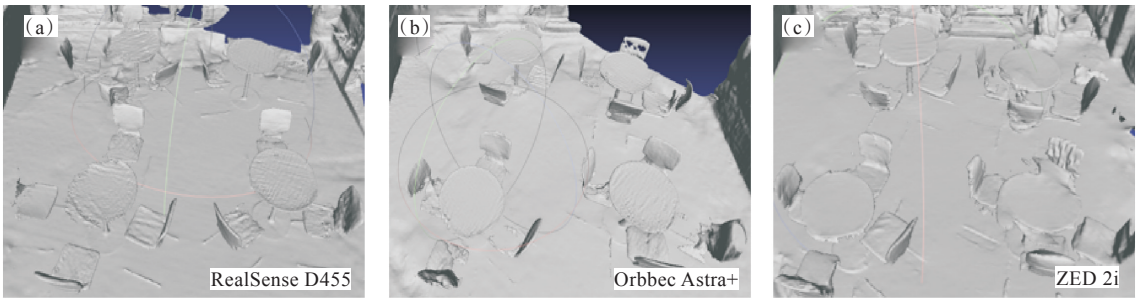

图3 不同相机数据集下基于单目深度估计先验和单目法向估计先验的MonoSDF方法重建结果

Fig. 3 Reconstruction results of the MonoSDF method based on monocular depth estimation prior and monocular normal estimation prior under different camera datasets

2 048。根据场景大小,在基于单个NVIDIA RTX 3090的硬件条件下,训练时间为4~6 h。SL-NeuS使用损失函数$L1$作为RGB损失来监督每条光线的体积渲染颜色,使用程函损失函数规范SDF正确表示三维点与其距离最近的曲面间的距离,使用单目深度一致性损失函数与单目法向一致性损失函数为图像添加几何约束,并使用测量深度作为额外的几何约束。该方法采用2个多层感知器(multi-layer perceptron,MLP)[19]分别编码SDF和颜色:SDF由1个MLP建模,该MLP由8个隐藏层组成,维度256,激活函数为Softplus;颜色由1个MLP建模,该MLP由4个隐藏层组成,维度256。颜色的MLP需要接收SDF值作为输入信息。

图4(a-c)分别代表3个相机拍摄的数据集在SL-NeuS进行训练后的重建结果。SL-NeuS以NeuS为基线,添加几何约束大幅提高重建效果,有效缓解基于测量深度监督的Neural RGB-D方法中重建结果噪点众多的问题。通过对比图2~图4可以看出,SL-NeuS提高了重建的精确度,降低了引入几何约束后产生的时间成本。

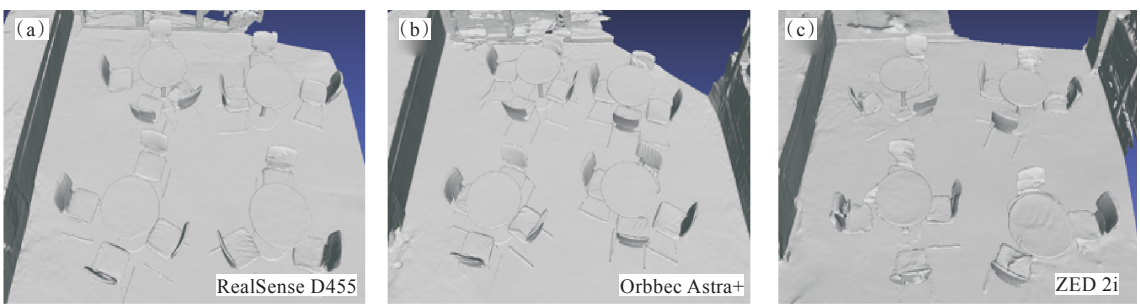

图4 不同相机数据集下SL-NeuS重建方法重建结果

Fig. 4 Reconstruction results of the SL-NeuS reconstruction method under different camera datasets

2.2.2 *渲染效果分析* 图5与图6分别是基于3种相机数据集使用3种方法的彩色与深度渲染效果对比图。如图5所示,3种相机数据在同种重建方法的彩色渲染过程中,渲染效果差别不大,但基于单目深度估计先验和单目法向估计先验的MonoSDF方法的彩色渲染效果与其他两种重建方法相比,渲染效果较差。由图6可知,基于测量深度监督的Neural RGB-D方法中,RealSense D455相机获取的数据深度渲染效果最佳。

峰值信噪比(peak signal-to-noise ratio,PSNR)是一种常用的图像质量评价指标,用于评估渲染后图像与原始图像之间的差异,PSNR越高,图像之间的差异越小。如表2所示,基于单目深度估计先验和单目法向估计先验的MonoSDF方法在渲染效果方面存在不足。而SL-NeuS重建方法在渲染效果方面与基于单目深度估计先验和单目法向估计先验的MonoSDF方法相比有一定提高,且通过图5~图6也可以发现,SL-NeuS重建方法的渲染效果有明显提升。

综上所述,基于深度测量监督的Neural RGB-D方法的训练速度快,但对相机要求较高,如图6所示,深度渲染效果很大程度上取决于相机性能。基于单目深度估计先验和单目法向估计先验的MonoSDF方法重建效果得到提升,但训练速度提过慢,渲染效果有待加强。SL-NeuS重建方法提升了重建精度、训练速度和渲染效果,获得最佳重建效果。

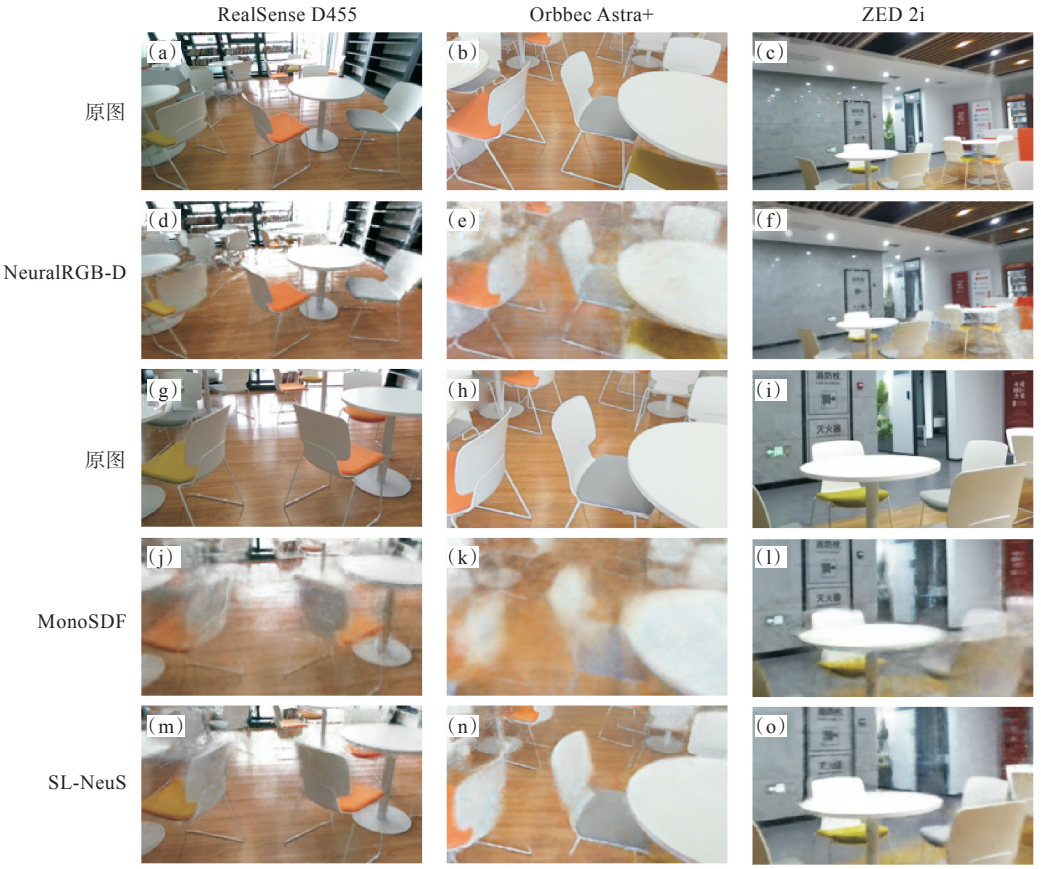

图 5　基于 3 种相机数据集使用 3 种方法的彩色渲染效果对比

Fig. 5　Comparison of color rendering effects using three methods based on three camera datasets

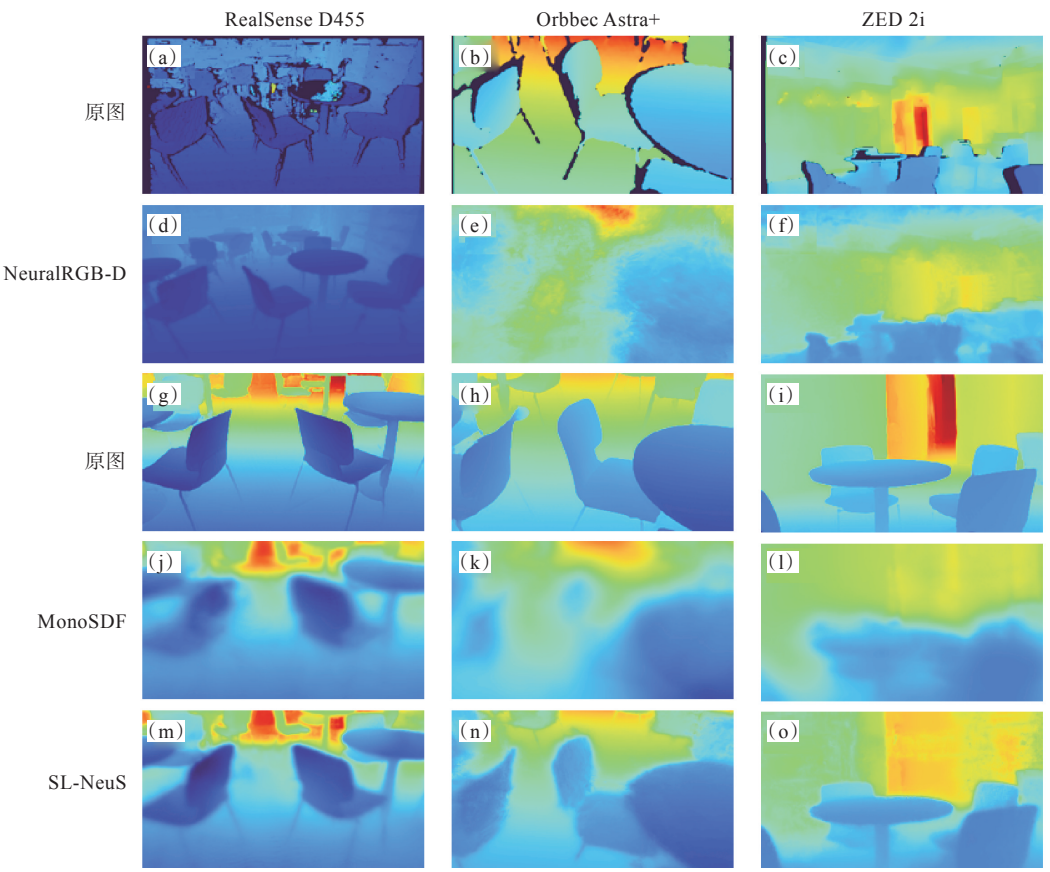

图 6　基于 3 种相机数据集使用 3 种方法的深度渲染效果对比

Fig. 6　Comparison of depth rendering effects using three methods based on three camera datasets

表2 彩色图与深度图渲染过程峰值信噪比

Tab. 2 PSNR during rendering of color and depth maps

重建方法	PSNR / dB					
	彩色图			深度图		
	RealSense D455	Orbbec Astra+	ZED 2i	RealSense D455	Orbbec Astra+	ZED 2i
Neural RGB-D	29.42	29.37	29.45	32.42	28.82	30.29
MonoSDF	28.85	29.30	29.32	29.33	29.34	28.57
SL-NeuS	29.17	30.07	31.26	29.89	30.70	29.65

3 结 论

基于以上研究与实验，发现在处理RGB-D数据集时，使用DROID-SLAM算法获取相机外参，结果更加精确而且获取相机外参时间更短。此外，使用结构光相机拍摄的真实深度图可以确保重建的模型尺寸与真实尺寸保持一致，有助于工程、建筑、医学等领域中的精确测量和分析。在基于测量深度监督的Neural RGB-D方法中，结构光相机与非结构光相机在深度信息的获取方面差异显著。结构光相机获取的深度信息具有良好的表面，非结构光相机获取的深度信息在物体表面呈凹凸不平状，这导致使用非结构光相机构建的数据集进行室内重建后，获得的点云中含有巨大噪声，难以获取平滑的物体表面。在基于单目深度估计先验和单目法向估计先验的MonoSDF方法中，加入单目深度估计先验和单目法向估计先验显著提高了重建结果的精度，但由于复杂的神经网络结构导致训练时间较长。SL-NeuS重建方法以NeuS作为基线，使用RGB-D数据集与单目先验信息作为输入，融合了双目结构光相机与NeuS重建，添加单目深度估计先验和单目法向估计先验提高重建效果，并使用测量深度作为额外的几何约束。通过两个MLP分别对SDF和颜色进行编码，既提升了重建结果的精度又降低了训练所需的时间成本。

参考文献

[1] 马兆敏，郭家祥. 果实定位的机器视觉技术研究现状[J]. 南方农机，2024，55(2)：6-9.

[2] 肖天行. 基于2D/3D信息融合的易拉罐外观缺陷识别方法研究[D]. 无锡：江南大学，2023.

[3] 杨航，陈瑞，安仕鹏，等. 深度学习背景下的图像三维重建技术进展综述[J]. 中国图象图形学报，2023，28(8)：2396-2409.

[4] 王维，高凤娇，刘彤军. 基于NeuS的大场景建筑物构建方法研究[J]. 黑龙江科学，2023，14(10)：51-55.

[5] YARIV L，KASTEN Y，MORAN D，et al. Multiview neural surface reconstruction by disentangling geometry and appearance [C]//Advances in Neural Information Processing Systems 33 (NeurIPS 2020). New York：Curran Associates，Inc. ，2020：2492-2502.

[6] 付雄. 无人车视觉定位及路径规划算法研究[D]. 南京：南京信息工程大学，2023.

[7] 张沛全，许威威. 哈希编码优化的IRON逆渲染模型：重建几何与材质[J]. 浙江大学学报(理学版)，2023，50(6)：754-760，769.

[8] 冯亚娟. 基于深度学习的MVS三维重建算法研究[D]. 太原：山西大学，2023.

[9] 李熙. 基于深度学习的图像三维重建算法研究[D]. 成都：电子科技大学，2022.

[10] AZINOVIĆ D，MARTIN-BRUALLA R，GOLDMAN D B，et al. Neural RGB-D surface reconstruction[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway：IEEE，2022：6280-6291.

[11] YU Z H，PENG S Y，NIEMEYER M，et al. MonoSDF：exploring monocular geometric cues for neural implicit surface reconstruction [C]//Advances in Neural Information Processing Systems 35 (NeurIPS 2022).New York：Curran Associates，Inc.，2022：25018-25032.

[12] WANG P，LIU L J，LIU Y，et al. NeuS：learning neural implicit surfaces by volume rendering for multi-view reconstruction [C]//Advances in Neural Information Processing Systems 34 (NeurIPS 2021). New York：Curran Associates，Inc.，2021：27171-27183.

[13] 王建森. 基于多任务联合学习的无监督图像逆渲染方法研究[D]. 天津：天津大学，2021.

[14] EFTEKHAR A，SAX A，MALIK J，et al. Omnidata：a scalable pipeline for making multi-task mid-level vision datasets from 3D scans[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway：IEEE，2021：10766-10776.

[15] 李森涛，万旺根，王旭智，等. 基于体积渲染和几何一致性的多视图三维重建[J]. 工业控制计算机，2023，36(10)：98-100.

[16] 胡荣林，付浩志，何旭琴，等. 基于胶囊卷积网络的多视图三维重建[J]. 南京师范大学学报(工程技术版)，2023，23(1)：46-55，92.

[17] 辛伟. 基于公共线面特征的异源点云配准方法[D]. 重庆：重庆交通大学，2023.

[18] 唐林瑞泽，白仲鑫，张晓雷. 基于受试者工作特征pAUC优化的人脸识别系统[J]. 太赫兹科学与电子信息学报，2023，21(9)：1150-1155.

[19] 袁浩瀚. 基于深度学习的图像编码与传输一体化技术研究[D]. 成都：电子科技大学，2023.

本文编辑：霍淑利