

基于光线碰撞检测的轻量化实时三维重建系统

陈昌川¹, 丁锡烽¹, 王亦何², 孙霞³, 乔飞^{2*}

(1. 重庆邮电大学通信与信息工程学院, 重庆 400065; 2. 清华大学电子工程系, 北京 100084;
3. 重庆工程学院电子信息学院, 重庆 400056)

摘要: 针对小型机器人在复杂环境中进行实时定位与建图时, 存在机载端 CPU (Central Processing Unit) 计算资源不足, 建图精度差、探索效率低的问题. 本文提出一种基于同时定位与建图 (Simultaneous Localization and Mapping, SLAM)、截断符号距离函数 (Truncated Signed Distance Function, TSDF) 实时的三维重建方法. 该方法基于深度相机或双目相机获取重建目标及场景的 RGB (Red Green Blue) 图和深度图, 同时基于 ORB_SLAM2 获取位姿信息; 采用基于特征点云数据的表面重建算法 TSDF 与深度图相结合, 实现一种实时三维场景重建; 为了降低三维重建模型与真实场景的误差, 提出一种采用光线碰撞检测融合特征点的方法, 并结合优化策略减小光线投影距离与体素到物体表面距离的误差. 通过优化后的 TSDF 值, 保证了重建场景的完整性. 在 ASL (Autonomous Systems Lab) 开源数据集上, 相比于 Voxblox、Voxfield 和 VDBblox, 该系统三维重建模型的均方根误差分别下降了 55.6%、47.11%、21.7%, 相比于 Voxblox、Voxfield, 系统地图更新速率分别提升了 9.7% 和 12.9%. 最后, 将该系统用于室内场景实验, 地图平均每帧更新速率为 7.35 ms/帧, 验证了所提系统的可行性和有效性.

关键词: 三维重建; TSDF; 同时定位与地图构建; ORB_SLAM2; 光线碰撞检测

基金项目: 新疆维吾尔自治区重点研发计划 (No. 2022B01008-3); 国家自然科学基金 (No. 92164203, No. 62334006); 重庆市教委重点项目 (No. KJZD-K202301901); 重庆市科委项目 (NO. 2023NSCQ-MSX4308); 陕西省重点研发计划 (No. 2023-YBNY-222)

中图分类号: TP242; TP391 文献标识码: A 文章编号: 0372-2112(2024)12-4074-13

电子学报 URL: <http://www.ejournal.org.cn> DOI: 10.12263/DZXB.20230983

Lightweight Real-Time 3D Reconstruction System Based on Light Collision Detection

CHEN Chang-chuan¹, DING Xi-feng¹, WANG Yi-he², SUN Xia³, QIAO Fei^{2*}

(1. School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2. Department of Electronic Engineering, Tsinghua University, Beijing 100084, China;
3. School of Electronic Information, Chongqing Institute of Engineering, Chongqing 400056, China)

Abstract: When small robots perform real-time positioning and mapping in complex environments, there are problems such as insufficient computing resources of the onboard CPU (Central Processing Unit), poor mapping accuracy, and low exploration efficiency. This paper proposes a real-time 3D reconstruction method based on simultaneous localization and mapping (SLAM) and truncated signed distance function (TSDF). This method obtains the RGB (Red Green Blue) image and depth image of the reconstructed target and scene based on a depth camera or a binocular camera, and obtains the pose information based on ORB_SLAM2. The surface reconstruction algorithm TSDF based on feature point cloud data is combined with the depth map to achieve a real-time 3D scene reconstruction. In order to reduce the error between the 3D reconstructed model and the real scene, a method of fusing feature points using light collision detection is proposed, and the error between the light projection distance and the distance from the voxel to the object surface is reduced by combining the optimization strategy. The integrity of the reconstructed scene is guaranteed by the optimized TSDF value. In ASL (Autonomous Systems Lab) the system is used in an open source dataset of the 3D reconstruction model of the proposed system. Compared with Voxblox, Voxfield and VDBblox, the root mean square error of the system's 3D reconstruction model is re-

duced by 55.6%, 47.11% and 21.7% respectively. Compared with Voxelbox and Voxfield, the system map update rate is increased by 9.7% and 12.9% respectively. Finally, the system is used in indoor scene experiments, and the average map update rate is 7.35 ms/frame, which verifies the feasibility and effectiveness of the proposed system.

Key words: 3D reconstruction; TSDF; simultaneous localization and mapping; ORB_SLAM2; ray collision detection

Foundation Item(s): Key Research and Development Program of Xinjiang Uygur Autonomous Region (No.2022B01008-3); National Natural Science Foundation of China (No.92164203, No.62334006); Science and Technology Research Project of Chongqing Education Commission under Grant (No. KJZD-K202301901); General Projects of Chongqing Science and Technology Commission under Grant (No.2023NSCQ-MSX4308); Key Research and Development Program of Shaanxi Province (No.2023-YBNY-222)

1 引言

随着计算机视觉技术快速发展,三维模型越来越地被应用于各种领域,例如医疗、电影、自动驾驶、智能家居等^[1-3],其中三维重建是构建三维模型的重要来源.三维重建是通过传感器对真实世界进行扫描从而构建三维模型.三维重建方式有主动式扫描重建和被动式扫描重建,这些方法通过采集图像特征点进行离线重建.同时定位与建图(Simultaneous Localization And Mapping, SLAM)是智能移动机器人实时定位与构建地图同步完成的系统算法^[4,5],是机器人实现自主探索的基础.为确保机器人在复杂环境中完成定位建图需求,实时构建高精度三维重建地图至关重要.

为了保证机器人在复杂的未知环境下的安全,需要构建高精度的三维重建地图.由于SLAM对地图的需求不一样,需要构建的地图也不一样.稀疏地图通常只能用于简单的定位,稠密地图则用于导航、避障等.点云拼接可以得到稠密点云地图,根据地图外观可以分类为体素(Voxel)、网格(Mesh)地图. Newcombe^[6]基于Mesh对物体表面进行TSDF表示,将新测量到的体素直接融合到TSDF地图中,但是没有回环检测消除累计误差,并且体素网格大小随内存按立方比例增长,无法构建大场景环境^[7]. Whelan等人^[8]通过相机位姿虚拟移动体素网格,针对移出重建体积的部分进行三角化,但依旧没有解决闭环问题. Zeng等人^[9]提出的八叉树地图可以显著降低体素网格内存. Endres等人^[10]利用RANSAC估计匹配特征点之间的转换,采用非线性优化位姿,同时引入波束的环境测量模型,从而可以处理复杂的场景. Zhang等人^[11]提出了一个三维点云重建方法FlowFusion,该算法使用光流残差与深度学习网络结合起来描述点云中的动态物体语义,可以同时完成场景中动态和静态物体分割、相机运动估计以及静态背景点云重建. Voxelbox利用传感器数据建立TSDF,同时过滤噪声,构建三维重建地图^[12]. Voxfield通过改进光线投影距离,对TSDF值进行优化从而构建精度更高的三维重建地图^[13]. Voxgraph在Voxelbox的基础上添加子图约束条件,并在系统后端进行全局地图优化,能够在

CPU上实时运行^[14]. Kimera-Multi将采用神经网络构建语义子图,同时对位姿进行全局优化从而实现分布式多机协同语义建图,并通过回环检测降低融合地图后的误差^[15]. Voxelbox-Multi-Agent计算固定的相机位姿,通过对稀疏点云添加前置滤波器进行压缩,将不同子端的位姿融合后建立三维重建地图^[16]. Coxgraph使用多个机器人作为子端,每个子端搭载视觉里程计,将采集的地图信息发送到服务端构建局部地图,并通过子图位姿约束和视觉里程计约束实现全局一致的体素级建图^[17]. Coxgraph++在Coxgraph的基础上改进了子图的优化方法,并提出双体素网格层存储地图信息,构建出带有语义信息的三维重建地图^[18]. VDBbloss在Voxfield的基础上添加权重更新因子,并通过光线非投影的方式进行三维重建,能够有效提高系统建图精度^[19]. 但上述文献只能以离线的方式构建三维重建地图,不能在现实场景中实时增量式构建三维重建地图.

实时构建和更新地图能够允许机器人完成更多的任务,为了在真实场景下实时构建三维重建模型, Klein等人^[20]提出的并行跟踪与建图(Parallel Tracking And Mapping, PTAM)利用FAST角点检测法对小空间增强现实进行地图构建. Song等人^[21]将快速定向与旋转描述子(Oriented FAST and Rotated BRIEF, ORB)特征用于虚拟现实用户视点跟踪的单目SLAM中. Mur-Artal提出ORB-SLAM2^[22],该算法基于ORB特征点提取与描述,实现定位与地图构建的算法,该方法可快速处理输入数据、跟踪特征点以及构建稀疏地图. Whelan等人^[23]提出的Elastic Fusion通过密集帧的相机跟踪方式和融合窗口曲面以及非刚性表面变形的形式,实现模型细化,以增量在线的方式实现地图构建. Dai等人^[24]提出的Bundle Fusion消除了跟踪相机位姿变换时对于时间的依赖,不断定位到全局优化的帧. 该算法可以从跟踪失败后实现重新定位,并实时重新估计3D模型. Sun等人^[25]提出了基于单目视频的实时三维重建框架NeuralRecon,基于学习的TSDF融合模块基于选通递归单元,该设计使得网络能够在连续重构三维曲面时,获得曲面的局部光滑性先验和全局形状先验. 虽然上述研究在真实场景下支持机器人实时定位与建图,但在

实时增量式构建三维重建地图上依然存在建图精度低、系统计算量大等问题。

确保小型机器人在复杂环境下进行实时探索的主要挑战是构建高精度的TSDF地图的同时高效地维护和更新TSDF地图,针对构建TSDF地图精度低和小型机器人计算资源受限导致实时建图效率低的问题,本文创新性地构建了1个全新的在线三维重建框架,该框架采用RealSensed435i相机获取重建物体的RGB图像和深度图像,为了满足系统的实时性,本文使用ORB-SLAM2作为前端实时采集图像特征点和计算位姿,ORB-SLAM2作为一种实时定位与建图算法,可快速高效提取图像特征点并计算相机位姿,将构建TSDF地图算法与深度图结合,实时进行三维重建;为了在不降低精度的情况下加速TSDF地图的构建,本文提出一种光线碰撞检测融合策略减少光线遍历体素数量,实现了在线三维模型重建.为了构建高精度三维重建模型,本文提出基于特征点法向量优化体素到物体表面真实距离,得到更精确的TSDF值,实现了一种实时高精度场景三维重建.相比于Voxblox、Voxfield在ASL开源数据集上TSDF地图平均更新速率分别提升了9.7%和12.9%,在ASL开源数据集上三维重建地图均方根误差相比于Voxblox、Voxfield和VDBblox分别下降了55.6%、47.11%、21.7%。

2 ORB-TSDF 三维重建系统框架

针对传统算法在线构建三维重建场景精度低、计算量大、实时性不稳定等问题,本文创新性地提出结合ORB_SLAM2与TSDF的ORB-TSDF系统框架,结构图如图1所示。

系统首先从视觉传感器(如单目相机、双目相机、

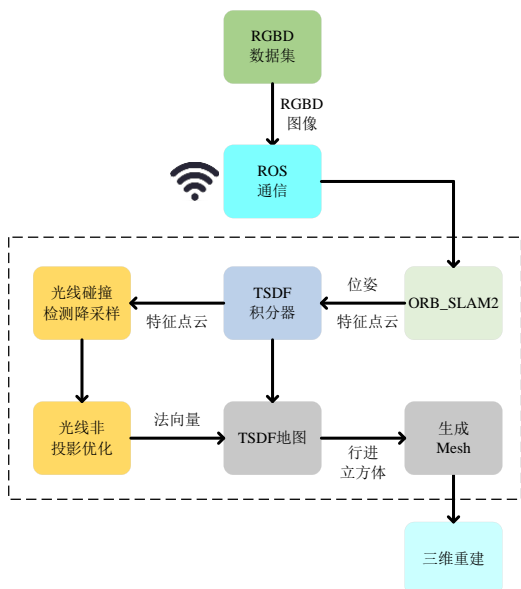


图1 三维重建模型系统框架

RGBD和视觉惯性里程计(如VIO),获得点云信息和位姿信息,然后通过机器人操作系统(Robot Operating System,ROS)在无线局域网的环境下利用不同的话题节点将前端采集的位姿信息和特征点云传输到TSDF积分器.传感器通过光线投影特征点的方式更新空间内每个体素,为了加速TSDF地图构建,本文提出光线碰撞检测策略减少光线遍历体素的数量,通过减少光线遍历体素的数量减轻系统计算量.同时为了提高TSDF地图精度,本文提出利用碰撞检测后的特征点构建局部平面获取特征点法向量信息,利用法向量更新体素梯度从而得到更准确的体素到物体表面真实距离,并以此更新TSDF地图.最后利用Merching Cube^[26]算法构建Mesh,利用Mesh进行在线三维重建。

本文的主要贡献如下:

(1)针对文献[12]和文献[13]建图速率低和实时性差等问题,本文提出一种通过光线碰撞检测降低光线遍历体素的数量并将体素内所有特征点融合的策略,减少每个体素融合的特征点数量,在保证系统精度的同时提升系统构建TSDF地图的效率。

(2)针对文献[12]、文献[13]和文献[19]建图精度低的问题,本文通过(1)筛选后的特征点,为每个特征点寻找其相邻特征点并构建局部平面,利用局部平面构建特征点法向量,并更新体素梯度信息,通过光线投影方向与体素梯度方向优化光线投影距离,构建更精确的TSDF地图。

(3)结合ORB-SLAM2,提出了一种能够实时构建高精度三维重建地图的框架。

2.1 基于光线碰撞检测构建TSDF地图

为了满足在线构建三维场景的实时性,本文通过限制光线遍历特征点的数量减小系统计算量,同时将体素内所有特征点融合为1个平均点,从而优化系统建图精度.本文提出的方法主要采用ORB_SLAM2的Tracking线程实时提取特征点来提高位姿估计的精度. Tracking线程的结构如图2所示,ORB_SLAM2作为目前主流的实时定位与建图系统,在实时提取目标特征点中有非常高的精度,同时还能通过匹配当前帧和相邻帧特征点输出位姿信息。

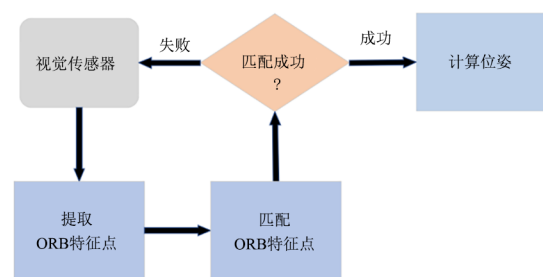


图2 Tracking线程流程

传感器利用光线投影将所有点云投射到全局地图的体素当中,然后将新的投影测量值更新到原有的体素当中. TSDF 是一种将特征点云结构化并以参数来表达物体表面的地图重建算法,可以高效地实时进行三维重建,其原理如图 3 所示. 图中,TSDF 首先将全局三维地图划分为等体积大小的体素立方体,每个小立方体存储体素值和权重. TSDF 根据相机内参、位姿、像素点坐标和深度值得到特征点 P 对应坐标,并根据式(1)~(5)计算体素的值和权重:

$$sdf = d_s - d_v \quad (1)$$

$$d(x, p, s) = sdf / trunc \quad (2)$$

$$w_{const}(x, p) = 1 \quad (3)$$

$$D_{i+1}(x) = \frac{W_i(x)D_i(x) + w(x, p)d(x, p, s)}{W_i(x) + w(x, p)} \quad (4)$$

$$W_{i+1}(x) = \min(W_i(x) + w(x, p), W_{max}) \quad (5)$$

式中,下标 i 表示当前帧; $i+1$ 表示下一帧; $w_{const}(x, p)$ 为体素的权重,默认为 1; W_{max} 为最大权重,默认为 2; s 为传感器原点, p 为输入传感器数据中 3D 点的位置; d_s 为视觉传感器到物体表面的距离, d_v 为传感器到体素的距离, sdf 为体素到物体表面的真实距离; $trunc$ 表示截断距离; $d(x, p, s)$ 为 sdf 除以截断范围的体素值; $W_i(x)$ 为当前体素融合更新前的体素立方体的权重, $W_{i+1}(x)$ 为当前体素融合更新后的体素立方体的权重, $D_i(x)$ 为体素融合更新前的体素立方体的值, $w(x, p)$ 为采集像素点 p 的权重, $D_{i+1}(x)$ 为体素融合更新后的体素立方体的值. 得到 $D_{i+1}(x)$ 后,计算所有 $D_{i+1}(x)$ 为 0 的等势面,即可得到三维重建的 Mesh 模型.

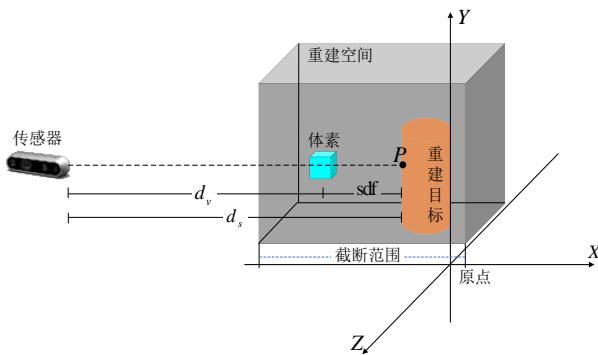


图3 ORB特征点投影体素示意图

传感器会发射光线到达特征点从而遍历地图中的每个体素,因此同一个体素可能会穿过多条光线,体素的尺寸越大则穿过该体素的光线越多,根据式(2)可以求得这个体素的有符号距离 d_v .

为了加快光线遍历体素的效率,本文创新性地提出了如图 4 所示光线碰撞检测融合策略. 已知传感器

在移动过程中会发射光线投影特征点并穿过空间中的体素,定义当 2 条及以上光线穿过同一体素时,将其视为光线碰撞. 光线碰撞检测即对穿过同一体素的光线进行检测,若穿过该体素的历史光线数量达到设定的阈值时,禁止后续光线穿过该体素;反之则允许光线穿过体素,并根据光线投影

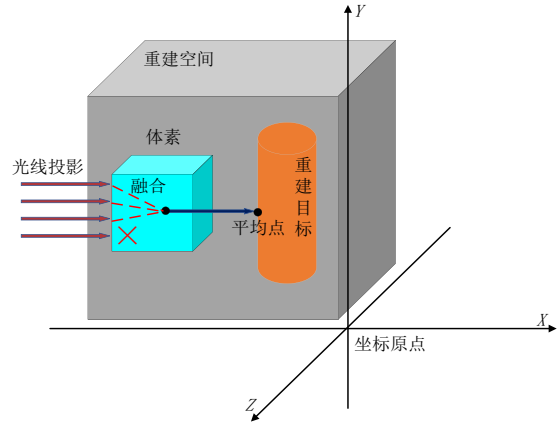


图4 光线碰撞检测融合像素点

距离对体素进行更新. 阈值的设定根据实验可知: 当体素内穿过光线数量达到 20 时,随着体素内光线数量的增加,建图精度没有出现明显提升,但系统计算开销会持续增大,不利于系统三维重建的实时性能,这表明合理地限制穿过体素的光线数量有利于提高系统实时性能,同时能够保证系统建图的精度不会出现显著下降. 如图 4 所示,设置体素内穿过的最大光线数量为 3,上面 3 条光线可正常穿过体素,最下层光线由于光线碰撞检测不能穿过体素. 根据已经穿过该体素的光线的所有投影点求其平均坐标,将其视为 1 个平均点,对此平均点投影射线,根据光线投影距离可得到此体素的 TSDF 值. 平均点的更新如式(6):

$$P_{i+1}(x) = \frac{P_i(x)W_i(x) + p_{i+1}(x)w_{i+1}(x, p)}{W_i(x) + w_{i+1}(x, p)} \quad (6)$$

式中, $p_{i+1}(x)$ 为更新后的平均点, $P_i(x)$ 为更新前的平均点, $W_i(x)$ 为当前体素融合更新前的体素立方体的权重, $p_{i+1}(x)$ 为新检测到的特征点, $w_{i+1}(x, p)$ 为新特征点的权重.

2.2 基于特征点法向量优化 TSDF 值

为了提高系统建图精度,本文创新性地提出了一种光线投影距离优化的策略如图 5 所示. 由于传感器处于移动状态,不同位置的传感器映射特征点时投射的光线方向不同,由于相机光线投影方向与体素到最近物体表面方向不一致,导致光线投影距离比实际体素值偏大,从而使得构建的三维模型存在一定的细缝合漏洞. 与 Voxfield 利用深度图计算所有特征点法向量

信息不同,本文遍历经过第2.1节筛选后的所有特征点 p_i ,寻找 p_i 最近的2个相邻点,构建1个局部平面,根据式(7)求得此局部平面法向量 n_i ,将此局部平面的法向量作为点 p_i 的法向量 n_i ,视觉传感器到点 p_i 的光线投影方向向量 q_i , d_i 为体素的光线投影距离,同时根据式(8)更新体素梯度值 \tilde{g}_j , g_j 表示体素梯度方向,根据式(9)得到真实的体素的有符号距离 d_i^* .

$$n_i = \overrightarrow{p_i p_{i+1}} \times \overrightarrow{p_i p_{i+2}} \quad (7)$$

$$g_j \leftarrow \frac{\tilde{g}_j}{\|\tilde{g}_j\|}, \tilde{g}_j = \frac{W_i(x)g_j + w_i(x,p)n_i}{W_i(x) + w_i(x,p)} \quad (8)$$

$$d_i^* = \left| \frac{\cos \theta_2 \sin \theta_1 + \cos \theta_1 \sin \theta_2 - \sin \theta_1}{\sin \theta_2} \right| \cdot d_i \quad (9)$$

$$\cos \theta_1 = \frac{q_i \cdot g_j}{\|q_i\|} \quad (10)$$

$$\cos \theta_2 = n_i \cdot g_j \quad (11)$$

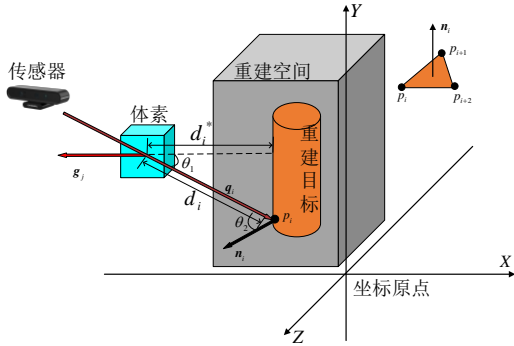


图5 优化体素到物体表面真实距离

传感器发出的光线会遍历从传感器原点到截断距离以内的所有体素,将每8个体素视为1个block,采用block的目的是提高遍历速率,按照block的索引计算每个平均点的sdf值,根据式(5)和式(9)更新前的体素权重 $w_i(x,p)$ 和有符号距离 d_i^* . 最后,融合后的TSDF地图有符号距离 D 和权重值 W 表示为

$$D_{i+1}(x) = \frac{W_i(x)D_i(x) + w_i(x,p)d_i^*}{W_i(x) + w_i(x,p)} \quad (12)$$

$$W_{i+1}(x) = \min(W_i(x) + w_i(x,p), W_{\max}) \quad (13)$$

式中, $D_{i+1}(x)$ 为更新后的有符号距离, $D_i(x)$ 为更新前的有符号距离, $W_i(x)$ 为当前体素融合更新前的体素立方体的权重, $W_{i+1}(x)$ 为当前体素融合更新后的体素立方体的权重, $w_i(x,p)$ 为更新前的体素权重, d_i^* 为真实的体素的有符号距离.

通过哈希表将每8个体素构建为1个形状为正方体的Block,根据每个Block中所有体素的值构建Mesh,通过3个TSDF值为0的体素构建1个三角形的等势面,等势面上所有体素值均为0,将空间中所有等势面进行

积分即可得到三维重建地图.为了得到带颜色的三维重建场景,将TSDF地图中的体素以及对应的颜色变量提取出来,以.mesh和.ply的点云文件进行保存.mesh文件用于生成三维模型,ply文件用于浏览带有RGB颜色的点云效果.使用meshlab三维点云可视化软件,可对保存的三维重建地图进行查看.

3 算法流程

本文提出的算法流程图如图6所示.

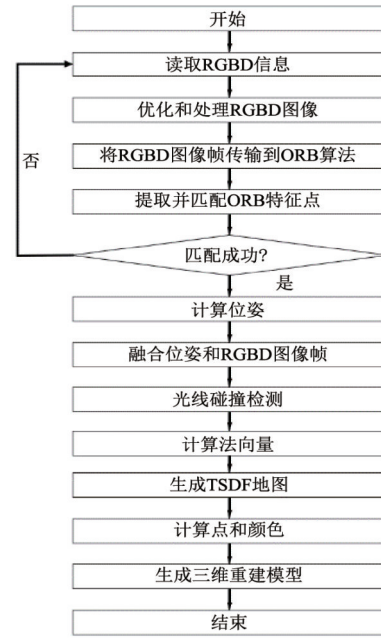


图6 算法流程图

为验证本文所提算法的可行性与适用性,本文在机器人操作系统(Robot Operating System, ROS)中进行实验,该系统包含视觉图像处理模块与地图构建模块之间的通信,并采用Rviz作为实时地图构建的显示平台.本文分别在公开RGB-D数据集和真实场景中进行实验,并将所提出的方法与先进的TSDF地图构建方法Voxblox、Voxfield以及VDBblox进行对比.为了进行对比实验,实验使用Cow&Lady和Flat数据集,数据集信息如表1所示.评价指标包括均方根误差(Root Mean Square Error, RMSE)、标准误差(Standard Error, SE)、均值误差(Mean Error, ME)、重建倒角距离、重建覆盖率、TSDF地图更新速率、特征点数量的选取.数据集包含RGB-D点云数据和位姿信息,实验中,ORB参数使用默认参数:nFeatures=1000, scaleFactor=1.2, iniThFast=20, minThFAST=7.实验采用的处理器为Intel(R) Core(TM) i7-8700K@3.70 GHz,内存为15 GB,操作系统为64位Ubuntu20.04,运行环境为python3.8,硬盘大小为2 TB.实验所涉及的基本仿真参数如表2所示.

表 1 数据集参数

数据集名称	数据集类型	传感器类型	数据集大小
FLAT	虚拟合成	视觉传感器/G	4.48
Cow&Lady	真实场景	视觉传感器/M	658.30

表 2 实验参数

参数名称	参数
图像分辨率	640×480
相机采集频率/Hz	30.00
传感器视距范围/m	5.00
特征点采集数量	1 000.00
数据集运行速率/(帧·s ⁻¹)	0.25

3.1 TSDF 地图精度评估

本文从 TSDF 建图的 RMSE 误差、均值误差、标准误差、重建倒角距离、重建覆盖率评估系统建图精度。

(1) 均方根误差:用于评估重建 TSDF 地图与真值误差大小为

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^n (Y_i - f(x_i))^2} \quad (14)$$

式中, N 为数据个数; Y_i 为 ground truth 提供的真实值; $f(x_i)$ 为重建出来的对象。

(2) 倒角距离:作为衡量重建质量的常用指标,重建地图 M 与真值点云 G 之间的倒角距离^[13] 计算为

$$d_{CD}(M, G) = \frac{1}{2|M|} \min_{g \in G} \|m - g\| + \frac{1}{2|G|} \sum_{m \in M} \min_{g \in G} \|g - m\| \quad (15)$$

(3) 重建覆盖率:定义为带有 mesh 的地面真值点的数量与总的真值点数量之间的比值。

(4) TSDF 地图更新速率:定义为 TSDF 地图平均每帧处理时间。

3.1.1 建图精度结果对比

图 7 所示为本文重建算法与 Voxblox、Voxfield 和 VDBblox 重建算法构建的三维重建模型与真值对比得到均方根误差、标准误差以及均值误差对比图。如图 7 所示, Voxel Size 设置为 25 cm 时, 相比于 Voxblox、Voxfield 和 VDBblox 这 3 种重建算法, 本文所提重建算法的 RMSE 在 Flat 数据集分别下降了 55.6%、47.11%、21.7%, 在 Cow&Lady 数据集下降了 22.3%、20.4%、11.9%。标准误差在 Flat 数据集下降了 47.5%、41.9%、21.2%, 在 Cow&Lady 数据集下降了 12.1%、6.6%、2.5%。均值误差在 Flat 数据集下降了 56.4%、42.7%、21.9%, 在 Cow&Lady 数据集下降了 20.2%、9.6%、6.7%。这表明本文所提重建算法是有效的, 在大场景下重建模型受到体素间重叠过多的影响误差仍然较大, 但相比于 Voxblox 和 Voxfield 重建精度有所提升。

3.1.2 倒角距离误差

倒角距离误差受体素大小的影响如图 8 所示, 图 8(a) 为 Flat 数据集, 图 8(b) 为 Cow&Lady 数据集。4 种重建算法的倒角距离误差随 Voxel Size 改变没有出现明显的差异, 这是由于用 mesh 重建的 Merching Cube 算法主要取决于体素值为 0 的相邻体素的 TSDF 值的比值, 而 4 种重建算法的 TSDF 积分方法之间几乎没有变化。

3.1.3 重建覆盖率

图 9 所示为本文重建算法与 Voxblox、Voxfield、VDBblox 重建算法构建的三维重建模型与真值对比得到重建覆盖率。图 9(a) 为 Flat 数据集, 图 9(b) 为 Cow&Lady 数据集。在小场景下本文所提重建算法覆盖率为 Voxfield、VDBblox 算法相似, 相比于 Voxblox 覆盖率有明显的提升。在大场景下本文所提算法重建覆盖率最高。这是由于采用特征点法向量信息优化了体素投影距离, Voxfield 利用深度图像获取图像中所有特征点法向量信息, VDBblox 在 Voxfield 基础上添加权重更新因子, 其重建方式依旧为非光线投影, 本文根据同一体素内相邻特征点构建局部平面获取法向量信息, 在减轻系统计算量的同时还能构建更精确的 TSDF 地图。在 Voxel Size 为 25 cm 时, 相比于 Voxblox、Voxfield、VDBblox 这 3 种重建算法, 本文所提重建算法的重建覆盖率在 Flat 数据集提升了 1.6%、1.4%、0.57%, 在 Cow&Lady 数据集提升了 1.24%、0.62%、0.27%。表明本文提出的算法有助于提高重建模型覆盖率。

3.2 系统实时三维重建性能评估

本文从 TSDF 地图更新速率、点云数量对地图平均更新速率的影响来评估系统建图实时性能, 由于 VDBblox 以 Open VDB 作为底层实现三维重建, 与传统三维重建算法具有不同的数据结构, 因此, 本文只与 Voxblox 和 Voxfield 对重建性能进行对比评估。

3.2.1 TSDF 地图平均更新速率

为评估本文重建算法实时构建 TSDF 地图的性能, 图 10 所示为本文提出的重建算法与 Voxblox、Voxfield 重建算法的 TSDF 地图平均更新时间对比示意图, 体素大小设置为 5 cm。其中蓝色柱状图代表 Cow&Lady 数据集, 绿色柱状图代表 Flat 数据集。可以看出本文提出的重建算法 TSDF 地图更新时间比 Voxblox 算法和 Voxfield 算法更短, 在 Flat 数据集上本文提出的算法比 Voxfield 重建速率提升了 12.9%, 比 Voxblox 重建速率提升了 9.7%。在 Cow&Lady 数据集上比 Voxblox 重建速率提升了 8.9%, 比 Voxfield 重建速率提升了 10.6%。这是因为 Voxblox 和 Voxfield 需要对图像中的每个特征点进行计算, 而本文只对光线穿过体素的部分特征点进行计算, 在保证建图精度的同时提高了系统运行时间。Flat 数据集 TSDF 地图平均更新时间明显大于 Cow&Lady 数

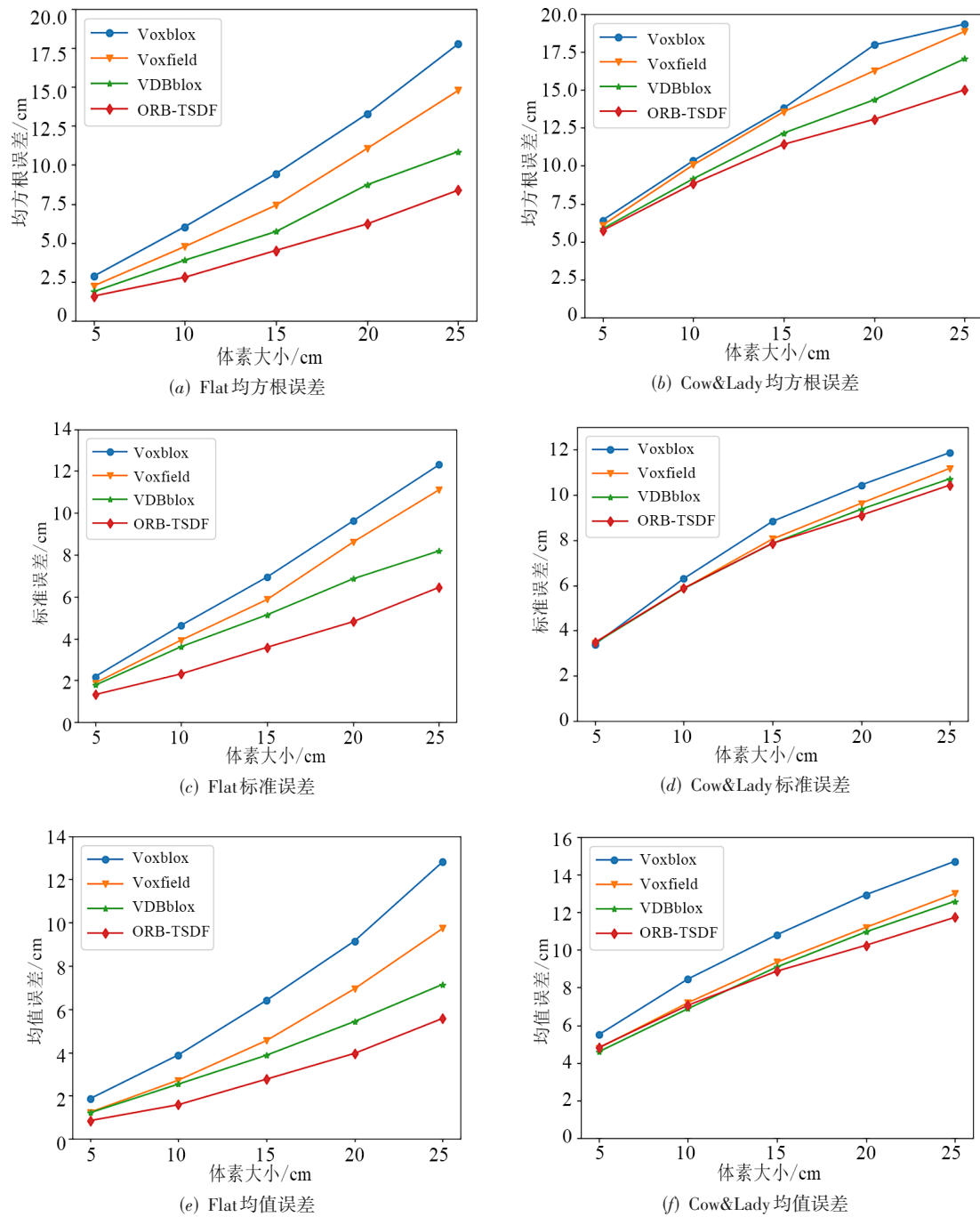


图7 均方根误差、标准误差、均值误差实验对比

据集,这是因为Flat数据集每一帧图像包含的特征点数量远高于Cow&Lady数据集,导致其平均更新时间更长。

3.2.2 体素内特征点数量对系统性能的影响

为验证特征点数量的选取对重建模型精度以及地图平均更新时间的的影响,本文在Flat数据集上测试了不同特征点数量三维重建模型的RMSE以及TSDF地图平均更新时间, Voxel Size为5 cm. 测试结果如图11所

示,随着体素内特征点数量的增多,地图平均更新时间会不断增加, RMSE不断下降. 体素内特征点数量在5~20时RMSE下降比较明显,当体素内特征点数量从5增加到20, RMSE下降了9.34%,而体素内特征点从20增加到40后, RMSE只下降了1.91%,但TSDF地图平均更新时间依然会成正比式上升,这表明每个体素内只需要部分特征点就能达到很好的建图效果,而多余的体素点只会增加系统计算量,不利于实时构建三维地图. 本系统

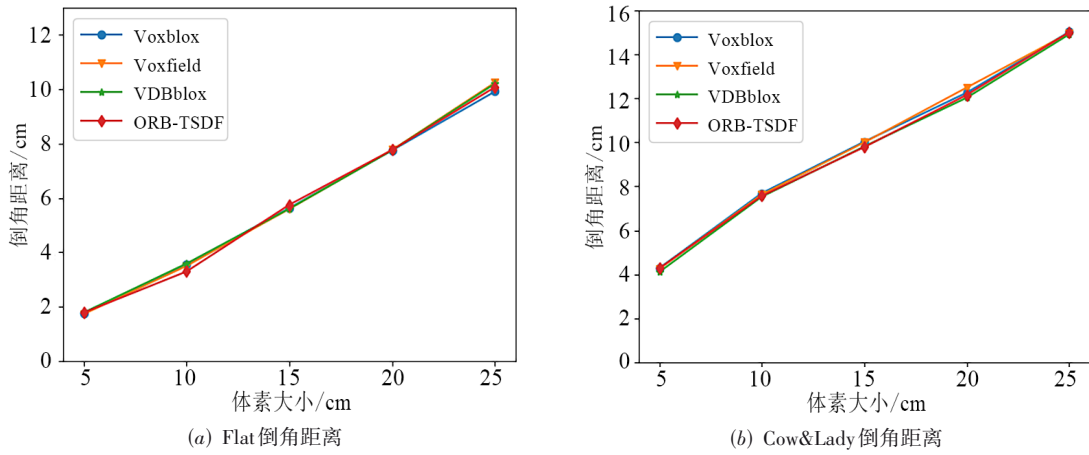


图8 倒角距离误差对比实验

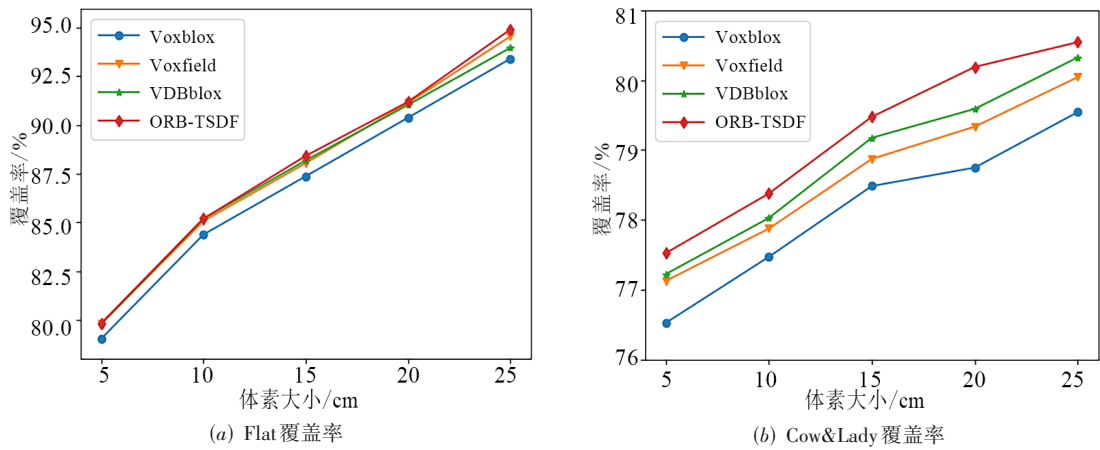


图9 重建覆盖率对比实验

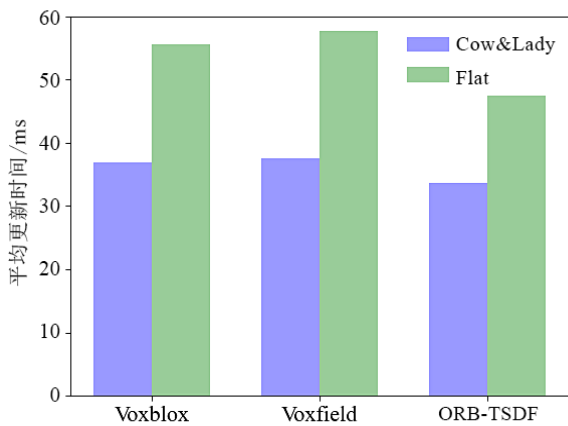


图10 TSDF地图平均更新速率

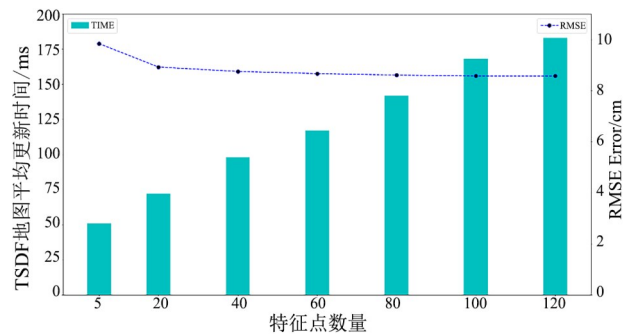


图11 特征点数量对系统性能的影响

将每个体素允许穿过的最大光线数量设定为20,即只随机选取20个点进行计算,在保证系统重建精度的同时提高系统实时性能.同时这也表明本文提出的算法是有效的,可以在保证建图精度的同时减轻系统计算量,达到实时构建高精度三维地图效果.

3.3 消融研究

为验证本文所提出的光线碰撞检测方法和光线投影距离优化策略对三维重建模型的影响,本文在Flat数据集上分别对每种方法进行测试,并与Voxblox、Voxfield进行对比实验.如表3所示,Voxel Size设置为25 cm,PN表示只对每个体素对应的点构建法向量信息方法,没有限制特征点数量.CC表示光线碰撞检测融合方

法,只限制特征点数量为 20. 从表 3 可以看出,CC 算法平均每帧图像更新速率为 32 ms,但 RMSE 与 Voxblox 相似,均值误差和标准误差也没有出现明显的上升,而且占用内存下降了 58.3%. PN 方法明显降低了 RMSE、均

值误差以及标准误差,但地图更新速率没有明显上升. 总而言之,本文提出的光线碰撞检测融合方法和基于特征点法向量更新 TSDF 值的方法不仅保证了系统运行的实时性,同时还提高了系统的对三维场景的重建精度.

表 3 消融研究实验

参数	Voxblox	Voxfield	PN	CC	本文算法
均方根误差/cm	17.746 8	14.763 5	8.379 3	17.607 7	8.535 1
标准误差/cm	12.301 2	11.105 9	6.445 5	12.469 5	6.461 4
均值误差/cm	12.808 6	9.730 5	5.574 4	12.458 8	5.583 8
更新时间/ms	59.000 0	53.600 0	55.700 0	32.000 0	45.200 0
占用内存/MB	483.100 0	428.100 0	483.600 0	201.000 0	206.300 0

注:加粗字体表示性能最好值.

3.4 数据集结果测试

为了更直观地展示本文提出的算法,特别是基于 TSDF 进行场景三维重建的效果,在 ASL 开源数据集进行三维重建,重建效果如图 12 和图 13 所示. 图 12(a)为 Voxblox 算法构建的三维重建模型,图 12(b)为 Voxfield 算法构建的三维重建模型,图 12(c)为 VDBblox 算法构建的三维重建模型,图 12(d)为本文算法构建的三维重建模型.

对比图 12(a)、图 12(b)以及图 12(c),可以看出,本文提出的算法具有更高重建覆盖率,在场景左下角的枕头上,Voxblox、Voxfield 和 VDBblox 3 种算法构建的三维重建模型出现了较大的漏洞,而本文构建的三维重建模型则出现了比较细微的缝隙. 此外本文对左下角

床尾的重建面积最大,这表明本文提出的重建算法是有效的.

对比图 13(a),场景中牛的后面没有完整重建,而 Voxfield、VDBblox 本文提出的方法可以很好地将该部分进行重建. 这表明利用法向量优化光线投影距离是有效的. 与 Voxfield 和 VDBblox 不同,Voxfield 通过将特征点云转化为深度图,在深度图中遍历所有特征点,并从深度图每个点构建法向量;VDBblox 在此基础上添加权重更新因子对体素进行更新,本文只针对体素内的点,寻找 2 个不同方向上的相邻点,从而构建局部平面,从局部平面构建法向量,在保证系统精度的同时可以提高地图更新速率. 图 13(b)中重建场景的地板上依旧会出现缝隙,图 13(c)中牛脚底会出现漏洞,而本

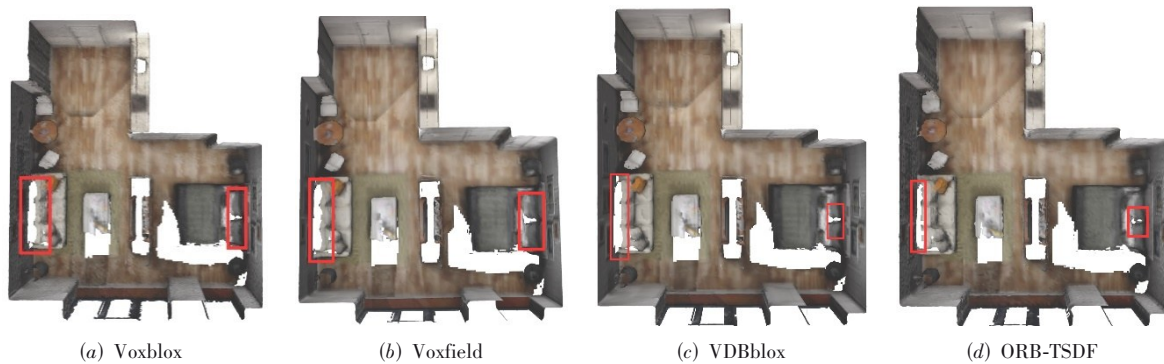


图 12 Flat 数据集重建效果对比

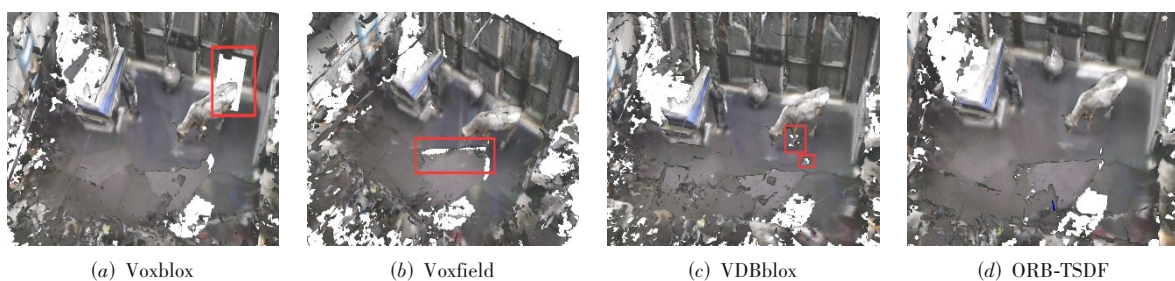


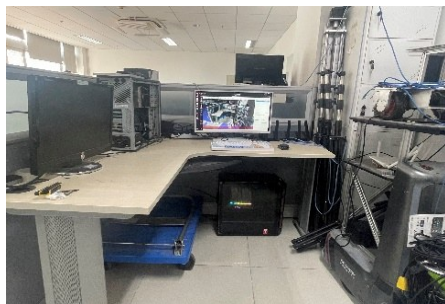
图 13 Cow&Lady 数据集重建效果对比

文构建的三维场景则较完美地将场景进行还原,这表明通过局部平面构建法向量优化光线投影距离的效果要好于采用深度图构建特征点法向量。

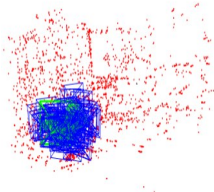
3.5 真实场景实验与分析

本文在真实室内环境下将 ORB-SLAM2 作为前端并移植在 CPU 为 Intel Core i7-12700K 的服务器上,并同时服务器作为后端用于接收关键帧数据,并完成三维重建的任务。重建对象为室内实验室工位,利用 RealSense D435i 相机实时采集图像并获取 ORB 特征点、深度图像帧以及 RGB 图像帧,并发送到服务器后端用于生成三维重建地图。

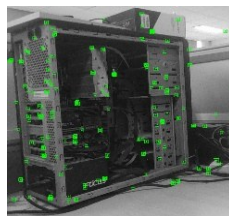
图 14 所示为本文三维重建方法在真实室内环境重建的效果。其中,系统的前端利用 ORB 特征点构建稀疏点云局部地图。后端利用带有颜色信息的 mesh 构建稠密体素地图。最终构建的场景保存为 ply 文件,并通过 meshlab 可视化软件查看。地图处理每一帧的时间仅为 7.35 ms,可以达到实现构建三维重建地图的效果。



(a) 实验场景



(b) 前端局部点云地图



(c) 后端特征提取



(d) 后端三维重建

图 14 真实室内环境三维重建效果

为了展示本系统在大场景下的实时三维重建效果,同时为了体现出相机的移动性和多角度场景建模,使用 CPU 为 Intel(R) Core(TM) i7-10750H 的 MACHENIKE 笔记本作为后端完成三维重建操作,并利用 RealSense D435i 相机采集三维重建所需的图像信息,相机采集频率为 30 Hz,图所获像分辨率为 640×480,使用 ORB-SLAM2 框架用于获取 ORB 特征点、深度图像帧以及 RGB 图像帧。本次采集数据的场地为实验室,真实场景实验及构建的三维重建地图如图 15 所示,其中图 15(a)为真实环境,图 15(b)为本文实时构建的三维重建地图,平均地图更新时间速率为 7.35 ms/帧,根据三维重建结果表明,本文提供的三维重建算法可以在大场景下很好地将场景进行还原,同时也表明本系统搭载的相机可自由移动并进行多角度场景建模。

3.6 轻量化多机协同三维重建

由于 ORBSLAM2 算法中具备回环检测线程,回环检测的目的在于检测出机器人走过一段距离后是否回到了曾经到过的位置,当检测到相同地图点或相同目标物体时,将地图进行全局更新。本文将回环检测进行简单的拓展,即通过 2 个轻量化子端搭载 Xavier 进行实时三维重建,并将局部地图和关键帧发送到服务端,服务端通过关键帧进行回环检测,检测到相同地图点时将 2 个局部地图进行融合,在服务端构建全局一致的三维重建地图。本次采集数据的场地为实验室,构建的三维重建地图 16(a)和图 16(b)分别表示子端 1 和子端 2 在实验室中移动并实时构建的三维重建地图,图 16(c)表示服务端构建的全局地图,其中黄线代表子端 1 的运动轨迹,绿线代表子端 2 的运动轨迹,地图融合发生在两个子端经过相同位置时,此后服务端将局部地图融合为全局地图,并根据子端采集的图像数据信息对全局地图进行实时更新。

表 4 为本文系统的定量实验结果。其中,第 1 列表示轻量化各项测试指标,包括功耗、帧率以及能效比,第 2 和第 3 列分别表示 2 个子端使用 Xavier 处理器在真实场景下运行轻量化实时三维重建算法的各项测试指标数值。第 4 列表示系统服务端将子端进行融合时各项测试指标数值。表中所有数据均为运行 5 次后取的平均值,其中空白处表示无需测试的性能指标。

从表 4 中的实验结果可知,2 个子端在真实场景中实际平均运行帧率为 17.56 fps,基本满足实时采集图像的最低帧率。此外,2 个子端的平均功耗为 12.58 W,证明了本文系统在真实场景下的部署和实时运行的可行性和轻量化特性。因此,本文系统中使用 2 个 Xavier 低算力平台处理器可以实现真实场景下实时构建三维重建地图的效果,满足实际需求。

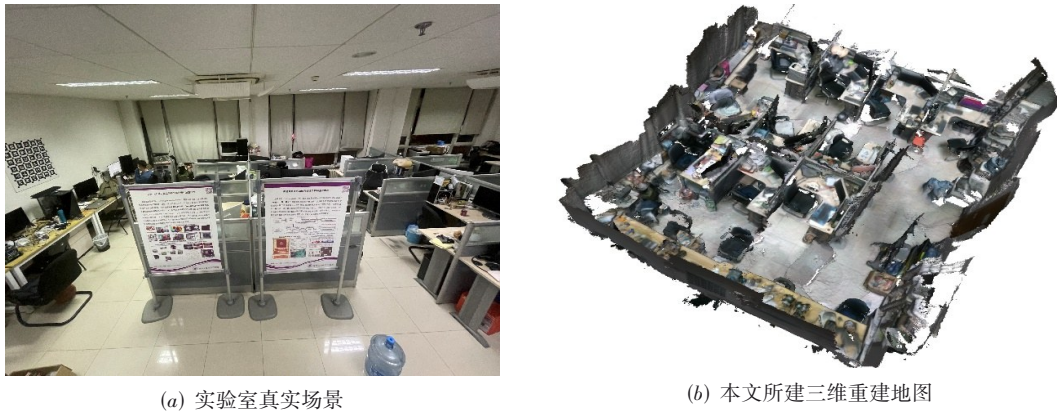


图 15 实验室场景及构建的三维重建地图

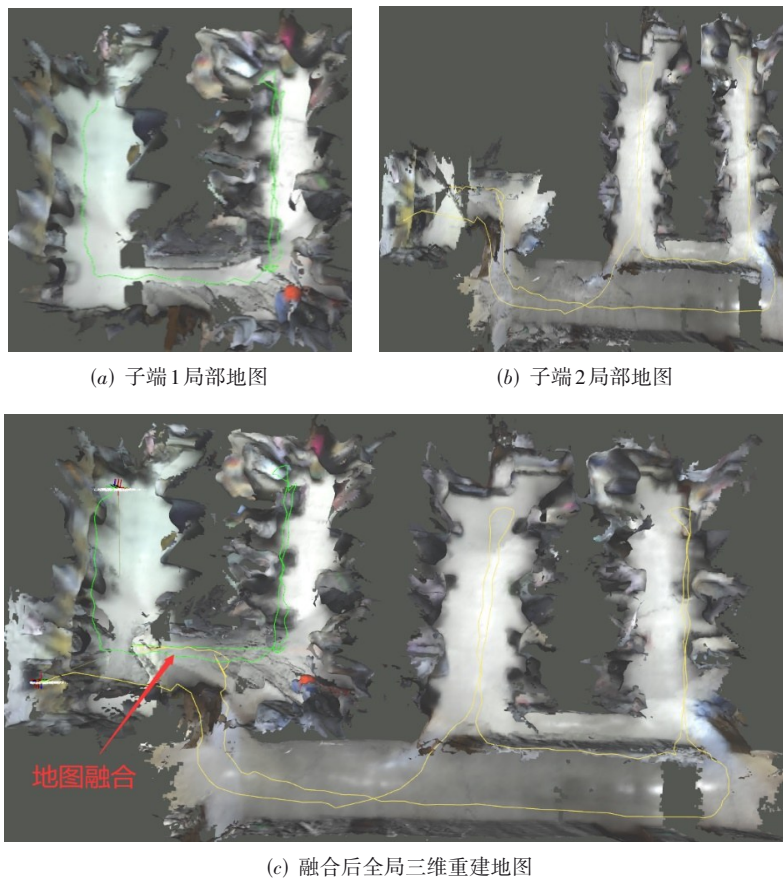


图 16 轻量化多机融合三维重建

表 4 轻量化性能测试

参数	子端 1	子端 2	服务端
功耗/W	12.54	12.62	42.89
帧率/fps	17.32	17.81	—
能效比/(fps·W ⁻¹)	1.38	1.41	—

4 结论

为了实现在线三维重建并端到端输出三维重建模

型,本文提出基于ORB-SLAM2和TSDF的在线重建算法ORB-TSDF. 同时在保证系统精度的同时提高系统建图效率,采用光线碰撞检测的方法,减少穿过体素的光线数量,提高系统实时构建TSDF地图的效率. 为了提高系统建图精度,对每个体素内剩余的特征点添加法向量信息,并更新体素梯度信息,通过体素梯度优化TSDF值,降低TSDF地图误差,并提高重建模型覆盖

率. 在 ASL 开源数据集上的对比实验结果以及真实场景下的实验表明, 本文所提出方法能够实时构建高精度三维重建地图. 在未来, 可以通过 TSDF 地图增量式构建 ESDF 地图, 实现在复杂场景下的导航和避障.

参考文献

- [1] JIA P S. Research on positioning and navigation of medical robot based on RGB-D visual SLAM[C]//2023 4th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT). Piscataway: IEEE, 2023: 627-632.
- [2] 代少升, 刘小兵, 赖智颖, 等. 网格化局部自适应 DBSCAN 聚类算法[J]. 重庆邮电大学学报(自然科学版), 2022, 34(2): 250-257.
DAI S S, LIU X B, LAI Z Y, et al. Gridded local adaptive DBSCAN clustering algorithm[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2022, 34(2): 250-257. (in Chinese)
- [3] 郭艺轩, 贾向东, 曹胜男, 等. 三维动态无人机网络覆盖性能与信道容量分析[J]. 重庆邮电大学学报(自然科学版), 2022, 34(4): 662-668.
GUO Y X, JIA X D, CAO S N, et al. Analysis of 3D dynamic UAV network coverage performance and channel capacity[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2022, 34(4): 662-668. (in Chinese)
- [4] 陈孟元, 韩朋朋, 刘金辉, 等. 动态遮挡场景下基于改进 Transformer 实例分割的 VSLAM 算法[J]. 电子学报, 2023, 51(7): 1812-1825.
CHEN M Y, HAN P P, LIU J H, et al. Improved transformer instance segmentation under dynamic occlusion based VSLAM algorithm[J]. Acta Electronica Sinica, 2023, 51(7): 1812-1825. (in Chinese)
- [5] 陈昌川, 全锐杨, 张谦, 等. 基于轻量化智能的多机协同 SLAM 系统[J]. 仪器仪表学报, 2022, 43(12): 188-198.
CHEN C C, QUAN R Y, ZHANG Q, et al. Lightweight intelligence-based multi-machine collaborative SLAM system[J]. Chinese Journal of Scientific Instrument, 2022, 43(12): 188-198. (in Chinese)
- [6] NEWCOMBE R A, IZADI S, HILLIGES O, et al. Kinect-Fusion: Real-time dense surface mapping and tracking[C]//2011 10th IEEE International Symposium on Mixed and Augmented Reality. Piscataway: IEEE, 2011: 127-136.
- [7] LIU K, SUN H X, YE P. Research on bundle adjustment for visual SLAM under large-scale scene[C]//2017 4th International Conference on Systems and Informatics (ICSAI). Piscataway: IEEE, 2017: 220-224.
- [8] WHELAN T, JOHANSSON H, KAESS M, et al. Robust real-time visual odometry for dense RGB-D mapping[C]//2013 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2013: 5724-5731.
- [9] ZENG M, ZHAO F K, ZHENG J X, et al. Octree-based fusion for realtime 3D reconstruction[J]. Graphical Models, 2013, 75(3): 126-136.
- [10] ENDRES F, HESS J, STURM J, et al. 3-D mapping with an RGB-D camera[J]. IEEE Transactions on Robotics, 2014, 30(1): 177-187.
- [11] ZHANG T W, ZHANG H Y, LI Y, et al. FlowFusion: Dynamic dense RGB-D SLAM based on optical flow[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2020: 7322-7328.
- [12] OLEYNIKOVA H, TAYLOR Z, FEHR M, et al. Voxblox: Incremental 3D euclidean signed distance fields for on-board MAV planning[C]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2017: 1366-1373.
- [13] PAN Y, KOMPIS Y, BARTOLOMEI L, et al. Voxfield: Non-projective signed distance fields for online planning and 3D reconstruction[C]//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2022: 5331-5338.
- [14] REIJGWART V, MILLANE A, OLEYNIKOVA H, et al. Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps[J]. IEEE Robotics and Automation Letters, 2020, 5(1): 227-234.
- [15] CHANG Y, TIAN Y L, HOW J P, et al. Kimera-multi: a system for distributed multi-robot metric-semantic simultaneous localization and mapping[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2021: 11210-11218.
- [16] BARTOLOMEI L, KARRER M, CHLI M. Multi-robot coordination with agent-server architecture for autonomous navigation in partially unknown environments[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2020: 1516-1522.
- [17] LIU X Y, YE W C, TIAN C R, et al. Coxgraph: Multi-robot collaborative, globally consistent, online dense reconstruction system[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2021: 8722-8728.
- [18] 李西亮. 基于子图的多机器人协同语义稠密建图方法

的研究[D]. 杭州: 浙江大学, 2023.

LI X L. Research on Multi-Robot Ccooperative Semantic Dense Mapping Method Based on Subgraph[D]. Hangzhou: Zhejiang University, 2023. (in Chinese)

- [19] BAI Y L, MIAO Z Q, WANG X K, et al. VDBblox: Accurate and efficient distance fields for path planning and mesh reconstruction[C]//2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2023: 7187-7194.
- [20] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Piscataway: IEEE, 2007: 225-234.
- [21] SONG S Y, CHANDRAKER M, GUEST C C. Parallel, real-time monocular visual odometry[C]//2013 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2013: 4698-4705.
- [22] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.
- [23] WHELAN T, SALAS-MORENO R F, GLOCKER B, et al. ElasticFusion: Real-time dense SLAM and light source estimation[J]. The International Journal of Robotics Research, 2016, 35(14): 1697-1716.
- [24] DAI A, NIEBNER M, ZOLLHÖFER M, et al. BundleFusion[J]. ACM Transactions on Graphics, 2017, 36(4): 1.
- [25] SUN J M, XIE Y M, CHEN L H, et al. NeuralRecon: Real-time coherent 3D reconstruction from monocular video [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 15593-15602.
- [26] RAJON D A, BOLCH W E. Marching cube algorithm: Review and trilinear interpolation adaptation for image-based dosimetric models[J]. Computerized Medical Imaging and Graphics, 2003, 27(5): 411-435.

作者简介



陈昌川 男, 1978年10月出生, 四川广安人. 重庆邮电大学通信学院副教授. 主要研究方向为智能信息处理、图像处理与人工智能、特高频局放检测、红外成像与测温.
E-mail: creditdegree@gmail.com



丁锡烽 男, 1999年8月出生, 重庆万州人. 重庆邮电大学通信学院硕士研究生. 主要研究方向为视觉SLAM和三维重建.
E-mail: S210131034@stu.cqupt.edu.cn



王亦何 男, 2006年5月出生, 甘肃白银人. 清华大学电子系本科生. 主要研究方向为SLAM算法、三维建图算法.
E-mail: wangyihe21@mails.tsinghua.edu.cn



孙 霞 女, 1976年9月出生, 河南周口人. 重庆工程学院电子学院副教授. 主要研究方向为多智能体系统、智能算法、数字信号处理.
E-mail: 51552654@qq.com



乔 飞 男, 1977年11月出生, 山西太原人. 清华大学电子系副研究员. 主要研究方向为低功耗集成电路设计、“感算共融”智能感知集成电路和系统.
E-mail: qiaofei@tsinghua.edu.cn