

DOI:

文章编号:

一种改进匹配点对选取策略的 ElasticFusion 室内 三维重建算法

王玮琦 游雄 杨剑 李钦

信息工程大学地理空间信息学院, 河南 郑州 450052

摘要: 对室内场景进行实时高质量的三维重建是机器人、增强现实等领域关注的重点。目前基于 RGB-D 传感器的三维重建方法存在局部模型重建效果差、点云模型存在孔洞等问题。而影响三维模型重建效果的主要因素有两个, 一是由点云配准解算出的位姿参数精度, 二是闭环检测准确程度。对此, 本文在保证算法实时性的基础上, 通过改进 ICP 算法中匹配点的选取策略, 提升模型重建效果。并利用径向基函数构建隐式曲面的方式对点云模型中存在的孔洞进行事后修补。最后选用 ICL-NUIM 等公开数据集进行实验验证, 实验结果表明, 改进后的算法在模型重建效果以及相机轨迹估计方面均有显著提升。

关键词: ElasticFusion 算法; 三维重建; ICP 算法; 孔洞修补

中图分类号: TP391

文献标识码: A

应用成本相对低廉的传感器获取场景的精细三维模型是机器人、增强现实等领域的研究热点。在机器人领域中, 为了实现机器人在未知场景的自主移动, 研究者们希望通过构建“地图”来指导机器人的行动, 其中代表性研究方向是同时定位与地图构建 (Simultaneous Localization and Mapping, SLAM)。在最初的研究中, 因计算资源的限制以及实时性要求, 经典 SLAM 算法基于 CPU 实现传感器定位定姿, 应用稀疏特征点法^[1]或是直接法^[2,3]实现地图构建, 仅能实时重建非常稀疏的三维点云或是半稠密的三维点云, 难以表达场景的细节信息。2010 年底, 微软公司发布了一种可实时获取彩色 RGB 图像以及深度图像的 RGB-D 深度摄像头“Kinect”, 其采集速度快、精度较高且成本低廉, 被迅速的应用于机器人与增强现实等领域。应用深度图像的优势在于不仅能辅助相机跟踪, 还可以提供稠密的场景三维信息。Newcombe 等人于 2011 年率先提出基于 RGB-D 传感器的室内实时三维重建算法 KinectFusion^[4], 引领了后续关于基于深度摄像头的三维重建工作以及稠密视觉 SLAM 的研究。以 KinectFusion 为基础, 相继涌现

了 Kintinuous^[5], ElasticFusion^[6], VoxelHashing^[7], BundleFusion^[8] 等优秀的室内场景三维重建算法。

在满足刚性变换以及室内静态场景的前提下, KinectFusion 算法使用深度信息进行位姿估计实现定位, 设计高效并行的算法在 GPU 上运行, 不同于传统 SLAM 算法的实现方式。KinectFusion 算法基于 TSDF (Truncated Signed-Distance Function, 截断符号距离函数) 表达三维模型, 使用 ICP 算法 (Iterative Closest Point Algorithm, 迭代最近点算法) 估计位姿并融合点云, 应用光线投影算法进行三维渲染。作为开创性的工作, 利用 GPU 首次实现了室内场景的实时稠密三维重建, 在小范围场景中具有不错的重建效果, 但其局限性有: ①基于 TSDF 预设的三维网格限制了重建场景的尺度, 仅能重建小范围场景; ②仅使用 ICP 算法进行位姿估计严重依赖于场景中几何特征的丰富程度, 当场景中出现光强度变化时会引起定位失效; ③没有闭环检测模块, 无法实现闭环优化; ④定位失效时无法进行重定位, 累积误差大等。为了解决上述问题, 研究者们开展了一系列研究工作: ①为了突破场景尺度的

收稿日期: 2018-07-16

项目资助: 国家重点研发计划(No.2017YFB0503500); 河南省科技创新(中原学者, No.142101510005)

第一作者: 王玮琦, 硕士研究生, 主要从事机器人地图方向研究。809741461@qq.com

通讯作者: 游雄, 博士, 教授。youarexiong@163.com

限制, 提出让三维网格随相机运动, 移出三维网格的部分不参与计算^[7]; 或是使用八叉树数据结构减弱对于显存的依赖^[9]; ②为了减小对于场景几何信息的依赖, 提出了融合 ICP, FOVIS, RGB-D 三种位姿跟踪方法, 减小误差累积^[5]; ③为了实现回环检测, 提出了使用 DBoW 词袋法检测闭环回路, 构建位姿图 (pose graph) 进行优化^[5]; ④为了改进定位失效并提升模型显示效果, 提出了利用边线特征中的点对点关系辅助原始 ICP 算法提高鲁棒性以及通过预设地面点云模型或添加极线及共面约束改进模型显示效果^[10, 11]; ⑤为了表达更大规模场景并提升模型的表达效果, 舍弃了基于预设网格表达场景的方式, 提出了使用面元 (surfels) 表达场景, 面元包含了三维点的位置、法向量等信息, 将传感器获取的信息直接作用到面元上而非通过位姿图间接作用于三维模型^[6, 12, 13]。

在上述的改进工作中, 2015 年提出的 ElasticFusion 算法融入借鉴了很多优秀的思想, 使用面元表达场景可以更好地重建大规模场景, 摆脱了以往基于 TSDF 对于场景尺度的限制, 并且由于随机藤算法的引入有效地减少了模型中的“重影”现象, 提升了模型的重建效果。但随机藤算法在闭环检测模块上的应用效果仍不及 DBoW 词袋法, 限制了 ElasticFusion 算法的重定位能力, 并且在初始点云配准阶段选用了经典 ICP 算法中基于投影法选取匹配点对的方式, 在初始配准时易出现误差造成位姿参数解算不准确, 在应用于小规模, 很少存在回环的场景时, 会出现模型局部重建效果差, 甚至应用在大规模场景重建时, 重建效果错乱等缺陷。

本文针对室内小规模场景局部模型重建效果差的问题, 在保证算法实时性的基础上, 通过改进原始 ICP 算法中匹配点的选取策略, 提升模型重建效果, 提高位姿参数解算精度。针对模型重建过程中因物体遮挡、物体反射率差异等造成模型中出现孔洞的问题, 利用径向基函数构建隐式曲面的方式

实现点云模型孔洞修补。

1. ElasticFusion 算法介绍

ElasticFusion 算法借鉴了以 KinectFusion 为代表的实时稠密视觉 SLAM 算法的经典框架, 在充分使用 GPU 的情况下选用 OpenGL 完成点云的更新、融合与显示。经典 SLAM 算法通过不断优化相机轨迹或是稀疏特征点等方式来提高模型精度, 而 ElasticFusion 算法则是通过不断优化已经重建好的模型来提高相机轨迹估计与位姿估计的精度。

不同于以往算法使用 TSDF 模型融合点云表达场景, ElasticFusion 算法使用面元模型表达场景。应用几何及光度一致性进行位姿估计并加以优化, 同时将已经重建好的模型根据预设的时间阈值 δ_t 划分为当前视角下的活动区域 (active area) 以及不在当前视角下的非活动区域 (inactive area)。并且引入随机藤算法实现回环检测与重定位, 提高了系统的鲁棒性^[14]。

ElasticFusion 算法主要分为四个模块:

①数据采集与处理, 通过深度传感器获取深度图像与 RGB 图像, 并将其转化为三维点云, 得到其三维坐标与法向量; ②位姿估计, 将当前帧的三维点云与已有重建模型进行配准, 几何上通过 ICP 算法计算位姿, 同时利用 RGB 图像在光度以及颜色一致性加以约束优化位姿估计结果; ③当前帧图像编码与回环检测, 利用随机藤算法中的编码机制对当前帧图像进行编码, 其编码值与当前存储关键帧编码值的数据库进行比对, 判断当前帧能否作为关键帧以及当前帧与已有模型是否存在回环 (若存在回环, 应用“形变图 (deformation graph)”^[12]) 对已重建模型以及位姿参数进行优化); ④点云融合, 将当前帧获取的点云与已有模型进行融合、更新及显示, 为下一帧配准提供依据。算法流程如图 1 所示。

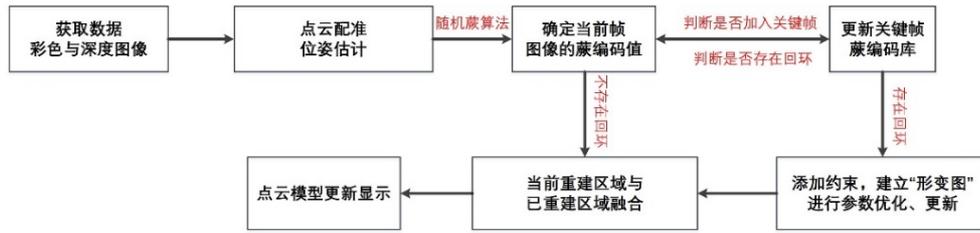


图 1 ElasticFusion 算法流程图

Fig.1 Workflow of the ElasticFusion algorithm

2. 改进的 ElasticFusion 算法

2.1 影响模型重建效果的因素分析

影响三维模型重建效果的原因主要有两点，一是初始点云配准精度，二是闭环检测准确程度。点云配准阶段，ElasticFusion 算法应用几何及颜色光度一致性约束实现配准。在闭环检测中应用随机藤算法，并没有选择目前最为主流的 DBoW 词袋法。其主要原因是算法在保证实时性的基础上基于 GPU 实现三维场景的稠密重建，算法的核心工作在于场景重建，若在闭环检测模块中改用词袋法，虽然精确性上可能会有提升，但会消耗较多计算资源，影响重建的实时性。因此，考虑通过减小点云初始配准阶段产生的累积误差来提高位姿参数解算精度进而提升模型重建效果。

2.2 ElasticFusion 中的 ICP 算法应用

ICP 算法广泛应用于点云配准中，在稠密 SLAM 中位姿估计的常用方法就是通过投影法确定点对的匹配关系，应用最小二乘法使点到平面的距离最小，进而完成位姿参数解算。利用 ICP 算法解算出的位姿参数是点云融合的重要参数。尽可能减小位姿参数解算的误差能够极大程度地帮助三维重建算法进行点云融合与更新显示。而相邻帧间点云对匹配点的选取直接影响着 ICP 算法进行点云配准的精度。本文在经典 ICP 算法匹配点对选取策略的基础上，改进匹配点选取方法，减小初始误差对于点云配准的影响。

2.2.1 匹配点选取方法

考虑到相邻两帧间位姿变化相对较小，通常采用投影法选取 ICP 算法所需要的匹配点对。图 2 为相邻两帧间点对选取示意图（以二维表示）。相机连续两帧采集的环境信息抽象表示为曲线 $K, K+1$ ，当前帧坐标

系原点为 O_{k+1} 。在选取匹配点对时，以 O_{k+1} 为基准，利用投影法近似选取 q_k 与 p_k 作为匹配点对进行位姿解算。

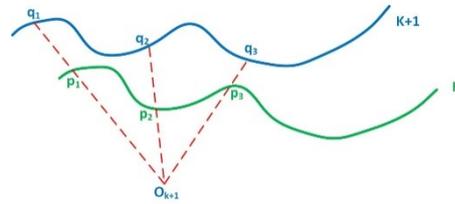


图 2 投影法选取匹配点示意图

Fig.2 Illustration of search of the matching points of the given pose point using projection method

2.2.2 位姿参数解算

在两帧间点云进行配准时，为了减小误差造成的影响，利用最小化点到平面距离的方式进行位姿参数的解算。相较于最小化点对点的距离进行优化，最小化点到平面距离在二三维点云配准上有着显著优势^[15]。最小化点到平面距离示意图如图 3 所示。

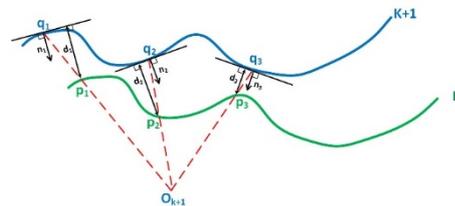


图 3 点到平面距离示意图

Fig.3 Illustration of the distance from the given pose point to the current plane

使目标函数 E 取得最小值求解位姿参数：

$$E = \sum_{l=1}^k ((Rp_l + t - q_l)^T \cdot n_l)^2 \quad (1)$$

R 为旋转矩阵， t 为平移参数，将 R 表示为绕 XYZ 三轴旋转角度分别为 α, β, γ ，

可得:

$$\mathbf{R} = \mathbf{R}_Z \mathbf{R}_Y \mathbf{R}_X \quad (2)$$

其中:

$$\mathbf{R}_Z = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R}_Y = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix}$$

$$\mathbf{R}_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix}$$

$$\mathbf{t} = [t_x \ t_y \ t_z]^T \quad (3)$$

根据算法中采用的小角近似处理方式,

即

$$\sin \theta \approx \theta, \quad \cos \theta \approx 1 \quad (4)$$

将 $[\mathbf{R} | \mathbf{t}]$ 表示为:

$$[\mathbf{R} | \mathbf{t}] = \begin{bmatrix} 1 & \gamma & -\beta & t_x \\ -\gamma & 1 & \alpha & t_y \\ \beta & -\alpha & 1 & t_z \end{bmatrix} \quad (5)$$

将 $[\mathbf{R} | \mathbf{t}]$ 改写为:

$$\mathbf{x} = [\alpha \ \beta \ \gamma \ t_x \ t_y \ t_z]^T \quad (6)$$

根据

$$\mathbf{R}\mathbf{p}_l + \mathbf{t}$$

$$= \begin{bmatrix} 0 & \gamma & -\beta \\ -\gamma & 0 & \alpha \\ \beta & -\alpha & 0 \end{bmatrix} \mathbf{p}_l + \mathbf{I}_{3 \times 3} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} + \mathbf{I}_{3 \times 3} \mathbf{p}_l$$

$$= \begin{bmatrix} 0 & -p_{lz} & p_{ly} & 1 & 0 & 0 \\ p_{lz} & 0 & -p_{lx} & 0 & 1 & 0 \\ -p_{ly} & p_{lx} & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ t_x \\ t_y \\ t_z \end{bmatrix} + \mathbf{p}_l$$

$$= \mathbf{A}(\mathbf{p}_l)_{3 \times 6} \mathbf{x}_{6 \times 1} + \mathbf{p}_{l_{3 \times 1}} \quad (7)$$

其中 $\mathbf{I}_{3 \times 3}$ 为三阶单位矩阵, 可得:

$$\begin{aligned} E &= \sum_{l=1}^k ((\mathbf{R}\mathbf{p}_l + \mathbf{t} - \mathbf{q}_l)^T \cdot \mathbf{n}_l)^2 \\ &= \sum_{l=1}^k ((\mathbf{A}(\mathbf{p}_l)\mathbf{x} + \mathbf{p}_l - \mathbf{q}_l)^T \cdot \mathbf{n}_l)^2 \\ &\stackrel{\text{def}}{=} \sum_{l=1}^k (\mathbf{x}^T_{1 \times 6} \mathbf{B}_{6 \times 6} \mathbf{x}_{6 \times 1} + 2\mathbf{A}_{1 \times 6} \mathbf{x}_{6 \times 1} + C) \end{aligned}$$

对 $f(\mathbf{x}) = \mathbf{x}^T_{1 \times 6} \mathbf{B}_{6 \times 6} \mathbf{x}_{6 \times 1} + 2\mathbf{A}_{1 \times 6} \mathbf{x}_{6 \times 1} + C$ 式中 \mathbf{x} 求导数, 令导数为 0 可求解 \mathbf{x} , 即

$$\begin{aligned} f'(\mathbf{x}) &= 2\mathbf{x}^T_{1 \times 6} \mathbf{B}_{6 \times 6} + 2\mathbf{A}_{1 \times 6} = \mathbf{0} \\ \mathbf{x}^T &= -\mathbf{A}_{1 \times 6} \mathbf{B}^{-1}_{6 \times 6} \end{aligned} \quad (8)$$

其中:

$$\begin{aligned} \mathbf{A}_{1 \times 6} &= (\mathbf{p}_l - \mathbf{q}_l)^T_{1 \times 3} \mathbf{n}_{l_{3 \times 1}} \mathbf{n}_l^T_{1 \times 3} \mathbf{A}(\mathbf{p}_l)_{3 \times 6} \\ \mathbf{B}_{6 \times 6} &= \mathbf{A}(\mathbf{p}_l)^T_{6 \times 3} \mathbf{n}_{l_{3 \times 1}} \mathbf{n}_l^T_{1 \times 3} \mathbf{A}(\mathbf{p}_l)_{3 \times 6} \\ C &= \mathbf{n}_l^T_{1 \times 3} (\mathbf{p}_l - \mathbf{q}_l)_{3 \times 1} (\mathbf{p}_l - \mathbf{q}_l)^T_{1 \times 3} \mathbf{n}_{l_{3 \times 1}} \end{aligned}$$

经多次迭代计算后, 可解算出位姿估计参数。

2.3 ICP 算法中匹配点选取策略的改进

ICP 算法虽然通过不断迭代求解出最优的位姿参数, 但在迭代求解过程中点对的对应关系并不改变, 因此, 匹配点对选取的恰当与否直接影响着位姿参数的解算精度, 影响着模型的重建质量。在原始算法中, 通过投影法选取匹配点对时, 虽粗略考虑法向量夹角以及点对间欧氏距离等因素, 但仍存在匹配点选取不恰当等问题。因相邻两帧间位姿变化通常相对较小 (正常情况下, 相机平移及旋转比较平滑), 在尽量保证算法实时性, 避免占用较大计算资源的基础上, 借鉴文献^[16]的思想, 考虑将法向量作为选取匹配点的重要参考指标。

在投影法选取出匹配点的基础上, 以一定半径确定圆形搜索区域, 遍历区域内包含该帧点云的所有候选点来寻找最佳匹配点, 如图 4 所示

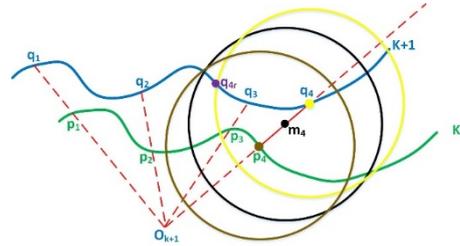


图 4 基于不同圆心选取匹配点示意图

Fig.4 Illustration of search of matching point using circles with different centers

以上图选取 p_4 的匹配点为例, 在已有匹配点 q_4 的基础上, 分别选取 p_4 , q_4 以及 p_4q_4 中点 m_4 为圆心, 以半径 $r = \frac{3}{2} |p_4q_4|$ 确定圆形搜索区域 Q (参数 $\frac{3}{2}$ 的选定使得圆形搜索区域

包含更多的候选点), 当前帧点云包含在圆形搜索区域内的所有点即作为候选点 $q_i(q_i \in Q)$, 求解公式:

$$f(q_i) = \lambda \cdot N(q_i) \cdot N(p_n) \quad (9)$$

其中, λ 为可调节权重值, $N(*)$ 为*点法向量。采用向量点乘的方式选取两帧间法向量最为接近的两点作为匹配点。根据 $f(q_i)$ 所得数值, 所选匹配点 q^* 可表示为:

$$q^* \leftarrow \max f(q_i) \quad (10)$$

根据各圆心确定的搜索区域内候选匹配点的情况, 经过实验结果对比, 确定选取线段中点作为圆心, 1.5 倍线段长度作为半径, 如图 4 黑色圆心及黑色圆形搜索区域所示。由起始点 p_n 及根据投影法确定的匹配点 q_n , 选取线段 $p_n q_n$ 中点 m_n 为圆心, 确定半径 $r = \frac{3}{2} |p_n q_n|$ 的圆形区域 Q_n , 在 Q_n 内的所有点作为候选点 q_i , 通过计算选取 p_n 的最佳匹配点 $q^* \leftarrow \max(\lambda \cdot N(q_i) \cdot N(p_n))$ 。

当根据投影法无法选出合适的初始匹配点时, 选取该帧点云中起始点 q_0 或终止点 q 任一点确定半径 $r = \frac{3}{4} |p_n q|$ 或 $r = \frac{3}{4} |p_n q_0|$ 圆形搜索区域, 如图 5 所示

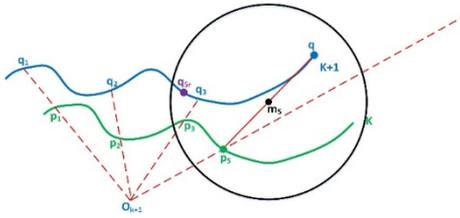


图 5 无初始匹配点时选取匹配点示意图

Fig.5 Illustration of search of matching points without an initial matching point

当相邻两帧出现大量点云平行的情况时(如图 6 所示), 减小 ICP 算法所占权重, 借鉴文献^[10]提出的提取深度数据中的边线特征并结合颜色光度一致性进行位姿估计。

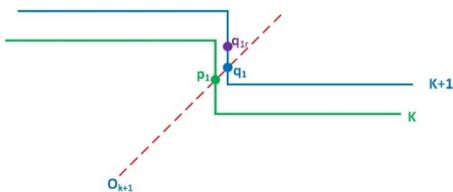


图 6 点云平行情况示意图

Fig.6 Illustration of search of the matching

point when parallel point clouds exist

2.4 利用径向基函数实现点云孔洞修补

ElasticFusion 算法应用于室内场景, 由于深度摄像头在采集数据时会由于物体遮挡、物体表面出现反射以及光强度变化等原因引起采集数据缺失造成在重建模型中出现孔洞。对于机器人来说, 模型中的孔洞影响着其对于场景的理解。以机器人导航为例, 在地面存在孔洞的情况下, 机器人难以判断孔洞部分是否可通行。为了更好地表达场景, 在已有重建模型的基础上, 利用径向基函数对重建点云模型中孔洞进行修补, 进一步提升模型重建效果。

对于孔洞中丢失的点, 无法获得其实测数据。孔洞修补最主要的是利用孔洞区域附近的邻域几何信息, 采用某种算法在保证几何特征的前提下最大程度地恢复出原有模型^[17]。利用径向基函数重建隐式曲面工作最早出现于 1995 年, 并迅速成为点云孔洞修补的主流方法^[18]。另一种主流方法是基于体素进行孔洞修补。由于基于体素的孔洞修复算法需要将点云模型进行体素离散化处理, 计算复杂耗时长且在数据量较大时效率低, 因此本文采用利用径向基函数建立隐式曲面的方法。

孔洞修补工作主要分为两个步骤, 一是提取并确定点云数据的边界与孔洞区域的边界, 二是对已有孔洞区域利用径向基函数进行修补。在提取并确定点云数据边界与孔洞区域边界后, 首先借鉴文献^[19]的思想, 利用法向量的特性, 将最外层点云数据中法向量相近的数据提取出来, 该部分数据有极大概率被认为是墙壁、地板或是天花板等。在此基础上, 以点云数据中散乱点云点为基础, 借鉴文献^[20]的思想, 对无法建立三角网格的散乱点进行标注, 并根据其空间拓扑关系确定孔洞边界点。

在孔洞边界确定后, 利用径向基函数实现孔洞修补。该方法对于较平坦区域效果显著, 但在曲率变化较大的地方难以满足复杂形状要求^[21]。因此, 针对于地面、墙壁及天花板等区域构建基于径向基函数的隐式曲面, 利用最小二乘平滑函数对孔洞多边形进行拟合。基于径向基函数的隐式曲面方程可

表示为:

$$F(n) = \sum_{i=1}^k \omega_i \varphi(n - q_i) + P(n) \quad (11)$$

n 表示隐式表面上的任一点, q_i 为建立曲面所需的采样点, ω_i 是采样点一一对应的权值, $P(n)$ 为根据 n 点坐标 n_x, n_y, n_z 构建的平面, $\varphi(n - q_i)$ 为径向基函数^[22]。最后根据梯度约束将拟合的孔洞多边形向隐式曲面进行调整, 完成孔洞修补。

3. 实验

采用公开数据集 ICL-NUIM^[23] 以及 ElasticFusion 算法提供的样例数据 dyson_lab.klg^[24] (由英国帝国理工学院戴森实验室提供), 对改进后的算法进行实验验证, 分别从模型重建效果和轨迹对比两方面进行评价。实验使用的笔记本电脑配置为 Intel Core i7-7820HK 2.9GHz 的 CPU, 16GB RAM 内存以及 Nvidia GeForce GTX1070 GPU(8GB 显存), 操作系统为 Ubuntu16.04。

3.1 三维重建算法改进实验

首先使用样例数据 dyson_lab.klg 进行三维重建实验, 因该样例数据没有提供轨迹的真值数据, 所以仅能从重建模型的显示效果上进行评价, 如图 7 所示(左侧为原始算法运行效果, 右侧为改进后算法运行效果)。可以明显看出, 在场景中心位置的办公桌附近, 原始算法重建模型出现杂乱不一致现象, 经改进后算法重建模型一致性更强, 显示效

果更逼真。

ICL-NUIM 数据集由英国帝国理工学院与爱尔兰国立梅努斯大学联合提供, 主要用于评价视觉里程计、室内三维重建以及 SLAM 算法等相关研究工作^[25]。选用该数据集进行实验的目的, 一是 ElasticFusion 算法及 ICL-NUIM 数据集均由英国帝国理工学院提供, 使用相同数据集对比算法改进效果; 二是要在避免使用算法闭环优化模块的前提下, 验证对于算法的改进效果。相比于 TUM-RGBD 数据集, ICL-NUIM 数据集的每组数据均为室内部分场景, 很少存在回环, 可以更好地满足实验需求。并且该数据集具备了室内场景的共性特点, 实验结果可以推广应用于其他数据集上。利用 ICL-NUIM 数据集提供的 RGB 图像、深度图像以及相机轨迹的真值文件(ground truth)进行实验对比。该数据集有 8 组数据, 分为卧室场景(livingroom, 简称为 lr_kt)以及办公室场景(officerroom, 简称为 or_kt), 每个场景对应 4 组数据。下面对各组数据进行实验, 分别从模型重建效果与相机轨迹进行分析。在表示模型重建效果图中, 左侧为原始算法运行效果, 右侧为改进后算法运行效果。在相机轨迹分析中, 为更好地对比相机轨迹恢复情况, 将相机轨迹映射至二维 x-y 平面进行比对, 红色为相机轨迹真值, 绿色为原始算法轨迹估计值, 蓝色为改进后算法轨迹估计值。

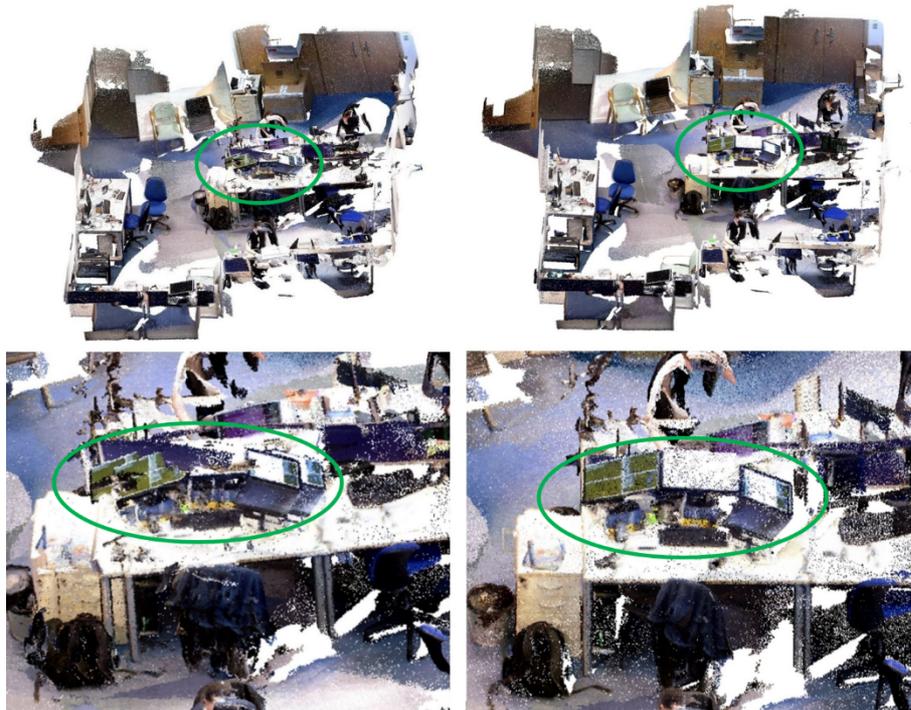


图7 利用 dyson_lab 数据实验对比图 (左: 原始算法效果图, 右: 改进后算法效果图)
 Fig.7 The comparison results using dyson_lab data (left: original algorithm result, right: advanced algorithm result)

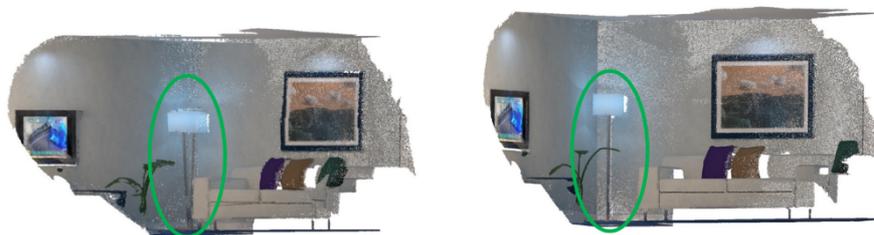


图8(a) 使用 lr_kt0 数据实验对比图
 Fig.8(a) The comparison results using lr_kt0 data

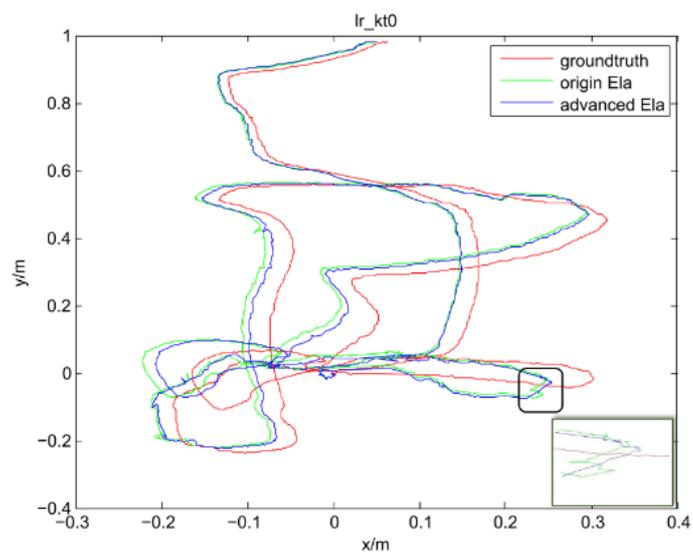


图 8(b) 相机轨迹对比图

Fig. 8(b) The comparison charts of the camera trajectories

图 8(a)所示为使用 lr_kt0 数据的实验效果对比, 图 8(b)所示为相机轨迹对比图。从模型显示效果可以看出, 对于立式台灯的模型重建有明显的提升; 在相机轨迹对比图中, 矩形方框内原始算法的轨迹估计出现跳跃, 该偏差范围所对应的图像帧为多帧立式

台灯图像, 其轨迹恢复的偏差对台灯的重建产生重影等影响。相较于轨迹真值, 虽原始算法与改进后的算法轨迹大致相似, 但对比原始算法在矩形区域的偏差可以看出, 改进后算法的相机轨迹与轨迹真值更加吻合。

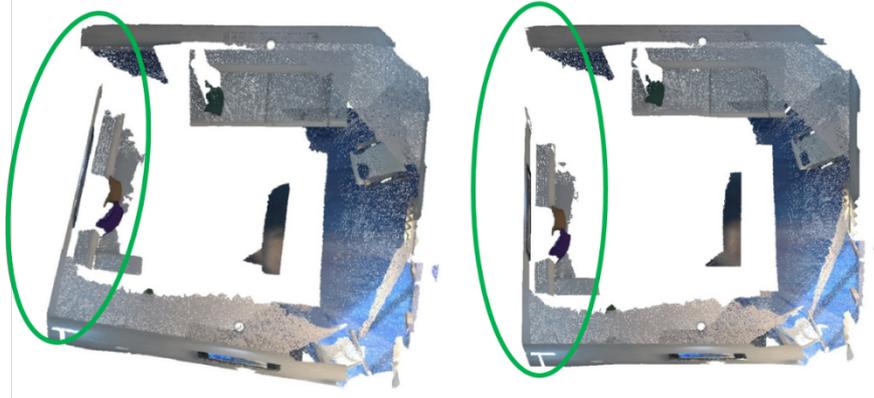


图 9(a) 使用 lr_kt1 数据实验对比图

Fig. 9(a) The comparison results using lr_kt1 data

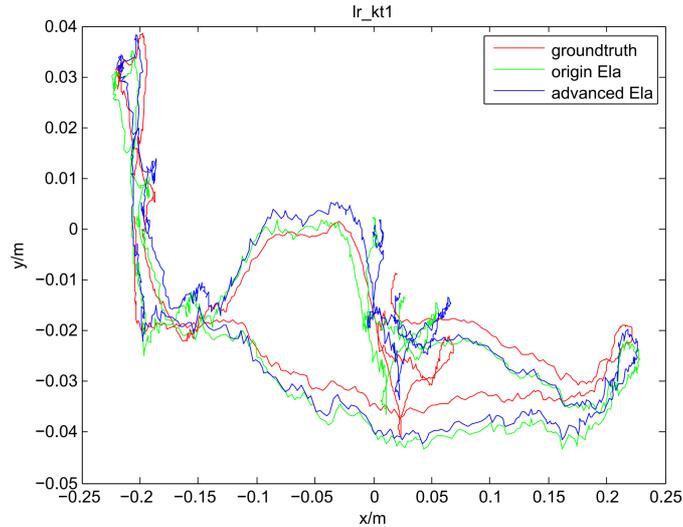


图 9(b) 相机轨迹对比图

Fig. 9(b) The comparison charts of the camera trajectories

图 9(a)所示为使用 lr_kt1 数据的实验效果对比图, 图 9(b)所示为相机轨迹对比图。在对于场景几何结构的重建中, 原始算法没能有效的恢复平面及立体上场景的矩形及

长方体结构, 改进后算法较好地场景进行表达。相机轨迹图中, 改进后算法的轨迹与真实轨迹更加吻合。

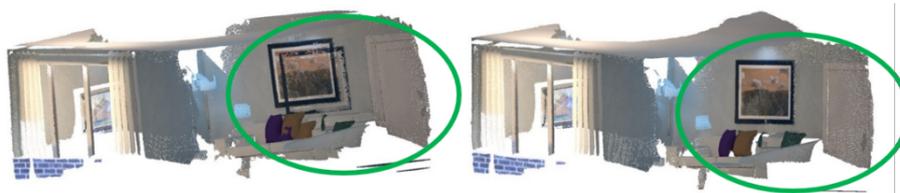


图 10(a) 使用 lr_kt3 数据实验对比图

Fig.10(a) The comparison results using lr_kt3 data

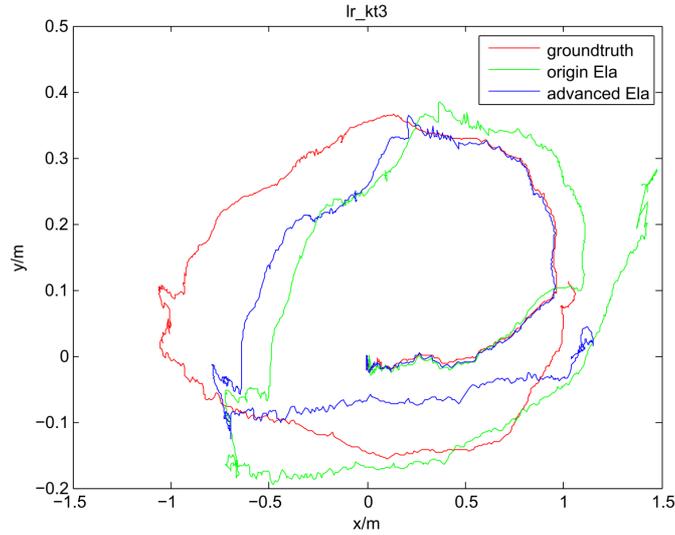


图 10(b) 相机轨迹对比图

Fig.10(b) The comparison charts of the camera trajectories

图 10(a)所示为使用 lr_kt3 数据的实验效果对比,图 10(b)所示为相机轨迹对比图。原始算法在壁画及门的重建效果上有重影,改进后的算法可很好地对场景进行表达。在相机轨迹对比上,随着模型重建的推进,原始算法的轨迹估计有明显偏差。在模型重建的后期,由于原始算法对于位姿估计的累积误差,造成了对同一场景的表达上出现重影。而改进算法的位姿估计则更接近于真值轨迹,使得重建模型表达效果更逼真。

在对 ICL-NUIM 数据集 8 组数据所进行的实验中,有 3 组数据的三维重建结果有明显的提升,避免了出现诸如同一物体重建过程中出现重影、闭合区域无法闭合等不一致现象;有 2 组数据重建效果提升不明显,仅是极小范围或是单一物体的重建效果改进;另 3 组数据的重建效果基本没有提升。但在相机轨迹估计方面,相较于原始算法,

8 组数据中,改进后的算法所估计的相机轨迹均与轨迹真值更加相近或吻合。

从上述实验表明,相比于原始算法,改进后的算法在模型显示效果以及相机轨迹估计方面都有提升。3 组重建效果基本没有提升的实验中,其重建的场景均是小范围、几何特征信息不丰富、较空旷的场景且不存在闭环约束。因此,在重建该类场景时,原始算法及改进后的算法的重建效果都有待提高。

3.2 点云模型孔洞修补实验

使用改进算法生成的点云模型作为实验数据,对平坦区域存在的孔洞进行修补实验。选择使用 dyson_lab.klg 数据生成的三维模型作为实验数据,实验对比如图 11 所示,左侧为原始点云模型,右侧为经过孔洞修补后较为完整的点云模型。

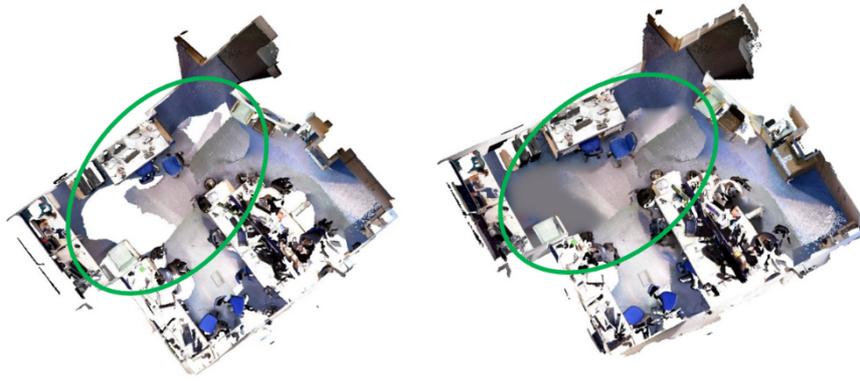


图 11 平坦区域点云孔洞修补实验对比图

Fig.11 The comparison results of points cloud hole repairing in the flat area



图 12 因物体遮挡产生的点云孔洞修补实验对比图

Fig.12 The comparison results of points cloud hole repairing caused by object occlusion

对于因物体遮挡产生的孔洞，选用由数据 lr_kt0 所生成的点云模型作为实验数据，实验对比如图 12 所示

由上述两个实验可以看出，利用径向基函数生成隐式曲面的方法可以较好地修补三维重建模型中存在的孔洞。该方法对于较为平坦的孔洞修补鲁棒性高，但是对于修补曲率较大或是不规则弯曲的物体等产生的点云孔洞有一定的局限性。

4. 结束语

本文工作主要针对 ElasticFusion 算法在三维重建过程中出现的局部模型重建出现错乱的情况进行改进，并对三维模型中因物体遮挡、物体反射率、光强度变化等原因造成的点云孔洞进行修补，以提升室内场景三维模型的重建效果。从原始算法中相机位姿估计的 ICP 算法入手，对匹配点对选取策略进行了改进。在利用投影法选取匹配点对的基础上，确定搜索区域，通过遍历计算区域内所有候选点与待匹配点的法向量乘积得

到匹配点对。在 ICP 算法的初始阶段通过引入简单的数学计算选取出更为准确的匹配点进行位姿解算，在不过多消耗计算资源的前提下保证了算法的实时性。对比实验表明，在几何特征信息丰富的场景中，改进后的算法无论在模型显示效果方面还是在相机轨迹估计方面，都有较为明显的提升。改进的算法在保证算法实时性的基础上，可以得到更加准确恰当的匹配点，提高了位姿解算的精度。但在场景中出现线面特征不明显（例如场景中大量边界平行于墙面、地面的物体）或是场景几何特征信息不丰富的情况下，仍无法达到理想效果。同时，使用改进后算法重建的模型，利用径向基函数构建隐式曲面的方法实现点云模型的孔洞修补，进一步提升了重建模型的显示效果。

参考文献 (References)

- [1] Mur-Artal R, Montiel J M M, Tardós J D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J]. *IEEE Transactions on Robotics*, 2017, 31(5):1147-1163.

- [2] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-Scale Direct Monocular SLAM[C]. *European Conference on Computer Vision*. Springer, Cham, 2014:834-849.
- [3] Forster C, Zhang Z, Gassner M, et al. SVO: Semidirect Visual Odometry for Monocular and Multicamera Systems[J]. *IEEE Transactions on Robotics*, 2017, 33(2):249-265.
- [4] Newcombe R A, Izadi S, Hilliges O, et al. KinectFusion: Real-time dense surface mapping and tracking[C]. *IEEE International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2011:127-136.
- [5] Whelan T, Kaess M, Fallon M, et al. Kintinuous: Spatially Extended KinectFusion[J]. *Robotics & Autonomous Systems*, 2012, 69(C):3-14.
- [6] Whelan T, Leutenegger S, Moreno R S, et al. ElasticFusion: Dense SLAM Without A Pose Graph[C]. *Robotics: Science and Systems*. 2015.
- [7] Izadi S, Stamminger M. Real-time 3D reconstruction at scale using voxel hashing[J]. *Acm Transactions on Graphics*, 2013, 32(6):169.
- [8] Dai A, Izadi S, Theobalt C. BundleFusion: real-time globally consistent 3D reconstruction using on-the-fly surface re-integration[J]. *Acm Transactions on Graphics*, 2017, 36(4):76a.
- [9] Zeng M, Zhao F, Zheng J, et al. A memory-efficient kinectfusion using octree[C]. *International Conference on Computational Visual Media*. 2012:234-241.
- [10] Zhu Xiaoxiao, Cao Qixin, Yang Yang, et al. An Improved KinectFusion 3D Reconstruction Algorithm[J]. *Robot*, 2014, 36(2):129-136. (朱笑笑, 曹其新, 杨扬,等. 一种改进的 KinectFusion 三维重构算法[J]. 机器人, 2014, 36(2):129-136.)
- [11] Ye Qin, Yao Yahui, Gui Popo. Kinect Point Cloud Registration Method Based on Epipolar and Point-to-Plane Constraints[J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(9):1271-1277. (叶勤, 姚亚会, 桂坡坡. 基于极线及共面约束条件的 Kinect 点云配准方法[J]. 武汉大学学报(信息科学版), 2017, 42(9):1271-1277.)
- [12] Newcombe R A, Fox D, Seitz S M. DynamicFusion: Reconstruction and tracking of non-rigid scenes in real-time[C]. *Computer Vision and Pattern Recognition*. IEEE, 2015:343-352.
- [13] Liu Haomin, Zhang Guofeng, Bao Hujun. A Survey of Monocular Simultaneous Localization and Mapping[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2016, 28(6):855-868. (刘浩敏, 章国锋, 鲍虎军. 基于单目视觉的同时定位与地图构建方法综述[J]. 计算机辅助设计与图形学学报, 2016, 28(6):855-868.)
- [14] Glocker B, Shotton J, Criminisi A, et al. Real-Time RGB-D Camera Relocalization via Randomized Ferns for Keyframe Encoding[J]. *IEEE Trans Vis Comput Graph*, 2015, 21(5):571-83.
- [15] Rusinkiewicz S. Efficient variant of the ICP algorithm[J]. *Proc Ofdim*, 2001.
- [16] Tan Xin. Global Optimization for 3D Reconstruction based on KinectFusion[D]. Zhejiang University. 2015. (谭歆. KinectFusion 三维重建的再优化[D]. 浙江大学, 2015.)
- [17] Liu Yongmei, Li Fengxia, Lei Zhengchao, et al. New Hole Filling Algorithm Based on Boundary-feature Growing[J]. *Journal of System Simulation*, 2014, 26(9):1916-1921. (刘咏梅, 李凤霞, 雷正朝,等. 基于边界特征增长的孔洞修补算法[J]. 系统仿真学报, 2014, 26(9):1916-1921.)
- [18] Savchenko V V, Pasko A A, Okunev O G, et al. Function Representation of Solids Reconstructed from Scattered Surface Points and Contours[J]. *Computer Graphics Forum*, 1995, 14(4):181-188.
- [19] Nüchter A, Hertzberg J. Towards semantic maps for mobile robots[J]. *Robotics & Autonomous Systems*, 2008, 56(11):915-926.
- [20] Yan Haiping, Wu Lushen, Chen Huawei. Hole repairing algorithm in scattered point cloud based on radial basis function[J]. *Computer Engineering And Design*, 2014, 35(4):1253-1257. (晏海平, 吴禄慎, 陈华伟. 基于径向基函数的散乱点云孔洞修复算法[J]. 计算机工程与设计, 2014, 35(4):1253-1257.)
- [21] Chen Mao. A Hole Repairing Algorithm for 3D Skull Model[D]. Zhejiang University of Tcehnology, 2009. (陈矛. 三维颅骨模型的孔洞修补算法研究[D]. 浙江工业大学, 2009.)

[22] Yan Haiping. Research of Boundary Extraction and Hole Repairing Algorithm Based on Scattered Point Cloud[D]. Nanchang University. 2014. (晏海平. 散乱点云边界提取及孔洞修复算法研究[D]. 南昌大学, 2014.)

[23] ICL-NUIMDataset.

<https://www.doc.ic.ac.uk/~ahanda/VaFRIC/iclnuim.html>

[24] dyson_lab.klg.

<https://wp.doc.ic.ac.uk/robotvision/>

[25] Handa A, Whelan T, McDonald J, et al. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM[C]. IEEE International Conference on Robotics and Automation. IEEE, 2014:1524-1531.

ElasticFusion for Indoor 3D Reconstruction with An Improved Matching Points Selection Strategy

WANG Weiqi, YOU Xiong, YANG Jian, LI Qin

Institute of Geospatial Information, Information Engineering University, Zhengzhou 450052, China

Abstract: Real-time and high-quality 3D reconstruction of indoor scenes has been a research focus in the field of augmented reality and robotics. However, the 3D reconstruction methods using RGB-D sensors suffer from weaknesses such as poor local model quality and producing holes in points cloud models. The two key factors that affect the quality of 3D reconstruction are the accuracy of the pose parameters derived from points cloud registration and the accuracy of loop closure. In this paper, an improved ElasticFusion algorithm is proposed to achieve better reconstruction quality in real time, which is achieved by improving the strategy of searching matching points in the ICP algorithm. Furthermore, radial basis functions are used to construct implicit surfaces in order to fill holes in the points cloud models that are generated in the previous step. In the end, benchmarks such as ICL-NUIM are used to evaluate the proposed algorithm. The experimental results have shown that our algorithm can significantly improve the quality of model reconstruction and the accuracy of the camera trajectory estimation.

Keywords: ElasticFusion; 3D reconstruction; ICP algorithm; Hole repairing

First Author: WANG Weiqi, postgraduate student, specialize in robotic mapping. E-mail: 809741461@qq.com

***Corresponding author:** YANG Jian, Ph D, Lecturer, specialize in spatial data mining and location-based services. E-mail: jyangtum@qq.com

Foundation Support: The National Key R&D Program of China, No.2017YFB0503500; The Project of Science and Technology Innovation of Henan Province, No.142101510005.