

# 基于 RGB-D 图像的具有滤波处理和位姿优化的同时定位与建图

熊军林, 王 婵

(中国科学技术大学自动化系, 安徽合肥 230027)

**摘要:** RGB-D 相机能够同时获得彩色图像和深度图像, 广泛用于同时定位与建图(SLAM)的研究. 本研究针对 RGB-D SLAM 方法进行了两方面的改进: 一方面, 改进点云滤波方法, 从而更有效地去除 RGB-D 相机数据中的噪声和冗余; 另一方面, 采用 ICP 算法提高相机位姿估计的精度, 从而提高估计的相机运动轨迹的精度. 在公开的数据集上对提出的 RGB-D SLAM 方法进行实验验证, 结果表明, 该方法能够有效提高移动机器人自主定位与建图的精度.

**关键词:** RGB-D 相机; RGB-D SLAM; ICP 算法; 滤波; 变换矩阵

**中图分类号:** TP391      **文献标识码:** A      doi: 10.3969/j.issn.0253-2778.2017.08.005

**引用格式:** 熊军林, 王婵. 基于 RGB-D 图像的具有滤波处理和位姿优化的同时定位与建图[J]. 中国科学技术大学学报, 2017, 47(8): 665-673.

XIONG Junlin, WANG Chan. Simultaneous localization and mapping based on RGB-D images with filter processing and pose optimization[J]. Journal of University of Science and Technology of China, 2017, 47(8): 665-673.

## Simultaneous localization and mapping based on RGB-D images with filter processing and pose optimization

XIONG Junlin, WANG Chan

(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

**Abstract:** RGB-D camera can capture color and depth images simultaneously, and is widely used for simultaneous localization and mapping (SLAM) research. In this article, The RGB-D SLAM method was improved from two aspects. Firstly, the point cloud filter method was improved to more effectively decrease the noise and redundancy of RGB-D camera data; secondly, an ICP algorithm was used to improve the estimated accuracy of the pose transformation matrix and the trajectories of camera movement. The proposed RGB-D SLAM method was verified on public datasets. The experimental results demonstrate that our RGB-D SLAM method can effectively improve the accuracy of the autonomous positioning and mapping of robots.

**Key words:** RGB-D camera; RGB-D SLAM; ICP algorithm; filter; transformation matrix

### 0 引言

移动机器人的同时定位与建图(SLAM)是指机

器人在未知环境中从一未知位置开始移动, 在移动的过程中根据控制信息与传感器观测到的数据进行自身位置估计, 同时构建环境地图<sup>[1]</sup>. 随着移动机器

收稿日期: 2016-09-01; 修回日期: 2017-01-01

基金项目: 国家自然科学基金(61374026)资助.

作者简介: 熊军林(通讯作者), 男, 1977年生, 博士/教授. 研究方向: 自动控制理论. E-mail: xiong77@ustc.edu.cn

人研究与应用的不断发展,SLAM 已经成为移动机器人领域研究的热点问题之一.SLAM 初期,由于激光雷达具有测量距离精度高的优势,故在 SLAM 研究领域备受青睐.如 Vazaquez-Martin 等通过结合自适应协方差矩阵和 EKF-SLAM 算法<sup>[2]</sup>,在机器人位姿和地图更新阶段采用更多的陆标,提高陆标更新频率,从而提高同时定位与建图的精度<sup>[3]</sup>.Hamzaoui 等通过使用单指令/多数据扩展指令集,对激光扫描匹配环节进行并行化处理,提高同时定位与建图的效率<sup>[4]</sup>.周武等提出 GOE-SLAM 算法,克服了系数特征地图环境信息不详尽的缺点,获得全局观测地图模型<sup>[5]</sup>.由于激光雷达存在以下缺陷:测量误差随测量距离增大而增大;二维激光数据的可读性差;三维激光体积较大,而且价格昂贵<sup>[6]</sup>;因此随着机器人应用场景复杂度的提高,激光雷达已经难以满足应用需求.

由于视觉传感器能够获得更加丰富的环境信息,体积小且价格低廉,因此基于视觉的同时定位与建图逐渐受到关注.基于视觉的同时定位与建图方法按照传感器类型可分为 4 类:立体相机<sup>[7]</sup>、单目相机<sup>[8]</sup>、全景相机<sup>[9]</sup>和 RGB-D<sup>[10-11]</sup>.使用立体相机存在尺度不确定、计算量较大的问题,使用单目相机很难获得地图的绝对尺度,使用全景相机获得的环境信息存在很大程度的畸变<sup>[8]</sup>,而 RGB-D 相机因其采集数据的速度快和测量的精度较高被广泛应用于基于视觉的同时定位与建图的研究领域.Henry 等于 2010 年首次提出 RGB-D SLAM<sup>[12]</sup>,并于 2012 年进行了优化<sup>[13]</sup>.Endres 等于 2012 年对 RGB-D SLAM 方法的准确性和鲁棒性进行了评估,同时使用输出的点云构建 3D 栅格地图,便于跟踪定位<sup>[14]</sup>、路径规划<sup>[15]</sup>和导航<sup>[16]</sup>直接使用<sup>[17]</sup>.通常情况下,RGB-D SLAM 方法包括前端和后端两部分<sup>[17]</sup>.前端的主要任务是对 RGB-D 相机获得的彩色图像和深度图像进行预处理,从而获得点云数据;首先根据 RGB-D 相机提供的彩色图像提取图像的特征点,进行特征匹配;然后根据特征匹配的结果以及

RGB-D 相机提供的深度图像估算两帧图像对应相机位姿的变换关系.后端的主要任务是通过图优化方法,减小轨迹估计过程中出现的非线性误差,从而获得更加精确的相机位姿.针对 RGB-D SLAM 前端,Henry 等并未对 RGB-D 相机数据的噪声和冗余进行处理<sup>[13]</sup>,Endres 等使用的 RGB-D SLAM 方法获得的相机位姿变换矩阵依然存在误差<sup>[18]</sup>,因此 RGB-D SLAM 前端仍有许多需要改进之处.

本文研究的重点:一是改进点云滤波方法,对预处理 RGB-D 相机获得的点云数据进行滤波处理,从而去除点云数据中的噪声和冗余;二是将经过滤波处理之后的点云数据和 RANSAC(random sample consensus)算法<sup>[19]</sup>估算的初始位姿变换矩阵作为 ICP(iteratively closet point)算法<sup>[20]</sup>的输入,进一步估算相机位姿变换矩阵,提高位姿变换矩阵的精度.文中首先对 RGB-D SLAM 方法进行理论介绍,然后在国际公认的由 Sturm 等<sup>[21]</sup>提供的数据集<sup>[22]</sup>上验证本文提出的 RGB-D SLAM 方法,实验结果表明,本文的 RGB-D SLAM 方法的精度更高.

## 1 算法介绍

本文 RGB-D SLAM 系统框架如图 1 所示.该方法的特点是在进行初始位姿变换估计和筛选关键帧之间采用点云滤波处理和 ICP 算法,从而进一步精确估计相机位姿变换矩阵.该方法包括以下 5 部分:①根据获取的彩色图像提取图像特征,进行特征点匹配,同时对彩色和深度图像进行预处理,获得点云数据.②根据特征匹配结果,结合特征点的深度图像,使用 RANSAC 算法剔除误匹配对和粗略估计两帧图像对应的相机位姿的初始变换矩阵.③使用滤波方法处理原始点云数据,并将初始变换矩阵和滤波点云数据作为 ICP 算法的输入,进一步精确计算相机位姿变换矩阵.④根据变换误差的大小,筛选出关键帧.⑤使用 G2O(general graph optimization)算法<sup>[23]</sup>减小匹配过程中的非线性误差,得到相机的图优化位姿.RGB-D SLAM 系统输出的是与真实环

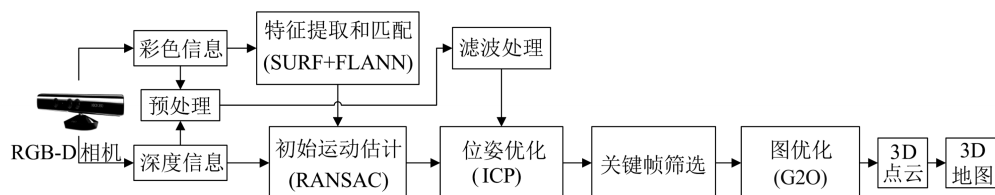


图 1 RGB-D SLAM 框架

Fig.1 Framework of RGB-D SLAM

境一致的 3D 彩色点云模型.使用 Octomap 库<sup>[24]</sup>,利用输出的点云数据生成一张 3D 栅格地图.下面对 RGB-D SLAM 方法中的各部分进行详细介绍.

### 1.1 数据预处理

首先根据彩色图和深度图的信息,即图像中的像素坐标  $[u, v, d]$  ( $u, v$  分别代表像素的列数和行数,  $d$  代表深度数据),计算出每一个点的 3D 信息.计算公式如下:

$$\left. \begin{aligned} x &= \frac{z}{f_y}(u - c_x) \\ y &= \frac{z}{f_x}(v - c_y) \\ z &= \frac{d}{c_{scal}} \end{aligned} \right\} \quad (1)$$

式中,  $c_x$  与  $c_y$  是指相机的光圈中心,  $f_x$  与  $f_y$  是指相机在  $x$  和  $y$  两个轴上的焦距,  $c_{scal}$  表示深度缩放因子,  $[x, y, z]$  代表点云在三维空间的坐标.

### 1.2 特征提取和匹配

特征提取可以使用 OpenCV<sup>[25]</sup> 中的 SIFT<sup>[26]</sup>、SURF<sup>[13]</sup>、ORB<sup>[27]</sup> 算法. SIFT 算法稳定性最好,但是其计算代价非常大; SURF 算法稳定性适中,实时性也比较好<sup>[11]</sup>; ORB 算法实时性最好,但是稳定性差<sup>[17]</sup>. 本文提出的 RGB-D SLAM 方法先选用 SURF 算法进行特征提取; 然后使用 FLANN (fast library for approximate nearest neighbors) 算法<sup>[28]</sup> 进行图像特征匹配. SURF 特征提取算法原理如下:

(I) 构建 Hessian 矩阵. 给定图像  $I$  中的一个点  $X(x, y)$ , 在点  $X$  处, 尺度为  $\sigma$  的 Hessian 矩阵如下:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (2)$$

$$L(X, t) = \frac{\partial^2 g(t)}{\partial X^2} * I(X) \quad (3)$$

式中,  $X(x, y)$  为图像  $I$  中的点,  $I(X)$  为像素值,  $L(X, \sigma)$  是一副图像在不同解析度下的表示,  $g(t)$  为高斯函数,  $t$  为高斯方差. Bay 等提出采用框式滤波器代替  $L$  在  $xx, yy, xy$  三个方向上的近似值, 分别记作  $D_{xx}, D_{yy}, D_{xy}$ <sup>[29]</sup>. Hessian 矩阵行列式计算简化如下:

$$\det(H_{appr}) = D_{xx}D_{yy} - \omega D_{xy} \quad (4)$$

式中,  $\omega$  是为平衡准确值与近似值间的误差采用的权值系数, 与尺度  $\sigma$  有关.

(II) 构建尺度空间. 图像的尺度空间是这幅图像在不同解析度下的表示, 如式(3). 尺度空间可以

表述为一个图像金字塔. 在构建尺度金字塔时, 图像的尺寸大小保持不变, 采用不同尺寸的框式滤波器对图像进行处理, 其尺寸计算公式如下:

$$F_{size} = 3(2^{oct}dis + 1) \quad (5)$$

式中,  $F_{size}$  为滤波器尺寸,  $oct, dis$  初始值为 1.

(III) 提取特征点. 使用与该尺度图层图像解析度相应大小的滤波器进行检测. 假设 Hessian 矩阵的行列式函数记为  $H(x, y, \sigma)$ , 且  $X = (x, y, \sigma)$ , 根据泰勒展开式可得

$$H(X) = H + \frac{\partial H^T}{\partial X}X + \frac{1}{2}X^T \frac{\partial^2 H}{\partial X^2}X \quad (6)$$

令  $H(x, y, \sigma) = 0$ , 可得极值如下:

$$\hat{x} = -\frac{\partial^2 H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad (7)$$

如果极值  $\hat{x}$  在  $x, y, \sigma$  三个方向中的值小于预设值, 则舍弃. 增大预设值, 使得检测到的特征点数量减少, 最终得到稳定特征点的位置和尺度.

FLANN 算法原理如下.

(I) 根据数据集的分布特点建立 Kd-tree 搜索树, 通过搜索树将数据点划分为特定的几个部分.

(II) 在 Kd-tree 中查询欧式距离最近的匹配对. 欧式距离计算公式如下:

$$d(p, q) = \langle D_p - D_q, D_p - D_q \rangle \quad (8)$$

式中,  $p$  和  $q$  是特征点,  $D_p$  和  $D_q$  是对应的特征分向量,  $d(p, q)$  是两个特征点之间欧式距离.

### 1.3 初始运动估计

为了提高 RGB-D SLAM 方法的精度, 本文使用 RANSAC 算法结合 RGB-D 相机提供的深度图像消除误匹配<sup>[30]</sup>; 同时得到一个粗略的相机位姿变换矩阵. RANSAC 算法的具体步骤如下:

(I) 从特征匹配对中随机选取三组匹配对, 估计参数模型.

(II) 用(I)中估计出的参数模型去测试其他所有的匹配对, 如果某个匹配对也适用于这个模型, 则将它作为局内点.

(III) 如果局内点足够多, 则估计的模型合理; 否则, 重复(I)-(III).

(IV) 用所有的局内点重新估计模型.

(V) 重复(I)-(V), 直至局内点数满足要求、模型错误率较低.

### 1.4 滤波处理

常用的点云滤波方法<sup>[31]</sup> 主要包括法向滤波、网格滤波、随机滤波和直通滤波 4 种方法. 其中, 直通滤波主要是作为一种辅助滤波方法与其他滤波方法

同时使用,去除点云中比较模糊的、不准确的数据信息,以提高后续相机位姿变换矩阵的精度.下面将详细介绍其他 3 种滤波方法.

法向滤波主要是根据每个点的法线在角度空间的位置进行角度均匀采样,防止出现失真.具体步骤如下.

(I) 计算每个点的法线.对于 3 维空间中的每一个点  $p$ ,先在其周围选取距离最近的  $k$  个点;然后计算协方差矩阵  $C$ .根据主元分析法,对应于协方差矩阵  $C$  最小特征值的特征向量即为  $p$  的法向量.

$$C = \frac{1}{k} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T \quad (9)$$

(II) 根据法线在角度空间的位置进行均匀采样.对法线在角度空间的均匀采样等价于对单位球面进行均匀采样.

网格滤波是用一个 3 维立体栅格的重心代表栅格中的所有像素点,达到降低点云数量和保留空间特征的目的.具体步骤如下:

(I) 输入的点云数据中随机选取一个点  $p$  作为一个顶点,构建给定边长的立方体,并进行扩展形成网格覆盖所有的点云数据.

(II) 根据每一个立方体  $A$  包含的点云数  $s$  计算该立方体的重心,用重心代替其包含的所有点云数据.重心计算公式如下:

$$\tilde{x} = \frac{1}{s} \sum_{(x,y,z) \in A} x \quad (10)$$

$$\tilde{y} = \frac{1}{s} \sum_{(x,y,z) \in A} y \quad (11)$$

$$\tilde{z} = \frac{1}{s} \sum_{(x,y,z) \in A} z \quad (12)$$

随机滤波主要采用均匀采样的思想,在很大程度上保留空间特征,且点云处理速度相对较快.假设从  $N$  个点云中随机抽取  $n$  个点云,其中  $0 \leq n \leq N$ ,且每次生成随机数  $V$  都是相互独立的.步骤如下:

(I) 生成随机数  $V (V \sim U[0,1])$ .

(II) 找出满足  $(N-n)^{s+1} \leq N^{s+1}V$  最小的变量  $s (s \geq 0)$ .

(III) 从  $N$  个点云中选取第  $s+1$  个点云,然后令  $N=N-s-1, n=n-1$ ,如果  $n > 0$ ,则重复步骤(I)-(III).

RGB-D 相机感知的清晰度与感知距离有关,即在一定距离范围内,距离越近,对环境的感知越清晰;距离越远,对物体的区别度越低.本文首先结合直通滤波方法去除深度距离较近和过远的点云数

据,然后用网格滤波或随机滤波方法处理点云数据,这样可以较好滤除点云数据中的噪声和冗余.

此外,因分段滤波可以根据光谱距离的大小分段处理原始光谱,可以去除随机噪声和减少冗余矿物质,很大程度上提高了矿物识别的精度<sup>[32]</sup>,故本文对点云数据进行分段滤波处理.分段滤波处理主要是在 RGB-D 相机感知距离范围内将深度距离等分为若干段,每一段对应一个随机滤波因子  $S(d)$ ,然后相应区间进行随机滤波,最终得到分段随机滤波后的结果.

$$S(d) = \alpha_i, d \in [d_i, d_{i+1}] \quad (13)$$

式中,  $i=1,2,\dots,n$ ,表示第  $i$  区间段,  $d$  表示深度距离.

### 1.5 位姿优化

本文以 RANSAC 算法估算出的变换矩阵作为 ICP 算法的初始预估矩阵,同时用滤波处理之后的点云作为 ICP 算法的输入,以达到精确计算变换矩阵的目的.ICP 算法简介如下:

(I) 根据源点云  $P_s$  中点的坐标,在曲面  $S$  中搜索对应的最近点集  $P_t$ .

(II) 计算两个点集的重心坐标,实行点集中心化从而生成新的点集.

(III) 根据新的点集计算正定矩阵  $N$ ,同时计算矩阵  $N$  的最大特征值和最大特征向量.

(IV) 因最大特征向量与残差平方和最小的旋转四元数等价,故将四元数转换成  $R$ .

(V) 在  $R$  确定之后,  $t$  可通过两个点集对应的坐标系的重心点与  $R$  来确定.

(VI) 根据  $R$  和  $t$ ,计算源点云  $P_s$  变换之后的点云  $\tilde{P}_t$ .

(VII) 计算  $P_t$  与  $\tilde{P}_t$  距离的平方和  $\gamma_{i+1}$ ,若连续两次迭代的距离平方和之差的绝对值  $\Delta\gamma = |\gamma_{i+1} - \gamma_i| < \epsilon$ ,则停止迭代,否则重复步骤(I)-(VII),直至满足  $\Delta\gamma < \epsilon$ .

### 1.6 关键帧筛选

由于用所有的图像进行拼接时会产生很大的累积误差<sup>[33]</sup>,因此本文考虑使用关键帧的方法进行拼接,提高系统的稳定性.同时,使用关键帧拼接还能显著地减小计算量和存储点云数据所消耗的空间.

首先,根据 ICP 算法的返回值计算相机位姿变化的幅度大小,然后判断变化幅度是否在本文设定阈值范围  $[a,b]$  内.如果变化幅度小于  $a$ ,或后续选取得关键帧与上一帧的关键帧是连续的两帧,则认

为变化的幅度太小,不予考虑.如果变化幅度大于  $b$ , 则变化的幅度太大,很有可能出现拼接错误的情况,也不予考虑.变化幅值在  $[a, b]$  之间的则作为关键帧,用于后续的后端优化环节.

### 1.7 图优化

图优化算法主要有 TORO 算法<sup>[34]</sup>和 G2O 算法.2012 年,Noda 等<sup>[35]</sup>用实验证明,无论是实时性还是算法的精度,G2O 算法都优于 TORO 算法.因此本文选用 G2O 算法进行图优化,其代价函数如下所示:

$$F(\mathbf{X}) = \sum_{i,j \in M} e(\mathbf{X}_i, \mathbf{X}_j, \mathbf{Z}_{ij})^T \Omega_{ij} e(\mathbf{X}_i, \mathbf{X}_j, \mathbf{Z}_{ij}) \quad (14)$$

$$\mathbf{X}^* = \underset{\mathbf{X}}{\operatorname{argmin}} F(\mathbf{X}) \quad (15)$$

式中,  $\mathbf{X} = (\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_n^T)^T$  是机器人位姿  $\mathbf{X}_i$  的向量,  $M$  是对应的时间序列,  $\mathbf{Z}_{ij}$  和  $\Omega_{ij}$  分别代表  $\mathbf{X}_i$  相对于  $\mathbf{X}_j$  的位姿约束和信息矩阵,  $e(\mathbf{X}_i, \mathbf{X}_j, \mathbf{Z}_{ij})$  代表位姿  $\mathbf{X}_i$  和位姿  $\mathbf{X}_j$  满足约束  $\mathbf{Z}_{ij}$  的程度.

RGB-D 相机的位姿模型下:

$$\mathbf{T} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (16)$$

针对该模型,RGB-D SLAM 的关键是找到一对  $(R, t)$ , 使得代价函数  $F(R, t)$  得到最小值,即

$$F(R, t) = \min_{R, t} \sum_i \|Rp_i + t - q_i\| \quad (17)$$

式中,  $p_i$  为相机运动之前一帧图像中的匹配点,  $q_i$  为相机运动之后一帧图像中的匹配点.

## 2 实验及结果分析

本文使用 Kinect(一种 RGB-D 相机)作为传感

器对文中的 RGB-D SLAM 方法进行分析 and 评估. Kinect 的可视范围描述如下:  $(0, 0.4 \text{ m}]$  和  $[8 \text{ m}, +\infty)$  为不确定区域;  $(0.4 \text{ m}, 0.8 \text{ m})$  表示红外摄像头与物体的距离太近;  $[0.8 \text{ m}, 4 \text{ m}]$  为 Kinect 正常的可视区域,数据比较精确;  $(4 \text{ m}, 8 \text{ m})$  表示红外摄像头与物体的距离太远.一般 Kinect 的最远可视距离是 4.096 m, 0 则表示深度值不确定.本文使用的 Kinect 深度值范围是 1.22~3.81 m.本文首先通过对连续两帧 Kinect 数据进行特征提取和匹配,验证特征提取和匹配算法;然后将 Kinect 深度距离等分成 3 段对点云数据进行分段滤波处理,同时使用其他滤波方法处理点云数据,并对滤波结果进行分析;最后对 RGB-D SLAM 方法进行整体评估.

### 2.1 特征提取和匹配结果

本文使用如图 2 和 3 所示的两个相邻相机位姿下的 Kinect 数据阐述特征提取和匹配的过程.其中,图(a),(b)和(c)分别表示 Kinect 获取的彩色图(印刷为灰度图)、深度图和经过预处理之后获得的点云数据.对于深度图像,颜色越深表示深度距离越小,越浅表示深度距离越大.首先,本文使用 SURF 算法对图 2(a)和图 3(a)进行特征提取,然后使用 FLANN 算法进行特征匹配,其结果见图 4(a).其中,圆圈表示提取的特征点,线段表示图 2(a)和图 3(a)中特征点的一一对应关系.实验结果表明,图 4(a)中第一个抽屉的把手误匹配到第二个抽屉的把手处,而使用 RANSAC 算法之后可以有效消除该误匹配对,如图 4(b).故本文同时使用 SURF 算法与 RANSAC 算法,提高了特征匹配的精度.

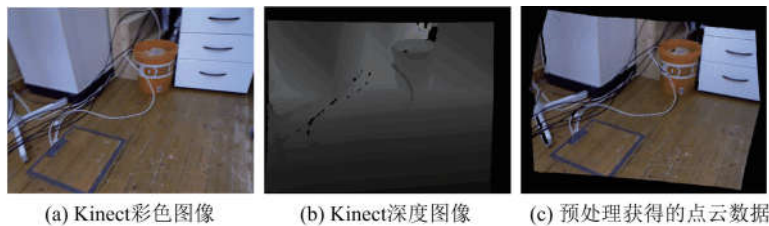


图 2 Kinect 数据 1  
Fig.2 Kinect data 1

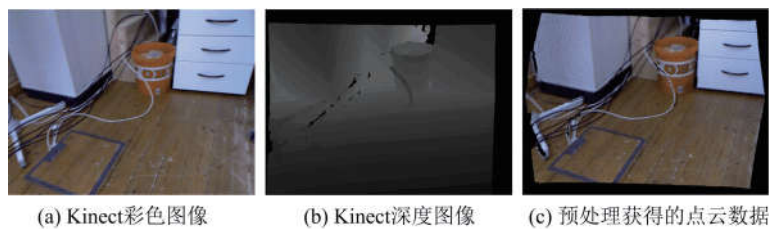


图 3 Kinect 数据 2  
Fig.3 Kinect data 2

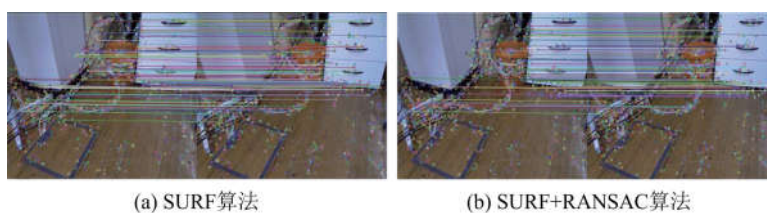


图 4 特征提取和匹配

Fig.4 Feature extraction and matching

## 2.2 滤波实验结果

本文对各种滤波方法进行了对比,如图 5 所示.其中,图 5(a)是 Kinect 拍摄到的真实环境的彩色图像(印刷为灰度图),图 5(b)是 Kinect 拍摄到的真实环境的深度图像,颜色越深表示物体距离相机的距离越远,图 5(c)是经过预处理之后获得的点云数据,图 5(d)是使用法向滤波之后得到的点云数据,

图 5(e)是使用网格滤波得到的点云数据,图 5(f)是使用随机滤波得到的点云数据,图 5(g)和(h)分别是先对原始点云直通滤波处理,去除模糊的点云数据,然后再分别进行网格滤波和随机滤波处理之后的结果,图 5(i)是使用分段滤波之后的点云数据.同时,本文还对使用各种滤波方法处理点云数据的效率进行了实验对比,结果如表 1 所示.

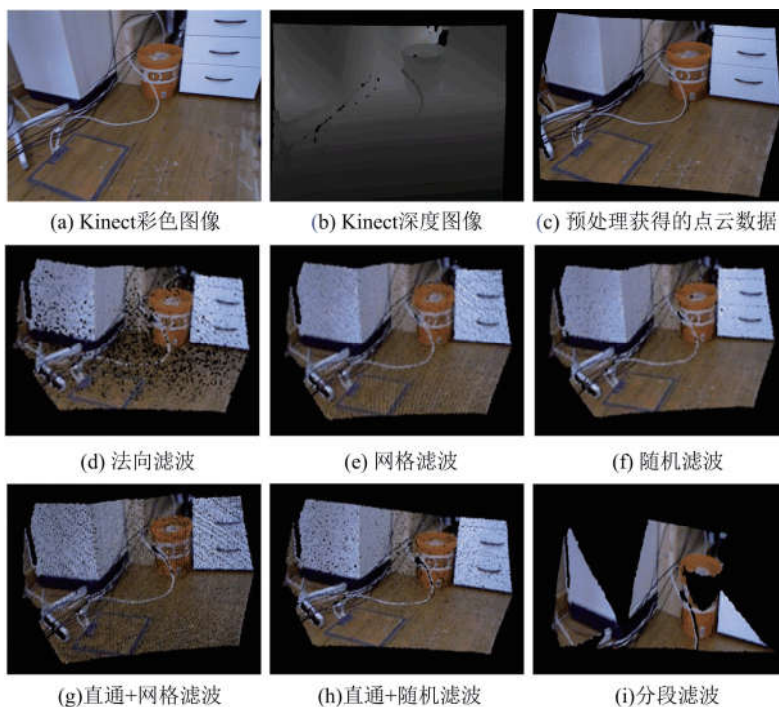


图 5 点云滤波处理

Fig.5 Point cloud filtering processing

表 1 滤波处理时间

Tab.1 The filter processing time

滤波方法	法向滤波	网格滤波	直通+网格滤波	随机滤波	直通+随机滤波	分段滤波
时间( $10^{-3}$ s)	27625.8	52.2	52.5	16.3	16.5	24.5

图 5 的实验结果表明,法向滤波<sup>[36]</sup>和随机滤波<sup>[31]</sup>的效果几乎相同,网格滤波<sup>[31]</sup>与直通滤波<sup>[31]</sup>结合网格滤波和随机滤波的效果相似,而分段滤波则可以较好地滤除深度距离较大的数据,从而获得更加精确的数据.表 1 的结果表明,使用法向滤波进

行点云滤波处理相比其他 5 种滤波方法需要付出更大的时间代价,因此本文将对采用其他 5 种滤波方法的 RGB-D SLAM 方法的精度分别进行评估.

## 2.3 算法评估

本文从 Sturm 等<sup>[21]</sup>提供的 Freiburg1 数据

集<sup>[22]</sup>选取三组序列,如 fr1\_xyz、fr1\_desk、fr1\_floor,验证本文提出的 RGB-D SLAM 方法.该数据集包括 Kinect 数据以及高精度、实时同步的相机真实轨迹.Kinect 数据包括彩色图像和深度图像.相机真实轨迹是由 8 个高速跟踪摄像机构成的运动捕捉系统获得.同时,该数据集提供了 RGB-D SLAM 系统的比较方法.本文使用 RMSE(root mean square error)作为评判 RGB-D SLAM 方法精度的标准.

$$RMSE = \left( \frac{1}{n} \sum_{i=1}^n \|\tilde{t}_i - t_i\| \right)^{1/2} \quad (18)$$

式中,  $\tilde{t}_i = (\tilde{t}_x, \tilde{t}_y, \tilde{t}_z)$  是使用本文方法计算出的相机位姿,  $t_i = (t_x, t_y, t_z)$  是真实的相机位姿.

图 6 是数据集 fr1\_desk 的部分实验结果.图 6 (a)是未采用滤波处理和 ICP 算法对应的实验结果,(b)是采用分段滤波和 ICP 算法处理之后的实验结果.根据实验结果对比可知:采用分段滤波和 ICP 算法之后,估计的相机轨迹(II)与真实轨迹(I)之间的误差(III)减小,证明了本文提出方法的有效性,并在数据集 fr1\_xyz 和 fr1\_floor 上验证了实验效果并进行了对比.此外,本文对 RGB-D SLAM 方法的运行时间进行对比,实验结果如表 2 所示.

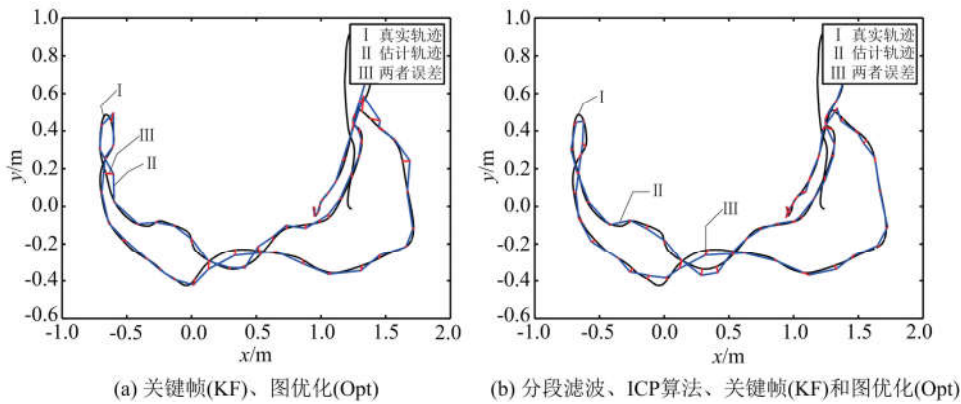


图 6 数据集 fr1\_desk 估计轨迹

Fig.6 Estimated trajectories of fr1\_desk dataset

表 2 RGB-D SLAM 数据集 RMSE 误差和运行时间

Tab.2 RMSE error and runtime based on RGB-D SLAM datasets

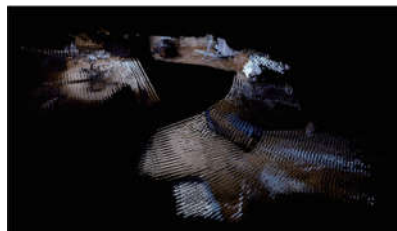
方法	RMSE/m			运行时间/s		
	fr1_xyz	fr1_desk	fr1_floor	fr1_xyz	fr1_desk	fr1_floor
SURF+KF +Opt <sup>[33]</sup>	0.0150	0.0310	0.0450	0.2455	0.2165	0.2283
SURF+网格滤波+ICP +KF +Opt	0.0141	0.0302	0.0446	0.2489	0.2197	0.2233
SURF+随机滤波+ICP +KF +Opt	0.0144	0.0291	0.0442	<b>0.2476</b>	<b>0.2181</b>	<b>0.2226</b>
SURF +直通+网格滤波+ICP +KF +Opt	0.0139	0.0297	0.0440	0.2494	0.2200	0.2242
SURF+直通+随机滤波+ICP +KF +Opt	0.0138	0.0279	0.0436	0.2489	0.2182	0.2228
SURF+分段滤波+ICP +KF +Opt	<b>0.0127</b>	<b>0.0260</b>	<b>0.0416</b>	0.2487	0.2190	0.2230
RGB-6D-SLAM <sup>[17]</sup>	0.0210	0.0490	0.0550	0.4630	0.3460	0.4020

本文方法运行在配置为四核 16 GB 内存的电脑上,文献[17]中 RGB-6D-SLAM 方法运行在配置为四核 8 GB 内存的电脑上,文献[33]中 RGB-D SLAM 方法运行在配置为双核 2 GB 内存的电脑上.通过对表 2 中运行时间数据分析可得:与文献[17]中方法相比,本文方法的运行时间提高了 42.50%,与文献[33]中方法相比,本文方法与文献[33]中方法的运行时间几乎相同.

分析表 2 中的 RMSE 数据可知,与文献[17]中方法相比,本文采用点云滤波处理和 ICP 算法之后,精度均明显提高;与文献[33]中方法相比,使用网格滤波和随机滤波之后系统的平均精度分别提高了 3.2%、4.0%;使用直通滤波结合网格滤波和随机滤波,系统的精度分别提高了 4.6%、7.0%;使用分段滤波之后,系统的精度提高了 13%;与不采用点云滤波处理和 ICP 算法相比,本文提出的 RGB-D

SLAM 方法估计出的相机轨迹的精度更高。

为了能够将实验成果应用于机器人的导航定位,本文将最终处理所得的点云数据进行栅格化处理。



(a) RGB-D SLAM输出的点云



(b) 栅格地图

图 7 RGB-D SLAM 建立的地图

Fig.7 Map built by RGB-D SLAM

### 3 结论

本文提出了提高 RGB-D SLAM 方法精度的措施。一方面,改进点云滤波方法,对预处理 RGB-D 相机数据获得的点云数据进行滤波处理,去除点云数据中的噪声和冗余;另一方面,用 RANSAC 算法计算的位姿变换矩阵作为 ICP 算法的初始值,同时用滤波之后的点云数据作为 ICP 算法的输入,使得 ICP 算法得到较为精确的相机位姿变换矩阵,从而提高估计的相机运动轨迹的精度。在国际公认的数据集上验证本文提出的 RGB-D SLAM 方法的精度,实验结果表明,在运行时间相同的条件下,这两方面的改进能够有效提高 RGB-D SLAM 方法的精度。

由文献[21]可知,采集数据集 fr1\_xyz 时 Kinect 的平移和旋转速度分别是 0.24 m/s、0.92 deg/s;采集数据集 fr1\_desk 时 Kinect 的平移和旋转速度分别是 0.41 m/s、23.33 deg/s;采集数据集 fr1\_floor 时 Kinect 的平移和旋转速度分别是 0.26 m/s、15.07 deg/s。由此可见,当 Kinect 运行速度比较小时,RGB-D SLAM 方法的精度较高;速度较大时,精度较低。下一步将考虑如何在 Kinect 运行速度比较大的情况下,提高 RGB-D SLAM 方法的精度。

#### 参考文献(References)

- [1] 李仁厚.自主移动机器人导论[M].西安:西安交通大学出版社,2006.
- [2] SMITH R C, CHEESEMAN P. On the representation and estimation of spatial uncertainty[J]. International Journal of Robotics Research, 1986, 5(4): 56-58.
- [3] VAZQUEZ-MARTIN R, NUNEZ P, DEL TORO J C, et al. Adaptive observation covariance for EKF-SLAM in indoor environments using laser data[C]// Proceedings of

理.对数据集 fr1\_floor 处理之后得到的点云数据如图 7(a)所示,使用 Octomap 库生成的栅格地图,如图 7(b)所示。

the IEEE Mediterranean Electrotechnical Conference. Malaga, Spain: IEEE Press, 2006: 445-448.

- [4] EL HAMZAOUI O, STEUX B. A fast scan matching for grid-based laser SLAM using streaming SIMD extensions [C]// Proceedings of the International Conference on Control Automation Robotics and Vision. Singapore: IEEE Press, 2010: 1986-1990.
- [5] 周武, 赵春霞, 沈亚强, 等. 基于全局观测地图模型的 SLAM 研究[J]. 机器人, 2010, 32(5): 647-654.
- [6] REINA G, UNDERWOOD J, BROOKER G, et al. Radar-based perception for autonomous outdoor vehicle [J]. Journal of Field Robotics, 2011, 28(6): 894-913.
- [7] TOMONO M. Robust 3D SLAM with a stereo camera based on an edge-point ICP algorithm [C]// Proceedings of the IEEE International Conference on Robotics & Automation. Kobe, Japan: IEEE Press, 2009: 4306-4311.
- [8] 陈伟, 吴涛, 李政, 等. 基于粒子滤波的单目视觉 SLAM 算法[J]. 机器人, 2008, 37(3): 242-253.
- [9] 顾爽, 陈启军. 基于全景视觉匹配的移动机器人蒙特卡罗定位算法[J]. 控制理论与应用, 2012, 29(5): 585-591.
- [10] 朱笑笑, 曹其新, 杨扬, 等. 基于 RGB-D 传感器的 3D 室内环境地图的实时创建[J]. 计算机工程与设计, 2014, 35(1): 203-207.
- [11] DRYANOVSKI I, VALENTI R G, XIAO J Z. Fast visual odometry and mapping from RGB-D data[C]// Proceedings of the IEEE International Conference on Robotics and Automation. Pisataway, USA: IEEE Press, 2013: 2305-2310.
- [12] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments[C]//Proceedings of the 12th International Symposium on Experimental Robotics. Delhi, India: CiteSeer, 2010.
- [13] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments[J]. International Journal of Robotics Research, 2012, 31(5): 647-663.



- [14] 欧美英, 孙海滨, 李世华. 具有外部扰动的非完整移动机器人的有限时间分钟控制[J]. 中国科学技术大学学报, 2012, 42(5): 405-414.  
OU Meiyang, SUN Haibin, LI Shihua. Finite time tracing control of a nonholonomic mobile robot with external disturbance [J]. Journal of University of Science and Technology of China, 2012, 42(5): 405-414.
- [15] 吕妍, 陈宗海. 不确定环境信息下基于方位关系的路径规划算法[J]. 中国科学技术大学学报, 2013, 43(10): 782-789, 829.  
LV Yan, CHEN Zonghai. Path planning algorithm based on directional relationship with uncertain environment information [J]. Journal of University of Science and Technology of China, 2013, 43(10): 782-789, 829.
- [16] 谷丰, 王争, 宋琦, 等. 空地机器人协作导航方法与实验研究[J]. 中国科学技术大学学报, 2012, 42(5): 398-404.  
GU Feng, WANG Zheng, SONG Qi, et al. Theoretical and experimental study of air-ground fluidized bed[J]. Journal of University of Science and Technology of China, 2012, 42(5): 398-404.
- [17] ENDRES F, HESS J, ENGELHARD N, et al. An evaluation of the RGB-D SLAM system [C]// Proceedings of the IEEE International Conference on Robotics and Automation. Minnesota, USA: IEEE Press, 2012: 1691-1696.
- [18] ENDRES F, HESS J, ENGELHARD N, et al. 3-D mapping with an RGB-D camera [J]. IEEE Transactions on Robotics, 2014, 30(1): 177-187.
- [19] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the Association for Computing Machinery, 1981, 24(6): 381-395.
- [20] BESL P J, MCKAY N D. Method for registration of 3-D shapes [C]// Proceedings of the International Society for Optics and Photonics. 1992: 586-606.
- [21] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB-D SLAM systems [C]// Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal; IEEE Press, 2012: 573-580.
- [22] Computer Vision Group. RGB-D SLAM Dataset and benchmark [EB/OL]. <http://vision.in.tum.de/data/datasets/rgbd-dataset>.
- [23] KÜMMERLE R, GRISETTI G, STRASDAT H, et al.  $g^2o$ : A general framework for graph optimization [C]// Proceedings of the IEEE International Conference on Robotics and Automation. Shanghai, China; IEEE Press, 2011: 3607-3613.
- [24] WURM K M, HORNUNG A, BENNEWITZ M, et al. OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems [C]// Proceedings of the International Conference on Robotics and Automation 2010 Workshop on Best Practice in 3D Perception and Modeling for Mobile Manipulation. 2010, 16(3): 403-412.
- [25] BRADSKI G, KAEHLER A. Learning OpenCV: Computer Vision with the OpenCV Library [M]. O'Reilly Media, Inc. 2008.
- [26] LOWE D G. Object recognition from local scale-invariant features [C]// Proceedings of the IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE Press, 1999: 1150-1157.
- [27] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF [C]// Proceedings of the IEEE International Conference on Computer Vision. Barcelona, Spain; IEEE Press, 2011: 2564-2571.
- [28] 冯亦东, 孙跃. 基于 SURF 特征提取和 FLANN 搜索的图像匹配算法 [J]. 图学学报, 2015, 36(4): 650-654.
- [29] BAY H, TUYTELAARS T, VAN GOOL L. SURF: Speeded up robust features [C]// Proceedings of European Conference on Computer Vision. Graz, Austria; Springer, 2006: 404-407.
- [30] 刘志斌, 吴显亮, 徐文立, 等. 视觉 SLAM 中的基于误匹配风险预测的特征选择 [J]. 机器人, 2010, 32(5): 635-641.
- [31] RUSU R B. Semantic 3D object maps for everyday manipulation in human living environments [J]. KI-Künstliche Intelligenz, 2010, 24(4): 345-348.
- [32] 王亚军, 蒯启忠, 王钦军, 等. 应用区域光谱库及分段滤波方法改进矿物识别精度的研究 [J]. 光谱学与光谱分析, 2012, 32(8): 2065-2069.
- [33] QUANG H P, QUOC N L. Some improvements in the RGB-D SLAM system [C]// Proceedings of the IEEE RVIF International Conference on Computing & Communication Technologies-Research, Innovation, and Vision for the Future. Can Tho, Vietnam; IEEE Press, 2015: 112-116.
- [34] SÜNDERHAUF N, PROTZEL P. Towards a robust back-end for pose graph SLAM [C]// Proceedings of the IEEE International Conference on Robotics and Automation. Saint Paul, Brazil; IEEE Press, 2012, 24(7): 1254-1261.
- [35] BRUGSLI D, BROENINK J F, KROEGER T, et al. Simulation, Modeling, and Programming for Autonomous Robots [M]. Berlin: Springer, 2008.
- [36] ZHANG L, SHEN P, DING J, et al. An improved RGB-D SLAM algorithm based on Kinect sensor [C]// Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics. Busan, Korea; IEEE Press, 2015: 555-562.