

## 基于 Kinect 和视觉词典的三维 SLAM

龙超\*, 韩波, 张宇

(浙江大学 航空航天学院, 杭州 310027)

(\* 通信作者电子邮箱 numkf11@126.com)

**摘要:**针对传统滤波器方法解决机器人同时定位与地图创建(SLAM)时的误差积累问题,提出了一种基于视觉词典(BOW)的三维 SLAM 算法,以有效解决机器人长时间运动下误差积累的问题。相比图优化 SLAM 中常用的随机检测和 Kd 树(Kd-Tree)算法,采用基于树结构的视觉词典闭环检测算法来提高相似场景的检索效率。首先采用基于 GPU 的特征提取算法提取图像特征,并利用交叉匹配和 k 最近邻(kNN)算法取得图像中鲁棒性较强的内点;然后通过基于随机抽样一致性奇异值分解(RANSAC SVD)算法计算出相邻帧的初始位姿变换,并利用通用迭代最近点(G-ICP)算法进行优化,得到高精度的位姿变换;最后利用增量平滑和建图(iSAM)图优化方法得出最终位姿,拼接出高精度的点云地图和运动轨迹。标准数据集的测试表明,所提算法在复杂情况下具有良好的鲁棒性和精度。

**关键词:**三维 SLAM; 闭环检测; 视觉词典; 图形处理器; 通用迭代最近点算法; 增量平滑和建图

**中图分类号:** TP24      **文献标志码:** A

### Three-dimensional SLAM using Kinect and visual dictionary

LONG Chao\*, HAN Bo, ZHANG Yu

(School of Aeronautics and Astronautics, Zhejiang University, Hangzhou Zhejiang 310027, China)

**Abstract:** Since traditional filter methods to solve Simultaneous Localization And Mapping (SLAM) problems will accumulate errors, a three-dimensional SLAM algorithm based on Bag-Of-Words (BOW) algorithm which can effectively solves the problem of accumulating errors was proposed. Compared to the common algorithms like random selection and k-Dimensional Tree (Kd-Tree), a tree structure visual bag of words loop detection algorithm was designed which could greatly increase the speed of similar scene detection. Firstly, a GPU based feature extraction algorithm was adopted. Through using cross matching and k-Nearest Neighbor (kNN) algorithm, robust inliers were got. Secondly, Random Sample Consensus Singular Value Decomposition (RANSAC SVD) algorithm was used to calculate the initial transformation between two frames. And then a Generalized-Iterative Closest Point (G-ICP) algorithm was used to optimize the transformation to get precise transformation. At last, incremental Smoothing And Mapping (iSAM) Graph optimization algorithm was used to calculate the camera pose and the point cloud map and trajectory were created. The test results on the standard dataset show that the algorithm can achieve good robustness and precision under complex environment.

**Key words:** three-dimensional Simultaneous Localization and Mapping (3D-SLAM); loop closure; Bag-Of-Words (BOW); Graphic Processing Unit (GPU); Generalized-Iterative Closest Point (G-ICP); incremental Smoothing and Mapping (iSAM)

## 0 引言

当机器人处于未知环境中时,利用自身传感器获取机器人所在环境的三维模型和机器人运动轨迹是一项基本需求。传统同时定位与地图创建(Simultaneous Localization and Mapping, SLAM)一般采用三维激光雷达或者声呐传感器来获取环境的深度信息,但此类传感器动辄上万的价格严重限制了其在机器人领域的广泛使用。2010 年微软推出的 Kinect 设备不仅可以获取环境的纹理信息,还可以廉价地获取环境的深度信息<sup>[1]</sup>,其低廉的价格使其在机器人领域得到了广泛的应用<sup>[2]</sup>。

传统基于滤波器求解 SLAM 问题的方法是利用递归贝叶斯估计的原理,在假定从 0 到  $t$  时刻的观测信息以及控制信息

已知的条件下,计算系统状态(机器人当前位姿以及所有地图特征位置)的后验概率<sup>[3]</sup>。依据后验概率的表示方式,主要有拓展卡尔曼滤波和粒子滤波等方法。基于滤波器的方法存在线性化导致长时间运行误差积累的问题和计算量大导致的更新效率低下的问题,这使其难以应用于大规模的地图创建。基于图优化解决 SLAM 问题的方法在 1997 年由 Lu 和 Milios 提出。基于图优化方法 SLAM 通过构建一张位姿图(pose graph),图的节点代表机器人的位姿或者路标,节点之间的边表示传感器测得的位姿约束信息。通过对图进行求解可以求得机器人的位姿信息。由于图优化方法利用了所有可用信息,其精度相比滤波器方法有很大的提高。Felix Endres 等提出了一种利用 g2o (a General Framework for Graph Optimization)<sup>[4]</sup>图优化框架的 3D-SLAM 方法<sup>[5]</sup>,利用 Kinect

收稿日期:2015-08-24;修回日期:2015-09-28。

**作者简介:**龙超(1992-),男,湖北石首人,硕士研究生,主要研究方向:计算机视觉、机器人导航; 韩波(1969-),男,浙江慈溪人,副教授,博士,主要研究方向:无人机的导航和控制; 张宇(1980-),男,浙江杭州人,讲师,博士,主要研究方向:计算机视觉、无人机的导航和控制。

等 RGB-D 传感器精确构建三维地图,利用 KD-Tree 进行闭环检测。但如果场景足够大,基于 KD-Tree 的闭环检测效率会降低。

如图 1 所示,本文首先通过对 RGBD 摄像头的 RGB 摄像头和深度摄像头进行对准,再对摄像头进行标定,接下来通过帧间匹配获取相机不同位置的相对位姿变换,同时为了提高

计算效率采取关键帧提取技术。在机器人实际运行中经常返回以前曾出现过地方,通过合理设计基于视觉词典(Bag-Of-Words, BOW)的闭环检测(Loop Closure)环节可以有效降低机器人运行过程中的累积误差。一旦机器人检测出闭环随即利用 iSAM 算法对姿态进行求解,然后将环境的点云实时拼接成点云地图,并绘出机器人运动路径。

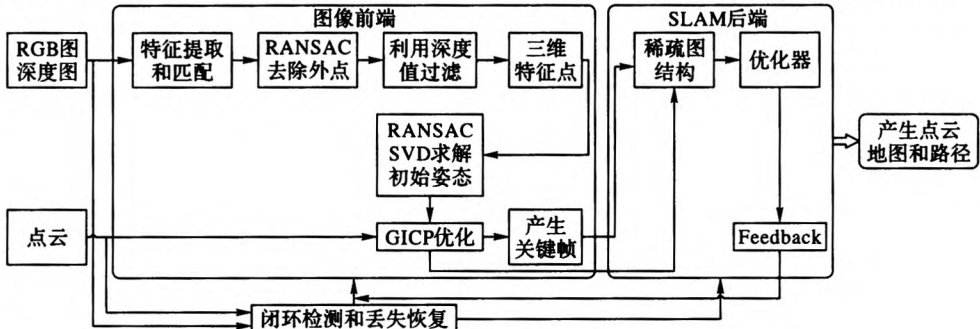


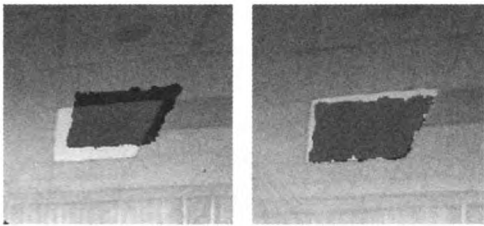
图 1 SLAM 算法原理

1 基本原理以及实现过程

1.1 OpenNI 图像采集和相机标定

SLAM 算法需要摄像头采集的 RGB 和深度图像具有相同的时间戳,同时 RGB 摄像头和深度摄像头之间外部参数表现为旋转和平移关系  $T = [R, t]$ ,需要将深度图和 RGB 图实现配准,使得通过索引 RGB 图的每个像素点可以准确获取其位置的深度值。可以利用 OpenNI 库实现图像的同步采集与深度图和 RGB 图的配准。

如图 2 所示,通过将 RGB 图和深度图进行融合(图中灰色色块为深度图中灯的位置),由于 Kinect 相机获得的 RGB 图和深度图并不是在同一个视角下,在配准前 RGB 图和深度图中同一盏灯的位置存在明显偏移,而经过校正后偏移得到了准确有效的补偿。



(a) 配准前 (b) 配准后  
图 2 配准前与配准后融合图像对比

图 3 为整个图像采集算法和配准的流程。

经过 OpenNI 将 RGB 和深度图片对准后,只要得到 RGB 中像素的索引就可以得到这个像素的深度值。RGB 相机一般采用张正友算法进行标定和获取相机的内参数矩阵。

如图 4 所示,其中  $O_{xy}$ -XY 为图片坐标系, $O-X_cY_c$  为 Kinect 坐标系,他们之间的相互转换关系为:

$$\begin{bmatrix} u_0 \\ v_0 \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (1)$$

其中: $s$  为深度单位, $O_{xy}(u_0, v_0)$  为图片坐标系中像素点的坐标,  $(X_c, Y_c, Z_c)$  为 Kinect 坐标系下点的三维坐标,

$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$  为相机的内参数矩阵,其中  $f_x$  和  $f_y$  为焦距在  $x$  和  $y$

方向上所占的像素数,  $(c_x, c_y)$  为基准点。



图 3 OpenNI 采集图像流程

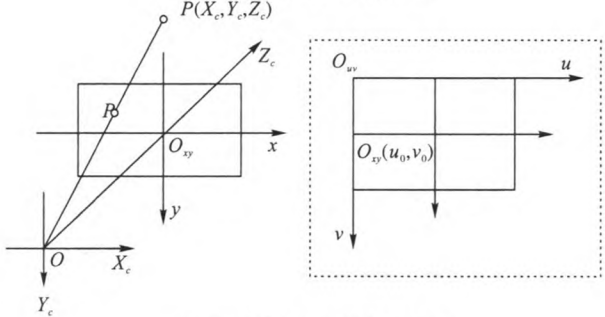


图 4 图片坐标系和相机坐标系

1.2 特征提取匹配与姿态计算

在一般三维 SLAM 中通常提取点云的 PFH(Point Feature Histogram)描述子并匹配,并以此来计算相机姿态。PFH 是通过参数化查询点与领域点之间的空间差异,并形成一对点  $k$  领域几何属性的直方图,PFH 的理论计算复杂度为  $O(nk^2)$ ,在不用 GPU 加速的情况下,并不适合实时 SLAM 的需求。

考虑到相机运动过程中的尺度变换,有 ORB(Oriented FAST and Rotated BRIEF)、SURF(Speeded Up Robust Feature)<sup>[6]</sup>、SIFT(Scale-Invariant Feature Transform)<sup>[7]</sup> 等特征描述子可供选择,ORB 计算速度非常快,但在纹理比较少的环境下提取特征点较少。SURF 利用 Haar 小波近似 SIFT 方法中的梯度操作,同时利用积分图技术进行快速计算,速度是 SIFT 的 3~7 倍。为了进一步提高特征提取的速度,采用 GPU-SURF(GPU accelerating Speeded-Up Robust Feature)。

由于 Kinect 传感器的特点,深度传感器会在物体边缘发

生散射,而在物体边缘往往有很多特征点,导致有些特征点深度值为 0,或者跳变。如果这些点不去除,将大大降低姿态计算的精度。因此获取鲁棒性较强的特征点至关重要。Dryanovski 等采取对深度特征点进行高斯混合建模,减小了跳变对姿态计算的影响<sup>[8]</sup>。

在用 SVD 算法对相机运动姿态进行估计时,对错误的匹配十分敏感。为了提高计算精度,采取三种方法:1)使用交叉过滤和 k 最近邻(k-Nearest Neighbor, kNN)匹配去除存在假阳性匹配和假阴性匹配,然后使用 RANSAC 算法估计相机的单应矩阵,得到鲁棒性较强的内点。2)使用改进的 RANSAC SVD 算法估计相机姿态。3)使用通用迭代最近点(Generalized Iterative Closest Point, G-ICP)<sup>[9]</sup>对 RANSAC SVD 得到的姿态进行进一步的优化。

算法步骤如下:

**步骤 1** 提取相邻帧的 GPU-SURF 特征和特征描述子,使用  $k(k=2)$  最近邻进行双向匹配得到 matches1 和 matches2。

**步骤 2** 移除最近邻(Nearest Neighbor, NN)比率大于阈值(1/1.5)的匹配,使用 RANSAC 移除匹配中的异常值,返回优质匹配 inlierMatches。

**步骤 3** 依据每个特征点的深度信息对 inlierMatches 进行滤波,去除深度值为 0 的特征点。

**步骤 4** 对于使用 RANSAC SVD 算法对滤波后的匹配点进行初始位姿估计。

**步骤 5** 利用 GICP 算法对初始姿态进行优化,得到最终的姿态 T。

交叉过滤原理简单,利用查询集和训练描述符相互匹配,只返回在这两个匹配中同时出现的匹配。由图 5 可知,交叉过滤起到的作用十分有限,得到的匹配中依然存在大量的错误匹配。在交叉过滤的基础上利用 kNN 比率测试,能够删除大量异常值,如图 6 所示,假阳性匹配依然能够通过 kNN 匹配,但匹配质量较之前有很大的提高。最后通过 RANSAC 算法估计单应性重投影误差,使得错误匹配完全去除,图 7 可知,RANSAC 过滤之后,内点规模依然保持 50% 以上。



图 5 交叉过滤

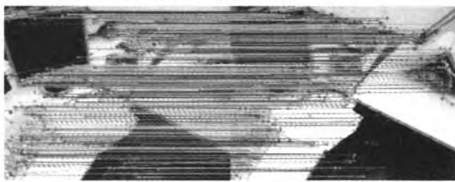


图 6 kNN 匹配

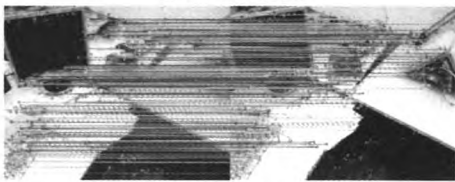


图 7 RANSAC 移除匹配异常值

### 1.3 闭环检测

在机器人实际运动过程中,机器人可能会回到之前访问

过的地方,通过识别机器人之前访问过的地方,可以修正里程计误差。正确的闭环检测可以增加实际位姿估计的精度。

可以将闭环看作弹簧-质点模型。如图 8 所示,当无闭环时,每一个关键帧的误差将会对往后的所有运动造成影响,这是经典的 SLAM 现象,机器人运动的时间越长,不确定性就会增加。通过检测当前帧和过去相似帧,增加两帧之间的约束,可以有效降低累积误差,增强了环内的稳定性,当检测到闭环时,整个环内结构将变得更加稳定。

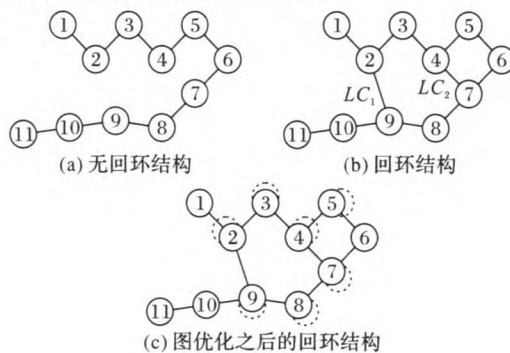


图 8 闭环示意图

图优化 SLAM 理论形成来,涌现了许多闭环检测的方法<sup>[10-12]</sup>。2006 年,David Nister 等提出了一种基于词汇树 BOW 的图片检索技术<sup>[13-14]</sup>。通过分层 K-Means 构建词汇树,可以实现海量图片的高效检索,单次检索平均速度可以达到 4 ms 左右。由于检索结果中的错误,并不能直接用于 SLAM 的闭环过程。基于该思想本文提出了一种基于 Ransac 检验的 BOW 闭环检测算法。

在分层 BOW 中,词汇是以树的结构存储的。树的构建过程十分直观,首先利用 K-Means 将所有训练集的特征描述子聚类成  $K$  类,然后再对各类中的特征描述子聚类成  $K$  类,此时原始节点便有了  $K$  个子节点,如此递归下去,直到构建到  $L$  层。最后可以获得  $W$  个叶节点。相比于其他 BOW 方法,分层构建词汇树的方法十分高效,时间复杂度为  $O(\log(N))$ ,其中  $N$  为叶节点的个数。图 9 为依照分层 BOW 原理构造的三层词典树模型。

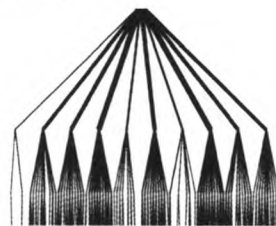


图 9 三层词典树

在构建树的过程中,需要利用逆向文本频率计算每个叶节点的权重  $w_i$ ,逆向文本频率可以表示每个叶节点的重要性,减小出现很频繁没有代表性的词汇对检索造成的影响。

$$w_i = \ln(N/N_i) \quad (2)$$

其中: $N$  表示数据库中图片的总数, $N_i$  表示至少有一个特征描述子通过节点  $i$  的图片的总数。

因此,检索图片向量  $q_i$  和数据库向量  $d_i$  可以表示成:

$$q_i = n_i w_i \quad (3)$$

$$d_i = m_i w_i \quad (4)$$

其中: $n_i$  和  $m_i$  为检索图片和数据库图片中的特征描述子通过节点  $i$  的数目。然后,数据库中的图片可以返回一个反映图片

相似程度的归一化分数  $s$ :

$$s(q, d) = \|q/\|q\| - d/\|d\|\| \quad (5)$$

计算归一化  $L_p$  范数时,如下可以使用式(6):

$$\|q - d\|_p^p = \sum_i |q_i - d_i|^p = \sum_{i: d_i=0} |q_i|^p + \sum_{i: q_i=0} |d_i|^p + \sum_{i: d_i \neq 0, q \neq 0} |q_i - d_i|^p = 2 + \sum_{i: d_i \neq 0, q \neq 0} (|q_i - d_i|^p - |q_i|^p - |d_i|^p) \quad (6)$$

当计算出检索的分数之后,可以通过对检索分数进行排序,选出其中分数最高的 10 张图片,然后利用 RANSAC 算法计算当前帧和检索帧的内点,当内点数目超过一定的阈值时,闭环检测成功。

#### 1.4 丢失恢复

机器人在实际运动过程中可能会遇到许多突发状况,比如镜头的遮挡、剧烈抖动等,导致持续出现位姿无法计算的情况,如果不作处理会后续的定位和构图造成严重影响。因此,本文设计了丢失恢复系统,每当检测到连续丢失  $N$  帧,立即启动丢失恢复,将当前帧设为关键帧,同时在已有的关键帧中逐帧进行两两匹配,检测当前位置是否之前出现过,尽量降低机器人丢失造成的影响。

#### 1.5 图优化求解相机姿态

利用图优化方法求解相机位姿,一般可以通过 Cholesky 解法来求解也可以使用 iSAM<sup>[15]</sup> 算法来求解。如图 10 所示,SLAM 的模型可以表示为:

$$P(X, L, U, Z) \propto P(x_0) \prod_{i=1}^M P(x_i | x_{i-1}, u_i) \prod_{k=1}^K P(z_k | x_{i_k}, l_{j_k}) \quad (7)$$

运动方程为:

$$x_i = f_i(x_{i-1}, u_i) + w_i \quad (8)$$

观测方程为:

$$z_k = h_k(x_{i_k}, l_{j_k}) + v_k \quad (9)$$

其中:  $P(x_0)$  为初始状态,  $X = \{x_i\}$  为机器人状态  $L = \{l_i\}$  为路标,  $U = \{u_i\}$  为控制输入,  $Z = \{z_k\}$  为路标测量。  $P(x_i | x_{i-1}, u_i)$  为运动模型,  $P(z_k | x_{i_k}, l_{j_k})$  为观测方程。  $w_i$  和  $v_k$  为运动和观测过程中的噪声。

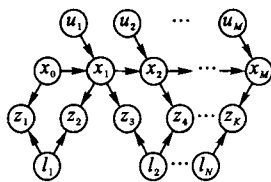


图 10 SLAM 问题的贝叶斯信念网表示

可以将这个问题转换为基于最大后验估计的最小二乘问题:

$$X^*, L^* = \arg \max_{X, L} P(X, L, U, Z) = \arg \min_{X, L} -\log P(X, L, U, Z) \quad (10)$$

结合运动和观测方程(8)、(9):

$$X^*, L^* = \arg \min_{X, L} \left\{ \sum_{i=1}^M \|f_i(x_{i-1}, u_i) - x_i\|_{A_i}^2 + \sum_{k=1}^K \|h_k(x_{i_k}, l_{j_k}) - z_k\|_{r_k}^2 \right\} \quad (11)$$

式(11)可以表示成:

$$\theta^* = \arg \min_{\theta} \|A\theta - b\|^2 \quad (12)$$

其中:  $\theta \in \mathbb{R}^n$ , 表示所有的位姿和路标变量,  $A \in \mathbb{R}^{m \times n}$  为稀疏雅克比测量矩阵,  $b \in \mathbb{R}^m$  为右端 (Right-Hand Side, RHS) 向量。

而稀疏雅克比测量矩阵  $A$  可以被 QR 分解为  $Q \begin{bmatrix} R \\ 0 \end{bmatrix}$ ,  $R \in \mathbb{R}^{n \times n}$  为上三角平方根信息矩阵 ( $R^T R = A^T A$ ),  $Q \in \mathbb{R}^{m \times m}$  为正交矩阵。式(12)可以转化为:

$$\|A\theta - b\|^2 = \left\| Q \begin{bmatrix} R \\ 0 \end{bmatrix} \theta - b \right\|^2 = \left\| Q^T Q \begin{bmatrix} R \\ 0 \end{bmatrix} \theta - Q^T b \right\|^2 = \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} \theta - \begin{bmatrix} d \\ e \end{bmatrix} \right\|^2 = \|R\theta - d\|^2 + \|e\|^2 \quad (13)$$

其中:  $[d, e]^T = Q^T b$ ,  $\|e\|^2$  为残差, 通过 QR 分解使得最小二乘问题转化为了一个线性系统:

$$R\theta^* = d \quad (14)$$

QR 分解获得  $R$  一般使用吉文斯旋转获得 (Givens Rotations)<sup>[16]</sup>。

$$\Phi := \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \quad (15)$$

其中:

$$(\cos \phi, \sin \phi) = \begin{cases} (0, 1), & \beta = 0 \\ \left( \frac{-\alpha}{\beta \sqrt{1 + (\alpha/\beta)^2}}, \frac{1}{\sqrt{1 + (\alpha/\beta)^2}} \right), & |\beta| > |\alpha| \\ \left( \frac{1}{\sqrt{1 + (\beta/\alpha)^2}}, \frac{-\beta}{\alpha \sqrt{1 + (\beta/\alpha)^2}} \right), & \text{其他} \end{cases} \quad (16)$$

其中:  $\alpha = a_{kk}, \beta = a_{ik}, a_{kk}, a_{ik}$  分别为  $A$  矩阵  $k$  行  $k$  列和  $i$  行  $k$  列的值。

当更新时,只是在  $A$  中加了一行新的测量  $w^T$ , RHS 加入了  $\gamma$ :

$$\begin{bmatrix} Q^T \\ 1 \end{bmatrix} \begin{bmatrix} A \\ w^T \end{bmatrix} = \begin{bmatrix} R \\ \gamma \end{bmatrix} \quad (17)$$

其中,新的 RHS 为  $\begin{bmatrix} d \\ \gamma \end{bmatrix}$ , 可以知道形式仍与之前的一样,同样可以用之前的方法求解。

## 2 实验及结果分析

本文设计了四组实验,均采用的是 Computer Vision Group 提供的数据集,最后一组为对比实验。RGBD 数据集提供 30 Hz 帧率采集的 RGB 图和深度图,分辨率为  $640 \times 480$ ,同时提供由 8 个高速运动跟踪摄像机 (100 Hz) 采集的真实轨迹以评估 SLAM 系统的性能。

第一组为 freiburg1\_floor 数据集,共 1 200 帧,为地面机器人室内所采。第二组为 freiburg2\_rpy 数据集,共 3 200 帧,为手持 Kinect 采集。第三组为 freiburg1\_xyz 数据集,共 570 帧,为手持 Kinect 采集,为了体现闭环检测作用,分别设计了有闭环和无闭环两组实验。图 11 ~ 14 分别为 4 组实验的三维点云地图与预估轨迹和真实路径的对比。

由实验结果可以知 SLAM 算法在复杂场景下取得了很好的结果,甚至能处理像 freiburg2\_rpy 这样超过 3 000 帧的数据集。对于 freiburg1\_xyz 数据集,设计了对比实验,由图 13 可以看出当添加了闭环检测环节,桌子和电脑的地图拼接得十

分精确,由图 14 可知一旦去除 SLAM 的闭环检测之后,由于累积误差的影响,拼接地图中书桌边缘以及电脑都出现明显错误,整个路径明显偏离。

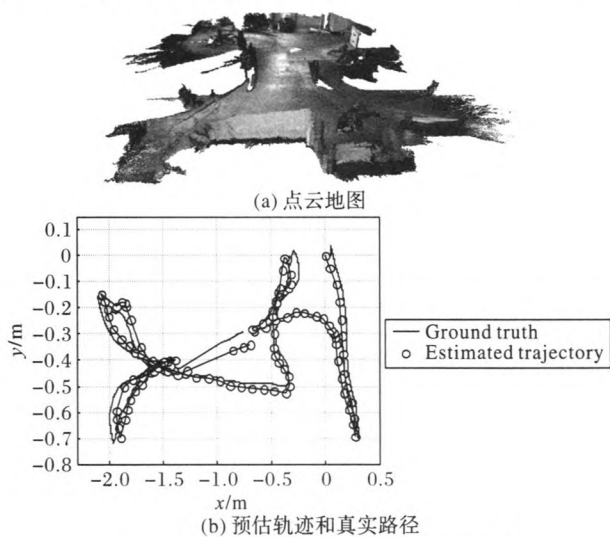


图 11 freiburg1\_floor 数据集点云地图及路线

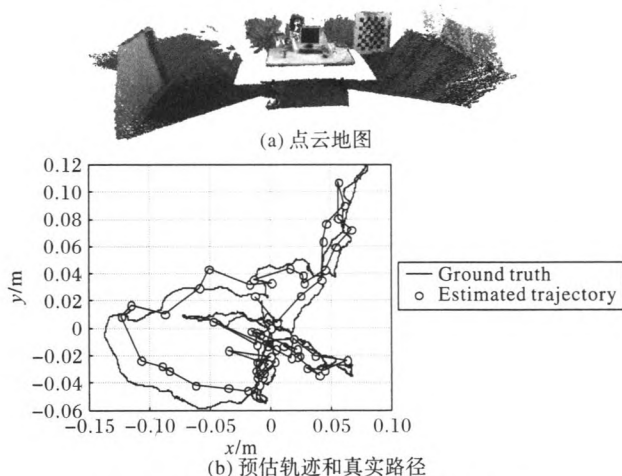


图 12 freiburg2\_rpy 数据集点云地图以及路线

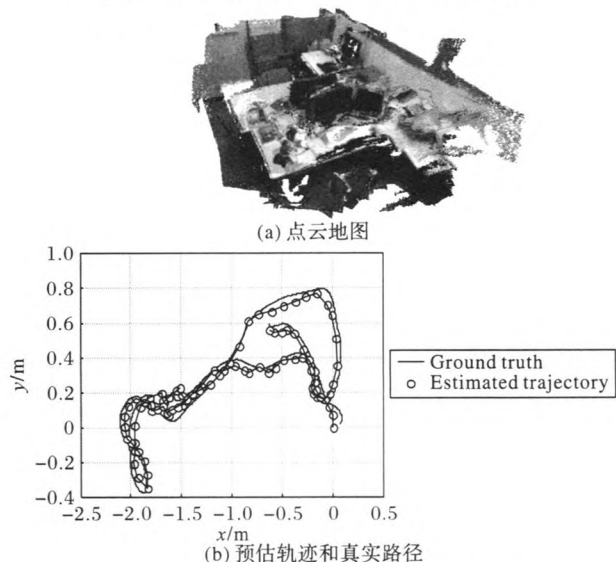


图 13 freiburg1\_xyz 检测闭环时地图及路线

### 3 结语

本文提出了一种基于图优化框架和 BOW 的三维 SLAM

算法,基于标准数据集的实验验证了本算法的高效性和鲁棒性。本算法未来计划从两个方面来进行改进:

1) 完全采用 CUDA、OpenMP 等语言实现,提高算法的实时性。

2) 和大多数基于特征提取的方法一样,算法对应用场景也有一定的要求,大面积的无纹理场景会对算法造成影响,未来可以结合点云特征来对算法进行改进。

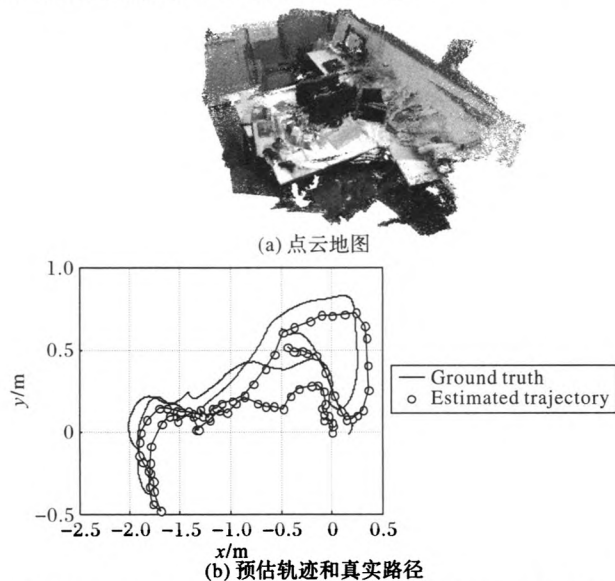


图 14 freiburg1\_xyz 不检测闭环时点云地图以及两种路线对比

### 参考文献:

- [1] ADAMS M, VO B N, MAHLER R, et al. SLAM gets a PHD: new concepts in map estimation [J]. IEEE Robotics and Automation Magazine, 2014, 21(2): 26-37.
- [2] GRIGORESCU S M, MACESANU G, COCIAS T T, et al. Robust camera pose and scene structure analysis for service robotics [J]. Robotics and Autonomous Systems, 2011, 59(11): 899-909.
- [3] 梁明杰, 闵华清, 罗荣华. 基于图优化的同时定位与地图创建综述 [J]. 机器人, 2013, 35(4): 500-512. (LIANG M J, MIN H Q, LUO R H. Graph-based SLAM: a survey [J]. Robot, 2013, 35(4): 500-512.)
- [4] GRISSETTI G, KÜMMERLE R, STRASDAT H, et al. g2o: a general framework for graph optimization [EB/OL]. [2015-02-09]. <http://ais.informatik.uni-freiburg.de/publications/papers/kuemmerle11icra.pdf>.
- [5] 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.
- [6] BAY H, TUYTELAARS T, van GOOL L. SURF: speeded up robust features [M]// LEONARDIS A, BISCHOF H, PINZ A. Computer Vision—ECCV 2006, LNCS 3951. Berlin: Springer, 2006: 404-417.
- [7] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] DRYANOVSKI I, VALENTI R G, XIAO J. Fast visual odometry and mapping from RGB-D data [C]// Proceedings of the 2013 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, 2013: 2305-2310.
- [9] SEGAL A, HÄHNEL D, THRUN S. Generalized-ICP [EB/OL]. [2015-02-11]. [https://www.researchgate.net/publication/221344436\\_Generalized-ICP](https://www.researchgate.net/publication/221344436_Generalized-ICP).

(下转第 788 页)



- [6] LIU L, DONG H, HUANG H, et al. No-reference image quality assessment in curvelet domain [J]. *Signal Processing Image Communication*, 2014, 29(4): 494 – 505.
- [7] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain [J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695 – 4708.
- [8] 夏裕建, 孙涵. 基于空域 NSS 的无参考图像质量评价[J]. *计算机与现代化*, 2015(2): 67 – 72. (XIA Y J, SUN H. No-reference image quality assessment based on spatial natural scene statistics [J]. *Computer and Modernization*, 2015(2): 67 – 72.)
- [9] BENOIT A, LE CALLET P, CAMPISI P, et al. Quality assessment of stereoscopic images [EB/OL]. [2015-01-17]. [https://www.researchgate.net/publication/254860233\\_Quality\\_Assessment\\_of\\_Stereoscopic\\_Images](https://www.researchgate.net/publication/254860233_Quality_Assessment_of_Stereoscopic_Images).
- [10] YOU J, XING L, PERKIS A, et al. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis [EB/OL]. [2015-02-03]. [https://www.researchgate.net/publication/200774635\\_Perceptual\\_Quality\\_Assessment\\_for\\_Stereoscopic\\_Images\\_Based\\_on\\_2D\\_Image\\_Quality\\_Metrics\\_and\\_Disparity\\_Analysis](https://www.researchgate.net/publication/200774635_Perceptual_Quality_Assessment_for_Stereoscopic_Images_Based_on_2D_Image_Quality_Metrics_and_Disparity_Analysis).
- [11] POTETZ B, LEE T S. Scaling laws in natural scenes and the inference of 3D shape [EB/OL]. [2015-02-03]. [http://www.cs.cmu.edu/~bpotetz/Publications/Potetz\\_NIPS05.pdf](http://www.cs.cmu.edu/~bpotetz/Publications/Potetz_NIPS05.pdf).
- [12] LIU Y, CORMACK L K, BOVIK A C. Statistical modeling of 3-D natural scenes with application to Bayesian stereopsis [J]. *IEEE Transactions on Image Processing*, 2011, 20(9): 2515 – 2530.
- [13] SU C C, CORMACK L K, BOVIK A C. Color and depth priors in natural images [J]. *IEEE Transactions on Image Processing*, 2013, 22(6): 2259 – 2274.
- [14] CHEN M J, CORMACK L K, BOVIK A C. No-reference quality assessment of natural stereopairs [J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3379 – 3391.
- [15] LEVELT W J M. On binocular rivalry [EB/OL]. [2015-02-05]. [http://pubman.mpg.de/pubman/item/escidoc:77195:8/component/escidoc:513080/Levelt\\_Binocular\\_Rivalry\\_1965.pdf](http://pubman.mpg.de/pubman/item/escidoc:77195:8/component/escidoc:513080/Levelt_Binocular_Rivalry_1965.pdf).
- [16] CHEN M J, SU C C, KWON D K, et al. Full-reference quality assessment of stereopairs accounting for rivalry [J]. *Signal Processing Image Communication*, 2013, 28(9): 1143 – 1155.
- [17] RUDERMAN D L. The statistics of natural images [J]. *Network Computation in Neural Systems*, 1994, 5(4): 517 – 548.
- [18] SHARIFI K, LEON-GARCIA A. Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1995, 5(1): 52 – 56.
- [19] 顾珊波, 邵枫, 蒋刚毅, 等. 基于支持向量回归的立体图像客观质量评价模型[J]. *电子与信息学报*, 2012, 34(2): 368 – 374. (GU S B, SHAO F, JIANG G Y, et al. Objective stereoscopic image quality assessment model based on support vector regression [J]. *Journal of Electronics and Information Technology*, 2012, 34(2): 368 – 374.)
- [20] MOORTHY A K, SU C C, MITTAL A, et al. Subjective evaluation of stereoscopic image quality [J]. *Signal Processing Image Communication*, 2013, 28(8): 870 – 883.
- [21] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600 – 612.
- [22] MOORTHY A K, BOVIK A C. Blind image quality assessment: from natural scene statistics to perceptual quality [J]. *IEEE Transactions on Image Processing*, 2011, 20(12): 3350 – 3364.

## Background

This work is partially supported by the National Science and Technology Support Program of China (2012BAH67F01), Key Projects of National Science Foundation of China (U1301257), the Research Project of Zhejiang Provincial Education Department (Y201327703), the Independent Design Project of Zhejiang Provincial Science and Technology Department/Creative Team (2012R10009-08), the Research Project of Ningbo Science and Technology Innovation Team (2011B81002).

**MA Yun**, born in 1991, M. S. candidate. Her research interests include multimedia communication and information security, image processing.

**WANG Xiaodong**, born in 1970, associate professor. His research interests include network communications, image processing, multimedia signal processing.

**ZHANG Linjun**, born in 1980, research fellow. His research interests include multimedia information security, image processing.

(上接第 778 页)

- [10] LABBE M, MICHAUD F. Appearance-based loop closure detection for online large-scale and long-term operation [J]. *IEEE Transactions on Robotics*, 2013, 29(3): 734 – 745.
- [11] CUMMINS M, NEWMAN P. Appearance-only SLAM at large scale with FAB-MAP 2.0 [J]. *IEEE Transactions on Robotics*. 2011, 30(9): 1100 – 1123.
- [12] HO K L, NEWMAN P. Detecting loop closure with scene sequences [J]. *International Journal of Computer Vision*, 2007, 74(3): 261 – 286.
- [13] NISTÉR D, STEWÉNUS H. Scalable recognition with a vocabulary tree[C]// *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2006, 2: 2161 – 2168.
- [14] SIVIC J, ZISSERMAN A. Video Google: a text retrieval approach to object matching in videos [C]// *Proceedings of the 2003 9th IEEE International Conference on Computer Vision*. Washington, DC: IEEE Computer Society, 2003, 2: 1470 – 1477.
- [15] KAESSE M, RANGANATHAN A, DELLAERT F. iSAM: incremental smoothing and mapping [J]. *IEEE Transactions on Robotics*, 2008, 24(6): 1365 – 1378.

## Background

**LONG Chao**, born in 1992, M. S. candidate. His research interests include computer vision and robot navigation.

**HAN Bo**, born in 1969, associate professor. His research interests include navigation and control of UAV (unmanned aerial vehicles).

**ZHANG Yu**, born in 1980, lecturer. His research interests include computer vision and navigation and control of UAV (unmanned aerial vehicles).