



## 基于kinect的改进RGB-D视觉里程计

朱俊涛, 陈强

引用本文:

朱俊涛, 陈强. 基于kinect的改进RGB-D视觉里程计[J]. 智能系统学报, 2020, 15(5): 943–948.

ZHU Juntao, CHEN Qiang. Improvement of kinect performance in RGB-D visual odometer[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(5): 943–948.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201903007>

## 您可能感兴趣的其他文章

### 基于LiDAR/INS的野外移动机器人组合导航方法

Integrated navigation approach for the field mobile robot based on LiDAR/INS

智能系统学报. 2020, 15(4): 804–810 <https://dx.doi.org/10.11992/tis.202008026>

### 基于Kinect的改进移动机器人视觉SLAM

Improved V-SLAM for mobile robots based on Kinect

智能系统学报. 2018, 13(5): 734–740 <https://dx.doi.org/10.11992/tis.201705018>

### 量子粒子群优化下的RBPF-SLAM算法研究

Research on RBPF-SLAM algorithm based on quantum-behaved particle swarm optimization

智能系统学报. 2018, 13(5): 829–835 <https://dx.doi.org/10.11992/tis.201705006>

### 基于图优化的移动机器人视觉SLAM

Visual-SLAM for mobile robot based on graph optimization

智能系统学报. 2018, 13(2): 290–295 <https://dx.doi.org/10.11992/tis.201612004>

### 多传感器的移动机器人可定位性估计与自定位

Self-localization of mobile robot in dynamic environments based on localizability estimation with multi-sensor observation

智能系统学报. 2017, 12(4): 443–449 <https://dx.doi.org/10.11992/tis.201607007>

### 粗匹配和局部尺度压缩搜索下的快速ICP-SLAM

Fast ICP-SLAM with rough alignment and local scale-compressed searching

智能系统学报. 2017, 12(3): 413–421 <https://dx.doi.org/10.11992/tis.201605029>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.201903007

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20191012.1719.006.html>

# 基于 kinect 的改进 RGB-D 视觉里程计

朱俊涛, 陈强

(上海工程技术大学 电子电气工程学院, 上海 201600)

**摘 要:** 针对 RGB-D 视觉里程计中 kinect 相机所捕获的图像深度区域缺失的问题, 提出了一种基于 PnP(perspective-n-point) 和 ICP(iterative closest point) 的融合优化算法。传统 ICP 算法迭代相机位姿时由于深度缺失, 经常出现特征点丢失导致算法无法收敛或误差过大。本算法通过对特征点的深度值判定, 建立 BA 优化模型, 并利用 g2o 求解器进行特征点与相机位姿的优化。实验证明了该方法的有效性, 提高了相机位姿估计的精度及算法的收敛成功率, 从而提高了 RGB-D 视觉里程计的精确性和鲁棒性。

**关键词:** kinect; 深度丢失; 融合算法; 特征点; ICP; PnP; 深度值; 位姿估计; BA 优化模型; g2o

**中图分类号:** TP242.6    **文献标志码:** A    **文章编号:** 1673-4785(2020)05-0943-06

中文引用格式: 朱俊涛, 陈强. 基于 kinect 的改进 RGB-D 视觉里程计 [J]. 智能系统学报, 2020, 15(5): 943-948.

英文引用格式: ZHU Juntao, CHEN Qiang. Improvement of kinect performance in RGB-D visual odometer[J]. CAAI transactions on intelligent systems, 2020, 15(5): 943-948.

## Improvement of kinect performance in RGB-D visual odometer

ZHU Juntao, CHEN Qiang

(Electrical and Electronic Engineering College, Shanghai University of Engineering and Technology, Shanghai 201600, China)

**Abstract:** Kinect is a 3D camera that gives you the depth values associated with every pixel. It uses structured infrared light to determine depth values. Apart from these, you also have access to raw RGB-D data, and even the raw infrared data. Aiming to solve the problem of insufficient depth values for the images captured by Kinect camera in RGB-D visual odometer, we propose a fusion optimization algorithm based on Perspective-n-Point and iterative closest point (ICP). Because of the lack of depth values, traditional ICP algorithm often loses feature points when iterating the camera pose; this results in excessive error, or we can say that the algorithm is unable to converge. This algorithm establishes bat algorithm optimization model by judging the depth of feature points and optimizes the feature point of poses and camera using g2o solver. Experiments show that the method is effective and improves the accuracy of camera pose estimation and the convergence success rate of the algorithm, thus improving the accuracy and robustness of RGB-D visual odometer.

**Keywords:** kinect; lack of depth; fusion algorithm; feature points; iterative closest point; perspective-n-point; depth value; pose estimation; BA optimization model; g2o

近年来, 视觉里程计 (visual odometry, VO)<sup>[1]</sup> 作为视觉 SLAM 的前端, 利用单个或多个视觉传感器, 获取图像的信息进而估计出相机运动, 给后端提供良好的初始值。视觉传感器因其获取信息丰富、成本低、易携带等诸多优点, 被广泛用于

机器人、自动驾驶<sup>[2]</sup>等领域。随着 RGB-D 深度相机的产生, 如微软的 kinect、Pmdtec 的 Camcube 等, 利用 RGB-D 的视觉里程计方法得到越来越广泛的研究。

视觉里程计的实现方法一般按是否需要提取特征分为两种: 1) 特征点法的前端即基于特征的视觉里程计 (FVO); 2) 不提特征的直接法前端即稠密视觉里程计 (DVO)。DVO 一般用图像帧的所有像素信息去估计相机位姿, 用到了图像一致

收稿日期: 2019-03-09. 网络出版日期: 2019-10-14.

基金项目: 国家自然科学基金项目 (61272097); 上海市科技委员会重点项目 (18511101600).

通信作者: 陈强. E-mail: [sues\\_chen@sues.edu.cn](mailto:sues_chen@sues.edu.cn).

性假设,最早由 Steinbruecker 等<sup>[3]</sup>提出,将问题转化为能量最小化处理,但是由于其数据及计算量巨大,无法达到实时目的。相对于直接法前端,基于特征点法的前端一直以来被认为是视觉里程计的主流方法,它具备运行稳定、计算量较小、对光照及动态物体相对不敏感等优点,是目前较为成熟的视觉里程计方案。

FVO 通过寻找图像帧的特征点,通过匹配特征点来估计相机位姿,特征点拥有稳定的局部图像特征,如著名的尺度不变特征变换 (scale-invariant feature transform, SIFT)<sup>[4]</sup>、加速鲁棒特征 (speed-up robust features, SURF)<sup>[5]</sup>、ORB<sup>[6]</sup> 等,根据匹配的特征采用迭代最近点 (ICP) 算法<sup>[7]</sup> 进行位姿估计,为了减少噪声及误匹配带来的影响,往往采用由 Fischler 和 Bolles<sup>[8]</sup> 提出的随机误差一致性 (RANSAC) 算法来进行优化,剔除误差匹配。SIFT 充分考虑了图像在变换过程中的尺度旋转等变换,但伴随的极大计算量是里程计无法接受的,至今普通的 CPU 还无法实时计算 SIFT 特征,而 ORB 特征通过改进 FAST 检测子<sup>[9]</sup>,并使用速度极快的二进制描述子 BRIEF<sup>[10]</sup> 使得实时 SLAM 变成了可能,但 ORB 也存在尺度光照不变性缺失的问题,文献 [11] 提出了改进 FAST 特征,将图像结合高斯金字塔模型实现了 FAST 角点的尺度不变性。文献 [12-13] 提出通过缩小抽样点总量来保证匹配点选取质量,有效地减少迭代次数。

然而,kinect 相机所捕获的深度图像由于遮

挡、吸收、散斑、反射等原因导致深度区域缺失,在实际测试中发现,缺失特征点数在最差情况下将近占总特征点数的三分之一,这就使得单纯使用 ICP 算法产生很大的误差。

针对以上,本文提出一种基于 kinect 相机的改进 RGB-D 视觉里程计方法,由改进 ORB 算法进行 RGB-D 图像特征匹配,对 kinect 所得深度图进行处理,先用 RANSAC 筛选出可靠内点,得到位姿初始估计,接着对深度缺失的 3D-2D 特征点对进行 PnP 优化,对其他 3D-3D 特征点对进行 ICP 优化,并将两者转化为一个 BA(bundle adjustment) 问题建立优化模型,使用 g2o 求解器同时进行特征点及相机位姿优化,最后实验结果证明了该方法的有效性。

## 1 改进的 RGB-D 视觉里程计

本文提出的基于 kinect 相机的改进 RGB-D 视觉里程计框架如图 1 所示,对相机捕获的彩色图用改进的 ORB 算法进行检测与匹配,并对所得深度图进行分析,当检测深度不为 0 时,利用 RANSAC 算法剔除 3D-3D 误差点得到初始位姿估计,采用 ICP 算法进行优化;当检测深度为 0 时,利用 RANSAC 算法剔除 3D-2D 误差点得到初始位姿估计,采用 PnP 算法进行优化。最后,将两者误差项混合转化为一个 BA 问题建立优化模型,使用 g2o 求解器进行特征点及位姿的优化。

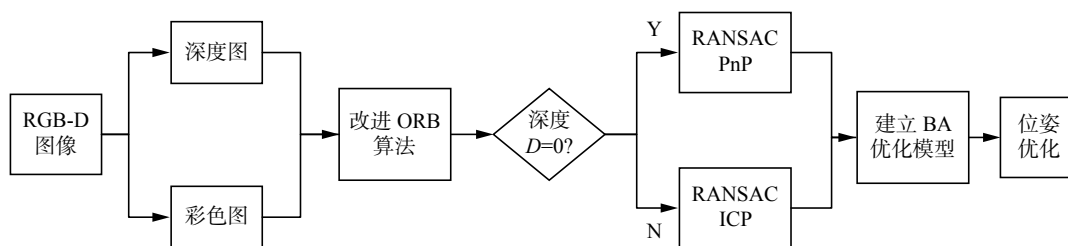


图 1 改进 RGB-D 视觉里程计框架

Fig. 1 Improved RGB-D visual odometer framework

### 1.1 改进的 ORB 算法

ORB 算法是由 Rublee 等提出的一种改进 FAST 角点和 BRIEF 描述子的描述算法,其在特征提取上具有良好的性能。它改进了 FAST 检测子不具有方向性的问题,采用二进制描述子 BRIEF 使得算法具有极快的运算速度,大大加速了图像的特征提取环节,其实时性能远远高于 SURF 与 SIFT,并且在质量与性能之间也有较好的折中,本文在其尺度与光照上的缺陷做了一些改进,使得 ORB 算法在尺度及光照变化上具有一定的健

壮性。

在旋转不变性上,ORB 特征采用灰度质心法 (intensity centroid) 实现,具体步骤如下:

1) 在图像块中定义图像矩为

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x,y), \quad p,q = \{0,1\} \quad (1)$$

式中:  $m_{pq}$  为区域的阶矩表示;  $I(x,y)$  为角点的灰度值。

2) 得到图像的质心  $C$ :

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

3) 连接几何中心与质心, 得到一个方向向量, 定义特征点的方向为

$$\theta = \arctan(m_{01}/m_{10}) \quad (3)$$

以上方法使得 FAST 角点具有了旋转描述, 大大提升了在不同图像之间表达的健壮性, 这种改进的 FAST 角点称为 Oriented FAST。

尺度问题上, 构建高斯图金字塔, 如图 2, 对图像进行不同层次的采样, 以获得不同分辨率图像, 并在金字塔的每一层上检测角点进而实现 ORB 对尺度的健壮性。

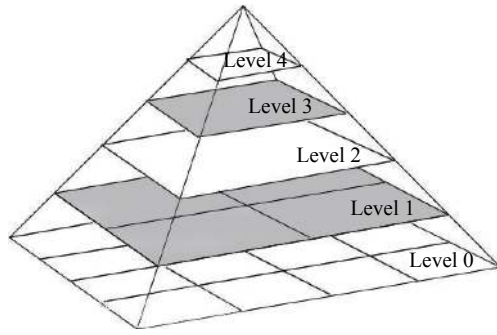


图 2 高斯图像金字塔

Fig. 2 Gaussian image pyramid

在光照问题上, FAST 关键点所选用的灰度阈值  $T$  是固定的, 这就使得当周边环境对比度变化时特征点的提取会受到很大的影响, 采用文献 [14-15] 的方法, 选用一种根据图像对比度的自适应阈值方法, 首先计算图像的对比度:

$$C = \sum_{\delta} \delta(i, j)^2 P_{\delta}(i, j) \quad (4)$$

式中:  $\delta(i, j) = |i - j|$ , 即相邻像素灰度差;  $P_{\delta}(i, j)$  为相邻像素间的灰度差为  $\delta$  的像素分布概率。

由得到的对比度  $C$  设计自适应阈值  $T^*$ :

$$T^* = \alpha C \quad (5)$$

## 1.2 RANSAC 算法完成初始化

PnP 和 ICP 的迭代优化算法都依赖良好的初值选取, 初值选取不当会导致迭代陷入局部最优或无法收敛, 不能达到全局最优效果。RANSAC 是一种随机参数估计算法, 通过反复选取数据中的一组子集来建立参数化模型, 并通过设置阈值来进行内点的筛选, 剔除误差数据, 从而获得最佳模型参数。本文通过 RANSAC 算法分别对 3D-2D 点对和 3D-3D 点对进行粗匹配与初始位姿估计, 采用直接线性变换 (DLT) 建立和求解参数模型。

### 1) 3D-2D 模型

现考虑空间点  $P$ , 它的齐次坐标为

$$P = [X \ Y \ Z \ 1]^T$$

由于在相邻帧中由于匹配点深度信息未知,

但由于相机的内参  $K$  一般都由厂家标定, 为已知信息, 故设其归一化平面齐次坐标为  $p = [u \ v \ 1]^T$ , 此时相机的位姿  $R, t$  未知, 定义增广矩阵  $[R|t]$  为  $3 \times 4$  矩阵, 包含了旋转与平移信息, 形式为

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 & t_3 & t_4 \\ t_5 & t_6 & t_7 & t_8 \\ t_9 & t_{10} & t_{11} & t_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (6)$$

假设有个 3D-2D 特征点, 则可以列出如下的线性方程组:

$$\begin{bmatrix} P_1^T & 0 & -u_1 P_1^T \\ 0 & P_1^T & -v_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_N^T & 0 & -u_N P_N^T \\ 0 & P_N^T & -v_N P_N^T \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = 0 \quad (7)$$

由于  $t$  一共有 12 维, 因此只需要 6 组匹配点即可实现矩阵的线性求解, 得到变换矩阵后测试其他匹配对, 通过设定阈值选取模型内点并剔除误差点, 当内点足够多时则完成特征点的优化并得到 3D-2D 的初始位姿估计  $\xi_P$  ( $\xi_P$  为李代数表达)。

### 2) 3D-3D 模型

相比于 3D-2D 点对, 3D-3D 点对的参数模型建立则简单了许多, 现有空间点  $Q_1 = [X_1 \ Y_1 \ Z_1]^T$ , 与之匹配的 3D 点  $Q_2 = [X_2 \ Y_2 \ Z_2]^T$ , 想要找一个欧式变换  $R, t$ , 使得:

$$Q_1 = RQ_2 + t \quad (8)$$

类似于 3D-2D 模型,  $R, t$  共有 12 维, 而 3D 点对存在 3 个约束, 故只需 4 对匹配点即可完成线性求解, 得到 3D-3D 位姿初始估计  $\xi_Q$ , 求解方法类似 3D-2D 模型。

## 1.3 BA 模型与位姿优化

PnP 和 ICP 算法都是指通过多对 3D 与 2D 或 3D 与 3D 匹配点, 在已知或者未知相机内参的情况下, 利用最小化误差函数来求解相机外参的算法。在得到由 RANSAC 进行优化后的点对和位姿初始估计后, 本文提出一种基于 PnP 和 ICP 的融合 BA 优化的算法。

在 PnP 问题中, 考虑  $N$  个三维空间点  $P$  及其对应的归一化平面点, 希望得到相机的位姿  $R, t$ , 在进行 PnP 问题时本文采用李代数的方法, 因为使用李代数可以构建无约束的优化问题, 从而更方便地通过高斯牛顿法、列文伯格—马夸尔特方法等优化算法进行求解。假设空间点坐标为  $P_i = [X_i \ Y_i \ Z_i]^T$ , 其对应像素坐标为  $p_i = [u_i \ v_i]^T$ , 像素位置与空间点的李代数关系如下:

$$s_i p_i = K \exp(\xi^{\wedge}) P_i \quad (9)$$

由于相机位姿及观测点的噪声, 式 (9) 存在一



个误差,将误差求和构建最小二乘问题,然后优化相机位姿:

$$\mathbf{e}_p = \min_{\xi} \sum_{i=1}^N \left\| \mathbf{p}_i - \frac{1}{s_i} \mathbf{K} \exp(\xi^\wedge) \mathbf{P}_i \right\|^2 \quad (10)$$

式中 $\mathbf{e}_p$ 为像素坐标误差,是二维向量。

在ICP问题中,考虑一组数量为 $M$ 的筛选后的匹配点对 $\mathbf{Q} = [X_i \ Y_i \ Z_i]^\top$ 和 $\mathbf{Q}' = [X'_i \ Y'_i \ Z'_i]^\top$ ,它的李代数位姿表达式为

$$\mathbf{Q}_i = \exp(\xi^\wedge) \mathbf{Q}'_i \quad (11)$$

类似式(10),可以构建误差函数:

$$\mathbf{e}_Q = \min_{\xi} \sum_{i=1}^M \left\| \mathbf{Q}_i - \exp(\xi^\wedge) \mathbf{Q}'_i \right\|^2 \quad (12)$$

式中 $\mathbf{e}_Q$ 为空间坐标误差,是三维向量。

将式(10)与(12)融合为一个BA问题,构成五维代价函数:

$$\mathbf{e} = \min_{\xi} \begin{pmatrix} \frac{1}{N} \mathbf{e}_p \\ \frac{1}{M} \mathbf{e}_Q \end{pmatrix} = \begin{pmatrix} \frac{1}{N} \sum_{i=1}^N \left\| \mathbf{p}_i - \frac{1}{s_i} \mathbf{K} \exp(\xi^\wedge) \mathbf{P}_i \right\|^2 \\ \frac{1}{M} \sum_{i=1}^M \left\| \mathbf{Q}_i - \exp(\xi^\wedge) \mathbf{Q}'_i \right\|^2 \end{pmatrix} \quad (13)$$

其中, $\mathbf{e}$ 为五维向量,下面对式(13)进行优化求解。

在使用高斯牛顿法求解之前,需要知道每个误差项关于优化变量 $\xi$ 的一阶导数,即雅克比矩阵 $\mathbf{J}$ 以及 $\xi$ 的初始位姿 $\xi_0$ 。 $\xi_0$ 的取值为

$$\xi_0 = \frac{N}{M+N} \xi_p + \frac{M}{M+N} \xi_Q \quad (14)$$

关于 $\mathbf{J}$ 的形式是本文关键所在,在 $\mathbf{e}_p$ 中,像素坐标误差为二维,位姿为六维, $\mathbf{J}_p$ 将是 $2 \times 6$ 的矩阵;在 $\mathbf{e}_Q$ 中,空间坐标误差为三维, $\mathbf{J}_Q$ 将是 $3 \times 6$ 的矩阵,故误差 $\mathbf{e}$ 的雅克比矩阵为 $5 \times 6$ 矩阵。

首先分析 $\mathbf{J}_p$ 的形式,记变换到相机坐标系下

$$\mathbf{J} = \begin{pmatrix} \sum_{i=1}^N \frac{f_x}{Z'_i} & 0 & -\sum_{i=1}^N \frac{f_x X'_i}{Z'^2_i} & -\sum_{i=1}^N \frac{f_x X'_i Y'_i}{Z'^2_i} & N f_x + \sum_{i=1}^N \frac{f_x X'^2_i}{Z'^2_i} & -\sum_{i=1}^N \frac{f_x Y''_i}{Z'^2_i} \\ 0 & \sum_{i=1}^N \frac{f_y}{Z'_i} & -\sum_{i=1}^N \frac{f_y Y'_i}{Z'^2_i} & -N f_y - \sum_{i=1}^N \frac{f_y Y'^2_i}{Z'^2_i} & \sum_{i=1}^N \frac{f_y X'_i Y'_i}{Z'^2_i} & \sum_{i=1}^N \frac{f_y X'_i}{Z'^2_i} \\ M & 0 & 0 & 0 & -\sum_{i=1}^M Z''_i & \sum_{i=1}^M Y''_i \\ 0 & M & 0 & \sum_{i=1}^M Z''_i & 0 & -\sum_{i=1}^M X''_i \\ 0 & 0 & M & -\sum_{i=1}^M Y''_i & \sum_{i=1}^M X''_i & 0 \end{pmatrix}$$

得到了初始位姿 $\xi_0$ 与雅克比矩阵 $\mathbf{J}$ 后调用高斯牛顿法,算法流程如下:

1) 给定初始值 $\xi_0$ 。

的空间坐标为 $\mathbf{P}' = [X' \ Y' \ Z']^\top$ ,它的相机投影模型为

$$u = f_x \frac{X'}{Z'} + c_x, \quad v = f_y \frac{Y'}{Z'} + c_y \quad (15)$$

其中, $f_x$ 、 $f_y$ 、 $c_x$ 、 $c_y$ 为相机内参,为相机出厂时标定。

在求导上本文使用李代数的左乘扰动模型<sup>[16]</sup>,对 $\xi^\wedge$ 左乘扰动量 $\delta\xi$ ,然后使用链式法则,列写如下:

$$\frac{\partial \mathbf{e}_p}{\partial \delta \xi} = \lim_{\delta \xi \rightarrow 0} \frac{\mathbf{e}_p(\delta \xi \oplus \xi)}{\delta \xi} = \frac{\partial \mathbf{e}_p}{\partial \mathbf{P}'} \frac{\partial \mathbf{P}'}{\partial \delta \xi} \quad (16)$$

式中: $\oplus$ 指李代数上的左乘扰动;右边两项,其中第1项是关于投影点的导数,由式(15)可得

$$\frac{\partial \mathbf{e}_p}{\partial \mathbf{P}'} = - \begin{pmatrix} \frac{f_x}{Z'} & 0 & -\frac{f_x X'}{Z'^2} \\ 0 & \frac{f_y}{Z'} & -\frac{f_y Y'}{Z'^2} \end{pmatrix} \quad (17)$$

第2项是误差关于位姿李代数的导数,由文献[12]可推导出:

$$\frac{\partial \mathbf{P}'}{\partial \delta \xi} = [\mathbf{I}, -\mathbf{P}'^\wedge] \quad (18)$$

将式(17)与(18)相乘即可得单个误差雅克比矩阵 $\mathbf{J}_p$ :

$$\frac{\partial \mathbf{e}_p}{\partial \delta \xi} = - \begin{pmatrix} \frac{f_x}{Z'} & 0 & -\frac{f_x X'}{Z'^2} & -\frac{f_x X' Y'}{Z'^2} & f_x + \frac{f_x X'^2}{Z'^2} & -\frac{f_x Y'}{Z'^2} \\ 0 & \frac{f_y}{Z'} & -\frac{f_y Y'}{Z'^2} & -f_y - \frac{f_y Y'^2}{Z'^2} & \frac{f_y X' Y'}{Z'^2} & \frac{f_y X'}{Z'^2} \end{pmatrix}$$

$\mathbf{J}_Q$ 的推导与之相似,空间点 $\mathbf{Q}' = [X' \ Y' \ Z']^\top$ 经位姿变换为 $\mathbf{Q}'' = [X'' \ Y'' \ Z'']^\top$ ,使用李代数扰动模型即可得

$$\frac{\partial \mathbf{e}_Q}{\partial \delta \xi} = - \begin{pmatrix} 1 & 0 & 0 & 0 & -Z'' & Y'' \\ 0 & 1 & 0 & Z'' & 0 & -X'' \\ 0 & 0 & 1 & -Y'' & X'' & 0 \end{pmatrix}$$

综上,可得由 $N$ 个3D-2D点对与 $M$ 个3D-3D点对组成的BA代价函数 $\mathbf{e}$ 的雅克比矩阵 $\mathbf{J}$ 为

2) 对于第 $k$ 次迭代,求出当前的雅克比矩阵 $\mathbf{J}$ 与误差 $\mathbf{e}$ 。

3) 求解增量方程 $\mathbf{H} \Delta \xi_k = \mathbf{g}$ 。

4)若  $\Delta\xi_k$  足够小,则停止。否则,令  $\xi_{k+1} = \xi_k + \Delta\xi_k$ , 返回 2)。

## 2 实验与结果

### 2.1 实验环境

为了对本文算法的有效性进行验证,实验计算机的配置为: Intel i5 双核 1.8 GHz CPU, 4 GB 内存, 500 GB 硬盘, NVIDIA 独立显卡, 操作系统为 64 位的 ubuntu14.04, 算法基于 C++、OpenCV3.0 和 g2o 库进行编写。

### 2.2 改进 ORB 特征点对比

在基于稀疏特征点法的里程计中,特征点的数量是反映计算准确性的重要指标之一。与传统 ORB 相比,本文采用改进 ORB 算法在 FAST 特征点提取上具有更好的适应性。原 FAST 特征点识别方法与改进后的特征点识别方法如图 3 所示。



(a) 原 ORB 特征点检测法



(b) 改进的 ORB 特征点检测法

图 3 特征点对比结果

Fig. 3 Feature point comparison result

从图 3 可以看出,对于同一关键帧,当光照强度不是很强时,改进之后的算法明显适应性更强,相比原算法多检测出了约 20% 的特征点,增强了整个系统的鲁棒性。

### 2.3 RANSAC 内点筛选

在基于稀疏特征点法的里程计中,特征点匹配精度决定了里程计的性能,图 4 为室外道路环境的两帧图像,图 4(a) 为未经 RANSAC 筛选的 ORB 特征点匹配,可以看到有许多特征点的误匹配,这会对优化算法造成极大的误差,图 4(b) 为经 RANSAC 筛选后的特征匹配,匹配精度得到很大提高。



(a) RANSAC 筛选前



(b) RANSAC 筛选后

图 4 RANSAC 筛选结果

Fig. 4 RANSAC screening result

### 2.4 直线运动实验

本实验是在 Turtlebot 移动机器人平台上进行的,设定直线运动速度 0.3 m/s,运动距离为 3 m。分别进行传统 ICP 算法、RANSAC+ICP 算法和融合算法估计移动机器人运动轨迹,并使用 g2o 求解器进行位姿优化求解。g2o 是一个在 SLAM 领域广为使用的优化库,可以将非线性优化与图论相结合,大大提高系统的实时性与精确性,运动轨迹如图 5 所示,结果表明,融合算法的平均误差为 1.98%,相比于传统 ICP 的 8.64% 与改进后的 RANSAC+ICP 算法的 5.28% 有着更小的误差。

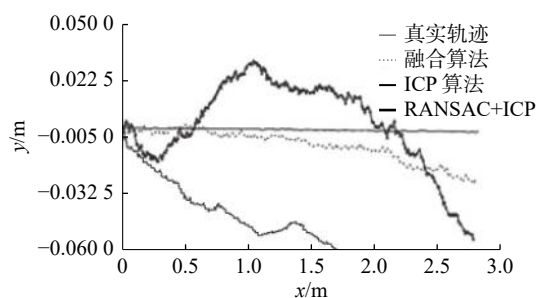


图 5 直线运动轨迹

Fig. 5 The trajectory of linear motion

从表 1 可以看出,在匹配点对上,直接 ICP 将所有 ORB 特征点都进行迭代优化, RANSAC 筛选后去除了近乎三分之一的误差点,本文算法在此基础上加上了被剔除的深度缺失的特征点;在迭代次数上可以看到相比直接 ICP 算法 RANSAC 筛选过的 ICP 与本文算法都有了迭代次数明显地减少;在收敛成功次数上本文算法存在显著优势,几乎达到了百分百成功率,大大提高了里程计的鲁棒性;在优化时间上本次实验取的为单一迭代优化的时间,并没有加上特征提取与匹配所花的时间,3 种方法优化时间相对比较接近,整体

里程计算法一般为 30 ms 左右;在 3 种方法的相对误差上可以看出,相比直接 ICP 与 RANSAC+ICP,本文算法在估计相机误差精度上有着显著提高。

表 1 3 种方法性能对比

Table 1 Performance comparison of three methods

位姿估计算法	100次实验平均值				
	匹配 点对/对	迭代 次数	收敛 成功/次	优化 时间/ms	相对 误差/%
直接ICP	128.74	21.35	65	2.256	8.64
RANSAC+ ICP	72.48	12.51	81	1.825	5.28
本文算法	89.60	11.73	98	1.934	1.98

### 3 结束语

本文提出一种融合 PnP 与 ICP 的 RGB-D 视觉里程计方法,有效针对 kinect 深度相机深度区域丢失的缺点,利用改进 ORB 算法进行特征点提取与匹配,再利用 RANSAC 算法得到良好的相机位姿估计初始值,根据特征点深度值的大小进行 PnP 和 ICP 模型建立,并得到一个 BA 优化模型,利用 g2o 进行迭代求解,实验证明了该方法优秀的收敛成功率以及匹配相机位姿的精度,大大提高了视觉里程计的精确性与鲁棒性。

### 参考文献:

- [1] SCARAMUZZA D, FRAUNDORFER F. Visual odometry: part I: the first 30 years and fundamentals[J]. IEEE robotics & automation magazine, 2011, 18(4): 80–92.
- [2] HÄNE C, HENG L, LEE G H, et al. 3D visual perception for self-driving cars using a multi-camera system: calibration, mapping, localization, and obstacle detection[J]. *Image and vision computing*, 2017, 68: 14–27.
- [3] STEINBRÜCKER F, STURM J, CREMERS D. Real-time visual odometry from dense RGB-D images[C]//Proceedings of 2011 IEEE International Conference on Computer Vision Workshops. Barcelona, Spain: 2011: 719–722.
- [4] NG P C, HENIKOFF S. SIFT: predicting amino acid changes that affect protein function[J]. *Nucleic acids research*, 2003, 31(13): 3812–3814.
- [5] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. *Computer vision and image understanding*, 2008, 110(3): 346–359.
- [6] RUBLEE E, RABAUDE V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of 2011 IEEE International Conference on Computer Vision. Barcelona, Spain: 2011: 2564–2571.
- [7] BESL P J, MCKAY N D. A method for registration of 3-D shapes[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1992, 14(2): 239–256.
- [8] 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 ACM*, 1981, 24(6): 381–395.
- [9] ROSTEN E, PORTER R, DRUMMOND T. Faster and better: a machine learning approach to corner detection[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2010, 32(1): 105–119.
- [10] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: binary robust independent elementary features[C]//Proceedings of the 11th European Conference on Computer Vision. Heraklion: Springer, 2010: 778–792.
- [11] 翟紫伶. 基于高斯尺度空间理论的 ORB 特征点检测[J]. *电子世界*, 2016(10): 192–194.
- [12] 葛山峰, 于莲芝, 谢振. 基于 ORB 特征的目标跟踪算法[J]. *电子科技*, 2017, 30(2): 98–100, 104.  
GE Shanfeng, YU Lianzhi, XIE Zhen. A target tracking algorithm based on ORB[J]. *Electronic science and technology*, 2017, 30(2): 98–100, 104.
- [13] 邢凯盛, 凌有铸, 陈孟元. ORB 特征匹配的误匹配点剔除算法研究[J]. *电子测量与仪器学报*, 2016, 30(8): 1255–1262.  
XING Kaisheng, LING Youzhu, CHEN Mengyuan. Mismatching points elimination algorithm for ORB feature matching[J]. *Journal of electronic measurement and instrumentation*, 2016, 30(8): 1255–1262.
- [14] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]//Proceedings of 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal, 2012: 573–580.
- [15] KERL C, STURM J, CREMERS D. Dense visual SLAM for RGB-D cameras[C]//Proceedings of 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, Japan, 2013: 2100–2106.
- [16] 高翔, 张涛. 视觉 SLAM 十四讲[M]. 北京: 电子工业出版社, 2017: 72–76.

### 作者简介:



朱俊涛, 硕士研究生, 主要研究方向为机器人。



陈强, 教授, 主要研究方向为城市管网探测理论与方法、仪器设备研发与应用。获得多项省部级科技进步奖与科技成果。发表学术论文 50 余篇。