

一种 SURF-FREAK 特征误差模型视觉里程计

刘军,柴宏旭,赵威

(兰州理工大学机械电子工程学院,甘肃兰州 730050)

摘要:针对深度相机在定位过程中存在的实时性差、定位精度低等问题,该文结合深度相机的特点,在传统 ICP 算法的基础上融合 FREAK-SURF 特征匹配点。首先,为了提高匹配速度,采用二进制描述符 FREAK 描述子简化传统的 SURF 特征点描述信息完成图像粗匹配,并结合对极几何约束预估相机初始位姿,为后续精匹配环节提供一个良好的初值。其次,在完成粗配准的基础上将特征点的空间位置和相机位姿误差模型引入点云配准算法中,并且为了弥补微小扰动和空间尺度对系统的影响,构造一种欧氏、马氏距离相融合的距离度量函数对匹配点进行度量,构建出视觉里程计累计误差模型。最后,为增强特征点空间位置准确性与降低视觉里程计累计误差,在精匹配后采用卡尔曼滤波对特征点空间位置进行更新,以提高特征点位置与相机位姿精度。实验验证表明,该研究对深度相机视觉里程计问题的处理上,实时性优于主流的 PnP-BA 算法和 SVD-ICP 算法,精度高于传统的 SVD-ICP 算法。

关键词:视觉里程计;特征点空间误差模型;SURF-FREAK;ICP 算法;先验概率模型;卡尔曼滤波

中图分类号:TP242

文献标识码:A

文章编号:1673-629X(2022)03-0008-07

doi:10.3969/j.issn.1673-629X.2022.03.002

A Visual Odometry Incorporating Uncertainty of SURF-FREAK Feature Points

LIU Jun, CHAI Hong-xu, ZHAO Wei

(School of Mechatronics Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aiming at the poor real-time performance and low positioning accuracy of the depth camera in the positioning process, we combine the characteristics of the depth camera and fuse the FREAK-SURF feature matching points on the basis of the traditional ICP algorithm. Firstly, in order to improve the matching speed, the binary descriptor FREAK is used to simplify the traditional SURF feature point description information to complete the rough image matching. In addition, the initial pose of the camera is estimated by combining with the epipolar geometric constraints to provide a good initial value for the subsequent fine matching. Secondly, based on the completion of the rough registration, the spatial position of the feature points and the camera pose error model are introduced into the point cloud registration algorithm, and in order to compensate for the influence of the small disturbance and the spatial scale on the system, a distance metric function combining Euclidean and Mahalanobis distance is constructed to measure the matching point pair, and a visual odometer accumulative error model is constructed. Finally, in order to enhance the spatial position accuracy of the feature points and reduce the cumulative error of the visual odometer, Kalman filter is used to update the spatial position of the feature points after fine matching, so as to improve the accuracy of the feature point positions and the camera pose. Experimental verification shows that the real-time performance of this research is better than the mainstream PnP-BA algorithm and SVD-ICP algorithm, and the accuracy is higher than the traditional SVD-ICP algorithm.

Key words: visual odometry; feature point spatial error model; SURF-FREAK; ICP algorithm; prior probability model; Kalman filtering

0 引言

近年来,随着自动化水平的提高,智能机器人已经成为生产生活中的一部分,如扫地机器人、无人机等。无人设备在未知环境下执行任务可通过自身携带的传

感器获取环境信息从而完成构建导航地图、路径规划等相关任务。而视觉传感器不仅能获得丰富的图像信息,且具有质量轻、价格低、体积小等优势,在机器人导航方面得到了广泛应用。

收稿日期:2021-03-10

修回日期:2021-07-13

基金项目:国家自然科学基金(71861025)

作者简介:刘军(1974-),男,教授,博士,研究方向为智能机器、深度学习、神经网络等;通信作者:柴宏旭(1993-),男,硕士研究生,研究方向为机器视觉导航。

视觉里程计 (visual odometry, VO) 是智能机器人同步定位与地图构建 (simultaneous localization and mapping, SLAM) 的关键技术之一^[1-2], 可实现移动机器人在未知环境下的自主定位, 为环境地图模型的建立与机器人导航提供位置信息。该技术起源于二十世纪八十年代, 自 2004 年 David Nister 发表了一篇以视觉里程计为标题的论文^[3], 并于同年成功应用于美国国家航空航天局 (NASA) 主导的火星漫游者勇气号和机遇号上后, 掀起了新一轮的研究高潮。经过众多学者的努力, 当前其主流技术主要发展为三大类: 直接法^[4-5]、特征点法^[6-7]、半直接法^[8]。较典型且已开源的 A. J. Davison 所提 MonoSLAM 方案的视觉里程计技术^[5] 即属于直接法, 其技术特点是直接通过计算像素灰度值的差异来确定相机运动。但直接法在过程中非常容易丢失特征点从而导致相机定位失败。特征点法主要基于特征点进行图像匹配, 相比于直接法更稳定。较典型的如 2015 年 Raul Mur-Artal 所提出的 ORB-SLAM^[7], 以 ORB 特征点贯穿于系统的特征处理、匹配、回环检测等方面, 在全局优化基础之上可以获得较准确、全局一致的相机运动轨迹。半直接法将特征点法和直接法相结合, 在关键帧中提取特征点, 但在普通帧上通过计算灰度值差异进行特征跟踪, 如文献^[8]等。该技术虽具有较快的单帧处理速率以及良好的实时性, 但在定位精度以及鲁棒性上存在不足之处。

上述视觉里程计方案主要基于 2D 图像模型, 因此容易受到环境因素的影响, 如相机旋转、图像模糊、弱纹理区域等等, 易导致定位失败, 因而难以处理复杂环境或难以满足复杂任务要求场景的建模要求。而深度相机近年来得到了广泛关注与应用^[9], 其除图像信息之外也提供深度信息, 可较好地弥补和解决上述问题^[10]。目前基于深度相机的视觉里程计有两种技术模型: 一种为 3D-2D 点位姿求解, 如 PnP (perspective-n-point) 模型, 通过已知三维空间点和图像中所对应投影点的位置关系来估计相机位姿; 另一种为 3D-3D 点云位姿求解, 即通过在不同坐标系下观测到的同一三维空间点集信息来估计相机位姿, 例如 Besl、Mckay 等人^[11]提出的迭代最近点算法等, 其算法是基于四元数来定义相机位置, 从测量点确定其对应就近点的点集后, 通过构建代价函数 (cost function) 进行迭代计算, 直至达到允许误差从而结束迭代过程。因此, 为提高处理效率, 许多学者结合深度相机的特点提出了诸多改进算法, 例如, Siemiatkowska 等人在 2D 图像上提取特征点进行粗配准, 然后将粗配准得到的变换矩阵作为精配准的初始值, 通过改进算法初始值的方式减小迭代次数以提高效率^[12]。Ji H 等人提出了一种

ColorICP 算法, 利用三维点的颜色信息, 在寻找最近点的时候, 只在颜色相近的区域寻找最近点作为匹配点, 可有效减少计算量, 提高匹配效率^[13]。其他围绕提高计算效率方面的努力如文献^[14-16]等, 但上述改进要么效率提升有限, 要么往往都以牺牲精度为代价, 存在不同的技术缺陷, 如光照变化影响大、空间旋转尺度难以保证等等, 很难兼顾精度与效率, 在此不再一一赘述。

基于此, 为兼顾精度与效率, 受文献^[17]启发, 该文提出了一种新的视觉里程计技术。该技术将 SURF 特征点提取方式与 FREAK 描述子相结合, 在粗配准确定特征点时, 为保证精度, 采纳 SURF 特征点提取方式以保持空间尺度不变性, 同时采用 FREAK 描述子, 即以先行检查前 128 位描述信息而非对全部描述信息进行检查的方式提高检查效率。同时在精配准时, 在概率与统计学意义下, 不同于传统采纳欧氏距离对特征点距离进行度量的方式, 而是结合实际情况将所有特征点的位置误差也引入到距离函数中, 即提出了一种欧氏、马氏距离相混合的距离函数对特征点距离进行度量, 该种方式更符合实际情况, 既有效度量了特征点的概率位置, 也考虑了其位置方差的影响; 并且, 为进一步增强特征点空间位置准确性与降低视觉里程计累计误差, 在精匹配后采用卡尔曼滤波对特征点空间位置进行更新, 以提高特征点位置与相机位姿精度。尤其不同于传统视觉里程计技术仅给出相机概率位置期望, 该文在上述改进的基础上, 给出了相机位置的方差计算公式。标准数据集测试实验表明, 该技术能有效兼顾实时性和准确性, 并能较好地对相机运动轨迹进行估计。

1 视觉里程计问题描述

视觉里程计是通过视觉方式还原出相机运动轨迹。该文在考虑视觉深度信息误差的基础之上, 提出了一种新型视觉里程计技术, 以解决当前基于深度相机的视觉里程计技术在小型自主移动设备上使用时的精度与效率协调问题。通过对相机采集到的图像进行处理, 计算出相机每一时刻运动位姿, 即公式(1), 从而构建出相机运动轨迹。

令世界坐标系 (用上标 w 表示) 是一个固定的参照坐标用来描述相机和物体的位置; 相机坐标系 (用上标 c 表示) 是相机视角下建立的坐标系; 像素坐标 (u, v) 是在相机采集的图像上建立的二维坐标。相机初始位姿坐标定义为世界坐标, 相机坐标后续不断变化。相机间后续位姿变化用一个 4×4 的变换矩阵表示:

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

每一时刻在不同位姿处,空间特征点与相机特征点的转换关系通过公式(2)呈现。设相机坐标系下的特征点为 P^c , 采集图像投影点为 p , 像素坐标记为 u_i , 此坐标可由变换矩阵 T 转换到世界坐标系下:

$$P_i^w = RP_i^c + t \quad (2)$$

其中, P_i^w 、 P_i^c 分别为世界坐标系和相机坐标系下的坐标。

2 一种融合 SURF-FREAK 特征点误差模型视觉里程计技术

传统的视觉里程计建立图像噪声扰动模型,但未考虑 Kinetic 相机采集深度数据时存在噪声,会导致深度值 d 存在误差,对位姿估计的准确性有较大的影响,甚至会导致机器定位失败。因此,受文献[17]启发,该文在其基础上采取高斯误差模型描述 Kinect 深度值的误差;并且为了快速获得准确的相机位姿,先行通过图像配准环节进行粗配准,为后续精配准环节提供初始值,减少精配准目标函数迭代次数,提高效率。

2.1 基于 SURF-FREAK 图像特征匹配

在粗配准环节,该文结合 SURF 特征点提取方式和 FREAK 描述信息的优势,提高特征点准确性和效率。相较于传统的 FAST 提取方式,因通过不同尺度的盒式滤波(box filter, BF)与图像卷积构建出多尺度空间,在图像中稳定的特征点进行检测筛选;并且 SURF 特征点描述子在匹配时计算量大,结合 FREAK 描述子通过二进制描述符,以先行匹配 128 位描述信息而非对全部描述信息进行检查的方式提高效率。

2.1.1 SURF 特征提取

SURF 特征检测是通过计算 Hessian 矩阵行列式,以局部曲率极大值来检测特征点的区域,并通过盒式滤波(box filter)构建多尺度空间,对特征点和原始图像做卷积,进行不同尺度空间的极值检测,从而生成图像稳定的边缘点(突变点),用于特征点提取,使其具备多尺度不变性。经过 Hessian 矩阵处理的每个像素点与其图像域和尺度域的所有相邻点进行比较寻找极值点,再通过盒式滤波器构建多尺度空间,滤除掉不具备多尺度不变性的极值点,筛选出最终稳定的特征点。传统 SURF 特征点在匹配时,建立特征点主方向,通过在特征点周围选取一个正方形框,框的边长为 20 倍特征点像素大小。然后将该框平分为 16 个子区域,在每个子区域计算 25 个像素的水平方向和垂直方向的 haar 小波特征,并对水平方向和垂直方向计算绝对值

之和,构造描述向量。以此作为图像特征区域描述信息,对图像进行匹配。

SURF 特征点匹配如图 1 所示。

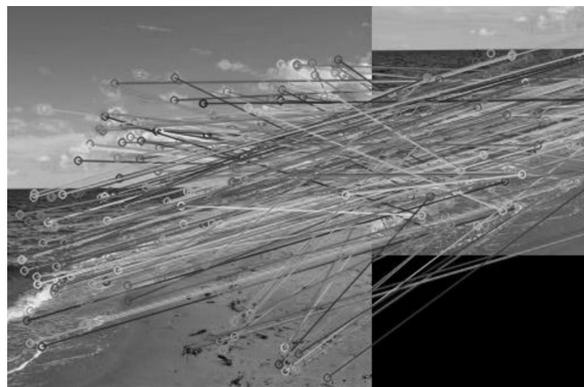


图 1 SURF 特征点匹配

SURF 特征点可以通过提高阈值来提高匹配准确性,但 SURF 特征需要计算每个尺寸的 haar 小波响应,计算量较大;此外图像配准时,通过计算特征点周围描述向量间的欧氏距离,确定特征点匹配度。SURF 特征点描述子虽然精度高,但存在计算量大、效率低等问题。基于此,该文通过 FREAK 二进制描述代替 SURF 特征描述的方式来提高效率。

2.1.2 FREAK 描述子及匹配

图像通过 SURF 完成特征提取,通过二进制描述符 FREAK 描述子代替 SURF 描述子,以降低计算量,提高效率。相较于 SURF 特征描述子,一方面通过二进制编码来代替 haar 小波响应,以提高运算速度;另一方面 FREAK 描述子模拟人眼,根据采样点与特征点的距离远近,采用了不同大小的高斯核函数进行平滑处理,通过重叠的感受野,可以获得更多的信息,使最终的描述符更具独特性和可辨识性。

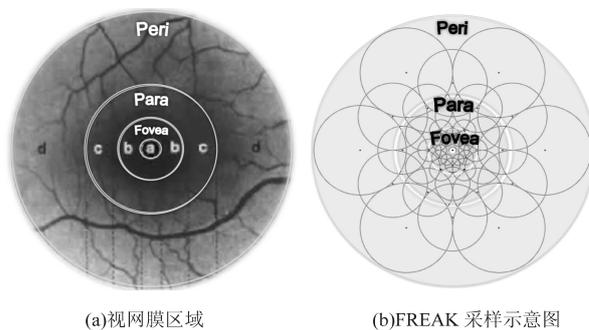


图 2 FREAK 描述子

FREAK 描述子^[18]是通过模拟人眼视网膜对特征点周围的图像信息进行处理。如图 2(a)所示,人眼的视网膜根据感光细胞的密度分成了三个区域,即中央凹(fovea)、旁中央凹(parafoveal)、周边区(perifoveal)。fovea 区域感光细胞密度大,能够获得较高的分辨率,主要用于细节识别;perifoveal 区域感光细胞密度较少可获得低分辨率的图像,可用于轮廓识别。FREAK 描

述子就是借鉴了人类视觉分区获取信息而提出的一种算法,FREAK 描述子采样像素分布在多个同心圆,采样像素的分布是按照指数形式的分布密度分布的,越接近同心圆的分布密度越大,如图 2(b)所示。

二进制编码 FREAK 特征描述算法,编码过程见公式(3):

$$F = \sum_{i=0}^N 2^i T(P_i)$$

$$T_i = \begin{cases} 1, & I(P_i^0) > I(P_i^1) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中, $I(P_i^j)$ 表示采样点经过高斯模糊后的灰度值。FREAK 的采样模式中一共有 43 个采样点,可以产生 $N = 43(43 - 1)/2 = 903$ 个采样点对,过多的采样点会造成特征冗余,因此需要对特征的描述向量进行筛选,也即降维。

最终,根据计算得到了 512 bit 的二进制描述符,该描述符的列是方差由高到低的排列,而高方差表征了模糊信息,低方差表征了细节信息。因此,选取前 128 bit 即 16 bytes 进行匹配,若两个待匹配的特征点前 16 bytes 距离小于设定的阈值,则再用剩余的比特位进行匹配。第一步匹配可以剔除掉 90% 的不相关匹配点,很大程度上提高了匹配的速度。通过融合 SURF-FREAK 特征提取描述算法,对室内区域进行环境匹配,如图 3 所示。



图 3 SURF-FREAK 图像匹配

至此相机通过图像匹配技术完成粗匹配,结合图像间对极几何约束预估相机初始位姿,为后续精匹配环节提供了一个较好的初值,有效缩短了迭代次数,降低了计算复杂度。

2.2 基于高斯分布的特征点云帧间配准模型

传统的视觉里程计只考虑图像噪声扰动模型,但未考虑 RGB-D 相机采集深度数据时存在噪声,导致深度值 d 存在误差,对位姿估计的准确性有较大的影响。因此,该文采取高斯误差模型描述 RGB-D 相机深度采集图像的误差,以便更好地估计相机的位姿变化。由于深度相机采集的数据存在噪声,会导致像素坐标 (u, v) 及深度值 d 存在误差,现作如下假设:

(1) u 、 v 相互独立,且服从正态分布,即 $u \sim N(\mu_u, \sigma_u)$, $v \sim N(\mu_v, \sigma_v)$;

(2) 相机相邻时刻观测值相互独立;

(3) 同一时刻图像提取不同特征点相互独立;

(4) 各个像素对应的深度值 z 服从正态分布,即

$$z \sim N(\hat{\mu}_z, \hat{\sigma}_z^2)。$$

像素 (u, v) 处 z 的均值和方差为:

$$\begin{cases} \mu_z = \sum_{i,j} G_{i,j} \hat{\mu}_z \\ \sigma_z^2 = \sum_{i,j} (\hat{\sigma}_{z_{i,j}}^2 + \hat{\mu}_{z_{i,j}}^2) - \mu_z^2 \end{cases}$$

$$i \in [\mu_u - 1, \mu_u + 1], j \in [\mu_v - 1, \mu_v + 1] \quad (4)$$

空间点 $P(x, y, z)$ 的位置期望为:

$$\mu = (\mu_x, \mu_y, \mu_z)^T$$

$$\begin{cases} \mu_x = \frac{\mu_z(\mu_u - c_x)}{f_x} \\ \mu_y = \frac{\mu_z(\mu_v - c_y)}{f_y} \end{cases} \quad (5)$$

且 μ_z 满足协方差矩阵,可表示为:

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 \end{bmatrix} \quad (6)$$

将图像特征点误差模型通过高斯表达,位置期望 μ 表示特征点位置理想情况,而协方差矩阵表达特征点位置在期望附近分布模型。在精配准环节,通过建立累积误差模型,并对其优化降低协方差矩阵数值,提高位置准确性。

3 点云图像帧间精配准模型

本节根据图像局部特征匹配为精配准环节中 ICP 算法提供初值,结合 RGB-D 相机图像高斯误差模型,改进传统 ICP 算法的距离函数,构建出新型的 ICP 算法目标函数,并推导出相机运动位置的期望和协方差。

3.1 ICP 算法帧间配准计算

首先,将已配准特征点带入 ICP 模型计算。使用特征描述向量 x , 匹配到新一帧与历史模型集中的相似特征点,进行状态估计。

该文结合 2.2 节世界坐标系下特征点位置方差的描述,引入 Kinect 相机位姿方差,使其局部特征点位置不确定性模型具有更加完整的描述:

$$\begin{cases} \mu_p^W = R\mu_p^C + \mu_t \\ \Sigma_p^W = R\Sigma_p^C R^T + \Sigma_t \end{cases} \quad (7)$$

其中, μ_p^W , μ_p^C , Σ_p^W , Σ_p^C 分别为世界坐标系 W 和相机坐标系 C 下特征点 P 的期望和协方差阵, μ_t , Σ_t 表示相机 t 时刻的期望和协方差阵。

然后,将相邻关键帧的特征匹配点集带入 ICP 算法进行配准。其中 ICP 匹配算法中的传统欧氏距离不能体现特征点位置的空间方差,而马氏距离是一种有

效计算相互匹配的未知样本集合之间的相似度计算方法。用马氏距离作为 ICP 算法配准准则时,对微小的扰动非常敏感,会夸大扰动或误匹配带来的影响,且缺乏描述空间距离的几何意义和直观性。因此,ICP 算法的配准特征点之间的距离函数为:

$$d(P_{k-1}^w, P_k^w) = \sqrt{\Delta_{P_{k-1}, P_k}^T (I + \Sigma_{P_{k-1}}^w + \Sigma_{P_k}^w) \Delta_{P_{k-1}, P_k}} \quad (8)$$

其中,

$$\begin{cases} \Delta_{P_{k-1}, P_k} = \mu_{P_{k-1}}^w - \mu_{P_k}^w \\ \mu_{P_k}^w = R_k \mu_{P_k}^c + \mu_{t_k} \\ \Sigma_{P_k}^w = R_k \Sigma_{P_k}^c R_k^T + \Sigma_{t_k} \end{cases}$$

式中, $\mu_{P_{k-1}}^w$ 、 $\mu_{P_k}^c$ 、 $\Sigma_{P_{k-1}}^w$ 、 $\Sigma_{P_k}^c$ 分别表示特征点 P 在 $k-1$ 时刻及 k 时刻世界坐标系、相机坐标系的位置期望和协方差阵。

3.2 视觉里程计累计误差优化模型

根据公式(8),构造世界坐标系下相机位姿累计误差模型,并求出最优的相机位置期望及协方差。

设当前场景和模型集中参与 ICP 配准的特征点一共为 n 对,其中 k 时刻的第 i 个特征点的位置期望和协方差矩阵分别用 $\mu_{P_{i,k}}$ 、 $\Sigma_{P_{i,k}}$ 表示。

根据 ICP 算法,相机位置期望 $\mu_{P_{i,k}}$ 应满足:

$$\min J = \sum_{i=1}^n \Delta_{P_{(i,k-1)}, P_{(i,k)}}^T (I + \Sigma_{P_{i,k-1}}^w + \Sigma_{P_{i,k}}^w)^{-1} \Delta_{P_{(i,k-1)}, P_{(i,k)}} \quad (9)$$

通过对上式用 μ_{t_k} 求偏导,计算期望极值点:

$$\frac{\partial J}{\partial \mu_{t_k}} = 2 \sum_{i=1}^n (I + \Sigma_{P_{i,k-1}}^w + \Sigma_{P_{i,k}}^w)^{-1} \quad (10)$$

$$(\mu_{P_{i,k-1}}^w - R_{i,k} \mu_{P_{i,k}}^c - \mu_{t_k}) = 0$$

整理可得,相机位置期望 μ_{t_k} 应满足:

$$\mu_{t_k} = \left[\sum_{i=1}^n (I + \Sigma_{P_{i,k-1}}^w + \Sigma_{P_{i,k}}^w)^{-1} \right]^{-1} \cdot \left[\sum_{i=1}^n (I + \Sigma_{P_{i,k-1}}^w + \Sigma_{P_{i,k}}^w)^{-1} (\mu_{P_{i,k-1}}^w - R_{i,k} \mu_{P_{i,k}}^c) \right] \quad (11)$$

根据协方差的定义,相机位置协方差阵 Σ_{t_k} 应满足:

$$\Sigma_{t_k} = E[(t_k - \mu_{t_k})(t_k - \mu_{t_k})^T] \quad (12)$$

已知特征点 P_i 在 $k-1$ 时刻世界坐标系的协方差矩阵为 $\Sigma_{P_{i,k-1}}^w = E[(P_{k-1}^w - \mu_{P_{i,k-1}}^w)(P_{k-1}^w - \mu_{P_{i,k-1}}^w)^T]$; 在 k 时刻相机坐标系下的协方差矩阵为 $\Sigma_{P_{i,k}}^c = E[(P_k^c - \mu_{P_{i,k}}^c)(P_k^c - \mu_{P_{i,k}}^c)^T]$; 并且根据假设(3),在同一时刻相机坐标系下不同特征点 $P_{i,k-1}^c$ 和 $P_{i,k}^c$ 相互独立且与 t_{k-1} 无关,即:

$$\Sigma_{t_k} = \left(\sum_{i=1}^n \Sigma_{P_i}^w \right)^{-1} \sum_{i=1}^n \Sigma_{P_i}^w E[\Sigma_{P_{i,k-1}}^w + R_k \Sigma_{P_{i,k}}^c R_k^T] (\Sigma_{P_i}^w)^T + \sum_{i=1}^n \sum_{j=1, i \neq j}^n \Sigma_{P_i}^w \Sigma_{P_j}^w (\Sigma_{P_i}^w)^T \quad (13)$$

3.3 基于 EKF 的特征点位置一致化更新

考虑到观测噪声的存在,采用卡尔曼滤波更新特征点空间位置的一致性,降低视觉里程计的累计误差。首先, $k-1$ 时刻 W 坐标系下的坐标表示 k 时刻特征点位置的预测值,即:

$$\begin{cases} \hat{\mu}_{P_k} = \mu_{P_{k-1}}^w \\ \hat{\Sigma}_{P_k} = \Sigma_{P_{k-1}}^w \end{cases} \quad (14)$$

特征点 p 在 k 时刻的世界坐标系下的观测值记为 P_k^w , 根据不确定性原则,特征点 p 的位置更新方程如下:

$$\begin{cases} K_k = \hat{\mu}_{P_k} (\hat{\mu}_{P_k} + \Sigma_{P_k}^w)^{-1} \\ \mu_{P_k}^w = \hat{\mu}_{P_k} + K_k (\mu_{P_k}^w - \hat{\mu}_{P_k}) \\ \Sigma_{P_k}^w = (I - K_k) \hat{\Sigma}_{P_k} \end{cases} \quad (15)$$

当用以表示特征点空间位置误差的协方差矩阵较小时,可以用式(16)对相机位姿的期望及协方差进行近似,相对于原式(11)、(13)一定程度上降低了算法的运算量:

$$\begin{cases} \mu_{t_k} = \frac{1}{n} \sum_{i=1}^n (\mu_{P_{i,k-1}}^w - R_{i,k} \mu_{P_{i,k}}^c) \\ \Sigma_{t_k} = (1 - \frac{1}{n}) \Sigma_{t_{k-1}} + \frac{1}{n^2} \sum_{i=1}^n (\Sigma_{P_{i,k-1}}^w + R_k \Sigma_{P_{i,k}}^c R_k^T) \end{cases} \quad (16)$$

至此,通过公式(16)可以得到相机当前时刻的位置期望及协方差,通过记录每一时刻相机位置期望,可以得到相机运动轨迹。

4 实验与分析

实验软硬件配置如下:

深度相机: Kinect v2;

计算机硬件配置: AMD A8 处理器, 3.0 GHz 主频, 不使用 GPU 加速, 4 G 内存;

计算机系统为 Ubuntu16.04, 运行平台 ROS kinetic。

为了验证算法的有效性和可对比性,采用 TUM 标准数据集,通过对办公桌前相机简单平行运动 frg1_xyz 数据集在保证精度的基础上测试算法实时性;此外在复杂室内办公室场景 frg1_room 数据集测试其准确性。同时作为比较,使用基于非线性 PnP 模型 BA 算法、基于 SVD-ICP 算法作为对比。

4.1 算法实时性测试

测试数据集 frg1_xyz, 主要用于程序实时性测试。基于此,该数据集将着重测试程序实时性,其中真实轨迹与估计轨迹对比见图 4,非线性 PnP-BA 算法、SVD-ICP 算法与文中算法的实时性对比见图 5,表 1 记录了不同算法下的误差数据。

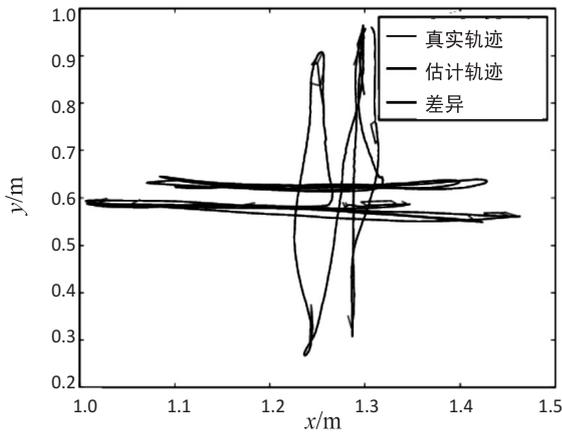


图4 真实轨迹与估计轨迹对比

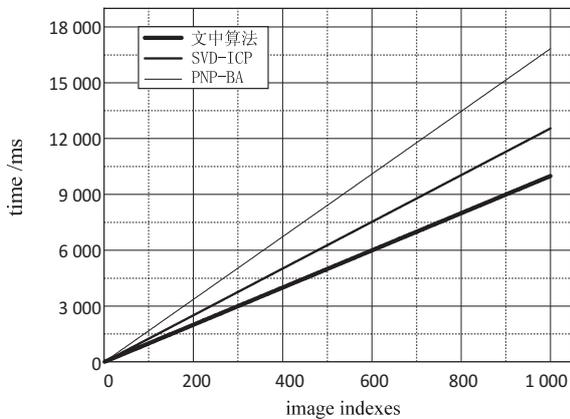


图5 不同算法实时性图像

表1 不同算法误差 mm

误差	PNP-BA	SVD-ICP	文中算法
rmse	10.4	20.6	12.8
mean	13.6	27.7	16.9
std	8.8	12.3	9.1

从图4可以看出真实轨迹与估计轨迹的误差情况,相机在做直线运动的情况下,没有产生误差,但在旋转过程中产生少许误差。图5所展示的是不同算法处理采集图像所使用的时间,PNP-BA模型算法处理精度最高但时间最慢,原因是PNP-BA算法采用了全局非线性优化,通过构造全局误差模型并对其优化,降低了整体误差;文中算法通过图像匹配技术结合对极几何约束提供初值,相较于SVD-ICP算法不但提高了精度而且缩短了时间。可以代替ICP算法在轻量级设备上计算相机运动轨迹。表1中数据表明,文中方法误差不超25mm,虽然PNP-BA算法误差很小,但是实时性较差,对处理器有较高的要求,不适用于小型化设备。

4.2 准确性对比实验

使用测试数据集 frg1_room,对文中算法进行测试。文中算法真实轨迹与估计轨迹对比见图6,SVD-ICP算法真实轨迹与估计轨迹对比见图7,表2记录了

不同算法模型下的误差。

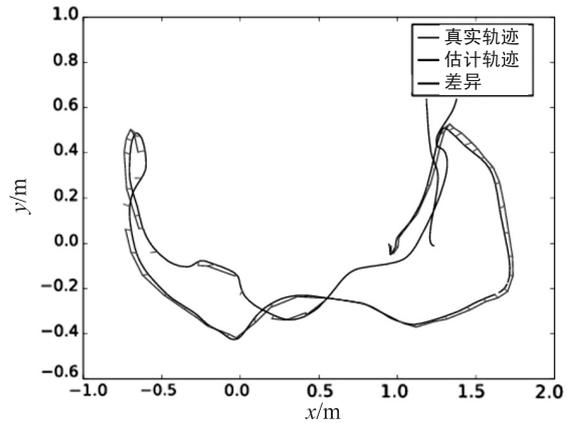


图6 文中算法真实轨迹与估计轨迹对比

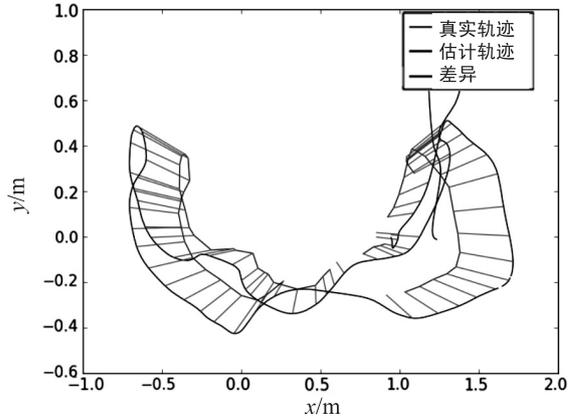


图7 SVD-ICP算法真实轨迹与估计轨迹对比

表2 算法误差对比 mm

误差	PNP-BA	SVD-ICP	文中算法
rmse	25.4	70.2	28.7
mean	24.7	86.3	26.6
std	10.6	33.4	10.8

用TUM的测试数据集 frg1_room,测试室内环境下相机定位准确性。从图6可以看出,在相机进行直线运动时产生的误差较小,但当相机出现旋转运动时会出现一定的误差,因为旋转运动时相机采集的图像信息会发生剧烈的变化,所提取的特征点会有一部分失效,此时深度图像估计运动轨迹比重增大,所以程序主要的误差都是从相机旋转运动中产生。

相较于SVD-ICP算法,图7展示的真实轨迹与估计轨迹存在较大差异,一方面SVD-ICP算法未对其建立高斯分布的特征点误差模型,会存在深度图的噪音干扰;另一方面,文中设计的距离函数,既有效度量了特征点的概率位置,也考虑了其位置方差的影响,因而提高了轨迹精度。表2可以通过数据展示不同算法估计轨迹产生的误差值,其中室内环境下相机定位误差不超过30mm。文中算法可以满足相机定位精度和实时性需求,能够较好地实现嵌入式机器视觉定位的

需求。

5 结束语

视觉里程计作为 SLAM 重要环节之一,对减少机器人运动轨迹误差、提高地图精度至关重要。该文提出一种基于 FREAK-SURF 特征点深度误差视觉里程计算法,旨在轻量级设备满足实时性要求的基础上提高相机运动轨迹精度,为构造地图提供准确的位置信息。首先,融合 SURF 特征提取算法和 FREAK 描述子完成特征匹配,提高匹配效率;其次,ICP 配准环节设计出一种新型的距离代价函数,融合马氏距离和欧氏距离的特点;然后,根据特征点和相机空间位置误差模型,建立视觉里程计累计误差并通过扩展卡尔曼滤波器优化。对比实验表明,该算法在实时性上强于非线性 PNP 模型 BA 优化算法和 SVD-ICP 算法,在精度上高于传统的 SVD-ICP 算法。

参考文献:

- [1] 李元,王石荣,于宁波.基于 RGB-D 信息的移动机器人 SLAM 和路径规划方法研究与实现[J].智能系统学报,2018,13(3):445-451.
- [2] CADENA C,CARLONE L,CARRILLO H,et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age[J]. IEEE Transactions on Robotics,2016,32(6):1309-1332.
- [3] NISTER D,NARODITSKY O,BERGEN J R. Visual odometry [C]//Proceedings of the 2004 IEEE computer society conference on computer vision and pattern recognition. Washington, D. C., USA: IEEE,2004:652-659.
- [4] ENGEL J,SCHOPS T,CREMERS D. LSD-SLAM: large-scale direct monocular SLAM [C]//13th European conference on computer vision. Berlin, Germany: Springer,2014:834-849.
- [5] DAVISON A J,REID I D,MOLTON N D,et al. MonoSLAM: real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2007,29(6):1052-1067.
- [6] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces [C]//6th IEEE and ACM international symposium on mixed and augmented reality. Nara, Japan: IEEE,2007:225-234.
- [7] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics,2015,31(5):1147-1163.
- [8] 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.
- [9] 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.
- [10] WASENMÜLLER O,STRICKER D. Comparison of Kinect V1 and V2 depth images in terms of accuracy and precision [C]//Computer vision - ACCV 2016 workshops. Taipei: Springer,2016:34-45.
- [11] BESL P J, MCKAY H D. A method for registration of 3-D shapes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1992,14(2):239-256.
- [12] SIEMIATKOWSKA B,ZYCHEWICZ A. The application of ICP and SIFT algorithms for mobile robot localization [J]. Robot Design, Dynamics and Control,2009,18(2):391-398.
- [13] JOUNG J H, AN K H, KANG J W, et al. 3D environment reconstruction using modified color ICP algorithm by fusion of a camera and a 3D laser range finder [C]//International conference on intelligent robots and systems. Louis, USA: IEEE, 2009:3082-3088.
- [14] HE L, WANG S, PAPPAS T N. 3D surface registration using Z-SIFT [C]//IEEE international conference on image processing. Evanston, USA: IEEE, 2011:1985-1988.
- [15] NEMRA A, SLIMANI S, BOUHAMIDI M, et al. Adaptive iterative closest SURF for visual scan matching, application to visual odometry [C]//International conference on systems and control. Morgantown, USA: IEEE, 2014:927-932.
- [16] RHEE S M, YONG B L, LEE H E. Two-pass ICP with color constraint for noisy RGB-D point cloud registration [C]//IEEE international conference on consumer electronics. Seoul, Korea: IEEE, 2014:89-90.
- [17] DRYANOVSKI I, VALENTI R G, XIAO J. Fast visual odometry and mapping from RGB-D data [C]//IEEE international conference on robotics and automation. Piscataway, USA: IEEE, 2016:2305-2310.
- [18] ALAHI A, ORTIZ R, VANDERCHEYNST P. FREAK: fast retina keypoint [C]//IEEE conference on computer vision and pattern recognition. Lausanne, Switzerland: IEEE, 2012:614-625.