

面向无人平台的室内精密定位与构图技术

蒋小强, 卢虎[✉], 贺磊南

(空军工程大学信息与导航学院, 西安, 710077)

摘要 针对无人作战平台在室内等复杂环境下的高精度定位与地图构建问题,提出了一种利用信标锚节点辅助的单目视觉同步定位与地图构建方法,该方法通过 WLAN 锚节点估计出单目视觉尺度因子并结合视觉传感参数,获取真实尺度空间下的定位位置,并利用因子图模型对 WLAN 定位和单目视觉定位结果进行最大后验概率准则下数据融合,从而获得精确的位置和构图信息。实验表明:文中所提出的方法较好地解决了传统单目视觉定位中的尺度不确定问题,定位精度达到分米级,扩展了单目视觉 SLAM 在真实尺度空间中的应用,在复杂环境下的无人侦察、搜救和打击等应用领域具有重要的工程和实践意义。

关键词 单目视觉;尺度因子;WLAN 定位;最大后验概率;图优化

DOI 10.3969/j.issn.1009-3516.2019.01.017

中图分类号 U666.1;TN966 **文献标志码** A **文章编号** 1009-3516(2019)01-0103-08

Indoor Precision Positioning and Mapping Technology for Unmanned Platforms

JIANG Xiaoqiang, LU Hu[✉], HE Leinan

(Information and Navigation College, Air Force Engineering University, Xi'an 710077, China)

Abstract: In order to realize the high-precision positioning and map construction for unmanned combat platforms in complex environments, we proposed a precision positioning and mapping method based on monocular visual SLAM with WLAN added. This method can estimate the scale factor of monocular VSLAM and get the position in real scale space. Furthermore, the result of the WLAN location and the monocular vision positioning is fused by the factor graph model, realizing accurate positioning and sparse mapping. The experiment shows that this method can estimate scale factor well and reach the positioning accuracy of decimeter, and has important engineering and practical significance in applications such as unmanned reconnaissance, search and rescue in complex environments.

Key words: monocular visual; scale factor; WLAN positioning; MAP; graph optimization

随着科学技术的进步,移动机器人越来越广泛地应用于军事领域,在监视、侦察、搜救等方面发挥着越来越大的作用。移动机器人的自主导航定位以及对未知环境的构建,成为了其完成其他一切工作

的前提。

在室外条件下,大多数机器人比如无人机的定位主要依靠卫星导航系统和惯性导航系统;对于室内机器人,定位主要利用激光雷达或者相机进行同

收稿日期: 2018-05-13

基金项目: 国家自然科学基金(61473308)

作者简介: 蒋小强(1995—),男,陕西宝鸡人,硕士生,主要从事智能导航与协同控制研究。E-mail: 18700827816@163.com

通信作者: 卢虎(1975—),男,陕西西安人,副教授,主要从事导航理论与技术研究。E-mail: sdkmsdn@qq.com

引用格式: 蒋小强, 卢虎, 贺磊南. 面向无人平台的室内精密定位与构图技术[J]. 空军工程大学学报(自然科学版), 2019, 20(1): 103-110.
JIANG Xiaoqiang, LU Hu, HE Leinan. Indoor Precision Positioning and Mapping Technology for Unmanned Platforms[J]. Journal of Air Force Engineering University (Natural Science Edition), 2019, 20(1): 103-110.

时定位与构图(SLAM)。由于相机具有成本低廉的优点,加上计算机技术的飞速发展,单目视觉SLAM技术在近些年逐渐成熟起来,由刚开始比较主流的基于卡尔曼滤波的算法发展到基于批处理的非线性优化方法,H, Strasdat 等人^[1]证明,在同等计算量的条件下,基于关键帧的非线性优化的方法能取得更优的结果。由 A. J. Davison 在 2007 年提出的 MonoSLAM 是第一个实时的单目视觉 SLAM 系统^[2],为后来的大部分 SLAM 系统提供了较好的研究思路,由于 MonoSLAM 是建立在 EKF 的基础上的,且采用了非常稀疏的特征点追踪,造成了其应用场景很窄,算法鲁棒性也较差。Klein 和 Murray 等人提出的 PTAM (Parallel Tracking and Mapping^[3]),实现了跟踪和建图过程的并行化,它引入了关键帧机制,是第一个使用非线性优化,摒弃了以传统滤波作为优化的方法,解决了应用非线性优化由于计算量过大而不能实时化的问题^[4]。Song 等人^[5]提出了用提取 ORB 特征点方法来处理图像前端问题,极大地提高了特征点的提取效率,进而用于跟踪和捆集调整(Bundle Adjustment, BA)优化。R. Mur-Artal 和 J. D. Tardos 等人提出了单目 ORB-SLAM 算法^[6-7],融合了 PTAM 算法的主要思想和 Strasdat 提出的闭环修正方法,在其整个系统中,均采用 ORB 特征点^[8-9]进行匹配,建图和回环检测,与上述提到的算法相比,计算实时性和精度有了很大的提高。

单目的尺度不确定性是单目视觉 SLAM 的固有问题,因此尺度估计一直是单目 SLAM 研究的热点。单目尺度的估计方法大体上可以分为 2 类,一类是利用先验知识来恢复尺度,例如,针对单目视觉在轮式车辆上的应用,S. Choi^[10]等人在假定相机到地面的距离恒定且地面绝对平整的前提下,提出通过测量地面上的点来估计真实尺度的方法;T. Botterill^[11]等人提出了利用环境中已知物体的尺寸来恢复单目尺度的方法,通过将事先学习好的物体分类及尺寸分布存储在词袋模型中,然后进行词袋匹配以实现尺度的恢复;这种方法依赖词袋规模和匹配的正确程度,由于词袋规模的有限性,当处于完全未知环境中时该方法会失效,因此适用范围较窄。另一类方法是通过增加额外的传感器如轮式编码器、惯性测量单元以及其他辅助定位手段等来获取真实尺度,基于滤波的方法 MSCKF^[12]将尺度信息加入到状态向量中,实现了利用 IMU 获得单目真实尺度的算法;Foster^[13]等人提出了对 IMU 信息进行预积分和基于非线性优化的增量式融合算法,并在因子图框

架上实现了 IMU 和单目视觉 SLAM 的高效融合;然而惯性导航和视觉 SLAM 均属于推算式定位方法,只能获得相对起始位置的位姿,且误差累积不可避免,要应用于实际还需要其他定位手段来辅助获得绝对位置;文献[14]提出利用 GPS 来估计单目尺度的方法,并将单目视觉与惯性导航进行了松组合,在室外有较好的应用价值。

针对室内环境,本文在文献[15]提出的 WiFi 多载波相位测距定位和 R. Mur-Artal 提出的 ORB-SLAM 算法的基础上,研究了利用 WiFi 定位来恢复单目 ORB-SLAM 尺度的算法,并将该问题转化为图模型进行了求解;最后,针对 WiFi 定位和视觉 SLAM 的异步融合问题,研究了基于因子图的 ORB-SLAM/WiFi 融合定位算法,实现了融合位姿的增量式求解。

1 单目视觉 ORB-SLAM 算法

ORB-SLAM 算法分为 3 个并行的线程:跟踪、局部地图构建以及闭环检测,算法总体框架参考文献[7]。

单目 ORB-SLAM 的工作过程如下:

1) 相机标定:针孔相机模型和透镜的畸变模型可以描述 1 个普通相机的整个投影过程,这 2 个模型把外部的三维点投影到相机内部成像平面,构成相机的内参数。因此,在实现单目定位之前必须对相机的内参数进行标定,本文中采用张正友棋盘标定方法来获得摄像机的内参数和畸变参数。

2) 提取图像的 ORB 特征,即改进的 FAST 角点和二进制描述子 BRIEF,ORB 的提取不像 SIFT 和 SURF 特征那样费时,因此可以满足 CPU 的实时计算。在跟踪阶段,实时提取当前帧的 ORB 特征点并与前一时刻的关键帧进行匹配,计算特征点的空间位置和当前相机的位姿。此外,ORB-SLAM 在开始运行之前必须首先导入线下已经训练好的 ORB 词典,用于比较快速的回环检测。

3) 将 Tracking 线程中得到的局部空间的粗略特征点位置和相机姿态放入局部优化线程,使用光束平差法(Bundle Adjustment)进行优化,从而得到更为精细的相机位姿和特征点空间位置。

4) 在闭环检测和后端优化线程中,对前面保存的全局地图进行闭环匹配,校正系统的尺度漂移并消除累积误差;在全局优化中,由于地图点数量庞大,如果将其加入优化线程必然会加重计算量,而我们此时只关心相机位姿的优化,因此该线程中只优化相机的位姿图。

综上,ORB-SLAM 的三线程结构能够取得很好的跟踪和稀疏建图效果,且得到的位姿和空间点位置精度高,鲁棒性强。由于单目无法获得深度信息,所以 ORB-SLAM 得到的相机运动和特征点的空间位置与实际空间相差一个尺度,因此尺度因子的计算是单目视觉用于导航首先要解决的问题。

2 WiFi 辅助的视觉传感参数标定

单目视觉 ORB-SLAM 的参考坐标系为初始化成功后第 1 帧图像的相机坐标系 $O_c-X_cY_cZ_c$,习惯上让 Z_c 轴指向相机的前方, X_c 轴向右, Y_c 轴向下, O_c 为相机的光心,相机位置和姿态信息都是相对于此坐标系得到。如图 2 所示,相机坐标系 $O_c-X_cY_cZ_c$ 和地理坐标系 $O_n-X_nY_nZ_n$ 之间是相似变换关系,可用相似变换群 $\text{Sim}(3)$ 表示:

$$\text{Sim}(3) = \left\{ \mathbf{S} = \begin{bmatrix} s\mathbf{R} & t \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \right. \quad (1)$$

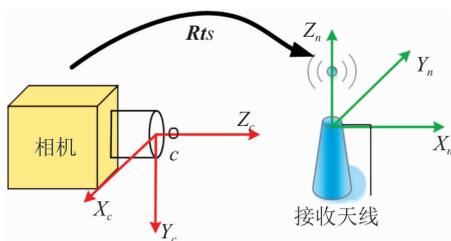


图 2 相机和 WiFi 坐标系

图中, \mathbf{R} 为旋转矩阵, t 为平移向量, s 为尺度因子,由于用相似变换表示位姿有较大的冗余度且在优化时求导不方便,我们引入了其对应的李代数, $\text{Sim}(3)$ 对应的李代数 $\hat{\text{Sim}}(3)$ 元素是一个 7 维向量 ζ ,通过指数映射可以找到李代数和李群的关系。由上可知,相机外参共有 7 个自由度。

$$\text{Sim}(3) = \left\{ \zeta \mid \zeta = \begin{bmatrix} \rho \\ \phi \\ \sigma \end{bmatrix} \in \mathbb{R}^7, \hat{\zeta} = \begin{bmatrix} \sigma I + \phi^\wedge & \rho \\ \mathbf{0}^T & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \right. \quad (2)$$

指数映射关系为:

$$\mathbf{S} = \exp(\hat{\zeta}) = \begin{bmatrix} e^\sigma \exp(\phi^\wedge) & J_s \rho \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (3)$$

外参标定流程可描述如下,规划一条固定的闭合路径,将设备连接到 AP 的无线网络,并执行 ping 指令,使客户端与 AP 产生数据交互;启动单目视觉 ORB-SLAM 程序,在完成初始化之后,沿着设定好的路径移动,并实时采集记录单目视觉和 WiFi 测距定位的结果以及当前时间戳,当 ORB-SLAM 检测到闭环之后,采集结束。由于视觉定位和 WiFi

定位的异步性,因此需要根据采集时间对数据进行一次时间上的对齐,原则上可以把采集时间相近于一个阈值的数据看成一对匹配点,这样得到 N 个定位点对。单目视觉的定位数据为 $\mathbf{p}'_i = [x'_i, y'_i, z'_i]^T$ ($i=1,2,\dots,N$),WiFi 定位数据为 $\mathbf{p}_i = [x''_i, y''_i, z''_i]^T$ ($i=1,2,\dots,N$)。定义第 i 对点的误差项为:

$$\mathbf{e}_i = \mathbf{p}_i - (s\mathbf{R}\mathbf{p}'_i + t) \quad (4)$$

然后构建最小二乘问题,求使误差平方和达到极小的 s, \mathbf{R}, t :

$$\min_{s, \mathbf{R}, t} J = \frac{1}{2} \sum_{i=1}^N \| (\mathbf{p}_i - (s\mathbf{R}\mathbf{p}'_i + t)) \|^2 \quad (5)$$

定义 2 组点的质心:

$$\mathbf{p} = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i, \mathbf{p}' = \frac{1}{N} \sum_{i=1}^N \mathbf{p}'_i \quad (6)$$

每个点的去质心坐标为:

$$\mathbf{q}_i = \mathbf{p}_i - \mathbf{p}, \mathbf{q}'_i = \mathbf{p}'_i - \mathbf{p}' \quad (7)$$

文献[12]中给出了利用三对点求解 s, \mathbf{R}, t 的解析求解方法,得到尺度因子 s 为:

$$s = \left(\frac{\sum_{i=1}^N \|\mathbf{q}_i\|^2 / \sum_{i=1}^N \|\mathbf{q}'_i\|^2}{\sum_{i=1}^N \|\mathbf{q}'_i\|^2} \right)^{1/2} \quad (8)$$

3 个定位点需要从采集的 N 个定位点对中选取,但由于存在误差过大或者误匹配的情况,将会使计算结果严重偏离真实值,使用随机采样一致性算法(Random Sample Consensus, RANSAC)[17]可以用来剔除外点而保留正确的匹配点,保证三点法和后面图优化估算结果的正确性。

在得到相机坐标系到地理坐标系的相似变换矩阵的初始解之后,再使用通用图优化工具 g2o(General Graph Optimization)[18]解决非线性最小二乘的优化问题,图优化模型见图 3。在这个图优化中,节点和边的选择如下:

1) 节点:相机到地理坐标系的相似变换矩阵 \mathbf{S} 对应的李代数 $\zeta \in \mathbb{R}^7$ 。

2) 边:3D 点从相机坐标系变换到地理坐标系,用如下方程描述:

$$\mathbf{p}_i = (s\mathbf{R}\mathbf{p}'_i + t) \quad (9)$$

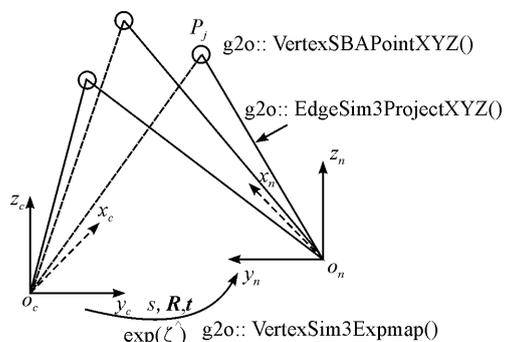


图 3 $\text{Sim}(3)$ 求解的图优化表示

本文中图优化选择的迭代方法为列文伯格-马夸尔特方法,在线性化过程中需要求解每个误差项关于优化变量的导数即雅克比矩阵,利用李代数的扰动模型^[19]对(4)式求导可得:

$$\frac{\partial e_i}{\partial \zeta} = -[I - (sR p'_i + t) \wedge (sR p'_i + t)] \quad (10)$$

上述矩阵为一个 3×7 的雅克比矩阵,其中 I 为单位矩阵,符号 \wedge 表示将一个向量变为反对称矩阵。

3 基于因子图的视觉/WiFi 定位融合算法

求得尺度因子 s 和两坐标系变换矩阵的最优值之后,就可以进行 ORB-SLAM 和 WiFi 的融合定位。由于 ORB-SLAM 的更新速率相对较快,而 WiFi 定位的速率较慢,一般情况下 ORB-SLAM 始终有输出,而 WiFi 只在某一时间段内有输出。针对传感器速率的异步性和实现融合所有数据进行导航状解算,本文采取了基于因子图(Factor Graph)的融合算法。因子图作为一种较新颖的方法,有着增长平滑的特性和全局优化的特点^[20],被逐渐用于传感器融合领域。

因子图将一个导航方位估计问题抽象为一个双射图 $G=(F, \Theta, \epsilon)$,其中含有 2 种节点类型:因子节点 $f_i \in F$ 和状态节点 $x_i \in \Theta$ 。当因子 f_i 与状态 x_i 有联系时,一个边 $e_{i,j} \in \epsilon$ 便存在于该状态与因子之间。将导航问题中的测量模型抽象为:

$$z_i = h_i(x_i) + v_i \quad (11)$$

式中: z_i 为传感器对于状态 x_i 利用测量函数 $h_i(x_i)$ 的测量; v_i 为测量噪声。假设 v_i 遵循高斯分布,那么可以得到一个因子化的最优化目标:

$$f_i(x_i) = d(h_i(x_i) - z_i) \quad (12)$$

函数 $d(\cdot)$ 是代价函数,对于误差服从高斯分布的估计问题,通常选择马氏距离作为其代价函数形式: $d(e) = e^T \sum^{-1} e$, \sum 为测量噪声的协方差矩阵。整个图模型则被定义为每个因子的乘积:

$$F(x) = \prod_i f_i(x_i) \quad (13)$$

其中每个因子代表了待优化求解的误差函数,通过最小化 $F(x)$ 便可以得到状态位姿的最优估计 \hat{x} :

$$\hat{x} = \arg \min_x (\prod_i f_i(x_i)) \quad (14)$$

因子图模型中可以根据新增传感器的性质随时扩容,简单传感器可以直接动态插入已知因子图,实现即插即用的融合滤波的目的。下面将分别对 ORB-SLAM 和 WiFi 定位模型进行因子图描述。

3.1 ORB-SLAM 节点

在单目视觉的针孔模型下,量测方程通过投影

方程得到:

$$p = K[R \ t]X \quad (15)$$

式中: p 为成像平面的像素坐标, X 为观测到的 3D 点的坐标, R 和 t 对应世界坐标系到当前相机的旋转矩阵和平移向量。为了避免将空间点作为优化变量,而只考虑对相机位姿的优化,需要用多视角约束来替代投影方程。相邻 2 个位姿节点的关系用李群表示如下:

$$T_{\Delta} = T_k^{-1} T_{k+1} \quad (16)$$

式中: T_{Δ} 表示前一帧到后一帧的变换矩阵, T_k 和 T_{k+1} 为当前变量节点 x_k 和 x_{k+1} 的位姿。利用式(16)可以构建相邻两帧的误差函数即因子节点:

$$f_{\text{SLAM}}(x_k, x_{k+1}) = d(T_{\Delta}^{-1} T_k^{-1} T_{k+1}) \quad (17)$$

3.2 WLAN 节点

文献[15]提出了采用多载波相位测距的最优二次定位,首先滤波处理 WLAN 传感器接收到的 WLAN 锚节点(Anchor Point, AP)的 CSI 得到原始距离,然后对原始距离进行预处理得到有效距离测量,记为 $z_{k, \text{WLAN}}$,最后将建立最优二次规划,按照最小误差平方准则定义目标函数为客户端到 AP 的测距误差平方和,定位方程为:

$$\begin{cases} \min_{x^{\text{SP}}, y^{\text{SP}}} g(l_x^{\text{SP}}, l_y^{\text{SP}}) = \sum_{j=1}^n |(l_x^{\text{SP}} - l_{x,j}^{\text{AP}}) + (l_y^{\text{SP}} - l_{y,j}^{\text{AP}})^2 - d_j^2| \\ \text{s. t. } (l_x^{\text{SP}}, l_y^{\text{SP}}) \in \text{AR} \end{cases} \quad (18)$$

式中: $(l_x^{\text{SP}}, l_y^{\text{SP}})$ 代表客户端的位置; $(l_{x,j}^{\text{AP}}, l_{y,j}^{\text{AP}})$ 代表第 j 个 WLAN 锚节点的位置; g 是目标函数; d_j 是客户端到第 j 个 AP 的测距距离; AR 是定位的可行域。用 $h_{\text{WLAN}}(z_{k, \text{WLAN}})$ 表示 WiFi 定位结果, x_k 为 t_k 时刻的位置估计, WLAN 节点定义如下:

$$f_{\text{WLAN}} = d(x_k - h_{\text{WLAN}}(z_{k, \text{WLAN}})) \quad (19)$$

WLAN 因子节点只与变量节点相连,当 WLAN 传感器的探测到的 CSI 更新时输出当前时刻的位置估计,将其加入图中,对导航状态的估计结果进行优化。

图 4 为本文提出的 ORB-SLAM/WiFi 融合模型,ORB-SLAM 使用的视觉传感器是基本单元,状态输出贯穿整个导航定位阶段, WiFi 定位在 $x_4 \sim x_7$ 状态有效,当只有 ORB-SLAM 输出时,因子图只对 SLAM 输出的位姿做优化,当 WLAN 传感器可用时,相应地在图中加入其测量值,对 ORB-SLAM 定位结果进行修正。

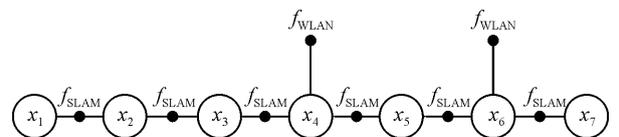


图 4 因子图融合模型

4 因子图增量式平滑算法

第 3 节介绍了因子图模型中各节点的具体描述,从优化的角度来看,优化一个因子图和普通的图优化并没有太大区别——最后都要转化为一个最小二乘问题来求解,因此在介绍因子图的增量式平滑之前,有必要先讨论下一般的批处理的优化方法。

4.1 批量优化 (Batch Optimization)

对于非线性优化问题,一般用迭代的方法来求解,从一个初始值出发,不断地更新当前的优化变量,使目标函数下降。具体步骤如下:

步骤 1 给定某个初始值 \mathbf{x}_0 。

步骤 2 对于第 k 次迭代,寻找一个增量 $\Delta \mathbf{x}_k$,使得 $F(\mathbf{x})$ 达到极小值。

步骤 3 若 $\Delta \mathbf{x}_k$ 足够小,则停止。

步骤 4 否则,令 $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$,返回步骤 2。

由上可知,求解增量 $\Delta \mathbf{x}_k$ 是迭代进行非线性优化的关键一步,对于常用的高斯牛顿法,通常求解如下方程:

$$\mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x}) \Delta \mathbf{x} = -\mathbf{J}(\mathbf{x})^T f(\mathbf{x}) \quad (20)$$

式(20)称为增量方程,其中 $\mathbf{J}(\mathbf{x})$ 为在线性化点 \mathbf{x} 处的雅克比矩阵,定义信息矩阵 $\mathbf{H} = \mathbf{J}(\mathbf{x})^T \mathbf{J}(\mathbf{x})$,其具有稀疏的特性, $f(\mathbf{x})$ 为代价函数,雅克比矩阵里的矩阵块结构能够反映出因子图构成,因此它等价于线性化后的因子图。

对于前面提到的因子节点,WiFi 定位节点对应的雅克比矩阵为 $\frac{\partial e}{\partial \mathbf{x}_k}$,ORB-SLAM 节点对应的雅克比矩阵为:

$$\frac{\partial e_{k,k+1}}{\partial \delta \mathbf{x}_k}, \frac{\partial e_{k,k+1}}{\partial \delta \mathbf{x}_{k+1}} \quad (21)$$

式中: $\mathbf{e} = \mathbf{x}_k - h_{\text{wLAN}}(z_k, \text{wLAN})$; $e_{k,k+1} = \mathbf{T}_{\Delta}^{-1} \mathbf{T}_k^{-1} \mathbf{T}_{k+1}$ 。

在批量优化方法中,通过对雅克比矩阵进行 QR 或者 Cholesky 分解来求解增量方程,这一过程相当于对因子图当中的变量节点进行消元,而每消去一个节点,就会在相邻的节点之间产生一条额外的边,从而导致信息矩阵的稀疏性变差,求解计算量增大。因此,对节点的不同消元顺序将会影响因子图后端数据结构——贝叶斯网的结构,进而影响解算效率,文献[22]提出的 SAM 方法使用 COLAMD (Column Approximate Minimum Degree Ordering)^[23] 选择最优的消元顺序,来保持信息矩阵稀疏性。

4.2 增量式优化

在因子图优化问题中,每当新的测量量产生时,

相应地会在图中产生一个因子节点,反映在雅克比矩阵当中就是新添加了一行或者一行矩阵块(考虑到状态为多维的情况)。

在 4.1 节的批量优化方法中,每当新的节点加进来时,都要对增量方程(20)进行完全求解,计算出所有节点的更新量。显然,这样做带来的计算量是巨大的,但是增量更新方式可以有效解决这一问题,下面将以图 5 为例来解释增量更新。当新的节点添加进来时,受影响的节点可以近似地看成只有最后一个与之相连的节点(图 5(a)中红色部分所示),而早先节点的估计值,可以近似看成没有变化,因此就没有必要对它们进行优化。这样每次计算更新量时只需考虑受影响的节点,而不必在每次新增节点时,对整个图进行优化。如果按照回环检测的方式添加节点(见图 5(b)),那么受影响的范围应该是回环开始到当前帧这一段中的所有节点,也就是回环之内的所有节点应该更新,但依然无须优化整张图。综上,在向图中添加节点时,通过分析受影响的区域,可以减少一些没必要的计算,进而加速求解过程。

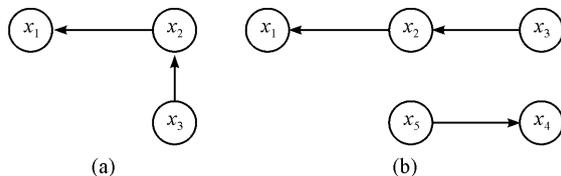


图 5 增量更新示意图

在此思想的基础上, Kaess 等人提出了 iSAM2^[25] 算法,当新的变量加入时,通过将贝叶斯网中受影响较大的部分转换到因子图表述,然后对这个局部的因子图进行求解更新,最后再把更新结果和原来未改变的贝叶斯网进行合并,进而得到和批量优化一样的结果。

5 实验验证分析

5.1 实验条件设置

以视觉 SLAM 公开数据集 TUM 的 freiburg3 为测试样本,其内容是室内环境下利用 30 Hz 深度相机采集的多帧彩色 RGB 图像及对应的深度图(注:实际算法在提取特征点之前需要将 RGB 三通道图转换为单通道的灰度图),并且该数据集中含有真实的运动轨迹,是通过 16 个高精度相机以 100 Hz 频率从不同角度拍摄计算得到。本文中单目视觉程序只用到了该数据集中的 RGB 图。

为了评估本文算法,实验根据数据集的具体场景设置了 WiFi 定位仿真条件,如图 6 所示,WiFi 定位环

境由4个位置已知的AP组成,图中虚线为相机运动的真实轨迹,运动方向为图中箭头所指方向,相机在运动时依次与4个AP进行测距通信来实现一次定位解算;实验中仿真的WiFi测量值是在已知参数轨迹条件下,通过文献[15]测距定位模型分析推导得到,其中仿真选取的定位误差 $\sigma=0.1$ m,更新频率0.75 Hz。

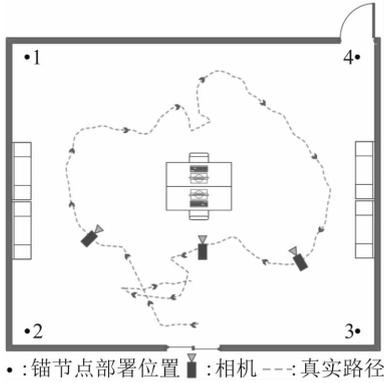


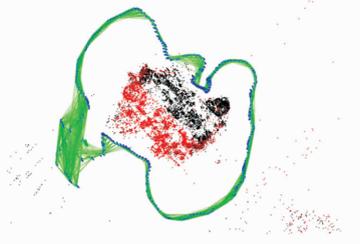
图6 实验场景设置

5.2 实验结果分析

ORB-SLAM的位姿测量和建图结果如图7所示,其中:图7(a)为ORB特征点提取过程,绿色方框为当前提取到的特征点;图7(b)代表ORB-SLAM输出的地图和相机位姿,红色点和黑色点为地图点,红色点表示当前地图点正在参与相机位姿的解算和优化过程,蓝色表示相机关键帧,绿色线条为相机轨迹。



(a) ORB-SLAM图像特征点



(b) ORB-SLAM地图构建和定位过程

图7 ORB-SLAM实验结果

5.2.1 轨迹对齐

由于定义参考系的不同,在比较轨迹误差前需要先进行轨迹对齐,计算出真实轨迹与估计轨迹之间的相对旋转和平移。

设真实轨迹为 T_g ,估计轨迹为 T_e ,二者皆以变

换矩阵 T_{WC} 格式存储。但是真实轨迹的坐标原点定义于外部某参考系中(取决于真实轨迹的采集方式),而估计轨迹则以相机出发点为参考系(在视觉SLAM中很常见),本文算法中估计轨迹以无线AP定义的坐标系为参考系。由于以上原因,理论上的真实轨迹点与估计轨迹点应满足:

$$T_{g,i} = T_{ge} T_{e,i} \quad (22)$$

式中: i 表示轨迹中的第 i 条记录, $T_{ge} \in SE(3)$ 为2个坐标系之间的变换矩阵,该矩阵在整条轨迹中保持不变, T_{ge} 可以通过2条轨迹数据估计得到。

本文采取的轨迹对齐方法为Umeyama^[20]提出的算法,在整条轨迹上通过最小二乘方法来计算 T_{ge} :

$$T_{ge} = \arg \min_{T_{ge}} \sum_{i=1}^n \| \log(T_{g,i}^{-1} T_{ge} T_{e,i}) \|_2 \quad (23)$$

文献[24]中Umeyama提出的算法具有较好的鲁棒性,即使数据中出现异常值,也能准确计算出2条轨迹的变换矩阵。图8为本文算法以及ORB-SLAM算法的定位轨迹在进行轨迹对齐之后与真实轨迹的比较。

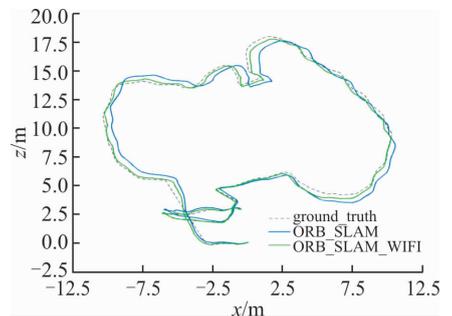


图8 ORB-SLAM和ORB-SLAM/WiFi定位轨迹对比

5.2.2 误差分析

经过轨迹对齐,估计值 $T_{e,i}$ 和真实值 $T_{g,i}$ 被转换到同一个参考坐标系中,为了评估全局一致性误差,本文中误差的计算采取了文献[21]中所提的绝对位姿误差APE(Absolute Pose Error),时刻 i 处的APE为:

$$F_i = T_{g,i}^{-1} T_{e,i} \quad (24)$$

同时可以定义2条轨迹的均方根误差为(Root-Mean-Square-Error, RMSE):

$$RMSE(F_{1:n}) = \left(\frac{1}{n} \sum_{i=1}^n \| \text{trans}(F_i) \|^2 \right)^{1/2} \quad (25)$$

式中: $\text{trans}(F_i)$ 表示绝对位姿误差矩阵 F_i 的平移部分。

图9为使用Umeyama配准方法进行轨迹对齐之后计算的每个时刻的绝对位姿误差,绿色为本文算法的APE曲线,蓝色为ORB-SLAM的APE曲线,从图中可以看出,在整个测试序列上,本文提出的融合算法的精度优于ORB-SLAM。

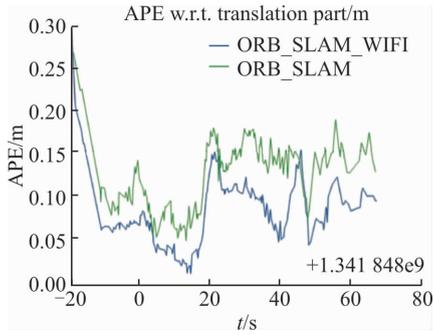


图 9 本文算法与 ORB-SLAM 的绝对位姿误差比较

由于计算的 APE 中包含了姿态误差和位置误差,而将算法应用于室内定位时,需要衡量的指标是算法的定位精度,因此需要将 APE 中的平移部分单独拿出来进行对比。如图 10 所示,图中分别对比了 ORB-SLAM 和本文算法的平移部分误差,可以看出,2 个算法在相机运行轨迹的起始部分均存在较大误差,这是单目初始化带来的不可避免的误差。而随着相机的运动,ORB-SLAM 算法由于误差的累积,误差会越来越大(轨迹由蓝趋向于绿色)。而本文所提的基于因子图的 WiFi 和 ORB-SLAM 融合算法,其融合过程相当于对累积误差不断矫正的过程,因此在整个轨迹中本文算法的平移误差均较小(除了起始阶段,轨迹误差均分布在蓝色附近)。

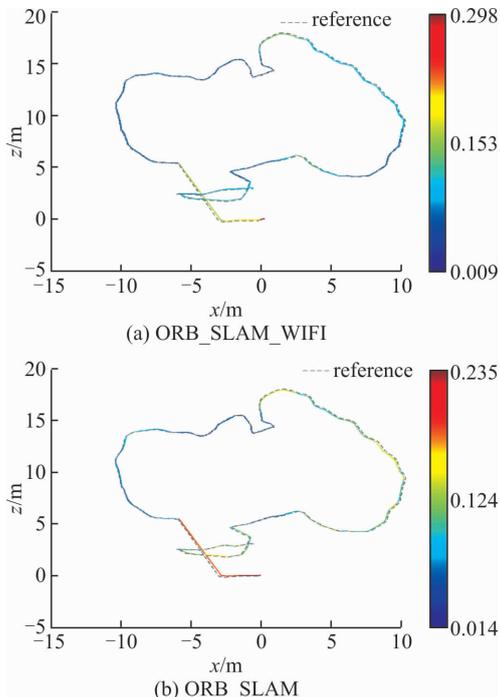


图 10 位姿配准之后的位置误差比较

为了更加直观地显示出误差的分布情况,本文还拟合了误差的分布曲线并画出了箱型图,见图 11 和图 12。从图 11 可以看出,ORB-SLAM 算法的误差大部分居于 0.12 m 左右,而本文所提算法的绝对位姿误差绝大多数小于 0.1 m,统计概率分布的

峰值在 0.05 m 左右,均要优于 ORB-SLAM。

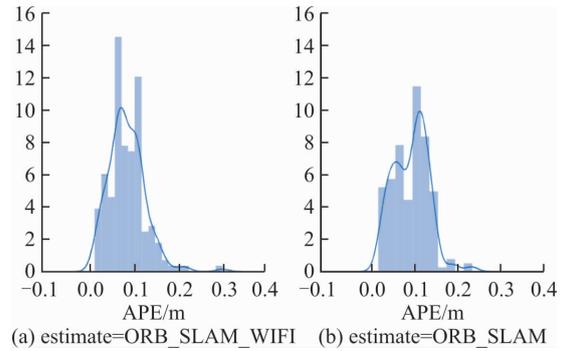


图 11 误差统计分布比较

根据图 12 可以看出,由于在融合之前没有对 WiFi 定位结果进行异常值的检测和剔除,融合算法误差的异常值要多于 ORB-SLAM,最大误差甚至达到 0.3 m,改进措施可以作为后续研究内容来进行研究,可从外点的检测及剔除策略、优化算法鲁棒核函数设置等方面考虑。

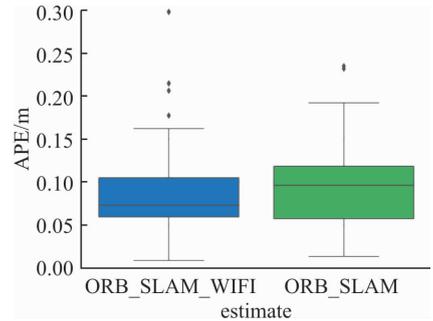


图 12 APE 箱型图

6 结语

本文将单目视觉 ORB-SLAM 与 WiFi 定位相结合,提出了基于因子图的视觉/WiFi 定位融合算法,并利用公开数据集设计了仿真实验进行了验证。结果表明,所设计的 ORB-SLAM/WiFi 融合定位算法具有较高的定位精度,具有一定的实用价值。目前提出的组合导航只包含了视觉和无线传感器,后面工作将考虑加入低成本 IMU 信息,以增强单目 SLAM 算法在大动态范围下的精度和鲁棒性。

参考文献 (References):

[1] STRASDAT H, MONTIEL J M M, DAVISON A J. Visual SLAM: Why Filter[J]. Image and Vision Computing, 2012, 30(2):65-77.

[2] DAVISON A J, REID I D, MOLTON N D, et al. MonoSLAM: Real-Time Singlecamera SLAM [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29(6):1052.

[3] KLEIN G, MURRAY D. Parallel Tracking and Map-

- ping for Small AR Workspaces[C]// IEEE and ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society, 2007:1-10.
- [4] 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.
- LIU H M, ZHANG G F, BAO H J. A Survey of Monocular Simultaneous Localization and Mapping[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(6): 855-868. (in Chinese)
- [5] SONG S, CHANDRAKER M, GUEST C C. Parallel, Real-Time Monocular Visual Odometry[C]//2013 IEEE International Conference on Robotics and Automation (ICRA). Karlsruhe, Germany:IEEE, 2013: 4698-4705.
- [6] MUR-ARTAL R, TARDOS J D. Fast Relocalisation and Loop Closing in Keyframe-Based SLAM[C]//2014 IEEE International Conference on Robotics and Automation. Hong Kong, China:IEEE, 2014: 846-853.
- [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] ROSTEN E, DRUMMOND T. Machine Learning for High-Speed Corner Detection [C]//2006 European Conference on Computer Vision. Graz, Austria: Springer, 2006: 430-443.
- [9] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: Binary Robust Independent Elementary Features [C]//2010 European Conference on Computer Vision. Heraklion, Crete: [s. n.], 2010: 778-792.
- [10] CHOI S, PARK J, YU W. Resolving Scale Ambiguity for Monocular Visual Odometry [C]// International Conference on Ubiquitous Robots and Ambient Intelligence. Jeju, South Korea:IEEE, 2013:604-608.
- [11] BOTTERILL T, MILLS S, GREEN R. Correcting Scale Drift by Object Recognition in Single-Camera SLAM[J]. IEEE Trans Cybern, 2013, 43(6):1767-1780.
- [12] LI M Y, MOURIKIS A I. High-Precision, Consistent EKF-Based Visual-Inertial Odometry[J]. International Journal of Robotics Research, 2013, 32(6):690-711.
- [13] FORSTER C, CARLONE L, DELLAERT F, et al. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry[J]. IEEE Transactions on Robotics, 2017, 33(1):1-21.
- [14] 周绍磊, 吴修振, 刘刚, 等. 一种单目视觉 ORB-SLAM/INS 组合导航方法[J]. 中国惯性技术学报, 2016, 24(5): 633-637.
- ZHOU S L, WU X Z, LIU G, et al. Integrated Navigation Method of Monocular ORB-SLAM/INS[J]. Journal of Chinese Inertial Technology, 2016, 24(5): 633-637. (in Chinese)
- [15] 贺磊南, 卢虎, 姬正洲. 采用多载波相位测距的 Wi-Fi 精密定位方法[J]. 西安交通大学学报, 2017, 51(8): 65-71.
- HE L N, LU H, JI Z Z. A Precise Positioning Method of Wi-Fi Using Multi-Carrier Phase Ranging[J]. Journal of Xi'an Jiaotong University, 2017, 51(8):65-71. (in Chinese)
- [16] HORN B K P. Closed-Form Solution of Absolute Orientation Using Unit Quaternions[J]. Journal of the Optical Society of America A, 1987, 4(4):629-642.
- [17] SCHNALBEL R, WAHL R, KLEIN R. Efficient RANSAC for Point-Cloud Shape Detection[J]. Computer Graphics forum. 2007, 26(2): 214-226.
- [18] KÜMMERLE R, GRISETTI G, STRASDAT H, et al. G2o: A General Framework for Graph Optimization[C]// IEEE International Conference on Robotics and Automation. IEEE, 2011:3607-3613.
- [19] PARK F C, BOBROW J E, PLOEN S R. A Lie Group Formulation of Robot Dynamics [J]. The International Journal of Robotics Research, 1995, 14(6): 609-618.
- [20] DELLAERT F. Factor Graphs and GTSAM: A Hands-on Introduction[R]. Georgia Institute of Technology, 2012.
- [21] STURM J, ENGELHARD N, ENDRES F, et al. A Benchmark for the Evaluation of RGB-D SLAM Systems[C]// IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal: IEEE, 2012:573-580.
- [22] DELLAERT F, KAESS M. Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing [M]. [S. l.]: Sage Publications Inc, 2005.
- [23] DAVIS T A, GILBERT J R, LARIMORE S I, et al. A Column Approximate Minimum Degree Ordering Algorithm [C]// ACM Trans Mathematical Software. 2004:353-376.
- [24] UMEYAMA S. Least-Squares Estimation of Transformation Parameters between Two Point Patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(4):376-380.
- [25] KAESS M, JOHANSSON H, ROBERTS R, et al. iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree[J]. International Journal of Robotics Research, 2012, 31(2):216-235.

(编辑: 徐楠楠)