

电缆隧道环境下工业智能巡检机器人定位研究

王玉涛^{1,2}, 安俊炜¹, 秦长生¹, 郭伟帆¹

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819;

2. 东北大学 流程工业数字化仪表教育部工程研究中心, 辽宁 沈阳 110819)

摘要: 电缆隧道封闭狭长, 存在重复布设的电缆架和相似的场景纹理, 属于退化场景. 针对该场景, 提出了1种基于点线特征融合的视觉惯导SLAM(simultaneous localization and mapping)算法. 该算法通过长度抑制和短线拟合来改进高维线特征, 使其能够更有效地描述结构化显著的隧道场景. 此外, 针对电缆隧道中特征相似导致的回环检测失败问题, 引入具有高效识别和精确位姿估计的ArUco标记, 限定回环发生区域, 并利用最小化位姿变换筛选最佳回环帧, 从而提升回环检测的准确度和定位精度. 最后, 在电缆隧道内采集数据集并进行实验验证. 结果表明, 相对于VINS-Mono(visual inertial system-Mono), 所提算法的绝对轨迹精度平均提升了69.73%, 满足了电缆隧道巡检的应用需求.

关键词: 同步定位与地图构建; 点线特征融合; 回环检测; 工业机器人; 电缆隧道

中图分类号: TP 242 文献标志码: A 文章编号: 1005-3026(2025)07-0049-10

Research on Localization of Industrial Intelligent Inspection Robots in Cable Tunnel Environment

WANG Yu-tao^{1,2}, AN Jun-wei¹, QIN Chang-sheng¹, GUO Wei-fan¹

(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China;

2. Engineering Research Center of Digital Instrument for Process Industries, Ministry of Education, Northeastern University, Shenyang 110819, China. Corresponding author: WANG Yu-tao, E-mail: wangyutao@ise.neu.edu.cn)

Abstract: The cable tunnel is closed and narrow, with repetitively laid cable racks and similar scene textures, which is a degraded scenario. To address this environment, a visual-inertial SLAM (simultaneous localization and mapping) algorithm based on point-line feature fusion is proposed. The algorithm improves the high-dimensional line features through length suppression and short line fitting to make it more effective in describing the structural features of tunnel scene. In addition, for the problem of loop closure detection failure due to feature similarity in cable tunnels, ArUco markers with efficient recognition and accurate pose estimation are introduced to limit the loop closure area, and the optimal loop closure frames are selected using the minimized pose transformation to improve detection accuracy and localization precision. Finally, dataset collection and experimental validation were conducted in actual cable tunnels. The results show that the absolute trajectory accuracy of the algorithm is improved by 69.73% on average relative to VINS-Mono (visual inertial system-Mono), which meets the application requirements of cable tunnel inspection.

Key words: simultaneous localization and mapping; point-line feature fusion; loop closure detection; industrial robot; cable tunnel

随着城市化进程的加速和电力需求的不断增长, 电缆隧道已成为城市电力输配系统的重要组成部分. 由于隧道环境容易积水、空气不流通

以及高压等复杂因素, 传统的人工巡检方式面临着效率低下和安全隐患的问题. 因此, 研发用于电缆隧道的工业智能巡检机器人显得尤为重要.

收稿日期: 2024-11-10

基金项目: 国家自然科学基金资助项目(62473092, 62303097); 辽宁省自然科学基金资助项目(2023JH2/101700356).

作者简介: 王玉涛(1970—), 女, 辽宁沈阳人, 东北大学教授.

同时,隧道的封闭环境阻碍了卫星信号的传输,导致在该场景下无法使用 GPS 进行定位.因此,自主定位技术成为实现电缆隧道智能化巡检的关键.

自主定位是指机器人在没有外界干预的情况下,通过传感器和算法配合,准确确定其在环境中的位置和姿态.目前,常用的定位方法包括视觉、激光雷达以及多传感器融合技术.视觉定位通过相机获取图像特征信息,结合图像处理算法,实现环境感知与定位.以具有代表性的 ORB-SLAM(oriented fast and rotated brief-simultaneous localization and mapping)^[1]为例,该算法涵盖特征追踪、局部地图构建、闭环检测及后端优化等模块,能够提供足够的定位精度.激光雷达通过扫描空间生成三维点云图,提供精确的距离和位置信息. LeGO-LOAM(lightweight and ground-optimized-lidar odometry and mapping)^[2]作为轻量级的定位算法,通过对点云信息进行滤波和配准实现位姿估计.考虑到电缆隧道封闭狭长、特征匮乏以及存在退化场景的特点,单一传感器无法满足工业智能巡检机器人的定位需求.因此,需要结合多传感器以及多特征融合的方法实现定位,以提高系统定位精度和鲁棒性.

在多传感器融合定位方法的研究中, MSCKF(multi-state constraint Kalman filter)^[3]以多状态约束扩展卡尔曼滤波为核心,实现了视觉惯导特征融合优化; VINS-Mono^[4]构建了单目视觉与 IMU(inertial measurement unit)的紧耦合系统,形成了完整的功能架构; Dang 等^[5]开发了基于激光雷达、视觉、IMU 及热学数据松耦合的多模态 SLAM 系统,该系统显著提升了地下环境的定位稳定性; Azpurua 等^[6]提出了适用于地下隧道勘测的双 SLAM 并行系统,分别采用视觉和激光方案; Campos 等^[7]设计了改进的 ORB-SLAM3 系统,该系统引入了基于最大后验推断的 IMU 初始化方法;为拓展系统适用性, GVINS(GNSS-visual-inertial system)^[8]整合全球导航系统,实现了室内外环境的平滑过渡; LE-VINS(lidar-enhanced-VINS)^[9]在因子图优化中融合视觉特征、激光雷达深度和惯性测量单位数据,特别适用于低速机器人应用; DGM-VINS(dynamic environments with joint geometry feature extraction and multiple object tracking-VINS)^[10]通过联合几何动态特征提取模块,有效提升了系统在动态场景下的定位精度; SM-VINS(stepwise marginalization-VINS)^[11]采用逐步边际化的

VINS 以及双向光流跟踪方法,在保证精度的同时降低了计算负担.在多特征融合方面,学者重点关注高维几何特征的应用. StructSLAM 框架^[12]利用线特征构建约束完成位姿估算;在 SVO(semi-direct monocular visual odometry)^[13]的基础上, Gomez-Ojeda 等^[14]提出了 PL-SVO(points and line segments-SVO),在最大似然估计中融合了点线特征; Fu 等^[15]在 VINS-Mono 框架中引入参数控制的线特征提取,开发了高效的 PL-VINS(point and line features-VINS)系统; Lim 等^[16]使用点线结合的方式来应对退化情况;此外, PLD-VINS(point-line features and depth information-VINS)^[17]和 PLF-VINS(point-line fusion and parallel-line fusion-VINS)^[18]同样将点线特征与 IMU 进行融合,进一步提高系统定位精度.

尽管多传感器融合与多特征融合定位技术已取得显著进展,但现有研究主要面向通用场景,鲜有针对电缆隧道场景的自主定位算法设计.在实际应用中,仍存在诸多问题,包括:在复杂光照条件、特殊场景结构下的特征提取、追踪问题;在相似纹理、结构信息下各模块的鲁棒性问题.为满足工业智能巡检机器人在电缆隧道环境下的自主定位需求,本文提出了基于点线特征融合的视觉惯导定位算法.该算法以图像信息和 IMU 信息作为输入,首先,在图像点特征的基础上引入表达结构信息更丰富的线特征,并结合电缆隧道特点对线特征提取进行改进,以保证前端视觉里程计特征匹配的正确率.随后,在回环检测中提出了 1 种融合 ArUco 标记的回环检测策略,以应对电缆隧道相似场景下的错误回环问题.最后,在实际电缆隧道场景下完成算法验证,实验结果表明,该算法能够满足工业智能巡检机器人在复杂隧道环境下的高精度自主定位需求.

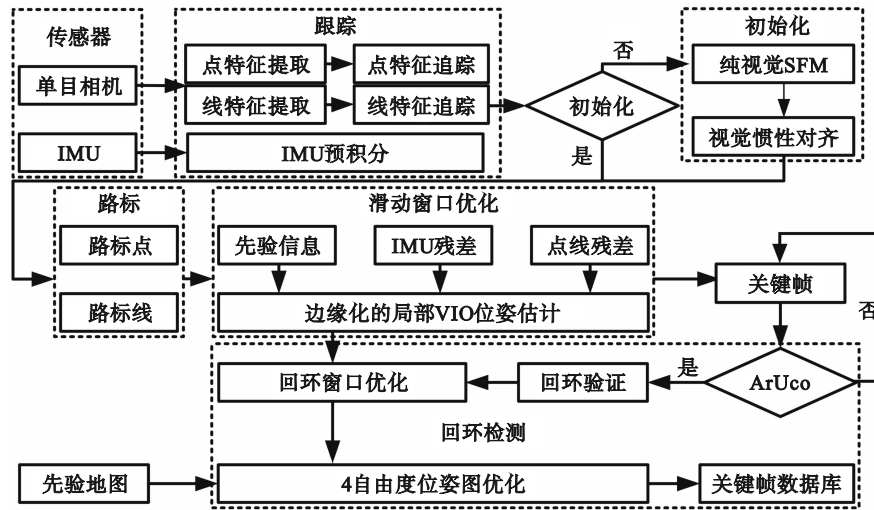
1 电缆隧道工业智能巡检机器人定位算法框架

本文提出的定位算法整体框架如图 1 所示,主要包括三大并行线程:跟踪、后端优化和回环检测.

在跟踪线程中,系统对相机和 IMU 采集到的信息进行预处理操作,包括点线特征提取与跟踪以及 IMU 预积分.具体而言,在点特征处理方面,选用 ORB^[19]特征点检测方法,并通过 RANSAC(random sample consensus)算法进行误匹配滤除.在线特征处理方面,基于改进的 ILSD(inside line segment detection)算法进行特征提取,并利用

LBD(line band descriptor)描述子实现稳定跟踪. 在后端优化阶段,系统根据前端提供的位姿信息,构建包含点线特征和IMU观测值的约束项,采用滑动窗口策略联合最小化残差函数完成位姿优

化.为提升系统在电缆隧道环境下的鲁棒性,回环检测模块融合了ArUco标记和词袋(bag of words, BoW)模型,通过筛选最佳回环匹配帧实现精准检测,最终完成所有关键帧的全局位姿优化.



注: VIO(visual-inertial odometry);SFM(structure from markup).

图 1 定位算法整体框架

Fig. 1 Overall framework of positioning algorithm

2 点线特征融合的视觉里程计

2.1 点特征跟踪

ORB 特征点整合了 FAST (features from accelerated segment test) 关键点检测算法和 BRIEF (binary robust independent elementary features)描述子,在保持尺度不变性的同时具有较低的计算成本和较高的检测速度,成为兼顾性能与效率的理想选择.基于这一特性,本文选取 ORB 特征点作为基础特征,并通过均匀化提取策略优化特征点分布.图 2 展示了电缆隧道环境下提取的 ORB 特征点分布情况.

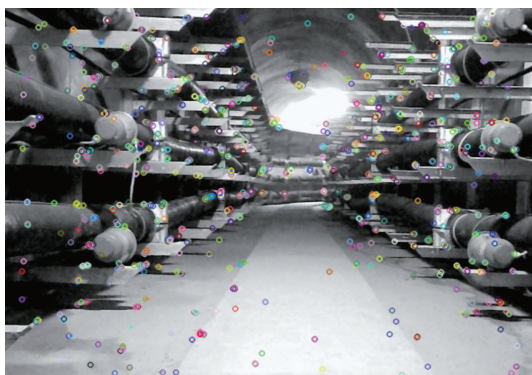


图 2 电缆隧道环境下 ORB 特征点提取效果

Fig. 2 Effectiveness of ORB feature point extraction in cable tunnel environment

特征提取后的正确匹配是运动估计的重要前提,因此,本文设计了基于RANSAC算法的误匹配滤除方法.该方法通过随机抽样和迭代选择匹配的特征点对,构建模型并评估其一致性,从而剔除误匹配特征点并保留正确匹配的特征点.如图 3 所示,经过 RANSAC 算法处理后的特征匹配结果显著提升了电缆隧道环境下的特征跟踪可靠性,为后续位姿估计提供了更准确的数据基础.

2.2 线特征跟踪

电缆隧道包含较为丰富的结构信息,可提取大量高维的结构化特征.高维特征具有信息量更大、约束更强等优势,在特征追踪过程中更不易丢失、成功率更高,符合 SLAM 前端特征稳定性的要求.针对当前场景,考虑在点特征提取基础上,进一步提取线特征,以点线结合的方式来进一步提升 SLAM 定位精度.

本文在提取线特征时选择 LSD(line segment detector)^[20]算法,尽管 LSD 算法在提取线特征的鲁棒性上表现良好,但在结构化电缆架众多的电缆隧道环境下,仍存在提取效率低、短线特征多等问题.因此,本文从两方面来处理短线段:一方面针对长线段因电缆架分割而产生的分段现象,将分段的线段拟合成实际场景原本的长线段;另一方面,为了减少噪声影响和提高匹配效率,对

短线特征进行长度抑制.改进后的 LSD (improved LSD,ILSO)算法流程图如图 4 所示.

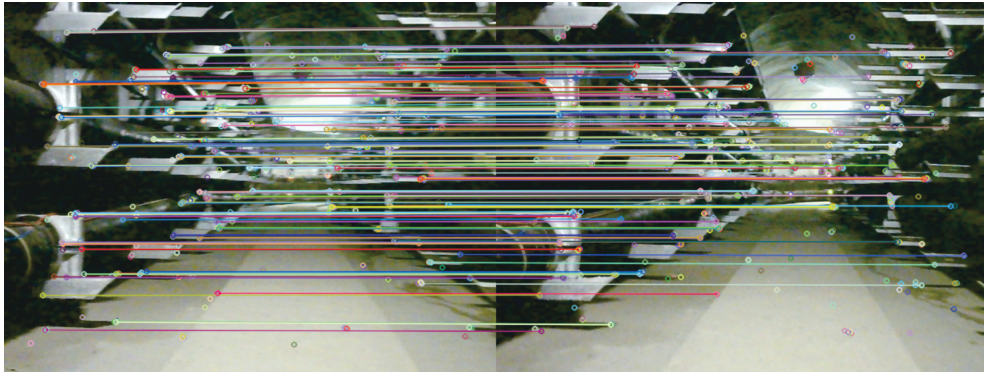


图 3 基于RANSAC算法的误匹配滤除效果

Fig. 3 Effectiveness of mismatch exclusion based on RANSAC algorithm

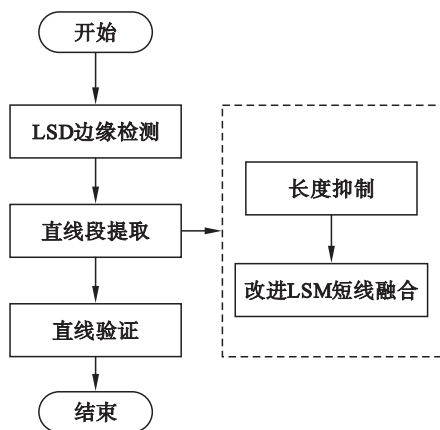


图 4 ILSD算法流程图

Fig. 4 Flowchart of ILSD algorithm

在电缆隧道场景中,LSD算法会提取大量短线段特征,从而增加线特征误匹配的概率和计算成本.因此,需要进行长度(l)抑制,长度抑制公式如式(1)所示:

$$l_i \leq \mu \times \min(W, H), i \in (1, n). \quad (1)$$

其中: $\min(W, H)$ 为图像宽度和高度的较小值; μ 为长度比例经验阈值.针对电缆隧道场景的图像特性,并结合大量实验验证,本文将 μ 设为 0.1,以确保线特征提取的质量和稳定性.

针对交叉线分段现象,Hamid等^[21]提出了1种基于线段合并策略的LSM(line segment merging)算法.该算法通过以下步骤实现线段拟合:首先选取2条线段中相距最远的2个端点作为新线段的端点;然后对拟合结果进行验证,若拟合后线段与基准线的角度偏差超过阈值,则取消本次拟合操作;最后对拟合后的线段在当前分组中迭代执行合并操作,直至无法继续合并为止.在电缆隧道场景的实际应用中,该算法存在一定的局限性:当处理电缆架边缘特征时,由于细微重叠现象的存在,原本应保持独立的多个电缆架线段被错误地合并为单一线段,如图5紫色边框区域所示.这种误拟合现象会影响后续特征匹配和位姿估计的准确性.

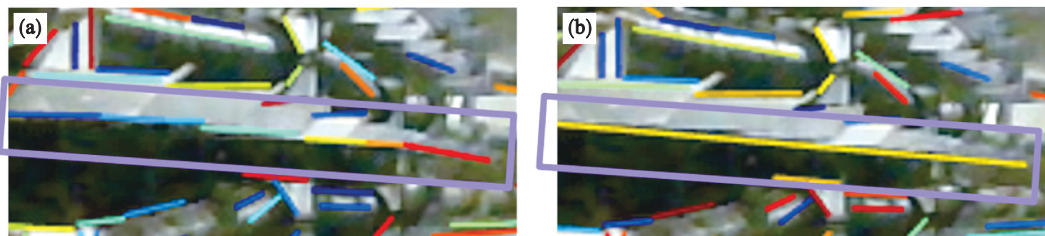


图 5 线段误拟合

Fig. 5 Wrong fitting of line segment

(a)—误拟合前线特征;(b)—误拟合后线特征.

误拟合的线段主要表现为以下两种情况:①基于同一基准线发生多次复拟合;②最终拟合线段长度明显超出初始基准线长度.针对这类误拟合线段,基于合并次数以及拟合线段与初始基

准线段的长度比值,进行筛选和滤除.此外,LSM算法的先分组再拟合策略会导致部分分段线段未发生拟合.因此,本文对该算法进行改进,取消其分组策略,针对每条基准线,直接在

排序线段集合中遍历能与其发生拟合的线段,同时对拟合过程进行追踪,若发现误拟合,则将当前拟合过程逆向还原,保证线段拟合的准确度.改进后的线段拟合结果如图6所示,其中红色线段为经过优化算法处理后的拟合结果.实验结果表明,ILSD算法显著提高了线段拟合的准确性和可靠性.

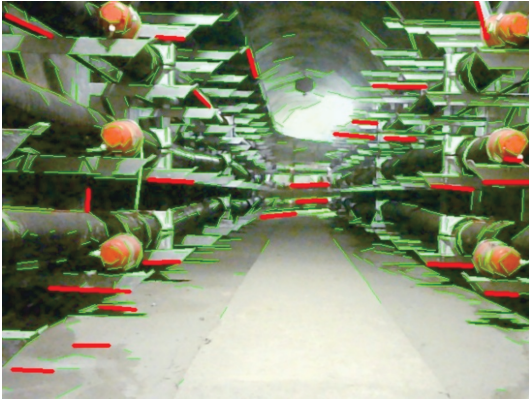


图6 改进后线段拟合结果

Fig. 6 Improved fitting result of line segment

提取线特征后,使用LBD描述子进行特征匹配.在构造LBD描述子的过程中,需要将线支持区域LSR(line support region)划分为多个相同宽度的条带 $B_j(j=1, 2, \dots, m)$.定义条带的2个方向,分别为直线方向向量 d_L 以及顺时针垂直于 d_L 的垂直方向向量 d_{\perp} .以线段中点构建局部坐标系原点,并将LSR内的每个像素梯度 (∇g) 投影到该局部坐标系下:

$$\nabla g = (\nabla g_{d_L}, \nabla g_{d_{\perp}}). \quad (2)$$

分别计算条带 B_j 和相邻条带 B_{j-1}, B_{j+1} 之间

的局部像素梯度,对于条带 B_j 中第 k 行,累积该行中4个方向上像素的梯度:

$$\left. \begin{aligned} v1_j^k &= \lambda \sum_{\nabla g_{d_L} > 0} \nabla g_{d_L}, v2_j^k = \lambda \sum_{\nabla g_{d_L} < 0} -\nabla g_{d_L}, \\ v3_j^k &= \lambda \sum_{\nabla g_{d_{\perp}} > 0} \nabla g_{d_{\perp}}, v4_j^k = \lambda \sum_{\nabla g_{d_{\perp}} < 0} -\nabla g_{d_{\perp}}. \end{aligned} \right\} \quad (3)$$

式中: $\lambda = f_g(k) f_l(k)$ 表示高斯权重; $v1_j^k, v2_j^k, v3_j^k$ 和 $v4_j^k$ 分别为 d_L 方向, d_L 反方向, d_{\perp} 方向和 d_{\perp} 反方向4个方向上的像素梯度和.

最后,将每行像素梯度和进行联立,构成线特征描述子矩阵 C_j :

$$C_j = \begin{bmatrix} v1_j^1 & v1_j^2 & \dots & v1_j^n \\ v2_j^1 & v2_j^2 & \dots & v2_j^n \\ v3_j^1 & v3_j^2 & \dots & v3_j^n \\ v4_j^1 & v4_j^2 & \dots & v4_j^n \end{bmatrix} \in \mathbf{R}^{4 \times n}, \quad (4)$$

$$n = \begin{cases} 2w, j=1 \text{ 或 } m, \\ 3w, \text{其他}. \end{cases}$$

通过计算 C_j 的均值方差 M_j^T 以及均值向量 S_j^T 可以得到每个特征向量 $D_j = (M_j^T, S_j^T)^T \in \mathbf{R}^8$,最终得到LBD描述子为

$$\text{LBD} = (M_1^T, S_1^T, M_2^T, S_2^T, \dots, M_m^T, S_m^T) \in \mathbf{R}^{8m}. \quad (5)$$

在进行回环检测时,需要频繁计算特征向量之间的距离,这严重影响了SLAM系统的实时性.因此,为提高计算效率,将LBD转换为二进制描述子.与BRIEF描述子类似,二进制的LBD特征向量由许多个0和1组成,编码了72维浮点型特征向量中元素之间的大小关系.采用二进制的LBD后,计算汉明距离作为2个特征向量之间的距离度量,可显著提高其匹配的效率.基于LBD描述子的线特征匹配效果如图7所示,线特征的提取有效地描述了电缆隧道轮廓和结构化信息.

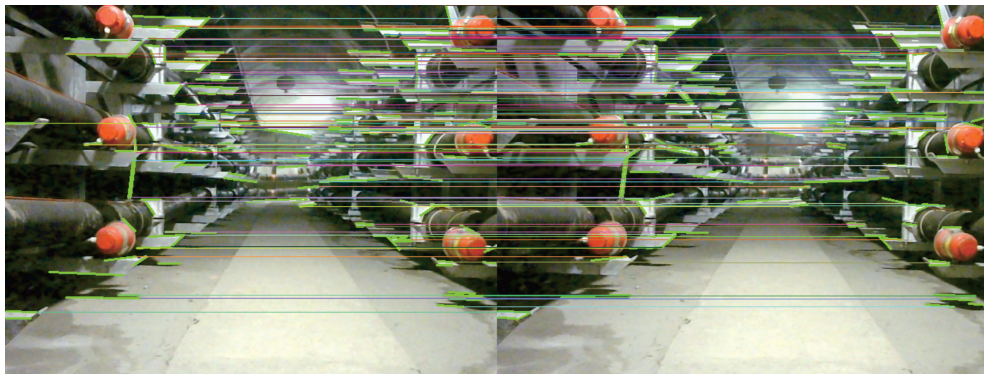


图7 基于LBD描述子的线特征匹配

Fig. 7 Line feature matching based on LBD descriptor

表1给出了改进前后线特征追踪耗时对比,相较于LSD算法,本文提出的线特征提取算法匹

配效率更高.

表 1 线特征追踪耗时对比
Table 1 Line feature tracking time consumption comparison

提取算法	线段总数/条	提取耗时/ms	描述耗时/ms	匹配耗时/ms	总耗时/ms
LSD算法	784	94.855	73.476	86.543	254.874
ILSD算法	150	36.238	28.172	4.751	69.159

3 后端优化与回环检测

3.1 基于滑动窗口的非线性优化

为保证优化效果和实时性,本文采用基于滑动窗口的非线性优化模型.该模型需要对滑动窗口中的状态向量进行联合优化,具体状态向量如下:

$$\min_{\chi} \left\{ \sum_{k \in B} \rho_b \left\| r_B(\hat{z}_{b_{k+1}}^{b_k}, \chi) \right\|_{P_{b_{k+1}}^{b_k}}^2 + \sum_{(i,j) \in F} \rho_f \left\| r_F(\hat{z}_{f_i}^c, \chi) \right\|_{P_{f_i}^c}^2 + \sum_{(i,j) \in L} \rho_l \left\| r_L(\hat{z}_{l_i}^c, \chi) \right\|_{P_{l_i}^c}^2 \right\}. \quad (7)$$

其中: $r_B(\hat{z}_{b_{k+1}}^{b_k}, \chi)$ 表示 IMU 测量残差; $r_F(\hat{z}_{f_i}^c, \chi)$ 表示点测量残差; $r_L(\hat{z}_{l_i}^c, \chi)$ 表示线测量残差. 具体而言, 滑动窗口内包含以下观测数据: B 为所有预积分的 IMU 测量值的集合; F, L 分别为观测帧中点、线的测量值集合; $P_{b_{k+1}}^{b_k}, P_{f_i}^c, P_{l_i}^c$ 分别表示 IMU、点、线的协方差矩阵. 针对点特征、线特征和 IMU 分别定义了对应的平均损失函数 ρ_f, ρ_l, ρ_b . 最后, 通过最小化目标函数来优化滑动窗口内所有的状态

$$\chi = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-1}, \varphi_1, \varphi_2, \dots, \varphi_{m-1}, O_1, O_2, \dots, O_{n-1}]. \quad (6)$$

其中: χ 表示整个状态向量; $\mathbf{x}_i = [p_{b_i}^w, q_{b_i}^w, v_{b_i}^w, b_a, b_g]$ 表示第 i 个滑动窗口中的状态, 由 IMU 位置信息 $p_{b_i}^w$ 、方向信息 $q_{b_i}^w$ 、速度 $v_{b_i}^w$ 、加速度计偏置 b_a 和陀螺仪偏置 b_g 组成; φ 表示点特征的逆深度; O 表示空间线的正交. 式(7)表示优化目标函数:

变量.

3.2 回环检测

在电缆隧道中, 由于场景结构的高度重复性, 相邻帧间图像极为相似. 除电缆箱、机器人充电桩等少数具有辨识度的特征点外, 系统极易在不同空间位置采集到高度相似的图像数据. 常用的基于词袋模型的回环检测方法主要通过计算图像相似度来完成回环候选帧的选择, 进而造成图 8 所示的回环误检测现象.

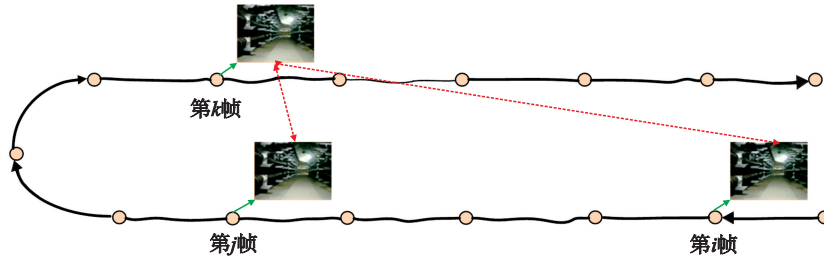


图 8 回环误检测结果

Fig. 8 Wrong loop closure detection results

针对这一问题, 本文引入额外约束以辅助回环检测, 并在基于词袋模型的回环检测基础上, 提出了融合 ArUco 标记的回环检测算法. 如图 9 所示, ArUco 标记提供了唯一的 ID 信息, 在电缆隧道内可以被快速检测和识别; 此外, 利用 ArUco 标记的 4 个角点坐标可以获得标记相对于相机的变换矩阵, 进而筛选最佳回环帧.

该算法的实现主要包括以下 2 个部分.

1) 回环匹配帧对集合筛选策略. 首先, 利用 ArUco 标记的唯一 ID 信息约束回环检测的发生条件: 仅当检测到 ArUco 标记时, 系统才触发回环检测, 且要求回环关键帧与候选帧必须包含相同的 ArUco 标记. 图 10 为利用 ArUco 标记完成的回环检测. 随后, 筛选回环匹配帧对集合, 筛选流

程如图 11 所示, 具体实现思想如下:

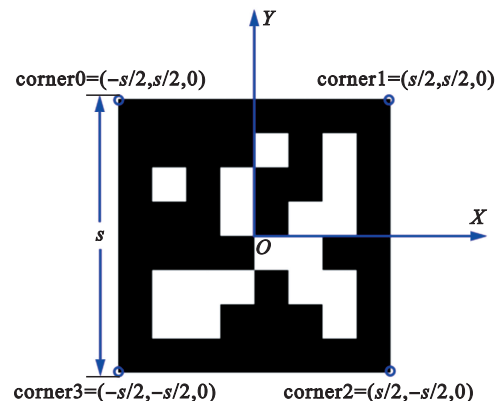


图 9 ArUco 标记图案

Fig. 9 ArUco marker pattern

数据库构建:当相机首次经过某 ArUco 标记附近时,其全部图像帧会存入基于词袋模型构建、对应标记 ID 编号的数据库中。

回环触发与相似度计算:相机再次经过该标记附近时,从检测到该标记的第一帧开始,系统利用词袋模型计算当前帧与其前一帧的图像相

似度 sim_1 ,并在对应标记 ID 的数据库中检索历史回环候选帧。

匹配帧对筛选:计算当前帧与历史最佳候选帧的图像相似度 sim_2 ,并将其与 sim_1 进行比较。若满足预设条件,则将该帧对加入回环匹配帧对集合;否则,予以剔除。

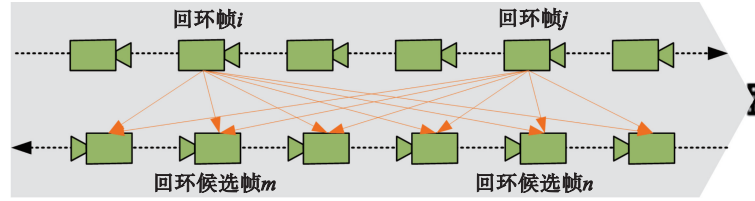


图 10 ArUco 标记多帧回环检测

Fig. 10 ArUco marker multi-frame loop closure detection

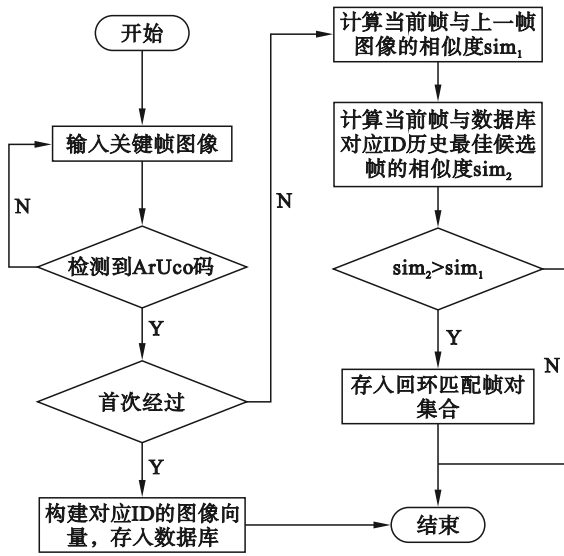


图 11 回环匹配帧对集合筛选策略

Fig. 11 Filtering strategy of loop closure matching frame pair ensemble

2) 最佳回环匹配帧对策略.冗余的回环检测不仅无法提升定位精度,还会增加计算开销.因此,需要对回环匹配帧对进行筛选,选取最佳匹配帧对.其原理是通过 ArUco 标记求解 PnP 问题 (perspective-n-point, 透视 n 点问题:指利用至少 n 个 3D-2D 点的对应关系解算相机位姿),进而利用位姿估计模型解算出匹配帧对的相对位姿,具体实现过程见算法 1.

算法 1 最佳回环匹配帧对选择算法

输入:回环匹配帧对组合 U , 帧间相机光心距离 d

输出:最佳回环帧及最佳回环候选帧 (f_i, f_j)

1 选取 U 中 1 对回环匹配帧组合 (f_m, f_n) ;

2 对该组合的 2 帧图像,提取特征并完成匹配,利用对极几何求解帧间位姿变换矩阵 T_1 ;

3 根据 ArUco 码对 2 帧图像的观测,分别利用 PnP 求解世界坐标系到 2 帧图像坐标系的位姿变换矩阵 T_i 和 T_j ,代换求解帧间位姿变换矩阵 T_2 ;

4 对于第 1 张图像, $d = (T_1 \cdot X)^T \cdot (T_2 \cdot X)$, 赋值 $(f_i, f_j) = (f_m, f_n)$;

5 对于后续图像,如果 $(T_1 \cdot X)^T \cdot (T_2 \cdot X) \leq d$,赋值 $(f_i, f_j) = (f_m, f_n)$;

6 循环执行上述步骤,直至 U 中组合全部被选取,返回 (f_i, f_j) ,算法结束。

ArUco 标记的位姿测量模型如图 12 所示, I_1 和 I_2 分别为当前帧图像和回环帧图像.假设 ArUco 标记某一点在世界坐标系下的坐标为 P_a , 在 O_1 相机坐标系下的坐标为 P_1 , 在 O_2 相机坐标系下的坐标为 P_2 .已知世界坐标系到 O_1 相机坐标系的变换矩阵为 T_1 , 到 O_2 相机坐标系的变换矩阵为 T_j , 则依据 ArUco 标记与相机坐标系间的位姿估算,可得

$$\begin{cases} P_1 = T_1 \cdot P_a; \\ P_2 = T_j \cdot P_a. \end{cases} \quad (8)$$

消去 ArUco 标记坐标 P_a , 则有

$$P_2 = T_j \cdot T_1^{-1} \cdot P_1. \quad (9)$$

最终得到相机坐标系 O_1 到相机坐标系 O_2 的位姿变换矩阵, $T_2 = T_j \cdot T_1^{-1}$.

得到帧间相机的位姿变换矩阵后,两相机光心 O_1 到 O_2 的距离 d 可通过变换矩阵的二范数值表述,令单位向量 $X = [1 \ 1 \ 1 \ 1]^T$, 则有

$$d = (T_1 \cdot X)^T \cdot (T_2 \cdot X). \quad (10)$$

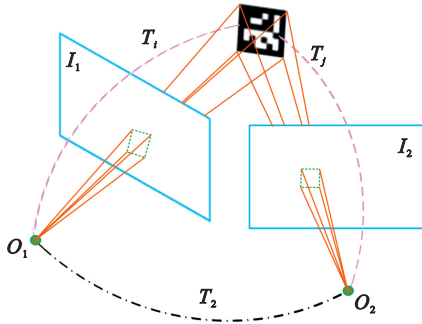


图 12 ArUco 标记位姿测量模型

Fig. 12 ArUco marker pose measurement model

对全部回环关键帧组合,利用 ArUco 标记位姿测量模型,分别计算得到多组帧间相机的光心距离,取距离最小的回环关键帧组合,即可得到最佳回环帧及其最佳回环候选帧。

4 实验结果与讨论

本文在自建的电缆隧道数据集上进行测试,以验证所提算法的准确性和鲁棒性.该数据集主要分为两类:直道数据和弯道数据.为了更好地展示算法评估结果,本文从数据集介绍、系统评价指标介绍和定位结果分析三方面展开论述。

4.1 数据集介绍

本算法的应用场景为电缆隧道,其中包括较长的直道场景及半径相对小的弯道场景,由于隧道空间狭窄,履带式机器人只能沿中间窄过道前进、后退及进行角度较小的偏航角旋转运动.因此,为获取符合本文算法的数据集,其数据采集操作流程如下:

1) 固定机器人起始位置(标记1),以该固定位置侧前方(偏x轴负方向约0.2 m,偏y轴正方向约5 m)放置 ArUco 标记,机器人的运动坐标系如图 13 所示;

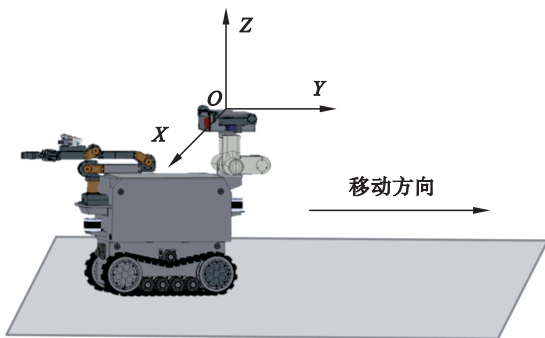


图 13 机器人运动坐标示意图

Fig. 13 Schematic of robot kinematic coordinates

2) 开始录制数据集,机器人在起始固定位置附近进行初始化,分别对 X, Y, Z 三轴进行 3 次平移后,机器人回到起始标记 1 位置;

3) 机器人开始分别沿直道、弯道行驶至固定位置(标记 2),然后按照来时路径回退折返至标记 1 位置,数据录制结束并保存。

4.2 系统评价指标介绍

在 SLAM 系统中,通用数据集包含了机器人运动轨迹每帧数据对应的真实位姿值.因此,可采用相对位姿误差 RPE(relative pose error)和绝对轨迹误差 ATE(absolute trajectory error)2 个指标评价定位精度.在电缆隧道场景中,由于无法使用 GPS 等外部信号且动作捕捉系统成本高达数百万元,难以实现真值数据的实时获取.为评估本文算法的有效性,针对上述采集数据集,从以下两方面对方案效果进行评估:

1) 绝对评价指标:机器人起始位置 $p_s = (x_s, y_s, z_s)$ 与结束位置 $p_e = (x_e, y_e, z_e)$ 的重合度,即端点距离,该距离越小表明优化轨迹的累积误差越小,定位精度越高:

$$d(p_s, p_e) = \sqrt{(x_s - x_e)^2 + (y_s - y_e)^2 + (z_s - z_e)^2}. \quad (11)$$

2) 相对评价指标:以机器人前进路径作为参考基准,对直道段轨迹进行直线拟合;随后计算返回路径中对应直道段的轨迹点到拟合直线的距离分布,包括均值距离、中值距离和最大距离.这 3 个距离越小,表明往返路径的轨迹重合度越高,间接反映系统位姿估计的精度越高.假设有 n 个 3D 空间点 $(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)$, 以及待拟合直线 $L: ax + by + cz + e = 0$, 基于点到直线距离(d_i)的平方和构建误差函数:

$$f(a, b, c, e) = \sum_{i=1}^n d_i^2 = \sum_{i=1}^n \left(\frac{ax_i + by_i + cz_i + e}{\sqrt{a^2 + b^2 + c^2}} \right)^2. \quad (12)$$

式中, a, b, c, e 表示待拟合直线方程中的待求解参数。

通过最小化该误差函数,可以得到最佳直线拟合参数.直道轨迹拟合后,对应该直道返回路径的空间点到拟合线段的均值距离、中值距离分别为

$$d_{\text{avg}} = \frac{\sum_{i=1}^k \frac{|ax_i + by_i + cz_i + e|}{\sqrt{a^2 + b^2 + c^2}}}{k}, \quad (13)$$

$$d_{\text{mid}} = \text{mid} \left(\sum_{i=1}^k \frac{|ax_i + by_i + cz_i + e|}{\sqrt{a^2 + b^2 + c^2}} \right). \quad (14)$$

其中: k 表示回时轨迹对应的 3D 点数量; (x_i, y_i, z_i)

表示第 i 帧图像对应的坐标.

4.3 定位结果分析

在电缆隧道的直道和弯道数据集上,将本文算法与 VINS-Mono 进行对比.表2给出了定量分析的结果.结果表明,线特征的引入和融合 ArUco 标记的回环检测使得本文算法获得了更好的指

标.与 VINS-Mono 相比,在直道场景的数据集下,绝对定位精度提升了 61.19%,相对评价指标中的均值距离、中值距离分别提升了 41.03%, 41.33%;在弯道场景的数据集下,本文算法的绝对定位精度提升了 78.26%,相对评价指标中均值距离、中值距离分别提升了 81.89%, 82.42%.

表2 直道和弯道数据集下评价指标对比

Table 2 Comparison of evaluation metrics under straight and curve datasets

m

评价指标	直道数据集			弯道数据集		
	d	d_{avg}	d_{mid}	d	d_{avg}	d_{mid}
VINS-Mono/no loop	0.572	0.234	0.225	1.095	0.718	0.745
Ours-SLAM/no loop	0.353	0.143	0.140	0.862	0.454	0.493
Ours-SLAM /loop	0.222	0.138	0.132	0.238	0.130	0.131

此外,图14给出了本文算法回环优化轨迹与 VINS-Mono 里程计轨迹的对比效果.根据轨迹对比效果可以看出,在2种数据集下,无论是起始

端点与终止端点的重合度,还是去时轨迹与回时轨迹的重合度,本文算法优化轨迹的起止端点更近、重合度更高.

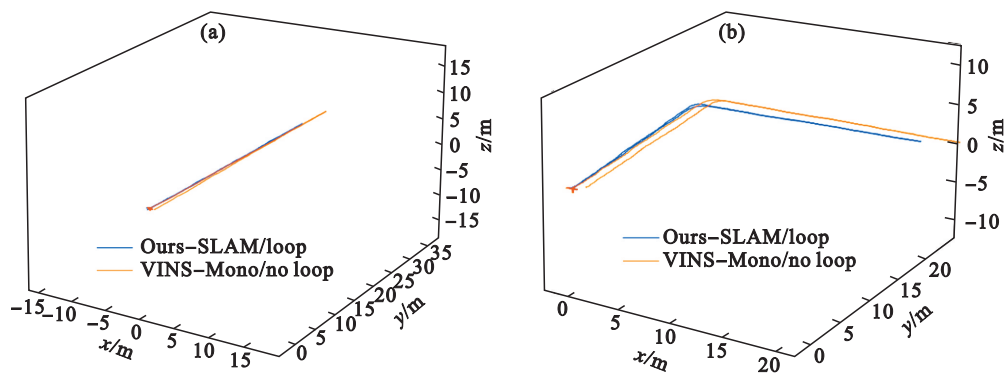


图14 轨迹结果比较

Fig. 14 Trajectory results comparison

(a)一直道数据集轨迹对比; (b)一弯道数据集轨迹对比.

5 结 语

针对电缆隧道复杂环境下工业智能巡检机器人的巡检需求,本文提出了基于点线特征融合的视觉惯导自主定位算法.该算法通过在视觉里程计中引入改进的高维线特征,使巡检机器人在低纹理、特征单一的隧道退化场景中能有效地获取图像特征并完成匹配,为后端优化提供充足条件.针对电缆隧道环境特征相似性导致的回环检测失效问题,本文通过引入 ArUco 标记,利用其 ID 信息和位姿测量模型,有效实现了机器人的正确回环识别,从而提升了系统的定位精度.利用自建的电缆隧道数据集进行实验,结果表明本文提出的算法与 VINS-Mono 相比,定位精度更高,满足隧道场景应用需求.在后续工作中,笔者计

划进一步提高线特征的提取效率,提高算法性能.

参考文献:

- [1] Mur-Artal R, Montiel J M M, Tardós J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. *IEEE Transactions on Robotics*, 2015, 31(5): 1147-1163.
- [2] Shan T X, Englot B. LeGO-LOAM: lightweight and ground-optimized lidar odometry and mapping on variable terrain [C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid: IEEE, 2018: 4758-4765.
- [3] Mourikis A I, Roumeliotis S I. A multi-state constraint Kalman filter for vision-aided inertial navigation [C]//Proceedings 2007 IEEE International Conference on Robotics and Automation. Rome: IEEE, 2007: 3565-3572.
- [4] Qin T, Li P L, Shen S J. VINS-Mono: a robust and versatile monocular visual-inertial state estimator [J]. *IEEE Transactions on Robotics*, 2018, 34(4): 1004-1020.
- [5] Dang T, Mascarich F, Khattak S, et al. Autonomous search for underground mine rescue using aerial robots [C]//2020

- IEEE Aerospace Conference. Big Sky: IEEE, 2020: 1–8.
- [6] Azpurua H, Campos M F M, Macharet D G. Three-dimensional terrain aware autonomous exploration for subterranean and confined spaces [C]//2021 IEEE International Conference on Robotics and Automation (ICRA). Xi'an: IEEE, 2021: 2443–2449.
- [7] Campos C, Elvira R, Rodríguez J J G, et al. ORB-SLAM3: an accurate open-source library for visual, visual-inertial, and multimap SLAM [J]. *IEEE Transactions on Robotics*, 2021, 37(6): 1874–1890.
- [8] Cao S Z, Lu X Y, Shen S J. GVINS: tightly coupled GNSS-visual-inertial fusion for smooth and consistent state estimation [J]. *IEEE Transactions on Robotics*, 2022, 38(4): 2004–2021.
- [9] Tang H L, Niu X J, Zhang T S, et al. LE-VINS: a robust solid-state-LiDAR-enhanced visual-inertial navigation system for low-speed robots [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 8502113.
- [10] Song B Y, Yuan X F, Ying Z M, et al. DGM-VINS: visual-inertial SLAM for complex dynamic environments with joint geometry feature extraction and multiple object tracking [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 8503711.
- [11] Li M L, Zhang H, Shen T A, et al. SM-VINS: a fast and decoupled monocular visual-inertial sensors SLAM system with stepwise marginalization [J]. *IEEE Sensors Journal*, 2024, 24(20): 33240–33251.
- [12] Zhou H Z, Zou D P, Pei L, et al. StructSLAM: visual SLAM with building structure lines [J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(4): 1364–1375.
- [13] Forster C, Pizzoli M, Scaramuzza D. SVO: fast semi-direct monocular visual odometry [C]//2014 IEEE International Conference on Robotics and Automation (ICRA). Hong Kong: IEEE, 2014: 15–22.
- [14] Gomez-Ojeda R, Briales J, Gonzalez-Jimenez J. PL-SVO: semi-direct monocular visual odometry by combining points and line segments [C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Daejeon: IEEE, 2016: 4211–4216.
- [15] Fu Q, Wang J, Yu H, et al. PL-VINS: real-time monocular visual-inertial SLAM with point and line features [J]. *arXiv preprint arXiv*, 2020: 2009.07462.
- [16] Lim H, Kim Y, Jung K, et al. Avoiding degeneracy for monocular visual SLAM with point and line features [C]//2021 IEEE International Conference on Robotics and Automation (ICRA). Xi'an: IEEE, 2021: 11675–11681.
- [17] Zhu Y Q, Jin R, Lou T S, et al. PLD-VINS: RGBD visual-inertial SLAM with point and line features [J]. *Aerospace Science and Technology*, 2021, 119: 107185.
- [18] Lee J, Park S Y. PLF-VINS: real-time monocular visual-inertial SLAM with point-line fusion and parallel-line fusion [J]. *IEEE Robotics and Automation Letters*, 2021, 6(4): 7033–7040.
- [19] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C]//2011 IEEE International Conference on Computer Vision. Barcelona: IEEE, 2011: 2564–2571.
- [20] Von Gioi R G, Jakubowicz J, Morel J M, et al. LSD: a line segment detector [J]. *Image Processing on Line*, 2012, 2(4): 35–55.
- [21] Hamid N, Khan N. LSM: perceptually accurate line segment merging [J]. *Journal of Electronic Imaging*, 2016, 25(6): 061620.