

文章编号 1004-924X(2026)13-2109-15

基于滑动窗口与共视关系增强的多机视觉 同时定位与建图回环检测

齐光峪¹, 李林阳^{1*}, 王安成¹, 李艺博¹, 郭文卓², 赵冬青¹

(1. 信息工程大学 地理空间信息学院, 河南 郑州 450001;

2. 中国人民解放军 61365 部队, 天津 300142)

摘要:多机视觉同时定位与建图(Simultaneous Localization and Mapping, SLAM)回环检测分为机器人内部回环和机器人间回环,分别用于局部轨迹优化和机器人间的相对位姿获取。针对多机视觉SLAM在光照变化与视角差异场景下回环检测效果不佳且易产生误匹配与漏检的问题,提出一种基于滑动窗口与共视关系增强的多机视觉SLAM回环检测方法。构建基于滑动窗口的候选帧筛选机制,利用图像相似度得分的均值和方差调整检测阈值,避免回环漏检问题。针对机器人内部回环具备时序连续性的特点,引入了基于滑窗的时序一致性检验,有效剔除由相似场景诱发的偶然性误匹配。最后,设计了共视关系增强的几何校验方法,通过引入共视帧构建局部子图扩展观测视野,有效提高了回环检测的质量。采用户外、隧道和地下场景的多份数据集进行了实验验证,结果表明,相比DBOW2方法,所提方法检测高质量回环的比例提升了6%以上,绝对轨迹误差降低了10.0%~63.7%。

关键词:自主导航;多机视觉同时定位与建图;机器人内回环;机器人间回环;滑动窗口;共视关系

中图分类号:TP391.4;TP242 **文献标识码:**A

doi:10.37188/OPE.20263413.2109

CSTR:32169.14.OPE.20263413.2109

Multi-robot visual SLAM loop closure detection using sliding window and co-visibility relationship enhancement

QI Guangyu¹, LI Linyang^{1*}, WANG Ancheng¹, LI Yibo¹, GUO Wenzhuo², ZHAO Dongqing¹

(1. School of Surveying and Mapping, Information Engineering University, Zhengzhou 450001, China;

2. 61365 Troops of the Chinese People's Liberation Army, Tianjin 300142, China)

* Corresponding author, E-mail: linyangli@whu.edu.cn

Abstract: In multi-robot visual Simultaneous Localization and Mapping (SLAM), loop closure detection consists of intra-robot and inter-robot loop closures, serving local trajectory optimization and inter-robot relative pose estimation, respectively. Nevertheless, it remains challenged by illumination changes and viewpoint differences, resulting in false matches and missed detections. Loop closure detection in multi-robot visual SLAM is challenged by illumination changes and viewpoint differences. These conditions lead to false matches and missed detections. This paper proposes a robust loop closure detection method with three contributions. First, a sliding-window candidate frame screening mechanism is designed. It adapts

收稿日期:2026-05-06;修订日期:2026-05-27.

基金项目:国家自然科学基金资助项目(No. 42474043);河南省自然科学基金优秀青年基金资助项目(No. 252300421205)

the detection threshold using the mean and variance of image similarity scores. This strategy effectively prevents missed loop closures. Second, a sliding-window temporal consistency test is introduced. It exploits the temporal continuity of intra-robot loop closures. This test eliminates spurious false matches caused by similar scenes. Third, a covisibility-enhanced geometric verification method is developed. It constructs a local submap by leveraging covisible neighbor frames. This expands the observation field and improves verification quality. Experiments are conducted on multiple datasets from outdoor, tunnel, and underground scenes. Compared with DBOW2, the proposed method increases the proportion of high-quality loop closures by more than 6%. The absolute trajectory error is reduced by 10.0% to 63.7%.

Key words: multi-robot visual simultaneous localization and mapping; intra-robot loop closure; inter-robot loop closure; sliding window; co-visibility relationship

1 引言

同时定位与建图技术 (Simultaneous Localization and Mapping, SLAM) 是赋予机器人环境感知与自主行动能力的关键。近年来,随着应用场景从单一封闭空间向大规模、复杂化环境拓展,如灾难救援^[1]、工业巡检和地下区域勘探^[2],单机器人 SLAM 在效率、稳定性与覆盖范围上的不足日益凸显^[3]。为了克服单机器人的局限性,广泛采用多机器人协同 SLAM 方案。多机器人通过群体协作,并行地探索和建模环境,显著提升了任务执行效率与系统的容错能力^[3-4]。然而,多机协同 SLAM 面临如何将各个机器人的局部轨迹和局部地图融合成全局轨迹和全局地图的挑战,其高度依赖单机器人内部回环和多机器人间回环的效率、准确性和稳定性^[5]。其中,机器人内部回环用于维护局部轨迹的精确性,机器人间回环实现了不同机器人之间的稳健关联^[6]。

在回环检测过程中,基于词袋模型 (Bag-of-Words, BOW)^[7]的方法因其高效性与轻量化特性,广泛应用于 ORB-SLAM 系列^[8-10]、Kimera^[11]等主流系统。然而,BOW 模型本质上依赖图像的表观相似性,在面对剧烈光照变化、视角变化等场景时,极易产生感知混淆,导致回环误报或漏报^[12]。为了解决这一问题,现有研究主要依赖深度学习特征 (如 NetVLAD^[13], SuperPoint^[14], Patch-NetVLAD^[15]) 或几何校验,但深度学习带来的高计算开销在无 GPU 支持的嵌入式平台上难以实时部署^[16-17],传统的几何校验多局限于帧对帧的匹配模式,当图像因运动模糊或遮挡导致特征点不足时,回环极易失败。

为应对通信受限与环境干扰^[18],当前主流多机视觉 SLAM 也采用轻量级的词袋模型,如集中式架构 CCM-SLAM^[19]采用 DBOW2 进行回环检测,虽实现了数据共享,但仍受回环方法的局限性影响且存在单点故障风险,难以适应地下或灾后等通信不稳定的环境^[20]。分布式架构的典型代表 Kimera-Multi^[21]虽增强了系统弹性,但其设计侧重于通信效率,缺乏强有力的几何校验环节,转而依赖后端优化。如成对一致最大化 (Pairwise Consistency Maximization, PCM)^[22]和渐进非凸性 (Graduated Non-Convexity, GNC)^[23],剔除错误匹配,但当环境相似度极高导致误匹配过多时,系统仍面临轨迹漂移风险。除词袋模型外,一些多机视觉 SLAM 系统也尝试采用深度学习的方法进行回环检测,如 DOOR-SLAM^[24]结合 NetVLAD 描述符进行回环检测,虽提升了识别能力,但是难以满足大范围建图的需求。因此,在保持轻量化的前提下,亟需提升回环检测的准确性和稳定性,以提升多机 SLAM 协同的性能。

考虑到视觉 SLAM 运行时会维护一个由关键帧和地图点构成的稀疏时空图,其共视关系蕴含了丰富的场景结构信息,这些内部状态隐式信息可以为外观相似性失效情形提供稳定的判别依据,最终增强与验证回环检测。基于此思想,提出了基于滑动窗口与共视关系增强的多机视觉 SLAM 回环检测方法。该方法构建了基于滑动窗口统计特性的候选帧筛选机制,用以调整阈值并避免漏检。在此基础上实施差异化的校验方法,针对机器人内部回环,额外引入时序一致性检验以剔除偶然性误匹配,并结合共视关系进行几何

校验;针对机器人间回环,则利用跨机共视关系构建局部子图扩展观测视野,克服多机异步观测带来的光照变化与视角差异问题,从而提高回环检测的质量。实验采用开源和自采数据集,验证了所提方法的优越性。

2 原 理

图 1 给出了基于本文改进回环检测方法的多

机视觉 SLAM 框架,由前端、回环检测和后端优化三部分组成,多个机器人之间通过通信交互模块连接。在回环检测部分,机器人内部回环和机器人间回环均采用了基于滑窗统计特性的候选帧筛选。在此基础上,机器人内部回环检测进一步融合基于滑窗的时序一致性检验与共视关系增强,以得到高质量的内部约束;而机器人间回环检测则专注于跨机共视关系增强的几何校验,以精确验证跨机回环并消除异步观测导致的误匹配。

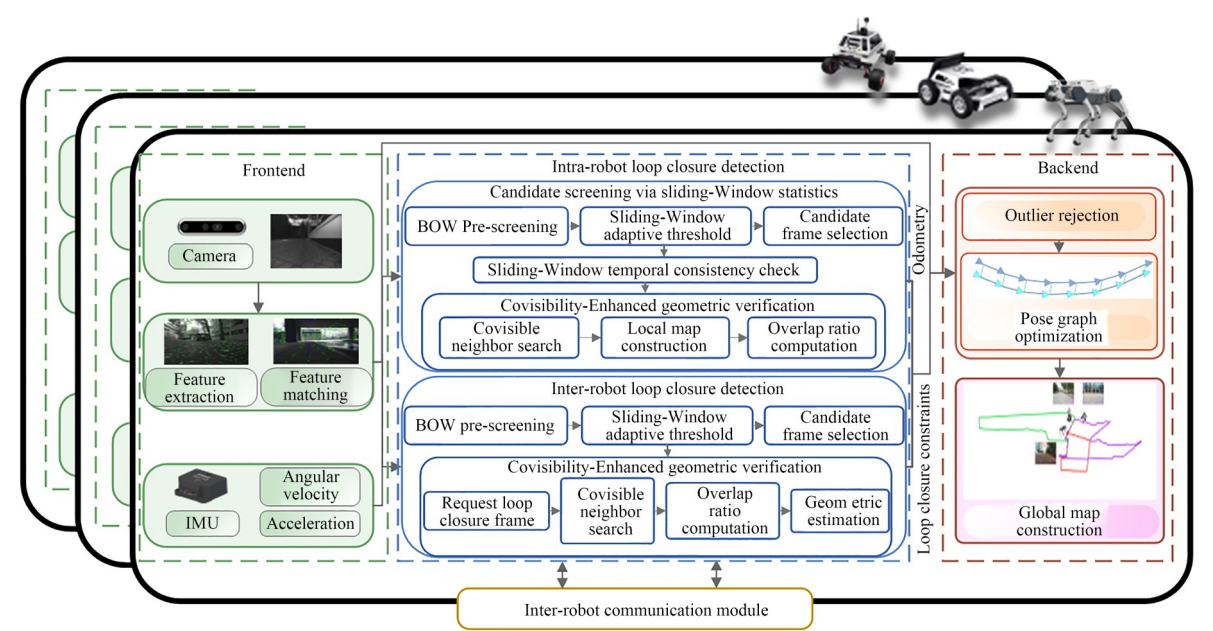


图 1 基于本文改进回环检测方法的多机视觉 SLAM 框架
Fig. 1 Multi-robot visual SLAM framework based on improved loop closure detection method

2.1 机器人内部回环和机器人间回环的特性分析

如表 1 所示,机器人内部回环与机器人间回

环虽同属地点识别问题,但在通信依赖、数据存储与校验方式,以及系统目标上存在本质差异。

表 1 两种回环方式的特性分析
Tab. 1 Characteristic analysis of two types of loop closure

特性	机器人内部回环	机器人间回环
相同点	均为地点识别问题	
通信	本地进行,无需通信	受通信带宽和范围限制
方式	本地存储完整信息 可以增加时序一致性检验	机器人间轻量化传输 无法增加时序一致性检验
目的	提升局部轨迹精度	构建全局坐标系

2.2 基于滑动窗口的机器人内部回环和机器人间回环筛选

2.2.1 基于BOW向量的候选帧初步筛选

采用BOW向量进行机器人内部和机器人间回环检测,将当前帧的局部特征转化为全局描述向量,并从历史关键帧数据库中初筛出潜在的匹配对。

首先对当前帧 I_q 提取ORB特征点集,利用预训练的视觉字典树结构,将特征描述子空间量化为 K 维的词袋向量 $\mathbf{v}_q$,向量 $\mathbf{v}_q$ 中的每一维分量 $\mathbf{v}_q^{(i)}$ 采用词频-逆向文件频率(Term Frequency-Inverse Document Frequency, TF-IDF)权重方案进行计算:

$$\mathbf{v}_q^{(i)} = \frac{m_i}{m} \cdot \log \frac{D}{D_i}. \quad (1)$$

通过该映射,当前帧 I_q 被表示为高维稀疏空间中的一个单位向量 $\mathbf{v}_q$ 。

在检索阶段,需要从数据库中查找与当前帧 I_q 表现最相似的历史帧 I_c ,完成候选帧的初步筛选。如图2所示,为提高检索效率,维护倒排索引结构,快速检索出数据库中与 I_q 至少共享一个视觉单词的所有历史关键帧,这些被检索到的帧构成一个候选集合 S_r 。

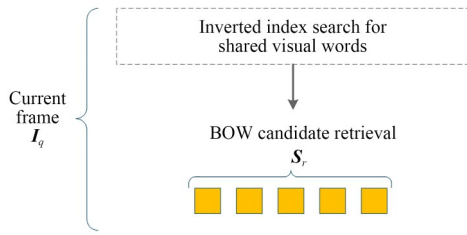


图2 基于BOW向量的候选帧初步筛选

Fig.2 Initial screening of candidate frames based on BOW vectors

2.2.2 基于滑动窗口统计特性的候选帧筛选

对候选集合 S_r 进一步进行精确筛选时,现有基于DBOW2回环检测模块的视觉SLAM系统通常采用固定阈值方法,如ORB-SLAM^[8],CCM-SLAM^[19]等框架通过计算初步筛选的候选帧与当前帧的相似度得分、上一帧与当前帧的相似度得分,仅保留两种得分的比值超过一定阈值的候选帧。但当机器人经过纹理重复区域时,极短时间内相似度得分会出现异常极值,导致检测

阈值过于严苛,最终造成漏检。为解决这一问题,提出了一种基于滑窗统计特性的候选帧筛选方法,旨在通过评估窗口内的平均外观响应来构建更稳定的判断基准。

2.2.1.1 机器人内部回环

对于候选集合 S_r 的每一个候选帧 I_c ,采用归一化的L1范数计算其与当前帧之间的相似度得分 $s(\mathbf{v}_q, \mathbf{v}_c)$:

$$s(\mathbf{v}_q, \mathbf{v}_c) = 1 - \frac{1}{2} \left\| \frac{\mathbf{v}_q}{\|\mathbf{v}_q\|} - \frac{\mathbf{v}_c}{\|\mathbf{v}_c\|} \right\|_1, \quad (2)$$

其中: $s(\mathbf{v}_q, \mathbf{v}_c) \in [0, 1]$,反映了外观相似度。随后,对于单个机器人,维护一个固定大小的滑动窗口,如图3所示,存储窗口内最近 N 帧与其上一个关键帧的相似度得分。

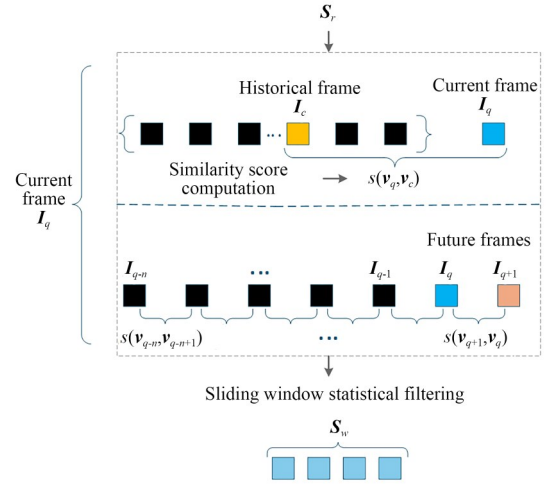


图3 基于滑窗统计特性的候选帧筛选流程

Fig.3 Sliding window statistics based candidate frame selection process

计算滑动窗口内相似度得分的均值 μ_w 与标准差 σ_w 为:

$$\begin{cases} \mu_w = \frac{1}{N} \sum_{k=0}^{N-1} s(\mathbf{v}_{q-k}, \mathbf{v}_{q-k-1}) \\ \sigma_w = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} (s(\mathbf{v}_{q-k}, \mathbf{v}_{q-k-1}) - \mu_w)^2} \end{cases} \quad (3)$$

通过候选帧与当前帧的相似度得分和滑动窗口内相似度得分的均值计算比值:

$$\tau(\mathbf{v}_q, \mathbf{v}_c) = \frac{s(\mathbf{v}_q, \mathbf{v}_c)}{\mu_w}. \quad (4)$$

得分比值的阈值设定为:

$$\tau_{\text{thres}} = 1 - \lambda \cdot \frac{\sigma_w}{\mu_w}. \quad (5)$$

其中: λ 为波动调节系数。在光照变化或者视角变化大的区域, σ_w 增大,阈值降低,此时可以保留更多的候选回环。当 $\tau(v_q, v_c) > \tau_{\text{thres}}$ 时,保留该候选帧,否则进行剔除。

2.2.1.2 机器人回环

如图4所示,当机器人 β 处于 α 的通信范围内(图中灰色圆圈)时,机器人 α 发送采用DBOW2库生成的、自上次相遇以来的新关键帧的BOW向量集合 V_c 。机器人 β 接收后,将查询向量 $v_c \in V_c$ 与本地关键帧的向量 v_q 进行比对,计算归一化相似度得分。对于机器人 β ,利用滑动窗口计算以 I_q 为中心,窗口大小为 N 帧的相邻关键帧相似度得分的均值和方差,随后利用式计算阈值,若相似度得分高于阈值,则判定 I_q 为候选帧。通过基于滑窗统计特性的策略,进一步筛选得到的机器人内部和机器人回环的候选集合为 S_w 。

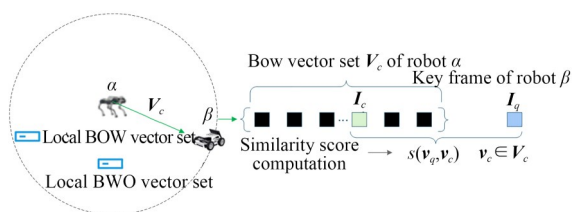


图4 通信范围内BOW向量集合传输与相似度得分计算
Fig.4 Transmission of BOW vector sets and similarity score calculation within communication range

2.2.3 基于滑窗的时序一致性约束检验

尽管可通过在当前帧周围统计滑动窗口特性,减少回环漏检,但在高度重复场景下,表现相似度引发的感知混淆仍会导致误匹配。考虑到真实的回环检测通常在时间序列上表现出较强的连续性,即当机器人经过之前访问过的区域时,应当检测到连续多帧的相似度均较高,而非偶然孤立的单帧匹配,因此,针对机器人内部回环候选帧,进一步引入了基于滑窗的时序一致性检验,避免单帧误匹配问题。图5展示了理想且直观的匹配对,对于已筛选的候选匹配对

(I_q, I_c) ,定义其符合时序一致性的支持集 S_t 为:

$$S_t(I_q, I_c) = \left\{ (I_i, I_j) \in H \mid \begin{array}{l} |t(I_i) - t(I_q)| \leq \Delta t \\ |t(I_j) - t(I_c)| \leq \Delta t \end{array} \right\}, \quad (6)$$

其中: H 为保存了通过滑窗统计筛选的 N_t 个候选对缓存, $t(\cdot)$ 表示关键帧的时间戳, Δt 为设定的时间滑动窗口半径。为了量化 I_c 的可靠性,由滑动内包含的匹配对数量,定义时序一致性评分 $F_t(I_q, I_c)$ 为:

$$F_t(I_q, I_c) = |S_t(I_q, I_c)|, \quad (7)$$

其中: $|\cdot|$ 统计支持集的数量。该评分直观反映了当前回环候选在时间维度上连续性的置信度。最终的有效性判定准则为:

$$\text{Valid}(I_q, I_c) = \begin{cases} \text{True}, & \text{if } F_t(I_q, I_c) \geq \zeta \\ \text{False}, & \text{otherwise} \end{cases}, \quad (8)$$

其中: ζ 为最小序列一致性阈值,通过基于滑窗的时序一致性检验的集合为 S_p 。通过该检验,系统能够有效滤除异常误匹配。

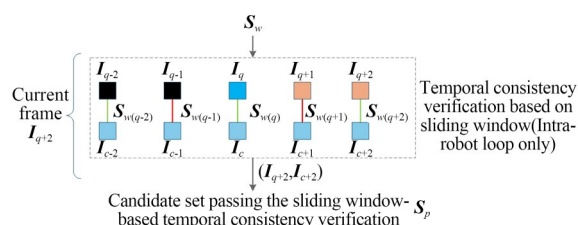


图5 基于滑窗的时序一致性检验

Fig.5 Sliding window based temporal consistency verification

2.3 共视关系增强的机器人内部和机器人回环检测

基于滑窗的机器人内部和机器人回环筛选时,均基于帧与帧之间的相似度得分。然而,当前帧与候选帧之间存在较大视角差异或局部遮挡,两帧之间的公共特征点极少,导致几何解算失败。为解决此问题,进一步利用与候选帧有共同观测的邻居帧,以共视图的方式扩大视野来增加共视关系。

2.3.1 机器人内部回环

如图6所示,对于候选帧 I_c ,在共视图中搜索

与其连接权重最高的 5 个邻居帧,其构成共视集合 $\xi(I_c)$:

$$\xi(I_c) = \{I_n \in M \mid w(I_c, I_n) > w_m\}, \quad (9)$$

其中: M 为关键帧的集合, $w(I_c, I_n)$ 为候选帧和邻居帧的共视权重, w_m 为最小共视权重阈值。

为了评估几何一致性,构建了一个包含 I_c 及其共视邻居 $\xi(I_c)$ 所有观测点的局部地图,并将其投影到查询帧 I_q 上进行几何校验。定义场景结构重合率 S_{intra} 为:

$$S_{\text{intra}}(I_q, I_c) = \frac{|P_{\text{inliers}}|}{|P_{\text{local}}^c|} = \frac{\left| \text{Inliers} \left(I_q, P(I_c) \cup \bigcup_{I_n \in \xi(I_c)} P(I_n) \right) \right|}{\left| P(I_c) \cup \bigcup_{I_n \in \xi(I_c)} P(I_n) \right|}, \quad (10)$$

其中: $|P_{\text{local}}^c|$ 代表候选帧 I_c 及其共视邻居帧观测到的局部地图点集合, $|P_{\text{inliers}}|$ 为该集合投影到当前帧 I_q 上成功匹配并符合几何约束的 3D 地图点数量, $P(I_c)$ 候选帧观测到的 3D 地图点, $P(I_n)$ 候选帧的邻居帧观测到的 3D 地图点。当 S_{intra} 低于阈值时,系统将拒绝该候选帧。通过引入共视关系,将几何校验从帧与帧之间的匹配升级为帧与局部子图的匹配,可以有效减少单一视角和光照突变引起的误匹配。

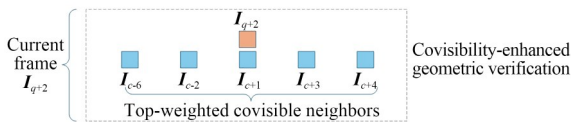


图 6 共视关系增强的几何校验

Fig. 6 Co-visibility enhanced geometric verification

2.3.2 机器人回环

机器人回环检测如图 7 所示。当判定 I_c 为候选帧,随后进入共视关系增强的几何校验阶段,如图中箭头所示,机器人 α 接收 β 发送的候选帧 I_c 及其共视邻居构成的局部地图点集 $|P_{\text{local}}^c|$,

计算场景结构重叠率 S_{inter} :

$$S_{\text{inter}}(I_q, I_c) = \frac{|P_{\text{inliers}}|}{|P_{\text{local}}^c|} = \frac{|\text{Inliers}(I_q, P_{\text{local}}^c)|}{|P_{\text{local}}^c|}. \quad (11)$$

当 S_{inter} 超过阈值时,利用 Nistér^[25] 五点法和 Arun^[26] 三点法计算相对变换。当机器人 α 的第 i 个位姿和机器人 β 的第 j 个位姿之间存在闭环时,相对变换记为 $\hat{Z}_{\beta_j}^{\alpha_i} \in SE(3)$,机器人 α 的局部坐标系记为 C_α ,第 i 个位姿的里程计信息在其坐标系中记为 $\hat{Z}_{\alpha_i}^{C_\alpha} \in SE(3)$;机器人 β 的局部坐标系记为 C_β ,第 j 个位姿的里程计信息在其坐标系中记为 $\hat{Z}_{\beta_j}^{C_\beta} \in SE(3)$ 。利用单个机器人局部坐标系下的位姿和回环得到的机器人之间的相对位姿,估计局部坐标系 C_α 和 C_β 之间的相对变换:

$$\hat{Z}_{C_\beta}^{C_\alpha} \triangleq \hat{Z}_{\alpha_i}^{C_\alpha} \hat{Z}_{\beta_j}^{\alpha_i} \left(\hat{Z}_{\beta_j}^{C_\beta} \right)^{-1}. \quad (12)$$

但是受到传感器噪声、回环检测误差等影响, $\hat{Z}_{C_\beta}^{C_\alpha}$ 存在误差。

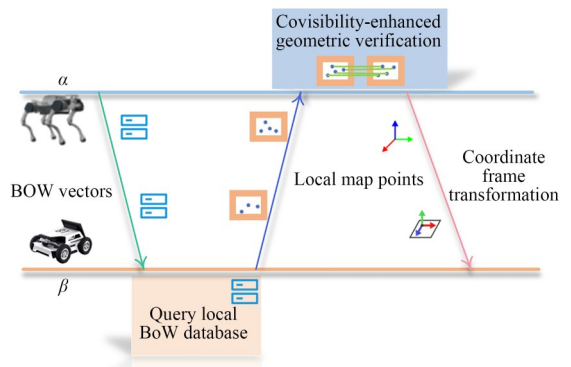


图 7 机器人回环检测过程

Fig. 7 Inter-robot loop closure detection

2.4 多机视觉 SLAM 后端优化

后端优化旨在从含噪声的观测中估计精确的相对变换与全局轨迹。为得到局部坐标系间的最优相对变换,构建以下优化方程:

$$\hat{Z}_{C_\beta}^{C_\alpha} \in \arg \min_{Z \in SE(3)} \sum_{(i,j) \in L_{\alpha,\beta}} \rho \left(\|r_{ij}(Z)\|_\Sigma^2 \right), \quad (13)$$

其中: ρ 是截断最小二乘鲁棒成本函数,采用 GNC 算法求解, $L_{\alpha,\beta}$ 是机器人 α 和 β 之间的回环

集合, $r_{ij}(Z)$ 为 $SE(3)$ 的加权测地线距离。

采用生成树结构进行多个机器人间的坐标系转换,利用变换矩阵的传递性,仅需解算生成树上的必要链路,即可将所有机器人的局部参考系统一至全局坐标系。

全局坐标系建立后,轨迹初值仍受里程计累积漂移影响,且尚未充分利用所有的回环约束。为了获取最优的轨迹估计,利用所有的里程计测量和回环约束信息来优化所有的位姿变量,构建以下全局优化目标函数:

$$\min_{Z_a \in SE(3), \forall a \in R, \forall i} \sum_{a \in R} \sum_{i=1}^{n_a-1} r(Z_{a_i}, Z_{a_{i+1}})^2 + \sum_{(a_i, \beta_j) \in L} \rho(r_{a_i}^{\beta_j}(Z_{a_i}, Z_{\beta_j})), \quad (14)$$

其中: R 为所有机器人合集, $Z_{a_i} = (R_{a_i}, t_{a_i}) \in SE(3)$ 表示机器人 a 在全局坐标系下的第 i 个位姿, n_a 是机器人位姿总数, L 是包括机器人内部回环和机器人间回环在内的所有回环集合。同样采用 GNC 算法求解上述目标函数。

3 实验与分析

为验证本文所提方法的有效性,基于户外、隧道和地下场景的多份开源和自采数据集,分析评估了回环数量、回环质量和定位精度。在回环质量分析时,参考 ORB-SLAM^[8-10],对于给定的候选回环对,通过五点法估计本质矩阵,并统计满足 Sampson 距离约束的特征对数量作为内点数,根据内点数的数量将回环划分为 4 个质量等级。其中,将内点数 < 61 记为 Poor, $61 \sim 84$ 之间记为 Fair, $85 \sim 103$ 之间记为 Good, 大于 103 记为 Excellent。采用绝对轨迹误差 (Absolute Trajectory Error, ATE) 评估定位精度。

3.1 实验数据介绍

3.1.1 开源数据集

采用麻省理工学院 (Massachusetts Institute of Technology, MIT) 2023 年发布的 Kimera Multi Data 多机数据集中的户外、隧道和户外-隧道 3 个场景。图 8 给出了 3 个场景图和轨迹,表 2 列出了 3 个场景的详细参数信息。



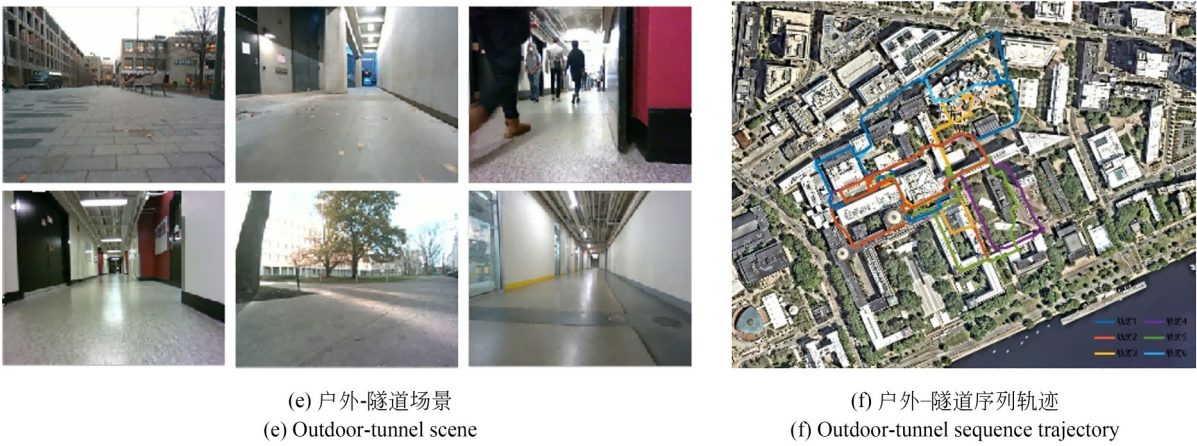


图 8 开源数据场景及轨迹
Fig. 8 Public data scenes and their trajectories

3. 1. 2 自采数据集

为进一步弥补开源数据场景的不足,利用如图 9(a)所示的 3 辆无人小车在光照条件差、纹理结构重复的地下防空洞进行了实验,并采用视觉动捕系统为实验提供轨迹参考值;但受限于是动捕系统的作用范围,如表 2 最后一行所示,自采数据的轨迹长度仅为 65 m 左右。表 3 给出了无人小

车配备的传感器型号和参数信息,自采数据轨迹如图 9(b)所示,图中蓝色、红色和绿色分别为 3 辆无人小车的轨迹(彩图见期刊电子版)。

ORB-SLAM 系列以及 CCM-SLAM^[19]等框架在回环检测时采用相同的固定阈值方法,将这种方法记为改进前方法,采用上一帧和当前帧归一化相似得分进行判定,且无共视关系增强;将

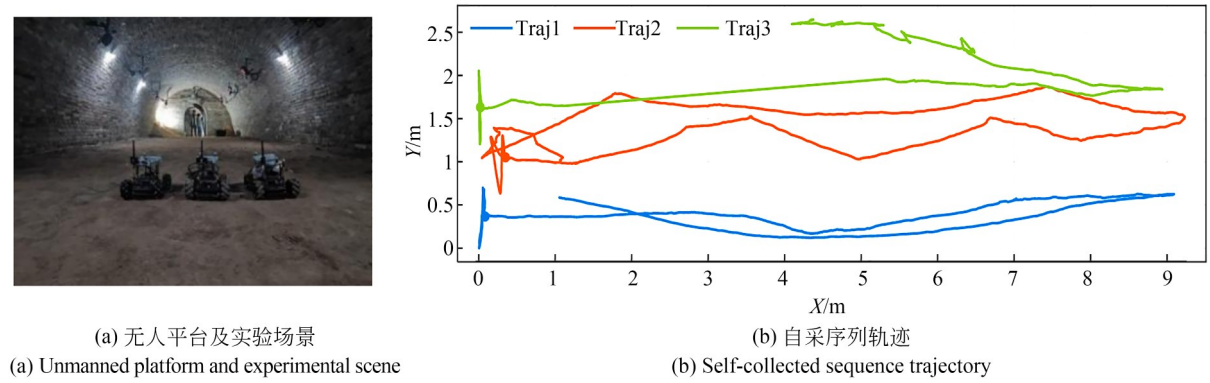


图 9 自采数据实验场景及轨迹
Fig. 9 Experimental scenes and trajectories on self-collected datasets

表 2 开源和自采数据集信息

Tab. 2 Information about public and self-collected datasets

数 据	轨迹长度/m	持续时间/min	采集平台数量	场景特性
户外序列	6 044	19	6	强弱光照变化、动态遮挡
隧道序列	4 522	28	6	结构单调、弱纹理、光照变化
户外-隧道序列	5 441	27	6	光照变化剧烈、视角变化
自采地下序列	65.475	2	3	光照条件差、重复纹理

表 3 自采平台配备的传感器型号和参数

Tab. 3 Unmanned platform device models and detailed specifications

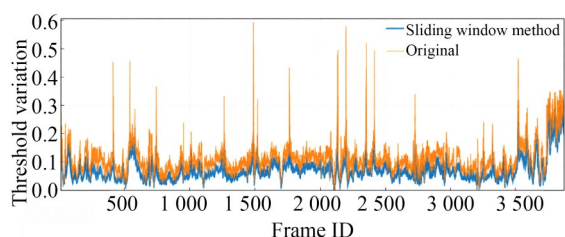
类型	型 号	参 数
相机	RealsenseD455	视场角:水平方向为 86° ,垂直方向为 57° 。 内置 IMU:提供 6DOF 的位姿数据
IMU	BMI055	陀螺仪和加速度计:三轴,12 位分辨率和 16 位分辨率

基于滑窗的候选帧筛选方法改为滑窗方法,并设定波动调节系数 $\lambda=2$,窗口大小 $N=5$,候选对数 $N_r=5$ 且无共视关系增强;将基于滑窗的候选帧筛选与共视关系增强几何校验相结合的方法记为滑窗+共视方法,采用 ORB-SLAM3^[10]的经验阈值设定 $\tau_m=15$,重合率阈值为 0.3,时序一致性约束检验中 $\Delta t=2, \zeta=3$ 。上述参数是本文基于户外、隧道等场景数据集,多次实验确定的,并且采用了部分经验阈值。在其他场景中,可以根据实际情况对参数取值适当微调。

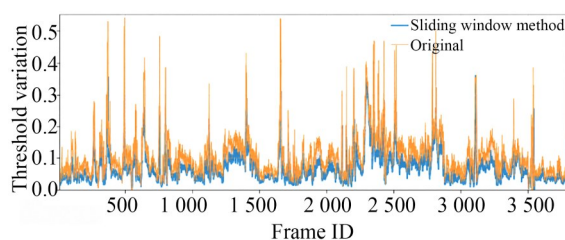
3.2 回环数量分析

首先分析了滑窗方法的优势,图 10 绘制了户外-隧道序列中平台 2 和 3 的候选帧筛选阈值 τ_{thres} 的变化曲线。可以看出,相比橙色线条对应的改进前方法,蓝色线条对应的滑窗方法更加平稳且数值更低,可以避免回环被漏检(彩图见期刊电子版)。其次,分别采用改进前、滑窗、滑窗+共视三种方法,图 11 展示了机器人内部和开源数据场景部分平台回环数量的变化。综合图和表可以得出:

(1)在多数平台中,相较于改进前方法,由于放宽了对重复纹理场景的判别限制,滑窗方法允许更多低分回环进入后续验证,从而检测到了更



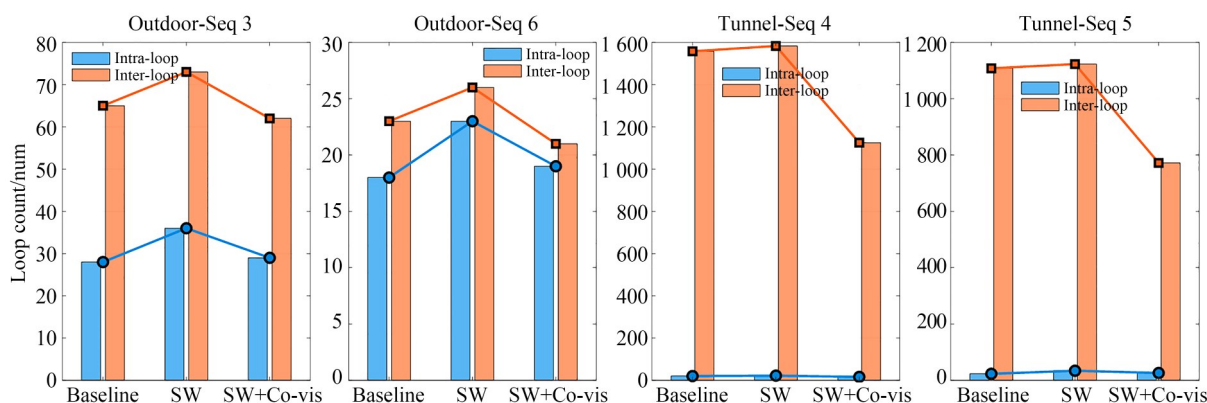
(a) 户外-隧道场景平台2
(a) Platform 2 in outdoor-tunnel scene



(b) 户外-隧道场景平台3
(b) Platform 3 in outdoor-tunnel scene

图 10 户外-隧道序列平台 2 和 3 的固定阈值与滑窗动态阈值变化曲线对比

Fig. 10 Comparison of fixed and sliding window dynamic threshold curves for platforms 2 and 3 in outdoor-tunnel scene



多候选回环,如对于自采地下数据中的平台 1,回环数量从改进前的 264 增加到 366 个。然而,该方法会引入部分由感知混淆引起的误匹配。

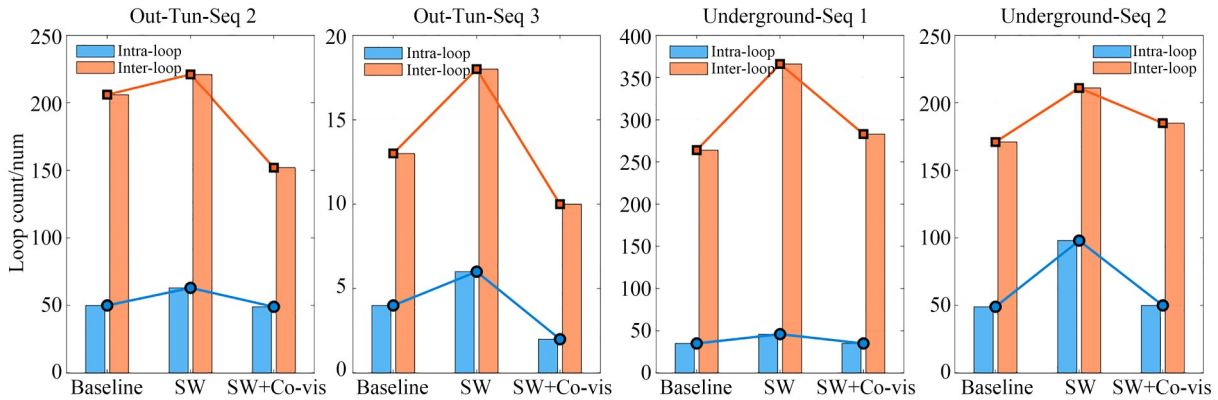


图 11 不同方法得到的回环数量对比

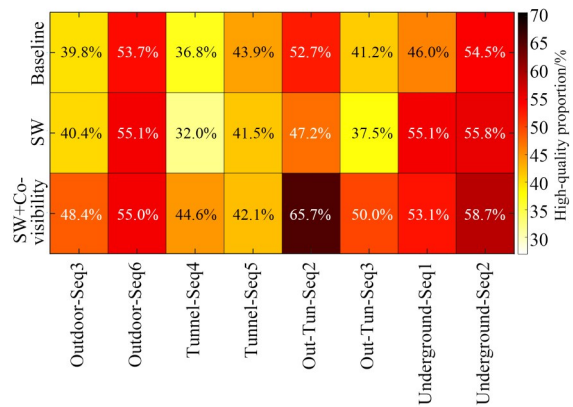
Fig. 11 Comparison of loop closure quantities using different methods

(2)通过进一步引入共视关系增强,滑窗+共视方法有效剔除了空间几何不一致导致的误匹配,使得回环数量有所下降,如开源隧道数据的平台 4,机器人回环数量由改进前及滑窗方法的 1 500 个以上,降低到 1 125 个,表明共视机制在抑制错误跨机关联方面具有显著效果。

3.3 回环质量分析

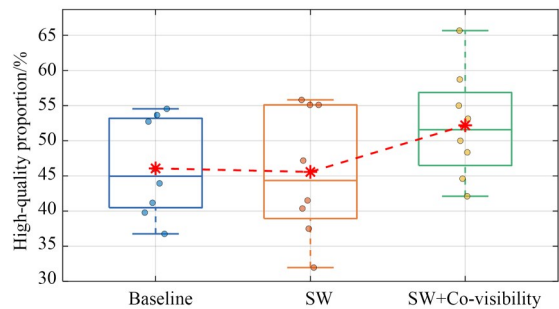
设 Good 和 Excellent 等级的回环为高质量回环,图 12(a)统计了开源和自采数据中部分平台的高质量回环占比,颜色越深,高质量回环占比越高;图 12(b)利用箱线图进一步展示了所有序列中高质量回环占比的统计分布,其中星号和虚线分别代表平均值和占比变化趋势。从图中可以得出:相比改进前方法和滑窗方法,户外、隧道、户外-隧道和地下等多种场景下,滑窗+共视方法对应区域的颜色更深,说明高质量回环的占比最高,证明了融合共视关系增强的几何校验能有效筛选出更高质量的回环。滑窗+共视方法得到的高质量回环占比的整体分布区间更靠上,且分布更为集中,表明该方法对回环质量的提升效果更稳定和显著。

表 4 进一步统计了不同回环质量的数量分布。其中,P,F,G 和 E 分别代表 Poor,Fair,Good 和 Excellent。由表可知:通过设定更加宽松合适的阈值,滑窗方法检测到了更多的回环信息,但由于感知混淆引起的误匹配,低质量、中等质量和高质量回环的比例分别提高了 1.425%、提高了 0.525%、下降了 2.05%。通过剔除低质量回



(a) 不同场景下高质量回环的比例

(a) Proportion of high-quality loop closures in different scenes



(b) 高质量回环统计

(b) High-quality loop closure statistics

图 12 三种方法得到的回环质量对比

Fig. 12 Comparison of loop closure quality obtained by three methods

环和保留高质量回环,相比改进前和滑窗方法,滑窗+共视方法得到的高质量回环的占比分别提高了 6.1% 和 8.15%,有助于得到更稳定的后端优化结果。

表 4 不同回环质量的数量分布
Tab. 4 Data distribution of different loop qualities

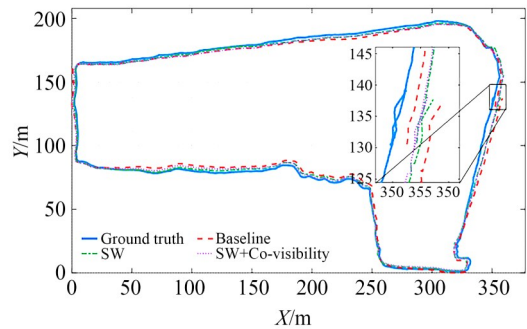
数据类型	平台	改进前方法 (P/F/G/E)	滑窗方法 (P/F/G/E)	滑窗+共视方法 (P/F/G/E)
户外序列	平台 3	26/30/19/18	24/41/24/20	20/27/24/20
	平台 6	8/11/13/9	10/12/13/14	8/10/12/10
隧道序列	平台 4	595/403/388/192	659/433/369/144	370/262/243/266
	平台 5	296/338/264/233	340/311/248/214	227/235/204/132
户外-隧道序列	平台 2	81/40/67/68	88/62/69/65	43/26/52/80
	平台 3	5/5/4/3	7/8/5/4	3/3/3/3
地下序列	平台 1	19/142/61/76	29/156/124/103	13/136/86/83
	平台 2	20/80/44/76	36/97/71/97	16/81/63/75

3.4 定位精度评估

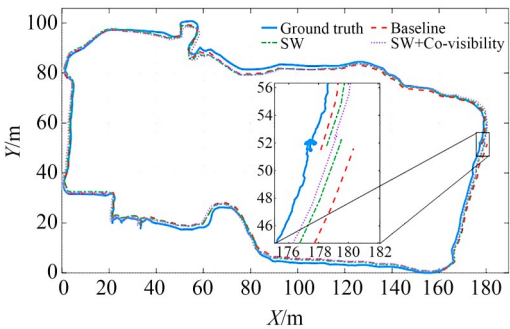
为充分评估本文方法,选取当前多机视觉SLAM中具有代表性的系统CCM-SLAM^[19]与COVINS-SLAM^[27]进行对比实验。然而,CCM-SLAM的设计限定最多接入3个平台,且其前端仅支持纯单目模式,在Kimera Multi Data提供的户外、隧道等大尺度、弱纹理序列中,前端频繁发生跟踪丢失,导致系统无法持续估计轨迹。COVINS-SLAM尽管允许双目-惯性组合,并以ORB-SLAM3为前端,具备更好的稳定性,但在上述数据集的长时间运行时,反复触发重定位,最终后端地图融合与位姿图优化模块无法收敛,未能输出有效的轨迹。鉴于此,后续轨迹精度对比主要围绕本文改进方法及原始Kimera-Multi系统展开。图13绘制了开源和自采数据的部分平台的轨迹,并放大了部分细节。

如图13(a)、13(b)和13(f)所示,改进前方法面临长轨迹情况下回环无法闭合的问题,虽然滑窗方法可以检测更多的回环信息,但仍然面临回环轨迹无法闭合问题;当加入共视关系增强的几何校验机制后,可以改善无法回环闭合的问题,图13(a)和(b)均成功检测到了正确回环。图13(f)中3种方法在轨迹末均没有成功回环,因场景切换导致无法获得足够多的特征匹配数量。

如图13(c)、13(d)和13(e)所示,光照变化和视角差异导致改进前方法轨迹明显偏移,且错误回环的持续影响难以通过个别正确回环消除;滑窗方法虽能检测更多回环,将轨迹进行纠正,但如前文所述,但因低质量回环混入,轨迹改善有限;滑窗+共视方法通过剔除低质量回环,可以改善回环的质量,进而提升轨迹精度,



(a) 户外序列平台3轨迹
(a) Trajectory of platform 3 on outdoor sequence



(b) 户外序列平台6轨迹
(b) Trajectory of platform 6 on outdoor sequence

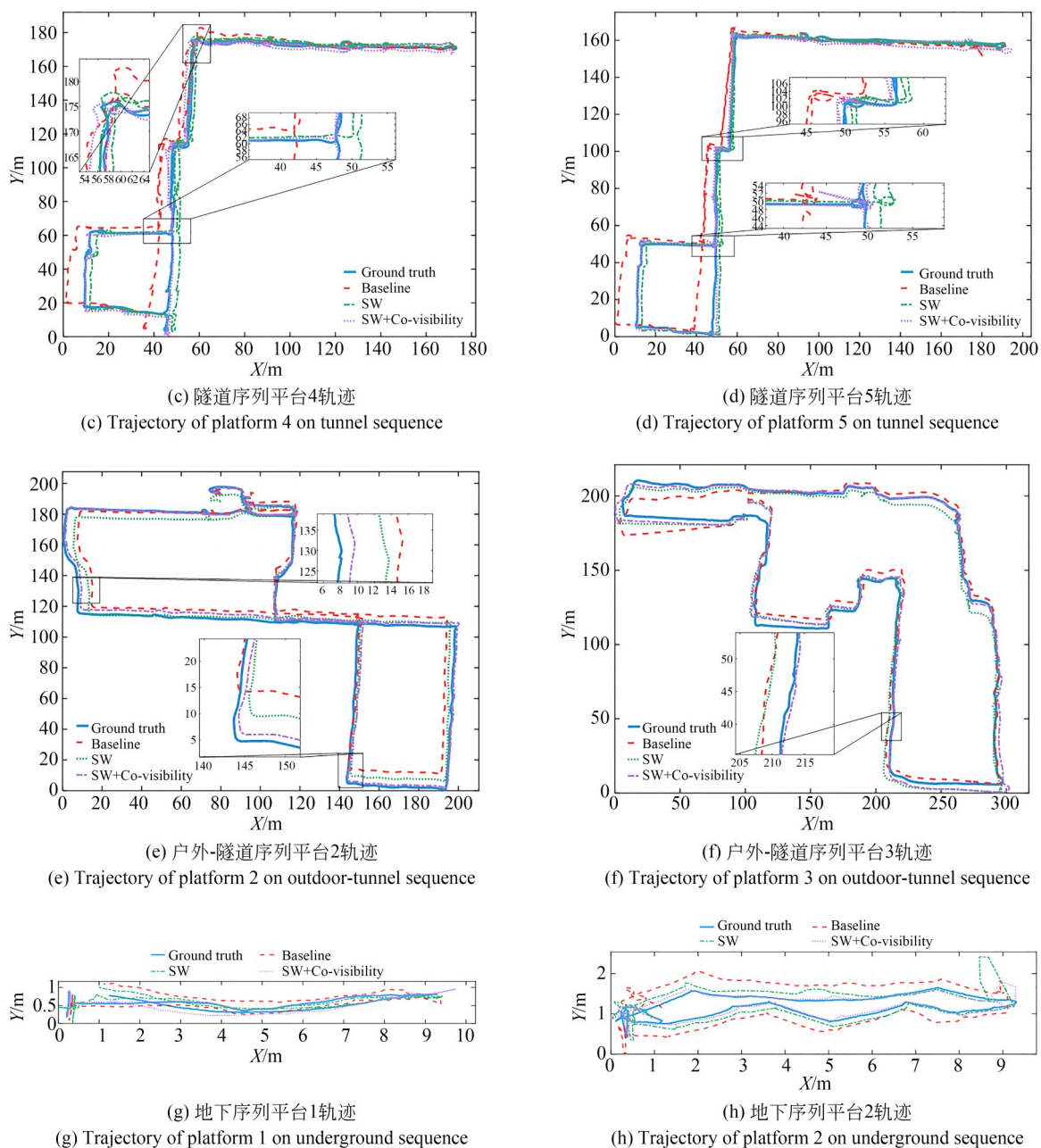


图 13 开源与自采数据部分平台轨迹

Fig. 13 Open source and self-collected data part of platform trajectory

尤其是在隧道和户外-隧道序列中这种低光照和弱纹理场景中。

如图 13(g)和 13(h)所示,在地下弱纹理场景中,改进前和滑窗方法与参考轨迹的一致性较差,而滑窗+共视关系方法能够有效地改善定位精度。

表 5 统计了各平台的 ATE,其中黑体表示最优,下划线次之。滑窗+共视方法在绝大多数序列的轨迹精度为最高,相比改进前方法,户外序列中平台 1、隧道序列中平台 2、户外-隧道序列中平台 3、地下数据中平台 1 和 2 的轨迹精度分别提升了 44.6%,25.8%,36.7%,63.7% 和 50.5%。

表 5 开源与自采数据的 ATE 统计结果

Tab. 5 Statistical results of ATE using public and self collected datasets

(m)

数据类型	平台	改进前方法	滑窗方法	滑窗+共视方法
户外序列	平台 1	8.774	6.590	4.855
	平台 2	4.329	3.530	2.797
	平台 3	3.620	2.499	2.486
	平台 4	5.757	8.323	3.696
	平台 5	5.387	5.843	3.977
	平台 6	1.621	1.512	1.492
隧道序列	平台 1	1.601	1.734	1.298
	平台 2	1.250	1.369	0.927
	平台 3	3.712	5.682	3.187
	平台 4	5.621	5.097	5.115
	平台 5	1.809	2.271	1.950
	平台 6	1.636	1.923	1.772
户外-隧道序列	平台 1	1.601	1.411	1.116
	平台 2	2.060	1.819	1.318
	平台 3	5.615	4.772	3.554
	平台 4	5.854	7.641	6.461
	平台 5	5.144	5.081	4.652
	平台 6	2.116	2.756	1.401
地下序列	平台 1	0.240	0.185	0.087
	平台 2	0.346	0.231	0.171
	平台 3	0.099	0.152	0.073

3.5 实时性分析

为分析本文方法的实时性,在本文实验平台与实现条件下对改进前后回环检测各模块的平均运行时间进行了统计,结果如表 6 所示。原始方法采用传统单帧几何验证,其主要时间开销集中在单帧几何验证阶段,平均耗时为 72.13 ms。本文方法在检测阶段引入了基于滑窗的自适应

阈值与时序一致性约束,但其额外开销较小。与原始方法不同,本文方法利用共视关系增强的几何验证替代原有单帧验证主流程,额外增加了 20.86 ms 耗时,成为验证阶段的主要计算开销来源。由此表明,本文方法在提升回环验证稳定性的同时,会引入可接受的额外时间开销,仍能够满足系统实时运行需求。

表 6 实时性分析结果

Tab. 6 Real-time analysis results

(ms)

方 法	BOW 查询	基于滑窗的自适应阈值	基于滑窗的时间一致性约束	单帧几何验证	共视关系增强的几何验证
改进前	0.65	0	0	72.13	0
改进后	0.68	0.02	0.05	0	92.99

4 结 论

在当前多机视觉 SLAM 回环检测中,多采用基于 DBOW2 的轻量化方法,但其易受光照和视角变化影响回环检测的准确性和稳定性。针

对此问题,提出一种多机视觉 SLAM 回环检测方法。首先,设计了基于滑窗统计特性的候选帧筛选,有效减少回环漏检;其次,对于具备时序连续性的机器人内部回环,融合时序一致性检验与共视几何校验,减少误匹配并且提升局

部轨迹精度;最后,利用跨机共视关系构建局部子图以扩展观测视野,有效克服了光照变化与视角差异问题,尤其是机器人回环。在开源数据集与自采数据集上的实验表明,所提出的方法能有效抑制错误回环,进而提高了回环检测的质量,最终显著降低了多机视觉 SLAM 的轨迹误差。

本文方法在很大程度上依赖于前端视觉里程计提供的共视地图点质量,且在不同场景存在差异较大的精度提升变化,未来将深入分析不同场景对定位精度造成的影响以及探索把语义信息融入回环模块,从而研究复杂场景下更加稳定

的回环检测。

作者贡献声明:

齐光峪:数据处理与实验,论文构思与撰写;
李林阳:论文构思与写作指导,实验指导与分析;
王安成:数据集采集,数据整理和分析;
李艺博:算法优化,论文审核与编辑;
郭文卓:方法论完善与理论分析,论文审核与编辑;
赵冬青:论文构思与写作指导,项目管理与组织协调。

参考文献:

- [1] EBADI K, BERNREITER L, BIGGIE H, *et al.* Present and future of SLAM in extreme environments: the DARPA SubT challenge [J]. *IEEE Transactions on Robotics*, 2024, 40: 936-959.
- [2] HE B, JI X X, LI G, *et al.* Key technologies and applications of UAVs in underground space: a review [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10 (3): 1026-1049.
- [3] WANG C Q, YU C, XU X, *et al.* Multi-robot system for cooperative exploration in unknown environments: a survey [EB/OL]. 2025: *arXiv*: 2503.07278. <https://arxiv.org/abs/2503.07278>.
- [4] 刘鑫,王忠,秦明星. 多机器人协同 SLAM 技术研究进展[J]. 计算机工程, 2022, 48(5): 1-10.
LIU X, WANG ZH, QIN M X. Research progress of multi-robot collaborative SLAM technology [J]. *Computer Engineering*, 2022, 48(5): 1-10. (in Chinese)
- [5] XU H, LIU P Z, CHEN X Y, *et al.* D²SLAM: decentralized and distributed collaborative visual-inertial SLAM system for aerial swarm [J]. *IEEE Transactions on Robotics*, 2024, 40: 3445-3464.
- [6] LAJOIE P Y, BELTRAME G. Swarm-SLAM: sparse decentralized collaborative simultaneous localization and mapping framework for multi-robot systems [J]. *IEEE Robotics and Automation Letters*, 2024, 9(1): 475-482.
- [7] GALVEZ-LÓPEZ D, TARDOS J D. Bags of binary words for fast place recognition in image sequences [J]. *IEEE Transactions on Robotics*, 2012, 28 (5): 1188-1197.
- [8] 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.
- [9] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras [J]. *IEEE Transactions on Robotics*, 2017, 33(5): 1255-1262.
- [10] 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.
- [11] ROSINOL A, ABATE M, CHANG Y, *et al.* Kimera: an open-source library for real-time metric-semantic localization and mapping [C]. 2020 *IEEE International Conference on Robotics and Automation (ICRA)*. May 31-August 31, 2020, Paris, France. IEEE, 2020: 1689-1696.
- [12] LOWRY S, SÜNDERHAUF N, NEWMAN P, *et al.* Visual place recognition: a survey [J]. *IEEE Transactions on Robotics*, 2016, 32(1): 1-19.
- [13] ARANDJELOVIĆ R, GRONAT P, TORII A, *et al.* NetVLAD: CNN architecture for weakly supervised place recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(6): 1437-1451.
- [14] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: self-supervised interest point detection and description [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recog-*

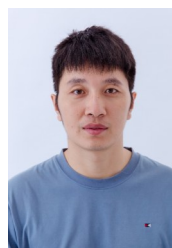
- nitition Workshops (CVPRW). June 18-22, 2018, Salt Lake City, UT, USA. IEEE, 2018: 337-33712.
- [15] HAUSLER S, GARG S, XU M, *et al.* Patch-NetVLAD: multi-scale fusion of locally-global descriptors for place recognition [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 14136-14147.
- [16] CHEN C H, WANG B, LU C X, *et al.* Deep learning for visual localization and mapping: a survey [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35 (12): 17000-17020.
- [17] ARSHAD S, KIM G W. SLGD-loop: a semantic local and global descriptor-based loop closure detection for long-term autonomy [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(12): 19714-19728.
- [18] ZHANG H C, ZHAO Y X, XU J M. Communication-efficient data exchange for decentralized loop-closure detection in collaborative SLAM [C]. 2025 American Control Conference (ACC). July 8-10, 2025, Denver, CO, USA. IEEE, 2025: 1185-1191.
- [19] SCHMUCK P, CHLI M. CCM-SLAM: Robust and efficient centralized collaborative monocular simultaneous localization and mapping for robotic teams [J]. *Journal of Field Robotics*, 2019, 36 (4): 763-781.
- [20] CIESLEWSKI T, CHOUDHARY S, SCARAMUZZA D. Data-efficient decentralized visual SLAM [C]. 2018 IEEE International Conference on Robotics and Automation (ICRA). May 21-25, 2018, Brisbane, QLD, Australia. IEEE, 2018: 2466-2473.
- [21] CHANG Y, TIAN Y L, HOW J P, *et al.* Kimera-multi: a system for distributed multi-robot metric-semantic simultaneous localization and mapping [C]. 2021 IEEE International Conference on Robotics and Automation (ICRA). May 30-June 5, 2021, Xi'an, China. IEEE, 2021: 11210-11218.
- [22] MANGELSON J G, DOMINIC D, EUSTICE R M, *et al.* Pairwise consistent measurement set maximization for robust multi-robot map merging [C]. 2018 IEEE International Conference on Robotics and Automation (ICRA). May 21-25, 2018, Brisbane, QLD, Australia. IEEE, 2018: 2916-2923.
- [23] YANG H, ANTONANTE P, TZOUMAS V, *et al.* Graduated non-convexity for robust spatial perception: from non-minimal solvers to global outlier rejection [J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 1127-1134.
- [24] LAJOIE P Y, RAMTOULA B, CHANG Y, *et al.* DOOR-SLAM: distributed, online, and outlier resilient SLAM for robotic teams [J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 1656-1663.
- [25] NISTER D. An efficient solution to the five-point relative pose problem [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(6): 756-770.
- [26] ARUN K S, HUANG T S, BLOSTEIN S D. Least-squares fitting of two 3-D point sets [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987, PAMI-9(5): 698-700.
- [27] SCHMUCK P, ZIEGLER T, KARRER M, *et al.* COVINS: visual-inertial SLAM for centralized collaboration [C]. 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). October 4-8, 2021, Bari, Italy. IEEE, 2021: 171-176.

作者简介:



齐光峪(2001—),男,河南辉县人,硕士研究生,主要从事多机器人协同同步定位与建图。E-mail:19715580939@163.com

通讯作者:



李林阳(1991—),男,河南确山人,博士,副教授,主要从事导航定位与位置服务。E-mail:linyangli@whu.edu.cn