arXiv:2603.11085v1 [cs.RO] 11 Mar 2026

通过高效通信实现边缘辅助多机器人视觉惯性同步定位与地图构建

Xin Liu，*IEEE* 学生会员，Shuhuan Wen，*IEEE* 高级会员，Jing Zhao、Tony Z. Qiu，*IEEE* 会员，Hong Zhang，*IEEE* 会士，

摘要——**云计算与边缘计算的融合是实现全局一致、实时多机器人同步定位与建图（SLAM）的有效途径。云计算有效解决了终端设备计算能力、通信能力和存储能力有限的问题。然而，带宽有限以及终端设备与云端之间通信链路极长，导致多机器人同步定位与地图构建系统的性能严重下降。为降低特征跟踪的计算成本并提升机器人的实时性能，本文提出一种基于金字塔式惯性测量单元预测的轻量级光流跟踪SLAM方法。在此基础上，提出一种基于机器人-边缘-云分层架构的集中式多机器人SLAM系统，以实现实时协同SLAM。该系统避免了机载计算资源有限和单机器人执行效率低下的问题。在这一架构中，仅传输特征点和关键帧描述子，并进行无损编码和压缩，从而在带宽资源有限的情况下实现实时远程信息传输。这一设计降低了数据传输过程中的实际带宽占用，且不会因数据压缩而造成SLAM精度损失。通过在EuRoC 数据集上的实验验证，与当前最先进的局部特征压缩方法相比，我们的方法能够以更低的数据量传输特征；与当前先进的集中式多机器人SLAM方案相比，在计算负载较低的情况下，我们的方法能够达到相同或更高的定位精度。**

致实践者：**本文旨在通过分别压缩并传输关键帧和非关键帧，降低云-边缘-机器人系统的通信负载。该方法适用于多机器人同步定位与建图系统，能够在高效通信的条件下实现多机器人联合定位和稀疏地图重建。目前，远程同步定位与地图构建或集中式多机器人同步定位与建图通常通过传输完整图像或图像的特征和描述子来实现。本文采用基于金字塔式惯性测量单元预测的轻量级同步定位与建图光流跟踪来跟踪非关键帧。在边缘服务器上，仅需传输关键点即可实现非关键帧之间的跟踪。对于关键帧，则通过传输压缩后的特征和描述子。通过关键帧特征信息在云端实现多机器人定位和地图融合。在公开数据集上的实验表明，该方法具有可行性，并且能够以较少的传输数据实现高精度联合定位。未来的研究中，我们将把这一框架应用于更多现实世界系统，同时利用更先进的特征实现丰富、准确的地图融合。**

索引词—**云边机器人协同、多机器人SLAM、特征压缩**

I. 引言

WITH 随着物联网（IoT）和智能系统的发展，复杂大规模场景的协同自主感知与定位已成为总体趋势。近年来，作为机器人智能与自主性的关键技术，同步定位与建图（SLAM）在全球蓬勃发展。研究人员正致力于将 SLAM 技术应用于现实场景，例如移动机器人[1]，[2]，车辆 [3]，和水下机器人 [4]。物联网将各种传感器设备与互联网相结合，形成一个能够随时随地广泛连接人与机器及物体的网络 [5]。边缘计算能够提升物联网的智能化水平，并广泛应用于物联网的诸多领域。它尤其适用于具有特殊服务需求的应用场景，例如低时延、高带宽、高可靠性、海量连接、异构聚合、本地安全和隐私保护。物联网边缘 [6] 是一组部署并运行于边缘节点设备上的智能系统。它将云计算能力扩展至边缘节点，并在现场执行智能计算，是分布式自主运行和自动控制的重要支撑。

通常，同步定位与建图主要使用两类传感器实现：相机和激光雷达。与使用激光雷达的同步定位与建图技术相比，视觉SLAM中使用的相机（尤其是单目相机）重量轻、成本低，且包含丰富的视觉信息，因此适用于成本和载荷受限的平台。然而，由于单目相机难以从单张图像中直接获取深度，因此通常无法观测真实场景的尺度。为解决这一问题，单目视觉惯性里程计可通过对齐视觉信息和惯性里程计测量来获取真实的尺度信息。事实上，单目视觉惯性里程计是能够满足轻量化和真实环境感知要求的最小组合套件 [7]。

尽管近年来同步定位与建图一直在发展，但单个移动机器人仍难以完成繁琐的任务

（通讯作者：*Shuhuan Wen*）电子邮箱：swen@ysu.edu.cn

This work was supported by the National Natural Science Foundation of China (Grant No. 62308130063), the Hebei Natural Science Foundation (Grant No. F2023203003), the Hebei Province Innovation Ability Improvement Plan Project (22567619H), the National Natural Science Foundation of China (Grant No. 52172332), and the Key Research and Development Project of Hainan Province (ZDYF2021GXJS015).

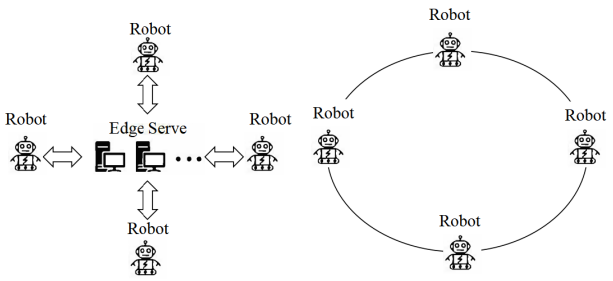
中国国家留学基金管理委员会（编号：202308130063）、河北省创新能力提升计划项目（22567619H）、国家自然科学基金委员会（编号：52172332）以及海南省重点研发项目（ZDYF2021GXJS015）。

Xin Liu、Shuhuan Wen 和 Jing Zhao 任职于河北省工业计算机控制工程重点实验室、教育部智能控制系统与智能装备工程研究中心、燕山大学，中国秦皇岛 066004。

Tony Z. Qiu 任职于中国武汉的武汉理工大学智能交通系统研究中心，以及加拿大阿尔伯塔大学土木与环境工程系。

Hong Zhang 任职于中国深圳的南方科技大学电子与电气工程系。

Manuscript received April 19, 2021; revised August 16, 2021年。



(a) 集中式架构 (b) 去中心化架构

图 1. 多机器人协同SLAM通信架构。

在复杂的大型场景中。多机器人系统通过通信、协作和信息共享来提升单个机器人的性能，例如任务执行效率、鲁棒性、灵活性和容错能力。由于单机器人同步定位与建图在任务上的局限性，以及多机器人系统强大的灵活性和适应性，多机器人协同SLAM受到了越来越多的关注。如图1所示，在多机器人协同SLAM中，构建局部地图时，单个机器人需要整合所有可用数据，以辅助其在全局环境中进行精确定位，并通过多个机器人系统之间的通信进行数据关联，从而协调并构建一致的全局地图。架构选择与任务分配、相对位姿确定、通信与数据关联、地图融合以及后端优化是多机器人SLAM的关键问题。借助物联网的智能边缘计算，机器人可以轻松接入边缘设备，并快速完成同步定位与建图云端上的工作以及传感器数据的传输。

图 1(b) 所示的分布式多机器人系统通常可以轻松添加更多机器人，同时每个机器人都可以并行计算，从而加快整个系统的建图和定位速度，例如 [8], [9]。然而，分布式多机器人SLAM系统需要大量数据共享和通信，因此对通信效率和网络负载有很高的要求。如图 1(a) 所示，集中式多机器人系统凭借其集中式架构，可以确保数据的全局一致性，这对于同步定位与建图系统至关重要，例如 [10], [11]。与此同时，中心节点可以通过任务分配平衡计算负载，大幅提升整个系统的效率。无论是分布式系统还是集中式系统，通信效率和数据量都可能造成性能瓶颈，影响整个系统的实时性和效率 [12]–[14]。

针对上述所有问题，我们提出了一种快速、轻量级、实时的多机器人单目视觉惯性SLAM系统。总而言之，我们的主要贡献如下：

1) 我们提出了一种基于 ORB-SLAM3 [15] 和 Lucas-Kanade（LK）稀疏光流法 [16] 的轻量级同步定位与建图方法。在该方法中，仅在关键帧中计算描述子；非关键帧之间的位姿估计则通过基于惯性测量单元预测的图像金字塔稀疏光流跟踪方法实现。这一设计大幅减少了特征跟踪过程中的计算开销，同时不会影响后端优化计算的精度。

2）基于上述快速SLAM，我们提出了一种基于机器人-边缘-云分层架构的集中式多机器人SLAM系统，以实现实时协同SLAM。我们对经典SLAM框架进行解耦，并将其分为特征提取、视觉惯性里程计（VIO）和全局优化三个层级。该设计避免了机器人机载计算资源有限、单个机器人执行效率低下的问题。

3）通过对机器人-边缘-云架构中的角点（非关键帧）和局部二进制描述符（关键帧）进行无损编码与压缩，我们能够在带宽资源有限的情况下实现实时远程信息传输。该设计降低了数据传输过程中实际占用的带宽，且不会因数据压缩而损失SLAM精度。此外，为了保证优化效率并减少存储开销，我们在方法中引入了地图骨干剖析（MBP）。

II. 相关工作

本节将从以下三个方面讨论相关工作：1）视觉惯性同步定位与建图（VINS）；2）多机器人SLAM；3）边缘辅助同步定位与建图。

A. 视觉惯性同步定位与建图

对于多机器人同步定位与地图构建系统，每个机器人都是独立移动的个体，因此单个机器人的同步定位与建图精度同样重要。经典的视觉SLAM方法主要分为基于直接法的同步定位与建图 [17]–[18]和基于特征点法的同步定位与建图 [19]–[21]。近年来，纯视觉SLAM方法发展迅速，此处不再赘述，而将重点放在VINS上。根据SLAM的优化方法，VINS/VIO可分为两类：一类是以扩展卡尔曼滤波器（EKF）为代表的滤波法，另一类是以图优化为代表的非线性优化法。

在计算资源有限的情况下，滤波法更为有效。Mourikis 提出的多状态约束卡尔曼滤波器（MSCKF）等 [22] 是最早的 VINS 系统之一。MSCKF 在 EKF 框架下融合惯性测量单元和视觉信息。与简单的视觉里程计算法相比，MSCKF 具有更高的鲁棒性。Li等 [23], [24]在 EKF 的基础上进一步确保所建立线性系统的正确可观测性，有效提高了大场景下的定位精度。在此基础上，[22], Sun等 [25] 进一步将 MSCKF 扩展到双目相机，提高了算法的鲁棒性，并使场景的尺度能够得到测量。ROVIO [26] 将像素块的光度误差转换为协方差信息，并在 EKF 中更新滤波状态。基于此， [26], Michael等 [27] 通过将 EKF 改进为 IEKF，进一步提高了算法的鲁棒性。在最新的

研究中，基于 EKF 的同步定位与建图方法仍广泛应用于资源受限场景，尤其是在基于雷达传感器进行感知的方案中 [28], [29], [30]。然而，其主要缺点之一是存储能力与状态数量呈平方增长关系，不适用于大规模场景。在视觉SLAM中，路标数量众多，因此滤波法效率较低。然而，基于图的优化同步定位与建图通常基于关键帧（KF），并使用光束法平差（BA） [31] 建立节点之间的相对变换关系，同时保证精度和实时性。

Leutenegger 等。 [32] 该方法通过维护一个滑动窗口，对最近一段时间内关键帧的状态进行优化，并进一步联合优化帧与惯性测量单元。结果表明，所提出的方法比松耦合和基于滤波的SLAM具有更高的精度。Qin 等[33] 提出的 VINS-Mono 是一种功能强大且通用的单目视觉惯性状态估计器。Qin 等 [34] 还提出了双目VINS，采用非线性优化法融合双目相机与惯性测量单元，显著提升了精度。VI ORB-SLAM [35]在 [20]‚[21], 的基础上加入 IMU 信息以辅助视觉测量，并采用 ORB（定向FAST与旋转 BRIEF） [36]特征进行特征提取。基于 [35], [37],, Campos等 [15] 提出的 ORB-SLAM3 是当前最先进的方法之一。该方法可在针孔或鱼眼模型下使用单目、双目和 RGB-D 摄像头稳定运行，仅依靠最大后验估计（初始化时包括惯性测量单元），并可在室内或室外运行。因此，我们采用 ORB-SLAM3 作为单机器人SLAM方案。

B. 多机器人*SLAM*

多机器人的全局感知可以通过地图复用技术实现，例如 [15], [33]。然而，地图复用需要预先构建地图，而这并非总能实现。此外，根据通信架构，多机器人SLAM可分为集中式架构和去中心化架构。

DDF-SLAM [38] 是最早用于去中心化多机器人SLAM的解决方案。在此基础上，Cunningham 等人 [39] 将局部地图与相邻地图融合，以避免 [38]中的重复计算问题。Choudhary 等人 [40], [41]采用分布式高斯-赛德尔（DGS）算法解决了多机器人协同SLAM问题。该方法所需的信息交换较少，且易于实现，因此可以很好地扩展到大型团队。在此基础上，Cieslewski 等人 [42] 将DGS算法整合到一个完整的视觉SLAM系统中。 [42] 通过预训练的NetVLAD模型 [43]实现机器人间的相对识别。尽管NetVLAD能够实现较高的场景识别精度，但该模型较为复杂，难以满足实时性要求。Lajoie[44], 、Tian [8] 和 Huang [9]等人通过位姿图优化（PGO）优化机器人的相对位姿，实现了比DGS更高的精度。分布式多机器人PGO可以部署在多个机器人上，以加快整个系统的建图和定位速度，但这可能会因机器人之间的不同步而导致数据异步问题，从而影响建图和定位的精度。同时，分布式并行计算会增加网络负载。总体而言，去中心化机器人具有很强的可扩展性，但在保证数据一致性和避免重复计算方面面临重大挑战。大多数协同视觉同步定位与建图系统采用集中式架构。CoSLAM [45] 是一种强大的集中式SLAM系统，仅依赖视觉信息。该系统根据帧的重叠情况进行分组，以处理动态场景。Forster 等人 [46] 提出了一种基于描述子的集中式SLAM系统，可通过匹配描述子完成地图融合。Bonin-Front等人 [47] 针对视角变化导致描述子难以匹配的问题，提出了一种基于会面的集中式系统。全球定位系统等全局信息也可用于估计多机器人的相对位置，但在无法使用全球定位系统的场景中，外部传感器的引入受到很大限制 [48]。CCM-SLAM [49] 提出通过服务器与机器人之间的双向通信，有效利用服务器协助完成高成本任务。Wen 等人 [50] 通过减少传输数据量，提高了CCM-SLAM在低带宽下的稳定性，但与当前先进的SLAM系统相比，它仍缺乏稳定的协同定位能力和更强的客户端可扩展性。CVI-SLAM [10], 、COVINS [11] 和 CVIDS [51]通过融合IMU信息增强了系统的协同定位效果，但这些系统均建立在机器人与服务器之间通信稳定的前提下。适用于不稳定通信环境的集中式SLAM系统仍有待研究。

C. 边缘辅助同步定位与建图

边缘计算 [52] 将工作负载从云端转移到与终端设备物理距离较近的服务器上。近年来，研究人员探索了利用边缘计算辅助同步定位与建图计算。Xu 等人 [53] 提出了一种边缘辅助移动语义视觉同步定位与建图框架，可实时执行视觉SLAM。Edge-SLAM [54] 从模块化视角分离同步定位与建图。然而，这些方法仅涉及单个机器人和一个边缘节点，无法直接扩展到多机器人与多个边缘节点之间的通信。Huang 等人 [55]通过向配备二维激光雷达的多台机器人分配服务器资源，优化边缘计算环境，从而加速同步定位与建图的计算过程。Lan 等人 [56] 将不同线程卸载到异构边缘服务器，以加速同步定位与建图过程，实现实时运行，但这对同步定位与建图系统的结构和线程架构有诸多限制。尽管这些系统通过边缘计算加速了同步定位与建图，但大多数系统聚焦于单机器人场景。同时，边缘同步定位与建图的通信带宽和通信稳定性仍有研究空间。

边缘SLAM中的传感器信息有两种不同的处理方法：先压缩后分析（CTA），即将视觉信息以压缩视频序列的形式传输。先分析后压缩（ATC），其中

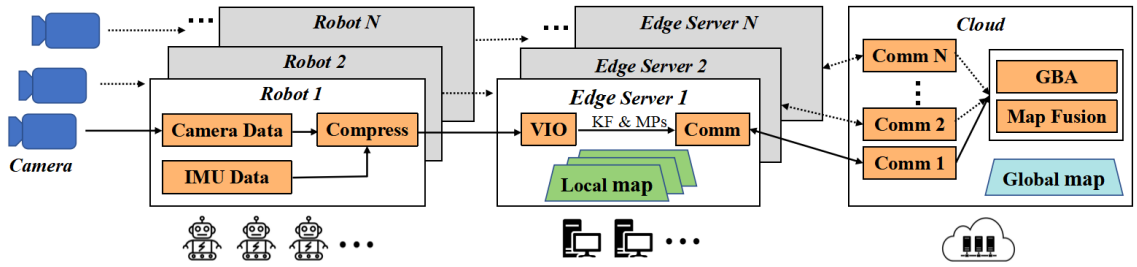


图2. 系统概述。每个机器人都配备了视觉惯性传感器套件。分析传感器信息后，压缩数据将传输至固定的边缘服务器。边缘服务器可通过实时、独立地运行视觉惯性里程计，在局部场景中进行定位和地图构建。通信接口用于在边缘服务器与云端之间交换数据（关键帧局部二进制特征和地图点）。云端可执行计算开销较大、全局且非实时的任务：冗余关键帧剔除、回环检测和全局地图融合。

视觉线索被提取、压缩，并以局部特征的形式传输。例如，Wang 等人 [57]构建了一个远程同步定位与地图构建系统来传输完整图像。然而，传输完整图像会占用更高的带宽。Van 等人 [58], [59] 通过分析图像后再压缩传输来降低信息带宽。但在间接法中评估关键点描述子通常非常耗时，并且除了使用关键帧 [60]之外，这些描述子不会被复用。非关键帧描述子的计算可能会导致边缘SLAM系统出现时延。

III. 系统概述

所提出系统的架构如图2所示。我们提出了一种用于集中式视觉惯性里程计的新架构。该架构主要分为三层：机器人端、边缘端和云端。该架构最初在[54] 中提出，用于运行配备二维激光雷达传感器的多机器人SLAM。由于实时高分辨率图像数据往往占用大量内存，在实际的多机器人应用中，数据传输和分析面临困难。在本工作中，我们通过分析并压缩传感器数据，将该架构扩展至可使用视觉惯性传感器的场景。在机器人端，我们针对非关键帧采用基于IMU辅助LK方法的快速特征检测与特征匹配（见第4节），并提取关键帧的ORB特征。在边缘端运行视觉惯性里程计前端，并通过非关键帧中的光流跟踪匹配结果求解当前帧的位姿。对于关键帧，当光流跟踪失败时，则采用ORB-SLAM3中的跟踪方法进行定位。同时，边缘端为每个智能体维护一张大小有限的局部地图，以确保每个智能体具备基本的自主性。此外，全局地图维护和计算密集型流程（位姿图优化和全局捆绑调整）被迁移至云端（见第6节）。我们在机器人端、边缘端和云端之间采用通用通信接口，实现多个智能体之间的协作。为减少传输数据量，我们通过后分析压缩分别对关键帧和非关键帧进行编码（见第5节）。

IV. IMU辅助非关键帧特征匹配

对于连续两帧之间的特征匹配，一种方法是计算描述子，以匹配相邻图像之间的关键点，例如 [19]。然而，描述子计算和匹配的计算成本较高。实际应用中，CPU 无法实时提取尺度不变特征变换（SIFT），而 ORB 特征的提取耗时近 20 毫秒，这对机载计算机的算力提出了挑战。另一种方法是采用直接法，通过最小化光度误差来完成匹配。利用最小化光度误差进行特征匹配的方法仅适用于相似的帧间匹配 [61]。因此，这种方法通常不包含回环检测。为了完成多机器人之间的回环检测和地图融合，同时提高计算效率，我们的方法在非关键帧之间进行 LK稀疏光流跟踪，以完成帧间跟踪，而描述子计算仅在关键帧上进行。

稀疏光流法可以快速估计相邻帧之间的对应关系，但它基于以下三个假设：1）相邻帧之间的亮度恒定；2）物体在相邻帧之间的运动幅度较小；3）相邻像素具有相同的运动。第 i帧和第 j帧的恒定亮度假设可表示为：

$$\mathrm{I}(u_1, v_1, i) = \mathrm{I}(u_2, v_2, j), \tag{1}$$

其中，$(u_1,\ v_1)$ 和 $(u_2,\ v_2)$ 分别是第 i帧和第 j帧中的匹配点，I 是对应点的灰度值。

在真实场景中，相应点的灰度值会随角度和曝光程度的不同而变化。为减少恒定光度假设带来的影响，FastORB-SLAM [60] 通过建立匀加速运动（UAM）模型提高匹配精度。UAM模型可以利用相机等外部传感器的信息预测关键点的位置。在静态场景中，与匀速运动模型（UMM）相比， [21] 能够提高匹配精度。然而，在动态场景中，外部传感器容易导致不可靠的估计，而且由相机计算的UAW模型的精度也会受到影响。如图3所示，我们将外部传感器（惯性测量单元）引入系统，通过对惯性测量单元进行预积分，可以快速、准确地预测位姿。同时，内部传感器不受动态物体的影响，因此能够提高系统的鲁棒性。

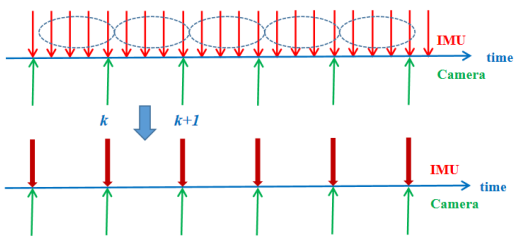


图 3. IMU预积分示意图。对第 k 帧与第 $(k+1)$ 帧之间的所有 IMU 信息进行积分，可得到第 $(k+1)$ 帧的位置、速度和姿态（PVQ），作为视觉估计的初始值。

在惯性测量单元初始化完成之前，我们使用匀速运动模型，根据过去连续的惯性测量单元信息估计当前帧的位姿。与 [21],不同，我们通过计算惯性测量单元坐标系中的平均角速度，来计算匀速运动模型中的速度。对于 i帧至j帧之间的 K 组速度数据，可以通过以下方式计算平均角速度:

$$\mathbf{v}_{mean} = \frac{\sum_{k=1}^{K} \mathbf{v}_i}{t_i - t_j}, s.t t_i < t_k < t_j, \tag{2}$$

其中， t_i 和 t_j 分别是第 i 帧和第 j 帧的时间戳， t_k 是 IMU 的时间戳。这里，由于短时间内平移向量可以忽略，因此我们只考虑旋转角速度。IMU 初始化后，通过 IMU预积分估计当前帧的位姿。IMU 的初始化通过最大后验估计（MPA） [15] 得到。

考虑由三轴加速度计和三轴陀螺仪组成的 IMU，其测得的加速度 $\hat{\omega}_B$ 和角速度 $\hat{\mathrm{a}}_B$ 会受到白噪声 η 和传感器偏置 b 的影响。其模型可描述为:

$$\begin{aligned} \hat{\omega}_B(t) &= \omega_B(t) + \mathrm{b}^g(t) + \eta^g(t) \\ \hat{\mathrm{a}}_B(t) &= \mathrm{R}_B^W(\mathrm{a}_B(t) - g_W) + \mathrm{b}^a(t) + \eta^a(t) \end{aligned}, \tag{3}$$

其中，$(\cdot)_B$ 和$(\cdot)_W$ 分别表示机体坐标系和世界坐标系。 ω_B 和a_B 分别表示世界坐标系中的真实角速度和加速度。g_W 表示世界坐标系中的重力加速度。 η^g 和 η^a 分别表示均值为0的陀螺仪噪声和加速度计噪声。b^g和b^a 分别表示陀螺仪偏置和加速度计偏置。它们的随机游走误差通常采用维纳过程建模。在流形 [62],上进行IMU预积分。第 i帧和第j帧之间的旋转增量 $\Delta\mathrm{R}_{ij}$、速度增量$\Delta\tilde{\mathrm{v}}_{ij}$ 和位置增量 $\Delta\tilde{\mathrm{p}}_{ij}$ 可表示为:

$$\begin{aligned} \Delta\tilde{\mathrm{R}}_{ij} &= \prod_{k=i}^{j-1} \mathrm{Exp}(\hat{\omega}(t_k) - \mathrm{b}^g(t_k) - \eta^g(t_k)\Delta t_{ij}) \\ \Delta\tilde{\mathrm{v}}_{ij} &= \sum_{k=i}^{j-1} \Delta\mathrm{R}_{ik}(\hat{a}(t_k) - \mathrm{b}^a(t_k) - \eta^a(t_k))\Delta t_{ij} \\ \Delta\tilde{\mathrm{p}}_{ij} &= \sum_{k=i}^{j-1} \left[\Delta\mathrm{v}_{ik}\Delta t_{ij} + \Delta\mathrm{R}_{ik}(\hat{a}(t_k) - \mathrm{b}^a(t_k) - \eta^a(t_k))\Delta t_{ij}^2\right] \end{aligned} \tag{4}$$

借助惯性测量单元辅助的位姿估计，我们可以准确预测关键点的位置，并在非关键帧之间完成稀疏光流跟踪，而无需计算描述子。完整步骤总结于算法1中。

算法1IMU辅助LK跟踪

输入： 参考帧：I_r，当前帧：I_c，其中的关键点 I_r： kps_r，IMU观测值： gyr, acc;
图像金字塔的尺度比例为 $s = 1.2$，层数为 $n = 8$ 用于描述 I_r 和 I_r;
如果 惯性测量单元尚未初始化，**则**
  $\longleftarrow$ I_r 与 I_c 之间的平均速度为 $\mathbf{v}_{mean}$ 由公式 (2) 估算;
  T_{rc} $\longleftarrow$ 从 I 得到的变换矩阵 $_r$ to I_c is 由 UMM 估计;
  返回 T_{rc};
否则如果惯性测量单元已初始化**则**
  $\Delta\mathrm{R}$ $\Delta\mathrm{v}$ $\Delta\mathrm{p}$ $\longleftarrow$ 通过以下方式计算 IMU预积分公式 (4);
  T_{rc} $\longleftarrow$ 从 I 得到的变换矩阵 $_r$ to I_c is 由 Δ估计 $\tilde{\mathrm{R}}$ $\Delta\tilde{\mathrm{v}}$ $\Delta\tilde{\mathrm{p}}$;
  返回 T_{rc};
结束 if
kps_{pred} $\longleftarrow$ T_{rc} 用于预测像素坐标对应于 I kps_r 中的c;
pts_{L8} $\longleftarrow$ 根据 kps_{pred}，对在 I_r 和 I_c 之间进行顶层金字塔的 LK 跟踪;
pts_{L1} $\longleftarrow$ 沿图像金字塔向下迭代 pts_{L8}，直至到达第 1 层;
根据旋转积分筛选对应点 IMU;
kps_c, kps_r $\longleftarrow$ 使用 RANSAC 剔除离群点;
通过透视-n-点求解 I_c 在世界坐标系中的位姿矩阵;
确保： 最终匹配点 kps_c, kps_r，位姿矩阵 I_c 在世界坐标系中的位姿矩阵： T_{cw};

V. 通信方法

得益于 IMU辅助的非关键帧特征匹配方法（第三节），在我们提出的云边分层架构中，无需传输非关键帧描述子即可实现精确定位和稀疏地图构建。在此模式下，我们能够以最少的信息传输完成多机器人SLAM。

A. 特征编码压缩

基于 [58],中提出的基于特征压缩的远程同步定位与地图构建系统，我们改进了特征压缩方法，并构建了基于特征压缩的多机器人SLAM系统。如图4所示，为实现无损压缩，我们采用帧内编码模式，并进一步将帧内编码模式分为关键帧模式和非关键帧模式。在非关键帧模式下，我们仅对关键点进行编码和传输。在关键帧模式下，我们对与ORB特征对应的视觉词、ORB特征的关键点以及描述子进行压缩和编码

ORB特征的。因此，在第 i个非关键帧中对特征 n 进行编码的代价可表示为：

$$C_{i,n}^{non-kf} = C_{i,n}^{non-kf,kp}, \tag{5}$$

并且，第 i个关键帧中特征 n 的编码成本可表示为：

$$C_{i,n}^{kf} = C_{i,n}^{kf,kp} + C_{i,n}^{kf,BoW} + C_{i,n}^{kf,des}, \tag{6}$$

其中， $C_{i,n}^{non-kf,kp}$、 $C_{i,n}^{kf,kp}$分别表示非关键帧和关键帧关键点的编码成本， $C_{i,n}^{kf,BoW}$表示视觉词索引的编码成本， $C_{i,n}^{kf,des}$表示描述子的编码成本。

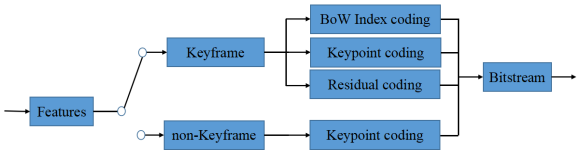


图4. 编码模式的选择。关键帧的判定取决于IMU辅助的LK光流的跟踪点数以及与前一关键帧的时间间隔。

1) 关键点的压缩编码：在非关键帧中，我们将关键点信息定义为仅需包含水平坐标、垂直坐标及其所在的图像金字塔层级。因此，非关键帧 i中的关键点n 可描述为：

$$kp_{i,n}^{non-kf} = [u, v, \sigma], \tag{7}$$

其中， u 和 v 分别表示图像坐标系中的水平坐标和垂直坐标， σ 表示关键点所在的图像金字塔层级。在关键帧中，为了使特征点对旋转具有鲁棒性，我们将关键点定义为具有以下 4 个属性的数组：

$$kp_{i,n}^{kf} = [u, v, \sigma, \theta], \tag{8}$$

其中， θ 表示通过灰度质心法计算得到的关键点方向。 N_θ 表示将 2π 弧度划分为 N_θ个部分，并按照逆时针方向从小到大为每个部分编号。因此，关键点的编码可以转换为 u、 v、 σ、 θ的编码。为了进一步降低带宽占用，我们使用图像金字塔坐标系中的水平坐标 u_σ 和垂直坐标 v_σ 来替代 u 和v。因此，在非关键帧 i上压缩关键点n 所需的比特数为：

$$C_{i,n}^{non-kf,kp} = \log_2(N_\sigma) + \log_2(H(\sigma)) + \log_2(W(\sigma)), \tag{9}$$

在关键帧 i 上压缩关键点 n 所需的比特数为：

$$C_{i,n}^{kf,kp} = \log_2(N_\sigma) + \log_2(N_\theta) + \log_2(H(\sigma)) + \log_2(W(\sigma)) \tag{10}$$

其中 N_σ 为图像金字塔的层数，赋值为8，以及 $N_\theta = 32$。 $H(\cdot)$和 $W(\cdot)$分别是对应图像金字塔的高度和宽度。

2) 视觉词索引与描述子的压缩编码：以ORB-SLAM3中计算得到的视觉词汇作为特征编码的参考，并将视觉词汇表作为机器人和边缘端均可访问的共享信息。对于图像 i中的特征 n ，在大小为 S的视觉词树中存在一个与 n 汉明距离最小的词 v。因此，我们对视觉词索引v进行编码，该部分占用的比特数为：

$$C_{i,n}^{kf,BoW} = \log_2(S), \tag{11}$$

为了无损表示特征描述子，我们进一步计算特征 i与最近词v之间的残差向量。在 [58], 中，对于 ORB 特征的描述子，残差向量通过异或运算计算得到。对于一个长度为D 、包含 h个非零元素的二进制描述子，对残差向量进行编码所需的比特数为：

$$C^{res}(h, p_0) = -(D-h)\cdot\log_2(p_0) - h\log_2(1-p_0), \tag{12}$$

其中， p_0 是二进制残差向量元素为 0 的概率。计算特征 n 的描述子与压缩图像帧 i 中最近词之间的残差向量所需的位数为：

$$C_{i,n}^{kf,res} = C^{res}(h_{i,n}^{kf}, p_0^{kf}), \tag{13}$$

B. 机器人-边缘-云通信

TCP 是一种可靠的协议，可确保数据在传输过程中不会丢失或损坏。TCP 通过建立、传输和释放连接来确保数据的完整性与可靠性。为避免系统通信中丢包，通信模块基于 TCP 协议进行套接字编程，并采用消息丢包重传机制进行通信 [50]。与其他方法相比，在套接字编程中，客户端创建连接套接字并向服务器发送请求，因此机器人、边缘服务器和云端之间的通信支持在任务执行过程中动态加入并发起通信请求。

1) 机器人-边缘：机器人与边缘端之间采用双向通信。我们采用 [50]中的高效通信模式，确保消息不会被重复发送。同时，检测到数据包丢失时，将重新发送数据。通过分别对关键帧和非关键帧进行数据压缩和编码，尽可能降低网络带宽占用。边缘服务器收到数据后，会对其进行解码，并利用这些数据维护本地视觉惯性里程计。

IMU 的初始化在 VIO 的局部建图线程中进行，从而获得更精确的惯性参数。为了减少离线标定参数的不准确性对机器人 IMU 预积分精度的影响。IMU 初始化完成后，边缘服务器会向机器人传输更优的惯性参数。

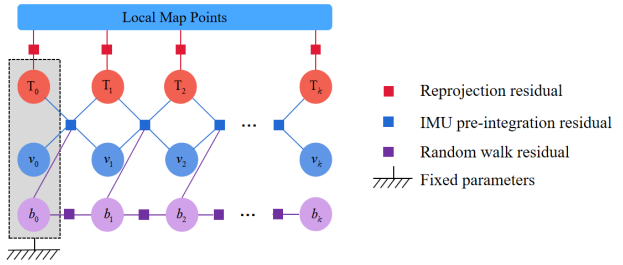
2) 边缘-云：为了将局部地图数据从边缘端共享到云服务器，并确保边缘服务器的局部里程计精度，我们建立了边缘-云双向通信。边缘服务器会将关键帧数据和其局部地图中地图信息的变化通知云服务器，并

持续发送关键帧和地图点的完整数据结构（关键帧数据包括从图像中提取的二维特征点压缩数据），这构成了边缘服务器与云服务器之间的主要通信流量。每个压缩的二维特征键约占 220 字节。其他占用流量较少的数据包括 IMU 相关信息（IMU测量以及当前偏置和速度估计值），以及用于确认关键帧和地图点消息更新的确认信息。云服务器收到这些消息后，通过全局图优化更新关键帧位姿和全局地图。全局图优化完成后，我们将更新后的关键帧位姿传回边缘端。由此可以量化当前姿态估计的漂移。

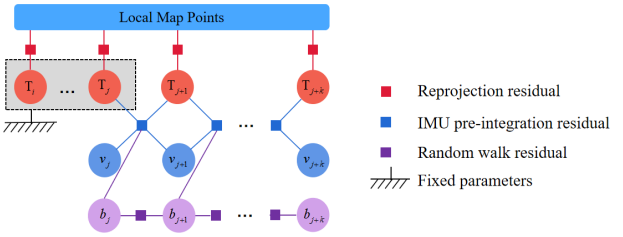
VI. 视觉惯性里程计与全局地图融合

A. 跟踪与局部建图

在初始化阶段，经过纯视觉优化估计和纯惯性优化估计后，初始地图点被准确表示，同时视觉参数信息和惯性参数信息也得到了相对准确的估计。在此基础上，如图5（a）所示，通过视觉惯性联合优化进一步精化这些参数，并且尺度作为显式变量迅速收敛。值得注意的是，此处我们固定了关键帧位姿与参数的初始化，仅优化尺度和重力信息。对于跟踪线程，采用第IV节所述方法进行帧间特征匹配。采用简单快速的滑动窗口优化来优化最新帧的位姿，而地图点保持不变。



（a）初始化过程中的视觉惯性联合优化因子图



(b) 局部建图线程中的视觉惯性联合优化因子图

图 5. 视觉惯性联合优化因子图。

所有关键帧都参与局部建图线程。随后，所提出的系统通过通信模块将生成的地图点、优化后的关键帧位姿和参数传输到云服务器。同时，

惯性测量单元初始化过程在局部建图线程中进行。在此过程中，基于公式（4）中的流形预积分，帧 i 与帧 j 之间惯性测量单元的旋转、速度和位置残差可表示为：

$$\begin{aligned} e_{imu}^{i,j} &= [e_{\Delta \mathrm{R}_{i,j}}^{i,j}, e_{\Delta \mathrm{v}_{ij}}^{i,j}, e_{\Delta \mathrm{p}_{ij}}^{i,j}] \\ e_{\Delta \mathrm{R}_{ij}}^{i,j} &= \mathrm{Log}\left[\Delta \mathrm{R}_{ij}^{\top} \mathrm{R}_i^{\top} \mathrm{R}_j\right] \\ e_{\Delta \mathrm{v}_{ij}}^{i,j} &= \mathrm{R}_i^{\top}(\mathrm{v}_j - \mathrm{v}_i - g_W \Delta t_{ij}) - \Delta \mathrm{v}_{ij} \\ e_{\Delta \mathrm{p}_{ij}}^{i,j} &= \mathrm{R}_i^{\top}(\mathrm{p}_j - \mathrm{p}_i - \mathrm{v}_i \Delta t_{ij} - \tfrac{1}{2} g_W \Delta {t_{ij}}^2) - \Delta \mathrm{p}_{ij} \end{aligned} , \tag{14}$$

其中，R_{ij}、v_{ij} 和p_{ij} 分别是惯性测量单元偏置更新后预积分的结果。其计算过程可表示为：

$$\begin{aligned} \Delta \mathrm{R}_{ij} &= \Delta \tilde{\mathrm{R}}_{ij} \cdot \mathrm{Exp}\left(\frac{\partial \Delta \tilde{\mathrm{R}}_{ij}}{\partial \mathrm{b}_i^g}\right) \\ \Delta \mathrm{v}_{ij} &= \Delta \tilde{\mathrm{v}}_{ij} + \frac{\partial \tilde{\mathrm{v}}_{ij}}{\partial \mathrm{b}_i^g} \partial \mathrm{b}_i^g + \frac{\partial \tilde{\mathrm{v}}_{ij}}{\partial \mathrm{b}_i^a} \partial \mathrm{b}_i^a \\ \Delta \mathrm{p}_{ij} &= \Delta \tilde{\mathrm{p}}_{ij} + \frac{\partial \tilde{\mathrm{p}}_{ij}}{\partial \mathrm{b}_i^g} \partial \mathrm{b}_i^g + \frac{\partial \tilde{\mathrm{p}}_{ij}}{\partial \mathrm{b}_i^a} \partial \mathrm{b}_i^a \end{aligned} . \tag{15}$$

帧 k 中地图点 l 的重投影误差可表示为：

$$e_{cam}^{l,k} = \mathrm{x}_k - \Pi_{cam}(\mathrm{R}_k \mathrm{X}^l + \mathrm{t}_k) \tag{16}$$

其中，$\Pi_{cam}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} f_x \frac{X}{Z} + c_x \\ f_y \frac{Y}{Z} + c_y \end{bmatrix}$ 是相机的投影函数，$\mathrm{X}^l \in \mathbb{R}^3$ 是地图点l 在世界坐标系中的三维坐标，$\mathrm{x}_k \in \mathbb{R}^2$ 是帧 k上对X^l 的观测点。如图 5 (b) 所示，局部BA优化问题的对应节点可表示为：

$$\begin{aligned} \chi &= [\chi_0, \chi_1, \chi_2, \cdots, \chi_k], \\ \mathcal{L} &= [\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_l], \end{aligned} \tag{17}$$

其中，$\chi_i = \{\mathrm{R_i}, \mathrm{t_i}, \mathrm{v_i}, \mathrm{b}^{\mathrm{g}}{}_{\mathrm{i}}, \mathrm{b}^{\mathrm{a}}_{\mathrm{i}}\}$ 表示与关键帧对应的状态向量。$\mathrm{b}^{\mathrm{g}}{}_{\mathrm{i}}$ 和$\mathrm{b}^{\mathrm{a}}_{\mathrm{i}}$ 分别表示时刻 i的陀螺仪偏置和加速度计偏置。用于连接相邻时刻的随机游走误差可用维纳过程建模。 $\mathcal{L}_i$ 是共视地图点的坐标集合。

相应局部BA优化问题中的边包括共视地图点的重投影残差（表示为 e_{cam}）以及关键帧之间的IMU预积分残差（表示为 e_{imu}）。因此，整体损失函数可表示为：

$$\min_{\chi,\mathcal{L}} \{ \sum_{j \in C} \rho(\left\| e_{imu}^{j,j+1} \right\|^2) + \sum_{l \in L} \sum_{k \in C} \delta(\left\| e_{cam}^{l,k} \right\|^2) \} \tag{18}$$

其中，C 表示局部BA窗口中新加入的关键帧集合，L 表示关键帧之间具有共视关系的地图点。为保持结果的鲁棒性，我们加入柯西损失函数 $\delta(\cdot)$ 和Huber损失函数 $\rho(\cdot)$，以提高损失函数的鲁棒性。

生成新的关键帧后执行局部建图，但求解式（18）时对所有关键帧进行优化，计算量很大。如图5（b）所示，在建图过程中，我们固定较早帧的位姿，仅优化固定局部窗口内关键帧所观测到的地图点。

B. 地图融合

边缘端完成局部建图后，云端会进行全局位姿图优化（GPGO）。在我们的方法中，不区分机器人间闭环与机器人内闭环。如图6所示，我们使用 DBoW 方法检查候选关键帧。当成功检测到候选关键帧 KF_c 时，通过特征匹配并优化重投影误差，求解共视关键帧 KF_q 和 KF_c 之间的相对位姿 T_{qc} 。当 KF_q 和 KF_c 来自同一机器人时，在机器人内进行闭环优化；当 KF_q 和 KF_c 来自不同机器人时，则优化机器人间的相对位置。尽管增加关键帧数量可以提高优化后的位姿精度，但随着机器人数量增加，关键帧数量也会大幅增加，从而显著增加计算负担。根据 [11], 中的方法消除冗余关键帧后，云端 GPGO 问题可描述为:

$$\min_{\chi,\mathcal{L}}\{\left\|e^{c}_{prio}\right\|^2+\sum_{j\in C}\rho(\left\|e^{j,j+1}_{imu}\right\|^2)+\sum_{l\in L}\sum_{k\in C}\delta(\left\|e^{l,k}_{cam}\right\|^2)\} \tag{19}$$

其中， e^{c}_{prio} 是用于固定第一帧关键帧的位姿约束。与局部优化不同，GPGO针对所有关键帧和地图点进行优化，而惯性测量单元测量施加的约束仅存在于同一机器人连续关键帧之间。

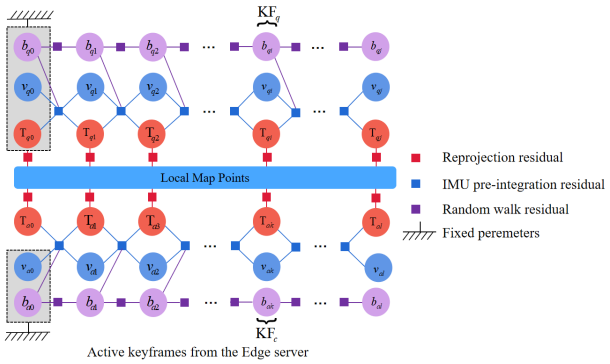


图6. 云端视觉惯性联合优化因子图。

尽管增加关键帧和地图点的数量可以提高优化结果的精度，但随着场景规模扩大，越来越多的关键帧和地图点会增加GPGO的计算负担。在 [11], 中，他们通过计算关键帧之间的相似度以及地图点的观测情况来进行剪枝。但这种方法有时会删除观测良好的地图点。同时，贪心移除关键帧会影响最终优化的精度。受 [63], 启发，我们通过MBP（提取地图骨干，并利用良好的观测创建虚拟关键帧）来减少地图存储量，并确保GPGO的运行速度。

VII. 实验

我们在公开可用的 EuRoC 数据集上评估了系统，该数据集包含三个场景：机器大厅（MH 01- MH 05序列）、室内房间1（V1 01-V1 03序列），以及室内房间2（V2 01-V2 03序列）。在该数据集中，相机数据由 Aptina MT9V034 全局快门相机采集，惯性测量单元数据由 ADIS16448 采集，地面真值则由 Leica MS50 激光跟踪仪采集。相机内参和相机-惯性测量单元外参通过 Kalibr [64]获取。算法在真实设备上进行测试。我们的服务器和移动机器人平台配备的型号如下:

- 机器人：Intel(R) Core(TM) i3-9100 CPU，8GB 内存
- 边缘服务器：Intel(R) Core(TM) i7-9750H CPU，8GB 内存。
- 云服务器：亚马逊云科技（AWS）云服务器。

此外，我们在真实场景中运行了所提出的算法。详情见第 VII-D 节。

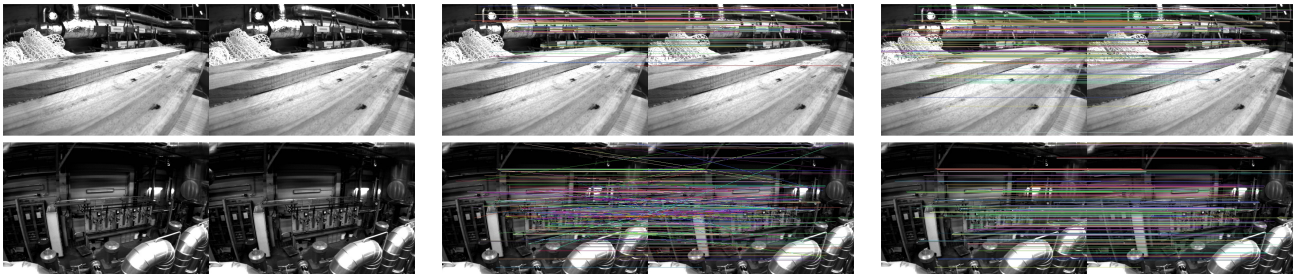
A. 光流跟踪精度与速度

我们首先测试了IMU辅助非关键帧特征匹配。我们使用第IV节中的方法计算非关键帧之间的匹配，无需计算描述子。我们方法的前端基于ORB-SLAM3算法和LK算法进行了改进。

图7展示了我们在EuRoC数据集上使用传统LK光流跟踪和IMU辅助光流跟踪得到的跟踪结果。在结果中，不同颜色的线连接两个匹配的特征点。可以看出，与传统的LK光流跟踪方法相比，所提出的跟踪方法能够匹配更多特征点，并减少误匹配点的数量。此外，在图8中，我们比较了两种方法在EuRoC数据集上连续20帧中的跟踪速度和跟踪点数。在某些帧中（例如第8帧至第11帧），由于光照发生显著变化，IMU预积分带来的增益受限。此外，由于RANSAC的随机选择过程具有不稳定性，我们无法保证在所有连续帧中都能获得稳定的增益。从结果可以看出，在大多数情况下，与LK光流法相比，我们的方法能够显著增加跟踪点数并提高跟踪速度。这主要是因为IMU预积分能够为待匹配特征点提供较好的初始值。因此，IMU辅助非关键帧特征匹配所提供的稳定跟踪能够提高定位的稳定性。

B. 通信速率

在这个云边端系统中，我们重点关注如何通过降低延迟和带宽占用来实现高定位精度。因此，我们从客户端通信延迟和带宽占用两个方面，与当前最先进的算法进行比较。与最先进的局部二进制特征压缩方法相比， [58],我们的方法将关键帧和非关键帧分类后进行传输。基于光流的非关键帧匹配无需使用描述子。表 I 展示了在 EuRoC 数据集中，所提方法和 [58] 传输一个特征所需的内存大小。传输关键帧时，我们的方法与帧内编码模式使用相同的 217 位内存，



(a) 原始图像 (b) 不使用 IMU 预积分的 LK (c) 使用 IMU 预积分的 LK（本文提出的方法）

图 7. EuRoC 上传统 LK 光流跟踪与 IMU 辅助光流跟踪的跟踪结果。原始图像数据 (a)、未使用 IMU 的传统 LK 光流跟踪结果 (b)、IMU 辅助的 LK 光流跟踪结果 (c)。

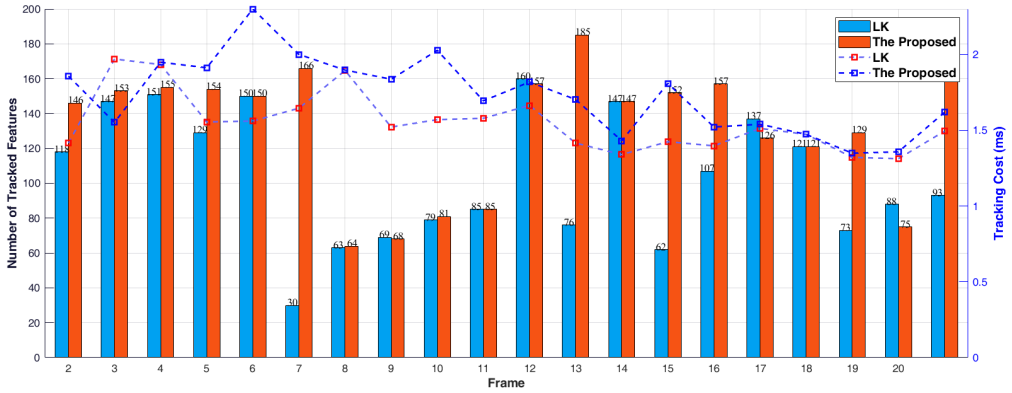


图 8. 跟踪点数及跟踪耗时。蓝色柱表示 LK 光流法跟踪的点数，橙色柱表示本文提出的方法跟踪的点数。紫色虚线表示 LK 跟踪耗时，绿色虚线表示本文提出的方法的跟踪耗时。

表 I 单个特征点所需内存的比较

方法	[58] 中的帧内编码	本文提出的方法	
		KF	非关键帧
内存	217 位	217 位	27 位

[58] 用于传输一个特征。在非关键帧中，我们的方法大幅减少了数据量，因为无需传输描述子和视觉词索引。图9详细展示了传输一个特征时特征点、视觉词索引和残差向量所占的比例。此外，我们还在 EuRoC 数据集上单独进行了数据传输实验。我们计算了 EuRoC 数据集中 8 个序列中从机器人到边缘服务器的平均通信速率。 [58] 方法的通信速率为 211.9kbits/s，而所提方法为 46.81kbits/s。

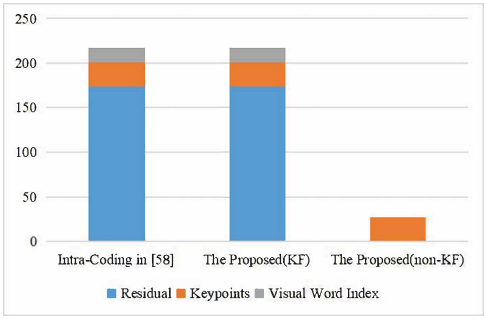


图9. 无损压缩下不同方法中各特征各组成部分所占的比例。

C. 定位与地图融合精度

对于集中式多机器人SLAM系统，主要评价指标是全局一致性和轨迹精度。因此，分别使用绝对轨迹误差（ATE）的均方根误差（RMSE）评估 MH 01-MH 03、MH 01- MH 05、V1 01-V1 03 和 V2 01-V2 03 的协同定位精度。

ATE 的 RMSE 可表示为：

$$\mathrm{ATE}_{\mathrm{RMSE}} = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N} \left\| trans(\mathrm{T}_{\mathrm{G}}^{-1}\mathrm{T}_{\mathrm{E}}) \right\|^2\right)} \qquad (20)$$

其中，T_{G} 和 T_{E} 分别表示真实轨迹和估计轨迹的位姿。N 表示轨迹包含的帧数。ATE 是估计位姿与真实位姿之间的直接差异，可直接反映算法精度和全局轨迹的一致性。

我们在机器人端（i3 CPU）、边缘服务器（i7 CPU）和AWS云服务器上运行我们的架构。表II展示了CVIDS与CCM-SLAM [49], VINS-mono [33], ORB-SLAM3 [15],所获得的绝对轨迹误差均方根误差的定量比较

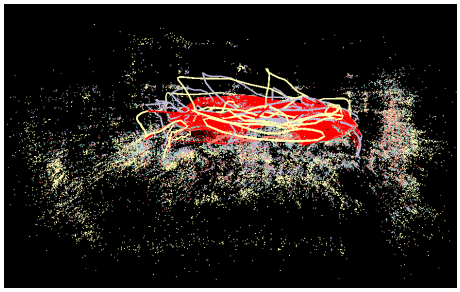
表 II EuRoC 数据集上的绝对轨迹误差均方根误差。我们的方法取五次运行后的平均值。请注意，粗体表示最佳结果，“-”表示文献中找不到相应方法的结果。

		序列			
		MH_01-MH_03	MH_01-MH_05	V1_01-V1_03	V2_01-V2_03
方法	CCM-SLAM [49]	0.077	-	-	-
	VINS-mono [33]	-	0.210	-	-
	ORB-SLAM3 [15]	0.037	0.065	0.040	0.048
	CVIDS [51]	-	0.074	-	-
	COVINS AWS [11]	**0.025**	**0.039**	0.049	-
	所提方法（不使用 MBP）	0.028	0.051	0.035	0.045
	所提方法（使用 MBP）	**0.025**	0.046	**0.031**	**0.033**

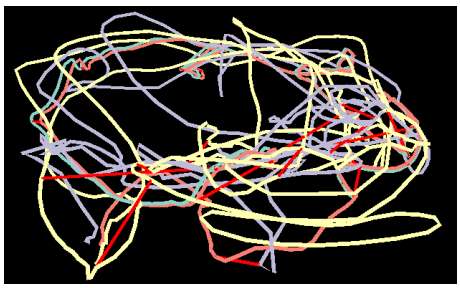
表 III 使用不同冗余关键帧移除方法的结果比较。

方法	绝对轨迹误差均方根	关键帧数量	地图点数量
所提出的方法（不使用 MBP）	0.051	1550	57169
所提出的方法（使用 MBP）	**0.046**	**925**	**41923**

CVIDS [51]以及COVINS AWS [11]。由于系统架构不同，我们从现有参考文献中获取了多个序列上多机器人定位的平均绝对轨迹误差均方根误差。从结果可以看出，我们的方法在所有序列上的平均均方根误差与当前最先进的集中式方法相当，甚至优于大多数结果，这证实了其定位的优越性。此外，带MBP的所提方法在全部四组实验中的表现均优于不带MBP的所提方法。这表明MBP能够为边缘辅助多机器人SLAM带来一定增益。值得一提的是，在实验中，我们使用了VINS-Mono和ORB-SLAM3的多会话模式。这些数据均证实了我们的架构为多机器人系统带来的改进。然而，在MH 01-MH 05序列中，所提方法相对COVINS的定位效果较差。不过，根据我们的验证 [3], , 基于光流法的前端相比COVINS中基于特征点法的前端，速度提升十分明显。此外，为了证明MBP在内存和运行速度方面的优势，我们比较了带MBP的所提方法与不带MBP的所提方法在MH 01-MH 05序列上的具体指标。结果表明，带MBP的所提方法能够通过优化更少的关键帧和地图点来保证高精度。



(a) 地图与共视关系



(b) 轨迹与闭环关键帧

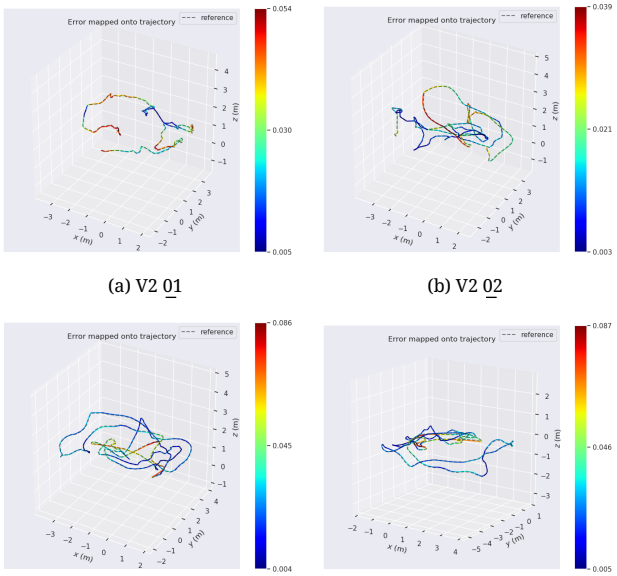
图 10. 所提出方法在 V2 序列中的协同定位与建图结果。

我们对 V2 场景中的序列进行建图，该场景中的光流颇具挑战性。图 10 展示了我们的协同 SLAM 在三个序列中的结果。不同颜色表示不同序列中的轨迹和地图点。图 10(a) 中的红色连线表示共视，图 10(b) 中的红色连线表示机器人之间的闭环关键帧。为清晰展示 V2 场景中三个序列的机器人轨迹，我们使用 evo 工具将生成的轨迹与参考真值对齐。如图 11 所示，我们分别绘制了 V2 01、V2 02 和 V2 03 的单独轨迹及协作式定位轨迹。结果表明，每个机器人的轨迹以及合并后的轨迹都能与参考真值良好对齐。

D. 真实场景实验

我们使用两台 TurtleBot 3（配备树莓派3）和一台 TurtleBot 2（配备联想Y7000），在图 12 所示的室内场景中运行我们的方法。机器人配备了 MYNT EYE深度相机，并使用 Kalibr 对相机和 IMU 进行联合标定。机器人的地面真值由 Marvelmind[1] 室内“全球定位系统”和 IMU 共同获得。机器人从起点出发绕行两周，并在轨迹终点会合。整个过程中，我们使用联想Y7000作为边缘服务器，并使用 AWS 作为云服务器。在真实场景实验中，我们使用 Kalibr 计算相机的内参以及相机-IMU的外参。我们使用三台不同的机器人行驶两

[1]https://marvelmind.com/pics/架构 比较.p df



(a) V2 01 (b) V2 02

(c) V2 03 (d) V2 01-V2 03

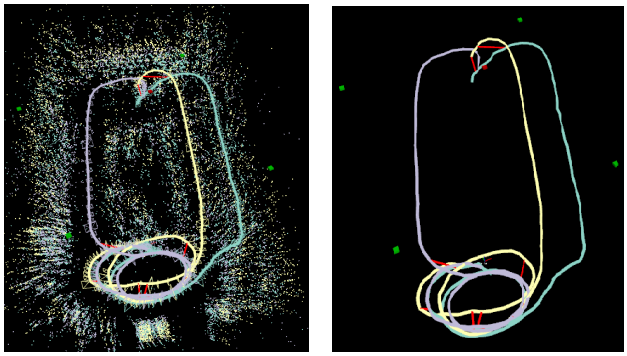
图 11. 所提方法在 V2 序列中的协同定位与建图轨迹。

周，从房间一端出发，沿不同轨迹行进，并通过不同路径在终点会合，且彼此不可见。

不同机器人的轨迹以及构建的地图点如图13所示。使用 Marvelmind 室内“GPS”测量地面真值时，我们使用四个固定锚点（图13(b)中的绿色方块）来定位移动锚点（图13(b)中的红色方块）。此外，通过融合 IMU 信息，计算得到参考真值。三台机器人在起点处绕行两周，随后沿不同轨迹在终点处形成闭环。结果表明，起点与终点之间形成了明显的共视关系。在图14中，我们将所提方法得到的估计轨迹与参考真值对齐。为便于直观比较，我们将估计轨迹与真实场景的二维栅格地图对齐，如图15所示。结果表明，我们的方法在实际场景中具有良好的估计效果。



图12. 实验场景及设备。



(a) 地图与轨迹 (b) 轨迹与闭环关键帧

图13. 所提方法在真实场景中的协同定位与建图结果。

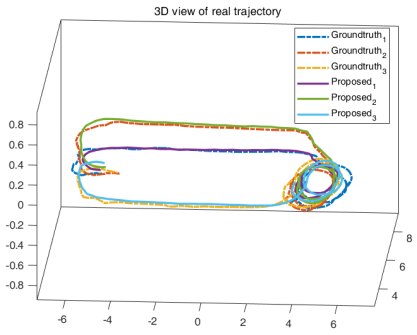


图14. 真实场景中测量轨迹与地面真值的比较。

VIII. 结论

本文提出了一种用于云端、边缘端和客户端协同计算的集中式多机器人SLAM系统。该系统可通过多个独立机器人的协作构建全局稀疏地图，并估计它们在全局地图中的位置。研究发现，在计算和传输过程中，非关键帧描述子并非必需。然而，描述子的计算和传输会消耗大量计算资源、内存和传输带宽。我们构建了一种无需描述子的相邻帧间惯性测量单元辅助匹配模型。同时，我们对传输数据进行剪枝和压缩。实验分析表明，我们的匹配方法能够实现更高的特征跟踪精度和更快的特征跟踪速度。此外，与类似系统相比，我们能够以最少的数据传输量完成协同SLAM，且不损失精度。云端完成全局地图融合后，实验对比证明，所提方法在精度、速度和数据传输量方面均可与当前最先进的多地图系统相媲美。未来，我们将整合激光雷达，以解决复杂场景中的视觉感知局限。此外，为避免动态环境、能量和电池限制、传感器故障等因素对系统造成干扰，未来我们将在系统中加入更多传感器组合，以及基于深度学习的识别方法。

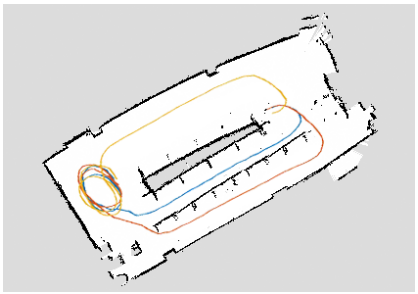


图 15. 轨迹与真实场景的二维栅格地图叠加，用于视觉比较。

参考文献

[1] X. Lin, J. Ruan, Y. Yang, L. He, Y. Guan, and H. Zhang, “针对检测缺失问题的鲁棒语义SLAM数据关联,” *IEEE* 自动化科学与工程汇刊, pp. 1–13, 2023. [2] Q. Sun, J. Yuan, X. Zhang, and F. Duan, “Plane-edge-slam：室内环境中平面与边缘的无缝融合,” *IEEE* 自动化科学与工程汇刊, vol. 18, no. 4, pp. 2061–2075, 2021. [3] S. Wen, X. Liu, H. Zhang, F. Sun, M. Sheng, and S. Fan, “基于双目VINS的移动载具稠密点云地图构建,” 《国际摄影测量与遥感学会摄影测量与遥感杂志》, vol. 178, 344, 2021 328pp. – . [4] J. McConnell, F. Chen, and B. Englot, “用于水下声呐同步定位与地图构建的俯视图像因子,” 《*IEEE*机器人与自动化快报》, vol. 7, no. 2, pp. 4901–4908, 2022. [5] P. Liu, K. Ma, J. Yang, B. Yang, Z. Liu, and X. Guan, “中继辅助智能电网通信中的联合干扰管理与功率分配,” *IEEE* 物联网期刊, vol. 7, no. 3, pp. 1927–1938, 2020. [6] B. D. Gouveia, D. Portugal, D. C. Silva, and L. Marques, “分布式机器人系统中的计算共享：一个SLAM案例研究,” vol. 12, no. 2, 2015, pp. 410–422. [7] S. Shen, N. Michael, and V. Kumar, “用于旋翼微型飞行器自主飞行的紧耦合单目视觉惯性融合,” in *2015 IEEE*机器人与自动化国际会议 *(ICRA)*, 2015, pp. 5303–5310.[8] Y. Tian, Y. Chang, F. Herrera Arias, C. Nieto-Granda, J. P. How, and L. Carlone, “Kimera-multi：面向多机器人系统的鲁棒分布式稠密度量语义SLAM,” *IEEE*机器人学汇刊, vol. 38, no. 4, pp. 2022–2038, 2022. [9] Y. Huang, T. Shan, F. Chen, and B. Englot, “Disco-slam：利用分布式扫描上下文的多机器人激光雷达SLAM及两阶段全局—局部图优化,” 《*IEEE*机器人与自动化快报》, vol. 7, no. 2, pp. 1150–1157, 2022. [10] M. Karrer, P. Schmuck, and M. Chli, “Cvi-slam—协同视觉惯性SLAM,” 《*IEEE*机器人与自动化快报》, vol. 3, no. 4, pp. 2762–2769, 2018. [11] P. Schmuck, T. Ziegler, M. Karrer, J. Perraudin, and M. Chli, “Covins：用于集中式协作的视觉惯性SLAM,” in *2021 IEEE*混合与增强现实国际研讨会（*ISMAR-Adjunct*）, 2021, pp. 171–176. [12] X. Song, P. Sun, S. Song, and V. Stojanovic, “互联系统的量化神经自适应有限时间预设性能控制,” *Neural Comput. Appl.*, vol. 35, no. 21, p. 15429–15446, apr 2023.[13] X. Song, C. Wu, V. Stojanovic, and S. Song, “基于1比特编码—解码的事件触发固定时间自适应控制：保证无人水面艇的跟踪性能,” 《控制工程实践》, vol. 135, p. 105513, 2023. [在线]。获取地址：https://www.sciencedirect.com/science/article/pii/S0967066123000825 [14] C. Zhou, H. Tao, Y. Chen, V. Stojanovic, and W. Paszke, “受约束系统的鲁棒点对点迭代学习控制：一种最小能量方法,” *International Journal of Robust and Nonlinear Control*, vol. 32, pp. 10 139–10 161, 2022. [15] C. Campos, R. Elvira, J. J. G. Rodr´ıguez, J. M. M. Montiel, and J. D. Tard´os, “Orb-slam3：用于视觉、视觉惯性和多地图SLAM的高精度开源库,” *IEEE*机器人学汇刊, vol. 37, no. 6, pp. 1874–1890, 2021. [16] B. Lucas 和 T. Kanade，“一种迭代图像配准技术及其在立体视觉中的应用（ijcai）”，第81卷，1981年4月，第121–130页。[17] J. Engel、T. Schoeps 和 D. Cremers，“Lsd-slam：大规模直接单目slam”，第8690卷，2014年9月，第1–16页。[18] J. Engel、V. Koltun 和 D. Cremers，“直接稀疏里程计”，IEEE模式分析与机器智能汇刊，第40卷，第3期，第611–625页，2018年。[19] A. J. Davison、I. D. Reid、N. D. Molton 和 O. Stasse，“MonoSLAM：实时单相机slam”，IEEE模式分析与机器智能汇刊，第29卷，第6期，第1052–1067页，2007年。[20] R. Mur-Artal、J. M. M. Montiel 和 J. D. Tard´os，“Orb-slam：一种灵活且准确的单目同步定位与建图系统”，IEEE机器人学汇刊，第31卷，第5期，第1147–1163页，2015年。[21] R. Mur-Artal 和 J. D. Tard´os，“Orb-slam2：一种适用于单目、立体和RGB-D摄像头的开源slam系统”，IEEE机器人学汇刊，第33卷，第5期，第1255–1262页，2017年。[22] A. I. Mourikis 和 S. I. Roumeliotis，“一种用于视觉辅助惯性导航的多状态约束卡尔曼滤波器”，载于《2007年IEEE国际机器人与自动化会议论文集》，2007年，第3565–3572页。[23] M. Li 和 A. I. Mourikis，“提高基于EKF的视觉惯性里程计的精度”，载于《2012年IEEE国际机器人与自动化会议》，2012年，第828–835页。[24] M. Li 和 A. Mourikis，“高精度、一致的基于EKF的视觉惯性里程计”，《国际机器人研究杂志》，第32卷，第690–711页，2013年5月。[25] K. Sun、K. Mohta、B. Pfrommer、M. Watterson、S. Liu、Y. Mulgaonkar、C. J. Taylor 和 V. Kumar，“面向快速自主飞行的鲁棒立体视觉惯性里程计”，《IEEE机器人与自动化快报》，第3卷，第2期，第965–972页，2018年。[26] M. Bloesch、S. Omari、M. Hutter 和 R. Siegwart，“采用直接式扩展卡尔曼滤波方法的鲁棒视觉惯性里程计”，载于《2015年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2015年，第298–304页。[27] M. Bloesch、M. Burri、S. Omari、M. Hutter 和 R. Siegwart，“采用直接光度反馈的迭代扩展卡尔曼滤波视觉惯性里程计”，《国际机器人研究杂志》，第36卷，第1053–1072页，2017年9月。[28] Y. Song、Z. Zhang、J. Wu、Y. Wang、L. Zhao 和 S. Huang，“一种用于基于对象的slam的右不变扩展卡尔曼滤波器”，《IEEE机器人与自动化快报》，第7卷，第2期，第1316–1323页，2022年。[29] M. Karimi、M. Oelsch、O. Stengel、E. Babaians 和 E. Steinbach，“Lola-slam：采用连续扫描切片的低延迟激光雷达同步定位与地图构建”，《IEEE机器人与自动化快报》，第6卷，第2期，第2248–2255页，2021年。[30] C. Xia、Y. Shen、Y. Yang、X. Deng、S. Chen、J. Xin 和 N. Zheng，“基于机载传感器和分层地图的自动驾驶汽车自主定位”，IEEE控制论汇刊，第1–14页，2022年。[31] B. Triggs、P. F. McLauchlan、R. I. Hartley 和 A. W. Fitzgibbon，“光束法平差——现代综述”，载于《视觉算法：理论与实践》。柏林，海德堡：施普林格柏林海德堡，2000年，第298–372页。[32] S. Leutenegger、S. Lynen、M. Bosse、R. Siegwart 和 P. Furgale，“采用非线性优化的基于关键帧的视觉惯性里程计”，《国际机器人研究杂志》，第34卷，2014年2月。[33] T. Qin、P. Li 和 S. Shen，“Vins-mono：一种鲁棒且灵活的单目视觉惯性状态估计器”，IEEE机器人学汇刊，第34卷，第4期，第1004–1020页，2018年。[34] T. Qin、J. Pan、S. Cao 和 S. Shen，“一种用于多传感器局部里程计估计的通用优化框架”，CoRR，第abs/1901.03638卷，2019年。 [在线]。网址：http://arxiv.org/abs/1901.03638 。[35] R. Mur-Artal 和 J. D. Tardos，“采用地图复用的视觉惯性单目同步定位与地图构建”，《IEEE机器人与自动化快报》，第2卷，第2期，第796–803页，2017年。[36] E. Rublee、V. Rabaud、K. Konolige 和 G. Bradski，“Orb：SIFT或SURF的一种高效替代方案”，载于《2011年国际计算机视觉会议》，2011年，第2564–2571页。[37] R. Elvira、J. D. Tard´os 和 J. Montiel，“Orbslam-atlas：一种鲁棒且准确的多地图系统”，载于《2019年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2019年，第6253–6259页。[38] A. Cunningham、M. Paluri 和 F. Dellaert，“Ddf-sam：采用约束因子图的完全分布式同步定位与地图构建”，载于《2010年IEEE/RSJ智能机器人与系统国际会议》，2010年，第3025–3030页。[39] A. Cunningham、V. Indelman 和 F. Dellaert，“Ddf-sam 2.0：一致的分布式平滑与建图”，载于《2013年IEEE国际机器人与自动化会议》，2013年，第5220–5227页。

[40] S. Choudhary、L. Carlone、C. Nieto、J. Rogers、H. I. Christensen 和 F. Dellaert，“兼顾隐私与通信约束的分布式轨迹估计：两阶段分布式高斯-赛德尔方法”，载于《2016年IEEE国际机器人与自动化会议（ICRA）》，2016年，第5261–5268页。[41] S. Choudhary、L. Carlone、C. Nieto-Granda、J. Rogers、H. Christensen 和 F. Dellaert，“兼顾隐私与通信约束的分布式建图：轻量级算法与基于对象的模型”，《国际机器人研究杂志》，第36卷，2017年2月。[42] T. Cieslewski、S. Choudhary 和 D. Scaramuzza，“数据高效的去中心化视觉slam”，载于《2018年IEEE国际机器人与自动化会议（ICRA）》，2018年，第2466–2473页。[43] R. Arandjelovi´c、P. Gronat、A. Torii、T. Pajdla 和 J. Sivic，“Netvlad：用于弱监督地点识别的卷积神经网络架构”，IEEE模式分析与机器智能汇刊，第40卷，第6期，第1437–1451页，2018年。[44] P.-Y. Lajoie、B. Ramtoula、Y. Chang、L. Carlone 和 G. Beltrame，“Door-slam：面向机器人团队的分布式、在线且具备离群值鲁棒性的slam”，《IEEE机器人与自动化快报》，第5卷，第2期，第1656–1663页，2020年。[45] D. Zou 和 P. Tan，“Coslam：动态环境中的协作式视觉SLAM”，IEEE模式分析与机器智能汇刊，第35卷，第2期，第354–366页，2013年。[46] C. Forster、S. Lynen、L. Kneip 和 D. Scaramuzza，“多个微型飞行器的协作式单目SLAM”，载于《2013年IEEE/RSJ智能机器人与系统国际会议》，2013年，第3962–3970页。[47] F. Bonin-Font 和 A. Burguera，“面向自主海洋航行器的多机器人视觉图优化slam探索”，《海洋科学与工程杂志》，第8卷，第6期，2020年。[48] L. Bartolomei、M. Karrer 和 M. Chli，“采用代理-服务器架构的多机器人协同：面向部分未知环境中的自主导航”，载于《2020年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2020年，第1516–1522页。[49] P. Schmuck 和 M. Chli，“Ccm-slam：面向机器人团队的鲁棒、高效、集中式协作单目同步定位与地图构建”，《野外机器人学杂志》，第36卷，第763–781页，2019年。[50] S. Wen、J. Chen、F. R. Yu、F. Sun、Z. Wang 和 S. Fan，“基于边缘计算的协作式车辆实时三维建图”，《IEEE车辆技术汇刊》，第69卷，第11期，第12 470–12 481页，2020年。[51] T. Zhang、L. Zhang、Y. Chen 和 Y. Zhou，“Cvids：面向多智能体视觉惯性slam的协作式定位与稠密建图框架”，《IEEE图像处理汇刊》，第31卷，第6562–6576页，2022年。[52] W. Shi、J. Cao、Q. Zhang、Y. Li 和 L. Xu，“边缘计算：愿景与挑战”，IEEE 物联网期刊，第3卷，第5期，第637–646页，2016年。[53] J. Xu、H. Cao、D. Li、K. Huang、C. Qian、L. Shangguan 和 Z. Yang，“边缘辅助移动语义视觉SLAM”，载于《IEEE INFOCOM 2020——IEEE计算机通信会议》，2020年，第1828–1837页。[54] A. Ali、M. Kouroshli、S. Semenova、z. s. Hashemifar、S. Ko 和 K. Dantu，“Edge-slam：边缘辅助视觉同步定位与建图”，《ACM嵌入式计算系统汇刊》，第22卷，2022年9月。[55] P. Huang、L. Zeng、X. Chen、K. Luo、Z. Zhou 和 S. Yu，“边缘机器人学：边缘计算加速的多机器人同步定位与建图”，IEEE 物联网期刊，第9卷，第15期，第14 087–14 102页，2022年。[56] D. Lan、A. Taherkordi、F. Eliassen、L. Liu、S. Delbruel、S. Dustdar 和 Y. Yang，“异构边缘平台上的任务划分与编排：以视觉应用为例”，IEEE 物联网期刊，第9卷，第10期，第7418–7432页，2022年。[57] Y.-P. Wang、Z.-X. Zou、C. Wang、Y.-J. Dong、L. Qiao 和 D. Manocha，“Orbbuf：一种用于远程视觉SLAM的鲁棒缓冲方法”，载于《2021年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2021年，第8706–8713页。[58] D. Van Opdenbosch、M. Oelsch、A. Garcea、T. Aykut 和 E. Steinbach，“面向远程视觉SLAM的局部二进制特征选择与压缩”，载于《2018年IEEE国际机器人与自动化会议（ICRA）》，2018年，第7270–7277页。[59] D. Van Opdenbosch 和 E. Steinbach，“采用压缩特征交换的协作式视觉SLAM”，《IEEE机器人与自动化快报（RAL）》，第4卷，第1期，第57–64页，2019年。[60] Q. Fu、H. Yu、X. Wang、Z. Yang、Y. He、H. Zhang 和 A. Mian，“无需关键点描述子的快速ORB-SLAM”，《IEEE图像处理汇刊》，第31卷，第1433–1446页，2022年。

[61] C. Forster、M. Pizzoli 和 D. Scaramuzza，《Svo：快速半直接单目视觉里程计》，载于*2014*年*IEEE*机器人与自动化国际会议（*ICRA*），2014年，第15–22页。[62] C. Forster、L. Carlone、F. Dellaert 和 D. Scaramuzza，《面向实时视觉惯性里程计的流形上预积分》，*IEEE*机器人学汇刊，第33卷，第1期，第1–21页，2017年。[63] J. Xu、H. Cao、Z. Yang、L. Shangguan、J. Zhang、X. He 和 Y. Liu，《{SwarmMap}：在边缘端扩展实时协作式视觉 {同步定位与建图}》，载于第*19*届*USENIX*网络系统设计与实现研讨会（*NSDI 22*），2022年，第977–993页。[64] P. Furgale、J. Rehder 和 R. Siegwart，《多传感器系统统一时间与空间标定》，载于*2013*年*IEEE/RSJ*智能机器人与系统国际会议，2013年，第1280–1286页。



Xin Liu 于2019年获得燕山大学自动化专业学士学位，目前正在燕山大学电气工程系攻读博士学位。他的研究兴趣包括同步定位与建图、边缘端计算和多传感器信息融合。



Shuhuan Wen 于2005年在中国秦皇岛燕山大学获得控制理论与控制工程博士学位。她现任燕山大学电气工程系自动控制教授。她合著了一本书，并发表了约40篇论文。她的研究兴趣包括同步定位与建图、计算机视觉、机器人学、三维物体识别和重建。2011至2013年，Wen博士曾在加拿大渥太华大学、卡尔顿大学和西蒙弗雷泽大学担任访问学者。

2021至2022年，她还担任加拿大阿尔伯塔大学客座教授。2021至2022年，她担任IROS副编辑。



Jing Zhao 是中华人民共和国秦皇岛燕山大学自动化系副教授。他于2008年在中华人民共和国济南山东大学获得自动化专业理学学士学位，2011年在中华人民共和国西安中国航天科技集团公司第四研究院获得测量与测试技术及仪器专业理学硕士学位，并于2016年在中华人民共和国天津天津大学获得控制科学与工程专业博士学位。他的研究兴趣包括脑机接口以及脑机器人协作。



Tony Z. Qiu 于2001年和2003年分别在中国北京清华大学获得控制科学与工程专业理学学士学位和智能交通系统专业理学硕士学位，并于2007年在威斯康星大学麦迪逊分校获得博士学位。2008年至2009年，他在加州大学伯克利分校的加州先进交通与公路合作伙伴计划（PATH）担任博士后研究员。目前，他在土木与环境工程系担任助理教授，加拿大阿尔伯塔省埃德蒙顿市阿尔伯塔大学。他的研究兴趣包括智能交通系统、交通建模与仿真、交通运行与控制、交通状态估计与预测，以及城市道路和高速公路的安全分析。



Hong Zhang 于1982年获得美国波士顿东北大学电气工程理学学士学位，并于1986年获得美国印第安纳州西拉斐特市普渡大学电气工程博士学位。1986年至1987年，他在宾夕法尼亚大学从事博士后研究，之后加入加拿大阿尔伯塔大学计算机科学系。他担任加拿大自然科学与工程研究委员会工业研究讲席。其当前研究兴趣包括机器人学、计算机视觉和图像处理。