

# 地图实验室 2.0—模块化和多模式映射框架

安德烈·克拉马留克<sup>1,\*</sup>, 卢卡斯·伯恩赖特<sup>1,\*</sup>, 弗洛里安·乔普<sup>2,\*</sup>, 马里乌斯·费尔<sup>3</sup>, 维克多·雷格瓦特<sup>1</sup>  
胡安·涅托<sup>4</sup>, 罗兰·西格瓦特<sup>1</sup>, 塞萨尔·卡德纳<sup>1</sup>

arXiv:2212.00654v2 [cs.RO] 2023年1月3日

**抽象的**—将多种传感器模式和深度学习集成到同步定位和地图构建(SLAM)系统中是当前研究的热门领域。多模态是实现具有挑战性的环境中稳健性和具有不同传感器设置的异构多机器人系统互操作性的垫脚石。**地图实验室 2.0**, 我们提供了一个多功能的开源平台, 便于开发、测试和将新模块和功能集成到成熟的 SLAM 系统中。通过大量实验, 我们证明了**地图实验室 2.0**的准确度与 HILTI 2021 基准上的最新水平相当。此外, 我们通过三个用例展示了我们系统的灵活性: i) 大规模 (~10公里) 多机器人多会话 (23 个任务) 地图绘制, ii) 非视觉地标的集成, 以及 iii) 将基于语义对象的回环模块纳入地图绘制框架。代码可在 <https://github.com/ethz-asl/maplab> 上开源获取。

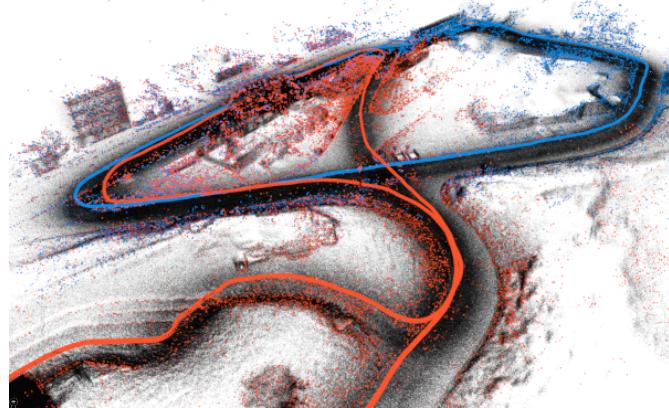


图 1: 我们建议地图实验室 2.0, 一个灵活、通用的多机器人和多模式框架。地图实验室 2.0 可以无缝集成多个机器人 (彩色路径)、视觉地标 (彩色点) 和 LiDAR 扫描 (黑点)。

## 我引言

**即时定位与地图构建 (SLAM)** 是各种机器人应用 (如自动驾驶 [1], 移动操控 [2] 和增强/混合现实) 的重要组成部分。在这些应用中, 机器人平台需要了解周围环境及其位置才能执行给定任务, 无论是自动驾驶到特定目的地还是拾取和运送物品。更进一步的是执行长期地图构建的能力, 这通常需要用于处理和合并多个地图的工具, 从而实现更广泛的多样化应用和任务。

近年来, 许多定制的 SLAM 解决方案已成功开发用于特定环境或传感器配置 [3]–[11]。然而, 在 SLAM 完全解决或在普遍的操作条件下普遍部署之前, 仍有许多挑战。由于多模态系统能够弥补单个传感器或方法的弱点, 因此最近在融合多种模态方面的努力获得了显著的关注。因此, 可以实现更稳健的机器人操作

这项工作得到了国家数字制造研究能力中心 (NCCR) 的部分支持, 并获得了欧盟“地平线 2020”研究与创新计划 (资助协议 871542) 的资助。

\*作者对这项工作做出了同等的贡献。

<sup>1</sup>作者是苏黎世联邦理工学院自主系统实验室的成员, 瑞士; /名字.姓氏/ @mavt.ethz.ch

<sup>2</sup>作者就职于英国 Arrival Ltd, 但这项工作是在作者是<sup>1</sup>; tschopp@arrival.com

<sup>3</sup>作者在瑞士 Voliro, 但作品是在作者是<sup>1</sup>; marius.fehr@voliro.com

<sup>4</sup>作者来自瑞士微软, 但这项工作是在作者是<sup>1</sup>; juannieto@microsoft.com

特别感谢 Thomas Schneider、Marcin Dymczyk、Juichung Kuo、Patrick Pfreundschuh、Nicolas Scheidt、Marius Brühlmeier 和 Benjamin Hahn 对该框架的概念化和实施所做的重大贡献。

在恶劣环境下, 甚至在传感器全部或部分故障的情况下, 仍能保持良好的性能。目前已有结合多种不同传感器的工作 [3]–[5], 并取得了显著的性能。然而, 与其他开源 SLAM 框架 [8]–[11] 一起, 这些系统紧密集成。更具体地说, 它们仅在特定的传感器配置下才能发挥作用, 基本模块 (例如传感器模块 (例如, 里程计、定位或特征提取) 相互交织。修改这些模块或添加新功能需要大量的工程工作, 这会给科学的研究和新产品的开发带来巨大的开销。因此, 需要能够无缝集成各种传感器设置并利用多种传感器模式的多功能系统。灵活支持多种模式也是异构多机器人系统的基础, 其中不同的机器人可以配备不同的传感器组合, 例如, 由于平台限制。

地图实验室 2.0 为多会话、多机器人和多功能多模式映射提供了开源平台。原始地图实验室 [11] 是一个开源工具箱, 用于创建和管理专用的视觉惯性地图。地图实验室 2.0, 我们通过集成多种新模式 (例如 LiDAR、GPS 接收器、车轮编码器、语义对象等) 将原始框架扩展到远远超出其初始范围。这些示例提供了轻松扩展到其他传感模式的模板。地图实验室 2.0 还提供了易于集成外部组件的接口, 例如添加任意数量的不同视觉特征或回环约束。这些功能使我们的新平台非常适合用作深度学习关键点检测器和回环引擎的开发和研究工具, 到目前为止, 这些引擎大多是单独测试的 [12]。此外, 现在可以使用在线协作 SLAM 地图实验室 2.0

由于新的子映射功能，可以从多个来源在线构建、优化和共同定位一个全局地图。这是通过我们实施一个新的集中式服务器节点实现的，该节点可以聚合来自多个机器人的数据，并可以将协作构建的地图传输回机器人以提高性能 [13]。我们在多个实验和数据集中展示了我们系统的功能和性能，为非视觉关键点、深度学习描述符集成和基于语义对象的循环闭合引擎提供了概念验证实现。

我们的贡献可以总结如下：

- 我们提供了一个开源、多模式和多机器人映射框架，与其他现有方法相比，该框架允许集成和融合无与伦比的大量不同数据。
- 一种在线协作制图系统，利用子地图和中央服务器来编译和分发全球一致、功能丰富的地图。
- 集成任意数量的自定义特征点、描述符和回环的接口。我们在 3D LiDAR 关键点和基于语义对象的回环的实验中展示了它们的灵活性。

## 二、R 兴高采烈西兽人

地图绘制可以定义为创建环境表征的挑战，过去几十年来，地图绘制已经出现了大量且多样化的解决方案，新传感器和新场景带来了重大变化 [14]。多模态技术已经超越了标准传感器融合 (IE、视觉惯性或立体相机) 以包括更复杂的组合，例如 LiDAR 和语义信息。另一个值得注意的主题是多机器人测绘，其中多个机器人同时探索环境并旨在创建一个全局一致的地图。多机器人测绘不同于多会话测绘，后者涉及在不同的时间间隔收集同一位置的测量值并启用会话之间和会话之间的离线操作。虽然多机器人框架可以通过顺序处理数据记录以多会话的方式使用，但这是低效的，因为由于缺乏地图管理工具，每次添加新记录时都需要重新处理所有以前的数据。表 I 显示了重要的 SLAM 框架及其特性的比较。

第一个版本 地图实验室 [11] 是一个为视觉惯性系统设计的多会话映射框架。其他类似的框架是 ORB-SLAM3 [8] 和 RTAB-Map [3]。ORB-SLAM3 是其前身 ORB-SLAM2 [15] 的扩展，增加了对 IMU 和多会话映射功能的支持。RTAB-Map 集成了来自 LiDAR 或 RGB-D 相机的视觉和深度测量。RTAB-Map 的扩展 [16] 支持各种手工制作的视觉特征和 SuperPoint [12]，但不允许轻松集成其他描述符。这两个框架都提供了类似的地图创建和管理功能 地图实验室，并在映射过程中添加了在线循环闭合和优化。上述三个框架都是

表一：最先进的地图绘制框架所支持特性的比较。(差分传感器：可以组合具有不同传感器配置的地图；分机：外部来源；信用证：环路闭合；\*：必须与姿势一起提供 IMU 偏差)

	运 动 联 网 多 机 器 人	多 会 话	多 线 程	在 线 的 差 分 传 感 器	大 外 部 程 里 计	扩 展 功 能	外 部 信 用 证	GPS 技 术	语 义 LC	漂 移 开
RTAB 地图 [3]	+	+	+		+					+
ORB-SLAM3 [8]	+	+	+							+
灯泡 2.0 [6]	二	+		XXX						+
CVI-SLAM [7]		+		+						
科文斯 [9]		+		+		*				+
猛地关门 [10]	二	+		+		+				+
基梅拉 [4]		+		+		*				+
Kimera-Multi [5]	二	+		+		*				+
地图实验室 [11]			+							+
地图实验室 2.0				XXXXXXXXXX						

为特定传感器配置设计的紧密集成系统。相反，我们允许轻松集成不同的传感器设置、视觉特征，并支持任意里程计输入地图实验室 2.0，这有助于异构机器人的使用并提供了新的灵活性。

Kimera [4] 是一个多模态映射框架，它提供带有语义注释的局部和全局 3D 网格以及基于视觉惯性 SLAM 的全局轨迹估计。与地图实验室 2.0，Kimera 不具备多会话功能，并且未使用带有语义注释的 3D 重建来提高 SLAM 估计的准确性。一般来说，语义信息有可能通过成为高级场景理解的催化剂来显著改善地图绘制 [14]。然而，以前利用语义进行地图绘制 [17]–[21] 的工作主要侧重于生成改进的描述符，而不是在完全语义的 SLAM 系统中利用它们。在这项工作中，我们建议利用图像描述符和简单的语义对象表示，这使我们能够使用众所周知的相对姿势误差进行优化。

Kimera-Multi [5] 是 Kimera [4] 的直接扩展，它使用完全分布式系统实现了多机器人场景，但没有改进原始 Kimera 中纯视觉惯性 SLAM 后端。我们的方法改为使用集中式服务器来收集子地图、对其进行优化并创建一个全局一致的地图。COVINS [9] 也探索了类似的集中式设置。然而，COVINS 仅限于视觉惯性用例，而地图实验室 2.0 可以整合多种传感器模式和配置。类似方法还有 LAMP 2.0 [6]、CVI-SLAM [7] 和 DOOR-SLAM [10]，它们提供机器人之间的协作映射，但都是紧密集成的系统，仅限于一种传感器模式，灵活性较差。

虽然存在各种其他 SLAM 框架，但它们主要侧重于特定的传感器或机器人环境配置，并且更改其中任何一个通常都很困难或不可能。据我们所知，所有现有方法 地图实验室 2.0 是最灵活的地图和定位框架，不仅支持各种传感器，还可以无缝适应新的需求。

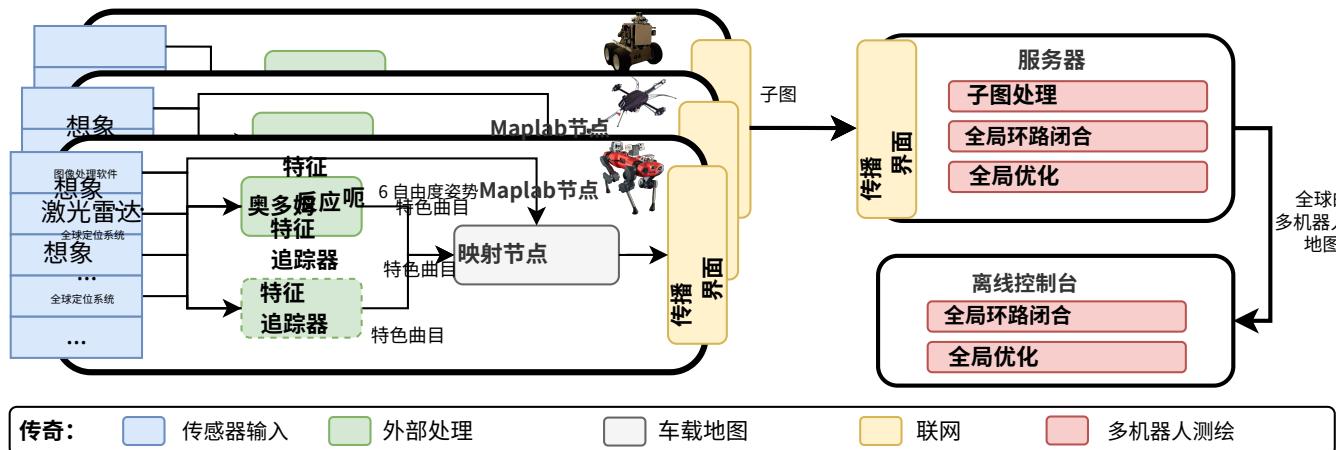


图2：概述地图实验室 2.0 框架及其三个主要组件，即映射节点、中央服务器和离线控制台。映射节点在每个机器人上运行，并将传感器数据收集到子地图中，这些子地图被传递到中央服务器，中央服务器将它们合并成一个全局一致的地图。然后，可以在稍后阶段使用离线控制台提供的工具对该地图进行细化或与其他地图合并。该图展示了一种可能的配置示例，说明如何地图实验室 2.0 可以使用。还有多种其他组合以及可以添加或排除的输入和模块，如第 III 节所述。

### 三、其他地图实验室2.0F模具

的总体结构地图实验室 2.0 框架如图 2 所示。整个框架可以分为三个主要组件：映射节点、映射服务器和离线控制台界面。我们首先概述一下地图实验室 2.0，之后我们更详细地讨论主要组件。

#### A. 地图结构

我们表示地图作为一个或多个的集合任务，其中每个任务都基于单个连续的地图绘制会话。地图的底层结构是一个由顶点和边组成的因子图，其中包含所有机器人信息和不同任务的测量值。机器人在某一时间点的状态被参数化为顶点（6 个自由度 (DoF) 姿势、速度、IMU 偏差）。地标也表示为图中的顶点，其状态定义为 3D 位置。3D 地标可用作环境中具有 3D 位置的任何事物的底层表示，例如、视觉地标、3D 地标、甚至语义对象。

1) 约束：顶点通过不同类型的边连接，这些边根据观察结果对其状态变量施加约束（例如、关键点、imu 测量值和回环）。IMU 边包含连接顶点之间的预积分 IMU 测量值，因此仅连接时间连续的顶点。相对姿势约束

边缘施加刚性 6 自由度 两者之间的信息  
 五埃蒂克并用于表示 相对运动 ( $I_E$ )  
 od奥德术或循环 关闭跨越较大时间间隔或  
 米伊西ns。边缘是 受让人 达科瓦里亚 噢丙去曲反对财年  
 合纳蚂蚁价值。c 奥康以使用 iance 来模拟程度  
 传感器可以观察到的运动，例如车轮里程表具有无限  
 丙卵子沿 z 轴运动，以及俯仰和直线运动  
 r 我们考虑循环 封闭边缘特定 亲属案件  
 位置约束边。为了提高稳健性并考虑

对于异常值，可以将回环边缘作为可切换约束 [22]。然后，如果边缘与其他约束冲突过多，优化器可以从图中丢弃这些边缘。最后，边缘将地标与观察到的姿态连接起来，并根据估计的地地标位置和观察到的地地标位置之间的差异施加误差。

在优化过程中，还可以直接对所选顶点的内部状态施加约束。例如，绝对约束在给定不确定性的情况下对顶点强制执行全局 3D 位置，并允许我们整合 GPS 测量值或绝对基准标记观测值。此外，固定某些状态可以灵活地选择要优化问题的哪些部分。

2) 地标：核心的视觉映射模块 地图实验室[11]仍然是地图实验室 2.0。它包括基于 ORB [23] 的特征检测，使用来自 BRISK [24] 或 FREAK [25] 的二进制描述符。基于描述符匹配建立连续帧之间的特征对应关系，其中出于稳健性考虑，匹配窗口受到集成陀螺仪测量的限制。然后将这些特征轨迹三角化为 3D 地标。

通过获取单个帧并使用特征描述符建立一组 2D-3D 匹配，可以实现全局定位和回环闭合。随后对匹配应用共视性检查以过滤异常值。然后，使用 RANSAC 方案中的 P3P 算法，使用剩余匹配来获得相对于地图参考框架的变换。然后可以将该变换作为回环闭合边添加到因子图中。我们还提供了一个

另一种方法是通过合并共视地标并最小化其重新投影误差来整合回环。这种方法可以避免调整显式回环的困难

重新边缘协方差 b ut 对  
 事实止图。

在地图实验室 2.0，我们添加了同时将任意数量的不同类型特征纳入地图的可能性。为了获得连续帧中的特征轨迹，我们  
 ers 可以使用 包括通用实现

Lucas-Kanade 跟踪器 [26] 或自己提供跟踪信息。此外，我们扩展了匹配引擎以支持浮点描述符，从而能够使用最新开发的描述符进行循环闭合。二进制描述符的匹配方式如下地图实验室，使用倒置多索引 [27]，而浮点描述符则使用近似最近邻快速库 (FLANN) [28] 进行匹配。然后，为了实现循环闭合，对匹配进行类似处理，如前所述。不过，出于调整目的，不同的特征类型可以具有单独的参数集，以解决质量和行为的差异。

地图实验室 2.0 还可以处理具有 3D 观测值的地标。这些地标可能来自 RGB-D 相机，其中视觉特征也具有相关的深度，或者来自直接在 3D 点云中检测到的特征。显着的区别是，这些地标的位置不是使用多视图几何进行三角测量，而是通过平均 3D 测量值进行三角测量。类似地，姿势图误差项不是基于重新投影误差，而是基于观察到的 3D 位置和地标的 3D 位置之间的欧几里得距离。另一个显着的区别是，回环被设置为 3D 到 3D RANSAC 匹配问题，没有 P3P 算法。

## B. 映射节点

映射节点在每个机器人上运行，并使用外部输入源和原始传感器数据以多模态因子图的形式创建地图。在地图构建过程中使用 6 DoF 里程计输入来初始化底层因子图的机器人姿势顶点。映射节点与里程计方法无关，具有简单的界面，因此可以轻松在各种机器人和传感器设置中使用。这与地图实验室 [11]，其中只有内置的视觉惯性估计器 ROVIOLI [29] 可用，而地图实验室 2.0 甚至不需要 IMU。但是，如果有 IMU，则将惯性约束添加到地图中，并且状态估计器还可以选择计算 IMU 偏差的初始估计。然后可以使用这些偏差估计来改进全局地图优化问题的初始化，这有利于其收敛速度和准确性。如果存在 IMU 但状态估计器未使用，则偏差估计也可以单独进行 [30]–[32]。

对原版进行重大修改 地图实验室 还制作了框架，以便可以使用自定义内部组件或易于配置的外部接口来处理和集成其他传感器模式。最值得注意的是，地图实验室 2.0 可以在运行时合并任意数量的不同 3D 地标类型。此外，相对约束（例如、里程计或外部环路闭合）和绝对 6 自由度约束（例如现在可以无缝添加位置信息（例如，GPS 或基准标记）。

原始摄像头图像或 LiDAR 点云可以作为资源附加到地图上，供后续模块随时使用，例如，用于计算额外的回环或检测物体。然后，可以将包含所有约束的结果地图传递到地图服务器进行在线处理，或存储并加载以供稍后在控制台中进行离线处理。

## C. 映射服务器

映射服务器是地图实验室 2.0，支持协作和在线制图。该方法已成功部署在 DARPA 地下挑战赛中，作为获胜团队 (CERBERUS) 多机器人制图系统的一部分 [33]。服务器节点可以与制图节点并行运行在专用机器或其中一个机器人上。制图节点定期将其地图分成块，称为子图。子图会立即传输到制图服务器，在那里进行预处理并与同一机器人先前传输的相应子图连接起来。在分割时，将每个子图的最后一个顶点复制到下一个子图中，从而进行簿记。这也避免了边和特征轨迹的不连续性。并行地，服务器不断循环关闭来自不同机器人的地图以形成全局一致的地图。值得注意的是，服务器和控制台共享相同的代码库，因此任何新功能都可以轻松集成到其中任何一个中。

1) 子图预处理：传入的子地图不会直接合并，而是先单独处理以确保局部准确性。具体来说，对每个机器人的子地图执行一组可配置的操作，其中包括局部地图优化（对所有传感器数据和约束进行全束调整）、特征质量评估和地图内环路闭合（视觉和激光雷达，取决于可用情况）。由于子地图彼此独立处理，因此映射服务器可以有效地同时处理多个地图。完成预处理步骤后，每个完成的子地图都会连接到来自同一机器人的上一个子地图。

2) 多机器人处理：映射服务器持续在全球多机器人地图上运行，并执行第二组可配置操作（循环闭合、特征质量评估、捆绑调整、可视化、绝对约束异常值拒绝、ETC。）。在这里，回环算法（视觉或激光雷达）尝试将所有不同的机器人置于同一参考系中并纠正漂移。与预处理步骤相比，多机器人地图上的操作始终在全局范围内执行，例如，采用机器人内和机器人间的方法检测回环，并对所有机器人地图进行全局优化。协作构建的全局地图也可以传回机器人，以提高其机载估计的准确性 [13]。对环境的认识增强不仅有利于定位准确性，还有利于其他任务，例如全局路径规划。

## D. 离线控制台

离线控制台从 地图实验室，旧工具已适应传感器和模态方面的新功能。有用于进一步处理地图的工具，例如批量优化、合并来自不同会话的地图、异常值拒绝、关键帧、地图稀疏化 [34]，ETC。现在，使用 LiDAR 进行回环闭合也成为可能，因为新模块包括 ICP [35] 和 G-ICP [36] 的实现，但不限于这些，而且可以轻松扩展。通过配准计算的变换

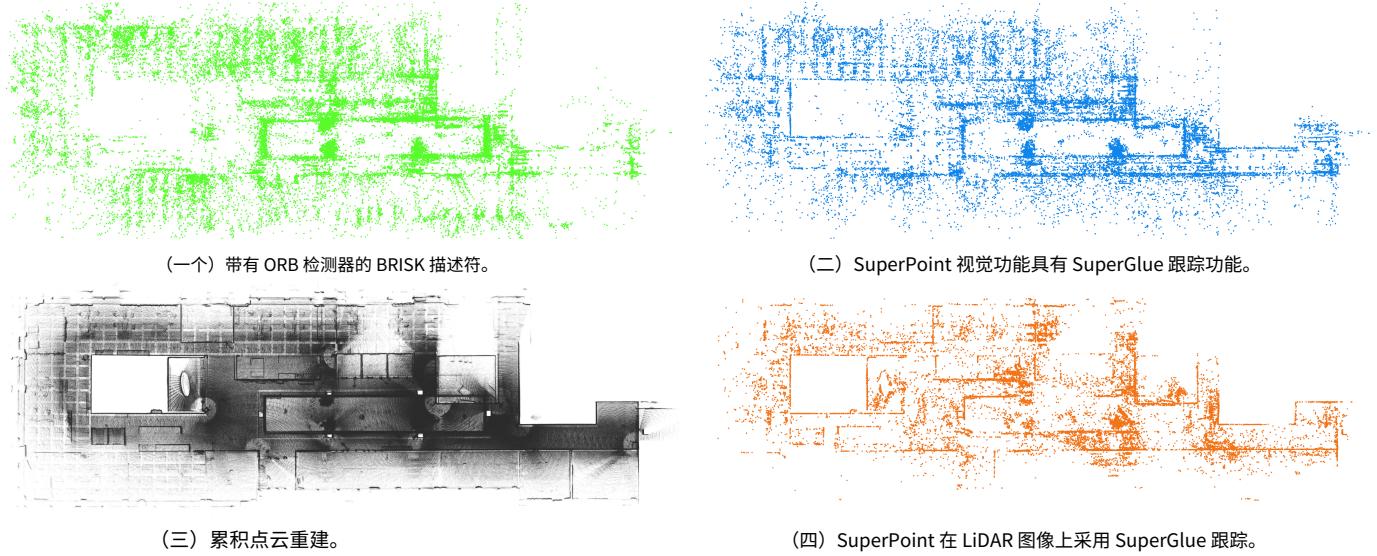


图3：特征和传感器数据的可视化地图实验室 2.0 在 HILTI 2021 数据集中的 Office Mitte 序列上，使用 OKVIS 和对特征进行全局捆绑调整。

模块被添加为具有可切换约束的回环闭合边（参见第 III-A 节）。对于每个传感器和方法组合，我们使用预定义的固定协方差，该协方差为每个平移和旋转分量单独设置。这些值是根据传感器噪声和配准方法的准确性经验选择的。密集重建也可以使用集成的 Voxblox [37] 插件完成。控制台还提供资源管理工具（操作或可视化附加的点云、图像和语义测量）或导出地图数据（姿势、IMU 偏差、地标、ETC.）。

最后，控制台可以通过插件轻松扩展，这些插件可以离线运行代码，并且独立于地图构建过程。例如，我们使用插件来实现上面提到的 LiDAR 注册模块和语义循环闭合模块（参见第 IV-D 节）。

#### IV. 乌东南-C东南大学

我们进行了多次实验，以全面评估我们提出的框架，并展示其易用性和高灵活性。具体来说，本节将展示四个数据集上的结果，以展示地图实验室 2.0。首先，我们在公共 HILTI SLAM 2021 数据集 [38] 上验证了我们提出的框架的性能和准确性，并将其与众所周知的最先进的方法进行了比较。接下来，我们展示了我们提出的框架在现实世界中的适用性，并展示了大规模多机器人多会话功能。然后，我们通过结合从投影点云中检测到的 3D LiDAR 特征来展示地标系统的多功能性。最后，我们在自定义室内数据集上展示了语义回环模块。所有数据集均使用硬件时间同步的传感器设置收集。

##### A. 验证与比较

我们使用 HILTI SLAM Challenge 2021 数据集 [38] 将我们提出的框架与最先进的方法进行比较。

该数据集包括 12 条记录，涵盖室内办公环境和具有挑战性的室外建筑工地。我们在实验中排除了三个太小且没有带来有趣挑战的序列。其余九个序列中的三对（建筑工地、地下室和校园）是在同一环境中拍摄的，可以共定位以提高准确性。

为了地图实验室 2.0，我们可以使用数据集中提供的五个摄像头、ADIS IMU 和 OSO-64 LiDAR。我们展示了三个用例地图实验室 2.0 使用三种不同的里程计源：ROVIO [29]、OKVIS [39] 和 FAST-LIO2 [40]。除了标准的 BRISK [24] 描述符外，我们还使用第 III-A2 节中的外部接口，包括 SuperGlue [41] 跟踪的 SuperPoint [12] 特征和 LK 跟踪 [26] 的 SIFT 特征 [42]。为了减小地图大小并加快描述符搜索速度，我们使用主成分分析 (PCA) 压缩 SuperPoint 和 SIFT 特征，256 佛罗里达州浮点 32. 使用所有可用特征计算全局回环，并合并匹配的地标。我们还使用来自 LiDAR 配准模块的 ICP [35]（参见第 III-D 节）来局部细化我们的姿势。回环边缘的协方差是根据经验预先定义的。序列可视化米特办公室如图 3a-c 所示。可以观察到，与 ORB 相比，SuperPoint 特征如何更好地遵循建筑结构。

表 II 还展示了里程计源单独以及其他 SLAM 基线 (LVI-SAM [43]、ORB-SLAM3 [8]、RTAB-Map [3] 和 地图实验室 [11])。地图实验室和地图实验室 2.0 是唯一能够使用所有五个摄像头进行回环闭合的方法。对于 ROVIO 和 OKVIS，我们仅使用正面摄像头或立体摄像头进行里程计。在所有使用视觉的方法中地图实验室 2.0 表现远超基线。仅使用 LiDAR-Inertial 的 FAST-LIO2 是最佳基线，甚至优于基于相同原理的视觉-LiDAR-惯性融合 LVI-SAM。然而，我们表明，我们也可以采用表现最佳的方法作为里程计，并进一步改进结果，尤其是在 停车处

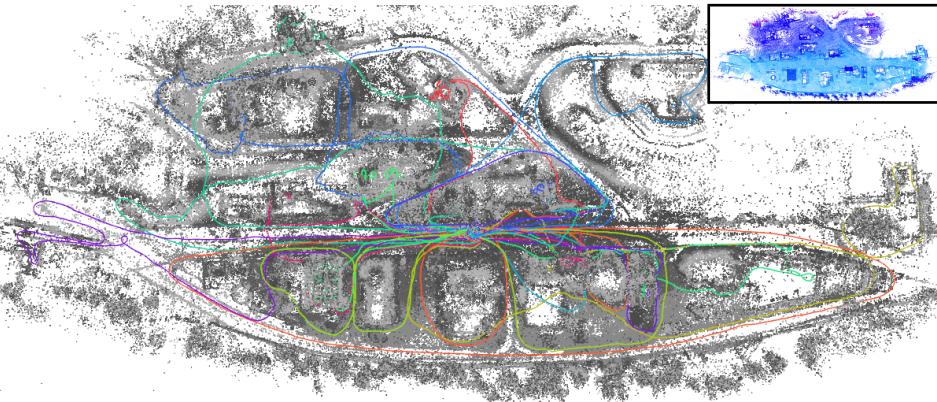


图4：全局多机器人地图的可视化测绘结果包含 23 次测绘运行。单个颜色表示机器人轨迹，灰点表示多机器人地图中的三角 BRISK 地标。右上图显示了通过将点云重新投影到优化姿势上而生成的 LiDAR 地图。

HILTI 2021 SLAM 数据集											
顺序	对象映射表 SLAM3	五十六山姆		实时同步广播 地图	地图实验室	ROVIO	奥科维斯	快速地 低碘酸	地图实验室 2.0		快充-LIO2 + SP + B
		图标	图标						图标	图标	
施工 1	1.55米	0.13米	0.36米	0.16米	0.98米	1.17米	0.04米	0.14米	0.08米	0.08米	0.04米
施工 2	2.77米	0.33米	0.67米	0.57米	1.50米	2.13米	0.07米	0.34米	0.19米	0.19米	0.07米
创意产业办公室	1.86米	0.12米	1.50米	0.09米	1.16米	1.27米	0.08米	0.08米	0.07米	0.07米	
米特办公室	1.70米	0.24米	0.94米	3.18米	0.86米	1.15米	0.12米	0.27米	0.18米	0.15米	0.10米
地下 3 层	1.55米	0.10米	0.38米	0.09米	3.05米	1.01米	0.05米	0.09米	0.09米	0.08米	0.05米
地下 4 层	1.71米	0.13米	0.38米	0.11米	2.90米	1.23米	0.04米	0.11米	0.10米	0.09米	0.04米
停车场	5.49米	4.43米	7.82米	0.39米	6.13米	3.36米	5.00米	0.31米	0.21米	0.21米	0.21米
校区 1	1.93米	0.12米	0.93米	0.60米	5.10米	2.41米	0.07米	0.38米	0.19米	0.17米	0.07米
校区 2	2.24米	0.14米	0.79米	0.47米	2.02米	2.23米	0.09米	0.28米	0.20米	0.18米	0.08米
总时间	61分钟	68分钟	163分钟	82分钟	58分钟	121分钟	52分钟	98分钟	236分钟	267分钟	194分钟

表二：绝对位置误差 (APE) 的 RMSE 与最先进方法的比较。SP + B 代表 SuperPoint 和 BRISK 视觉特征。图标代表所使用的传感器：单目、多摄像头、LiDAR 和 IMU。数据集的总时长 152 分钟。

序列。我们还展示了 ROVIO 和 SIFT 的融合，展示了 ROVIO 的多功能性地图实验室 2.0 实现快速增量改进，独立于更好的深度学习视觉特征。所有方法的计时也在配备 Intel i7-8700 和 Nvidia RTX 2080 GPU 的机器上呈现。

### B. 大规模多机器人多会话建图

我们通过在瑞士的一个大型训练设施中部署我们提出的框架来展示其对复杂现实场景的适用性。环境包括城市街道和建筑物，以及倒塌建筑物和狭窄空间等恶劣环境。在这次实验中，我们使用带有五个摄像头的手持设备和 Ouster OS0-128 记录了 23 次单独的运行，其中包含大约两个小时的数据 10 公里以及多个室内室外转换。每次运行都使用 OKVIS [39] 进行里程计。前五张地图用于构建全局多机器人地图，使用地图

服务器，以及其余的使用多会话合并地图 n 映射于 t 他安慰。所有任务之间的一致性是通过全局视觉环路闭合来强制执行的，如果可用，则使用 RTK GPS 的额外绝对姿态约束。此外，单个轨迹

通过执行任务内和任务间 LiDAR 配准来完善地图。图 4 显示了最终的多机器人地图。

为了定量评估多机器人服务器，我们在公共 EuRoC 基准 [44] 上进行了测试。由于控制台和服务器使用相同的底层映射框架（除了操作顺序等细节外），预期准确度相同。我们在多机器人实验中使用映射服务器、ROVIO 和 BRISK 同时运行所有 11 个序列。之后，我们重复实验，使用映射节点和控制台按顺序处理每个任务，然后将它们合并在一起。两种场景的平均 RMSE APE 均为 0.043 米，但并行映射服务器只需要 3 分钟 27 秒，而不是 35 分钟 56 秒用于顺序多会话工作流程。对于这两种场景，时间都包括里程计、优化和地图合并。

### C. 投影 LiDAR 图像上的视觉跟踪

为了展示地标系统的灵活性 地图实验室 2.0，我们整合了 3D LiDAR 关键点<sup>1</sup> 我们从 Streiff 的作品中汲取灵感等[45] 并投射

<sup>1</sup>请注意，可以为其他模式实施类似的工作流程，例如 RGB-D 摄像机、雷达或声纳成像。

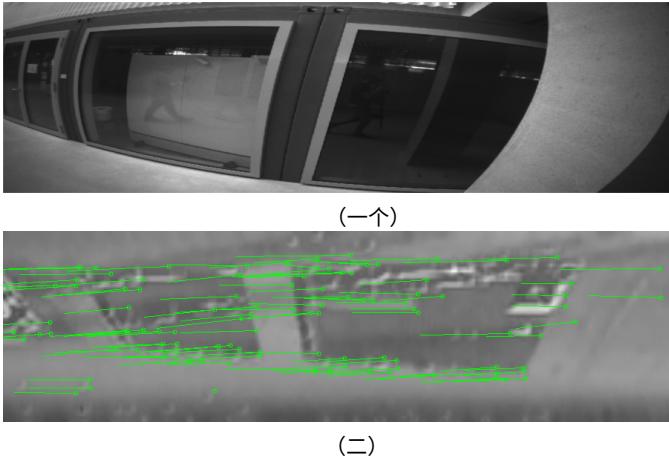


图5：跟踪 LiDAR 图像上的关键点。 (a) 仅显示相关相机图像，仅用于说明目的。 (b) 显示 LiDAR 图像（已裁剪以便于可视化）~40 °。其中绿色圆圈和线条代表 SuperPoint 检测以及它们对前一帧的跟踪运动。

LiDAR 点云到 2D 平面。我们使用对数刻度对 LiDAR 范围和强度值进行归一化，并使用 Mertens 融合 [46] 合并两个通道。使用相邻值修复不良 LiDAR 返回中缺失的像素。图 5 显示了生成的 2D 投影的示例图像，旁边是显示环境的同一视角的相机图像。然后，我们将 LiDAR 图像视为相机图像，并将 SuperPoint 与 SuperGlue 一起应用以获得点特征和轨迹，如图 5 所示。由于对于每个特征观察，我们都有来自 LiDAR 的深度信息，因此我们可以更有效地初始化和循环关闭这些 3D LiDAR 地标，如第 III-A2 节所述。最后，我们使用这些 LiDAR 关键点绘制出 HILTI 2021 数据集中的一个序列，并在图 3d 中可视化生成的地图。从墙壁的直线度可以看出，LiDAR 地标比视觉关键点更准确地映射到结构上。然而，它们也会受到由于缺失点或环境中的移动物体而导致的 LiDAR 图像噪声所造成的影响。

#### D. 基于语义的映射

本节展示了地图实验室 2.0 通过用语义信息扩充地图并说明其在现实世界场景中的潜在应用，我们实现了语义对象跟踪。最初，使用 Mask R-CNN [47] 在图像中检测语义对象，对于每次检测，我们使用 NetVLAD [48] 在掩码实例分割上提取描述符。我们使用 Deep SORT [49] 而不是内置跟踪器来跟踪所有检测到的对象，它使用外观项扩展了典型的空间数据关联度量，可以直接利用先前提取的对象描述符。与视觉地标类似，语义对象是空间中的 3D 地标地图实验室 2.0 但有一个关联的类标签，可以用于，例如，语义循环闭合检测。

最后，通过直接比较同一类的对象描述符来找到候选语义循环闭包。首先，

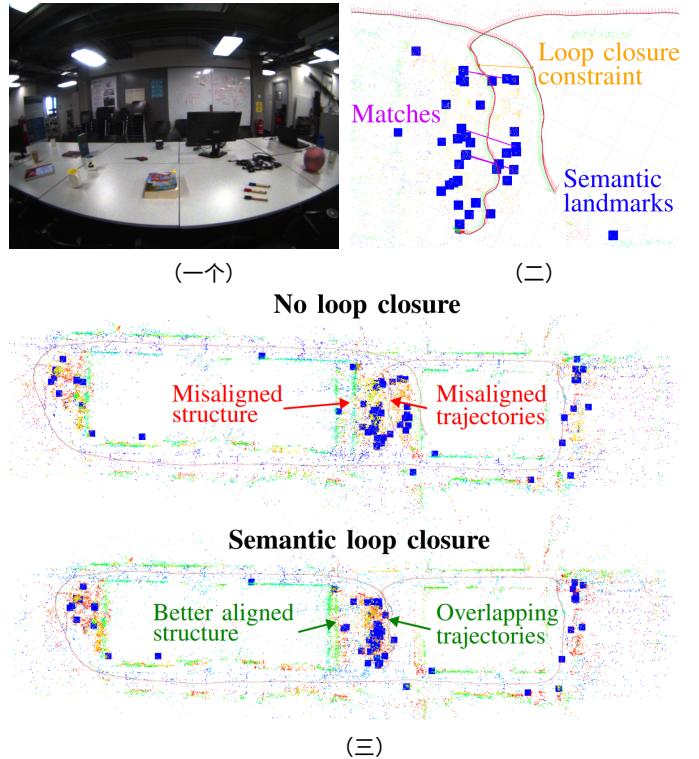


图6：语义映射管道。 (a) 实验设置：具有多个语义对象的表。 (b) 语义标志（蓝色）之间的回环匹配（洋红色），产生回环约束（橙色）。 (c) 语义回环前后的视觉惯性语义图。

应用独特的可见性过滤器， $IE$ ，单幅图像中观察到的两个地标无法匹配。在对候选者进行几何验证并对可共视地标进行聚类后，利用 Horn 的方法 [50] 获得的两个 3D 地标簇之间的相对坐标变换，在最接近两个匹配地标簇的两个机器人顶点之间构建 6-DoF 约束。最后，使用 Manoj 提出的方法计算相应因子图约束的协方差等[51]。

我们使用 RGB 惯性传感器 [52] 收集了办公室环境中具有多个物体的室内数据集。我们在两次观察带有物体的办公桌时留出一些时间来累积漂移（见图 6a 和 6c）。图 6b 显示了语义地标聚类和检测到的回环候选。将语义对象的回环边缘添加到全因子图后，漂移显著减少，改进后的地图可以在图 6c 中看到。

#### V、C 结论

我们提出了一个多模态和多机器人地图绘制研究平台，支持地图的在线和离线处理。我们在大规模 SLAM 基准测试和多个实验用例中展示了最先进的性能地图实验室 2.0。我们提出的映射框架的灵活和模块化设计促进了各种机器人的应用的研究，并在学术界和工业界产生了重要影响。重现实验的代码和教程可在存储库的 wiki 上找到。

## R参考文献

- [1] G. Bresson、Z. Alsayed、L. Yu 和 S. Glaser, “同步定位与地图绘制：自动驾驶当前趋势调查”，第四型，第 2 卷，第 3 期，第 194-220 页，2017 年。
- [2] K. Blomqvist、M. Breyer、A. Cramariuc、J. Förster、M. Grinvald, F. Tschopp、JJ Chung、L. Ott、J. Nieto 和 R. Siegwart, “Go Fetch：非结构化环境中的移动操作”，感知、行动、学习研讨会，ICRA，2020 年。
- [3] M. Labb   和 F. Michaud, “RTAB-Map 作为用于大规模和长期在线运行的开源激光雷达和视觉同步定位与地图绘制库”，杰弗瑞，第 36 卷，第 2 期，第 416-446 页，2019 年。
- [4] A. Rosinol、M. Abate、Y. Chang 和 L. Carlone, “Kimera：一个用于实时度量语义定位和映射的开源库”，免疫学与细胞免疫学协会，5 2020，第 1689-1696 页。
- [5] Y. Tian、Y. Chang、F. Herrera Arias、C. Nieto-Granda、J. How 和 L. Carlone, “Kimera-Multi：适用于多机器人的稳健、分布式、密集度量语义 SLAM”，反渗透，2022 年。
- [6] Y. Chang、K. Ebadi、CE Denniston、MF Ginting、A. Rosinol、A. Reinke、M. Palieri、J. Shi、A. Chatterjee、B. Morrell、等，“灯 2.0：适用于具有挑战性的大规模地下环境的强大机器人 SLAM 系统”，右心室，2022 年。
- [7] M. Karrer、P. Schmuck 和 M. Chli, “CVI-SLAM——协作视觉惯性 SLAM” 右心室，第 3 卷，第 4 期，第 2762-2769 页，2018 年。
- [8] C. Campos、R. Elvira、JJG Rodr  guez、JM Montiel 和 JD Tard  os, “ORB-SLAM3：用于视觉、视觉惯性和多地图 SLAM 的精确开源库”，反渗透，2021 年。
- [9] P. Schmuck、T. Ziegler、M. Karrer、J. Perraudin 和 M. Chli, “COVINS：用于集中协作的视觉惯性 SLAM”，ISMAR-兼职，2021 年，第 171-176 页。
- [10] 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、在线和异常弹性 SLAM”，右心室，第 5 卷，第 2 期，第 1656-1663 页，2020 年。
- [11] T. Schneider、M. Dymczyk、M. Fehr、K. Egger、S. Lynen、I. Gilitschenski 和 R. Siegwart, “maplab：视觉惯性测绘与定位研究的开放框架”，右心室，第 3 卷，第 3 期，第 1418-1425 页，2018 年 11 月。
- [12] D. Detone、T. Malisiewicz 和 A. Rabinovich, “SuperPoint：自监督兴趣点检测与描述”，CVPR 研讨会，12 2018，第 337-349 页。
- [13] L. Bernreiter、S. Khattak、L. Ott、R. Siegwart、M. Hutter 和 C. Cadena, “使用光谱图分析进行协作机器人绘图”，免疫学与细胞免疫学协会，2022 年。
- [14] C. Cadena、L. Carlone、H. Carrillo、Y. Latif、D. Scaramuzza、J. Neira、I. Reid 和 JJ Leonard, “同步定位与地图绘制的过去、现在和未来：迈向稳健感知时代”，T-RO，第 32 卷，第 6 期，第 1309-1332 页，2016 年。
- [15] R. Mur-Artal 和 JD Tard  os, “ORB-SLAM2：适用于单目、立体和 RGB-D 相机的开源 SLAM 系统”，反渗透，第 1-8 页，2017 年。
- [16] M. Labb   和 F. Michaud, “室内环境中光照不变重新定位的多会话视觉 SLAM”，机器人和人工智能前沿，第 115 页，2022 年。
- [17] R. Dub  、A. Cramariuc、D. Dugas、H. Sommer、M. Dymczyk、J. Nieto、R. Siegwart 和 C. Cadena, “SegMap：使用数据驱动描述符进行基于段的映射和定位”，伊利诺伊，第 39 卷，第 2-3 期，第 339-355 页，2020 年。
- [18] A. Gawel、C. Del Don、R. Siegwart、J. Nieto 和 C. Cadena, “X-View：基于图的语义多视图定位”，右心室，第 3 卷，第 3 期，第 1687-1694 页，2017 年。
- [19] A. Cramariuc、F. Tschopp、N. Alatur、S. Benz、T. Falck、M. Br  hlmeier、B. Hahn、J. Nieto 和 R. Siegwart, “SemSegMap - 基于 3D 片段的语义定位”，IROS，2021 年。
- [20] F. Taubner、F. Tschopp、T. Novkovic、R. Siegwart 和 F. Furrer, “LCD - 位置识别的线聚类和描述”，载于 3DV，2020 年，第 908-917 页。
- [21] L. Bernreiter、A. Gawel、H. Sommer、J. Nieto、R. Siegwart 和 C. Cadena, “稳健数据关联的多假设语义映射”，右心室，第 4 卷，第 4 期，第 3255-3262 页，2019 年。
- [22] N. S  nderhauf 和 P. Protzel, “稳健姿态图 SLAM 的可切换约束”，IROS，2012 年，第 1879-1884 页。
- [23] E. Rublee、V. Rabaud、K. Konolige 和 G. Bradski, “ORB：SIFT 或 SURF 的有效替代方案”，独立控制计算机视觉，2011 年，第 2564-2571 页。
- [24] S. Leutenegger、MM Chli 和 RY Siegwart, “二元稳健不变可扩展关键点”，独立控制计算机视觉，2011 年，第 2548-2555 页。
- [25] A. Alahi、R. Ortiz 和 P. Vandergheynst, “FREAK：快速视网膜关键点”，计算机视觉与图像处理，2012 年，第 510-517 页。
- [26] BD Lucas 和 T. Kanade, “一种应用于立体视觉的迭代图像配准技术”，影像，第 130 卷，第 674-679 页，1981 年。
- [27] S. Lynen、T. Sattler、M. Bosse、J. Hesch、M. Pollefeys 和 R. Siegwart, “走出我的实验室：大规模实时视觉惯性定位”，在 RSS，2015 年，第 18 页。
- [28] M. Muja 和 DG Lowe, “具有自动算法配置的快速近似最近邻”。维萨普，卷。2、没有。331-340, p. 2, 2009。
- [29] M. Bloesch、M. Burri、S. Omari、M. Hutter 和 R. Siegwart, “基于迭代扩展卡尔曼滤波器的直接光度反馈视觉惯性里程计”，伊利诺伊，第 36 卷，第 10 期，第 1053-1072 页，2017 年。
- [30] S. Lynen、MW Achtelik、S. Weiss、M. Chli 和 R. Siegwart, “一种适用于 MAV 导航的稳健模块化多传感器融合方法”，在 IROS，2013 年，第 3923-3929 页。
- [31] SO Madgwick、AJ Harrison 和 R. Vaidyanathan, “使用梯度下降算法估计 IMU 和 MARG 方向”，异丙醇，2011 年。
- [32] RG Valenti、I. Dryanovski 和 J. Xiao, “保持良好心态：基于四元数的 IMU 和 MARG 方向滤波器”，传感器，第 15 卷，第 8 期，第 19 302-19 330 页，2015 年。
- [33] M. Tranzatto、T. Miki、M. Dharmadhikari、L. Bernreiter、M. Kulkarni、F. Mascalich、O. Andersson、S. Khattak、M. Hutter、R. Siegwart 和 K. Alexis, “DARPA 地下挑战中的 CERBERUS” 科学机器人，第 7 卷，第 66 期，2022 年。
- [34] M. Dymczyk、S. Lynen、T. Cieslewski、M. Bosse、R. Siegwart 和 P. Furgale, “地图的要点——总结终身定位的经验”，免疫学与细胞免疫学协会，2015 年，第 2767-2773 页。
- [35] F. Pomerleau、F. Colas、R. Siegwart 和 S. Magnenat, “在真实数据集上比较 ICP 变体开源库和实验协议”，自主机器人，第 34 卷，第 3 期，2014 年。
- [36] A. Segal、D. Haehnel 和 S. Thrun, “Generalized-ICP”，机器人技术：科学与系统，卷。2、没有。2009 年 4 月，第 4 页。435。
- [37] H. Oleynikova、Z. Taylor、M. Fehr、R. Siegwart 和 J. Nieto, “Voxblox：用于机载 MAV 规划的增量式 3D 欧几里得有向距离场”，IROS，2017 年，第 1366-1373 页。
- [38] M. Helmberger、K. Morin、B. Berner、N. Kumar、G. Cioffi 和 D. Scaramuzza, “喜利得大满贯挑战数据集” 右心室，2022 年。
- [39] S. Leutenegger、S. Lynen、M. Bosse、R. Siegwart 和 P. Furgale, “基于关键帧的非线性优化视觉惯性里程计”，伊利诺伊，第 34 卷，第 3 期，第 314-334 页，2015 年。
- [40] W. Xu、Y. Cai、D. He, J. Lin, 以及 F. Zhang, “FAST-LIO2：快速直接激光雷达-惯性里程计” 反渗透，第 38 卷，第 4 期，第 2053-2073 页，2022 年。
- [41] P.-E. Sarlin、D. DeTone、T. Malisiewicz 和 A. Rabinovich, “SuperGlue：使用图神经网络学习特征匹配”，肺动脉高压，2020 年，第 4938-4947 页。
- [42] DG Lowe, “通过局部尺度不变特征进行对象识别” 独立控制计算机视觉，第 2 卷，1999 年，第 1150-1157 页。
- [43] T. Shan、B. Englot、C. Ratti 和 D. Rus, “LVI-SAM：通过平滑和制图实现紧密耦合的激光雷达-视觉-惯性里程计”，免疫学与细胞免疫学协会，2021 年，第 5692-5698 页。
- [44] M. Burri、J. Nikolic、P. Gohl、T. Schneider、J. Rehder、S. Omari、MW Achtelik 和 R. Siegwart, “EuRoC 微型飞行器数据集”，伊利诺伊，第 35 卷，第 10 期，第 1157-1163 页，2016 年。
- [45] D. Streiff、L. Bernreiter、F. Tschopp、M. Fehr 和 R. Siegwart, “3D3L：用于 LiDAR 的深度学习 3D 关键点检测和描述”，免疫学与细胞免疫学协会 IEEE，2021 年，第 13 064-13 070 页。
- [46] T. Mertens、J. Kautz 和 F. Van Reeth, “曝光融合”，载于太平洋图形公司 IEEE，2007 年，第 382-390 页。
- [47] K. He、G. Gkioxari、P. Dollar 和 R. Girshick, “Mask R-CNN”，独立控制计算机视觉，2017 年，第 2980-2988 页。
- [48] R. Arandjelovic、P. Gronat、A. Torii、T. Pajdla 和 J. Sivic, “NetVLAD：用于弱监督地点识别的 CNN 架构”，帕金森病，第 40 卷，第 6 期，第 1437-1451 页，2018 年。
- [49] N. Wojke、A. Bewley 和 D. Paulus, “使用深度关联指标进行简单的在线和实时跟踪”，集成电路设计专业委员会，2017，第 3645-3649 页。
- [50] BK Horn, “利用单位四元数实现绝对方向的闭式解”，约萨，第 4 卷，第 4 期，第 629-642 页，1987 年。
- [51] PS Manoj、L. Bingbing、Y. Rui 和 W. Lin, “3D ICP 协方差的闭式估计”，机动车增值业务，日本东京，2015 年。
- [52] F. Tschopp、M. Riner、M. Fehr、L. Bernreiter、F. Furrer、T. Novkovic、A. Pfunder、C. Cadena、R. Siegwart 和 J. Nieto, “VersaVIS - 开放式多功能多摄像头视觉惯性传感器套件”，传感器，卷。20、没有。5, p. 1439, 3 2020.