

LVI-SAM：通过平滑和映射的紧密耦合的激光雷达-视觉-惯性测距法

Tixiao Shan, Brendan Englot, Carlo Ratti, and Daniela Rus

rus}@mit.edu。

摘要

我们提出了一个通过平滑和映射实现紧密耦合的激光雷达-视觉-惯性测距的框架，即LVI-SAM，它可以实现实时的状态估计和具有高精度和稳健性的地图构建。LVI-SAM建立在一个因子图之上，由两个子系统组成：一个视觉-惯性系统（VIS）和一个激光雷达-惯性系统（LIS）。这两个子系统是以紧密耦合的方式设计的，其中VIS利用LIS的估计来促进初始化。通过使用激光雷达测量提取视觉特征的深度信息，VIS的精度得到了提高。反过来，LIS利用VIS估计的初始猜测来支持扫描匹配。环路闭合首先由VIS识别，然后由LIS进一步完善。当两个子系统中的一个出现故障时，LVI-SAM也能发挥作用，这增加了它在无纹理和无特征环境中的稳健性。LVI-SAM在各种规模和环境从几个平台收集的数据集上进行了广泛的评估。我们的实现可在<https://git.io/lvi-sam>。

I. 简介

同步定位和测绘（SLAM）是许多移动机器人导航任务所需的一种能力。在过去的20年里，利用SLAM在具有挑战性的环境中通过单一的感知传感器，如激光雷达或照相机，进行实时状态估计和测绘，已经取得了巨大的成功。基于激光雷达的方法可以在远距离捕获环境的精细细节。然而，这种方法在无结构的环境中操作时通常会失败，如长长的走廊或平坦的空地。尽管基于视觉的方法特别适用于地点识别，并且在纹理丰富的环境中表现良好，但其性能对光照变化、快速运动和初始化很敏感。因此，基于激光雷达的方法和基于视觉的方法通常都与一个惯性测量单元（IMU）结合在一起，以提高它们的再稳定性和准确性。激光雷达-惯性系统可以帮助纠正点云失真，并说明在短时间内缺乏特征。度量衡和姿态可以通过IMU测量恢复，以协助视觉-惯性系统。为了进一步提高系统性能，激光雷达、照相机和IMU测量的融合正在引起越来越多的关注。

我们的工作与视觉-惯性测绘（VIO）、激光雷达-惯性测绘（LIO）和激光雷达-视觉-惯性测绘（LVIO）关系最为密切。我们注意到，我们在本文中不考虑非惯性系统，尽管我们知道

T. Shan和C. Ratti在美国麻省理工学院城市研究与规划系，{shant, ratti}@mit.edu。

B. 恩格罗特是史蒂文斯学院机械工程系的学生。技术，美国，benglot@stevens.edu。

T. Shan和D. Rus是美国麻省理工学院计算机科学与人工智能实验室的成员，{shant, rus}@mit.edu。

有一些成功的非惯性激光雷达-视觉系统，如[1], [2]。视觉-惯性测距（VIO）可分为两大类：基于滤波器的方法和基于优化的方法。基于滤波器的方法通常使用前卡尔曼滤波器（EKF）来传播系统状态，使用来自摄像机和IMU的测量值。基于优化的方法保持一个滑动窗口估计器，并将视觉重投影误差和IMU测量误差降到最低。在我们的工作中，我们只考虑到单眼相机。在最流行的公开可用的VIO管道中，MSCKF[3]、ROVIO[4]和OpenVINS[5]是基于滤波器的，而OKVIS[6]和VINS-Mono[7]是基于优化的。尽管OKVIS在使用立体相机时表现出卓越的性能，但它没有针对单眼相机进行优化。VINS-Mono在滑动窗口设置中进行了非线性优化，并在单眼相机上达到了最先进的精度[8]。

根据他们的设计方案，激光雷达-惯性测距也可以分为两大类：松散耦合的方法和紧密耦合的方法。LOAM[9]和LeGO-LOAM[10]是松耦合系统，因为在优化步骤中不使用IMU的测量值。紧密耦合系统，通常提供更好的精度和鲁棒性，是目前研究的重点[11]。在公开的紧耦合系统中，LIO-mapping[12]适应了[7]的优化管道，并使IMU和激光雷达测量的残差最小。由于LIO-mapping被设计为优化所有的测量值，因此不能实现实时性能。LIO-SAM[13]，通过引入激光雷达关键帧的滑动窗口来限制计算的复杂性，利用因子图进行IMU和激光雷达的联合约束优化。专门为地面车辆设计的LINS[14]使用误差状态的卡尔曼滤波器来递归地修正机器人的状态。

最近，激光雷达-视觉-惯性系统由于其传感器退化任务中的鲁棒性而引起了越来越多的关注[15]。[16]提出了一个紧密耦合的LVIO系统，其特点是有一个顺序处理管道，从粗到细解决状态估计问题。粗略的估计从IMU预测开始，然后由VIO和LIO进一步细化。[16]目前在KITTI基准测试中达到了最先进的精度[17]。基于MSCKF的框架，[18]具有在线空间和时间多传感器校准的特点。[16]和[18]的实现并不公开。我们的工作与上述工作不同，我们利用因子图进行全局优化，可以通过循环闭合检测定期消除机器人发生的漂移。

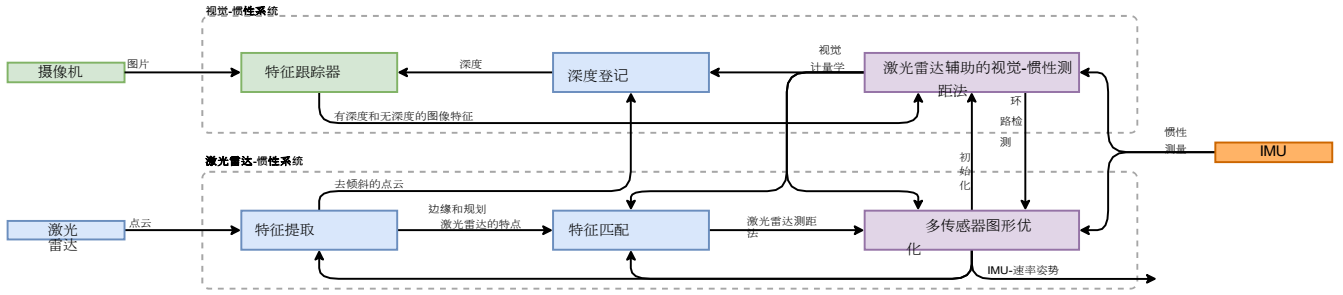


图1：LVI-SAM的系统结构。该系统接收来自三维激光雷达、照相机和IMU的输入，可分为两个子系统：视觉-惯性系统（VIS）和激光雷达-惯性系统（LIS）。视觉-惯性系统和激光雷达-惯性系统可以独立运作，同时利用彼此的信息来提高系统的准确性和稳健性。该系统以IMU速率输出姿态估计。

在本文中，我们提出了一个通过平滑和映射的紧密耦合的激光雷达-视觉-惯性测距框架，即LVI-SAM，用于实时状态估计和映射。LVI-SAM建立在一个因子图之上，由两个子系统组成，一个视觉-惯性系统（VIS）和一个激光雷达-惯性系统（LIS）。当其中一个系统检测到故障时，这两个子系统可以独立运作，或者当检测到足够的特征时，可以联合运作。VIS执行视觉特征跟踪，并可选择使用激光雷达帧提取特征深度。视觉测距，是通过以下方式获得的

优化视觉重投影和IMU测量的误差，作为激光雷达扫描匹配的初始猜测，并将约束引入因子图。在使用IMU测量对点云进行去偏后，LIS提取激光雷达的边缘和平面特征，并将其与保持在一个滑动窗口中的特征图相匹配。LIS中的估计系统状态可以被发送到VIS中，以促进其初始化。对于循环闭合，候选匹配首先由VIS识别，并由LIS进一步优化。来自视觉测距、激光雷达测距、IMU预集成和环路闭合的约束在因子图中被联合优化。最后，优化的IMU偏置项被用来传播IMU测量，以IMU速率进行姿态估计。我们工作的主要贡献可以归纳为以下几点。

- 一个建立在因素图之上的紧密耦合的LVIO框架，在地点识别的帮助下，实现了多传感器融合和全局优化。
- 我们的框架通过故障检测绕过失败的子系统，使其对传感器的退化具有鲁棒性。
- 我们的框架通过在不同规模、平台和环境中收集的数据得到了广泛的验证。

从系统的角度来看，我们的工作是最先进的，代表了VIO和LIO的最先进技术的独特整合，以实现LVIO系统提供更好的稳健性和准确性。我们希望我们的系统可以作为一个坚实的基线，其他人可以很容易地在此基础上推进激光雷达-视觉-惯性测距的最先进水平。

II. 通过激光雷达视觉惯性测距法平滑化和制图

A. 系统概述

拟议的激光雷达-视觉-惯性系统的概述，该系统接收来自三维激光雷达、单眼相机的输入。

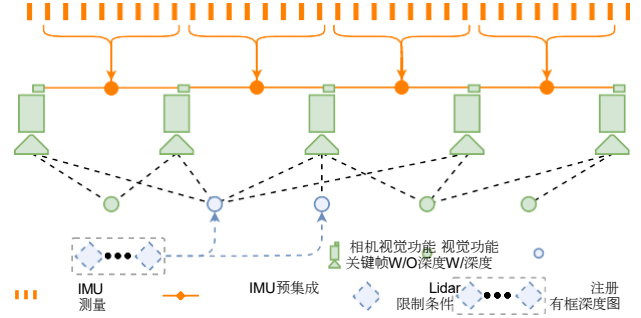


图2：我们的视觉-惯性系统的框架。该系统优化了来自IMU预集成、无深度视觉测量和有深度视觉测量的残差。

图1显示了我们的框架：一个视觉惯性系统和一个IMU。我们的框架由两个关键的子系统组成：一个视觉-惯性系统（VIS）和一个激光雷达-惯性系统（LIS）。视觉惯性系统处理图像和IMU测量，而激光雷达测量是可选的。通过最小化视觉和IMU测量的联合残差来获得视觉测距。LIS提取激光雷达特征，并通过将提取的特征与特征图相匹配来执行激光雷达测距。该特征图以滑动窗口的方式进行维护，以实现实时性能。最后，状态估计问题可以被表述为最大后验（MAP）问题，通过使用iSAM2[19]联合优化因子图中IMU预集成约束、视觉测距约束、激光测距约束和环路闭合约束的贡献来解决。请注意，LIS中采用的多传感器图的优化是为了减少数据交换，提高系统效率。

B. 视觉-惯性系统

我们将文献[7]中的处理流程用于我们的VIS，如图2所示。视觉特征是用角检测器[20]检测出来的，并由Kanade-Lucas-Tomasi算法[21]进行跟踪。在VIS初始化时，我们使用视觉测距法注册激光雷达帧，并获得一个稀疏的深度图像用于特征深度估计。

系统在滑动窗口设置中进行捆绑调整，系统状态 $\mathbf{x} \in \mathbf{X}$ 可以写成。

$$\mathbf{x} = [\mathbf{R}, \mathbf{p}, \mathbf{v}, \mathbf{b}]$$

$\mathbf{R} \in \text{SO}(3)$ 是旋转矩阵， $\mathbf{p} \in \mathbb{R}^3$ 是位置向量， \mathbf{v} 是速度， $\mathbf{b} = [\mathbf{b}_a, \mathbf{b}_w]$ 是IMU的偏置。

\mathbf{b}_a 和 \mathbf{b}_w 分别是加速度和角速度的偏置矢量。从传感器体框架 \mathbf{B} 到世界框架 \mathbf{W} 的变换 $\mathbf{T}_{SE(3)}$ 被表示为 $\mathbf{T} = [\mathbf{R} \ \mathbf{p}]$ 。在下面的章节中，我们给出了改进VIS初始化和特征深度估计的详细程序。由于篇幅有限，我们请读者参考[7]以了解更多细节，如残差的实现。

1) 初始化。基于优化的VIO通常会受到影响

由于在初始化时解决了一个高度非线性的问题，导致分歧。初始化的质量在很大程度上取决于两个因素：初始传感器的运动和IMU参数的准确性。在实践中，我们发现[7]在传感器以小速度或恒定速度行驶时，往往不能初始化。这是由于当加速度激励不够大时，度量衡是无法观察到的。IMU的参数包括缓慢变化的偏差和白噪声，它们对原始加速度和角速度测量都有影响。在初始化时对这些参数的良好猜测有助于优化的快速收敛。

为了提高我们的VIS初始化的稳健性，我们利用了来自LIS的估计系统状态 \mathbf{x} 和IMU偏置 \mathbf{b} 。因为深度是可以从激光雷达上直接观测到的，所以我们首先初始化LIS并获得 \mathbf{x} 和 \mathbf{b} ，然后我们根据图像的时间戳将它们插值并与每个图像关键帧相关联。注意，IMU的偏差被假定为在两个图像关键帧之间是恒定的。最后，来自LIS的估计的 \mathbf{x} 和 \mathbf{b} 被用作VIS初始化的初始猜测，这大大改善了初始化速度和鲁棒性。有和没有LIS帮助的VIS初始化的比较可以在补充视频中找到¹。

2) 特征深度关联。在VIS的初始化过程中，我们使用估计的视觉测距法将激光雷达帧注册到摄像机帧上。由于现代三维激光雷达经常产生稀疏的扫描，我们将多个激光雷达帧堆叠起来以获得密集的深度图。为了将一个特征与一个深度值联系起来，我们首先将视觉特征和激光雷达深度点投射到以摄像机为中心的单位球体上。然后对深度点进行降采样，并使用它们的极坐标存储在球体上的恒定密度。我们通过使用视觉特征的极坐标搜索二维K-

D树，为一个视觉特征找到球体上最近的三个深度点。最后，特征深度是由视觉特征和摄像机中心 O_c 形成的线的长度，它与笛卡尔空间中三个深度点形成的平面相交。图3(a)是对这一过程的可视化描述，其中特征深度是虚线的长度。

我们通过检查最近的三个深度点之间的距离来进一步验证相关的特征深度。这是因为堆叠不同时间戳的激光雷达帧可能会导致不同物体的深度不明确。图3(b)是这种情况的一个例子。在时间 t_i ，观察到的深度点被描述为绿色。在 t_j ，摄像机被移动到一个新的位置，并观察到新的深度。

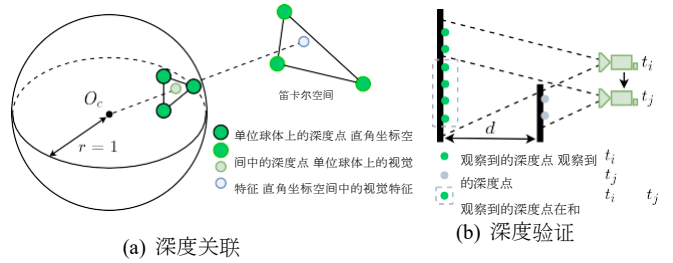


图3：视觉特征深度关联。

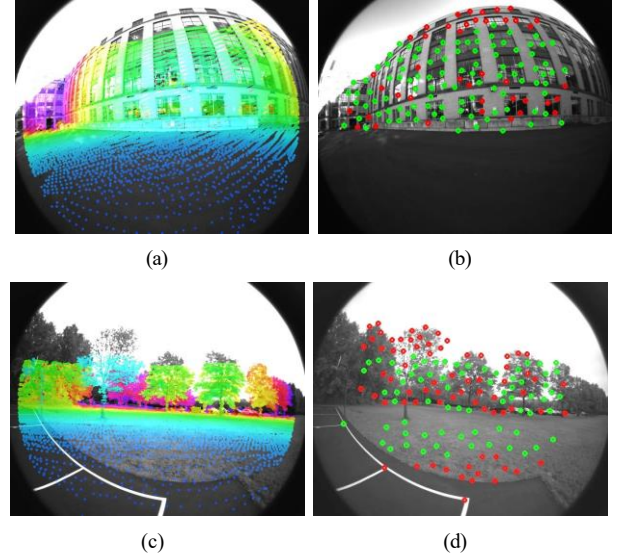


图4：注册深度图和视觉特征。在 (a) 和 (c) 中，深度图的颜色变化表示深度变化。在 (b) 和 (d) 中，绿色点是与深度成功关联的视觉特征。深度关联过程失败的特征为红色。

灰色的点。然而，在 t_i 的深度点，由灰色虚线圈出，由于激光雷达帧的堆叠，在 t_j ，可能仍然可以观察到。使用来自不同物体的深度点来关联特征深度会导致不准确的估计。与[16]类似，我们通过检查一个特征的深度点之间的最大距离来拒绝这种估计。如果最大距离大于2米，则该特征没有深度关联。

图4显示了注册的深度图和视觉特征的演示。在图4 (a) 和 (c) 中，使用视觉测距法登记的深度点被投影到摄像机图像上。在图4 (b) 和 (d) 中，与深度成功关联的视觉特征被染成绿色。请注意，虽然深度图覆盖了图4 (a) 中的大部分图像，但4 (b) 中许多位于窗口角落的特征由于验证检查失败而缺乏深度关联。

3) 故障检测。由于激烈的运动、光照变化和无纹理的环境，VIS遭受了失败。当机器人经历了激烈的运动或进入无纹理的环境时，跟踪的特征数量会大大减少。不足的特征可能会导致优化分歧。我们还注意到，当VIS失败时，会估计出一个大的IMU偏差。因此，当跟踪的特征数量低于一个阈值时，我们报告VIS失败。

¹<https://youtu.be/8CT107D6lbc>

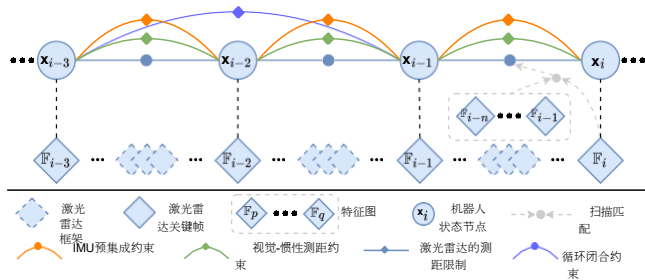


图5：我们的激光雷达-惯性系统的框架。该系统维护着一个有四种类型约束的因子图。

或当估计的IMU偏差超过一个阈值时。对我们的系统来说，主动的故障检测是必要的，这样它的故障就不会破坏LIS的功能。一旦检测到故障，VIS将重新初始化并通知LIS。

4) 循环闭合检测。我们利用DBoW2[22]进行循环闭合检测。对于每个新的图像关键帧，我们提取BRIEF描述符[23]，并与之前提取的描述符进行匹配。由DBoW2返回的循环闭合候选图像的时间戳被发送到LIS进行进一步验证。

C. 激光雷达-惯性系统

如图5所示，所提出的激光雷达-惯性系统，改编自[13]，为全局姿态优化保持一个因子图。四种类型的约束，即IMU预集成约束、视觉测距约束、激光雷达测距约束和环路闭合约束，被添加到图中并共同优化。激光雷达测距约束来自于扫描匹配，我们将当前激光雷达关键帧与全球特征图相匹配。循环闭合约束的候选人首先由VIS提供，然后通过扫描匹配进一步优化。我们为特征图保持一个滑动的激光雷达关键帧窗口，这保证了计算复杂性的约束。当机器人姿势的变化超过一个阈值时，就会选择一个新的激光雷达关键帧。位于关键帧对之间的间歇性激光雷达帧被丢弃。在选择一个新的激光雷达关键帧时，一个新的机器人状态 \mathbf{x} 作为一个节点被添加到因子图中。以这种方式添加关键帧，不仅实现了内存消耗和地图密度之间的平衡，而且还有助于保持一个相对稀疏的因素图，以便进行实时优化。由于篇幅有限，我们请读者参考[13]的实现细节。在下面的章节中，我们重点讨论提高系统鲁棒性的新程序。

1) 初始猜测：我们发现初始猜测对扫描匹配的成功起着至关重要的作用，特别是当传感器发生激烈的运动时。初始猜测的来源在LIS初始化之前和之后是不同的。

在LIS初始化之前，我们假设机器人从一个静态位置开始，速度为零。然后，我们对原始IMU测量值进行整合，假设偏差和噪声为零值。两个激光雷达关键帧之间的综合平移和旋转变换产生了扫描匹配的初始猜测。我们发现这种方法可以成功地初始化

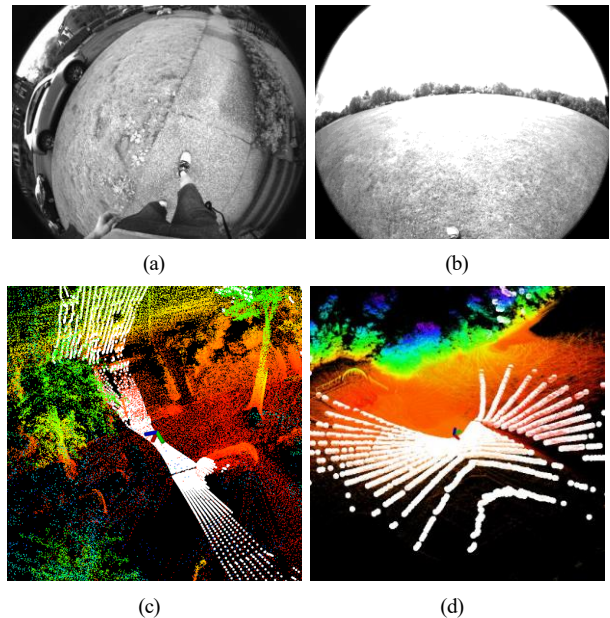


图6：扫描匹配不受限制的退化制图场景。在(a)和(c)中，激光雷达是面向地面放置的。在(b)和(d)中，激光雷达处于一个平坦、开放、无结构的环境中。白色的点表示激光雷达扫描的内容。颜色的变化表示海拔的变化。

当初始线速度小于10米/秒，角速度小于 $180^\circ/\text{s}$ 时，系统在挑战性的条件下。一旦LIS被初始化，我们在因子图中估计IMU偏差、机器人姿势和速度。然后，我们将它们发送给VIS以帮助其初始化。

在LIS初始化之后，我们可以从两个来源获得初始猜测：带有修正偏差的综合IMU测量值，以及VIS。当视觉-惯性测距法可用时，我们使用它作为初始猜测。如果VIS报告失败，我们就改用IMU测量作为初始猜测。这些程序提高了初始猜测的准确性和在富于纹理和没有纹理的环境中的稳健性。

2) 故障检测。尽管激光雷达可以在远距离捕获环境的精细细节，但它仍然会遇到扫描匹配受到限制的退化情况。图6中描述了这种情况。我们将[24]中的方法用于LIS故障检测。扫描匹配中的非线性优化问题可以被表述为迭代地解决一个线性问题。

$$\min_{\mathbf{T}} \|\mathbf{A}\mathbf{T} - \mathbf{b}\|_2. \quad (1)$$

其中 \mathbf{A} 和 \mathbf{b} 是在 \mathbf{T} 处线性化得到的。当 \mathbf{A} 的最小特征值 $\lambda_{\min}(\mathbf{A})$ 小于优化的第一次迭代的阈值时，LIS报告失败。当失败发生时，激光雷达测距约束不会被添加到因子图中。关于这些假设所依据的详细分析，我们请读者参考[24]。

III. 实验

我们现在描述了一系列的实验，在三个自我收集的数据集上验证了所提出的框架。

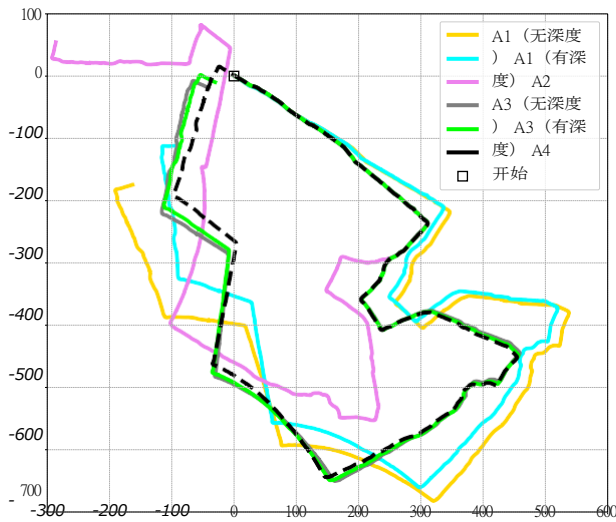


图7：使用城市数据集的消融研究轨迹。

表一：端对端翻译和旋转误差。

错误类型	A1 (不含深度)	A1 (含深度)	A2	A3 (不含深度)	A3 (含深度)	A4
翻译 (米)	239.19	142.12	290.43	45.42	32.18	0.28
旋转 (度)	60.72	39.52	116.50	6.41	7.66	5.77

这些数据集被称为“城市”、“豺狼”和“手持机”。这些数据集的细节将在以下章节中提供。我们收集数据的传感器套件包括Velodyne VLP-16激光雷达、FLIR BFS-U3-04S2M-CS相机、MicroStrain 3DM-GX5-25 IMU和Reach RS+ GPS（用于地面实况）。我们将提议的框架与开源的解决方案进行比较，其中包括VINS-Mono、LOAM、LIO-mapping、LINS和LIO-SAM。所有的方法都是用C++实现的，并在一台装有英特尔i7-10710U的Ubuntu Linux笔记本上执行。我们对LVI-SAM的实现和数据集可在下面的链接²。

A. 消融研究

我们使用城市数据集展示了系统中每个模块的设计是如何影响所提出的框架的性能的。这个数据集的特点是建筑物、停放和移动的汽车、行人、骑自行车的人和植被，是由操作员步行和携带传感器套件收集的。我们还特意将传感器套件放置在具有挑战性的位置（图6(a)），以验证系统在退化情况下的稳健性。由于高空植被茂密，该地区没有GPS。我们在同一位置开始和结束数据采集过程，以验证端到端的平移和旋转误差，见表一。

1) A1 - 包括来自激光雷达的特征深度信息对视觉-惯性测距的影响。我们禁用LIS中的扫描匹配，仅依靠VIS进行姿势估计。启用和不启用深度注册的结果轨迹被标记为A1。



(a) Jackal数据集环境

(b) 手持数据集环境

图8：收集Jackal和Handheld数据集的卫星图像。白点表示数据集中的GPS可用性。

图7.轨迹方向为时钟方向。表一中显示的端到端姿势误差，在将深度与视觉特征联系起来时，大大减少。

2) A2 - 包括视觉-惯性测距的效果。我们禁用VIS，只使用LIS进行姿势估计。在图7中标记为A2的轨迹，在遇到退化的情况时，会有几次发散。

3) A3 - 包括激光雷达的特征深度信息对激光雷达-视觉-惯性测距的影响。我们现在使用VIS和LIS一起，切换VIS中的深度注册模块，比较所得到的LVIO轨迹。在视觉特征的深度帮助下，翻译误差进一步减少了29%，从45.42米减少到32.18米。注意，在这个测试中禁用了环形闭合检测，以验证在纯测距模式下的系统。

4) A4 - 包括视觉闭环检测的效果。我们通过启用VIS中的闭环检测功能来消除系统的漂移。当框架中的每个模块都被启用时，最终的轨迹在图7中被标记为A4。

B. 豺狼数据集

Jackal数据集是通过将传感器套件安装在Clearpath Jackal无人驾驶地面车辆（UGV）上收集的。我们在一个特征丰富的环境中手动驾驶机器人，在同一位置开始和结束。该环境如图8（a）所示，具有结构、植被和各种路面。可用GPS接收的区域用白点标记。

我们比较了各种方法，并在图9(a)中显示了它们的运行轨迹。我们通过手动禁用和启用，进一步验证具有循环关闭功能的方法的准确性。基准测试结果显示在表II.LVI-SAM相对于GPS测量值实现了最低的平均均方根误差（RMSE），GPS测量值被视为地面真实。LINS实现了最低的端到端平移误差，它改编自LeGO-LOAM[10]，专门为UGV操作设计。最低的端到端旋转误差也是由LVI-SAM实现的。

C. 手持式数据集

手持式数据集是由携带传感器套件的操作人员在几个空地上行走收集的，这

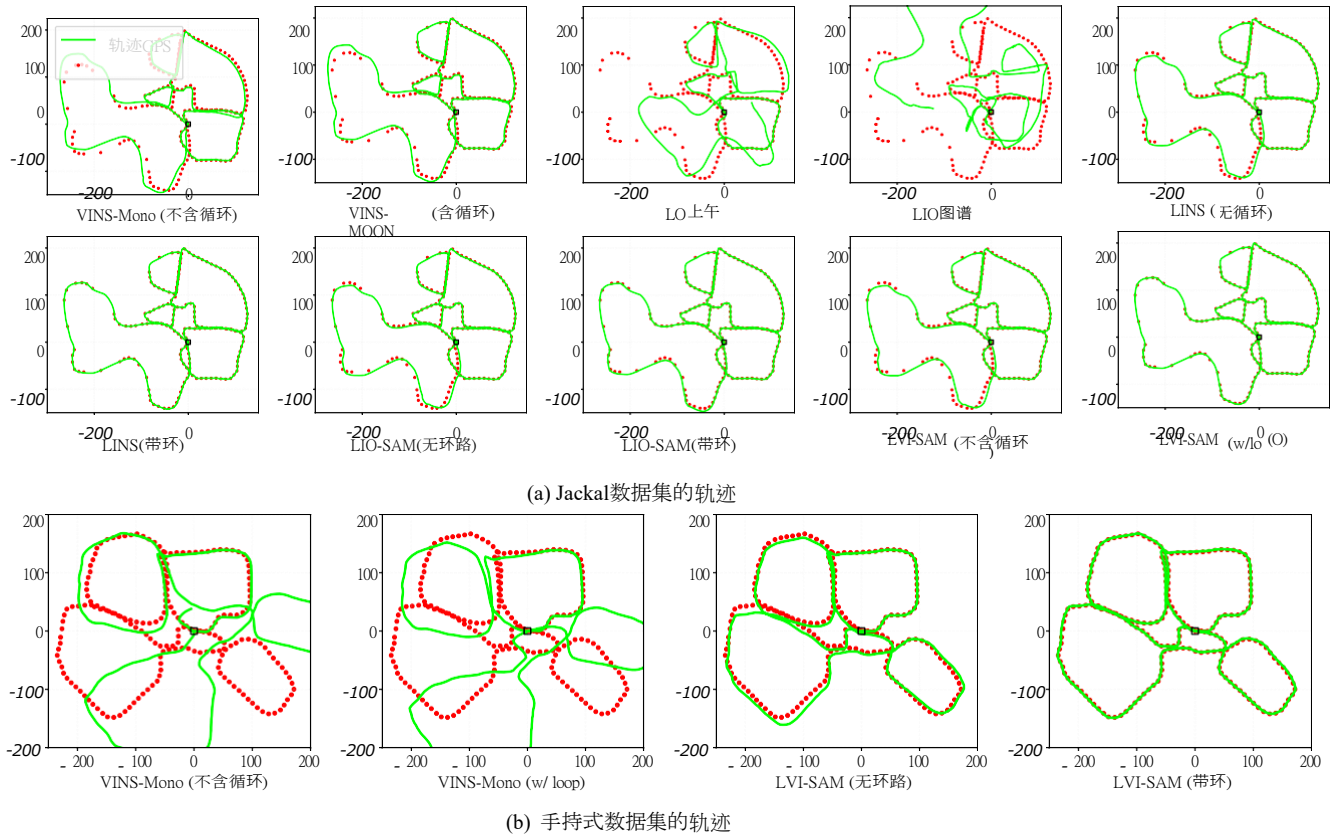


图9：使用Jackal和Handheld数据集的各种方法的轨迹。绿线描述了每种比较方法的结果轨迹。作为地面实况的GPS定位测量值用红点表示。LOAM、LIO-mapping、LINS和LIO-SAM的轨迹没有显示，因为它们未能产生有意义的结果。

表二：使用各种方法对Jackal和Handheld数据集进行定量比较。

数据集	误差类型	VINS (不含循环)	VINS (含循环)	LOAM	LIO-mapping (不含循环)	辽宁省 (含循环)	LIO-SAM (不含循环)	LIO-SAM (含循环)	LVI-SAM (不含循环)	LVI-SAM (含循环)
豺狼	RMSE与GPS的关系 (m)	8.58	4.49	44.92	127.05	3.95	0.77	3.54	1.52	0.67
	翻译 (米)	10.82	11.86	61.73	123.22	7.37	0.09	5.48	0.12	4.69
	旋转 (度)	9.98	12.79	61.41	139.23	4.80	2.07	2.18	2.64	1.52
掌上型	RMSE与GPS的关系 (m)	87.53	73.07	失败	失败	失败	失败	53.62	失败	7.87
	翻译 (米)	40.65	1.87	失败	失败	失败	失败	58.91	失败	7.57
	旋转 (度)	53.48	52.09	失败	失败	失败	失败	20.17	失败	24.82

视觉-惯性测距紧密耦合的框架。

在图8(b)中显示。该数据集也在同一位置开始和结束。我们通过一个开放的棒球场来增加这个数据集的挑战性，这个棒球场位于图像的顶部中心。当经过这个场地时，摄像机和激光雷达收集的主要观测数据分别是草地和地平面（图6 (b) 和 (d)）。由于上述的退化问题，所有基于激光雷达的方法都不能产生有意义的结果。所提出的框架LVI-SAM在启用或不启用环路闭合的情况下都成功地完成了测试，在表二中列出的所有三个基准标准中实现了最低误差。

IV. 结论

我们提出了LVI-SAM，这是一个通过平滑化和视觉化来实现激光雷达-

在复杂环境中进行实时状态估计和地图绘制。所提出的框架由两个子系统组成：一个视觉-惯性系统和一个激光雷达-惯性系统。这两个子系统被设计为以紧密耦合的方式互动，以提高系统的稳健性和准确性。通过对各种规模、平台和环境的数据集的评估，我们的系统显示出与现有的公开方法相当或更好的准确性。我们希望我们的系统将作为一个坚实的基线，其他人可以很容易地在此基础上推进激光雷达-视觉-惯性测距技术的发展。

鸣谢

这项工作得到了荷兰阿姆斯特丹先进城市解决方案研究所的支持。

参考文献

- [1] J.Graeter, A. Wilczynski, and M. Lauer, "LIMO:Lidar-Monocular Visual Odometry, " *IEEE/RSJ 智能机器人和系统国际会议 (IROS)*, 第7872-7879页, 2018年。
- [2] Y.-S. Shin, Y. S. Park, and A. Kim, "DVL-SLAM: Sparse Depth Enhanced Direct Visual-LiDAR SLAM," *Autonomous Robots*, vol. 44, no. 2, pp.115-130, 2020.
- [3] A.I. Mourikis和S. I. Roumeliotis, "A Multi-state Constraint Kalman Filter for Vision-aided Inertial Navigation," *IEEE International Conference on Robotics and Automation (ICRA)*, pp.3565-3572, 2007.
- [4] M.Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust Visual Inertial Odometry using A Direct EKF-based Approach," *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp.298-304, 2015.
- [5] P.Geneva, K. Eickenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation, " *IROS Workshop on Visual-Inertial Navigation: 挑战与应用*, 2019年。
- [6] S.Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based Visual-Inertial Odometry using Nonlinear Optimization, " *The International Journal of Robotics Research*, vol. 34, no.3, pp.314-334, 2015.
- [7] T.Qin, P. Li, and S. Shen, "VINS-Mono:稳健和通用的单眼视觉惯性状态估计器", *IEEE Transactions on Robotics*, vol. 34, no.4, pp. 1004-1020, 2018.
- [8] J.Delmerico和D. Scaramuzza, "飞行机器人的单眼视觉惯性测距算法的基准比较", *IEEE国际机器人和自动化会议 (ICRA)*, pp.2502-2509, 2018.
- [9] J.Zhang and S. Singh, "Low-drift and Real-time Lidar Odometry and Mapping," *Autonomous Robots*, vol.41, no. 2, pp.401-416, 2017.
- [10] T.Shan和B. Englot, "LeGO-LOAM, Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.4758-4765, 2018.
- [11] C.Chen, H. Zhu, M. Li, and S. You, "A Review of Visual-Inertial Simultaneous Localization and Mapping from Filtering-based and Optimization-based Perspectives," *Robotics*, Vol. 7, no.3, p. 45, 2018.
- [12] H.Ye, Y. Chen, and M. Liu, "Tightly Coupled 3D Lidar Inertial Odometry and Mapping," *IEEE International Conference on Robotics and Automation (ICRA)*, pp.3144-3150, 2019.
- [13] T.Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp.4758-4765, 2020.
- [14] C.Qin, H. Ye, C. E. Pranata, J. Han, S. Zhang, and M. Liu, "LINS: A Lidar-Inertial State Estimator for Robust and Efficient Navigation," *IEEE International Conference on Robotics and Automation (ICRA)*, pp.8899-8906, 2020.
- [15] C.Debeunne和D. Vivet, "A Review of Visual-LiDAR Fusion based Simultaneous Localization and Mapping," *Sensors*, vol. 20, no. 7, p. 2068, 2020.
- [16] J.Zhang and S. Singh, "Laser-Visual-Inertial Odometry and Mapping with High Robustness and Low Drift," *Journal of Field Robotics*, vol. 35, no. 8, pp.1242-1264, 2018.
- [17] A.Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision Meets Robotics:KITTI数据集", 《国际机器人研究杂志》, 第32卷, 第11期, 第1231-1237页, 2013年。
- [18] X.Zuo, P. Geneva, W. Lee, Y. Liu, and G. Huang, "LIC-Fusion:LiDAR-Inertial-Camera Odometry," *arXiv preprint arXiv:1909.04102*, 2019.
- [19] M.Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F.Dellaert, "iSAM2: Incremental Smoothing and Mapping using the Bayes Tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp.216-235, 2012.
- [20] J.Shi et al., "Good Features to Track," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.593-600, 1994.
- [21] B.D. Lucas, T. Kanade 等人, "一种应用于立体视觉的迭代式图像注册技术", 1981。
- [22] D.Ga'lvéz-Lo'pez和J. D. Tardos, "用于图像序列中快速位置识别的二进制词袋", *IEEE 机器人学反应*, 第28卷, 第5号。5, pp. 1188-1197, 2012.
- [23] M.Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief:Binary robust independent elementary features," *European conference on computer vision*, pp.778-792, 2010.

- [24] J.Zhang, M. Kaess, and S. Singh, "On Degeneracy of Optimization-based State Estimation Problems," *IEEE International Conference on Robotics and Automation (ICRA)*, pp.809-816, 2016.