面向高速 GPS拒止飞行的嵌入式硬件实时热惯性里程计

Austin Stone、Mark Petersen 和 Cammy Peterson

arXiv:2603.02114v1 [cs.RO] 2 Mar 2026

摘要—**我们提出了一种实时单目热惯性里程计系统，专为嵌入式硬件上的高速、GPS拒止飞行而设计。该系统在固定滞后因子图中融合来自 FLIR Boson+ 640 长波红外相机、高速率惯性测量单元、激光测距仪、气压计和磁力计的测量值。为在运动模糊、低对比度和视角快速变化的情况下保持可靠的特征轨迹，我们采用了轻量级热成像优化前端，并进行多阶段特征过滤。激光测距仪的测量值提供逐特征深度先验，在运动可观测性较弱时稳定尺度。高速率惯性数据首先经过切比雪夫 II 型无限脉冲响应（IIR）滤波器预滤波，然后进行预积分，从而提升系统在激烈机动时抵抗机体振动的鲁棒性。为解决高速空速引起的气压高度误差，我们训练了一个能够感知不确定性的门控循环单元（GRU）网络，用于建模静压畸变的时序动态，其性能优于多项式和多层感知机（MLP）基线。该完整系统部署于 NVIDIA Jetson Xavier NX，支持四旋翼以 30米/秒进行闭环飞行，在公里级轨迹上的漂移低于 2%。这些贡献拓展了热惯性导航的运行范围，使飞行器能够在视觉条件恶化且无 GPS 环境中可靠地高速飞行。**

I. 引言

在 GPS 拒止和视觉退化环境中进行准确的状态估计，仍然是空中机器人领域的核心挑战。视觉退化环境通常具有照明不足或烟雾、雾等遮蔽物，广泛存在于各种现实应用中，包括军事行动、夜间城市导航、基础设施检查、地下探索以及搜救任务。在许多此类场景中，快速机动能力同样是关键要求，这对估计系统的鲁棒性和响应能力提出了更高要求。在本研究中，我们针对这些挑战的交汇点，提出了一种单目热惯性里程计（TIO）系统，专为嵌入式平台上的高速 GPS拒止飞行而设计。

为应对 GPS 拒止环境带来的挑战，视觉惯性里程计（VIO）已成为一种主流解决方案。视觉惯性里程计系统利用相机与惯性测量单元（IMU）的互补特性 [1]。通常，相机能够提供漂移较低的结构观测，但缺乏度量尺度；惯性测量单元则能提供高频的度量运动估计，尽管估计结果会受到积分误差和偏置漂移的影响。我们在这一核心视觉惯性里程计框架中增添了三种额外的低成本传感器：激光测距仪（LRF）、气压计和磁力计。我们采用 VINS-Fusion 式的方案，通过向因子图中添加相应因子来融入这些传感器的测量值 [2]。激光测距仪可直接提供度量深度信息，有助于在单目流程中初始化特征深度，并在恒定加速度期间维持尺度一致性 [3]。气压计是大多数四旋翼无人机的标准配置，可提供高频伪高度测量，但在高速机动时容易受到压力扰动的影响而产生误差 [4]。最后，磁力计可约束偏航漂移。惯性测量单元能够观测滚转角和俯仰角，但无法观测偏航角 [5]。偏航漂移是视觉惯性里程计误差的重要来源，即使使用质量相对较低的磁力计，也能减轻这一问题。

尽管传统视觉惯性里程计（VIO）算法在中等速度和良好光照条件下已取得极大成功，但有几个因素限制了其在高速、低照度或遮蔽物密集环境中的适用性。首先，标准 RGB相机在低光照条件下信噪比较低，而缩短曝光时间以减轻运动模糊，往往会导致图像进一步曝光不足。其次，RGB相机无法穿透烟雾或雾等环境遮蔽物。然而，长波红外（LWIR）热像仪为视觉退化条件提供了一种颇具吸引力的替代方案，因为它们能够捕捉不受环境照明影响而持续存在的辐射对比度。它们还能够穿透遮蔽物。将其集成到视觉惯性框架中，便形成了热惯性里程计，它结合了惯性传感的鲁棒性与长波红外图像的可见性优势 [6]。

尽管长波红外相机具有显著优势，但并非万能。在以温度均一为主的场景中——例如雪、雨或刚被淋湿的表面——热对比度会降低，反射或冷却效应也会扭曲辐射线索，因此相机可能难以发挥作用。在这些情况下，电光（EO）相机的表现可能优于热视觉。热图像的对比度通常也低于可见光谱图像，因为大多数自然场景呈现的是平滑的温度梯度，而非纹理鲜明的特征。与电光相机的光电二极管电荷能在微秒内完全释放不同，长波红外相机热敏电阻的时间常数约为数毫秒 [7]。热敏电阻中的残余热量意味着先前场景的部分信息仍会影响当前场景，从而造成额外模糊。此外，长波红外相机还会受到显著的固定模式噪声（FPN）影响，这种噪声源于探测器阵列中像素间随温度变化的非均匀性。尽管现场非均匀性校正（NUC）程序能够减弱固定模式噪声，

A. Stone 和 C. Peterson 就职于杨百翰大学，美国犹他州普罗沃，邮编 84602。

但这些程序会造成间歇性的“数据丢失”时段，通常持续 250 毫秒至 1 秒；在此期间，传感器会停止输出有效帧 [7]。在高速飞行中，这类数据丢失可能导致状态估计明显漂移，因为特征会消失，且 IMU偏置无法得到适当校正。尽管非均匀性校正通常能有效去除大部分固定模式噪声，但伪影和低频噪声仍可能残留。在寒冷环境中运行时，这种情况尤为明显。

我们的系统所针对的高速低空飞行环境带来了额外挑战。与高空飞行器不同，后者可以采用地形相对导航（TRN）技术，将雷达或光学剖面与存储的数字地形图进行匹配，以实现漂移校正；低空飞行则无法采用这类方法——视场（FOV）过窄，而且相对于机载地图的分辨率，地形变化过快 [8]。因此，导航必须依靠基于局部观测结构的实时视觉惯性估计，并且即使在贴近地面高速飞行时也必须保持稳定。虽然 S-MSCKF [9] 等先前工作已展示了速度最高达 17.5 米/秒的双目 RGB 视觉惯性里程计，但我们的系统使用单目热成像，以 30米/秒的速度运行。在如此高的速度下，新问题随之出现，已有问题也会加剧。运动模糊更加明显，影响视觉特征跟踪。由于特征留在视场中的时间缩短，特征轨迹也随之变短，从而加大数据关联的难度，并降低单个特征对状态估计的影响。激烈机动要求以低延迟的状态估计来维持控制稳定性，因此处理延迟会造成更大影响。因此，机载嵌入式处理器必须在严格的计算资源限制下实时执行稳健估计。此外，高速飞行会加剧机械振动，降低惯性测量单元（IMU）的测量质量。

高速飞行的另一项挑战是保持准确的高度估计。尽管视觉惯性里程计（VIO）能够可靠地估计短期垂直运动，但在较长的轨迹上，高度不可避免地会发生漂移，因此，全局测量对于持续飞行至关重要。气压测量通常承担这一作用，用于稳定视觉惯性里程计系统中的垂直状态。然而，气压计极易受到气动扰动的影响：随着车辆空速增加，气压计端口处的静压偏低会导致报告的高度虚高 [4]。在我们的实验中，这些高速飞行扰动造成的高度误差超过 15 m，足以降低估计器性能，并在低空飞行期间带来严重风险。据我们所知，此前尚无研究探讨静压位置误差（SPPE）对基于视觉惯性里程计的状态估计的影响。

本文探讨了使用热像仪进行高速飞行时，视觉惯性里程计（VIO）面临的关键挑战。我们提出了一种统一的热惯性里程计估计器（TIO），可应对（1）高速飞行、（2）单目热成像以及（3）动态气压计补偿，同时满足嵌入式硬件严格的计算与功耗限制。通过解决高速飞行、热成像和动态气压扰动交织带来的挑战，我们拓展了自主四旋翼的运行范围，使其能够在以往无法到达的环境中可靠导航。

我们的主要贡献如下：

- 我们提出了一种高效、鲁棒的特征跟踪与滤波流水线，专为应对高速热成像的独特挑战而设计，包括运动模糊、低对比度和固定模式噪声（FPN）。
- 我们提出了一种新颖的气压补偿算法，该算法对气压计误差的时变动态进行建模，相较于现有技术实现了更优的性能。
- 我们提出了一种 IMU 预滤波步骤，可在激烈机动期间提高估计器的稳定性。
- 我们展示了首个经验证可支持以最高 30米/秒速度进行闭环飞行的单目热惯性里程计系统。我们通过严格的硬件实验验证了该系统。

本文其余部分安排如下。第二节回顾高速视觉惯性里程计、热惯性里程计和气压补偿方面的相关工作。第三节介绍系统设置，包括传感器同步、校准流程和非均匀性校正调度。第四节介绍算法设计，从前端图像增强和特征跟踪到后端平滑。第四节还详细介绍气压计补偿模型及其在平滑框架中的集成。第五节报告了大量实验结果，比较了不同飞行速度和环境条件下的特征检测精度、IMU 预处理效果、气压计误差降低情况、轨迹精度和计算性能。最后，第六节总结全文并展望未来研究方向。

II. 背景

本研究处于两个主要领域的交汇点——热惯性里程计（TIO）和气压补偿，同时借鉴了视觉惯性里程计（VIO）的基础性见解。尽管 TIO 沿用了 VIO 框架的算法结构，但传感模态从根本上改变了问题：热成像与可见光图像的表现不同，因此在低纹理或噪声场景中，需要专门的前端处理、跟踪和鲁棒性策略。虽然这些领域都已分别得到研究，但在高速、低能见度航空导航背景下对它们进行整合仍未得到充分研究。本节回顾各领域的最新进展，并着重介绍其发展历程与局限性。

A. 视觉惯性里程计

视觉惯性里程计已发展成为无 GPS 环境下进行状态估计的主要解决方案，依赖相机和惯性测量单元的互补优势。紧耦合方法在该领域占主导地位，可分为两大类：基于滤波的方法和基于优化的方法。

基于滤波器的方法是最早投入实际应用的视觉惯性里程计（VIO）实现，它们利用卡尔曼滤波器的各种变体，将高频惯性数据与低频相机观测进行融合。这些方法通常采用扩展卡尔曼滤波器（EKF）或无迹卡尔曼滤波器。在这两种情况下，惯性测量单元（IMU）都用于随时间向前传播状态，而

相机导出的特征轨迹则为约束IMU漂移提供了依据。

这一类别中的开创性系统是多状态约束卡尔曼滤波器（MSCKF） [1]。MSCKF 引入了将一段较短的相机位姿历史纳入滤波器状态的概念。当某个特征在多个图像帧中被观测到时，其测量值会被投影到这些位姿上，以生成多视图约束。这些约束会在特征轨迹终止后形成，并随后作为延迟测量更新用于滤波器。因此，视觉信息无法立即使用。通过在纳入约束后边缘化特征状态，MSCKF 在利用多帧几何的同时避免了状态无限增长，从而能够在嵌入式处理器上实时运行。此后，许多衍生方法对这一架构进行了改进，例如用于双目相机的 S-MSCKF [9]，以及可降低对线性化漂移敏感性的机器人中心表述（R-VIO） [10] 。

另一项值得关注的贡献是 ROVIO [11]，，它将直接光度误差纳入扩展卡尔曼滤波器的更新步骤，有效地将像素强度对齐与惯性传播结合起来。这项创新表明，基于滤波器的框架不仅能够处理几何残差，也能进行直接图像对齐，从而提高低纹理场景中的精度。类似地，也有人提出松耦合表述，通过扩展卡尔曼滤波器将现有视觉里程计流程（例如单目SfM）的输出与惯性估计融合，具有模块化和易于实现的优点 [12],[13]。

基于滤波的视觉惯性里程计方法在嵌入式环境中具有多项实用优势。其更新结构能够以极低的延迟生成状态估计，这对于敏捷空中机器人的快速控制回路至关重要。由于状态维度较小且更新是局部的，这类方法计算开销低，非常适合资源有限的处理器。此外，其模块化设计使得只需对底层滤波器方程做最小改动，便可集成磁力计、气压计、GPS伪距或测距仪等附加传感器。

尽管具有这些优势，但依赖围绕单一工作点进行线性化也会带来局限 [2]。在较长的轨迹上，线性化中的微小误差会不断累积，导致不一致和漂移。滤波器也难以妥善处理回环或延迟更新，因为过去的状态一经边缘化便无法再访问。因此，尽管基于滤波的视觉惯性里程计在嵌入式和时间关键型应用中仍广受欢迎，但其精度通常不及现代基于优化的方法；后者会重新访问并联合优化过去的状态。尽管如此，MSCKF [1]及其后续方法开创的理念仍然影响深远：滑动窗口边缘化、多帧特征约束以及轻量级传感器融合策略，至今仍为新一代系统提供借鉴。

随着机载计算能力不断提升，基于优化的方法逐渐成为视觉惯性里程计的主流范式。这些系统将估计建模为针对近期状态滑动窗口的非线性最小二乘问题，并在活动集合中保留多个位姿、速度和偏置 [5]。测量值以残差因子的形式纳入，联合优化会同时细化所有变量。通过重新审视过去的状态，优化能够减少线性化误差，增强对离群值的鲁棒性，并自然地适应延迟更新。

OKVIS [14] 证明了实时紧耦合优化视觉重投影和惯性约束的可行性，其漂移显著低于基于扩展卡尔曼滤波的方法。VINS-Mono [5] 将这一方法扩展到单目相机，并通过回环和位姿图优化实现全局一致性；VINS-Fusion [2] 则将因子图推广到支持双目视觉和附加传感器，展示了该框架的模块化特性。

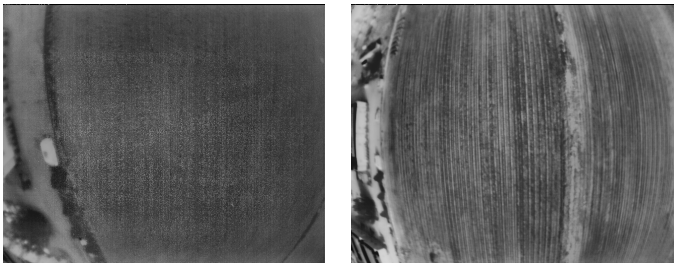
这些系统的一项关键使能技术是IMU预积分。将相机帧之间高频率的惯性数据汇总为单个复合测量值的理念，至少可以追溯到Lupton和Sukkarieh [15],，他们提出了用于基于图的平滑的预积分惯性测量模型。Forster等人 [16]通过在SO(3)流形上推导闭式表达式，进一步完善了这一概念，在避免反复对原始IMU数据进行重新积分的同时，兼顾了精度与效率。

近期研究着重于可扩展性和效率。ICE-BA [17] 提出了增量式重新线性化，仅更新受较大状态变化影响的雅可比矩阵，从而即使在扩展窗口下也能实现低延迟运行。XR-VIO [18] 强调了鲁棒初始化和混合特征匹配的重要性，表明前端可靠性与后端优化精度密不可分。

基于优化的方法具有多项优势，包括通过联合优化降低漂移、灵活融合异构传感器模态，以及在视差足够时再纳入约束，从而延迟一致性处理。然而，这些优势需要付出计算成本：每一步都要求解一个非线性最小二乘问题，计算负担很重，尤其是在处理预算有限的嵌入式平台上。

为保持实时性能，实际系统必须有所取舍。许多基于优化的视觉惯性里程计框架会限制每个更新周期的迭代次数，这意味着求解器在每个时间步都可能无法完全收敛。此外，还需要采取激进的边缘化策略来限制计算成本，但这会重新引入优化框架试图避免的部分线性化误差。这些问题会导致延迟增加，并在长距离轨迹中造成漂移。

我们的系统直接建立在以 VINS-Fusion [2]和 ICE-BA [17]为代表的基于优化的方法之上。我们采用固定滞后平滑方案，联合优化位姿、速度和偏置；使用 Forster等人 [16], 提出的基于流形的预积分模型，并采用 ICE-BA 风格的增量式重新线性化，以维持实时性能。此外，我们还在因子图中加入额外的传感器模态——由惯性测量单元和磁力计推导的姿态、激光测距仪深度先验以及气压高度——从而针对高速热惯性导航面临的独特挑战，定制优化框架。



(a) 劣化的长波红外图像 (b) 良好的长波红外图像

图1：劣化红外图像与良好红外图像的对比。劣化图像呈现出典型的交叉网格图案。

B. 热惯性里程计

热惯性里程计能够在退化环境（例如，低照度、烟雾）中实现导航，而 RGB相机在这类环境中会失效。长波红外相机不受光照影响，且能够穿透许多遮蔽物，但热惯性里程计面临以下挑战：

1) 空间分辨率和对比度较低：与可见光相机相比，长波红外传感器通常像素分辨率较低，视场也较窄（例如，FLIR Boson+提供的最大水平视场约为 95°）。此外，自然热场景的辐射梯度平滑、纹理有限，因此可跟踪特征的数量和可靠性都会降低。**2) 固定模式噪声：**固定模式噪声由像素响应非均匀性引起，传感器会定期通过非均匀性校正流程进行补偿，该流程将传感器输出与均匀温度源（通常为快门）进行比较。固定模式噪声示例见图1。**3) 图像丢失：**进行非均匀性校正时（通常持续250–1000毫秒），相机可能会冻结输出或生成无效帧，导致估计器暂时失去视觉约束。在此期间，无法校正不断累积的惯性测量单元偏置漂移或加速度计噪声。

4)热滞后：长波红外相机使用微测辐射热计（将红外辐射作为热量进行测量的热探测器），其热时间常数有限，约为5–15毫秒。因此，时刻 t的输出包含来自先前图像帧的残差信息。在激烈机动或高振动期间，热滞后表现为图像中模糊或拖影加剧，从而进一步降低特征稳定性 [19]。

近期的热惯性里程计（TIO）研究涵盖了学习式与手工设计方法。TP-TIO [20],等学习式描述子流程通过合成热成像增强来改进 SuperPoint [21]，提高了特征对长波红外（LWIR）伪影的鲁棒性。其他研究则侧重于选择信息量更丰富的区域：Khattak等人 [6] 利用空间熵指导低对比度场景中的直接光度跟踪。Self-TIO [22] 等基于光流的方法将 KLT初始化器与RAFT [23] 结合，以实现亚像素运动估计，并通过基于SVD的去噪来减轻固定模式噪声（FPN）。DeepTIO [24]等端到端回归器则完全绕过特征提取，直接预测位姿根据热成像和IMU数据进行预测，不过这类方法仍然计算开销较大，泛化能力也较弱。

手工设计的流水线提供了一种强调效率与可解释性的替代方案。诸如ETIO [25] 等系统采用边缘检测、二值化和KLT跟踪，而相关方法则采用受同类视觉惯性里程计文献启发的直接对齐或距离变换目标 [26], [27]。由于延迟低且计算需求适中，这些方法仍然是最适合嵌入式部署的方案。

尽管基于深度学习的方法在热成像上已展现出有前景的性能提升，但其推理时间仍无法满足机载无人机硬件严格的延迟与功耗约束。例如，Self-TIO在笔记本级图形处理器上每图像帧需要超过30毫秒，而TP-TIO仅跟踪特征就需要30毫秒。因此，手工设计的流水线仍然是快速闭环飞行中嵌入式热惯性里程计最实用的解决方案。我们的系统遵循这一理念：受XR-VIO [18],启发，我们将对比度受限自适应直方图均衡（限制对比度自适应直方图均衡化）、基于梯度的滤波和描述子验证相结合，以便在运动模糊和固定模式噪声下保持稳定轨迹。这种混合设计在保留实时高速导航所需计算效率的同时，直接应对长波红外图像的视觉挑战。

C. 高速导航

高速导航已在采用地形相对导航（TRN）的高空平台上得到广泛研究。TERCOM[28], [29]和 SITAN [30] 等系统将雷达高度测量与数字地形图相结合，为以高速度飞行的固定翼飞机和导弹提供无漂移的状态估计。近年来，基于视觉的地形相对导航方法也取得了令人瞩目的成果。Maggio 等人 [31] 报告称，在一次蓝色起源的亚轨道飞行中，利用朝下的相机图像和惯性传感，将观测到的地形与卫星地图进行匹配，在速度超过 240 m/s 时，平均误差低于 55 m。这些方法在高空表现最佳，因为此时可以看到大片地形特征，并将其与已知地图进行匹配。

相比之下，低空导航面临着不同的挑战——视差更强，但可见区域有限。在这一环境下，视觉惯性里程计（VIO）更为常用 [8]。然而，现有视觉惯性里程计系统在高速下会因运动模糊、特征轨迹缩短以及控制回路延迟增加而性能下降。Sun 等人 [9] 证明，S-MSCKF 可支持最高 17.5 m/s 的基于双目 RGB 的视觉惯性里程计，但该方法依赖双目视觉，并假设光照条件良好。大多数公开数据集不包含热成像，且运行速度较低，例如，EuRoC 微型飞行器数据集的速度很少超过 5 m/s [32], , SwiftBase [33] 则提供以 200 Hz 采集的中速双目单色成像数据。竞速无人机数据集的速度最高可达 28 m/s [34], , 但通常使用RGB相机，且飞行轨迹位于结构化的室内环境或较短的室外场地。

据我们所知，我们的工作首次展示了在嵌入式硬件上以30米/秒的速度进行低空飞行时，实现实时单目热惯性里程计并进行闭环控制。

D. 气压计高度补偿

气压传感器通过测量大气压力来提供高度估计。虽然成本低、结构简单颇具吸引力，但它们容易受到高速运动引发的误差影响。当气流经过气压计的端口时，会产生局部负压，导致测得的高度出现虚假波动，这种现象即为SPPE。该效应会随飞行器速度增加，并可能在激烈机动过程中严重降低估计器性能。

现有的补偿策略采用启发式方法。例如，PX4 [35]和 ArduPilot [36]建议根据人工采集的校准数据进行多项式拟合。较新的方法使用多层感知机（MLP）神经网络学习从速度到气压计校正的映射[4]。然而，这类前馈模型忽略了 SPPE 的时间结构。

相比之下，我们的系统采用门控循环单元（GRU）网络 [37] 对气压误差的时间依赖关系进行建模。随后，将学习到的校正量作为高度约束融入因子图，从而改善状态估计并减少长期漂移。据我们所知，这是首次展示在视觉惯性框架中应用循环式气压计校正。

III. 设置

要确保空中机器人实时状态估计的可复现性，必须仔细处理硬件同步和系统配置。本节详细介绍实验平台所依赖的设置流程和校准规程。

A. 硬件配置

所有估计过程均在 NVIDIA Jetson Xavier NX 上执行。由于使用的是非实时操作系统，精确标记传感器数据的时间戳并非易事。为此，我们采用专用微控制器，并将其称为传感器管理器（SM），用于处理时间要求严格的传感器数据并进行同步。

惯性测量单元由硬件以 1200 Hz 的频率触发。原始数据在设备端使用 IV-C 节所述的无限脉冲响应（IIR）滤波器流水线进行滤波，随后降采样至 120赫兹，再传送至 Xavier NX。热像仪（ FLIRBoson+ 640 ）以 30 Hz 运行，并由 SM 通过硬件触发。每次相机触发均与相应的惯性测量单元触发在时间上对齐，并将关联的惯性测量单元采样标记为图像同步。Xavier NX 读取图像，并使用与图像对齐的惯性测量单元测量时间戳为其标记时间戳。

为了在 SM 与 Xavier NX 之间保持一致的时间基准，我们采用了一种硬件辅助同步方案。SM 每秒输出一个同步脉冲，并发送相应的带时间戳消息，该消息由 NX 捕获，用于估计时钟偏移和漂移。该偏移被用于将所有图像和惯性测量单元时间戳转换到统一的 SM 时间框架中。这确保了各传感器之间亚毫秒级的时间一致性，同时避免了仅靠软件同步通常带来的延迟和抖动。

附加传感器由 SM 加盖时间戳，并异步转发至 Xavier NX。为提供度量深度信息，激光测距仪被刚性安装并与相机光轴对准，从而使其测量值表示图像中心处的深度。

PX4 实例运行在独立的微控制器上，并作为我们车辆的飞行控制器。进行闭环飞行时，它可使用专用的惯性测量单元，并在 TIO 流水线于 Xavier NX 上生成位置信息后（30 Hz）接收该信息。

B. 相机标定

我们使用 Kalibr [38]和专为长波红外（LWIR）成像设计的定制标定靶标进行热红外-惯性标定。该靶标由一块反射板构成，板上用吸收材料勾勒出一组 ArUco标记 [39]。靶标平放在地面上，以反射寒冷的天空。我们执行标准的标定运动序列，以激发实现可观测性所需的模式。然而，在标定运动序列期间，天空受到遮挡，在一定程度上限制了标定质量。我们遵循 Kalibr 建议的标定传感器频率：相机为 20 Hz，惯性测量单元为 200 Hz。

C. 平场校正调度

Boson+ 热像仪需要定期进行平场校正（FFCs），以补偿内部温度漂移和非均匀性伪影。这会造成一个无法获取可用于跟踪的特征的数据中断窗口。因此，时机不当的平场校正可能会中断特征连续性，加剧估计器漂移。为缓解这一问题，我们实施了动态平场校正调度器，并采用以下策略:

- Boson+根据内部指标请求执行 FFC。
- 如果当前线速度低于 5 m/s 且角速度低于 0.5 rad/s，Boson+一旦提出请求，便会立即触发 FFC。
- 否则，FFC 最多延迟 2 分钟，或等到运动降至阈值以下时触发——以先发生者为准。
- 起飞前始终会触发一次初始 FFC，以确保图像质量一致。

该策略兼顾了热校正的需求与不间断视觉惯性跟踪的要求，尤其是在高速飞行阶段。

IV. 方法

本节将概述我们的整体算法架构，随后详细介绍各个组成部分，这些部分共同确保在前文所述的严苛条件下实现稳健的状态估计。

A. 算法概述

图2展示了我们完整的估计流程。来自 FLIR Boson+ 相机的原始热成像首先经过高斯滤波和限制对比度自适应直方图均衡化（ CLAHE） [40]预处理，以增强图像对比度并减少固定模式噪声。特征

随后使用 FAST 角点检测器 [41]检测特征。通过光流跟踪在连续帧之间维持特征轨迹，并辅以本质矩阵异常值剔除 [42]。通过基于梯度的滤波和 ORB描述子匹配 [43]增强特征跟踪的鲁棒性。

同时，高频IMU测量值（1200Hz）经过切比雪夫II型IIR滤波器[44] 预处理，以减弱机械振动，随后降采样至120Hz。滤波和降采样后的IMU数据经过预积分 [16], ，为连续的优化窗口提供惯性约束。

随后通过三角测量验证特征轨迹。经过验证的特征和IMU预积分结果被输入我们的Levenberg–Marquardt预条件共轭梯度（LMPCG）后端优化 [17], [45]–[48], 。该优化被表述为滑动窗口非线性最小二乘优化，联合估计导航状态（位姿、速度、偏置）和特征逆深度。

为解决高速机动期间SPPE在气压测量值中引入的高度误差，我们采用不确定性感知GRU神经网络。GRU通过捕捉时间依赖关系，动态校正气压高度测量值；校正后的测量值则作为额外的高度约束纳入优化框架。

最后，将优化后的状态估计融合到轻量级高频率EKF中，以便为闭环控制提供低延迟位姿更新。EKF由惯性测量单元进行传播，并由后端优化器更新，从而实现快速且经过漂移校正的估计，适用于激进飞行。这种结构化方法结合了精心设计的预处理、稳健的特征过滤和新颖的气压补偿，显著提高了估计器在严苛飞行条件下的精度和可靠性。

B. 初始化

为确保安全部署和状态估计的一致性，我们假设在估计器初始化阶段车辆保持静止——鉴于许多空中机器人应用中使用发射平台，这是一个合理的假设。在此假设下，我们将初始位置设为原点，并使姿态与重力方向和磁北对齐，以此初始化局部参考坐标系。

我们需要估计五个量：陀螺仪偏置、加速度计偏置、初始横滚角、初始俯仰角和初始偏航角。鉴于车辆处于静止条件，我们利用高频率传感器数据的统计特性来估计这些量。

陀螺仪偏置：陀螺仪偏置$\mathbf{b}_g$ 估计为 500 个陀螺仪测量值 $\{\boldsymbol{\omega}_i\}$的平均值，这是利用了真实旋转速度应为零这一事实。

滚转角、俯仰角与加速度计偏置：我们构建一个非线性优化，以联合估计加速度计偏置$\mathbf{b}_a$、滚转角 ϕ和俯仰角 θ。令$\mathbf{a}$表示500次加速度计测量值的平均值，$\mathbf{g} = [0, 0, g]^\top$ 表示惯性坐标系中的重力向量。则我们模型中的误差可表示为：

$$\mathbf{e} = \mathbf{a} - \mathbf{b}_a + \mathbf{R}_{\text{pitch}}(\theta)\mathbf{R}_{\text{roll}}(\phi)\mathbf{g}$$

我们可以进一步将加权目标函数定义为：

$$J(\mathbf{x}) = \|\mathbf{e}(\mathbf{x})\|^2 + w\|\mathbf{b}_a\|^2,$$

其中 $\mathbf{x} = [\mathbf{b}_a^\top,\ \phi,\ \theta]^\top$ 且 w为可配置权重。第一项根据估计的姿态和偏置，约束测得的加速度与重力向量对齐。然而，由于仅凭加速度计数据无法在静止场景中充分观测到滚转角、俯仰角和加速度计偏置，因此该问题约束不足。为此，我们加入第二项作为正则化惩罚项，以最小化加速度计偏置$\mathbf{b}_a$的幅值。我们计算该目标函数的雅可比矩阵$\mathbf{J}$和海森矩阵$\mathbf{H}$ ，并使用截断牛顿信赖域求解器求解由此得到的非线性系统。

误差函数关于状态$\mathbf{x}$的雅可比矩阵$\mathbf{J} \in \mathbb{R}^{3\times5}$ 为：

$$\mathbf{J} = \begin{bmatrix} -1 & 0 & 0 & 0 & -c_\theta g \\ 0 & -1 & 0 & c_\phi c_\theta g & -s_\phi s_\theta g \\ 0 & 0 & -1 & -s_\phi c_\theta g & -c_\phi s_\theta g \end{bmatrix},$$

其中 s_ϕ =sin(ϕ)， c_ϕ =cos(ϕ)。

海森矩阵$\mathbf{H} \in \mathbb{R}^{5\times5}$ 定义为：

$$\mathbf{H} = 2w\mathbf{J}^\top\mathbf{J} + 2\begin{bmatrix} \mathbf{I}_{3\times3} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + 2w\left(\frac{\partial \mathbf{J}}{\partial \phi}^\top \mathbf{e} + \frac{\partial \mathbf{J}}{\partial \theta}^\top \mathbf{e}\right)$$

偏航角初始化：仅凭加速度计和陀螺仪无法观测到偏航角 ψ 。为确定全局姿态，我们计算 PX4 提供的 14 个姿态测量值的平均值（时间间隔与 500 个惯性测量单元测量值相同），并将偏航角分量与磁北对齐。

上述过程为状态估计提供了可靠的初始化，使惯性坐标系与地球重力和磁北对齐，同时补偿传感器偏置。

当车辆静止在地面上时，可能没有可靠的视觉特征可用于约束位置。为防止估计器在这种约束不足的情况下发生漂移，我们会在添加状态时对其施加软先验，将其锚定在原点。这可以防止估计器在起飞前缓慢漂移。

*C. IMU*预处理

我们的系统以 1200 赫兹对机载惯性测量单元进行采样，但受通信管线带宽限制，只有降采样至 120 赫兹的数据会被转发到前端进行估计。尽管惯性测量单元安装在隔振平台上，残余机械噪声仍可能污染加速度计和陀螺仪读数。为减轻机体振动引入的高频噪声，以及降采样时由此产生的混叠，我们首先对原始IMU数据应用 IIR 滤波器。

我们采用截止频率为50 Hz的一阶切比雪夫II型IIR滤波器。滤波器系数为： A =[1, −0.97104422] 和 B = [0.01447789, 0.01447789]。滤波器系数通过scipy signal软件包[44]获得。该滤波器分别独立应用于加速度计和陀螺仪的每个轴。如第V-C节所示，

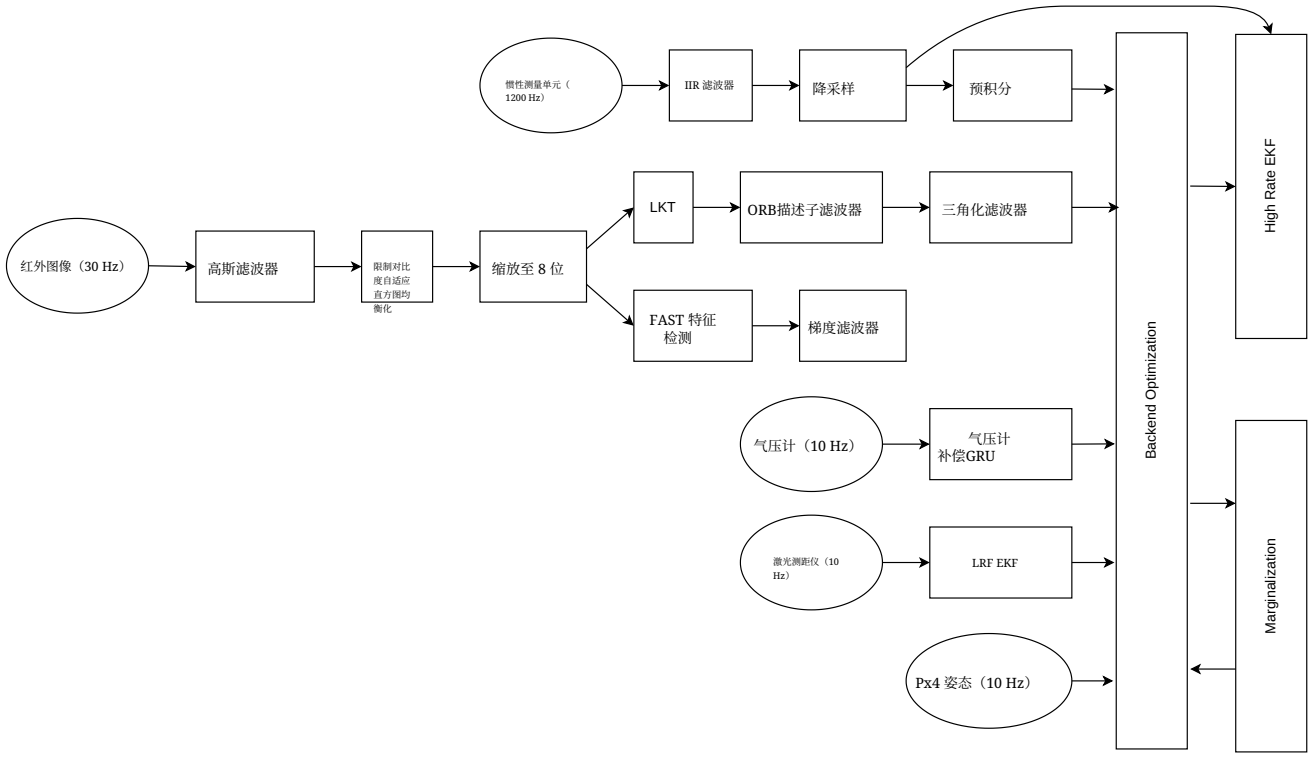


图2：我们的单目热惯性里程计流程概览。

这一滤波步骤显著提升了高速情况下的预积分性能。

经过滤波和降采样后，我们使用120赫兹数据流，按照 Forster等人 [16] 所述的方法进行惯性测量单元预积分。

D. 零速度检测与特征一致性

为进一步提高速度和IMU偏置的可观测性，我们采用了一个零速度检测模块，持续监测惯性测量单元和光流数据，其方式类似于 OpenVINS [49]。该检测器计算加速度计和陀螺仪测量值的滑动窗口统计量，以及重力补偿加速度的滑动窗口统计量。当这些量的均值和标准差均处于预设阈值范围内时，将平台判定为处于恒速状态。如果检测到恒速状态，则按照下文所述的相同方式添加姿态因子，但使用重力作为参考向量，并将观测加速度作为测量值。当恒速状态与接近于零的光流幅值和协方差同时出现时，即可推断为零速度。检测到零速度后，将零速度因子添加到因子图中，以减少静止区间内的漂移。

E. 图像预处理

我们从 Flir boson+热像仪的预自动增益控制（pre-AGC）16 位数字视频输出获取图像。该配置可最大限度降低处理延迟，并在机载重缩放之前提供对原始热信号完整动态范围的访问。

尽管我们的特征跟踪流程依赖于 OpenCV 的 8 位图像处理例程，我们仍首先在原生 16 位域中进行预处理，以尽可能保留更多信息。预处理流程包含三个主要阶段：

1) **固定图样噪声抑制：** 对原始 16 位图像应用具有 3×3卷积核且 $\sigma = 1$ 的高斯滤波器。这会削弱高频固定模式噪声，同时保留粗略的热结构。2) **对比度增强：** 我们直接在平滑后的 16 位图像上应用 CLAHE [40] 。CLAHE 可在局部增强辐射对比度，从而提高热梯度较弱区域的特征可检测性。然而，在热强度快速变化的场景中，CLAHE 会加剧这种差异，进而影响特征稳定性。3)**量化：** 随后对增强对比度后的图像进行缩放，并将其转换为 8 位灰度图像，以供 OpenCV 特征检测和跟踪算法使用。

这种分阶段的方法在增强图像的同时保留场景细节，并确保与标准计算机视觉工具兼容。尤其是，使用 AGC 前的图像可避免截断伪影，并确保在整个处理流程中保留高动态范围的热变化。

F. 特征跟踪

当已跟踪特征数量低于阈值时，使用 FAST 角点检测器 [41] 对预处理后的 8 位图像进行热特征检测。这

将被称为检测帧。为确保空间分布均匀并防止聚集，检测过程受分箱掩码约束，该掩码将图像划分为固定网格。仅允许在当前未被占用的分箱中添加新特征，从而确保特征间距不小于最小值。如果检测过程中向某个分箱添加了多个特征，则仅保留检测得分最高的特征。

检测后，每个候选特征都会接受类似于 Harris 检测器 [50]所采用的基于梯度的质量评估。该筛选器使用 5×5 Sobel 核 [51] ，沿水平方向和垂直方向计算检测帧中每个特征位置的梯度幅值。若某个特征的梯度幅值超过预设阈值，则予以保留；这是因为热成像中的显著特征通常位于边缘或角点上。这一启发式方法利用边缘和角点的结构连贯性，提高特征的可靠性。

特征通过筛选后，使用 KLT 光流法 [52]在图像帧之间进行跟踪。每个特征都会与其在检测帧中检测时计算得到的 ORB 描述符 [43] 相关联。跟踪过程中，通过将存储的描述符与当前图像帧中新位置处的描述符进行比较，来验证特征的新位置。如果描述符不匹配程度超过阈值，则将该特征标记为未在当前图像帧中找到。仅将 ORB 描述符用作检查项，可以避免比较所有描述符这一耗时的组合问题。凡是未能在图像帧之间成功跟踪的特征，都会从集合中剔除。如果跟踪到同一网格单元中的特征数量超过指定数目，也会剔除多余特征。

为进一步抑制错误匹配，采用基于本质矩阵的七点 RANSAC 检验 [42] [53]，以识别相对于主导场景运动的离群值。为利用更大的视差，该验证步骤在上一个检测帧与当前帧之间进行。若某个特征在连续数帧中有超过 50 % 被判定为离群值，则将其永久移除。未通过本质矩阵检验的观测将不纳入后端滤波器更新。

为适应场景变化并保持足够的特征密度，系统会定期检测新特征。检测到新特征后，每个已有特征的当前描述子都会更新为后续匹配所用的新参考描述子，其位置也会被缓存，以供本质矩阵检验使用。光流以各特征在上一帧中的位置作为初始值，以改善收敛性并保持时间连续性。为保持基于 LRF 的尺度线索，当图像中央 20% 区域内没有活跃轨迹时，系统会在该区域检测一个新特征。

G. 特征过滤

由于热成像跟踪本身存在挑战，即使满足上述标准，特征轨迹也常常会出现错误或不可靠的情况。为提高鲁棒性，我们在将特征纳入后端之前采用多阶段过滤流程，只选择最可靠的特征测量值。

我们分批收集并评估候选特征。对于每个候选特征，都会执行一系列状态检查，以评估其几何一致性。随后，根据一组质量指标（详见下文）对通过验证的特征集进行排序，每次更新时仅将排名靠前的子集传递给后端，且新特征数量不得超过上限。

主要的过滤机制是基于三角测量的一致性检查。在特征积累至少十次有效观测，或连续 0.5 秒未收到新的观测之前，不会将其提交给后端。达到上述条件后，将使用参照 OpenVINS [49], [53] 的基于 RANSAC 的算法尝试进行三维三角测量。此步骤旨在剔除错误匹配，同时保持多图像帧间的几何一致性。三角测量始终包含该特征的首次记录测量值。只有 RANSAC 得到的内点子集会传递给优化后端。特征一旦被后端接受，后续测量便会跳过此三角测量检查，以降低计算开销和延迟。

三角测量和特征验证过程中会应用以下几项标准：

- **深度符号一致性：** 如果估计深度为负，则丢弃该特征的所有关联测量值。
- **与激光测距仪的深度一致性：**如果图像中心附近特征的三角化深度与机载激光测距仪的测量值相差显著，则将其排除。这项跨传感器一致性检查可提高低视差或存在歧义的配置下深度的可靠性。
- **三角测量质量评分：** 为量化三角测量得到的三维点的可靠性，我们根据其观测射线的空间分布计算一个评分。具体而言，我们通过评估表征多视角观测几何关系的矩阵的条件数，判断特征深度受到约束的程度。评分较高表示观测射线近乎平行，导致深度估计不可靠；评分较低则表示视差较强，三角测量的约束良好。如果三角测量得到的特征数量超过预设阈值，则仅将质量评分排名前 80% 的特征传递给后端。
- **重投影误差：** 在因子图构建过程中，如果轨迹中的任何单个测量值产生的因子误差超过规定阈值，则不会将该测量值纳入后端。如果剩余的有效测量值过少，则排除该特征。

只有在特征的观测中检测到足够的视差时，才会进行三角测量，以确保深度可观测性。然而，为了兼顾可能无法持续足够长时间以完成全面验证的瞬态特征，以及悬停期间缺乏视差的情况，如果某个特征在超时窗口内未满足观测数量或视差标准，且没有更好的替代特征，仍会考虑将其纳入后端。这种回退机制可确保在悬停或剧烈机动期间，及时纳入其他方面可行的特征。

所有阈值和排序权重（例如，最小观测数、视差阈值、LRF 一致性等）均经过人工

调优，覆盖 30 多组具有代表性的高速飞行日志，以平衡估计鲁棒性、延迟和吞吐量约束。

这一多层过滤阶段显著提高了传递至后端的特征轨迹的可靠性，尤其是在高速和低纹理条件下，而这正是单目热成像的典型情况。

H. 特征纳入

经过验证的特征通过两个相机帧之间基于投影的残差纳入后端。每个特征均由一个单位方位向量 $\mathbf{s}_i$ 和一个逆深度 ρ 参数化；前者取自首次检测到该特征的图像帧（原始根帧），后者则表示在该图像帧中的逆深度。这种逆深度形式避免了过度参数化，并为单目测量提供了紧凑的表示。

通过边缘化进行根帧跟踪：由于与原始根帧关联的导航状态之后将被边缘化，系统无法假设该特征的锚定帧会持续存在于因子图中。因此，当这种情况发生时，我们不会重新三角化或重新定义该特征，而是维护一个累积的SE(3)变换

$$^{\text{origRoot}}\mathbf{T}_{\text{currRoot}},$$

其中，当前根帧是包含该特征观测的最早活动关键帧。每当边缘化移除先前的根帧时，就会更新这一变换。该策略使特征能够继续以其原始坐标系表示，同时仅调整投影因子所使用的参考变换。

投影因子公式：令 $^{w}\mathbf{T}_{x}$ 表示导航状态 x 在世界坐标系中的机体位姿，并令 $^{b}\mathbf{T}_{c}$表示相机到机体的外参。若 x_i 为当前根状态， x_j 为观测状态，则从当前根相机坐标系到观测相机坐标系的相对变换为

$$^{c_i}\mathbf{T}_{c_j} = {}^{b}\mathbf{T}_{c}^{-1}\,{}^{w}\mathbf{T}_{x_j}^{-1}\,{}^{w}\mathbf{T}_{x_i}\,{}^{\text{currRoot}}\mathbf{T}_{\text{origRoot}}\,{}^{b}\mathbf{T}_{c}.$$

记

$$^{c_i}\mathbf{T}_{c_j} = \left({}^{c_i}\mathbf{R}_{c_j},\ {}^{c_i}\mathbf{t}_{c_j}\right),$$

观测坐标系中的预测三维点为

$$\mathbf{z} = {}^{c_i}\mathbf{R}_{c_j}\,\mathbf{s}_i + \rho\,{}^{c_i}\mathbf{t}_{c_j}.$$

为得到尺度不变残差，我们将 $\mathbf{z}$ 投影到与观测方位 $\mathbf{s}_j$正交的平面上：

$$\mathbf{P} = \mathbf{I} - \mathbf{s}_j\mathbf{s}_j^{\top}, \qquad \mathbf{r}_{\text{feature}} = \mathbf{P}\,\mathbf{z}.$$

该残差反映了预测方位方向与观测方位方向之间的偏差。这种形式无需显式引入尺度，也避免了过度参数化，因此非常适合用于优化。

深度初始化和基于*LRF*的先验：逆深度值采用混合策略进行初始化：

- 靠近光学中心的特征使用来自激光测距仪的深度，并获得较强的先验；
- 具有足够视差的特征使用三角化深度；
- 否则，将指定一个标称深度。

对于激光测距仪约束特征，深度先验的形式为

$$r_{\text{prior}} = \rho - \rho_{\text{LRF}},$$

直接将一维距离测量与每个特征的逆深度相耦合。

以往研究已证明一维距离先验在低速飞行中的有效性，而我们的实验将这一验证扩展到了高速飞行，在这种情况下，尺度漂移可能迅速累积。我们观察到，基于激光测距仪的深度先验显著提升了此类场景下的度量一致性，尤其是在直线运动和悬停阶段。这与可观测性分析的理论结论一致 [3]，该分析表明，距离约束有助于在其他情况下弱可观测的运动区间中确定尺度。

局限性：尽管这种方法有许多优点，但也存在不足。该方法的一个局限是依赖准确的IMU数据和可靠的逆深度估计——二者在高敏捷机动或悬停期间都可能退化。

I. 气压计补偿模型

为缓解高速飞行导致的气压高度估计中的静压位置误差，我们采用基于门控循环单元（GRU）神经网络[37]的数据驱动校正模型。由于 SPPE 存在时间相关性，循环架构天然适用于这项任务。我们选择 GRU，是因为它能够捕捉时间依赖关系，同时参数复杂度低于长短期记忆（LSTM）网络 [54]，且经验表现优于普通循环神经网络。

模型的输入由多个部分组成：滚转角、俯仰角、机体系加速度、机体系速度以及机体系速度平方。通过在机体坐标系中表示运动学信息并纳入滚转角和俯仰角，只要训练数据中包含方向相关的气动效应，模型就能学会这些效应。由于不纳入偏航角，我们确保模型不受训练数据采集方法的影响，成为偏航无关模型。GRU 有 128 个隐藏单元和三层。网络有两个输出头——一个输出标量修正项，另一个输出其对数方差（$\log\sigma^2$）。这样便可根据模型的置信度动态调整因子图中的权重。

该模型使用经过精心筛选的数据集进行训练，该数据集旨在充分激发影响 SPPE 的气动状态。下面详细介绍数据采集流程：

- **基准轨迹：** 沿南北方向直线飞行，每次往返均保持恒定航向。车辆航向在 180° 的前方弧段内以 15° 为增量变化，并重复相同的轨迹。

- **速度扫描：**重复基准轨迹，速度以 5 m/s 为步长递增，并包含持续时间较长的（30+ s）路段。
- **双向重复：**每条出程轨迹之后，都会沿反向路径进行返程。
- **时间多样性：**在不同日期和风速条件下重复进行各类场景，以捕捉环境变化。

气压计高度和GPS高度均经过低通滤波，并使用前50个样本进行对齐，以消除静态偏置。所有输入均经过归一化。训练所用的真值取自GPS高度。模型使用Adam优化器进行训练（$\mathrm{lr} = 10^{-3}$），序列长度为4096，批量大小为64。我们使用焦点高斯负对数似然损失训练了7个轮次，其定义如下：

$$\mathcal{L}_{\text{focal}} = \frac{1}{2N}\sum_{i=1}^{N}\left(1 - e^{-\frac{(y_i-\hat{y}_i)^2}{2\sigma_i^2}}\right)^{\gamma}\left[\log\sigma_i^2 + \frac{(y_i-\hat{y}_i)^2}{\sigma_i^2}\right],$$

其中，y_i 为真实高度校正值，$\hat{y}_i$ 为预测校正值，σ_i^2 为预测方差，γ 为焦点指数，用于控制对高残差（“困难”）样本的强调程度。预热计划在前三个训练轮次中将 γ从0线性增加至1，从而兼顾训练稳定性，并增强对较大标准化残差的加权。

运行时，经门控循环单元校正的气压测量值通过以下残差与状态估计融合：

$$\mathbf{r}_{baro} = \mathbf{r}_3^{\top}(z_{\text{baro}} - b - z_{\text{est}}),$$

其中，$\mathbf{r}_3$ 是旋转矩阵的第三行（将机体Z轴映射到世界坐标系），z_{baro} 是校正后的气压测量值，b是用于关联世界坐标系高度与惯性坐标系高度的偏置项，z_{est} 是当前估计高度。该模型使气压因子能够根据学习得到的不确定性动态调整其影响，从而在不影响估计器稳定性的前提下实现高速运行。

由于我们的门控循环单元是在单一机体的数据上训练的，其校正具有特定机型适用性，并假设地速≈ 空速；如果这些假设不再成立，性能可能会下降。我们确实在数据集中纳入了大风天气下的几次飞行（30+ 英里/小时），补偿模型仍然相当有效。我们通过经验测试了将角速度加入输入，但发现（与 [4]一致）SPPE主要由速度和加速度决定，因此当前配置在精度和模型简洁性之间实现了最佳权衡。

J. 激光测距仪预处理

机载激光测距仪提供相对于地形的高频（≈10 Hz）高度测量值。虽然总体上较为精确，但其读数可能因取决于表面的反射率、短暂数据丢失，或在高对比度边界附近的载具运动而出现瞬态尖峰。为提高鲁棒性并抑制这些伪影，我们采用基于简单匀速运动模型的门控卡尔曼滤波器。该滤波器估计高度及其变化率，并利用统计门控剔除残差超过马氏距离阈值的测量值，从而有效去除高频噪声和严重离群值。

由此得到的高度估计在时间上平滑且可靠，适合用于后端优化和特征验证。其逆值既用于初始化新特征的逆深度，也用于为图像帧中心附近观测到的特征提供先验。

K. 姿态处理

我们的系统并未直接纳入原始磁力计测量值，而是利用 PX4 提供的融合姿态估计。该估计由 PX4 的机载扩展卡尔曼滤波器得出，该滤波器紧密融合磁力计和惯性测量单元的测量值，从而生成经过漂移校正的姿态。除了提供磁参考的偏航角外，PX4 姿态先验还有助于保持机载控制器所采用的姿态与视觉惯性里程计估计之间的一致性。

尽管 PX4 姿态包含完整的三维姿态信息，我们仅纳入其偏航角分量，因为滚转角和俯仰角已受到惯性预积分和热惯性里程计流程的充分约束。PX4 得出的偏航角先验采用通常用于磁力计对准的相同残差结构纳入：

$$\mathbf{r}_{att} = \mathbf{v}_{\text{px4}} - ({}^{w}\mathbf{R}_b)^{\top}\mathbf{v}_{\text{wld}},$$

其中，$\mathbf{v}_{\text{px4}}$ 是 PX4 姿态中磁北方向在机体坐标系下的水平面投影，$\mathbf{v}_{\text{wld}} = [1,0,0]^{\top}$ 表示世界坐标系中的磁北，而 ${}^{w}\mathbf{R}_b \in SO(3)$ 是估计得到的机体到世界坐标系的旋转。从概念上讲，该残差用于惩罚 PX4 磁参考姿态与视觉惯性里程计姿态估计之间的偏航角偏差。

L. 关键帧与状态结构

为控制计算复杂度并支持实时运行，我们采用滑动窗口方法，并使用混合关键帧结构。每隔 20 个图像帧创建一个新的关键帧。每个关键帧都包含完整的惯性导航状态：位姿、速度、姿态和 IMU偏置（$\mathbf{b}_g$,$\mathbf{b}_a$）。中间的 20 个普通图像帧仅包含位置、旋转和速度状态，不包含偏置项。我们将中间图像帧之前的关键帧称为它们的父关键帧。

这种结构既能高效地对 IMU动力学建模，又能限制高维状态的数量。关键的是，所有图像帧之间的 IMU预积分因子都共享并引用最近关键帧的 IMU偏置。假设这些偏置在每个 20 帧组内保持不变，并以此为基础对普通图像帧之间以及直至下一个关键帧的惯性测量值进行积分。

在后端优化过程中，利用先验关键帧的共用偏置估计，在连续帧之间构建预积分残差。这使得惯性建模能够保持一致，而无需在每一图像帧进行偏置估计，从而同时降低内存占用与计算量。

M. 后端优化

后端采用非线性最小二乘优化，通过最小化来自多种传感器模态的加权残差之和，联合估计导航状态与特征逆深度。整体优化问题表述如下：

$$\min_{\mathbf{x}} \sum_{i} \|\mathbf{r}_i(\mathbf{x})\|^2_{\mathbf{W}_i}, \tag{1}$$

其中，$\mathbf{x}$ 表示导航状态的集合（位置 $\mathbf{p}$、惯性速度 $\mathbf{v}$、姿态 $\mathbf{R}$、IMU偏置 $\mathbf{b}_g$,$\mathbf{b}_a$），以及特征逆深度 ρ。权重矩阵 $\mathbf{W}_i$ 反映了逆协方差，并根据测量的确定性对残差贡献进行平衡。

为确保计算效率而不牺牲精度，我们采用了一种受 ICE-BA [17] 启发的选择性重线性化策略。我们不会在每次迭代时都对所有变量进行重线性化，而只对变化超过预设阈值的变量子集进行重线性化。这种方法减少了雅可比矩阵计算和残差计算的次数，从而即使在高频率估计需求下，也能满足实时运行要求。

每次优化迭代都会执行以下步骤：1）评估每个变量的线性化点变化。2）标记变化超过阈值的变量。3）应用变化，重新计算雅可比矩阵，并重新评估已标记子集的残差。4）继续进行 LMPCG 优化。

优化采用 LMPCG [17], [45]–[48], ，该方法结合了信赖域列文伯格-马夸尔特（LM）阻尼与预处理共轭梯度（PCG）的高效性。这种组合对于稀疏结构化系统尤为有益。LM正则化引入阻尼项，通过确保海森矩阵保持正定来稳定解，即使某些参数（例如特征逆深度）约束不足时也是如此。线性化系统如下：

$$(\mathbf{J}^\top \mathbf{W} \mathbf{J} + \lambda \mathbf{I}) \Delta \mathbf{x} = -\mathbf{J}^\top \mathbf{W} \mathbf{r},$$

其中，$\mathbf{J}$ 是残差相对于状态向量$\mathbf{x}$的雅可比矩阵，$\mathbf{W}$ 是块对角权重矩阵，λ 是LM阻尼参数，$\Delta\mathbf{x}$是待求的更新向量。

我们定义阻尼海森矩阵$\mathbf{H} = \mathbf{J}^T\mathbf{W}\mathbf{J} + \lambda\mathbf{I}$ 和梯度向量$\mathbf{d} = -\mathbf{J}^T\mathbf{W}\mathbf{r}$，并将其分块如下：

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{C} \end{bmatrix}, \quad \Delta\mathbf{x} = \begin{bmatrix} \Delta\mathbf{x}_B \\ \Delta\mathbf{x}_C \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} \mathbf{d}_B \\ \mathbf{d}_C \end{bmatrix},$$

其中，$\mathbf{B}$ 和 $\Delta\mathbf{x}_B$ 对应导航状态，$\mathbf{C}$ 和$\Delta\mathbf{x}_C$ 对应特征逆深度，而$\mathbf{E}$ 对应交叉项。

通过应用舒尔补，我们消去特征深度：

$$\begin{aligned} \mathbf{S} &= \mathbf{B} - \mathbf{E}\mathbf{C}^{-1}\mathbf{E}^\top, \\ \mathbf{v} &= \mathbf{d}_B - \mathbf{E}\mathbf{C}^{-1}\mathbf{d}_C, \\ \Delta\mathbf{x}_C &= \mathbf{C}^{-1}(\mathbf{d}_C - \mathbf{E}^\top \Delta\mathbf{x}_B), \end{aligned}$$

这很高效，因为$\mathbf{C}$是块对角的，而且很容易求逆。

然后，我们应用 PCG 迭代求解约化系统$\mathbf{S}\Delta\mathbf{x}_B = \mathbf{v}$。对于特征值高度聚集的系统，PCG 的效率优于 QR分解，因此更受青睐。此外，作为一种迭代求解器，PCG 可以在固定迭代次数后提前终止，从而确保运行时间有界。

这种方法保留了稀疏性、提高了鲁棒性，并支持可扩展优化，适用于嵌入式高频率视觉惯性里程计系统。优化后的状态会与IMU测量值在高频率EKF中融合。随后，EKF状态以30 Hz的频率发送至飞行控制器，用于闭环控制。

N. 边缘化策略

为保持计算复杂度有界，我们使用基于舒尔补的先验对旧状态进行边缘化。令 x_{oldest} 表示最旧的关键帧，x_{next} 表示其后继关键帧。具体步骤如下：

1)选择要移除的状态。我们收集父关键帧为 x_{oldest}（包括 x_{oldest} 本身）的所有导航状态和IMU偏置状态。我们还收集根关键帧为 x_{oldest}的所有特征状态。

2) 复制在之后仍会继续使用的特征状态 x_{next}。如果某个特征在边缘化窗口内有测量值，同时也与 x_{next} 或更晚的关键帧相关联，我们就会拆分其表示。创建该特征状态的克隆副本（如果存在先验，也一并复制），并仅将其关联到边缘化窗口内的因子；随后对这个克隆状态进行边缘化。保留原始特征状态，但不重新三角测量其逆深度，而是通过插入这些关键帧之间已知的相对变换，将其根导航状态从 x_{oldest} 移至 x_{next} 。然后，将所有剩余因子重新关联到 x_{next}。

3) 隔离临时边缘化块。凡涉及任何待边缘化导航状态或任何克隆特征状态的因子，均从全局容器中移除并放入临时集合。提取之后，这些因子不再连接到任何被保留的状态，x_{next}除外。因此，全局海森矩阵可置换为

$$\mathbf{H}_{\text{global}} = \begin{bmatrix} \mathbf{H}_{\text{marg}} & 0 \\ 0 & \mathbf{H}_{\text{keep}} \end{bmatrix}, \qquad \mathbf{r}_{\text{global}} = \begin{bmatrix} \mathbf{r}_{\text{marg}} \\ \mathbf{r}_{\text{keep}} \end{bmatrix}.$$

因此，边缘化完全在$\mathbf{H}^{\text{marg}}$上进行。

4) 构建临时海森矩阵与残差。在应用所有延迟状态更新之后，我们对临时容器中的全部因子进行线性化。由此得到一个分块结构系统

$$\mathbf{H}_{\text{marg}} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \mathbf{H}_{13} \\ \mathbf{H}_{12}^\top & \mathbf{H}_{22} & \mathbf{H}_{23} \\ \mathbf{H}_{13}^\top & \mathbf{H}_{23}^\top & \mathbf{H}_{33} \end{bmatrix}, \qquad \mathbf{r}_{\text{marg}} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix},$$

其中：

- $\mathbf{H}_{11}$，$\mathbf{r}_1$ 仅包含与正在被边缘化的导航/惯性测量单元状态相关的贡献，
- $\mathbf{H}_{33}$,$\mathbf{r}_3$ 包含仅涉及克隆特征状态的贡献，

- $\mathbf{H}_{22}$,$\mathbf{r}_2$ 包含涉及$x_{\text{下一个}}$时刻的状态（其导航状态和IMU偏置）的贡献，以及
- $\mathbf{H}_{12}$、$\mathbf{H}_{13}$、$\mathbf{H}_{23}$ 是交叉项。

5) **消除克隆特征状态**。由于$\mathbf{H}_{33}$ 是块对角矩阵，因此求逆开销很低，我们应用舒尔补来移除克隆特征状态。用波浪号表示更新后的量，

$$\begin{aligned}\tilde{\mathbf{r}}_1 &= \mathbf{r}_1 - \mathbf{H}_{13}\mathbf{H}_{33}^{-1}\mathbf{r}_3,\\ \tilde{\mathbf{r}}_2 &= \mathbf{r}_2 - \mathbf{H}_{23}\mathbf{H}_{33}^{-1}\mathbf{r}_3,\\ \tilde{\mathbf{H}}_{11} &= \mathbf{H}_{11} - \mathbf{H}_{13}\mathbf{H}_{33}^{-1}\mathbf{H}_{13}^{\top},\\ \tilde{\mathbf{H}}_{22} &= \mathbf{H}_{22} - \mathbf{H}_{23}\mathbf{H}_{33}^{-1}\mathbf{H}_{23}^{\top},\\ \tilde{\mathbf{H}}_{12} &= \mathbf{H}_{12} - \mathbf{H}_{13}\mathbf{H}_{33}^{-1}\mathbf{H}_{23}^{\top}.\end{aligned}$$

6) **消除关键帧之间的导航/惯性测量单元状态**。接下来，我们使用相同的方法移除与x_{oldest} 相关的其余导航状态：

$$\begin{aligned}\bar{\mathbf{H}}_{22} &= \tilde{\mathbf{H}}_{22} - \tilde{\mathbf{H}}_{12}^{\top}\tilde{\mathbf{H}}_{11}^{-1}\tilde{\mathbf{H}}_{12},\\ \bar{\mathbf{r}}_2 &= \tilde{\mathbf{r}}_2 - \tilde{\mathbf{H}}_{12}^{\top}\tilde{\mathbf{H}}_{11}^{-1}\tilde{\mathbf{r}}_1.\end{aligned}$$

$^{-}$二元组（$\mathbf{H}_{22}$、$^{-}\mathbf{r}_2$）将被移除的导航、惯性测量单元和特征状态的全部信息，作为关键帧 x_{next}的先验。

7) **正则化与先验因子构造**。我们$^{-}$对$\mathbf{H}_{22}$ 进行对称化并截断其奇异值，以确保数值稳定性，然后计算隐含的先验均值

$$\boldsymbol{\mu}_{\text{prior}} = \bar{\mathbf{H}}_{22}^{-1}\bar{\mathbf{r}}_2.$$

使用$^{-}$（$\mathbf{H}_{22}$，$\boldsymbol{\mu}_{\text{prior}}$）构建 Nav+IMU 偏置先验因子，并将其添加到 x_{next}。随后丢弃边缘化后的状态及其所有因子。

V. 结果

本节将介绍所述方法的实验验证，重点展示本文提出的贡献。我们通过专门设计的实验，系统地评估各个组件对系统性能的影响。首先，我们通过与基线进行比较，展示特征跟踪流程的有效性和鲁棒性。其次，我们展示 SPPE 补偿器对估计轨迹带来的改进，并将其与现有方法进行比较。接着，我们展示对高频 IMU 数据进行降采样前预滤波所带来的改进。最后，我们考察闭环飞行性能，包括估计器漂移和时序特性。为便于直接比较，除闭环实验外，其他飞行的记录数据均在笔记本电脑上重放，并分别禁用各个流程组件。真值轨迹由机载 GPS 得出。图3展示了飞行轨迹类型的示例。

为量化估计器性能，我们报告每条轨迹的均方根误差（RMSE）和漂移率。计算RMSE时，首先对估计轨迹和由GPS测量值推导出的轨迹进行时间对齐。我们将GPS测量值插值到估计器的时间戳，然后计算每个匹配样本的欧几里得位置误差。报告的RMSE是整个飞行过程中位置误差平方均值的平方根。漂移以最终位置误差除以总行驶距离来衡量，并以百分比表示，从而提供一种经过尺度归一化的长期一致性指标。这些指标共同刻画了估计器的瞬时精度和长时间尺度下的稳定性。

A. 实验*1*：热优化特征处理流程

我们将热优化特征跟踪流程与基线进行比较。基线由高斯滤波、限制对比度自适应直方图均衡化[40]，FAST 特征检测 [41]，Lucas–Kanade光流 [52],以及基于本质矩阵的异常值剔除组成。我们逐步引入改进项——基于梯度的滤波、ORB 描述符 [43]滤波和基于三角测量的滤波，以量化各项改进单独及组合后的效果。为确保测试具有多样性，我们选取了五份具有代表性的飞行日志，其峰值速度范围为 20米/秒至30米/秒。其中两份日志因固定模式噪声（FPN）水平较高而被特别选中，其噪声水平超过了典型条件。表 I 汇总了所有日志的均方根误差均值和标准差。如表所示，每增加一个滤波阶段，都会降低整体误差幅度或减少误差波动；完整的滤波配置在所有数据集上始终表现最佳。虽然表中报告的是累积改进效果，但我们注意到，单独应用每个滤波步骤时也能提升性能；为简洁起见，此处省略了这些结果。

表 I：实验1：特征处理流程消融

方法	均方根误差均值 (m)	均方根误差标准差 (m)
基线	19.35	7.03
+ 梯度滤波	16.63	7.63
+ ORB描述子滤波	12.43	5.23
+ 三角测量滤波	12.83	3.65

B. 实验*2*：气压补偿效果

为了评估我们的基于GRU的气压补偿器，我们使用一组留出的测试数据集，将其与传统的多项式拟合方法（如PX4[35]和 Ardupilot [36]）以及一种基于多层感知机的方法进行比较。图4展示了一个示例轨迹中，GPS高度和低通滤波后的气压计高度与三种补偿方法的对比结果。此外，我们使用五次示例飞行来比较这些方法对估计轨迹的影响。如图4所示，在一条具有代表性的高速轨迹上，GRU校正得到的结果与GPS高度最为吻合。在这段序列中，均方根误差从 4.66 m（滤波气压计）降至使用GRU时的 2.19 m，改善了 2.47 m。此外，基于GRU的方法不仅在均值方面表现更好，其不确定性的三倍标准差界限也几乎涵盖了误差，从而进一步提升了性能。在整个验证集上也观察到了类似的改进（表 II）。

表II展示了留出测试集的均方根误差，以及平均Z轴均方根误差、平均轨迹均方根误差、平均端点误差（端点位置误差）和平均轨迹漂移率

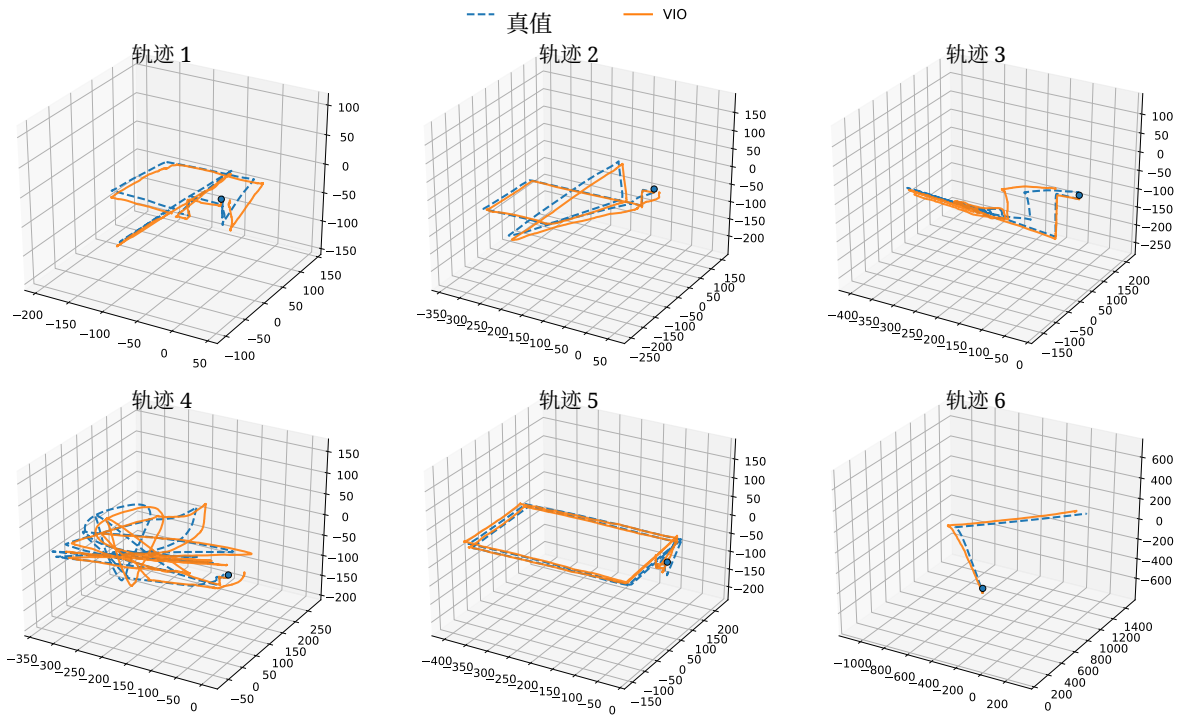


图3：飞行实验中的示例轨迹　真值参考由机载GPS测量获得

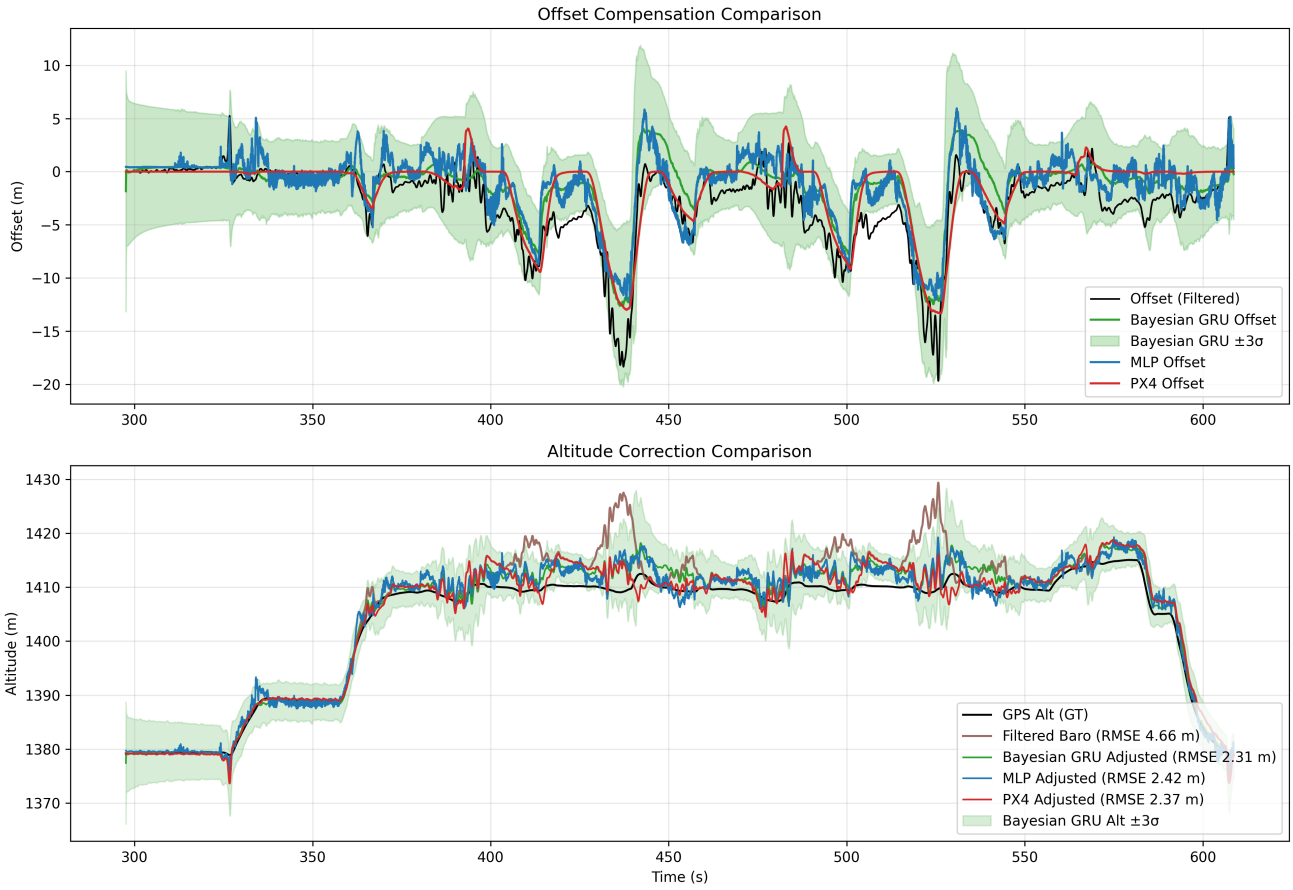


图4：滤波、多项式、多层感知机和门控循环单元校正的气压计误差随时间变化情况。

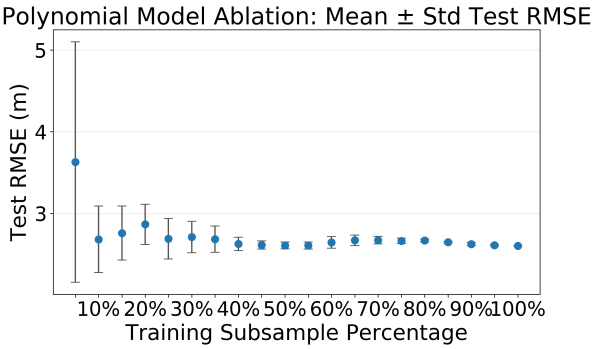


图5：使用训练数据的随机子样本训练的多项式模型的测试误差均值和标准差

在所考察的五次高速飞行中，每种情况下，我们提出的基于GRU的补偿器都优于其他方法。

表II：气压补偿方法比较

方法	测试均方根误差	Z轴均方根误差	轨迹均方根误差	漂移（%）
滤波气压计	3.35	4.07	38.93	3.20
多项式拟合	2.60	2.93	29.14	2.11
MLP	2.59	3.58	31.99	2.12
GRU（我们的方法）	**2.00**	**2.82**	**25.03**	**1.27**

基于深度学习的方法有一个缺点：训练有效模型需要大量数据。由于多项式拟合方法的参数较少，直觉上会认为它所需的数据更少，同时仍能获得预期的性能提升。为了检验这一判断，我们设计了一个蒙特卡洛训练场景。对于每次训练测试，从整个训练集中随机选取一定比例的训练数据作为一个数据块。然后，使用这些子采样数据训练模型，并在留出的验证集上进行测试。对于每个比例，这一过程重复30次，并计算分布的均值和标准差。如图5所示，至少需要使用采集数据的45%，结果才会变得可靠。这表明，多项式方法看似具有的数据效率优势，在实际中远没有预想的那么明显。这很可能是因为需要足够丰富的姿态变化，才能构建具有代表性的空气动力学模型。我们也对基于GRU的方法重复了这一实验，结果相似。

尽管与多项式拟合相比，我们基于门控循环单元的方法具有更多可调参数，但它仍然轻量且兼容实时运行。我们的实现仅使用 252,546 个参数，运行时开销可以忽略不计，并且训练完成后精度更高。

C. 实验 *3*：高频*IMU*预滤波的影响

我们使用120赫兹的原始惯性测量单元测量值，与我们提出的IIR滤波并降采样方法进行比较，以评估估计器性能。表III中的结果展示了一条具有多个急剧转向的典型4公里轨迹的端点位置误差、漂移率和均方根误差。该轨迹的图示以及我们估计的轨迹，可见图3中的轨迹4。这些结果表明，在激进飞行状态下，我们的预滤波和降采样方法比不滤波的方法性能更佳。该方法还可避免有限的处理器间通信带宽达到饱和，从而提高可靠性。

表III：IMU 处理方法比较

方法	端点位置误差（m）	漂移（%）	均方根误差（米）
120赫兹原始预积分	150.07	3.60	77.68
1200赫兹 + 降采样	**31.24**	**0.75**	**24.96**

D. 实验 *4*：闭环飞行与延迟

我们在 NVIDIA Jetson Xavier NX 上，使用一条以30米/秒 飞行的箱形轨迹，验证了所提出估计器的实时闭环飞行性能。机载估计器以 30 赫兹的频率向飞行控制器提供位姿更新，以实现完全闭环运行。表 IV 总结了各主要处理流程组件的平均执行延迟。整个估计循环的平均周期时间为 42 毫秒，即使在激烈机动期间也能实现闭环飞行。

表 V 报告了相对于 GPS真值 的最终轨迹精度。在 3.1 千米的箱形轨迹上，估计器的端点位置误差为 31.5 米，对应的漂移率为 1.0%。这些结果表明，该系统能够在嵌入式硬件上全程实时运行，同时为高速 GPS拒止飞行提供状态估计。

表 IV：闭环延迟（Xavier NX）

组件	均值（毫秒）	Std
图像预处理	11.63	2.91
特征检测	1.74	1.08
特征跟踪	4.66	1.60
图优化 + 边缘化	23.9	23.4
总循环时间	**41.93**	**28.99**

表 V：闭环误差

端点位置误差（m）	漂移（%）	均方根误差（m）
31.51	1.00	19.23

VI. 结论

本文提出了一种单目热惯性里程计系统，专为在视觉条件恶劣、无 GPS 的环境中高速飞行而设计。我们采用全面的方法应对热成像和高速度所特有的关键挑战，将稳健的热成像优化特征跟踪流程、高频IMU预滤波以及新颖的基于 GRU 的气压补偿模型整合在一起。

通过严格的实验评估，我们展示了各个组件带来的显著性能提升。优化后的特征跟踪流程显著

降低了估计漂移和端点误差，相较于传统方法表现更优。所提出的基于GRU的气压补偿器始终优于传统的多项式校正和基于MLP的校正，能够有效捕捉时间依赖关系，提升高度估计精度。此外，我们的IMU预滤波策略显著提高了剧烈机动过程中的估计器稳定性，并在不影响性能的前提下实现了高效的带宽利用。

总体而言，这些进展实现了飞行速度达30米/秒的稳健实时热惯性里程计。该系统保持了较低的漂移（<1.3%）和较小的平均均方根误差（<26米），验证了其适用于要求严苛的空中机器人应用。

未来的工作包括将我们的系统扩展至涉及可见光/红外组合的多传感器融合方法，利用IMU数据更好地筛选特征轨迹，并改进热成像相机与IMU的标定。

参考文献

[1] A. I. Mourikis 和 S. I. Roumeliotis，“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。[2] T. Qin、J. Pan、S. Cao 和 S. Shen，“一种利用多传感器进行局部里程计估计的通用优化框架”，2019年。[3] J. Delaune、D. S. Bayard 和 R. Brockers，“xvio：一种测距视觉惯性里程计框架”，2020年。 [在线]。网址：https: //arxiv.org/abs/ 2010 06677 。N. Xuan-Mung 和 S.-K. Hong，“用于改进四[4]旋翼飞行器高度控制的气压高度测量故障诊断”，载于《2019年第19届国际控制、自动化与系统会议（ICCAS）》，2019年，第1359–1364页。[5] T. Qin、P. Li 和 S. Shen，“VINS-Mono：一种稳健且通用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。[6] S. Khattak、C. Papachristos 和 K. Alexis，“视觉退化环境中的视觉热成像地标与惯性融合导航”，载于《2019年IEEE航空航天会议》，2019年，第1–9页。[7] FLIR Systems Inc.，“Flir boson热成像核心组件产品数据表”，http://elimec.co.il/ Uploads/dbsAttachedFiles/Boson.pdf，2019年，访问日期：2023年9月22日。[8] I. Jarraya、A. Al-Batati、M. B. Kadri、M. Abdelkader、A. Ammar、W. Boulila 和 A. Koubaa，“无GNSS无人机导航：计算复杂度、传感器融合与定位方法分析”，《卫星导航》，第 6, 1, 9, 卷，2025 [ ] 10 1186 43020 025 00162 期，页码。在线。网址：https://doi.org/ . /s - - -z [9] K. Sun、K. Mohta、B. Pfrommer、M. Watterson、S. Liu、Y. Mulgaonkar、C. J. Taylor 和 V. Kumar，“适用于快速自主飞行的稳健双目视觉惯性里程计”，《IEEE机器人与自动化快报》，第3卷，第2期，第965–972页，2018年。[10] Z. Huai 和 G. Huang，“机器人中心视觉惯性里程计”，《国际机器人研究杂志》，第41卷，第7期，第667–689页，2022年。[11] M. Bloesch、S. Omari、M. Hutter 和 R. Siegwart，“使用基于直接EKF的方法实现稳健的视觉惯性里程计”，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。[12] A. Spaenlehauer、V. Fremont、A. Sekercioglu 和 I. Fantoni，“一种用于单目视觉惯性系统度量尺度估计的松耦合方法”，2017年11月。[13] S. Weiss、M. W. Achtelik、S. Lynen、M. Chli 和 R. Siegwart，“未知环境中微型飞行器的实时机载视觉惯性状态估计与自标定”，载于《2012年IEEE国际机器人与自动化会议》，2012年，第957–964页。[14] S. Leutenegger、S. Lynen、M. Bosse、R. Siegwart 和 P. Furgale，“基于关键帧的视觉惯性里程计：采用非线性优化”，《国际机器人研究杂志》，第34卷，2014年2月。[15] T. Lupton 和 S. Sukkarieh，“在无初始条件的建筑环境中实现高动态运动的视觉惯性辅助导航”，《IEEE机器人学汇刊》，第28卷，第1期，第61–76页，2012年。[16] C. Forster、L. Carlone、F. Dellaert 和 D. Scaramuzza，“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。[17] H. Liu、M. Chen、G. Zhang、H. Bao 和 Y. Bao，“ICE-BA：用于视觉惯性同步定位与地图构建的增量式、一致且高效的光束法平差”，载于《2018年IEEE/CVF计算机视觉与模式识别会议》，2018年，第1974–1982页。[18] S. Zhai、N. Wang、X. Wang、D. Chen、W. Xie、H. Bao 和 G. Zhang，“XR-VIO：面向 XR 应用、具有快速初始化功能的高精度视觉惯性里程计”，2025年。 [在线]。网址：https://arxiv.org/abs/ 2502 01297 。J. Delaune、R. Hewitt、L. Lytle、[19]C. Sorice、R. Thakker 和 L. Matthies，“适用于全天候夜间自主飞行的热惯性里程计”，载于《2019年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2019年，第1122–1128页。[20] S. Zhao、P. Wang、H. Zhang、Z. Fang 和 S. Scherer，“TP-TIO：一种具有深度热特征点的稳健热惯性里程计 ”，载于《2020年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2020年，第4505–4512页。[21] D. DeTone、T. Malisiewicz 和 A. Rabinovich，“SuperPoint：用于兴趣点检测与描述的自监督方法”，载于《2018年IEEE/CVF计算机视觉与模式识别会议研讨会（CVPRW）》，2018年，第337–33 712页。

[22] J. Lee、T. Ando、M. Shinozaki、T. Kitajima、Q. An 和 A. Yamashita，“Self-TIO：通过自监督 16 位特征提取器和跟踪器实现热惯性里程计”，《IEEE机器人与自动化快报》，第 10 卷，第 2 期，第 1003–1010 页，2025年。[23] Z. Teed 和 J. Deng，“RAFT：用于光流的循环全对场变换”，2020 [ ] 2003 12039 。在线。[24]可访问：https://arxiv.org/abs/。M. R. U. Saputra、P. P. B. de Gusmao、C. X. Lu、Y. Almalioglu、S. Rosa、C. Chen、J. Wahlstr¨om、W. Wang、A. Markham 和 N. Trigoni，“DeepTIO：一种具有视觉幻觉的深度热惯性里程计”，《IEEE机器人与自动化快报》，第 5 卷，2020年，第 2 期，第 1672–1679 页。[25] Y. Wang、H. Chen、Y. Liu 和 S. Zhang，“视觉退化环境中的基于边缘的单目热惯性里程计”，《IEEE机器人与自动化快报》，第 8 卷，第 4 期，第 2078–2085 页，2023年。[26] J. J. Tarrio 和 S. Pedre，“单目相机的实时基于边缘的视觉里程计”，载于《2015年IEEE国际计算机视觉会议（ICCV）》，2015年，第702–710 页。[27] F. Schenk 和 F. Fraundorfer，“使用机器学习检测的边缘实现鲁棒的基于边缘的视觉里程计”，载于《2017年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2017年，第 1297–1304 页。[28] W. Baker 和 R. Clem，“地形等高线匹配（TERCOM）入门”，航空系统部，赖特-帕特森空军基地，技术报告 ASP-TR-77-61，1977年，航空系统部，赖特-帕特森空军基地，俄亥俄州。[29] J. P. Golden，“地形等高线匹配（TERCOM）：一种巡航导弹制导辅助系统”，载于《导弹制导图像处理》，第 238 卷。SPIE，1980年，第 10–18 页。[30] J. Hollowell，“Heli/SITAN：一种用于直升机的地形参照导航算法”，载于《IEEE定位与导航研讨会：导航科学领域卓越十年》，1990年，第 616–625 页。[31] D. R. Maggio、C. Mario、B. Streetman、T. Steiner 和 L. Carlone，《高空气球和亚轨道火箭上的基于视觉的地形相对导航》。 [在线]。可访问：https://arc.aiaa.org/doi/abs/10.2514/6.2023-0875[32] M. Burri、J. Nikolic、P. Gohl、T. Schneider、J. Rehder、S. Omari、M. W. Achtelik 和 R. Siegwart，“EuRoC微型飞行器数据集”，《国际机器人研究杂志》，2016年。 [在线]。可访问：
h
tt
p://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.abstract
[33] Z. Zou、Y. Lyu、C. Zhao、X. Kao、J. Liu 和 H. Chai，“SwiftBase：一个基于高频视觉测量、用于高速运动场景下视觉惯性定位的数据集”，载于《2024年IEEE/RSJ智能机器人与系统国际会议（IROS）》，2024年，第 10 206–10 213 页。[34] E. Kaufmann、L. Bauersfeld、A. Loquercio、M. M¨uller、V. Koltun 和 D. Scaramuzza，“利用深度强化学习实现冠军级无人机竞速”，《自然》，第 620 卷，第 7976 期，第 982–987 页，2023年。[35] PX4自动驾驶仪开发者，“PX4自动驾驶仪”，https://px4.io/，2025年，版本 1.14.3，访问日期：2025-10-28。[36] ArduPilot开发团队，“ArduPilot开源自动驾驶仪软件”，https://ardupilot.org/，2025年，版本 4.5，访问日期：2025-10-28。[37] K. Cho、B. van Merri¨enboer、D. Bahdanau 和 Y. Bengio，“论神经机器翻译的特性：编码器—解码器方法”，载于《SSST-8论文集：第八届统计翻译中的句法、语义与结构研讨会》，D. Wu、M. Carpuat、X. Carreras 和 E. M. Vecchi 编。卡塔尔多哈：计算语言学协会，2014年10月，第 103–111 页。 [在线]。可访问：https://aclanthology.org/W14-4012/ [38] J. Rehder、J. Nikolic、T. Schneider、T. Hinzmann 和 R. Siegwart，“扩展Kalibr：标定多个惯性测量单元的外参及单独轴”，载于《2016年IEEE国际机器人与自动化会议（ICRA）》，2016年，第 4304–4311 页。[39] S. Garrido-Jurado、R. Mu˜noz-Salinas、F. Madrid-Cuevas 和 R. Medina-Carnicer，“使用混合整数线性规划生成基准标记字典”，《模式识别》，第 51 卷，2015年10月。[40] S. M. Pizer、E. P. Amburn、J. D. Austin、R. Cromartie、A. Geselowitz、T. Greer、B. ter Haar Romeny、J. B. Zimmerman 和 K. Zuiderveld，“自适应直方图均衡化及其变体”，《计算机视觉、图形学与图像处理》，第 39 卷，第 3 期，第 355–368 页，1987年。[在线]。可访问：https://www.sciencedirect.com/science/article/pii/S0734189X8780186X[41] E. Rosten 和 T. Drummond，“高速角点检测的机器学习方法”，载于《计算机视觉—ECCV 2006》，A. Leonardis、H. Bischof 和 A. Pinz 编。柏林、海德堡：Springer Berlin Heidelberg，2006年，第 430–443 页。[42] R. Hartley，“为八点算法辩护”，《IEEE模式分析与机器智能汇刊》，第 19 卷，第 6 期，第 580–593 页，1997年。

[43] 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。[44] P. Virtanen、R. Gommers、T. E. Oliphant、M. Haberland、T. Reddy、D. Cournapeau、E. Burovski、P. Peterson、W. Weckesser、J. Bright、S. J. van der Walt、M. Brett、J. Wilson、K. J. Millman、N. Mayorov、A. R. J. Nelson、E. Jones、R. Kern、E. Larson、C. J. Carey、 ̇I. Polat、Y. Feng、E. W. Moore、J. VanderPlas、D. Laxalde、J. Perktold、R. Cimrman、I. Henrik-sen、E. A. Quintero、C. R. Harris、A. M. Archibald、A. H. Ribeiro、F. Pedregosa、P. van Mulbregt 和 SciPy 1.0 贡献者，“SciPy 1.0：Python科学计算的基础算法”，《自然·方法》，第 17 卷，第 261–272 页，2020年。[45] M. R. Hestenes 和 E. Stiefel，“求解线性方程组的共轭梯度方法”，《美国国家标准局研究杂志》，第 49 卷，第 409–436 页，1952年。[46] S. Agarwal、N. Snavely、S. M. Seitz 和 R. Szeliski，“大规模光束法平差”，载于计算机视觉——*ECCV 2010*，K. Daniilidis、P. Maragos 和 N. Paragios 编。柏林、海德堡：Springer Berlin Heidelberg，2010年，第 29–42 页。[47] M. Byr ̈od 和 K. ̊Astr ̈om，“共轭梯度光束法平差”，载于计算机视觉——*ECCV 2010*，K. Daniilidis、P. Maragos 和 N. Paragios 编。柏林、海德堡：Springer Berlin Heidelberg，2010年，第 114–127 页。

[48] Y. Jeong、D. Nister、D. Steedly、R. Szeliski 和 I.-S. Kweon，“推动光束法平差现代方法的极限”，载于 *2010*年*IEEE*计算机学会计算机视觉与模式识别会议，2010年，第1474–1481页。[49] P. Geneva、K. Eckenhoff、W. Lee、Y. Yang 和 G. Huang，“OpenVINS：视觉惯性估计研究平台”，载于 *IEEE*国际机器人与自动化会议论文集，法国巴黎，2020 [ ] \ { } 。在线。获[50]取地址：url https://github.com/rpng/open vins C. Harris 和 M. Stēphens，“一种结合角点与边缘的检测器”，载于 *Alvey*视觉会议论文集，1988年，第147–151页。[51] I. Sobel 和 G. Feldman，“用于图像处理的 3×3 各向同性梯度算子”，《模式分类与场景分析》，第271–272页，1973年1月。[52] B. D. Lucas 和 T. Kanade，“一种迭代图像配准技术及其在双目视觉中的应用”，载于 国际人工智能联合会议（*IJCAI*），1981年，第674–679页。[53] M. A. Fischler 和 R. C. Bolles，“随机抽样一致性：一种模型拟合范式及其在图像分析和自动制图中的应用”，《美国计算机协会通讯》，第24卷，第6期，第381–395页，6月 1981 [ ] 10 1145 358669 358692 。在线。[54]获取地址：https://doi.org/。/。S. Hochreiter 和 J. Schmidhuber，“长短期记忆”，《神经计算》，第9卷，第8期，第1735–1780页，1997年。