

一种基于信息熵的关键帧优化方法

张熙¹, 谭年清²

¹中国民用航空飞行学院/航空电子电气学院, 四川 广汉

²中国民用航空飞行学院/航空电子电气学院, 四川 广汉

发布日期: 2025年4月11日

摘要

针对VINS-Fusion算法在关键帧选取过程中主要依赖视差大小和特征跟踪质量的启发式策略, 难以有效度量地图全局不确定性, 可能导致定位精度不足的问题, 本文提出一种基于信息熵理论的关键帧选取优化方法。首先, 以四旋翼无人机为硬件平台, 系统分析了VINS-Fusion算法中前端特征跟踪与后端滑动窗口优化的关键流程; 进而在关键帧插入环节, 引入信息熵度量来准确刻画地图点及位姿估计的不确定性, 通过设置熵值阈值, 实现关键帧插入时机的自适应优化, 从而有效降低全局累积误差。最后, 基于EuRoC开源数据集进行仿真实验验证, 并与原始VINS-Fusion算法进行系统性对比分析。实验结果表明, 融合信息熵的关键帧选择策略能够显著减少冗余关键帧数量, 在一定程度上降低地图点的不确定性, 提升系统的全局轨迹估计精度, 并且对算法实时性仅产生较小的影响。

关键词

无人机室内定位; 视觉惯导融合; 视觉惯性里程计; 信息熵

A Keyframe Optimization Method Based on Information Entropy

Xi Zhang¹, Nianqing Tan²

¹Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan, Sichuan.

²Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan, Sichuan.

Abstract

In response to the issue that the VINS-Fusion algorithm primarily relies on heuristic strategies based on parallax magnitude and feature tracking quality during keyframe selection, making it difficult to effectively measure global map uncertainty and potentially leading to insufficient localization accuracy, this paper proposes a keyframe selection optimization method based on information entropy theory. First, by using a quadrotor UAV as the hardware platform, the key processes of front-end feature tracking and back-end sliding-window optimization in VINS-Fusion

本文通讯作者: 张熙, 中国民用航空飞行学院高级实验师, 硕士生导师, 主要从事电子技术应用与机载设备智能检测方面的研究。

E-mail: 1815065038@qq.com

文章引用: 张熙, 谭年清. 一种基于信息熵的关键帧优化方法[J]. 计算机科学辑要, 2025, 1(1): 44-53. ISSN Print: 3079-5362; ISSN Online: 3079-5370.

DOI: 10.63313/CS.8004

are systematically analyzed. Then, during the keyframe insertion stage, an information entropy metric is introduced to accurately characterize the uncertainties of map points and pose estimates. By setting an entropy threshold, the timing of keyframe insertion is adaptively optimized, effectively reducing the global cumulative error. Finally, the proposed method is validated on the open-source EuRoC dataset and comprehensively compared with the original VINS-Fusion algorithm. Experimental results demonstrate that incorporating information entropy into the keyframe selection strategy significantly reduces redundant keyframes, decreases map-point uncertainty to some extent, enhances global trajectory estimation accuracy, and has only a minimal impact on the real-time performance of the algorithm.

Keywords

UAV indoor localization; Vision-inertial fusion; Visual-Inertial Odometry; Information entropy

Copyright © 2025 by author(s) and Erytis Publishing Limited.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



1. 引言

视觉惯性导航系统 (Visual-Inertial Navigation System, VINS) 通过紧耦合融合相机与惯性测量单元 (Inertial Measurement Unit, IMU) 的观测数据^[1], 能够在 GPS 拒止环境或室内外复杂场景中实现精确且鲁棒的自主定位与导航, 已被广泛应用于无人机^[2]、机器人^[3]以及增强现实等领域。其中, 基于视觉惯性融合的同步定位与建图 (Visual-Inertial Simultaneous Localization and Mapping, VI-SLAM) 技术通过视觉特征提取与 IMU 数据的紧密融合, 实时完成环境建图与自定位任务, 逐渐成为当前无人机自主导航与轨迹估计的重要研究方向。

近年来, 视觉惯性融合 SLAM 算法发展迅速。Qin 等提出的 VINS-Mono 算法首次实现了单目相机与 IMU 的紧耦合融合^[4], 在保证实时性的基础上显著提升了定位精度。随后, 其团队进一步提出了 VINS-Fusion 算法^[5], 扩展了多传感器融合框架, 能够同时融合单目、双目以及深度相机等多种视觉传感器数据, 进一步提升了系统的鲁棒性与适应性。VINS-Fusion 算法在关键帧选择策略上^[6], 主要通过启发式方法进行判定。当当前帧与最近关键帧之间的平均视差超过某个预设阈值, 或者当特征跟踪质量显著下降时, 则插入新的关键帧进行后端优化。然而, 现有 VINS 系列算法的关键帧选择策略仍然采用的这种基于视觉特征点视差阈值及特征跟踪质量的启发式判定方法虽然在常规场景下能够满足定位与导航需求, 但本质上未直接对地图点或位姿的不确定性进行量化评估, 难以有效适应高动态或纹理稀疏的复杂环境, 可能导致系统关键帧插入的冗余度增加或重要关键帧漏选, 从而限制了全局定位性能的进一步提升。

信息熵^[7]作为信息论中度量随机变量不确定性的重要指标 (Shannon, 1948), 近年来在机器人导航与 SLAM 领域的研究中得到了广泛关注^[8]。Strasdat 等首次提出了基于信息熵的关键帧选择策略^[9], 通过量化地图点信息量减少冗余关键帧插入, 显著提升了单目 SLAM 系统的性能。Das 与 Waslander 将信息熵关键帧策略推广到多摄像头视觉 SLAM 系统 (MCPTAM)^[10], 证明了其在复杂环境下能有效提高定位精度和减少关键帧冗余。

针对上述问题与挑战, 本文提出一种基于信息熵理论的关键帧选择优化方法, 创新性地

入 VINS-Fusion 算法框架中, 利用熵值定量表征地图点和相机位姿的不确定性, 动态地选择并插入最能减少系统全局不确定性的关键帧。本文以四旋翼无人机为硬件平台, 基于开源 EuRoC 数据集进行仿真实验验证, 对比分析了所提出方法相较于原始 VINS-Fusion 算法在定位精度方面的优化效果。实验结果表明, 相较于原始 VINS-Fusion 算法, 引入信息熵的关键帧选择方法能够有效提高系统在复杂场景下的全局轨迹估计精度与鲁棒性。

2. 算法实现框架

VINS-Fusion 算法是由 Qin 等在 VINS-Mono 基础上进一步发展提出的一种通用视觉惯性导航融合框架。该算法通过紧耦合方式^[11]融合相机和惯性测量单元 (IMU) 数据, 实现实时、高精度的定位和建图功能。VINS-Fusion 算法框架主要包括前端视觉特征提取与跟踪、后端滑动窗口非线性优化和回环检测三个主要环节, 其算法框架的流程图如图 1 所示。

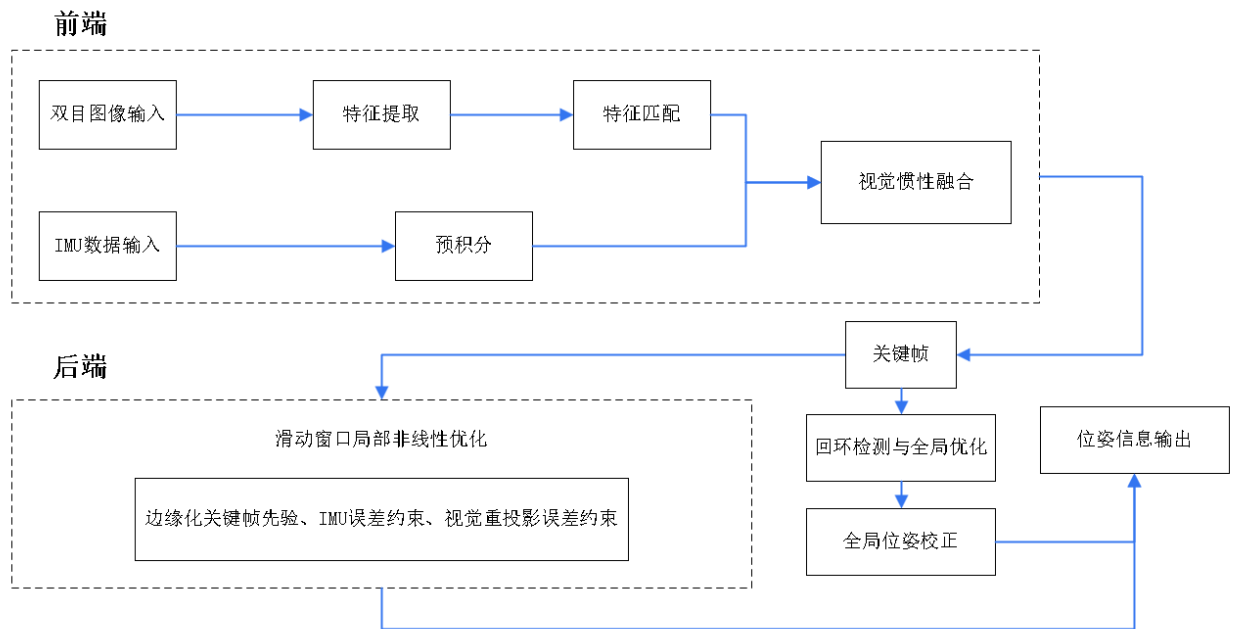


Figure 1. Flowchart of the VINS-Fusion algorithm

图 1. VINS-Fusion 算法流程图

2.1. 前端视觉特征提取与跟踪

在 VINS-Fusion 算法中, 前端模块的核心目标是从连续输入的图像序列中获取并持续跟踪高质量特征点, 为后端优化提供精确而稳定的几何约束^[12]。具体而言, 系统常借助角点检测在多种光照与视角变化场景下提取稳定且可追踪的特征^[13]。随后, 为在帧间建立可靠的特征对应关系, 通常结合多尺度金字塔光流 (Kanade-Lucas-Tomasi, KLT) 算法对这些特征点进行逐帧跟踪。若特征点在相邻帧间能够被成功关联, 则后端可获得高精度的视觉观测约束, 从而在后续的非线性优化中精确估计相机姿态与环境地图。此外, 为尽量减少外点或误匹配带来的负面影响, 系统会根据特征点的跟踪时长、视差变化等指标动态监测跟踪质量, 一旦发现特征出现明显漂移、丢失或潜在外点, 便及时进行剔除或更新, 以确保输

入数据的纯净度与有效性。

2.2. 后端滑动窗口非线性优化

完成特征提取与匹配后, VINS-Fusion 算法的后端通过滑动窗口方法, 将视觉观测与 IMU 测量紧耦合为一个统一的非线性最小二乘问题进行迭代优化^[14]。具体而言, 系统同时考虑了重投影误差与 IMU 预积分误差这两类约束, 从而在窗口中联合估计相机姿态、速度、IMU 偏置以及三维环境中的关键特征点坐标。滑动窗口技术既能保留若干关键帧以保证地图构建的连续性与一致性, 又能通过剔除窗口内时序最久远或贡献较低的帧, 避免随着时间累积而导致的优化规模持续膨胀。在此过程中, 每当插入新的关键帧, 系统便及时截断部分旧观测, 既能适时捕捉新的几何信息、减缓漂移积累, 也能在确保实时性与计算效率的前提下兼顾全局定位精度。然而, VINS-Fusion 目前主要依赖视差阈值与跟踪特征数等启发式条件来决定何时插入关键帧, 从而构建滑动窗口中的观测数据。这种基于局部几何指标的判定方式, 虽然实现简便且能满足一般场景的实时需求, 却缺乏对系统全局不确定性的定量评估, 在场景复杂或环境动态较大的情况下, 过多或过少的关键帧插入都可能造成地图估计偏差, 进而导致全局精度下滑。因此, 如何在滑动窗口框架内结合全局不确定性度量, 进一步优化关键帧的插入策略, 便成为提升 VINS-Fusion 算法整体性能的重要研究方向。

2.3. 回环检测

在 VINS-Fusion 框架中, 回环检测^[15]主要通过将视觉词袋 (Visual Bag-of-Words, BoW)^[16]检索与特征级配准相结合来识别并验证可能存在闭环关系的关键帧, 从而在系统后端构建全局约束。具体而言, 当滑动窗口内加入新的关键帧后, 回环检测模块会提取其图像特征并转化为 BoW 向量, 再与历史关键帧保存的 BoW 向量进行比对, 以迅速筛选出潜在的闭环候选。若匹配结果表明当前关键帧与某一历史关键帧之间存在足够多且稳定的特征对应, 则判定回环成立, 并对这两帧进行相对位姿估计。

在此基础上, 系统将回环约束纳入后端的全局位姿优化。通过该全局优化环节, 一方面能够确保所有关键帧位姿在宏观层面上相互一致, 进而维持地图结构的全局准确性; 另一方面, 则可以有效削减因飞行时间或运动范围较大而不断累积的漂移误差, 从而在平滑局部轨迹的同时进一步强化整体定位精度。得益于回环检测和全局优化的协同作用, VINS-Fusion 在长时或大尺度应用场景中亦能维持对相机位姿与环境地图的高精度估计, 为无人机自主导航、机器人路径规划等任务提供更加稳健的运行保障。

3. 改进的关键帧选择设计

3.1. 信息熵理论基础

除美国数学家 Shannon (1948) 在研究通信系统的信息传输效率问题时首次提出信息熵的概念^[17], 用以度量一个随机变量所包含“信息量”的大小。对于离散随机变量 X , 若其可能取值的概率为 $p(x_i)$, 则信息熵定义为:

$$H(X) = -\sum_i p(x_i) \log p(x_i) \quad (1)$$

其中, $\sum p(x_i) = 1$, 对数底通常取 2 或自然常数 e 。该值反映了随机变量的不确定性, 即随机变量越混乱或不可预测, 其熵值越大。对于连续随机变量, 则可采用微分熵 (Differential Entropy) 进行类比, 或者在具体应用中将其近似离散化处理。在机器人导航与 SLAM 等实际系统中, 很多观测或状态呈现高斯分布特性时, 可使用多维高斯分布熵的封闭形式进行快速计算。

在 SLAM 问题中, 无论是基于视觉、激光雷达, 还是视觉惯性融合, 都需要对机器人自身位姿及环

境地图进行估计。然而, 测量噪声、数据不完整等因素会使估计存在一定程度的不确定性。信息熵提供了一种信息论层面的度量手段: 通过考察观测或估计分布的混乱度, 可以量化当前系统对于环境地图和机器人状态的不确定程度。当熵值过高时, 意味着系统尚有大量未知或难以确定的信息; 当熵值降低时, 则表明系统对场景的把握和对自身状态的估计更精确。

在许多基于卡尔曼滤波或非线性优化的 SLAM 框架中, 都能输出状态的协方差矩阵。若设状态分布为多维高斯, 则可用如下公式估算其熵:

$$H(\mathbf{x}) = \frac{1}{2} \ln((2\pi e)^n |\Sigma|) \quad (2)$$

其中, Σ 为为状态协方差矩阵, $|\Sigma|$ 表示其行列式, n 为状态分量的维度, 当协方差越大时, 行列式越大, 对应的熵值也随之增大, 意味着系统对该状态掌握程度更低、不确定性更高。该公式为 SLAM 在使用高斯近似时提供了直接计算不确定性的手段。

在 SLAM 系统中, 信息熵也可用于度量地图特征或传感器观测分布的丰富度。当特征分布均匀、观测充分时, 熵值会相应降低, 表明系统对环境结构已较为明晰; 若场景纹理贫乏或运动视角单一, 熵则可能较高, 提示系统需要增加新的观测或更新估计方法。

相较于传统依赖启发式阈值的策略, 信息熵能够从概率与信息论角度, 对当前估计或地图构建的不确定度进行精确度量。当熵值超出某个阈值时, 就能提示系统需进一步优化观测或补充新约束。同时, 以熵为指标可帮助系统在关键帧插入时机以及冗余关键帧剔除方面做出更具自适应性的决策, 从而在全局角度降低地图或位姿估计的混乱度。

3.2. 基于信息熵的关键帧选择设计与实现

在 SLAM 中, 关键帧的插入决定了系统对新观测的利用程度: 若插入的关键帧数量过少, 系统将难以通过足够的几何约束来降低不确定性; 插入关键帧的数量过多, 则会增加冗余数据与运算负担, 可能引入不必要的误差。基于信息熵的方法可将关键帧插入这一过程定义为在能最大程度降低系统不确定性的时刻或状态引入新的观测, 从而提升整体精度与效率。

在视觉惯性 SLAM 中, 每个地图点通常在后端优化后拥有协方差矩阵, 用于表征系统对该点三维位置的不确定性。当假定该分布接近高斯时, 其熵可由下式表示:

$$H(\mathbf{p}_j) = \frac{1}{2} \ln((2\pi e)^3 \cdot |\Sigma_j|) \quad (3)$$

其中, $\mathbf{p}_j$ 表示三维地图点, Σ_j 表示对应协方差矩阵, 熵值越高, 代表该地图点位置越不确定。当我们在后端插入新的关键帧并进行相关的优化更新时, 其协方差变为 $\bar{\Sigma}_j$, 对应的熵则变为:

$$\bar{H}(\mathbf{p}_j) = \frac{1}{2} \ln((2\pi e)^3 \cdot |\bar{\Sigma}_j|) \quad (4)$$

为了衡量新关键帧对该地图点带来的不确定性下降程度, 定义熵减少量 $\Lambda(\mathbf{p}_j)$ 如下:

$$\Lambda(\mathbf{p}_j) = H(\mathbf{p}_j) - \bar{H}(\mathbf{p}_j) \quad (5)$$

若插入新的关键帧后, 该地图点的协方差得到显著改善, 则其协方差矩阵会减小, $\Lambda(\mathbf{p}_j)$ 数值随之增大。这一减少量不仅直观反映了地图点从不够清晰到更为明确的过程, 也为系统在全局层面衡量关键帧价值提供了单点参考基准。

假设某个候选关键帧 $\mathbf{K}_i$ 能够观测到一组地图点 $\{p_1, p_2, \dots, p_m\}$, 则可将所有这些点带来的熵减少量进行累加, 得到该关键帧对地图整体不确定性降低的总贡献为:

$$\Psi(\mathbf{K}_i) = \sum_{j=1}^m \Lambda(\mathbf{p}_j) \quad (6)$$

一旦关键帧 $\mathbf{K}_i$ 被选入后端优化, 将对每个被观测到的点带来不确定性下降, 进而实现整体地图更清晰、误差分布更紧凑的效果。与传统基于视差或特征数的启发式不同, 这里的指标是严格从概率论与信息论出发, 对关键帧插入进行定量评估。

在实际系统中, 我们通常维护一批候选关键帧 $\Phi = \{K_1, K_2, \dots, K_r\}$, 若目标是一次仅插入一个关键帧, 则可直接比较所有 $\mathbf{K}_i$ 的 $\Psi(\mathbf{K}_i)$, 选取能产生最大熵减少量的关键帧 $\mathbf{K}^*$, 其数学表达为:

$$\mathbf{K}^* = \arg \max_{\mathbf{K}_i \in \Phi} \Psi(\mathbf{K}_i) \quad (7)$$

这一过程确保在每个插入周期内, 系统总能优先选择最有价值的关键帧。实践中, 考虑到实时性与复杂度, 系统可能只对少量具备较高潜在贡献的帧进行评估, 而非所有候选帧逐一比对。

除了对地图点的不确定性进行信息熵分析外, 在实际视觉惯性导航系统中, 还有一项极为关键的状态变量是相机或无人机自身的位姿, 即六维空间中的平移 (x, y, z) 与旋转 $(roll, pitch, yaw)$ 分量。由于位姿估计直接决定了轨迹重构的精度和系统对地图观测的几何一致性, 因此对其不确定性可以进行有效监测。

在信息熵框架下, 位姿的估计结果同样可被近似为一个高斯分布, 其协方差矩阵通常用于度量系统在六维位姿空间中的不确定性。为了确保系统在滑动窗口优化过程中能够持续获得足够的几何约束, 有必要设置一个熵值上限作为位姿不确定性阈值, 通常记为 ε , 当 $H(x) > \varepsilon$ 时, 系统将判定当前观测信息不足以有效约束位姿空间的状态演化, 从而触发关键帧的插入操作。这一策略能够动态适应不同运动状态下的观测变化, 适用于高速飞行或高动态环境下的无人系统定位任务。

此外, 协方差矩阵反映了相机在三个空间轴向和平移方向上的联合不确定性, 因此其行列式不仅受到单一方向误差的影响, 还反映了各分量之间的协方性与耦合性。在实际操作中, 除了考虑整体熵值 $H(x)$ 外, 也可以进一步分解协方差矩阵, 对六维位姿的各主对角元素进行逐维监控。例如, 在激烈加速度变化时, 平移部分协方差可能显著增大, 而旋转部分相对稳定; 此时系统可以根据不同维度设定独立阈值, 更加具体地决定是否需要插入关键帧。

在 VINS-Fusion 原有的框架中, 滑动窗口非线性优化通过最小化视觉重投影误差和 IMU 预积分误差来联合估计相机姿态、速度、IMU 偏差以及部分地图点坐标。改进思路的主要工作是, 在此优化过程结束后, 保留并输出对各状态量的协方差近似, 以便后续计算信息熵。

在非线性最小二乘算法中, 后端可通过信息矩阵与一次逆运算近似获取协方差矩阵。其数学表达式为:

$$\Sigma = \Lambda^{-1} \quad (8)$$

对于特定地图点, 可从中提取其对应的 3×3 块; 对于相机姿态, 则可提取 6×6 块。在高斯假设下, 可使用公式 (3) 分别计算地图点与位姿的熵值。如果某状态熵超过预设阈值或特定候选帧能显著降低该熵值, 则触发关键帧插入操作。插帧后, 系统再次执行后端优化, 以便充分利用新增观测, 从而进一步降低不确定性。

4. 实验与分析

4.1. 四旋翼无人机硬件平台设计

本研究构建了基于自主开发的四旋翼无人机系统的飞行验证平台, 其核心架构采用多传感器融合与异构计算协同设计。感知模块集成工业级 Intel RealSense D435i 深度视觉传感器与高带宽 IMU, 通过时间

同步机制构建多模态传感阵列。无人机的机载计算单元基于 RK3588 芯片构建异构计算架构, 采用并行流水线架构实现特征提取算法与滑动窗口优化的协同计算, 其算力满足实时性约束下的复杂场景处理需求。

系统通信架构采用分层分布式设计, 通过多协议串行总线与并行接口建立确定性通信网络, 实现机载计算机与传感器、飞控模块间的数据交互。飞行控制采用 PX4 开源飞控系统, 结合 QGroundControl 地面站完成参数整定与状态监控。此外, 无人机硬件系统搭建完成后, 研究利用动态捕捉场地提供的毫米级定位精度完成了 PID 控制器参数的迭代优化, 确保无人机飞行的稳定性和可控性。开发完成的无人机平台如图 2 所示。



Figure 2. UAV hardware platform diagram

图 2. 无人机硬件平台图

4.2. 基于 EuRoC 数据集的仿真实验

EuRoC MAV 数据集作为视觉惯性导航领域公认的基准测试集, 其数据采集系统配备全局快门双目相机与工业级 IMU, 在工业厂房、室内办公环境以及特征稀疏等典型场景中, 通过微型飞行器完成了包含悬停、高速旋转、急加速等复杂运动模式的数据采集。该数据集具有三个显著优势: 首先, 场景覆盖低纹理走廊、强光照变化及动态物体干扰等典型挑战性环境; 其次, 传感器数据包含运动模糊、IMU 量测噪声等真实扰动因素; 最后, 其提供的毫米级精度运动轨迹真值, 可支持绝对轨迹误差 (ATE) 与相对位姿误差 (RPE) 等指标的定量评估。

实验通过多维度测试方案验证算法性能, 选取 EuRoC MAV 数据集中的三个代表性序列作为数据来源。其中 MH_01_easy 序列作为工业场景测试用例, 在布满机械臂、金属管道等工业设施的低照度环境中进行数据采集, 场景呈现典型结构化特征与丰富纹理分布, 重点检验系统在弱光条件下的位姿估计稳定性。V1_01_easy 序列作为光照变化测试用例, 其场景涵盖从强逆光到阴影区域的快速切换, 配合低纹

理特征分布的走廊环境，用于量化评估算法在极端光照波动下的位姿跟踪连续性。而针对动态干扰场景设计的 V2_01_easy 序列，其数据采集区域包含可移动桌椅与密集书架构成的半结构化环境，通过引入人员走动造成的动态遮挡和视角突变，系统化验证算法在特征退化场景下的位姿恢复能力。通过构建包含静态结构、光照扰动和动态遮挡的多模态测试场景，实现了对算法定位性能的全面测试。

本研究构建了双算法对比验证框架，通过多维度误差评价体系对 VINS-Fusion 算法与改进后的算法进行性能评估。在评估指标设计层面，采用均方根误差（RMSE）作为核心评价准则，同步构建包含最大偏差（Max）、期望误差（Mean）、中位数偏差（Median）及误差分布离散度（Std）的四维统计量。为确保评估结果的统计一般性，所有测试序列均进行 30 次实验，通过均值计算排除偶然性的影响。

仿真实验的定量评估结果如表 1 所示，改进后的算法（记为 Ous）在多维度误差指标上呈现一定程度的优化，其中，最大偏差（Max）呈现 11.37% 的降幅，期望误差（Mean）下降幅度达 13.36%，中位数偏差（Median）降低 10.32%，轨迹均方根误差（RMSE）取得 11.23% 的优化增益，标准差（Std）也实现 11.63% 的收敛改善。实验结果表明，引入基于熵的关键帧选择机后的算法相比于原始 VINS-Fusion 算法展现出更好的定位性能。

Table 1. Quantitative evaluation results of simulation experiments
表 1. 仿真实验定量评估结果

		起飞位置	终点位置	轴向实际距离	估计位置	轴向误差
基于 VINS-Fusion 算法的实飞实验	X 轴	10.85 mm	1567.06 mm	1556.21 mm	1502 mm	54.21 mm
	Y 轴	15.62 mm	1273.97 mm	1258.35 mm	1203 mm	55.35 mm
	Z 轴	170.02 mm	1205.43 mm	1035.41 mm	1001 mm	34.41 mm
基于改进算法的实飞实验	X 轴	10.63 mm	1545.89 mm	1535.26 mm	1503 mm	32.26 mm
	Y 轴	15.11 mm	1261.70 mm	1246.59 mm	1201 mm	45.59 mm
	Z 轴	170.02 mm	1194.89 mm	1024.87 mm	999 mm	25.87 mm

4.3. 无人机飞行实验

为验证改进后视觉惯导算法的实际性能与适用性，设计并开展了严格的无人机实飞验证实验。实验在事先建立的标准测试环境中进行，该环境配备了高精度的视觉动捕摄像系统，以提供无人机飞行过程中的真实位置坐标信息。

首先，将搭载视觉惯导系统的四旋翼无人机放置于动捕系统视场范围内的指定起飞位置。随后对动捕系统进行初始化检查，确认摄像机覆盖整个测试飞行区域，确保系统功能正常且无视场盲区或遮挡等影响因素。

在实验正式开始前，需要通过地面站终端向无人机系统发送预先规划的飞行路径终点坐标信息。飞行任务下达后，无人机从指定起飞位置自主起飞，分别利用改进前后的视觉惯导系统实时估计自身位置并控制飞行轨迹，直至到达规划的目标终点并自主降落。飞行过程中，动捕系统同步采集无人机实际飞行轨迹与终点位姿的精确坐标信息，以此作为无人机飞行的轨迹真值参考基准。为进一步限制实验过程中的人为因素扰动，整个实验采用指点飞行策略，通过程序控制无人机自初始点起飞，经过各途经点后在目标终点（1.5，1.2，1）（单位：米）处悬停指定时间后降落。

待无人机抵达目标终点并飞行稳定后，分别记录视觉惯导定位系统给出的无人机终点估计位置坐标 (x, y, z) 以及动捕系统测得的无人机终点实际坐标 (x', y', z') ，根据所得坐标，使用欧氏距离公式计算无人机导航终点的定位误差 ρ ：

$$\rho = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2} \quad (9)$$

为确保实验结果的有效性与可靠性, 以上飞行实验重复实施多次, 并统计多次实验获得的定位误差数据, 最后使用数次实验数据的均值作为最终结果。需要注意的是, 无人机会默认机体起飞位置为 (0, 0, 0), 但是动捕系统提供的位置是考虑到动捕球位置以及无人机机体结构几何尺寸的结果, 所以本实验记录了无人机在动捕场地中的实际起飞以及终点的全局位置, 将各轴的轴向差值与算法位姿反馈结果进行对比, 以此获得相比于真值轨迹的无人机实飞位姿估计误差, 实验结果如表 2 所示。根据表格数据和式 (9) 计算得到: 同样的无人机, 搭载传统 VINS-Fusion 算法时的实飞定位误差 ρ_1 为 8.475cm, 搭载改进后的定位算法时的实飞定位误差 ρ_2 为 6.156cm。因此得出结论, 本研究所提出的改进定位算法相较于原始 VINS-Fusion 算法在无人机实际飞行过程中可以实现更好的定位精度, 在实时性方面也可以满足无人机的应用需求。

5. 结论

本研究针对现有 VINS-Fusion 算法关键帧选择策略仅基于启发式规则, 缺乏对系统全局不确定性的有效量化评估, 易导致定位性能下降的问题, 提出了一种基于信息熵理论的关键帧选择的优化方法。该方法通过构建地图点及位姿状态的信息熵数学模型, 以熵值阈值为决策依据动态调整关键帧的插入策略, 从而有效降低系统的全局累计误差。在 EuRoC 数据集上进行的仿真对比实验表明, 改进后的算法在最大偏差、均方根误差等多个指标上均优于原始 VINS-Fusion 算法, 其中 RMSE 平均降幅为 11.23%。此外, 四旋翼无人机的实飞实验进一步验证了本文方法的实际有效性, 相比传统 VINS-Fusion 算法, 改进后的方法将无人机终点定位误差也有一定程度的下降, 提升了实际应用场景下系统的定位精度与鲁棒性。

致 谢

在此, 我谨向我的导师致以最诚挚的谢意。在本研究的整个过程中, 他们以渊博的学识、严谨的学术态度以及建设性的反馈为我提供了无可替代的指导与支持。他们的专业洞见不仅帮助我明确并完善了研究目标和方法, 也在不断交流与打磨之中引领我将论文的学术质量推向更高水平。

同时, 我也要衷心感谢同学们的陪伴与助力。从日常的切磋讨论到关键时刻的鼓励与协作, 他们的热情与启发性建议让我在学术之路上倍感动力, 也为我的研究视野注入了新的活力。正是因为他们的坚定支持与富有建设性的意见, 我才能在整个研究过程中始终保持专注与热情。

最后, 我要将最深切的感谢献给我的父母。父母始终如一的支持、理解与无私的爱, 为我勇敢追求学术理想与个人成长提供了不可或缺的情感基石。正是因为有了他们的坚实后盾与悉心鼓励, 我才能不惧挑战、不断前行, 努力在学术探索的道路上收获新的高度。

参考文献

- [1] 苏东彦. 基于视觉惯性里程计的无人机室内导航系统研究[D]. 南京: 南京信息工程大学, 2023.
- [2] 王子卉. 无人机视觉/惯性紧耦合导航系统关键技术研究[D]. 南京: 东南大学, 2024.
- [3] 王立强. 轮式机器人运动增强的视觉/惯性/GNSS 融合导航算法研究[D]. 武汉: 武汉大学, 2025.
- [4] Qin T, Cao S, Pan J, et al. A General Optimization-based Framework for Global Pose Estimation with Multiple Sensors[A]. arXiv, 2019.
- [5] Qin T, Pan J, Cao S, et al. A General Optimization-based Framework for Local Odometry Estimation with Multiple Sensors[A]. arXiv, 2019.
- [6] 李丹, 李俊祥, 赵文杰, 等. 基于改进的点线融合和关键帧选择的视觉 SLAM 方法[J]. 机器人, 2025, 34(2): 123-129.
- [7] 杨宇瞳, 王海珍. 基于信息熵和注意力机制的 Bi-LSTM 物联网入侵检测方法[J]. 齐齐哈尔大学学报(自然科学

- 版), 2025(2): 28-33, 69.
- [8] 席含. 基于 RGB 信息熵的 VSLAM 的算法研究[D]. 南昌: 南昌大学, 2024.
 - [9] Strasdat H, Montiel J M M, Davison A J. Visual SLAM: Why filter?[J]. *Image and Vision Computing*, 2012, 30(2): 65-77.
 - [10] Das A, Waslander S L. Entropy based keyframe selection for Multi-Camera Visual SLAM[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2015: 3676-3681.
 - [11] 王风华, 徐志城, 赵冷锐. 基于紧耦合 IMU 的辐射场视觉惯性 SLAM 算法[J]. *电子测量技术*, 2025, 44(5): 112-118.
 - [12] 周益鹏. 基于双目视觉的视觉惯性里程计算法改进研究[D]. 苏州: 苏州大学, 2024.
 - [13] 尚硕, 曹建荣, 汪明, 等. 机器视觉中角点检测算法研究[J]. *计算机测量与控制*, 2024(1): 217-225.
 - [14] 刘高. 基于点线特征的视觉惯性里程计研究[D]. 昆明: 昆明理工大学, 2022.
 - [15] 黄良沛, 胡志豪, 杨天龙, 等. 基于优化回环检测的激光雷达 SLAM 算法[J]. *信息技术*, 2024(12): 26-31, 38.
 - [16] 谷学静, 刘秋月, 马冠征, 等. 基于双边滤波和词袋模型的图像匹配算法[J]. *激光杂志*, 2022(7): 70-74.
 - [17] Shannon C E. A mathematical theory of communication[J]. *The Bell System Technical Journal*, 1948, 27(3): 379-423.