

# 视觉惯性导航系统初始化方法综述\*

刘哲<sup>1</sup> 史殿习<sup>2</sup> 杨绍武<sup>1</sup> 李睿豪<sup>2</sup>

(1. 国防科技大学 计算机学院, 湖南 长沙 410073; 2. 军事科学院 国防科技创新研究院, 北京 100071)

**摘要:** 视觉惯性导航系统通过初始化,对尺度信息、重力向量、速度、惯性传感器偏差等一系列状态估计所需参数进行快速求解,以提升系统后续导航定位与环境感知的准确性。根据传感信息耦合方式,视觉惯性导航系统初始化方法可以分为三类:联合初始化、非联合初始化和半联合初始化。基于现有研究工作,从基础理论、发展与分类、现有方法、性能评估四个方面展开,对目前主流的初始化方法进行综述,并总结视觉惯性导航系统初始化领域未来的发展趋势,有利于对视觉惯性导航系统初始化方法形成总体性了解并把握其发展方向。

**关键词:** 视觉惯性导航系统; 初始化; 传感融合; 惯性数据处理

中图分类号: TN18 文献标志码: A 开放科学(资源服务)标识码(OSID):

文章编号: 1001-2486(2023)02-015-12



听语音  
与作者互动  
聊科研

## Review of visual – inertial navigation system initialization method

LIU Zhe<sup>1</sup>, SHI Dianxi<sup>2</sup>, YANG Shaowu<sup>1</sup>, LI Ruihao<sup>2</sup>

(1. College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China;

2. National Innovation Institute of Defense Technology, Academy of Military Sciences, Beijing 100071, China)

**Abstract:** VINS (visual-inertial navigation systems) solve a series of parameters required for state estimation through initialization, such as scale, gravity vector, velocity, inertial measurement unit's bias, etc., to improve the accuracy of navigation positioning and environmental perception of the system. The initialization methods of VINS can be divided into three categories according to the sensing information fusing mode: joint initialization, disjoint initialization and semi-joint initialization. Based on the existing research work, the current mainstream initialization methods of VINS were reviewed from four aspects: basic theory, development and classification, existing methods, performance evaluation, and the future development trends were summarized, which is helpful to have a general understanding of VINS initialization methods and grasp its development direction.

**Keywords:** visual-inertial navigation systems; initialization; sensor fusion; inertial measurement processing

同时定位和构图(simultaneous localization and mapping, SLAM)以及视觉里程计(visual odometry, VO)是机器人导航、计算机视觉等领域的新兴研究热点<sup>[1]</sup>,允许单个机器人仅通过视觉传感器(如单目相机<sup>[2]</sup>、双目相机<sup>[3]</sup>、深度相机<sup>[4]</sup>或事件相机<sup>[5]</sup>)或多机器人通过视觉传感器和信息交互<sup>[6]</sup>进行未知环境感知和自身定位导航,目前已经被广泛应用于虚拟现实(virtual reality, VR)、增强现实<sup>[7]</sup>(augmented reality, AR)、场景重建<sup>[8]</sup>以及自动驾驶<sup>[9]</sup>等领域。

由于视觉传感器与惯性测量单元(inertial measurement unit, IMU)具有较好的互补关系,近年来,加入IMU的视觉惯性SLAM<sup>[10-12]</sup>(visual

inertial SLAM, VI-SLAM)和视觉惯性里程计<sup>[13-15]</sup>(visual inertial odometry, VIO)越来越受到人们的关注,VI-SLAM与VIO也被统称为视觉惯性导航系统<sup>[16]</sup>(visual-inertial navigation system, VINS)。仅使用单目相机和低成本IMU的VINS,可以提供小型、紧凑、廉价、准确、低功耗的定位导航与环境感知解决方案。

VINS执行定位导航任务时,往往需要在启动之初对一些特定参数进行初始估计,这一参数估计的过程被称为VINS初始化。准确的初始化结果可以提升VINS位姿估计、地图构建的精度和求解速度;反之,错误或不准确的初始化结果将影响VINS性能<sup>[17]</sup>。VINS初始化参数一般可以分

\* 收稿日期: 2022-11-01

基金项目: 国家自然科学基金资助项目(91948303, 61903377)

作者简介: 刘哲(1994—),男,山西柳林人,博士研究生, E-mail: liuzhe16@nudt.edu.cn;

史殿习(通信作者),男,山东龙口人,研究员,博士,博士生导师, E-mail: dxshi@nudt.edu.cn

为以下三类:

1) 尺度信息、重力向量与系统初始速度: 单目视觉系统无法准确估计的参数, 可以通过引入 IMU 测量数据对其进行求解, 进而获得世界坐标系中、绝对尺度下的系统位姿与地图构建结果。

2) IMU 偏差: 低成本 IMU 由于加工精度低、易受环境温度影响等特性, 存在一定的零漂和温漂, 这些误差被统称为 IMU Bias。通过融合视觉信息与 IMU 测量数据, 可以对 IMU Bias 进行求解, 以降低其对系统状态估计的影响。

3) 系统外参: 系统外参指相机坐标系与 IMU 坐标系之间的变换矩阵。当无法事先获得时, VINS 需要通过初始化来获取系统外参估计值。

目前, 对 VINS 的研究被认为既有理论价值又具实用意义<sup>[18]</sup>, 并有不少综述文献可供参考<sup>[16, 19]</sup>。但作为 VINS 的重要组成部分, VINS 初始化却鲜有文献单独对其进行整理和总结。因此, 本文将对现有 VINS 初始化方法进行综述, 从 VINS 初始化所依赖的基础工作、VINS 初始化发展与分类、VINS 初始化现有方法、VINS 初始化方法性能评估四个方面进行介绍, 最后对 VINS 初始化未来的发展趋势进行讨论和展望。图 1 对本文重要概念及名称之间的关系进行梳理, 其中部分概念将在后续章节进行介绍。

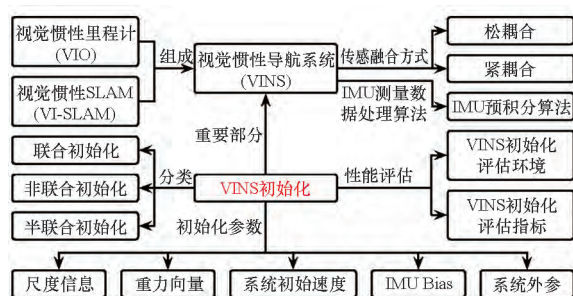


图 1 重要概念及名称关系

Fig. 1 Relationship between important concepts and definitions

## 1 VINS 初始化基础工作

初始化是 VINS 的重要组成部分, 其核心与 VINS 状态估计相似, 均是对一系列系统状态参数的估计。因此, VINS 状态估计中 IMU 测量值与视觉信息之间的融合方式在 VINS 初始化中同样适用。随着 IMU 预积分算法的提出与完善<sup>[20-23]</sup>, 2016 年以来, 新提出的初始化方法基本上都基于预积分推导而来。本节主要从如下两个方面对 VINS 初始化基础工作进行介绍: 传感融合方式和 IMU 预积分算法。

### 1.1 传感融合方式

VINS 需要对不同传感器数据进行融合, 并基于传感融合结果对系统状态进行估计, 以得到系统的定位信息、运动信息以及周边环境的地图信息。目前 VINS 有两种传感融合方式: 松耦合和紧耦合<sup>[24]</sup>。

图 2(a) 为松耦合示意图。松耦合采用独立的惯性定位模块和视觉定位模块, 两个模块的更新频率通常是不一致的, 模块之间存在一定的信息交互, 通过融合两个模块的结果获得系统最终状态估计结果。松耦合算法较为简单并易于实现。

图 2(b) 为紧耦合示意图。紧耦合使用两种传感器的原始数据共同估计系统状态, 不同传感器的噪声会相互影响, 算法实现上相对复杂。但相比较于松耦合, 紧耦合充分利用并结合了异构传感器数据, 可以得到更精确的状态估计结果。

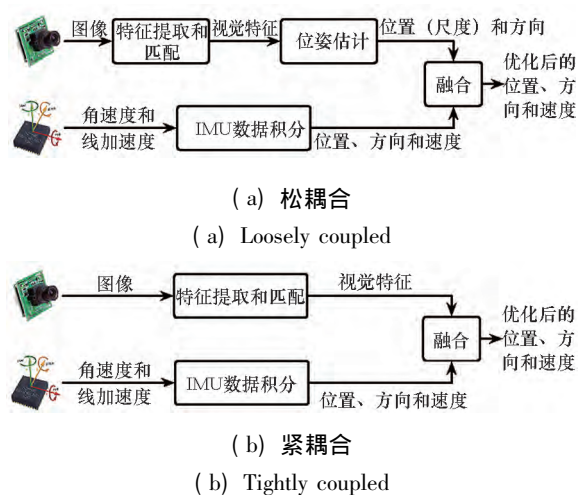


图 2 VINS 松耦合与紧耦合示意图

Fig. 2 Loosely and tightly coupled paradigms for VINS

VINS 初始化所使用的传感数据融合方法, 基本都是对松耦合和紧耦合的延伸与拓展。

### 1.2 IMU 预积分算法

IMU 预积分算法最先在文献 [20] 中被提出, 后经文献 [21-22] 完善, 目前有流形上<sup>[22]</sup>和基于运动方程<sup>[23]</sup>两种表征方式。

在 VINS 中, IMU 的工作频率远高于视觉传感器的工作频率。传统 VINS 进行状态估计时, 往往以图像帧时刻作为时间节点, 每当有新图像输入时, 就会对系统状态向量进行更新。状态向量的频繁更新导致了对 IMU 测量数据的重复积分, 而高频、大量的 IMU 测量数据又使这一过程的时间开销骤增。为解决此问题, 研究者提出了 IMU 预积分算法<sup>[20-23]</sup>。

IMU 预积分算法的核心思想是将一段时间内 IMU 测量数据中与状态向量更新无关的数据进行积分处理,作为这段时间内的 IMU 预积分项,避免这部分数据在后续过程中的重复积分,最终达到简化 VINS 状态估计的目的。图 3 为预积分算法示意图。IMU 预积分算法的提出,不仅提升了 VINS 状态估计性能,同时也大大提升了 VINS 初始化算法的效率和实时性。

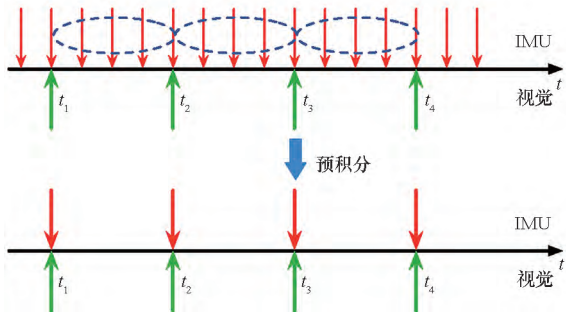


图 3 IMU 预积分算法示意图

Fig. 3 Illustration of IMU pre-integration algorithm

## 2 VINS 初始化发展与分类

VINS 初始化的发展是伴随 VINS 发展而逐渐兴起的。早期 VINS 由于受到硬件的限制,多使用基于滤波的方法进行状态估计,这种方法虽然计算效率高,但精度较低。随着计算机硬件的发展,基于优化的状态估计方法逐渐在 VINS 领域展现优势。这种方法虽然计算开销大,但是精度高。与之一起蓬勃发展的便是 VINS 初始化算法。准确的初始化结果,可以为 VINS 提供一个良好的初始状态,大大提升 VINS 启动之初状态优化求解的收敛速度,同时有效提升状态估计的

准确性。此外,IMU 预积分算法<sup>[20-23]</sup>的提出对于 VINS 初始化的发展也起到了至关重要的作用。

在文献[25]中,VINS 初始化按照传感融合方式被分为两大类——联合初始化和非联合初始化,可以分别看作是基于紧耦合和基于松耦合的初始化。本文在其分类之上,又进一步总结细分出一类初始化方法,称之为半联合初始化。三类不同初始化方法定义如下:

1) 联合初始化:不单独处理视觉信息,一步求解初始化参数的初始化方法。

2) 非联合初始化:单独处理视觉信息,之后分多步依次求解初始化参数的初始化方法。

3) 半联合初始化:单独处理视觉信息,之后一步求解初始化参数的初始化方法。

对于上述三类初始化方法,联合初始化由于需要一步求解所有初始化参数,往往存在计算开销大、优化收敛难的问题;非联合初始化虽然通过分步求解、参数解耦的方式降低了整体计算开销,但分步求解的过程中往往忽略了不同初始化参数之间的相互影响,使得初始化精度有所损失;而半联合初始化可以看作是联合初始化和非联合初始化之间的一种平衡,它对初始化求解过程中耦合程度最低的纯视觉部分单独处理,不仅简化了后续的求解过程,而且保证了初始化的精度。目前,三类初始化方法都在继续发展完善,并相互借鉴和补充。表 1 对 VINS 初始化典型工作所属初始化分类、是否基于预积分算法和主要贡献进行了概况和总结。一般来说,目前的初始化工作重点关注的是初始化方法求解精度和在具体 VINS 中的求解耗时。

表 1 VINS 初始化典型工作

Tab. 1 Typical work of VINS initialization

方法	提出年份	来源	所属初始化分类	基于预积分算法	主要贡献
闭式求解方法 <sup>[26]</sup>	2014	IJCV	联合初始化	否	首次提出联合初始化方法
M-K 方法 <sup>[27]</sup>	2017	R-AL	联合初始化	否	分析 IMU Bias 对初始化的影响,基于非线性优化求解初始化参数
文献[28]	2016	T-ASE	半联合初始化	是	首次提出半联合初始化方法,首次引入预积分算法,支持外参估计
VI-ORB 方法 <sup>[11]</sup>	2017	R-AL	非联合初始化	是	首次提出非联合初始化方法
VINS-Mono 方法 <sup>[29]</sup>	2017	IROS	非联合初始化	是	简化重力向量细化过程,代码开源
改进 VI-ORB 方法 <sup>[30]</sup>	2018	ICRA	非联合初始化	是	支持外参估计,最全面的初始化方法之一
改进 M-K 方法 <sup>[31]</sup>	2019	ICRA	联合初始化	是	将预积分算法引入联合初始化,初始化性能评估方法
ORB-SLAM-3 方法 <sup>[25]</sup>	2020	ICRA	半联合初始化	是	基于最大后验概率,代码开源

### 3 联合初始化

#### 3.1 闭式求解方法

联合初始化方法最早由 Martinelli 于 2014 年在文献 [26] 中提出。在文献 [32] 的基础上, Martinelli 提出了一种针对 VINS 初始化的闭式求解方法,该方法在不考虑系统外参和陀螺仪 Bias 的情况下,基于系统运动方程和特征点在不同图像帧之间的空间约束,构造单个特征点在两帧图像中的初始化方程,再通过联立所有特征点初始化方程得到初始化求解线性系统,最终求解获得重力向量、系统速度、特征点深度(尺度信息)、加速度计 Bias 这四项初始化参数。

#### 3.2 M-K 方法与改进 M-K 方法

2017 年, Martinelli 和 Kaiser 对闭式求解方法进行改进<sup>[27]</sup>,通过分析不同 IMU Bias 对初始化造成的影响,去除了闭式解中对加速度计 Bias 的估计,增加了对陀螺仪 Bias 的估计,并采用非线性优化的方式对初始化参数进行求解。这种改进的闭式求解方法被称为 M-K 方法。

2019 年, Campos 等在文献 [33] 的基础上,对 M-K 方法进行了全面改进,提出了融合 IMU 预积分算法的改进 M-K 方法<sup>[31]</sup>。Campos 等认为 M-K 方法存在三个问题:①关于路标点可观测性的假设不合理;②在精度和时间开销上难以达成平衡;③无法判断初始化结果是否可用。

针对以上三个问题,改进 M-K 方法首先拓展了路标点可观测性假设的适用范围,使得更多路标点可以参与初始化过程;其次引入预积分算法,并加入重力向量细化算法,在提升初始化求解精度的同时,降低了求解耗时;最后增加了两项针对初始化结果的测试——可观测性测试和一致性测试,以提升初始化求解结果的精度和可靠性。

#### 3.3 M-K 方法的后续改进

2020 年, Martinelli 等将 M-K 方法推广到多机器人协同系统中,提出一种针对多机器人的改进 M-K 方法<sup>[34]</sup>,并对初始化以及状态估计中相关状态的可观测性进行研究。

2021 年, Evangelidis 等对 M-K 方法进行改进<sup>[35]</sup>。该方法将同一个路标点同时与多帧图像中的特征观测信息进行匹配,构造初始化优化函数进行求解。该求解函数结构简单,并可通过消元法有效降低求解系统维度,最终可获得比 M-K 方法快 50% 的初始化求解速度。同年, Micusik 等将改进 M-K 方法推广到基于多卷帘快门相机

的 VINS 中,并在初始化求解过程中提出一套完整的重整化求解方法<sup>[36]</sup>。

2022 年, Liu 等在改进 M-K 方法的基础上,首次将特征线引入 VINS 初始化<sup>[37]</sup>;基于普吕克坐标线表示方法<sup>[38]</sup>推导出特征线初始化方程,并在改进 M-K 方法的基础上构造结合点线特征的初始化非线性优化损失函数,通过求解获得初始化相关参数以及点线特征参数。

### 4 非联合初始化

#### 4.1 VI-ORB 方法

2017 年,非联合初始化方法首次由 Mur-Artal 等在文献 [11] 中提出。Mur-Artal 等在 ORB-SLAM<sup>[39]</sup>、ORB-SLAM-2<sup>[40]</sup> 的基础上加入惯性测量数据,提出首个基于关键帧的视觉惯性 SLAM 系统——VI-ORB-SLAM,并为其设计了一种全新的初始化方法(下文称为 VI-ORB 方法)。VI-ORB 方法引入预积分算法,并分步求解陀螺仪 Bias、尺度信息、重力向量、加速度计 Bias 和系统速度这五项初始化参数。

VI-ORB 方法可分为四步:①基于纯视觉信息对陀螺仪 Bias 进行估计;②在不考虑加速度计 Bias 的情况下,基于预积分算法和纯视觉信息构建融合尺度信息、重力向量的运动约束方程组,求解方程组对尺度信息和重力向量进行初步估计;③在第二步运动约束方程组中加入加速度计 Bias 和重力细化项重新求解,获得加速度计 Bias 估计值以及细化后的尺度信息和重力向量;④将前三步求得的初始化参数代入系统运动方程中,求解获得不同时刻的系统速度向量。

VI-ORB 方法相对于当时主流的 M-K 方法,具有耗时低、鲁棒性强的优势。

#### 4.2 VINS-Mono 方法

同样在 2017 年, Qin 基于其团队已有研究成果<sup>[41]</sup>在文献 [29] 中提出了一种非联合初始化方法(下文称为 VINS-Mono 方法),该方法后续被使用在 VIO 系统 VINS-Mono<sup>[15]</sup> 中,为 VINS-Mono 提供了良好的初始化估计结果。

VINS-Mono 方法与 VI-ORB 方法思路基本一致,主要区别在于:①不单独对系统速度求解,而是与尺度信息、重力向量一同求解;②通过仿真实验发现,在大部分情况下,加速度计 Bias 对于初始化精度影响较小,因此忽略了对于加速度计 Bias 的求解;③简化重力向量细化过程,将重力细化项定义为正切平面上的正交分量,增加求解

约束。

#### 4.3 改进 VI-ORB 方法

2018 年, Huang 等对 VI-ORB 方法进行改进, 提出了一种更全面的非联合初始化方法<sup>[30]</sup> (下文称为改进 VI-ORB 方法)。该方法相比于 VI-ORB 方法, 增加了对系统外参的估计。

与 VI-ORB 方法相同, 改进 VI-ORB 方法也分为四步, 其与 VI-ORB 方法的区别在于: ①参照文献 [28] 在 VI-ORB 方法第一步中加入对系统外参旋转分量的估计; ②在第二步和第三步中, 外参平移分量从已知项变为运动约束方程组中的求解项, 相应的求解方式也从线性方程组求解变为非线性优化求解。改进 VI-ORB 方法是目前求解项最全面的初始化方法之一。

#### 4.4 其他非联合初始化方法

由于非联合初始化耗时低、鲁棒性强, 再加上 VINS-Mono、ORB-SLAM 和 ORB-SLAM-2 代码均开源, 近年来, 很多 VINS 初始化工作均是基于 VI-ORB 方法和 VINS-Mono 方法改进而来。

2018 年, Mu 等对 VI-ORB 方法和 VINS-Mono 方法进行整合<sup>[42]</sup>, 在 VI-ORB 方法的基础上, 利用 VINS-Mono 方法对重力向量进行细化。2019 年, Li 等在文献 [43] 中对 VI-ORB 方法进行改进, 利用人造物体边缘多与地面垂直/平行的特点, 通过图像中与地面垂直的线段对重力向量进行细化, 一方面提升重力方向的估计精度, 另一方面加速初始化求解过程, 从而有效提升初始化性能。同年, Jung 等将 VI-ORB 方法推广到双目 VINS 中<sup>[44]</sup>, 通过引入双目相机约束直接获得尺度信息, 降低初始化求解复杂度。2020 年, Gong 等在半直接法视觉惯性里程计框架<sup>[45]</sup>中, 提出一种基于尺度收敛的初始化评估方法, 以提升初始化算法可靠性。2021 年, Cheng 等在文献 [46] 和文献 [47] 中对 VI-ORB 方法进行改进, 提出一种基于信赖域的迭代优化求解方法, 提升初始化求解性能。同年, Huang 等在改进 VI-ORB 方法的基础上, 提出了一种考虑时间-空间约束的 VINS 初始化方法<sup>[48]</sup>, 可以对 IMU 和相机之间的传感器时延进行估计。

此外, 非联合初始化还被推广到了更复杂的传感融合系统中。2020 年, Wang 等提出一种视觉-惯性-磁力融合的定位导航系统<sup>[49]</sup>, 并将非联合初始化方法应用到该系统当中, 在求解初始化参数的同时, 还实现了视觉-惯性-磁力对齐。2021 年, Jin 等在 VI-ORB 方法的基础上加入全球

卫星导航系统 (global navigation satellite system, GNSS) 测量数据<sup>[50]</sup>, 提出并实现了一种面向视觉-惯性-GNSS 融合场景的初始化方法。2022 年, Xiong 等针对自动驾驶场景, 在 VINS-Mono 方法的基础上加入 GNSS 测量数据<sup>[51]</sup>, 同样设计并实现了一种面向视觉-惯性-GNSS 融合场景的初始化方法。

### 5 半联合初始化

#### 5.1 早期半联合初始化方法

2016 年, Yang 等首次在文献 [28] 中提出了半联合初始化方法。该方法首次将 IMU 预积分算法引入 VINS 初始化, 可以有效对尺度信息、重力向量、系统速度和系统外参进行估计, 但是该方法忽略了对 IMU Bias 的估计。

该半联合初始化方法可分为两步: ①基于 SfM 方法<sup>[52]</sup>和陀螺仪测量数据, 分别获得相机坐标系与 IMU 坐标系中的系统旋转状态, 并在此基础上求解获得系统外参旋转分量; ②基于预积分公式和特征点重投影误差, 构造初始化线性优化损失函数, 求解获得系统速度、重力向量、特征点深度 (尺度信息) 以及系统外参平移分量。

该半联合初始化方法由于没有考虑 IMU Bias, 因此可用性并不高, 但是其提出的初始化线性优化方法以及后续状态估计非线性优化方法, 为日后 VINS-Mono 等一系列 VINS 的提出提供了良好的理论基础。

#### 5.2 ORB-SLAM-3 方法

2020 年, Campos 等提出了一种基于最大后验概率估计的半联合初始化方法<sup>[25]</sup>, 该方法被用于 ORB-SLAM-3<sup>[53]</sup>中, 作为视觉惯性部分的初始化方法 (下文称为 ORB-SLAM-3 方法)。ORB-SLAM-3 方法可分为三步: ①仅基于视觉信息的最大后验概率估计, 即通过 ORB-SLAM 算法获得不考虑尺度的系统状态; ②仅基于 IMU 测量数据的最大后验概率估计, 即通过最大化后验概率分布构造仅考虑 IMU 测量数据的非线性优化损失函数, 通过优化求解获得尺度信息、重力向量、系统速度、加速度计 Bias、陀螺仪 Bias 的初始化估计值; ③基于视觉-惯性数据的联合最大后验估计, 即以第二步结果为初始值, 分两次进行视觉惯性联合优化, 以获得更加精确的初始化估计值 (该步骤与 VINS 状态估计并无本质区别)。

ORB-SLAM-3 方法在实时性和准确性之间取得了良好的平衡, 是目前效果最好的 VINS 初始



化方法之一。

### 5.3 其他半联合初始化方法

2021 年,文献[54]基于 ORB-SLAM-3 方法,通过引入拉格朗日乘子法提出一种针对 VINS 初始化的解析解方法,该方法在保证初始化求解实时性的同时提升了初始化求解精度。同年,Yathirajam 等提出一种基于闭式求解方法的半联合初始化方法<sup>[55]</sup>,该方法可以看作是对闭式求解方法的全面改进。此外,吴涛等也提出了类似文献[55]中的改进闭式求解方法<sup>[56]</sup>。刘刚等提出一种基于 VINS-Mono 方法的半联合 VINS 初始化方法<sup>[57]</sup>,并提出一种基于费歇尔信息矩阵的初始化评估方法。

## 6 VINS 初始化方法性能评估

随着 VINS 理论快速发展和应用日益广泛,对于 VINS 初始化的研究也日渐增多。如何评估初始化方法的性能,对于 VINS 初始化研究至关重要。本节主要从评估环境和评估指标两个部分对 VINS 初始化方法性能评估进行介绍,并在最后分析比较了目前主流 VINS 初始化方法的性能。

### 6.1 VINS 初始化评估环境

VINS 初始化方法性能评估环境可以分为两类:仿真环境和真实环境。

一般来说,VINS 初始化方法会在仿真环境中对特定条件下的初始化进行模拟仿真验证和评估。例如在文献[29]中,作者在仿真环境下,通过设置不同的加速度计 Bias 进行初始化测试。测试结果表明,加速度 Bias 对初始化结果影响并不大。因此文献[29]所提 VINS-Mono 方法放弃了对加速度计 Bias 的估计。同样,文献[27]也在仿真环境中,探究了不同初始化参数对初始化结果的影响。

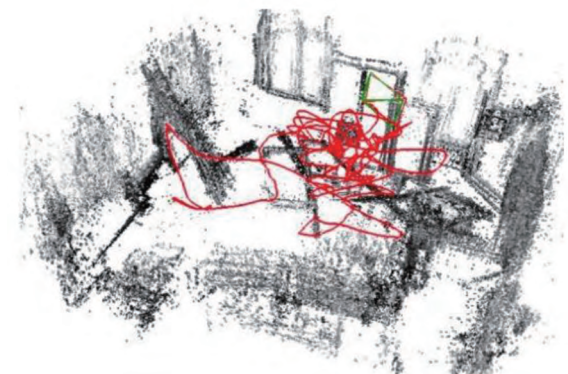
VINS 是一类面向真实环境的实际系统,因此绝大部分初始化工作都会基于真实环境中的数据,进行实验验证和评估。目前,采用最广泛的初始化数据来自公共数据集 EuRoC MAV Dataset<sup>[58]</sup>(以下简称 EuRoC 数据集)。EuRoC 数据集属于室内视觉惯性数据集,提供由无人机采集的双目相机和 IMU 数据,同时提供由 VICON 运动捕捉系统采集的无人机运动轨迹和六自由度位姿真值。EuRoC 数据集包含两个场景共计 11 条地图序列,其中,工厂场景(Machine Hall)包含 5 条地图序列,房间场景(Vicon Room)包含 6 条地图序

列。图 4(a)展示了 EuRoC 数据集中不同场景的图像序列,图 4(b)展示了 VINS 工作 VI-DSO<sup>[59]</sup>在 EuRoC 数据集房间场景中的导航结果。由于 EuRoC 数据集可以提供 VINS 初始化所需的运动激励和传感信息,大多数初始化工作都选择 EuRoC 数据集进行性能评估,包括改进 M-K 方法<sup>[31]</sup>、VI-ORB 方法<sup>[11]</sup>、VINS-Mono 方法<sup>[29]</sup>、改进 VI-ORB 方法<sup>[30]</sup>、ORB-SLAM-3 方法<sup>[25]</sup>等。此外,部分 VINS 初始化方法还会采用实物实验的方式,对算法的有效性进行验证,例如 M-K 方法<sup>[26]</sup>、VINS-Mono 方法<sup>[29]</sup>、文献[28]中的半联合初始化方法等。



(a) 不同场景图像序列

(a) Image sequences of different scenes



(b) VI-DSO<sup>[59]</sup> 导航结果(轨迹与三维重建)

(b) Navigation results of VI-DSO<sup>[59]</sup>

(trajectory and 3D reconstruction)

图 4 EuRoC 数据集图像序列展示与 VI-DSO<sup>[59]</sup>

在其房间场景中导航结果示意图

Fig.4 Image sequences of EuRoC Dataset and navigation results of VI-DSO<sup>[59]</sup> in Vicon Room scene

除了 EuRoC 数据集,文献[19]还介绍了目前主流的、基于不同平台载体的 VINS 数据集,理论上来说,这些数据集均可应用于不同场景下 VINS 初始化方法的性能评估。

### 6.2 VINS 初始化评估指标

VINS 初始化方法性能评估指标可以分为两类:实时性评估指标和准确性评估指标。

实时性评估指标主要评估 VINS 初始化耗

时,一般会基于初始化过程中的两类时间开销进行评估—— $t_{\text{Init}}$ 和 $t_{\text{Tol}}$ 。其中 $t_{\text{Init}}$ 是初始化数据所占时长,即初始化求解框架内最后一帧数据时间戳与第一帧数据时间戳的差值; $t_{\text{Tol}}$ 是初始化总耗时,即初始化完成时刻与初始化求解框架内第一帧数据时间戳的差值。因为 VINS 初始化往往需要额外时间对数据进行处理和求解,因此 $t_{\text{Tol}} \geq t_{\text{Init}}$ 。对于 VINS 来说,初始化方法耗时越短越好。

准确性评估指标可以分为两类:初始化参数准确性评估指标和状态估计准确性评估指标。初始化参数准确性评估主要对比初始化参数求解结果与参数真值之间的差距,差距越小,说明初始化方法准确性越高。一般主要对比 IMU Bias、尺度信息求解结果和真值之间的差距。状态估计准确性评估是指将初始化方法耦合到现有 VINS 中,通过对比不同初始化方法(或有/无初始化)下 VINS 后续状态估计的精度来对初始化方法准确性进行评估。后续状态估计精度越高,说明选用的初始化方法准确性越高。对于 VINS 来说,初始化方法准确性越高越好。

### 6.3 VINS 初始化方法性能对比

文献[25]对目前最受关注且性能优异的联合初始化方法——改进 M-K 方法<sup>[31]</sup>、非联合初始化方法——VINS-Mono 方法<sup>[29]</sup>和半联合初始化方法——ORB-SLAM-3 方法<sup>[25]</sup>的性能进行了对比,主要对比了各方法初始化耗时以及初始化尺度信息估计误差。同时,也加入了不经初始化而直接进行视觉惯性联合优化(即 VINS 状态优化)后系统的尺度误差作为参考项。

该对比实验基于 EuRoC 数据集进行,将 EuRoC 数据集中的每一条地图序列都切分为众多子地图序列以扩大测试对象数量,并在所有子地图序列中测试上述三种初始化方法和无初始化的参考方法,最终得到每种方法在所有地图序列初始化结果的平均值。实验结果如表 2 所示。由于改进 M-K 方法与 VINS-Mono 方法处理图像帧的频率不同(分别为 4 Hz 与 10 Hz),因此表 2 中的实验结果分为两部分:上半部分为改进 M-K 方法与 ORB-SLAM-3 方法性能对比结果,其中 ORB-SLAM-3 方法与无初始化参考方法(即不经初始化而直接进行视觉惯性联合优化,下文同)均以 4 Hz 的频率处理图像帧;下半部分为 VINS-Mono 方法与 ORB-SLAM-3 方法性能对比结果,其中 ORB-SLAM-3 方法与无初始化参考方法均以 10 Hz 的频率处理图像帧。

表 2 改进 M-K 方法、VINS-Mono 方法与 ORB-SLAM-3 方法在 EuRoC 数据集上性能对比结果<sup>[25]</sup>

Tab. 2 Performance comparison results of the improved M-K method, VINS-Mono method and ORB-SLAM-3 method on EuRoC MAV Dataset<sup>[25]</sup>

方法	尺度误差 / %	$t_{\text{Init}} / \text{s}$	$t_{\text{Tol}} / \text{s}$
改进 M-K 方法 <sup>[31]</sup>	5.98	1.95	13.54
ORB-SLAM-3 方法 <sup>[25]</sup> (4 Hz)	5.29	2.16	3.75
视觉惯性联合优化 (4 Hz)	13.82		
VINS-Mono 方法 <sup>[29]</sup>	22.05	1.39	2.32
ORB-SLAM-3 方法 <sup>[25]</sup> (10 Hz)	9.09	1.26	2.63
视觉惯性联合优化 (10 Hz)	19.60		

根据表 2 中对比实验的结果,可以得出以下结论:

- 1) 不经初始化而直接进行视觉惯性联合优化的系统尺度误差仅小于 VINS-Mono 方法,说明初始化对 VINS 系统状态估计是有意义的;
- 2) 改进 M-K 方法准确性较高,但是耗时较久,反映了联合初始化的特点,即求解精度高,但求解过程太过复杂从而导致实时性较差;
- 3) VINS-Mono 方法耗时最少,但是准确性最差,反映了非联合初始化的特点,即求解实时性强,但分步求解忽略了后求解参数对先求解参数的影响从而导致准确性较差;
- 4) ORB-SLAM-3 方法不仅有着最好的准确性,实时性方面也与 VINS-Mono 方法接近,反映了半联合初始化的特点,即在不影响求解实时性的前提下保证了良好的准确性。

## 7 VINS 初始化未来发展趋势

VINS 研究已有数十年的历史,自 2012 年预积分算法被提出并推广以来,VINS 研究又焕发了新的活力,尤其是初始化算法,借助预积分算法取得了新一轮的进展和成果。此外,深度学习方法与更多新型传感器也逐渐被 VINS 研究所接纳和融合。综合目前 VINS 的相关研究现状,本文认为 VINS 初始化存在以下发展趋势:

- 1) 运动激励。在不同的运动激励下,VINS 状态以及尺度信息的可观测性有所不同,而潜

在的运动退化(例如悬停或匀速运动)风险,也有可能导致初始化结果出现偏差。初始化研究需要对初始化运动激励相关问题进行深入研究,探究各种瞬时/持续的运动激励以及运动状态(例如运动退化)对初始化结果可能造成的影响。

2) 性能评估。目前,初始化性能评估方法主要通过将初始化估计值与数据真值对比进行评估。但是在初始化完成阶段,系统通过非真值信息判断初始化结果是否满足性能要求,对于VINS来说更为重要。

3) 持续初始化。近年来,越来越多的工作认为初始化不应只存在于系统初始阶段,而应在系统运行的全阶段持续进行,以应对尺度信息随时间变化带来的挑战。文献[59-61]所提出的视觉惯性系统,均在系统运行的全过程中加入类似初始化的模块,对系统尺度信息进行持续更新。然而,需要注意的是,持续初始化在提升状态估计精度的同时,不可避免地会带来时间开销的增长。

4) 深度学习融合。目前,已有一系列方法利用深度学习技术对IMU测量数据进行降噪处理<sup>[62-65]</sup>,以获得无噪声或低噪声IMU数据并应用在不同领域中。基于此思路,可以提出基于深度学习的IMU Bias估计算法,结合目前已有的初始化方法,提升初始化性能。此外,文献[66]中提出了一种基于单目深度估计网络、结合深度信息的全新VINS初始化框架,同样值得关注。

5) 特征处理。在VO与SLAM领域,多项工作<sup>[3, 67-69]</sup>已经表明,更丰富的视觉特征可以带来更精确的状态估计结果。对于初始化,更多视觉特征的加入,例如线特征,也可以有效提升视觉信息处理的准确性<sup>[37]</sup>。但是视觉特征的增加会带来计算开销的增长,需要在性能和实时性之间寻求平衡。

6) 时空参数。VINS需要同时处理来自IMU和视觉传感器的信息,如果无法对两者真实采样时间之间的关系进行标定,那么VINS状态估计精度则有可能受到影响。目前,多数时间标定方法相对于初始化过程独立<sup>[70-71]</sup>,仅文献[48]所提初始化方法可以在估计初始化参数的同时进行相机-IMU时间标定。未来还需要更多更为完善的、包含传感器时间标定的VINS初始化方法。

7) 事件相机。事件相机(也称作神经形态视

觉传感器<sup>[72]</sup>)是一种全新的仿生视觉传感器,具有高时域分辨率、高动态范围、低数据冗余和低功耗等优势<sup>[72-73]</sup>,近年来在机器人视觉领域引起了广泛关注。目前,虽然已经有一些基于事件相机的VINS工作<sup>[74-76]</sup>,但是这些工作并没有提供相应的初始化方法。基于事件相机的VINS初始化方法将成为未来发展趋势。

8) 更多传感信息。虽然视觉传感器与IMU具有强互补关系,但将VINS应用于大规模复杂场景中时,仍可能存在鲁棒性不佳、误差积累等问题。融合更多传感信息成为解决此类问题的一种思路,例如加入GNSS<sup>[50-51, 77-78]</sup>或磁力计<sup>[49]</sup>的VINS等。对于这类融合更多传感信息的VINS,针对不同传感组合、涉及更多初始化求解参数的初始化方法将成为未来发展趋势。

9) 预积分算法。目前常用的IMU预积分算法主要包括流形上<sup>[22]</sup>和基于运动方程<sup>[23]</sup>两种,均在离散空间对预积分进行推导,在求解过程中使用离散近似,虽然简化了计算过程,但降低了求解精度。文献[79]中提出了一种基于连续空间的预积分推导方法,给出了更为精确的预积分封闭式解,但是由于计算过程相对复杂,目前仅被部分方法<sup>[80-81]</sup>使用。如何在保证实时性的前提下,通过改进预积分算法提升初始化精度,也将是未来发展趋势之一。

10) 面向任务场景的应用适配。随着VINS越来越多的广泛应用于自动驾驶、VR/AR、无人系统定位导航等领域,其面对不同任务场景和应用需求时,对初始化算法的要求也越来越复杂和多元。一方面,初始化算法在面对不同传感器平台以及不同任务场景时,应具备更强的适应性;另一方面,如何针对特殊场景或特殊应用需求对初始化过程进行改进优化,使其满足不同的需求,是未来VINS和初始化研究工作亟须解决的问题。

11) 数据集。目前适合VINS初始化研究的数据集并不算多,VINS初始化未来的发展需要更多更具针对性的VINS初始化数据集作为支撑。这类数据集不仅需要提供准确的视觉、IMU传感数据以及系统状态真值,还需要提供不同运动激励、不同传感器观测条件下的真实场景数据。

## 8 结论

视觉惯性导航系统一直是机器人导航、计算机视觉等领域不可忽视的研究热点。作为其



中的关键部分,初始化可以看作是一种求解参数更多、先验信息更少的视觉惯性传感信息融合过程。本文通过对近年来众多研究工作的回顾和分析,对 VINS 初始化方法进行了分类梳理与综述报告,并对其未来发展趋势进行了总结、讨论和展望。可以预见,随着新技术的发展、机器人硬件水平的提升,初始化方法将得到进一步发展,VINS 也会取得更加精确的自主定位导航以及环境感知效果,为人类探索未知、改善生活带来更多便利。

## 参考文献(References)

- [1] CADENA C, CARLONE L, CARRILLO H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age [J]. IEEE Transactions on Robotics, 2016, 32(6): 1309–1332.
- [2] YANG N, VON STUMBERG L, WANG R, et al. D3VO: deep depth, deep pose and deep uncertainty for monocular visual odometry [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 1281–1292.
- [3] GOMEZ-OJEDA R, MORENO F A, ZUÑIGA-NOËL D, et al. PL-SLAM: a stereo SLAM system through the combination of points and line segments [J]. IEEE Transactions on Robotics, 2019, 35(3): 734–746.
- [4] 王浩,卢德玖,方宝富. 动态环境下基于增强分割的 RGB-D SLAM 方法[J]. 机器人, 2022, 44(4): 418–430.  
WANG H, LU D J, FANG B F. RGB-D SLAM method based on enhanced segmentation in dynamic environment [J]. Robot, 2022, 44(4): 418–430. (in Chinese)
- [5] REBECQ H, HORSTSCHAEFER T, GALLEGÓ G, et al. EVO: a geometric approach to event-based 6-DOF parallel tracking and mapping in real time [J]. IEEE Robotics and Automation Letters, 2017, 2(2): 593–600.
- [6] 史殿习,杨卓越,金松昌,等. 面向数据共享的多无人机协同 SLAM 方法[J]. 计算机学报, 2021, 44(5): 983–998.  
SHI D X, YANG Z Y, JIN S C, et al. A multi-UAV collaborative SLAM method oriented to data sharing [J]. Chinese Journal of Computers, 2021, 44(5): 983–998. (in Chinese)
- [7] LI P L, QIN T, HU B T, et al. Monocular visual-inertial state estimation for mobile augmented reality [C]//Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2017.
- [8] 于雅楠,卫红,陈静. 基于局部熵的 SLAM 视觉里程计优化算法[J]. 自动化学报, 2021, 47(6): 1460–1466.  
YU Y N, WEI H, CHEN J. Optimization algorithm of visual odometry for SLAM based on local image entropy [J]. Acta Automatica Sinica, 2021, 47(6): 1460–1466. (in Chinese)
- [9] QIN T, CHEN T Q, CHEN Y L, et al. AVP-SLAM: semantic visual mapping and localization for autonomous vehicles in the parking lot [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021.
- [10] KASYANOV A, ENGELMANN F, STÜCKLER J, et al. Keyframe-based visual-inertial online SLAM with relocalization [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017.
- [11] MUR-ARTAL R, TARDÓS J D. Visual-inertial monocular SLAM with map reuse [J]. IEEE Robotics and Automation Letters, 2017, 2(2): 796–803.
- [12] USENKO V, DEMMEL N, SCHUBERT D, et al. Visual-inertial mapping with non-linear factor recovery [J]. IEEE Robotics and Automation Letters, 2020, 5(2): 422–429.
- [13] BLOESCH M, OMARI S, HUTTER M, et al. Robust visual inertial odometry using a direct EKF-based approach [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015.
- [14] LEUTENEGGER S, LYNEN S, BOSSE M, et al. Keyframe-based visual-inertial odometry using nonlinear optimization [J]. The International Journal of Robotics Research, 2015, 34(3): 314–334.
- [15] QIN T, LI P L, SHEN S J. VINS-mono: a robust and versatile monocular visual-inertial state estimator [J]. IEEE Transactions on Robotics, 2018, 34(4): 1004–1020.
- [16] HUANG G Q. Visual-inertial navigation: a concise review [C]//Proceedings of International Conference on Robotics and Automation (ICRA), 2019.
- [17] 罗世彬,刘海桥,胡茂青,等. 无人飞行器异源图像匹配辅助惯性导航定位技术综述[J]. 国防科技大学学报, 2020, 42(6): 1–10.  
LUO S B, LIU H Q, HU M Q, et al. Review of multi-modal image matching assisted inertial navigation positioning technology for unmanned aerial vehicle [J]. Journal of National University of Defense Technology, 2020, 42(6): 1–10. (in Chinese)
- [18] 吴禹彤,张林,陈善秋,等. 先进视觉惯性导航系统综述 [C]//惯性技术与智能导航学术研讨会论文集, 2019: 120–127.  
WU Y T, ZHANG L, CHEN S Q, et al. A survey of advanced visual-inertial navigation system [C]//Proceedings of the Symposium on Inertial Technology and Intelligent Navigation, 2019: 120–127. (in Chinese)
- [19] 司书斌,赵大伟,徐婉莹,等. 视觉—惯性导航定位技术研究进展 [J]. 中国图象图形学报, 2021, 26(6): 1470–1482.  
SI S B, ZHAO D W, XU W Y, et al. Review on visual-inertial navigation and positioning technology [J]. Journal of Image and Graphics, 2021, 26(6): 1470–1482. (in Chinese)
- [20] LUPTON T, SUKKARIEH S. Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions [J]. IEEE Transactions on Robotics, 2012, 28(1): 61–76.

- [21] FORSTER C, CARLONE L, DELLAERT F, et al. IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation [C]//Proceedings of Robotics: Science and Systems, 2015.
- [22] FORSTER C, CARLONE L, DELLAERT F, et al. On-manifold preintegration for real-time visual: inertial odometry[J]. IEEE Transactions on Robotics, 2017, 33(1): 1–21.
- [23] SHEN S J, MICHAEL N, KUMAR V. Tightly-coupled monocular visual-inertial fusion for autonomous flight of rotorcraft MAVs [C]//Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2015.
- [24] 施俊屹, 查富生, 孙立宁, 等. 移动机器人视觉惯性SLAM研究进展[J]. 机器人, 2020, 42(6): 734–748.
- SHI J Y, ZHA F S, SUN L N, et al. A survey of visual-inertial SLAM for mobile robots[J]. Robot, 2020, 42(6): 734–748. (in Chinese)
- [25] CAMPOS C, MONTIEL J M M, TARDÓS J D. Inertial-only optimization for visual-inertial initialization [C]//Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2020.
- [26] MARTINELLI A. Closed-form solution of visual-inertial structure from motion [J]. International Journal of Computer Vision, 2014, 106(2): 138–152.
- [27] KAISER J, MARTINELLI A, FONTANA F, et al. Simultaneous state initialization and gyroscope bias calibration in visual inertial aided navigation [J]. IEEE Robotics and Automation Letters, 2017, 2(1): 18–25.
- [28] YANG Z F, SHEN S J. Monocular visual-inertial state estimation with online initialization and camera-IMU extrinsic calibration [J]. IEEE Transactions on Automation Science and Engineering, 2017, 14(1): 39–51.
- [29] QIN T, SHEN S J. Robust initialization of monocular visual-inertial estimation on aerial robots [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017.
- [30] HUANG W B, LIU H. Online initialization and automatic camera-IMU extrinsic calibration for monocular visual-inertial SLAM [C]//Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2018.
- [31] CAMPOS C, MONTIEL J M M, TARDÓS J D. Fast and robust initialization for visual-inertial SLAM [C]//Proceedings of International Conference on Robotics and Automation (ICRA), 2019.
- [32] MARTINELLI A. Vision and IMU data fusion: closed-form solutions for attitude, speed, absolute scale, and bias determination [J]. IEEE Transactions on Robotics, 2012, 28(1): 44–60.
- [33] DOMÍNGUEZ-CONTI J, YIN J F, ALAMI Y, et al. Visual-inertial SLAM initialization: a general linear formulation and a gravity-observing non-linear optimization [C]//Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2019.
- [34] MARTINELLI A, RENZAGLIA A, OLIVA A. Cooperative visual-inertial sensor fusion: fundamental equations and state determination in closed-form [J]. Autonomous Robots, 2020, 44: 339–357.
- [35] EVANGELIDIS G, MICUSIK B. Revisiting visual-inertial structure-from-motion for odometry and SLAM initialization [J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1415–1422.
- [36] MICUSIK B, EVANGELIDIS G. Renormalization for initialization of rolling shutter visual-inertial odometry [J]. International Journal of Computer Vision, 2021, 129(6): 2011–2027.
- [37] LIU H, QIU J Y, HUANG W B. Integrating point and line features for visual-inertial initialization [C]//Proceedings of International Conference on Robotics and Automation (ICRA), 2022.
- [38] BARTOLI A, STURM P. Structure-from-motion using lines: representation, triangulation, and bundle adjustment [J]. Computer Vision and Image Understanding, 2005, 100(3): 416–441.
- [39] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147–1163.
- [40] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras [J]. IEEE Transactions on Robotics, 2017, 33(5): 1255–1262.
- [41] LIN Y, GAO F, QIN T, et al. Autonomous aerial navigation using monocular visual-inertial fusion [J]. Journal of Field Robotics, 2018, 35(1): 23–51.
- [42] MU X F, CHEN J, ZHOU Z X, et al. Accurate initial state estimation in a monocular visual-inertial SLAM system [J]. Sensors, 2018, 18(2): 506.
- [43] LI J Y, BAO H J, ZHANG G F. Rapid and robust monocular visual-inertial initialization with gravity estimation via vertical edges [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019.
- [44] JUNG J H, CHUNG J Y, CHA J, et al. Rapid initialization using relative pose constraints in stereo visual-inertial odometry [C]//Proceedings of IEEE 15th International Conference on Control and Automation (ICCA), 2019.
- [45] 龚赵慧, 张霄力, 彭侠夫, 等. 基于视觉惯性融合的半直接单目视觉里程计 [J]. 机器人, 2020, 42(5): 595–605.
- GONG Z H, ZHANG X L, PENG X F, et al. Semi-direct monocular visual odometry based on visual-inertial fusion [J]. Robot, 2020, 42(5): 595–605. (in Chinese)
- [46] CHENG J, ZHANG L Y, CHEN Q H. Fast monocular visual-inertial initialization with an improved iterative strategy [J/OL]. Journal of Sensors, 2021, 2021: 5565158.
- [47] CHENG J, ZHANG L Y, CHEN Q H. An improved initialization method for monocular visual-inertial SLAM [J]. Electronics, 2021, 10(24): 3063.
- [48] HUANG W B, WAN W W, LIU H. Optimization-based online initialization and calibration of monocular visual-inertial odometry considering spatial-temporal constraints [J]. Sensors, 2021, 21(8): 2673.

- [49] WANG J Z, LI L L, YU H, et al. VIMO: a visual-inertial-magnetic navigation system based on non-linear optimization [J]. *Sensors*, 2020, 20(16): 4386.
- [50] JIN R H, LIU J N, ZHANG H P, et al. Fast and accurate initialization for monocular vision/INS/GNSS integrated system on land vehicle [J]. *IEEE Sensors Journal*, 2021, 21(22): 26074–26085.
- [51] XIONG L, KANG R, ZHAO J Q, et al. G-VIDO: a vehicle dynamics and intermittent GNSS-aided visual-inertial state estimator for autonomous driving [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 11845–11861.
- [52] NISTÉR D. An efficient solution to the five-point relative pose problem [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(6): 756–777.
- [53] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: an accurate open-source library for visual, visual-inertial, and multimap SLAM [J]. *IEEE Transactions on Robotics*, 2021, 37(6): 1874–1890.
- [54] ZUÑIGA-NOËL D, MORENO F A, GONZALEZ-JIMENEZ J. An analytical solution to the IMU initialization problem for visual-inertial systems [J]. *IEEE Robotics and Automation Letters*, 2021, 6(3): 6116–6122.
- [55] YATHIRAJAM B, MEENAKSHISUNDARAM V S, MUNIYAPPA A C. An efficient approach to initialization of visual-inertial navigation system using closed-form solution for autonomous robots [J]. *Journal of Intelligent & Robotic Systems*, 2021, 101(3): 59.
- [56] 吴涛, 谢志军. 一种改进的视觉/IMU 导航系统初始化方法 [J]. *压电与声光*, 2021, 43(6): 863–868.  
WU T, XIE Z J. An improved initialization method of visual/IMU navigation system initialization [J]. *Piezoelectrics & Acousto-optics*, 2021, 43(6): 863–868. (in Chinese)
- [57] 刘刚, 葛洪伟. 视觉惯导 SLAM 初始化算法研究 [J]. *计算机科学与探索*, 2021, 15(8): 1546–1554.  
LIU G, GE H W. Research on initialization algorithm for visual-inertial SLAM system [J]. *Journal of Frontiers of Computer Science and Technology*, 2021, 15(8): 1546–1554. (in Chinese)
- [58] BURRI M, NIKOLIC J, GOHL P, et al. The EuRoC micro aerial vehicle datasets [J]. *International Journal of Robotics Research*, 2016, 35(10): 1157–1163.
- [59] VON STUMBERG L, USENKO V, CREMERS D. Direct sparse visual-inertial odometry using dynamic marginalization [C]// *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2018.
- [60] HONG E, LIM J. Visual-inertial odometry with robust initialization and online scale estimation [J]. *Sensors*, 2018, 18(12): 4287.
- [61] VON STUMBERG L, CREMERS D. DM-VIO: delayed marginalization visual-inertial odometry [J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 1408–1415.
- [62] ESFAHANI M A, WANG H, WU K Y, et al. OriNet: robust 3-D orientation estimation with a single particular IMU [J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 399–406.
- [63] BROSSARD M, BONNABEL S, BARRAU A. Denoising IMU gyroscopes with deep learning for open-loop attitude estimation [J]. *IEEE Robotics and Automation Letters*, 2020, 5(3): 4796–4803.
- [64] LIU W X, CARUSO D, ILG E, et al. TLIO: tight learned inertial odometry [J]. *IEEE Robotics and Automation Letters*, 2020, 5(4): 5653–5660.
- [65] CHEN D P, WANG N, XU R S, et al. RNIN-VIO: robust neural inertial navigation aided visual-inertial odometry in challenging scenes [C]// *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2021.
- [66] ZHOU Y W, KAR A, TURNER E, et al. Learned monocular depth priors in visual-inertial initialization [C]// *Proceedings of the European Conference on Computer Vision*, 2022.
- [67] 王丹, 黄鲁, 李焱. 基于点线特征的单目视觉同时定位与地图构建算法 [J]. *机器人*, 2019, 41(3): 392–403.  
WANG D, HUANG L, LI Y. A monocular visual SLAM algorithm based on point-line feature [J]. *Robot*, 2019, 41(3): 392–403. (in Chinese)
- [68] LI X, HE Y J, LIN J L, et al. Leveraging planar regularities for point line visual-inertial odometry [C]// *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.
- [69] LIU Z, SHI D X, LI R H, et al. PLC-VIO: visual-inertial odometry based on point-line constraints [J]. *IEEE Transactions on Automation Science and Engineering*, 2022, 19(3): 1880–1897.
- [70] QIN T, SHEN S J. Online temporal calibration for monocular visual-inertial systems [C]// *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.
- [71] QIU K J, QIN T, PAN J, et al. Real-time temporal and rotational calibration of heterogeneous sensors using motion correlation analysis [J]. *IEEE Transactions on Robotics*, 2021, 37(2): 587–602.
- [72] 李家宁, 田永鸿. 神经形态视觉传感器的研究进展及应用综述 [J]. *计算机学报*, 2021, 44(6): 1258–1286.  
LI J N, TIAN Y H. Recent advances in neuromorphic vision sensors: a survey [J]. *Chinese Journal of Computers*, 2021, 44(6): 1258–1286. (in Chinese)
- [73] GALLEGO G, DELBRÜCK T, ORCHARD G, et al. Event-based vision: a survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 44(1): 154–180.
- [74] ZHU A Z, ATANASOV N, DANILIDIS K. Event-based visual inertial odometry [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [75] VIDAL A R, REBECQ H, HORSTSCHAEFER T, et al. Ultimate SLAM? Combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios [J]. *IEEE Robotics and Automation Letters*, 2018, 3(2): 994–1001.
- [76] LE GENTIL C, TSCHOPP F, ALZUGARAY I, et al. IDOL:

- a framework for IMU-DVS odometry using lines [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems ( IROS) ,2021.
- [77] LEE W , GENEVA P , YANG Y L , et al. Tightly-coupled GNSS-aided visual-inertial localization [C]//Proceedings of International Conference on Robotics and Automation ( ICRA) ,2022.
- [78] CAO S Z , LU X Y , SHEN S J. GVINS: tightly coupled GNSS-visual-inertial fusion for smooth and consistent state estimation [J]. IEEE Transactions on Robotics , 2022 , 38( 4) : 2004 – 2021.
- [79] ECKENHOFF K , GENEVA P , HUANG G Q. High-accuracy preintegration for visual-inertial navigation [C]//Proceedings of the 12th International Workshop on the Algorithmic Foundations of Robotics( WAFR) ,2016.
- [80] ECKENHOFF K , GENEVA P , HUANG G Q. Direct visual-inertial navigation with analytical preintegration [C]//Proceedings of IEEE International Conference on Robotics and Automation ( ICRA) ,2017.
- [81] ECKENHOFF K , GENEVA P , HUANG G Q. Closed-form preintegration methods for graph-based visual-inertial navigation [J]. The International Journal of Robotics Research ,2019 ,38( 5) : 563 – 586.