中图分类号：V219 论文编号：1028701 26-S

学科分类号：082500

学术学位硕士学位论文

室内环境下四旋翼视觉惯性融合

定位导航研究

研究生姓名 王天舒

学科专业 航空宇航科学与技术

研究方向 飞行控制与仿真

指导教师 史志伟 教授

南京航空航天大学

研究生院 航空学院

二О二六年三月

Nanjing University of Aeronautics and Astronautics

The Graduate School

College of Aerospace Engineering

**Research on Vision-Inertial Integrated Localization and Navigation for Quadrotors in Indoor Environments**

A Thesis in

Aeronautical and Astronautical Science and Technology

by

Wang Tianshu

Advised by

Prof. Shi Zhiwei

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Master of Engineering

March, 2026

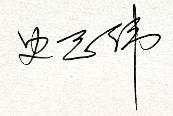
独创性声明

本人声明所呈交的硕士学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得南京航空航天大学或其他教育机构的学位或证书而使用过的材料。

研究生签名：  日期： 2026.3.6

使用授权声明

本人完全了解南京航空航天大学有关保留、使用学位论文的规定，即学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权南京航空航天大学可以将本学位论文的全部内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

研究生签名：导师签名：日期： 2026.3.6

摘要

**关键词：**

**ABSTRACT**

**Key Words:**

目录

[第1章 绪论 1](#_Toc215067060)

[1.1 背景和意义 1](#_Toc215067061)

[1.2 国内外研究现状 2](#_Toc215067062)

[1.2.1 视觉惯性融合定位研究现状 2](#_Toc215067063)

[1.2.2 无人机运动规划研究现状 6](#_Toc215067064)

[1.3 本文主要工作 8](#_Toc215067065)

[1.4 论文组织结构 8](#_Toc215067066)

[第2章 自主定位导航问题建模与相关原理 9](#_Toc215067067)

[2.1 引言 9](#_Toc215067068)

[2.2 视觉惯性融合定位原理 9](#_Toc215067069)

[2.2.1 视觉前端处理 9](#_Toc215067070)

[2.2.2 IMU预积分 10](#_Toc215067071)

[2.2.3 初始化与运动估计 11](#_Toc215067072)

[2.3 三维感知与局部地图构建 12](#_Toc215067073)

[2.4 四旋翼动力学模型与微分平坦特性分析 15](#_Toc215067074)

[2.4.1 坐标系及坐标变换 15](#_Toc215067075)

[2.4.2 四旋翼刚体模型 16](#_Toc215067076)

[2.4.3 四旋翼微分平坦特性分析 17](#_Toc215067077)

[2.5 基于ROS的仿真环境搭建 19](#_Toc215067078)

[2.6 本章小结 20](#_Toc215067079)

[第3章 四旋翼轨迹规划与跟踪控制研究 21](#_Toc215067080)

[3.1 引言 21](#_Toc215067081)

[3.2 基于B样条优化的轨迹规划方法 21](#_Toc215067082)

[3.2.1 轨迹规划整体框架 21](#_Toc215067083)

[3.2.2 轨迹参数化 22](#_Toc215067084)

[3.2.3 前端路径搜索与控制点重分配 23](#_Toc215067085)

[3.2.4 基于梯度的后端轨迹优化 25](#_Toc215067086)

[3.3 基于PID的轨迹跟踪控制方法 27](#_Toc215067087)

[3.4 仿真实验与验证 29](#_Toc215067088)

[3.4.1 Octave轨迹跟踪数值仿真 29](#_Toc215067089)

[3.4.2 轨迹规划与跟踪全流程软件在环仿真 31](#_Toc215067090)

[3.5 本章小结 31](#_Toc215067091)

[第4章 基于IMU输入的神经网络四旋翼短时位姿预测 32](#_Toc215067092)

[4.1 引言 32](#_Toc215067093)

[4.2 视觉短时失效下视觉惯性融合定位结果分析 32](#_Toc215067094)

[4.3 四旋翼视觉惯性位姿估计可观测性分析（待定） 32](#_Toc215067095)

[4.4 仿真数据集及神经网络模型设计 32](#_Toc215067096)

[4.4.1 LSTM网络数学模型 32](#_Toc215067097)

[4.4.2 LSTM网络模型结构设计 33](#_Toc215067098)

[4.5 IMU位姿预测神经网络模型训练 35](#_Toc215067099)

[4.5.1 数据集构建 35](#_Toc215067100)

[4.5.2 LSTM神经网络建模 36](#_Toc215067101)

[4.6 IMU位姿预测神经网络模型验证 37](#_Toc215067102)

[4.6.1 LSTM预测结果与分析 37](#_Toc215067103)

[4.6.2 软件在环仿真 38](#_Toc215067104)

[4.7 本章小结 40](#_Toc215067105)

[第5章 室内环境下四旋翼自主定位与导航实验 41](#_Toc215067106)

[5.1 引言 41](#_Toc215067107)

[5.2 四旋翼无人机硬件平台 41](#_Toc215067108)

[5.3 自主定位与导航系统架构 43](#_Toc215067109)

[5.4 常规工况下四旋翼轨迹规划与跟踪控制实验验证 44](#_Toc215067110)

[5.5 视觉短时失效下四旋翼轨迹规划与跟踪控制实验验证 44](#_Toc215067111)

[5.6 本章小结 44](#_Toc215067112)

[第6章 总结与展望 45](#_Toc215067113)

[6.1 研究总结 45](#_Toc215067114)

[6.2 研究展望 45](#_Toc215067115)

[参考文献 46](#_Toc215067116)

[致谢 49](#_Toc215067117)

[在学期间的研究成果及发表的学术论文 50](#_Toc215067118)

[附录 附录名称 51](#_Toc215067119)

图表清单

[图1.1 室内旋翼无人机应用场景 1](#_Toc215067120)

[图2.1 VINS-Fusion视觉前端处理示意图 10](#_Toc215067121)

[图2.2 IMU预积分示意图[4] 11](#_Toc215067122)

[图2.3 重投影误差示意图[] 11](#_Toc215067123)

[图2.4 视觉惯性对齐初始化[] 12](#_Toc215067124)

[图2.5 射线投射示意图 14](#_Toc215067125)

[图2.6 微分平坦输出和参考坐标系 17](#_Toc215067126)

[图2.7 基于ROS的软件在环仿真架构 19](#_Toc215067127)

[图3.1 基于B样条优化的轨迹规划方法示意图 21](#_Toc215067128)

[图3.2 均匀B样条凸包性质示意图[] 22](#_Toc215067129)

[图3.3 障碍物内的控制点重分配示意图 25](#_Toc215067130)

[图3.4 PID轨迹跟踪控制结构图 29](#_Toc215067131)

[图3.5 四旋翼螺旋上升轨迹跟踪位置曲线 30](#_Toc215067132)

[图 3.6 四旋翼速度阶跃轨迹跟踪位置曲线 31](#_Toc215067133)

[图4.1 LSTM神经网络单元 33](#_Toc215067134)

[图4.2 IMU位姿预测LSTM深度神经网络结构 34](#_Toc215067135)

[图4.3 伯努利双扭线飞行轨迹 36](#_Toc215067136)

[图4.4 模型误差随迭代次数变化曲线 37](#_Toc215067137)

[图4.5 不同IMU高斯噪声下模型预测位姿误差曲线 38](#_Toc215067138)

[图4.6 不同IMU高斯噪声和飞行周期下模型预测位姿误差仿真曲线 39](#_Toc215067139)

[图4.7 视觉短时失效下模型预测位姿自主导航轨迹图 39](#_Toc215067140)

[图5.1 四旋翼无人机硬件平台 41](#_Toc215067141)

[图5.2 英特尔RealSense D435i双目相机 42](#_Toc215067142)

[图5.3 英特尔13代NUC 42](#_Toc215067143)

[图5.4 Pixhawk 2.4.8飞行控制处理器 43](#_Toc215067144)

[图5.5 无人机自主定位与导航系统软件框架 44](#_Toc215067145)

[表2.1 占据栅格地图算法 13](#_Toc215067146)

[表3.1 A\*寻路算法 24](#_Toc215067147)

[表3.2 四旋翼关键物理参数 29](#_Toc215067148)

[表4.1 数据集参数 36](#_Toc215067149)

注释表

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

缩略词

|  |  |
| --- | --- |
| 缩略词 | 英文全称 |
| UAV | Unmanned Aerial Vehicle |
| GPS | Global Positioning System |
| INS | Inertial Navigation System |
| SLAM | Simultaneous Localization And Mapping |
| IMU | Inertial Measurement Unit |
| VO | Visual Odometry |
| VIO | Visual-Inertial Odometry |
| ROS | Robot Operating System |

# 绪论

## 背景和意义

无人飞行器（Unmanned Aerial Vehicle，UAV），简称无人机，专指通过远程操控或者 机载程控的不载人飞行器。根据气动原理的不同，传统无人机可以分为固定翼无人机、旋翼无人机、扑翼无人机和复合翼无人机四大类。对于场景受限的室内环境来说，旋翼无人机因其结构简单、可悬停、机动性高的特点，受到国内外无人机研究领域的极大关注。

|  |  |
| --- | --- |
| 湖北首次开展变电站室内无人机智能巡视 - 湖北日报新闻客户端 | 无人机室内巡检_智能巡检系统_无人机智能系统开发-系统定制公司-上海魁鲸科技 |
| (a) 服务器巡检 | (b) 仓库物流管理 |
| 湖北首次开展变电站室内无人机智能巡检_设备_进行_运行 | 工单巡检智能系统和无人技术-上海魁鲸科技 |
| (c) 电力设备巡检 | (d) 工单智能巡检 |

图1.1 室内旋翼无人机应用场景

提供实时、准确的无人机位姿（位置、姿态）信息，是智能多旋翼无人机自主决策、规划以及底层集成控制的必要前提。目前，已有的室内无人机定位方法主要为相对定位方法，包括惯性导航系统（Inertial Navigation System，INS）、激光雷达（Light Detection and Ranging，LIDAR）里程计、视觉里程计（Visual Odometry，VO）、视觉同时定位与地图构建（Visual Simultaneous Localization and Mapping，V-SLAM）、视觉惯性里程计（Visual Inertial Odometry，VIO）。INS利用惯性测量单元（Inertial Measurement Unit，IMU）的三轴加速度计和三轴陀螺仪数据，通过滤波融合算法实现无人机在惯性坐标系下的局部相对位置、速度以及姿态估计，具有数据更新频率高和短时间内精度高等优点，但随时间的累积误差发散[1]，虽然高精度INS可以维持较长一段时间的定位精度，但成本较高。激光雷达里程计通过激光雷达对无人机周围动态环境进行三维扫描，利用三维点云匹配与位姿估计算法实现无人机的相对定位，通过融合 IMU可以进一步提高系统的鲁棒性，具有实时性好、定位精度高、算法成熟等优点，具有市场应用的潜力，但由于激光雷达尺寸偏大导致在无人机上安装部署受限，同时多线激光雷达的成本也相对较高。相比之下，视觉传感器具有体积小、成本低、感知信息丰富等优点，被广泛应用于自主无人机的目标识别与定位导航等领域。VO和V-SLAM都是基于视觉传感器的相对定位方法，VO是通过相机采集的图像序列进行机器人的位姿估计，V-SLAM是在VO的基础上拓展了地图构建和回环检测功能，在具有回环的封闭场景下可以提高VO的定位精度，但在室外开放环境下VO的适用性更为普遍。但是，环境纹理、光照变化以及载体的高机动运动工况导致的图像模糊等情况对VO的鲁棒性造成了巨大挑战，相关研究表明基于深度学习的VO和基于多源传感器融合的VIO可以明显改善VO的鲁棒性，提高定位精度。不过，基于深度学习的VO计算量较大，需要依赖GPU的加速，相比之下，VIO通过融合消费级IMU和视觉传感器,具有鲁棒性好、隐蔽性高、微型化、成本低等优势，在无人机和移动机器人的位姿估计领域得到了大量应用。然而，将VIO直接应用于室内无人机定位目前还是面临一些问题：由于环境单一和快速运动情况下导致图像特征的缺失，VIO系统的可观测性会发生变化，导致无人机的定位精度低甚至定位失效[2]。

因此，改善VIO系统应用于无人机的可观测性，提高无人机的定位精度，是目前亟需解决的科学问题。本文以四旋翼无人机为研究对象，提出了一种基于IMU输入的神经网络四旋翼短时位姿预测方法，在存在视觉特征模糊甚至失效等不确定观测情况下实现四旋翼无人机的高精度相对定位，同时设计了一种基于均匀B样条优化的轨迹规划方法和基于PID的轨迹跟踪控制方法，最终实现复杂室内环境下四旋翼无人机的自主导航与避障。

## 国内外研究现状

### 视觉惯性融合定位研究现状

单一传感器难以适应日益复杂的态势环境，如惯性导航系统的误差随时间累积特性、视觉里程计受可见光和纹理特征限制等，使得仅使用单一传感器的无人系统难以提供正确和鲁棒的导航信息。随着机载飞行器计算能力的不断提高，利用互补融合的优势，对多源传感器测量的多源异构数据进行融合，可以显著提高导航系统的鲁棒性和环境适应性。经过30多年的发展，数据融合导航系统的框架已经日趋成熟，可以分为基于滤波的、基于优化的和基于学习的。

（1）基于滤波的视觉惯性融合研究进展

1960年，Kalman[3]提出了离散随机线性系统的Kalman滤波器，这是一种高效的递归滤波器，用于从一系列不完全和含噪声的测量中估计动态系统的状态，广泛应用于组合导航领域。利用泰勒展开对非线性方程进行线性化的拓展卡尔曼滤波器（Extended Kalman Filter，EKF）在非线性估计[4]中得到了广泛的应用。Mourikis和Roumeliotis[5]在2007年提出了多状态约束卡尔曼滤波器（MSCKF），首次将EKF应用于视觉辅助下的INS。它是目前许多基于滤波的视觉惯性同步定位与映射（VI-SLAM）的基础，它解决了EKF-SLAM算法的尺寸爆炸问题，但它导致了偏航通道的观测不一致。Li和Mourikis[6]进一步对其进行了优化，在不增加额外计算成本的情况下保证了正确的可观测性，进一步提高了算法估计的一致性和准确性。Sun等人[7]进一步将MSCKF算法扩展到多视点融合框架，以获得比优化方案更高的计算效率和旋翼机18m/s的高速导航。开源视觉惯导系统（VINS）是Geneva等人[8]于2019年提出的基于MSCKF的开源VINS算法。主要特征包括基于KF的流形滑动窗口、在线相机内外参数校准、视觉惯性时间偏移校准、具有不同表示和一致的第一估计雅可比矩阵的SLAM地标处理、状态管理的模块化系统、可扩展的视觉惯性系统模拟器以及用于算法评估的广泛工具包。对于低成本的无人机导航问题，EKF算法框架已在工程领域广泛应用，用于融合多源传感器数据，为无人机提供准确的姿态信息。Martin和salan[9]使用23-D EKF融合IMU/GPS/磁力计和气压计，并测试了导航系统在单个传感器故障情况下的鲁棒性。作为无人机的开源工程项目，PX4估计和与控制库（ECL）作为一个面向无人机的开源工程项目，广泛应用于许多项目中，利用EKF框架对无人机的状态进行估计。它采用松散耦合的分布式框架，根据测量值对维持的状态变量进行EKF校正，并使用互补滤波校正EKF中不同传感器融合引起的时间延迟。ECL、EKF算法的特点是采用多源传感器组合进行状态估计，使系统具有较高的容错性，但也大大增加了系统的计算量和参数整定难度。

为了以更高的精度和更快的计算速度处理非线性系统滤波问题，Julier和Uhlmann[10]提出了基于无迹变换的无迹卡尔曼滤波器（UKF）。其关键思想是生成许多表示状态向量及其转移模型的sigma点，经过非线性变换后，这些sigma采样点能够以至少二阶精度近似系统状态的后验期望和协方差（Taylor展开式）。UKF比EKF具有更好的非线性滤波精度和稳定性。为了进一步提高强非线性场景下状态估计的精度，van der Merwe等[11]提出在粒子滤波框架下通过EKF/UKF生成重要度分布函数，有效调制重要样本及其权值，提高算法的滤波精度。然而，该方法的计算量远远高于EKF或UKF。

松耦合滤波器法VIO是分别将IMU运动学位姿估计结果与视觉里程计（VO）[12]的位姿估计结果通过EKF或UKF等滤波器进行融合，具有计算效率高的优点，但没有考虑到视觉特征数据和IMU数据的关联性从而导致有效信息丢失。Lynen等提出的MSF-EKF（Multi-Sensor-Fusion Extended Kalman Filter）[13]是基于松耦合滤波器法的VIO，采用松耦合滤波器机制可以灵活扩展多种传感器数据，在无人机上取得较好的实验效果。

（2）基于优化的视觉惯性融合研究进展

多源传感器融合问题可以表述为使用贝叶斯图网络计算系统状态随时间的最大后验（MAP），最优解决方案一旦接收到新度量就执行批处理优化。与只考虑前一时刻状态和协方差的滤波框架不同，当加入新状态时，前一状态会被边缘化，这一过程会导致信息丢失[14]。非线性优化算法基于接收到的所有可用测量值，通过求解非线性最小二乘问题，迭代算法得到系统的最优状态估计，它为多源传感器融合[15][16]提供了一种新的思路。Dellaert[17]回顾了因子图算法在高级机器人优化问题中的应用，阐述了因子图算法在状态估计、定位构建、轨迹规划等领域的应用和优势。

求解效率的提高为因子图算法奠定了基础。基于优化方法的最初目的是解决关键构图帧中状态特征点的视觉同步定位和尺寸展开问题，代表作品有ORB-SLAM[18]和LSD-SLAM[19]。两者的主要区别在于前端视觉预处理部分的特征提取方法。由于纯视觉受纹理特征限制，易受光照影响，难以应对快速旋转场景，研究人员在数据融合结构中引入惯性测量，形成互补优势。面对IMU的高速率采样，最简单的方法是在IMU测量接收时，在因子图中添加一个新的状态。但是，这种方法会严重影响实时性能的优化。为了解决这一问题，Lupton和Sukkarieh[20]首先提出使用无初值的增量INS方法求解高动态下的VIO问题。通过重新参数化相对运动约束，可以避免因初始条件不同而导致的重复积分。IMU预积分最重要的理论价值在于，通过计算视觉关键帧间隔内状态的增量值，可以方便地直接将IMU增量因子添加到因子图中。

随着预积分理论的建立，基于IMU优化的方法得以成功应用，并出现了Leutenegger等人[21]提出的基于关键帧的视觉惯性SLAM（OKVIS）的经典工作，提出了一种将惯性测量与基于关键帧的视觉惯性SLAM紧密结合的方法，并首次使用关键帧技术进行滑动窗口优化，适用于短时间导航。Forster等[22]进一步改进了基于李群的IMU预积分算法，分析了基于雅可比扰动模型的误差传播规律，使得预积分和不确定性在李群中的传播趋于成熟。Qin等[23]基于IMU预积分理论，提出了VINS-Mono。作为一种完整的VIO方法，它以视觉束平差和IMU预积分残差作为代价函数，利用滑动窗口将单目视觉和IMU紧密耦合。此外，该方法还具有鲁棒初始化、闭环检测和边缘化等功能，增强了长期导航的适应性。VINS-Fusion进一步扩展了双目立体视觉和惯性融合[16]。VI-DSO[24]将直接稀疏里程计（DSO）扩展到VIO。作为面向地图构建的SLAM算法，ORB-SLAM、ORB-SLAM2和ORB\_SLAM3[25]增加了一种IMU融合算法，并基于新改进的重定位模块构建了混合地图，从而提高了ORB-SLAM3在特征点较少场景下的鲁棒性，提出了完全融合建图的概念：Atlas[26]。如果在系统运行过程中丢失了跟踪，则建图不会停止，而是会开始一个新的子建图，当访问历史公共段建图时，它可以与以前的建图完美融合。

（3）基于学习的视觉惯性融合研究进展

传统的基于滤波器和基于优化的方法仍然使用传感器的数学模型和误差传播过程，更依赖于系统的初始化精度和传感器数据的精度。对于强非线性问题，非线性优化算法比基于滤波器的方法更有优势。由于深度人工智能算法进一步增强了非线性拟合能力，并且深度学习过程与人类的学习过程非常吻合，因此受到越来越多的关注。然而，深度神经网络（DNN）算法也面临着非常困难的挑战，主要是因为DNN内部复杂的多层非线性结构使得人工将自适应机制整合到黑盒模型中变得困难[27]。

随着深度学习技术的发展，近年来出现了许多基于深度神经网络的端到端状态估计算法。视觉解自运动的核心是估计不同特征点的深度，然后用PnP法求解摄像机的自运动。在过去的十年中，许多工作都在使用深度神经网络进行视觉深度估计。2014年，Eigen等[28]提出了一种双尺度DNN，显示了其像素深度估计的可行性，该方法需要使用图像和相应的深度真值进行训练，算法模型直接学习原始像素，不需要任何人工标记的特征。该算法在训练过程中依赖于现有的高质量、像素对齐的校准深度真值。近年来，Goodfellow等人[29]因将计算机视觉问题视为两个网络之间的对抗任务而受到欢迎。有些工作使用经典的生成对抗网络（GAN）从单个图像估计深度。Wu等[30]提出了一种新的空间对应(SC)-GAN，不依赖于相机姿态估计精度的影响。Aleotti等[31]提出了一种深度估计模型，该模型使用递归GAN以无监督的方法从立体图像中估计深度。这些工作证明了GAN在深度图估计中的有效性。

一般来说，在神经网络中使用IMU测量的线性加速度和角速度值进行惯性里程测量的工作较少。首先，惯性导航的数学模型及其误差更加完备，而神经网络带来的不可解释性反而会成为不可预测的风险；其次，神经网络需要大量各种条件下的实验数据进行训练，降低了其适用性。然而，对于低成本的MEMS惯性器件，其误差具有强非线性和高时变的特点，在GNSS拒止或视觉纹理不足的环境下，仅依靠双惯性积分会迅速积累误差。许多学者尝试使用深度学习框架来解决IMU等长时间序列数据的建模问题，以提高数据融合的精度。Chen等[32]基于递归深度神经网络（RDNN）提出了一种端到端惯性里程计网络（IONet）框架来预测惯性里程计，通过大样本数据训练实现了基于室内低成本IMU的位置估计；实验结果证明，与捷联惯导系统相比，精度有了显著提高。然而，在更动态的环境中或者训练数据和测试数据不在同一领域时，网络可能会降低算法的准确性，这也是神经网络固有的不足之一。Li等[33]从IMU误差建模的角度，构建了基于长短期记忆（LSTM）的深度神经网络来提高IMU模型的表示能力，并构建了多种综合激励训练集来训练LSTM模型，通过实验验证可以获得良好的补偿效果。

使用深度神经网络处理高维特征（如图像）具有更多优势，并且许多端到端的工作都是基于纯视觉或VIO完成的。VINet[34]是第一个基于端到端学习的VIO算法。与传统方法相比，VINet消除了相机与IMU之间繁琐的手动同步和校准。近年来，已有几种成功的无监督深度估计方法，如Mahjourian等[35]通过构建无监督学习框架，解决了里程计和深度估计问题。Zhou等[36]提出了一种联合无监督学习方法，从多个未标记的RGB图像帧中获得车辆的姿态估计和深度估计；该模型输入连续的视觉RGB帧，输出每帧图像之间的姿态估计和深度估计。这些VO算法仅通过帧间的空间信息来估计车辆的自运动，没有充分利用帧间的时间信息，因此估计不够连续和准确。UnDeepVO[37]是另一种无监督深度和自我运动估计方法，但它只能在立体图像对数据集上进行训练，难以直接集成IMU数据。VIO-Learner[38]是最近提出的使用无监督学习框架的多RGB深度相机，它使用学习优化器来最小化自我运动估计的光度损失。虽然不需要地面真实里程数据，但系统的深度输入有时可以为网络提供外部监督。SelfVIO[39]提出了一种新的基于自监督深度学习的VIO和深度图重建方法，该方法可以在没有IMU固有参数和IMU-相机空间参数的情况下运行VIO，并通过开源数据集实验证明了该算法的优点。

日益复杂的民用/军用需求将要求未来的无人机配备更多的多源、异构传感器。在相对理想的环境下，基于滤波器的方法和基于优化的方法都可以提供高精度的输出；然而，在复杂环境和高速运动场景中，鲁棒性对融合算法是一个很大的挑战。针对不准确的传感器测量、不准确的系统建模、复杂的环境动力学和不符合实际的约束所带来的模型不确定性，未来可以使用安全可靠和可解释的数据驱动的方法来进行探索研究。

### 无人机运动规划研究现状

#### 基于采样的运动规划

基于采样的运动规划器通过探索和开发，专注于问题的全局解决方案，其中复杂性主要来自配置空间。概率动态图（PRM）和快速探索随机树（RRT）都是概率完备的，因为它们的失败概率随着样本数量趋近于无穷大而指数性地衰减至零。Karaman和Frazzioli提出了PRM和RRT的渐近最优变体，称为PRM\*和RRT\*，确保随着样本数量的增加收敛到全局最优解。还有一些算法进一步提高了随机运动规划的效率或适用性。我们的方法利用基于采样的规划器来克服环境带来的复杂性。它实现了与给定低维无碰撞路径同伦的动态可行轨迹的优化。它被设计成灵活地纳入系统状态-输入约束，而这不是基于采样方法的优势。通过这种方式，环境和动态带来的复杂性都被分解并征服。

#### 基于优化的运动规划

基于优化的规划者通过利用问题的高阶信息来专注于局部解决方案。他们依赖特定环境的预处理方法，以便将障碍物信息编码到优化中。轨迹优化在控制领域的广泛系统中已经研究了很长时间[23]。许多通用方法被设计用于高质量的解决方案，例如基于配点的方法GPOPS-II[24]和基于射击的方法ACADO[25]。它们将原始问题转录为一个非线性规划(NLP)，使用大量的等式和变量，然后依赖成熟的NLP求解器，如SNOPT[26]或IPOPT[27]。然而，机器人中的轨迹规划可能会施加难以表述的约束、非光滑性和整数变量。此外，通用方法通常需要较长的计算时间，使得它们不适合时间关键的任务。例如，Bry等人[28]报告称，使用SNOPT的直接配点方法需要几分钟来优化一个在圆柱形障碍物之间飞行的12状态飞机的4.5米轨迹[29]。因此，专门的方法正在被呼唤以克服这些困难。

对于微分平坦的多旋翼无人机，运动规划可以被转换为低维平坦输出轨迹的优化问题。Mellinger和Kumar[12]采用固定时长的样条来表征四旋翼的平坦轨迹。他们将轨迹四阶时间导数的平方范数积分作为二次规划（Quadratic Programming）的代价函数来保证轨迹的平滑性，并以线性不等式来保证轨迹的安全性。尽管如此，该方案缺乏样条各段时间的有效优化。为了获得代价函数对时间变量的梯度，该二次规划的摄动问题需要被多次求解以获得对时间变量的灵敏度，这极大地提高的问题求解的计算复杂度。此外，该方案仅支持过度简化的动力学约束。Bry等人[43]随后给出了该二次规划在仅有高阶连续性约束和航点约束时候的闭式解。他们采用RRT\*算法作为可行路径生成的前端，并不断地在后端无约束轨迹发生碰撞的时候，从该可行路径上取出路径点作为后端轨迹生成的航点。这种方法固然高效，但是难以在障碍物稠密的环境中生成高质量的轨迹。此外，无人机的飞行速度等动力学约束无法被显式地纳入到轨迹生成中。更重要的是，该闭式解中待求逆矩阵的奇异性并没有被很好地讨论清楚。Deits和Tedrake[53]首先用凸多面体来近似可行空间，这样将多项式轨迹约束在该多面体的线性不等式中可以完全被一组多项式平方和（Sum of Squares）条件等价地描述。他们采用混合整数二阶锥规划（Mixed Integer Second Order Conic Programming）来求解轨迹各段的分配与优化问题。该方法可以得到固定段数分辨率下的全局最优轨迹，然而其计算效率仅适用于离线计算。Gao等人[21]同样采用凸多面体来描述可行区域。他们交替地对轨迹的几何形状以及时域特性进行优化。轨迹的安全性通过Bézier曲线控制点上的线性约束得到保障。轨迹的时域信息则利用时间最优路径重参数化[54]（Time Optimal Path Parameterization）来完成。此外，许多现有方法可以归结为以上方法的变体。例如，Tordesillas等人[55]将Deits等人的方法中的多项式平方和条件替换成了Bézier控制点的线性约束，从而避免了二阶锥约束，对应的混合整数二次规划（Mixed Integer Quadratic Programming）通过牺牲了约束的零间隙特性提高了计算求解的效率。Sun等人[56]则采用双层规划（Bilevel Programming）的灵敏度分析来获得轨迹各段时间关于底层约束二次规划最优值的解析灵敏度，进而做到直接时域优化，然而该灵敏度的存在性、光滑性及其导致的数值问题没有被很好地解决。

## 本文主要工作

本文针对xxx，介绍xxx，接着就xxx进行讨论，然后针对xxx进行介绍，重点阐述xxx，最后介绍xxx。

## 论文组织结构

本论文的组织结构如下：

第一章 阐明课题的研究背景，介绍课题的研究现状和来源，并概述论文的主要工作和创新点。

第二章；

第三章

第四章；

第五章 总结全文，对今后的研究做出总结和展望。

# 自主定位导航问题建模与相关原理

## 引言

无人机自主定位导航系统包含自主定位、局部感知与运动规划等模块，模块之间相互联系，存在共同的理论前提。本章首先介绍无人机视觉惯性融合定位原理，之后对本系统所使用的局部感知和建图方法进行了介绍，感知建图是无人机轨迹规划的必要前提，接着对四旋翼无人机的动力学模型和微分平坦特性进行了描述和分析，该特性证明了四旋翼无人机各状态量之间的联系，为减少无人机的规划变量提供了理论依据，简化了无人机在未知环境中的轨迹规划问题。最后搭建了基于机器人操作系统（Robot Operate System，ROS）的仿真环境用于四旋翼自主定位与导航算法的验证

## 视觉惯性融合定位原理

实现室内环境下四旋翼的自主定位是其完成自主导航的前提，传统VO是通过相机采集的图像序列进行四旋翼的位姿估计，然而，环境纹理、光照变化以及载体的高机动运动工况导致的图像模糊等情况对VO的鲁棒性造成了巨大挑战， VIO通过融合消费级IMU和视觉传感器，具有鲁棒性好、隐蔽性高、微型化、成本低等优势。本节从双目视觉惯性里程计的角度，并以代表开源算法VINS-Fusion为例，系统分析了视觉惯性融合的定位原理及其特性。

### 视觉前端处理

VINS-Fusion视觉前端处理流程图如图2.1所示。首先使用Shi-Tomasi 角点[]检测器提取图像特征点，Shi-Tomasi角点检测通过计算图像局部窗口的自相关矩阵特征值，选取较小特征值作为角点响应值来识别图像中的显著角点，能稳定地检测适用于光流跟踪的优质特征点。然后对左右目图像进行立体匹配，利用视差和三角相似原理恢复提取的图像特征点的深度信息。接着采用随机采样一致性算法（Random Sample Consensus，RANSAC）剔除误匹配点，提高特征点匹配的鲁棒性。最后采用 LK (Lucas-Kanade) 光流法在连续帧间跟踪特征点：通过假设相邻帧间局部图像块的光流恒定，利用最小二乘法求解特征点的像素位移，从而建立特征点的运动轨迹。该方法结合图像金字塔实现多尺度跟踪，能够适应特征点的大幅度运动，同时利用 IMU 预积分提供运动先验，进一步提高光流跟踪的准确性和鲁棒性，为后续的视觉-惯性紧耦合优化提供可靠的特征对应关系。

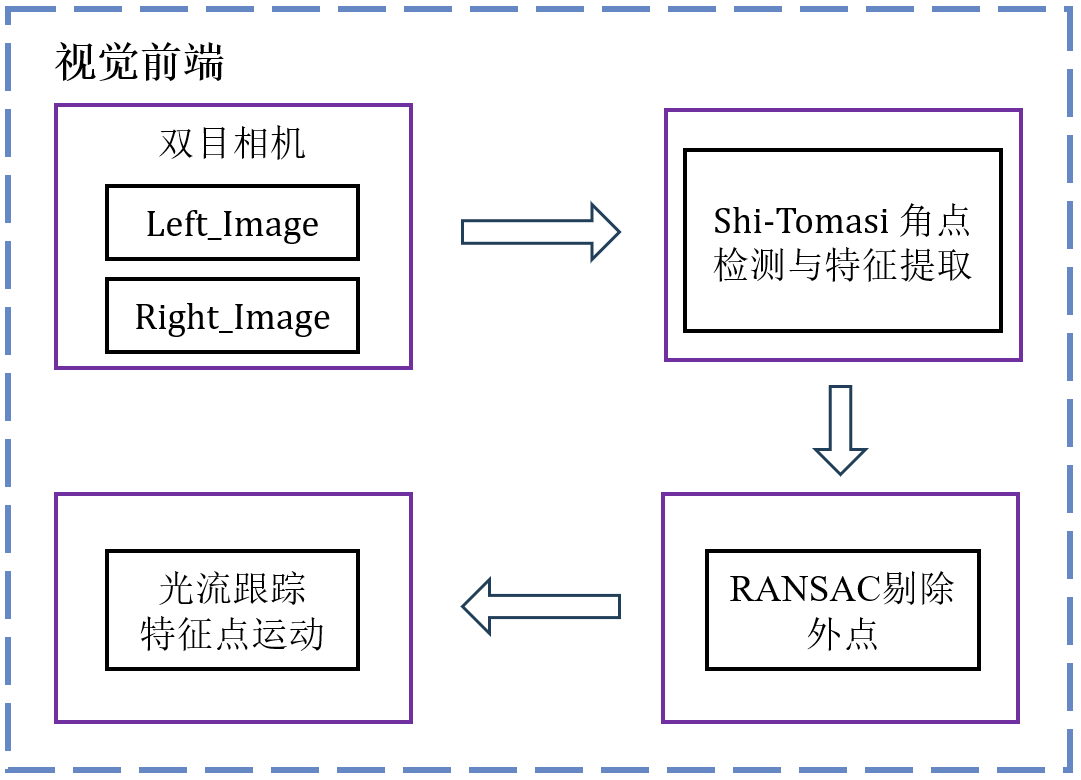


图2.1 VINS-Fusion视觉前端处理示意图

### IMU预积分

IMU能够测量传感器本体的角速度和加速度，但这些数据都存在明显的漂移，使得积分两次得到的位姿非常不可靠，随着传感器运行时间的推移，累积误差会不断增大。但是对于短时间内的快速运动，IMU能提供较好的运动估计，能弥补视觉传感器无漂移但频率低的缺点。于短时间内的快速运动，IMU能提供较好的运动估计，能弥补视觉传感器的缺点。由于相机成像在快速运动时可能会出现运动模糊，导致无法进行特征提取和跟踪，从而丢失了一些视觉传感器的观测数据。IMU传感器和相机不同，IMU的观测频率是相机的几倍甚至上十倍，也就是在连续两帧RGB图像之间，存在多个IMU传感器的观测值。所以为了获得每帧图像对应的IMU数据，就需要对两个图像帧之间的IMU数据进行积分，才能实现图像帧和IMU数据的配对。采用IMU预积分还有一个原因，在后端全局优化的框架下，经常需要对历史状态进行更新，如果不使用预积分的话，每当一个状态发生变化时，就需要从头往后，运用每一帧IMU数据进行计算，直至更新完所有的状态量为止，这样会导致后端优化的耗时大幅度增加。IMU预积分可采用中值积分法，通过上个时刻求解出来的位姿变化数据，来推测下一时刻的位姿。IMU预积分示意图如图2.2所示。在图2.2中，最上方的坐标轴代表时间，RGB以较低的频率发布出来，而IMU数据在两帧RGB之间已经有多个观测值。IMU预积分将第i时刻的图像帧和第j时刻的图像帧之间的IMU数据进行预积分。在视觉信息不足或质量不高的情况下，如在动态或纹理贫乏的环境中，纯视觉的初始化可能变得困难。而IMU预积分提供的连续运动估计可以减轻视觉初始化的负担，提供额外的无人机运动信息来支持系统的初始化。

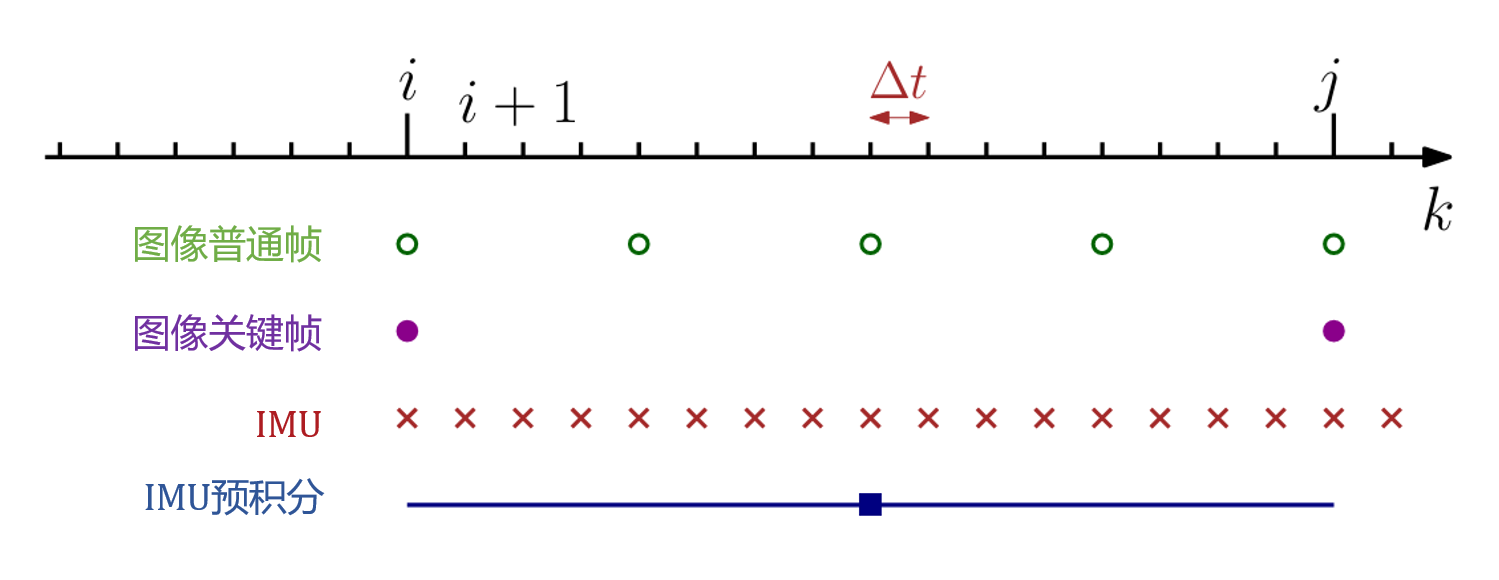


图2.2 IMU预积分示意图[4]

### 初始化与运动估计

紧耦合视觉惯性里程计是一个高度非线性的系统，故在运动估计前需要对系统进行准确的初始化。VINS-Fusion的双目初始化过程通过联合优化视觉与惯性测量数据，实现系统状态量的鲁棒估计。该过程首先基于双目视觉几何恢复无尺度的场景结构，通过特征匹配与三角化获取相邻帧间的相对位姿，相机初始位姿的求解可以使用PnP算法[56]进行求解。PnP是一种求解3D到2D点对运动的方法，通过图像中的2D特征及其在空间中的3D点来估计相机位姿。求解PnP问题的方法有许多，本文采取最小化重投影误差求解PnP，图2.3展示的是一个空间点在两帧相机图像的重投影误差示意图。随后，利用IMU预积分得到的运动约束与视觉结构进行对齐，构建线性方程组求解尺度因子、速度、重力向量及传感器偏差，如图2.4所示。其中，陀螺仪偏置通过最小化预积分旋转与视觉旋转的差异进行标定，而重力方向则通过引入幅值约束在切线空间内进一步优化。最终，系统将重力向量与世界坐标系z轴对齐，完成所有状态量的初始化。这一过程充分融合了双目相机的尺度观测能力与IMU的高频运动信息，显著提升了初始化的精度和鲁棒性。

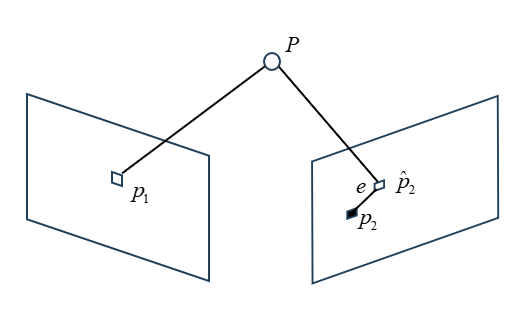


图2.3 重投影误差示意图[]

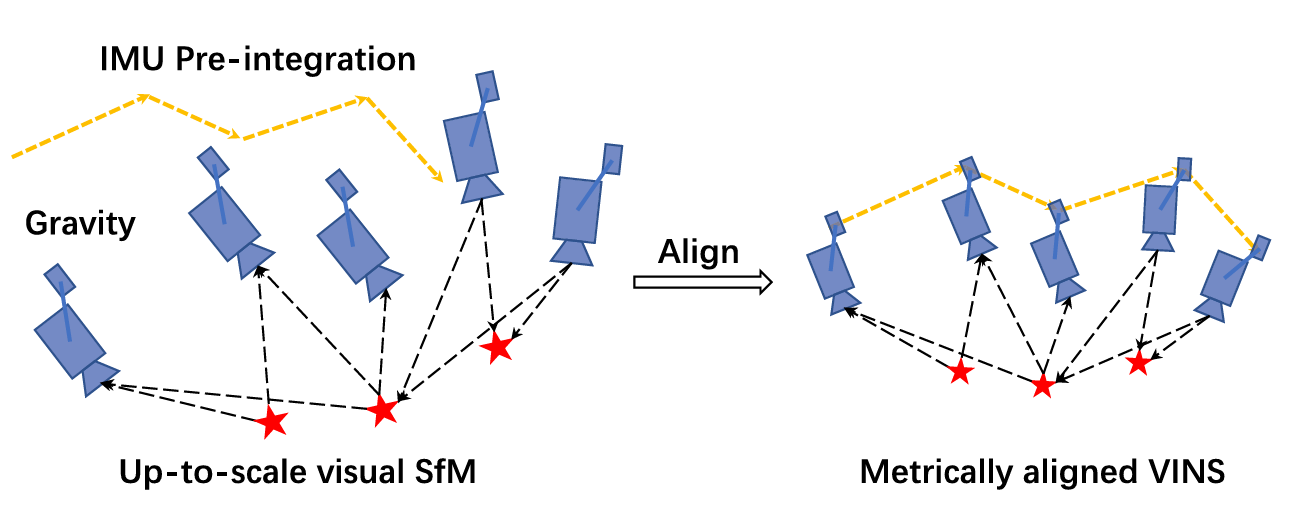


图2.4 视觉惯性对齐初始化[]

在初始化完成后，基于滑动窗口的紧耦合双目VIO继续接收RGB图像流和IMU数据流，以实现高精度和鲁棒的状态估计。滑动窗口中的变量如下：



式中，是捕捉第帧图像时对应的IMU状态，包括世界坐标系下IMU的位置、速度和方向以及机体坐标系下的加速度和陀螺仪偏置。代表滑动窗口中的关键帧总数，是滑动窗口中的特征总数。代表第一次观测中所有特征点的逆深度，即深度值的倒数。里程计采取滑动窗口对窗口内的变量进行优化，里程计的初始值是初始化流程的最后一个状态。采取滑动窗口的原因是为了保持优化变量的个数在一定范围内，需要使用滑动窗口算法动态增加或移除优化变量来保证系统的实时性。

## 三维感知与局部地图构建

鉴于双目相机得到的点云存在较多的噪点，无法准确表达环境特征，难以直接用于导航任务。而简洁高效的地图可以简化轨迹规划的难度，以及降低算法实现难度，因此本系统选择使用三维占据栅格地图[]表征未知环境。栅格地图作为机器人学中最常用的地图形式，通过将一个封闭而连续的空间细分为等大的小格，实现对环境的离散化描述。这种离散化的地图使得环境描述和碰撞检测变得更直接、更简单，因为判断一条给定路径是否可行仅需检验路径所通过的栅格是否为可通行状态。在栅格地图里，一个位置要么被标记为无障碍的自由空间，要么是被障碍物占据的空间，且这些状态是确定的。尽管如此，确定一个点的具体状态依赖于传感器对该点的观测结果，由于测量误差的存在，会使得单次测量结果的准确性受到影响，进而可能导致对该点状态的误识别。

占据栅格地图构建算法的黄金定律是根据给定的有噪声和不确定的传感器测量数据计来算整个地图的后验概率。



式中，为地图；为直到时刻的所有测量值；为所有无人机位姿定义的路径；表示第个栅格单元，占用栅格地图将空间分割为有限多个栅格单元，将地图中的一个栅格单元被占据的概率表示为，自由的概率表示为。将二值贝叶斯滤波器应用于占据栅格地图构建可得算法如下：

表2.1 占据栅格地图算法

|  |
| --- |
| 算法1：占据栅格地图算法(,,) |
| for all cells  do |
| if  in perceptual field of  then |
|  |
| else |
|  |
| endif |
| endfor |
| return |

若栅格状态仅通过概率值来表达，多次观测的累加或累减操作会使得其超出区间，为克服此问题，本系统选择以概率的对数形式来描述每个栅格状态的置信度，这样做将概率值的取值范围扩展为，确保了在连续观测中，栅格状态的更新保持在合理的范围内，的定义如公式(2.3)所示：



并由此计算后验概率为



表1给出的占据栅格地图算法遍历所有的栅格单元，并更新所有传感器测量锥内的测量值。对于传感器测量锥内的栅格单元，用第4行中的函数的结果来更新占用值，否则，占用值保持不变。常数是用对数让步比所表示的先验占用概率，即



函数应用了反演测量模型的对数型式，即



依据以上更新公式，每当接收到新的测量数据时，便可对地图内的栅格状态进行刷新。通过将最新的观测信息应用于地图上相应点的概率状态，采用递推公式进行更新。对于在当前数据中被直接观测到的点，使用逆观测模型进行状态更新，而那些未在最新观测中出现的点，其状态保持为前一次的状态不变。在构建三维栅格地图时，一旦确定了更新机制，接下来的关键任务便是将从深度相机收集到的点云数据转化为栅格的测量信息。深度相机生成的点云数据是激光光束遇到障碍后反射回传感器，从而确定障碍表面的坐标位置。在单次测量中，光线经过的点都视为未被占据，而与障碍物相交的终点为占据状态。通过对深度相机的光束进行三维空间投射，可以计算出它所通过的空闲栅格和被占据栅格的坐标。这一过程被称为射线投射[53]。如图2.5所示，蓝色方格为光线途径的没有障碍物的自由栅格，黄色方格为光线末端遇到障碍物的占据栅格，基于无人机的当前位置以及检测到的障碍物表面位置，可以确定路径上经过的栅格坐标。假设在深度图像中观测到某个像素深度为，则在对应的三维空间点上该点是被占据的，从相机光心到该空间点的线段上不存在障碍物，不然该线段会被遮挡。利用射线投射技术获取的信息可以进行地图更新。

三维栅格地图更新的输入包括无人机里程计信息、深度图像以及坐标变换，根据射线投射后更新当前无人机位姿状态下的栅格状态，然后将栅格地图进行膨胀。因为在本系统中，四旋翼无人机看成是一个质点模型，忽略其体积大小，这样便于对无人机与环境障碍物的关系表达，但引入了不安全性，因此对栅格地图进行膨胀处理，通常是以无人机轴距的一半为膨胀系数进行膨胀。



图2.5 射线投射示意图

## 四旋翼动力学模型与微分平坦特性分析

### 坐标系及坐标变换

地面坐标系记为，用于研究飞行器相对于地面的运动状态，确定机体的空间位置坐标。它忽略地球曲率，即将地球表面假设成一张平面。通常以飞行器起飞位置或者地心作为坐标原点。轴指向地平面某任意选定方向；轴铅垂向下；轴垂直平面，按右手定则确定。

机体坐标系记为，用于研究四旋翼飞行器相对于重心的旋转运动，其原点取在飞行器的重心位置上。轴在飞行器对称平面内并指向飞行器的前进方向。轴亦在飞行器对称平面内，垂直轴向下。轴垂直于平面，按右手定则确定。

机体坐标系相对于地面坐标系的方位，或者说飞机在空中的姿态，常用三个欧拉角表示。偏航角：机体轴在水平面上的投影与轴之间的夹角。规定飞行器右偏航时形成的角度为正。滚转角：飞行器对称平面与包含轴的铅垂平面之间的夹角。规定飞行器向右滚转时形成的角度为正。俯仰角：机体轴与水平面之间的夹角。规定当飞行器头部上仰时形成的角度为正。

地面坐标系先绕轴方向转过角，然后绕当时的轴方向转过角，最后绕当时的轴方向转过角，就可与重合。三次旋转对应的旋转矩阵分别为：







按坐标系转换一般法则，可得出由到的转换矩阵为：



设机体旋转角速度为，四旋翼姿态角为，由坐标变换过程可得机体角速度与欧拉角的导数之间的转换关系为：



即：



### 四旋翼刚体模型

为了简便，在建立四旋翼飞行器的刚体运动学模型和刚体动力学模型时，作如下假设：

假设1 四旋翼是刚体。

假设2 四旋翼质量和转动惯量不变。

假设3 四旋翼几何中心与重心一致。

假设4 四旋翼飞行速度较慢，因此忽略空气阻力的影响，认为四旋翼无人机只受重力和旋翼拉力的影响，其中重力沿轴正方向，拉力沿轴负方向。

假设5 四旋翼奇数号桨叶逆时针转动，偶数号桨叶顺时针转动。

本文采用基于欧拉角的运动学模型，基于假设1构建四旋翼的刚体运动学模型具体形式如下：



其中，表示四旋翼在地面坐标系下的空间位置，表示四旋翼在地面坐标系下的速度。

基于假设4，本文只考虑桨盘水平的多旋翼，在地面坐标系内构建位置动力学模型如下：



其中，表示四旋翼无人机的总质量，为重力加速度，表示螺旋桨总拉力的大小。

基于假设1~假设3，在机体坐标系内构建姿态动力学方程如下：



其中，表示四旋翼的转动惯量，表示螺旋桨在机体轴上产生的力矩。

### 四旋翼微分平坦特性分析

微分平坦是非线性系统中的一个概念，可以用来描述无人系统的一种特性。考虑如下的仿射非线性系统:



其中，满秩输入映射矩阵，系统状态为，控制输入为。一个微分平坦的动力系统必定存在一组由状态和控制的有限阶导数唯一决定的平坦输出（Flat Output），同时和也可以被该平坦输出及其有限阶导数参数化：





其中平坦变换和均被系统方程和唯一决定。形象地说，该动力系统的状态和控制可以从的瞬态信息中直接计算出来，而不需要像一般动力系统对微分方程做依赖于初值的积分。

本文按照前述Z-X-Y旋转顺序进行世界坐标系到机体坐标系的变换，如图2.6所示，图中表明了无人机机体坐标系和世界坐标系、中间坐标系的关系。

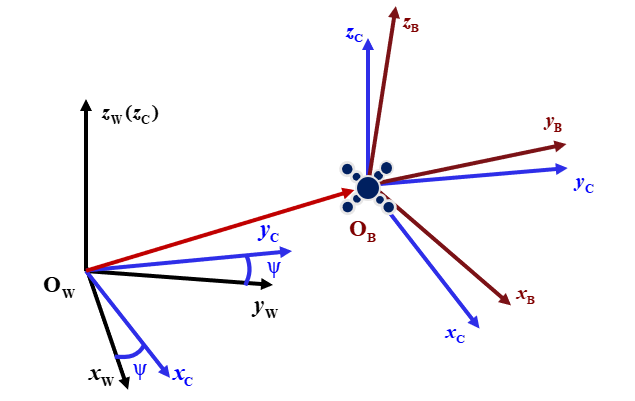


图2.6 微分平坦输出和参考坐标系

四旋翼无人机的状态由其位置、姿态角、速度以及姿态角速度组成，可以表示为，控制量，选择微分平坦输出，分别表示无人机质心在世界坐标系下的三轴坐标和航向角，则有，。

机体角速度在世界坐标系下可以表示为：



基于公式(2.14)，无人机机体z轴在世界坐标系下的表达式如下：



记，将用单位向量表示后，进一步简化为：



根据图2.6中给出的中间坐标系，记偏航角为，则中间坐标系的轴在世界坐标系中可描述为：



那么根据右手法则可以得出无人机机体坐标系的和可以定义为：



其中，至此，机体坐标系在世界坐标系下的三轴单位向量已经被表示成平坦输出的函数，又由于姿态角和旋转矩阵之间存在相互转换的关系，因此三个姿态角也可以表示成的函数。进一步对无人机所受合力公式(2.14)两边求导：



将该表达式沿着投影，并利用，同时记，可得到：



将式(2.19)两边同时叉乘可得：



再根据式(2.25)可以得到：



至于机体角速度分量可以通过直接获得，由于中间坐标系相对世界坐标系只有绕的一个偏航旋转，并且观察到不含分量，因而可以将表示为：



同理，角加速度沿和的分量可以通过对式(2.14)两边求二阶导数并按照上面的步骤获得，至于的分量可以通过推到得到，将两边同时乘以后，结合和，则的分量为。同时根据式(2.20)和(2.21)可得。这样，四旋翼无人机系统的状态变量和控制量都可以通过所定义的平坦输出及其有限阶导数来表示，从而证明了四旋翼无人机系统具备微分平坦特性。在做无人机轨迹规划的过程中，难以在十二维的全状态空间规划，但可以找到一组微分平坦输出变量，其他的状态变量都可以由这四个状态变量组合表示，这样的四个变量就是无人机的位置以及偏航角。在对无人机进行轨迹规划时，只需要规划这四个变量即可，微分平坦特性的存在降低了轨迹规划的难度。

## 基于ROS的仿真环境搭建

在自主无人机实飞之前，需要通过软件在环仿真（SITL）对四旋翼的自主定位与导航功能进行验证，本节搭建了基于ROS、PX4和Gazebo的仿真架构，介绍了各部分的功能以及互相之间的通信形式，具体如图2.7所示。

整个仿真环境都是在基于ROS的通讯架构下运行的，Gazebo物理引擎定义四旋翼无人机的物理信息，并发布无人机上视觉传感器和IMU传感器的数据信息；VIO接收来自Gazebo的传感器数据并发布无人机位姿信息和局部地图点云信息；轨迹规划部分接收全局目标点信息和局部地图点云信息并发布期望轨迹信息，PX4飞控软件接收来自轨迹规划的当前局部目标点信息和来自VIO的位姿信息并发布无人机的控制作动指令；最后由Gazebo接收并执行无人机的控制作动指令，完成四旋翼自主定位与导航的闭环仿真。

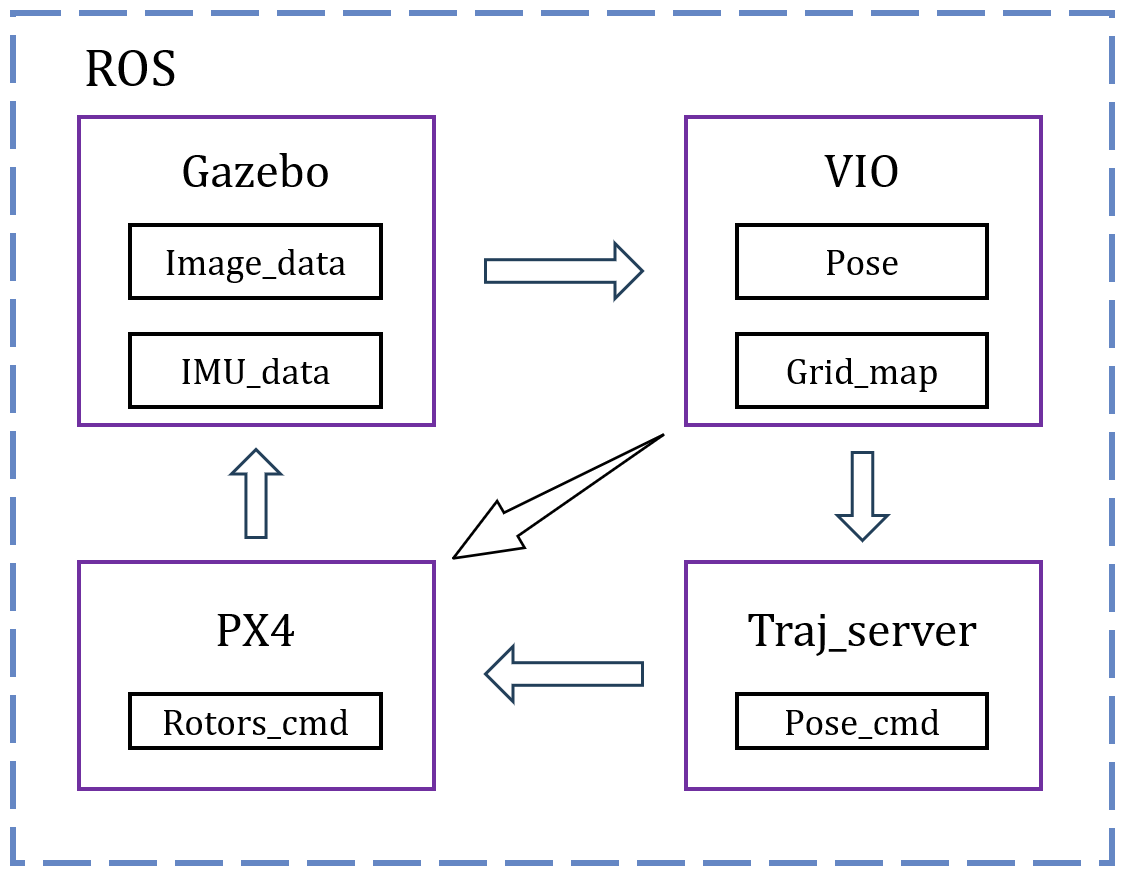


图2.7 基于ROS的软件在环仿真架构

## 本章小结

本章对自主定位与导航系统的基础原理和技术进行了阐述，在紧耦合VIO定位框架的基础上，首先对本系统所使用的双目相机和IMU观测数据处理进行了阐述；然后，根据双目相机产生的深度点云信息，描述了本系统对三维世界的感知过程和未知环境下的局部地图建立过程；接着，定义了本文使用的坐标系和坐标变换，建立了四旋翼的刚体动力学模型，并针对欠驱动的四旋翼无人机系统状态量较多，直接进行轨迹规划难度大的问题，分析了作为非线性系统的四旋翼无人机拥有的微分平坦特性，从状态量和控制量两个方面都证明了可以通过一组特定平坦输出来表达状态量和控制量，从而将四旋翼无人机轨迹规划的问题转换成对无人机位置及其偏航角进行规划的问题，避免在全状态空间规划。最后，搭建了基于ROS、PX4和Gazebo的仿真架构，为无人机自主定位与导航算法的验证提供了先验平台。

# 四旋翼轨迹规划与跟踪控制研究

## 引言

在四旋翼无人机的自主导航系统中，轨迹规划与跟踪控制是实现高效、安全飞行的关键环节。轨迹规划负责生成满足动力学约束、避障要求及任务目标的可行路径，而跟踪控制则确保无人机能够准确、稳定地跟随规划轨迹。本章围绕四旋翼的轨迹规划与跟踪控制展开研究，构建了基于B样条优化的轨迹生成与避障算法，建立了基于PID和微分平坦的轨迹跟踪控制方法，并在基于Gazebo和ROS仿真环境验证了算法有效性。

## 基于B样条优化的轨迹规划方法

### 轨迹规划整体框架

本文采用的无人机轨迹规划方法框架为浙江大学Fast-LAB实验室开源框架EGO-Planner[]，如图3.1所示，图中阐述了本方法的整体流程。该方法根据目标点数量得到不同的初始全局轨迹，初始全局轨迹不考虑障碍物。当目标点唯一时，直接通过多项式直线连接起点和目标点，是一条满足终端约束的平滑轨迹。当有若干个目标点时，采取Minimum Snap轨迹生成算法生成初始全局轨迹，该全局轨迹同样也是忽略环境障碍物的。然后采取均匀B样条参数化拟合轨迹，对发生碰撞的局部轨迹重新规划。为保证实时性，该方法采取无ESDF的规划框架，只通过前端路径搜索方法和基于锚点的控制点重分配方法计算障碍物距离信息，从而计算避障代价函数。本方法的轨迹优化分为两阶段，一阶段考虑避障代价、平滑代价以及动力学可行性代价，第二阶段引入时间跨度约束，优化无人机速度。

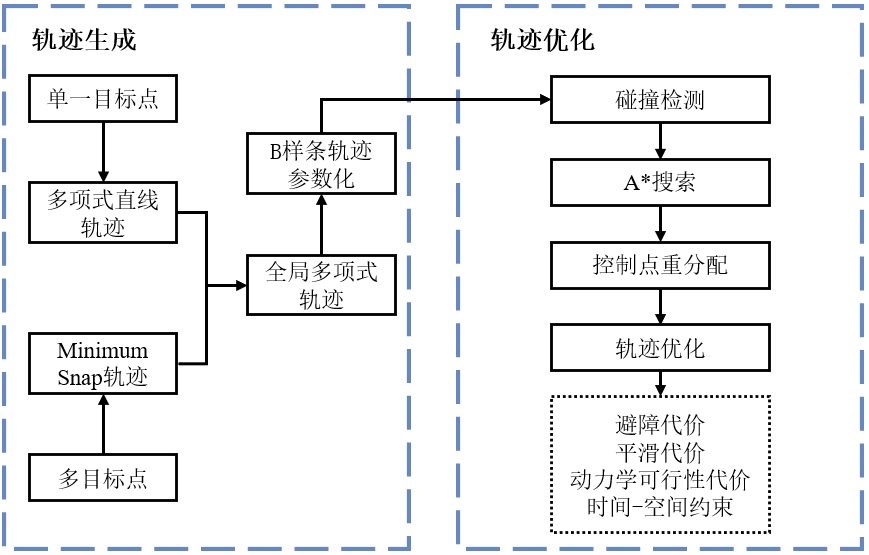


图3.1 基于B样条优化的轨迹规划方法示意图

### 轨迹参数化

在轨迹规划里，通常用一系列离散点表示路径，通过轨迹参数化，可以精确地描述物体的运动状态，如位置、速度、加速度等，从而实现对物体运动的精确控制。轨迹参数化不仅可以描述轨迹，还可以优化轨迹，使其满足某些特定的约束条件，如最短时间、最小能耗等。在无人机规划里面，需要实时计算和调整轨迹，轨迹参数化提供了这种实时性。B样条曲线是通过逼近一组控制点生成的曲线，通常曲线不通过其控制点，B样条曲线形状由它的阶数，个控制点和时间节点向量唯一确定。其中，，，并且长期成立。均匀B样条则是在B样条定义基础上，还拥有相同的时间跨度，时间跨度。一段B样条曲线，若其时间节点向量是均匀分布的，相邻时间节点的时间跨度恒等于，则该曲线为均匀B样条曲线。实际中，时间向量需要先进行归一化处理才能够表示时间上的位置：



然后无人机位置可以通过矩阵形式来表达：







其中，矩阵代表由B样条曲线阶数确定的矩阵，在本方法中，，即使用三阶均匀B样条来对无人机轨迹进行参数化表达。

本章采取均匀B样条参数化轨迹的原因主要是因为其三点重要性质。首先是均匀B样条曲线具有凸包性，对于阶均匀B样条曲线，凸包性指的是单个时间跨度的轨迹仅由个连续控制点所控制，并且该轨迹包含于这些控制点所形成的凸包内，如图3.2所示：

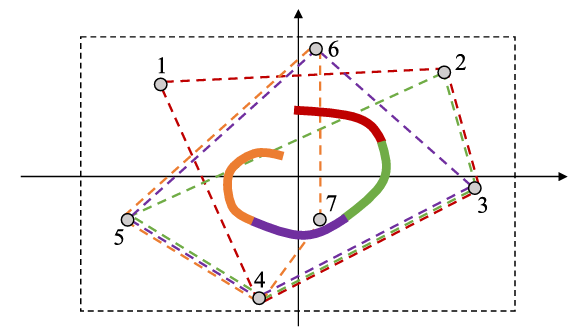


图3.2 均匀B样条凸包性质示意图[]

在图3.2中，轨迹由四段不同颜色的分段轨迹组成，且不同的分段轨迹与对应的四个控制点构成的凸包相对应，凸包用对应颜色的虚线表示。通过该图可以清楚看到，每段轨迹均位于对应四个控制点所形成的凸包内。该性质对于后端轨迹优化阶段十分关键，尤其是对于轨迹优化中轨迹安全约束的构建。第二个性质是B样条曲线的局部性，B样条基函数的局部支撑性决定了B样条的局部性。阶曲线上的一点至多与个控制点有关，与其它控制点无关。故移动曲线上第i个控制顶点，最多影响到定义在此点对应区间上的对应部分曲线的形状，对曲线的其余部分不发生影响，这点性质也可以从图3.2中看出来，分段轨迹只与对应凸包的四个顶点相关。第三个性质是阶B样条曲线的阶导数是阶B样条曲线，因此以B样条曲线表征无人机轨迹时，对轨迹求微分得到的是低阶B样条曲线，可以参数化轨迹速度曲线、加速度曲线、加加速度曲线等等，简化后续对轨迹平滑性约束项的建立，通过控制点可以得到对应的速度、加速度、加加速度的表达式：



因此，在得到全局原始轨迹后，通过均匀B样条参数化轨迹表达，可以简化后续轨迹规划问题。

### 前端路径搜索与控制点重分配

本系统使用三维栅格地图作为局部地图，其将无人机运动空间划分为规则的网格，这使得基于栅格的计算和数据处理变得简单和高效。未知环境中，由于没有先验地图，无人机自主导航通常只能感知到局部环境，维护的栅格地图规模较小，基于采样的搜索算法，如RRT、RRT\*等由于其采样策略的随机性容易出现搜索效率低的情况，基于采样的算法更适合大规模和较低障碍物密度的情景，因此本系统采取基于图搜索的寻路算法实现前端路径搜索，得到初步满足避障约束的离散路径点。

A\*算法是目前基于图搜索算法中使用最为广泛的寻路算法，特别是在无ESDF的规划框架中，其应用十分广泛，具体算法伪代码如表2所示。A\*通过评估已知和预测的成本来选择路径，A\*算法首先创建两个列表，开放列表(Open List)和封闭列表(Closed List)。开放列表用于存储待评估的节点，封闭列表用于存储已经评估过的节点。初始时，只将起始节点放入开放列表。之后循环检查开放列表，在开放列表中查找具有最低分数的节点作为当前节点。分数是分数和分数之和：，其中分数是从起点到当前节点的实际成本，分数是当前节点到终点的预估成本，也称启发函数。

A\*寻路算法可以搜索得到一条无碰撞的离散路径点。若直接将这些路径点作为均匀B样条的控制点，这样得到的一段轨迹无法与上一段轨迹连接，因此前端路径点无法直接作为控制点，因此需要重新分配控制点。

表3.1 A\*寻路算法

|  |
| --- |
| 算法2：A\*寻路算法  初始化open\_list和close\_list； |
| 将起点加入open\_list中，并设置优先级为0（优先级最高）； |
| 如果open\_list不为空，则从open\_list中选取优先级最高的节点n： |
| 如果节点n为终点，则： |
| 从终点开始逐步追踪parent节点，一直达到起点； |
| 返回找到的结果路径，算法结束； |
| 如果节点n不是终点，则： |
| 将节点n从open\_list中删除，并加入close\_list中； |
| 遍历节点n所有的邻近节点： |
| 如果邻近节点m在close\_list中，则： |
| 跳过，选取下一个邻近节点 |
| 如果邻近节点m也不在open\_list中，则： |
| 设置节点m的parent为节点n |
| 计算节点m的优先级 |
| 将节点m加入open\_list中 |

鉴于避障功能仅在特定条件下激活，即当检测到无人机轨迹可能与障碍物发生碰撞时，才需启动轨迹优化机制，本文采用基于障碍物表面锚点的梯度计算方法来获取必要的避障梯度信息。轨迹碰撞检测通过分析轨迹控制点在栅格地图中的占据状态实现。具体流程包括：遍历当前轨迹的所有控制点，将其坐标转换至地图坐标系，并查询该坐标点是否处于占据状态，从而判定该控制点是否发生碰撞。通常，若一段轨迹发生碰撞，将对应一个控制点进入障碍物区域，另一控制点离开障碍物区域。而在初始轨迹中，可能存在多个碰撞段，因而会形成多对进出不同障碍物的控制点。如图3.3所示，针对发生碰撞的轨迹段，选取进入障碍物的控制点作为起点，离开障碍物的控制点作为终点，在此两点间进行A\*搜索。绿色点位为A\*算法搜索得到的无碰撞离散路径点；圆柱障碍物内部的蓝色线段表示当前发生碰撞的无人机轨迹；轨迹上的红色点代表位于障碍物内部的原始控制点。根据B样条曲线的凸包性质，由这些红色控制点所定义的轨迹段必然完全位于该障碍物内部。

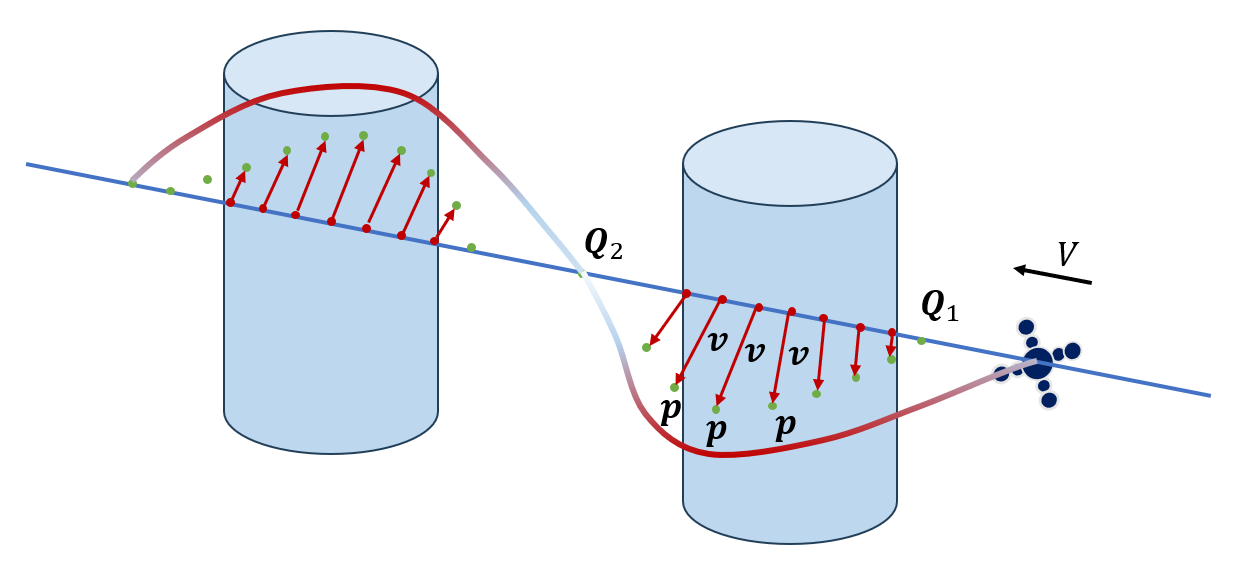


图3.3 障碍物内的控制点重分配示意图

对于、之间处于障碍物内部的控制点，需要重新配置，将内部的碰撞控制点通过优化的方式推离到障碍物外面，因此需要得到控制点与障碍物的距离信息来计算避障梯度。锚点是连接路径点和障碍物内部控制点的线段与障碍物表面的交点。根据障碍物内部控制点的数量，可以将前端搜索得到的离散路径点均匀分为份，则对应有个节点，从对应的节点以栅格地图分辨率迭代回退，直到障碍物表面，此时障碍物上的点为锚点，故发生碰撞的控制点可以得到对应的锚点。如图3.3所示，通过为障碍物内部发生碰撞的控制点重新分配锚点，然后生成从该控制点到锚点的方向向量，从而可以计算避障优化使用的距离信息：



式中，表示第个控制点到障碍物的距离，表示障碍物内部的第个控制点，表示对应的锚点，表示从到的单位方向向量。与障碍物的距离信息后续用来设计碰撞代价函数。

### 基于梯度的后端轨迹优化

针对传感器感知范围内的轨迹优化，本系统设计了一种两阶段轨迹优化方案，第一阶段轨迹优化旨在得到无碰撞且连续光滑的控制点集，第二阶段引入时间跨度约束取代常规的时间重分配优化无人机速度。基于锚点法得到轨迹控制点与环境障碍物距离信息后，根据无人机的微分平坦特性降低待规划的变量，构建轨迹优化问题形式如下：



其中，表示平滑损失函数，表示碰撞损失函数，表示可行性损失函数。，，为各个损失函数对应的权重。

平滑损失函数最小化轨迹控制点的加速度和加加速度，最小化轨迹的高阶导数能够使整段轨迹光滑，所以平滑损失函数由加速度和加加速度的平方和构成。由于均匀B样条曲线的凸包性质，只要最小化轨迹的二阶和三阶控制点的平方和就能够有效地减小加速度和加加速度的平方和：



碰撞损失函数通过采用安全间距将控制点推离障碍物。本系统构建了一个二次连续可微的惩罚函数，并在减小时抑制其斜率，从而得到分段函数如式3.9所示。







在第二章中已证明四旋翼无人机的动力学是微分平坦的，所以本系统可以通过限制B样条轨迹每一个维度上的高阶导数来确保动力学可行性：



式中，，和分别是各高阶损失项的权重系数，为在轨迹每一维度上的损失函数之和：



其中，，函数定义如下：



其中，，，，，和为保证二次多项式连续性设置的系数，是自定义的各项动力学上限，是二次区间与三次区间之间的分界点。

在第一阶段优化之前为初始轨迹分配精确的时间曲线是不合理的，因为此时规划器尚未掌握最终轨迹的任何信息。因此，为确保动力学可行性，额外的时间重分配过程至关重要。在第二阶段轨迹优化中，本系统引入了时间跨度约束项，充分考虑无人机的动力学限制，实现更安全稳定的飞行。由于两个控制点之间距离相同的均匀B样条曲线，扩展时间跨度不会改变B样条曲线的空间形状，但会影响整个轨迹的速度和加速度。因此，在第一次优化确定了曲线空间位置之后，继续优化轨迹，以保证的高阶导数在设定的动力学极限以下，而不必担心由于第二次优化而产生的碰撞。

首先检查轨迹各控制点处的速度 ​、加速度 和加加速度 是否超过最大允许值 、和。然后计算一个违背比例系数 ：



其中，，，；代表坐标轴分量。该系数量化了当前轨迹超出动力学限制最严重的程度。根据违背比例系数，对整条轨迹的时间进行均匀拉伸；通过增大时间间隔，可以直接降低轨迹各阶导数的幅值，从而保证新的时间分配能够满足动力学约束。新的时间间隔  为



## 基于PID的轨迹跟踪控制方法

四旋翼无人机是一个典型的欠驱动、强耦合的非线性系统，这为其高精度控制带来了挑战。在众多控制方法中，PID（比例-积分-微分）控制器凭借其结构简洁、参数物理意义明确、无需依赖精确的系统模型，以及在各种工程实践中被反复验证的卓越鲁棒性，至今仍在四旋翼无人机，尤其是对实时性要求极高的在线轨迹跟踪场景中，占据着不可替代的地位。尽管现代控制理论催生了诸如滑模控制、模型预测控制、反步控制等先进算法，并在特定条件下展现出优越性能，但其复杂的计算负担与对模型精确性的高度依赖，在一定程度上限制了其在计算资源紧张的嵌入式平台上的实时应用。

鉴于此，本系统立足于工程实践的可靠性与实时性需求，提出并实现了一种基于串级PID控制器的四旋翼无人机轨迹跟踪控制策略。首先，从牛顿-欧拉方程出发，建立了完整四旋翼动力学模型。随后，针对轨迹跟踪任务的内外环特性设计了串级控制结构：外环位置控制器负责解算期望姿态角，内环姿态控制器则精准跟踪该指令，以此实现对欠驱动系统的有效控制。

在第二章中已证明四旋翼无人机的动力学是微分平坦的，四旋翼的状态量可以由微分平坦输出以及其高阶导数表示。依据时标分离原理[]，本文将四旋翼无人机的轨迹跟踪控制问题解耦为位置控制与姿态控制两个子系统。首先，基于牛顿-欧拉方程建立了包含质心平动和绕质心转动的完整动力学模型，具体形式如式3.1、式3.2所示。





其中，表示四旋翼从机体坐标系到世界坐标系的转换矩阵；为四旋翼螺旋桨到质心的距离即半轴距；和为各螺旋桨产生的拉力和反扭矩大小；。

在位置控制子系统中，根据上层规划得到的期望轨迹与偏航角，设计了一个PD控制器来计算控制量以及四旋翼的期望滚转角和俯仰角，，以确保四旋翼跟踪期望轨迹，由此，可得到期望姿态角。具体控制率形式如式3.20~式3.23所示。









其中，表示期望加速度控制量；表示比例增益系数；表示微分增益系数。

在姿态控制子系统中，设计了一个PID控制器以获得控制量、和，从而确保四旋翼在外部扰动下跟踪期望的姿态角，积分项有助于保证系统的稳定性以及在扰动下实现零误差。具体控制率形式如式3.24~式3.25所示。





其中，，，为期望角加速度控制量；为滚转角比例增益系数，为滚转角微分增益系数，为滚转角积分增益系数，并且类似的定义了俯仰角和偏航角的比例、微分、积分增益系数。整体控制策略如图3.4所示。

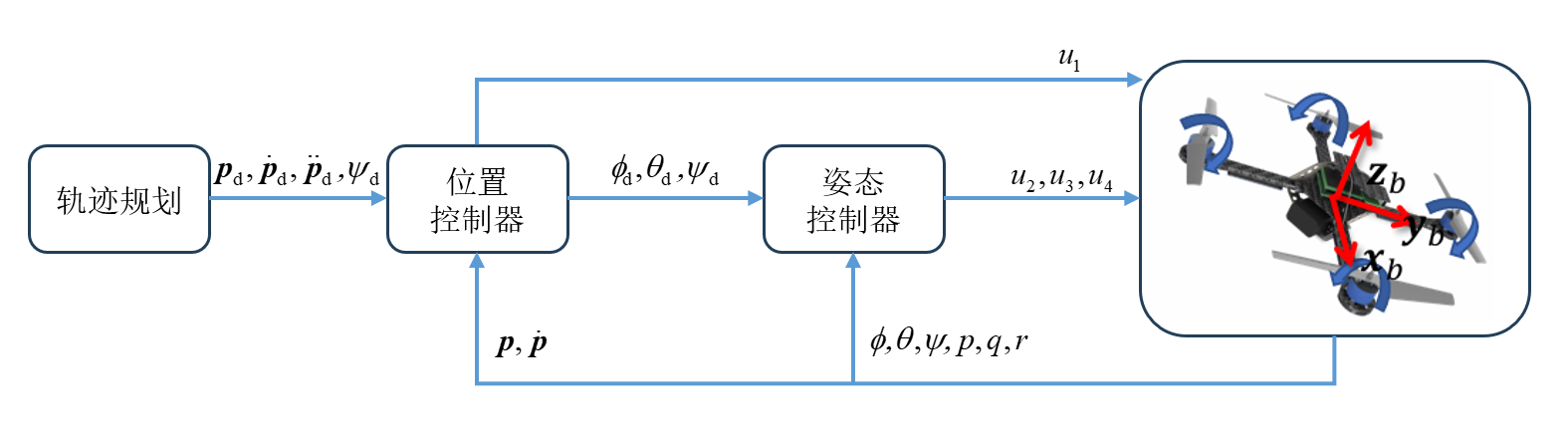


图3.4 PID轨迹跟踪控制结构图

## 仿真实验与验证

### Octave轨迹跟踪数值仿真

为验证本文所提出的PID控制器在轨迹跟踪控制的有效性和鲁棒性，本章搭建了基于Octave的数值仿真环境，并设计了多种典型场景轨迹下的跟踪效果分析实验。本文所建立的动力学模型（式3.17~式3.18）以及PID控制器（式3.19~式3.24）均可在Octave中实现。仿真采用固定步长0.002s的ODE45（龙格-库塔）求解器，以保证计算的精度与实时性。四旋翼无人机模型的关键物理参数如表3.2所示。

表3.2 四旋翼关键物理参数

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 参数 | 质量 | 半轴距 | 转动惯量  （X轴） | 转动惯量  （Y轴） | 转动惯量  （Z轴） |
| 符号/单位 |  |  |  |  |  |
| 数值 | 1.5 | 0.175 | 0.0347563 | 0.0458929 | 0.0977 |

为全面评估控制器性能，本文设定了两条具有代表性的参考轨迹进行测试：

（1）螺旋上升轨迹

该轨迹用于检验控制器在平稳、连续运动下的跟踪精度与稳态性能。同时为了验证轨迹跟踪全程控制器的鲁棒性，对无人机动力学模型施加了幅值为的三向扰动。轨迹方程如下式所示。



螺旋上升轨迹跟踪整体效果与各轴位置跟踪误差如图3.5所示，真实轨迹与期望轨迹基本重合，可以看出四旋翼无人机能够稳定的跟踪期望轨迹，具备足够的稳态精度。并且在初始时刻存在阶跃位置误差时，能够快速稳定的收敛到期望轨迹。

|  |  |
| --- | --- |
|  |  |
| （a）整体轨迹示意图 | （b）x轴位置跟踪 |
|  |  |
| （c）y轴位置跟踪 | （d）z轴位置跟踪 |

图3.5 四旋翼螺旋上升轨迹跟踪位置曲线

（2）速度阶跃轨迹

该轨迹在X和Y方向均包含幅值为的速度阶跃信号，同时与螺旋上升轨迹相同，对无人机动力学模型施加了幅值为的三向扰动。该轨迹用于考核控制器在参考指令突变时的动态响应速度、超调量与抗干扰能力。

速度阶跃轨迹跟踪整体效果与各轴位置跟踪误差如图 3.6所示，真实轨迹与期望轨迹比较重合，在期望轨迹速度发生突变时，四旋翼跟踪轨迹虽然有一定的超调量，但所提出的PID控制器仍具有足够的跟踪稳定性。

|  |  |
| --- | --- |
|  |  |
| （a）整体轨迹示意图 | （b）x轴位置跟踪 |
|  |  |
| （c）y轴位置跟踪 | （d）z轴位置跟踪 |

图 3.6 四旋翼速度阶跃轨迹跟踪位置曲线

### 轨迹规划与跟踪控制全流程软件在环仿真

为验证本文所提出的基于B样条优化的轨迹规划方法，并进一步检验PID轨迹跟踪控制器的有效性，本章基于2.5节搭建的ROS仿真环境架构进行四旋翼的轨迹规划与跟踪控制全流程软件在环仿真实验。

本章基于开源物理引擎Gazebo搭建了地图大小为40m×40m×5m的四旋翼无人机飞行仿真环境，其中放置了多种大小不同圆柱体与长方体以模仿四旋翼在室内飞行过程中可能遇到的各种障碍物情况。具体场景如图 3.7所示。四旋翼无人机物理参数与上一小节数值仿真中表3.2的参数一致。

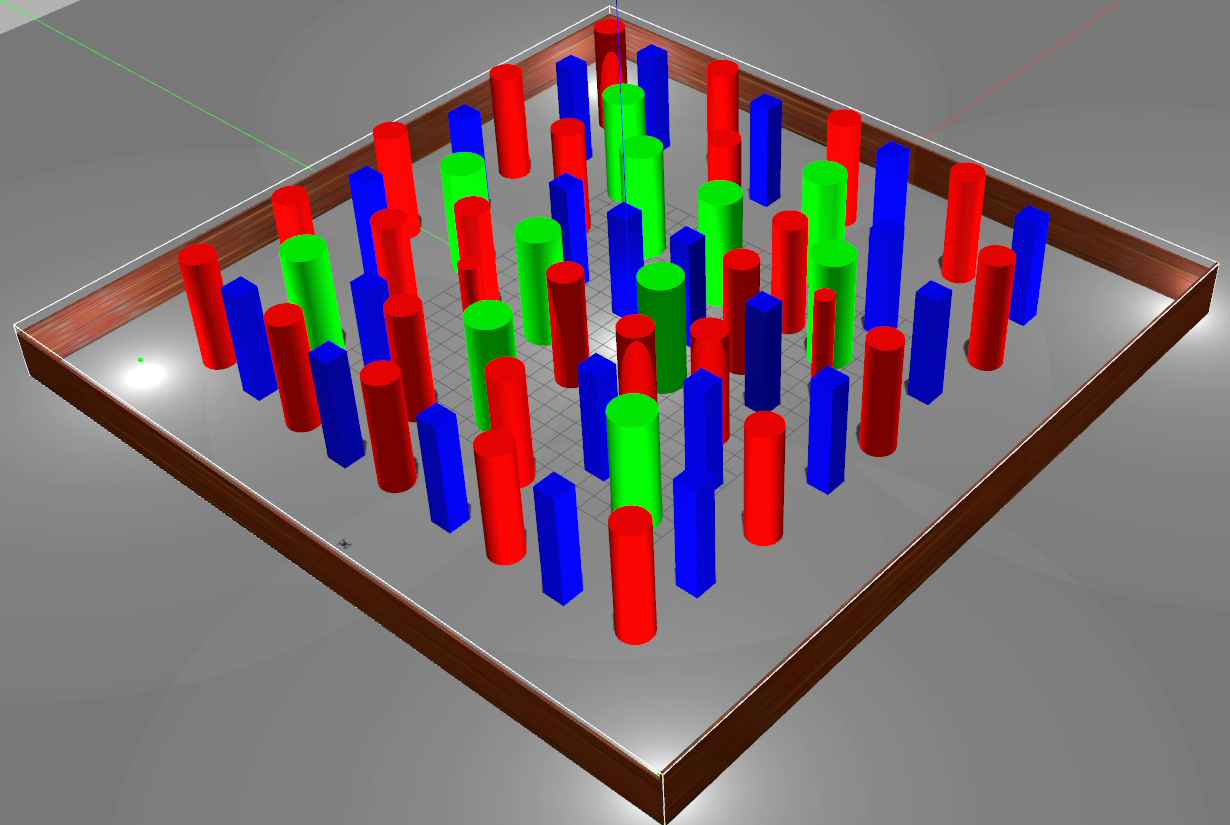


图 3.7 四旋翼轨迹规划与跟踪控制软件在环仿真场景

设定仿真中四旋翼自主导航避障任务：起始位置为，自主起飞至，依次飞至设置的三个途经点，，，并最后返回降落在起始位置。

碰撞代价项权重0.5，光滑代价项权重，动力学可行性代价项权重0.1，在预先设置的动力学限制下，最大速度，1以及最大加速度2。避障安全距离为0.5m。

## 本章小结

# 基于IMU输入的神经网络四旋翼短时位姿预测

## 引言

在四旋翼快速运动或环境视觉纹理不足的情况下，视觉特征的缺失和失效会使得VINS退化为仅靠IMU积分获取无人机位姿，而对于低成本的MEMS惯性器件，其误差具有强非线性和高时变的特点，仅依靠双惯性积分会迅速积累误差，对VIO的定位精度造成极大的影响。本章对视觉短时失效时四旋翼不同飞行工况下的VIO定位结果进行了分析讨论，并尝试使用深度学习框架来解决IMU长时间序列数据的建模问题，以提高VIO在视觉短时失效时的可观测性和无人机定位精度。

## 视觉短时失效下视觉惯性融合定位结果分析

基于前文搭建的ROS仿真环境，设计在四旋翼无人机悬停和盘旋飞行两种典型工况下注入视觉失效故障，测试分析无人机VIO在没有视觉图像来源时的定位结果。

（1）悬停工况

（2）盘旋飞行工况

## 仿真数据集及神经网络模型设计

### LSTM网络数学模型

深度神经网络是一类具有多个隐藏层、能够自动学习层次化特征表示的神经网络模型，在理论上，具有比传统的浅层神经网络更强的特征学习能力和非线性处理能力。深度神经网络使得模型具备扩展到大输入的处理能力，发展出了面向长时间序列建模的深度循环网络（Recurrent Neural Network，RNN）。惯性测量单元的数据是反映载体运动的时间序列数据，数据中的误差也是跟运动、环境等因素有关的时间序列，因此，使用深度循环神经网络符合IMU位置预测建模的问题特征。

循环神经网络是一类具有记忆机制的用于处理序列问题的神经网络，能够将隐藏层前一时刻的输出作为当前时刻的输入信息[40][41]。理论上RNN能够记忆任意长度的序列信息，然而实际上因为训练过程中权值矩阵的重复使用导致的梯度消失与梯度爆炸问题，RNN只能记忆很短的序列信息，实际应用中，最有效的序列模型是门控RNN（Gated RNN），长短时记忆网络是一类主要的门控RNN[42]。

LSTM的结构由多个重复的神经网络单元构成[43]，如图4.1所示，每个单元内部都具有自循环结构，分别为t时刻的输入信息、隐藏层输出和记忆单元状态。

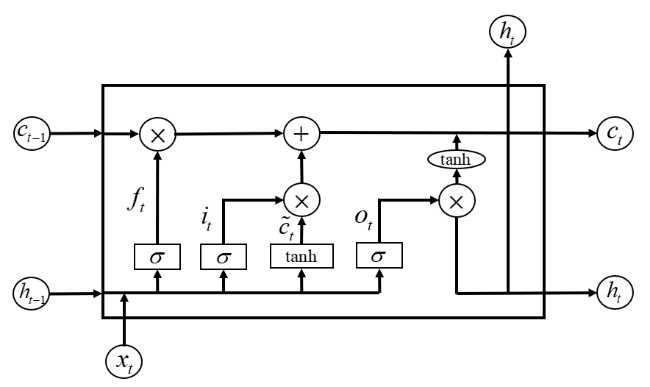


图4.1 LSTM神经网络单元

每个LSTM单元都有控制信息流动门控单元系统，通过“输入门”、“遗忘门”、“输出门”来控制细胞状态的更新和丢失，以提取输入信息中的有效信息并抛弃无效信息，保护输入信息中的时序性。输入门、输出门和遗忘门公式分别如式(1)、(2)、(3)所示。







其中：、、分别是时刻输入门输出、输出门输出、遗忘门输出；为激活函数；、、为偏置值；是时刻的隐藏层输出；其余为相应的权值系数矩阵。

输入门、遗忘门和上一时刻的记忆状态决定了当前时刻记忆状态，公式如式(4)所示。



在输出门的作用下得到最终的输出值，公式如式(5)所示。



LSTM网络中的状态和参数及循环结构为IMU位姿预测的时间序列建模奠定了网络结构与参数记忆、特征表示的基础。

### LSTM网络模型结构设计

LSTM神经网络的结构取决于待学习对象的复杂度，对LSTM网络输入、输出层进行设计，使其能对线运动、角运动进行表征。本文所设计的IMU位姿预测LSTM深度神经网络结构如图4.2所示。网络的输入信息包括：

(1) 三轴加速度计的三维加速度序列

(2) 上一时间序列的位置

(3) 三轴陀螺仪的三维角速度序列

(4) 上一时间序列的姿态

IMU位姿预测深度神经网络输出包括：

(1) 当前时间序列的位置

(2) 当前时间序列的姿态

其中，加速度量纲为，角速度量纲为，位置量纲为，角度用四元数表示。

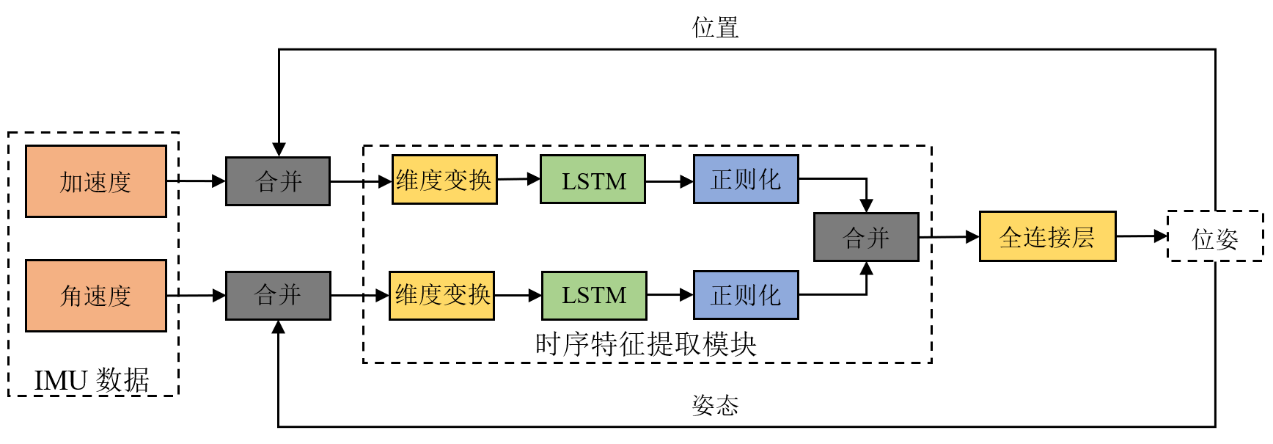


图4.2 IMU位姿预测LSTM深度神经网络结构

LSTM网络的训练算法是反向传播算法，主要有5个步骤：①前向计算每个神经元的输出值； ②计算网络输出的误差；③反向计算每个神经元的误差值（沿时间的反向传播，即从当前t时刻开始，计算每个时刻的误差项；将误差项向上一层传播）；④计算每个权重的梯度；⑤权值更新

通过梯度下降法对参数进行迭代更新的关键之处就在于对基于损失函数的所有参数的偏导数的计算。本文设计网络总损失函数由位置损失函数、姿态损失函数和速度损失函数以及各部分的对应权重参数组成，具体如式(6)所示：



其中权重参数设置为，，。此权重配置主要基于以下两方面的考虑与实验依据：

（1）任务重要性导向：在无人机状态估计中，精确的位置和姿态估计是位姿预测最核心的输出，直接关系到导航的成败，因此赋予它们最高的基础权重。而速度损失在此框架中作为一个物理约束正则项，其目的在于提升轨迹的平滑性与动力学合理性，而非提供一个独立的速度观测。因此，将其权重数值至于较低水平。

（2）消融实验验证：本章进行了消融实验，在位置损失和姿态损失权重保持不变的情况下，尝试提升速度积分损失的权重，对比了不同权重组合在验证集上的性能。实验结果表明，当前采用的(1, 1, 0.5)权重配置能够使模型在位置精度和姿态精度上取得最佳的综合性能。过度调整权重（如显著增大某一项的权重）会导致模型优化失衡，反而使整体性能下降。这从实证角度验证了当前权重配置的有效性。

为了防止数据中的异常值对网络训练的影响，本研究使用Huber损失函数构建了如式(7)所示的位置损失函数：



其中，为位置真值，为神经网络的预测位置输出，为可选参数。与位置误差相反，姿态误差是一个几何量，常见欧拉角表示形式因存在奇点导致不稳定，旋转矩阵计算量大，且在无人机系统中姿态通常以四元数形式传递，因此本文使用四元数来描述姿态误差。如果两个四元数之间的相对旋转角度为0，则四元数的内积等于1。因此，姿态损失函数如式(8)所示：



其中，表示姿态四元数的真实值，表示神经网络的预测姿态四元数输出。由于并不总是提供真实速度值，为了提高本文提出的方法的适用性，本文使用位移差分的速度计算速度损失，如式(9)所示：



其中，为位置真值，为神经网络的预测速度输出。

## IMU位姿预测神经网络模型训练

### 数据集构建

本文数据集通过采集开源物理引擎Gazebo中自主飞行四旋翼的IMU数据和位姿真值数据产生。四旋翼选用3D Robotics开源的Iris构型，设计四旋翼飞行轨迹为伯努利双扭线，其表达式如式(10)所示：



其中，取。每隔记录一段的IMU和位姿真值数据，如图4.3所示。数据集具体参数如表4.1所示。随机选取60%的数据集作为训练集，其余40%为测试集。

表4.1 数据集参数

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 编号 | 单圈飞行时间T/(s) | 无人机最小速度/(m/s) | 无人机最大速度/(m/s) | IMU高斯噪声 | 段数 | 时间/(s) |
| 01 | 25 | 0.5 | 0.9 | 0.0001 | 20 | 10 |
| 02 | 25 | 0.5 | 0.9 | 0.001 | 20 | 10 |
| 03 | 25 | 0.5 | 0.9 | 0.01 | 20 | 10 |
| 04 | 20 | 0.7 | 1.2 | 0.0001 | 20 | 10 |
| 05 | 20 | 0.7 | 1.2 | 0.001 | 20 | 10 |
| 06 | 20 | 0.7 | 1.2 | 0.01 | 20 | 10 |
| 07 | 15 | 0.9 | 1.7 | 0.0001 | 20 | 10 |
| 08 | 15 | 0.9 | 1.7 | 0.001 | 20 | 10 |
| 09 | 15 | 0.9 | 1.7 | 0.01 | 20 | 10 |

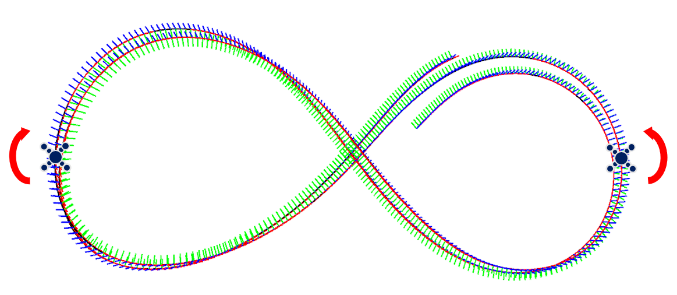


图4.3 伯努利双扭线飞行轨迹

### LSTM神经网络建模

基于Pytorch提供的LSTM神经网络框架，编程实现所需深度神经网络模型。为使建立的模型最优,在神经网络建模过程中，需根据数据的情况选择不同的参数。

（1）网络层数及每层神经元个数。网络共包含8层：1个输入层，4个隐藏层，2个并行输出层。其中，输入层包含13个神经元，隐藏层每层各含32个神经元，每个输出层含64个神经元。

（2）激活函数。为避免梯度消失/爆炸，激活函数选为Relu。

（3）优化器。图4.4为采用Adam优化算法和随机梯度下降算法（SGD）训练时，在训练集上模型误差（）随迭代次数（N）变化曲线。由于Adam 的收敛速度及模型误差远远好于SGD优化器,因此训练时选择Adam优化器。

|  |  |
| --- | --- |
|  |  |
| (a) Adam优化算法 | (b) 随机梯度下降算法 |

图4.4 模型误差随迭代次数变化曲线

设置模型损失函数达到10e-4结束迭代。

## IMU位姿预测神经网络模型验证

为验证模型的有效性，考虑到IMU数据的时变特性和噪声扰动特性，按照表1中数据集相同的配置，重新采集四旋翼自主飞行的IMU和位姿真值数据，将新采集的数据作为测试集合，验证LSTM模型对不同数据的适应性。此外，为了验证模型的实时预测性能，本文采用机器人操作系统（Robot Operating System, ROS）和Gazebo联合仿真，验证四旋翼以LSTM模型预测位姿作为定位数据时的自主导航的可靠性和鲁棒性。

### LSTM预测结果与分析

在新数据集下，LSTM深度神经网络的预测位姿与真实值之间的误差曲线如图4.5所示。其中，Avg表示平均值，单位为m；均方根误差（Root Mean Square Error，RMSE）能反映模型在较大偏差时的表现，公式如下所示：



其中，表示位姿数据的个数，表示位姿真值，表示模型预测的位姿值。

结果表明，随着数据集中IMU高斯噪声的增大，LSTM深度神经网络模型预测的位姿误差平均值增大。IMU高斯噪声越大，导致在输入网络模型训练时的扰动越大，网络模型学习到IMU数据与位姿之间关系的不确定性就越大，因此训练出网络模型的预测效果会下降。

在相同IMU高斯噪声配置下，随着四旋翼飞行速度增大，模型预测的位姿误差平均值减小。在本文数据集中，四旋翼绕着伯努利双扭线轨迹飞行，根据轨迹特性可知：飞行速度越大，飞行过程中的最大加速度和最大角速度均有所增加，此时IMU信噪比减小，网络模型学习到IMU数据与位姿之间关系的不确定性减小，因此训练出网络模型的预测效果会上升。

|  |  |
| --- | --- |
|  |  |
| (a) 高斯噪声= 0.0001 | (b) 高斯噪声= 0.001 |
|  | |
| (c) 高斯噪声= 0.01 | |

图4.5 不同IMU高斯噪声下模型预测位姿误差曲线

### 软件在环仿真

基于ROS-Gazebo联合仿真框架，编写神经网络节点接收IMU输入数据，可得到四旋翼实时位姿预测输出和位姿地面真值对照误差，如图4.6所示。

由图4.6可见，轨迹预测平均误差始终保持在0.25m以下，并且随着IMU高斯噪声增大，预测轨迹误差增大；随着四旋翼飞行速度增大，预测轨迹误差减小，与数据集验证结果一致。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (a) 高斯噪声= 0.0001, T = 25s | (b) 高斯噪声= 0.0001, T = 20s | (c) 高斯噪声= 0.0001, T = 15s |
|  |  |  |
| (d) 高斯噪声= 0.001, T = 25s | (e) 高斯噪声= 0.001, T = 20s | (f) 高斯噪声= 0.001, T = 15s |
|  |  |  |
| (g) 高斯噪声= 0.01, T = 25s | (h) 高斯噪声= 0.01, T = 20s | (i) 高斯噪声= 0.01, T = 15s |

图4.6 不同IMU高斯噪声和飞行周期下模型预测位姿误差仿真曲线

|  |  |
| --- | --- |
|  |  |
| (a) VIO位姿轨迹误差，高斯噪声= 0.0001 | (b) 模型预测位姿轨迹误差，高斯噪声= 0.0001 |
|  |  |
| (c) VIO位姿轨迹误差，高斯噪声= 0.001 | (d) 模型预测位姿轨迹误差，高斯噪声= 0.001 |
|  |  |
| (e) VIO位姿轨迹误差，高斯噪声= 0.01 | (f) 模型预测位姿轨迹误差，高斯噪声= 0.01 |

图4.7 视觉短时失效下模型预测位姿自主导航轨迹图

最后，为验证所提出模型实时预测的位姿是否能够支持四旋翼无人机的自主导航，本文设计了在四旋翼飞行过程中注入视觉失效故障的仿真试验。在试验中，当视觉传感器发生模拟失效时，神经网络模型预测节点介入提供四旋翼导航所需的位姿数据。仿真采用开源视觉惯性里程计框架VINS-Fusion，四旋翼飞行周期为25s，对应飞行速度区间为，IMU噪声配置分别为0.0001，0.001和0.01，以模拟不同传感器性能条件下的系统表现，仿真结果如图4.7所示。

由图4.7可见，在不同IMU噪声配置的飞行仿真中，视觉失效故障注入后，视觉惯性里程计VINS-Fusion提供的位姿很快趋于发散而无法用于四旋翼自主导航；LSTM深度神经网络模型提供的实时预测位姿尽管与真实轨迹存在一定偏差，其平均误差，均方根误差，仍保持在可接受范围内，且预测轨迹趋势与真实轨迹相符，使得四旋翼能够稳定跟踪既定的伯努利双扭线轨迹，并完成全周期飞行。

综合分析可知,针对室内四旋翼无人机设计的LSTM深度神经网络位姿预测方法能够根据IMU数据输入生成实时位姿信息，在不同噪声配置和飞行状态中误差始终保持在合理范围内，验证了所设计位姿预测模型的有效性。

## 本章小结

本章针对视觉惯性里程计在视觉失效时位姿估计失准的问题，采用一种基于IMU输入的LSTM深度神经网络位姿预测方法。在新数据集上进行网络性能测试，测试结果表明，所设计的网络模型在IMU高斯噪声较小，四旋翼飞行速度较快时的数据集上预测位姿误差较小。最后基于ROS和Gazebo的模块设计进行软件在环仿真，实现四旋翼在视觉失效下的短时自主定位与导航。仿真结果表明，所设计的位姿预测方法能够在短时间内预测生成连续可靠的位姿数据, 验证了此方法在视觉失效情形下用于四旋翼自主导航的可靠性和鲁棒性。

# 室内环境下四旋翼自主定位与导航实验

## 引言

现阶段的自主旋翼无人机多采用机载计算机作为系统的控制中心控制飞行器的形式，机载计算机可以控制四旋翼飞行器执行很多复杂的任务。本章节针对室内环境下四旋翼自主定位与导航的任务需求，对飞行器进行系统方案设计与主要硬件选型。实验采用机载计算机作为上位机、飞行控制处理器作为下位机的形式，针对室内环境下的几种典型工况，对本文所提出的四旋翼自主定位与导航功能进行验证。

## 四旋翼无人机硬件平台

如图5.1所示，本实验选取S500“X”型四轴机架，该机架对角线轴距为500mm，机身采用碳板加树脂的结构，保证结构强度的同时为机载计算机、双目视觉传感器和飞行控制处理器保留了空间。桨叶选取为9寸自锁桨叶，供电模块为5300mAh的锂电池，整机重量为3kg，能为无人机提供充足的动力和大约十分钟的续航。



图5.1 四旋翼无人机硬件平台

本实验选取英特尔实感D435i双目RGB摄像头如图5.2所示，与常规的RGB相机不同，D435i作为一款RGBD相机，是附带有深度信息的，在获取图像的同时，也能获取到图像中每个像素点到相机的距离。因此，D435i除了获取周围环境的图像信息之外，还能够实现障碍物感知以及三维地图重建等功能。



图5.2 英特尔RealSense D435i双目相机

室内未知环境下，无人机上的数据处理与计算任务须在无人机上进行，因为采取其他无线通信方式，均存在不确定性和高延时性，无法满足任务要求，因此无人机携带机载电脑实时处理传感器数据。本系统搭载的板载计算资源为英特尔13代NUC，如图5.3所示，其提供丰富的外部接口，可以实现直连双目传感器等设备的功能，拥有强大的算力，足以满足神经网络预测、位姿估计以及轨迹规划与跟踪控制的性能需求。



图5.3 英特尔13代NUC

本实验选用Pixhawk 2.4.8作为飞行控制处理器，如图5.4所示。其内置磁力计、加速度计、陀螺仪等微型传感器，其通过EKF算法融合处理这些传感器信息，可以得到无人机的姿态信息，并以此对无人机的姿态进行增稳控制。该飞控通过Micro-USB转接口与机载计算机相连，并通过MAVROS协议进行通信。



图5.4 Pixhawk 2.4.8飞行控制处理器

## 自主定位与导航系统架构

无人机机载电脑系统版本为Ubuntu20.04，软件算法实现基于机器人操作系统（Robot Operating System，ROS）Noetic版本，进程通过ROS中的节点话题通信机制来实现通信。系统软件框架示意图如图5.5所示，图中展示了系统的主要软件模块，箭头表示模块之间的消息和数据传递方向。首先D435i的驱动程序从相机获取原始RGB图像、原始深度图以及IMU数据，并发布在对应的ROS话题中。机载计算机上运行的VINS-Fusion程序通过订阅图像与IMU话题接收实时的传感器信息，发布无人机里程计信息，本文在第二章已经对该方法进行了详细的介绍。深度图像和里程计信息话题由栅格建图节点和轨迹规划节点订阅。建图节点接收到里程计信息和深度图像后，将深度图像转换到点云数据，启动地图更新步骤，发布栅格地图话题。原始深度信息是基于像素坐标的，要获取点云数据，首先需要使用相机内参将每个像素的深度值转换到相机坐标系中，然后利用相机在世界坐标系中的位置信息，将这些点从相机坐标系转换到世界坐标系。轨迹规划节点订阅到栅格地图话题和目标点后，通过第三章描述的EGO-Planner规划算法得到无人机实时轨迹，并将轨迹话题发布。接下来，ROS中的轨迹话题通过Mavros与无人机通信，Mavros允许ROS系统与兼容MAVLink协议的无人机进行通信，最终飞控得到飞行指令后解算得到各个电机的转速，控制无人机沿着期望轨迹飞行。在真机实验环节中，本文借助远程客户端NoMachine以SSH协议连接无人机上的机载电脑，便于在无人机上启动传感器驱动程序、发布目标点以及可视化调试工作。

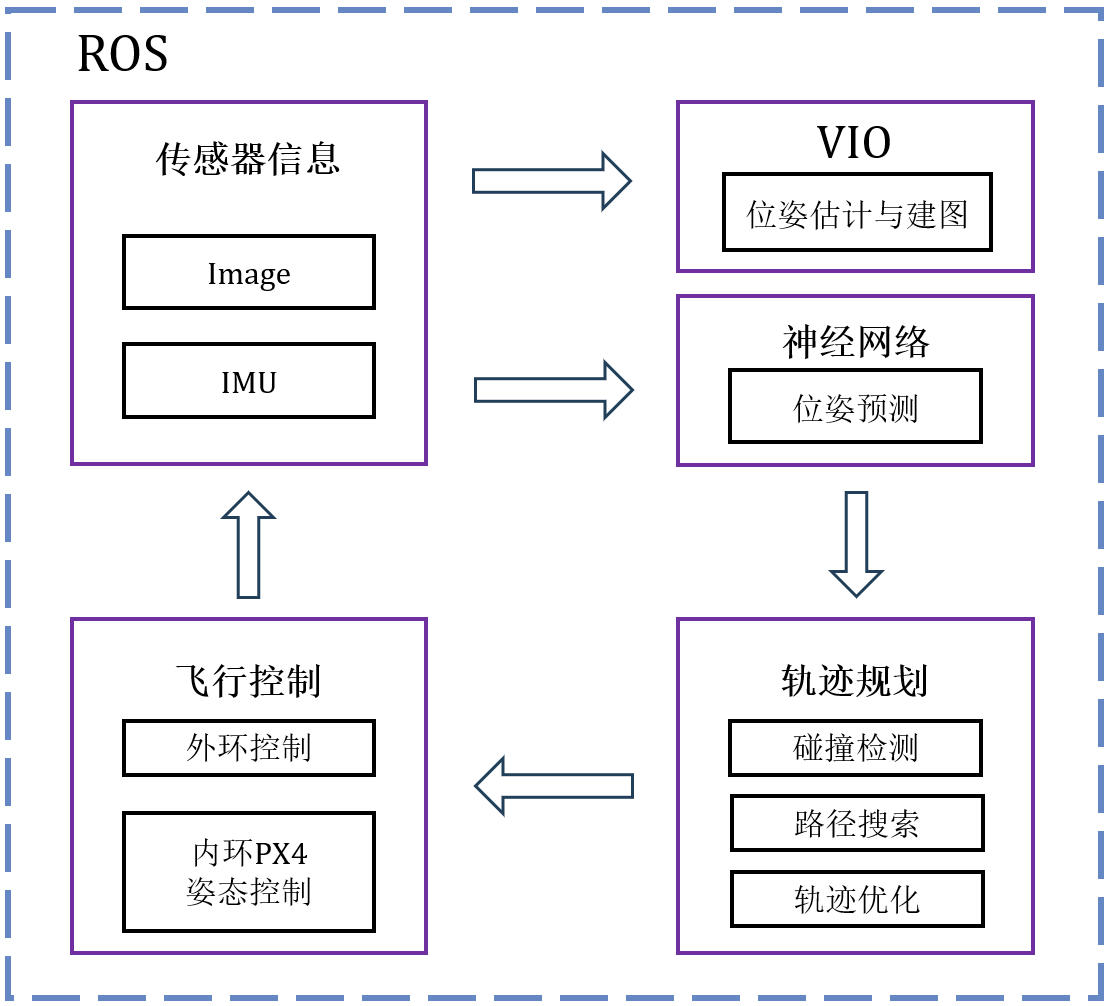


图5.5 无人机自主定位与导航系统软件框架

## 常规工况下四旋翼轨迹规划与跟踪控制实验验证

## 视觉短时失效下四旋翼轨迹规划与跟踪控制实验验证

## 本章小结

# 总结与展望

## 研究总结

xxx。

## 研究展望

xxx。

参考文献

1. Liu Z, Wang L, Li K, et al. An improved rotation scheme for dual-axis rotational inertial navigation system[J]. IEEE Sensors Journal, 2017, 17(13): 4189-4196.
2. 史津竹.车辆运动约束的视觉惯性位姿估计及可观测性分析[D]. 吉林大学, 2020.
3. R. E. Kalman. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering, 1960, 82(1): 35-45.
4. Bristeau P J, Dorveaux E, Vissière D, et al. Hardware and software architecture for state estimation on an experimental low-cost small-scaled helicopter[J]. Control Engineering Practice, 2010, 18(7): 733-746.
5. Mourikis A I, Roumeliotis S I. A multi-state constraint Kalman filter for vision-aided inertial navigation[C]//Proceedings 2007 IEEE international conference on robotics and automation. IEEE, 2007: 3565-3572.
6. Li M, Mourikis A I. Improving the accuracy of EKF-based visual-inertial odometry[C]//2012 IEEE International Conference on Robotics and Automation. IEEE, 2012: 828-835.
7. Sun K, Mohta K, Pfrommer B, et al. Robust stereo visual inertial odometry for fast autonomous flight[J]. IEEE Robotics and Automation Letters, 2018, 3(2): 965-972.
8. Geneva P, Eckenhoff K, Lee W, et al. OpenVINS: A research platform for visual-inertial estimation[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 4666-4672.
9. Martin P, Salaün E. Design and implementation of a low-cost observer-based attitude and heading reference system[J]. Control engineering practice, 2010, 18(7): 712-722.
10. Julier S J, Uhlmann J K. New extension of the Kalman filter to nonlinear systems[C]//Signal processing, sensor fusion, and target recognition VI. Spie, 1997, 3068: 182-193.
11. Van Der Merwe R, Doucet A, De Freitas N, et al. The unscented particle filter[J]. Advances in neural information processing systems, 2000, 13.
12. Younes G, Asmar D, Shammas E, et al. Keyframe-based monocular SLAM: design, survey, and future directions[J]. Robotics and Autonomous Systems, 2017, 98: 67-88.
13. Lynen S, Achtelik M W, Weiss S, et al. A robust and modular multi-sensor fusion approach applied to MAV navigation[C]//2013 IEEE/RSJ international conference on intelligent robots and systems. IEEE, 2013: 3923-3929.
14. Chiu H P, Zhou X S, Carlone L, et al. Constrained optimal selection for multi-sensor robot navigation using plug-and-play factor graphs[C]//2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014: 663-670.
15. Indelman V, Williams S, Kaess M, et al. Information fusion in navigation systems via factor graph based incremental smoothing[J]. Robotics and Autonomous Systems, 2013, 61(8): 721-738.
16. Qin T, Cao S, Pan J, et al. A general optimization-based framework for global pose estimation with multiple sensors[J]. arXiv preprint arXiv:1901.03642, 2019.
17. Dellaert F. Factor graphs: Exploiting structure in robotics[J]. Annual Review of Control, Robotics, and Autonomous Systems, 2021, 4(1): 141-166.
18. Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE transactions on robotics, 2015, 31(5): 1147-1163.
19. Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[C]//European conference on computer vision. Cham: Springer International Publishing, 2014: 834-849.
20. Lupton T, Sukkarieh S. Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions[J]. IEEE Transactions on Robotics, 2011, 28(1): 61-76.
21. Leutenegger S, Furgale P, Rabaud V, et al. Keyframe-based visual-inertial slam using nonlinear optimization[J]. Proceedings of Robotis Science and Systems (RSS) 2013, 2013.
22. Forster C, Carlone L, Dellaert F, et al. On-manifold preintegration for real-time visual--inertial odometry[J]. IEEE Transactions on Robotics, 2016, 33(1): 1-21.
23. Qin T, Li P, Shen S, et al. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator[J]. IEEE Transactions on Robotics, 2018, 34(4): 1004-1020.
24. Von Stumberg L, Usenko V, Cremers D. Direct sparse visual-inertial odometry using dynamic marginalization[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 2510-2517.
25. 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.
26. Elvira R, Tardós J D, Montiel J M M. ORBSLAM-Atlas: a robust and accurate multi-map system[C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019: 6253-6259.
27. Agrawal P, Nair A V, Abbeel P, et al. Learning to poke by poking: Experiential learning of intuitive physics[J]. Advances in neural information processing systems, 2016, 29.
28. Eigen D, Puhrsch C, Fergus R. Depth map prediction from a single image using a multi-scale deep network[J]. Advances in neural information processing systems, 2014, 27.
29. Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
30. Wu Z, Wu X, Zhang X, et al. Spatial correspondence with generative adversarial network: Learning depth from monocular videos[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 7494-7504.
31. Wu Z, Wu X, Zhang X, et al. Spatial correspondence with generative adversarial network: Learning depth from monocular videos[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 7494-7504.
32. Chen C, Rosa S, Miao Y, et al. Selective sensor fusion for neural visual-inertial odometry[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 10542-10551.
33. Li R, Yan J, Liu G, et al. MEMS-IMU error model and calibration method based on LSTM deep neural network[J]. Journal of Chinese Inertial Technology, 2020, 28(2): 165-171.
34. Clark R, Wang S, Wen H, et al. VINet: Visual-inertial odometry as a sequence-to-sequence learning problem[C]//Proceedings of the AAAI conference on artificial intelligence. 2017, 31(1).
35. Mahjourian R, Wicke M, Angelova A. Geometry-based next frame prediction from monocular video[C]//2017 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2017: 1700-1707.
36. Zhou T, Brown M, Snavely N, et al. Unsupervised learning of depth and ego-motion from video[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1851-1858.
37. Almalioglu Y, Santamaria-Navarro A, Morrell B, et al. Unsupervised deep persistent monocular visual odometry and depth estimation in extreme environments[C]//2021 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2021: 3534-3541.
38. Shamwell E J, Lindgren K, Leung S, et al. Unsupervised deep visual-inertial odometry with online error correction for RGB-D imagery[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 42(10): 2478-2493.
39. Almalioglu Y, Turan M, Saputra M R U, et al. SelfVIO: Self-supervised deep monocular visual–inertial odometry and depth estimation[J]. Neural Networks, 2022, 150: 119-136.
40. Jordan M I. Serial order: A parallel distributed processing approach[M]//Advances in psychology. North-Holland, 1997, 121: 471-495.
41. Elman J L. Finding structure in time[J]. Cognitive science, 1990, 14(2): 179-211.
42. Zaremba W, Sutskever I, Vinyals O. Recurrent Neural Network Regularization[C] // ICLR International Conference on Learning Representations (ICLR), San Diego: OpenReview.net, 2015(116): 1-8.
43. Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.

致谢

致谢宜以简短的文字对课题研究与论文撰写过程中曾直接给予帮助的人员（如指导教师、答疑教师及其他人员）表达自己的谢意。致谢不仅是一种礼貌，也是对他人劳动的尊重，是治学者应当遵循的学术规范，内容一般限一页。

在学期间的研究成果及发表的学术论文

**攻读硕士学位期间发表（录用）论文情况**

1.王天舒，史志伟\*.基于IMU输入的LSTM深度神经网络四旋翼短时位姿预测[J].（录用）

**攻读硕士学位期间申请专利情况**

【填写专利类型、专利名称、专利号、本人排名；软件著作权、著作权人（前三）、登记号、本人排名，分条目列举】