中国科学技术大学

硕士学位论文

融合人机信任的人-无人机协同方法研究

作者姓名： 黄康杰

学科专业： 控制科学与工程

导　　师： 赵云波 教授

完成时间： 二〇二六年五月二十七日

University of Science and Technology of China

A dissertation for master's degree

Research on Human-UAV Collaboration Methods Integrating Human-Machine Trust

Author: Huang Kangjie

Speciality: Control Science and Engineering

Supervisor: Prof. Zhao Yunbo

Completion date: May 27, 2026

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名：黄康杰　　签字日期：2026年5月27日

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定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，可以将学位论文编入《中国学位论文全文数据库》等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。本人提交的电子文档的内容和纸质论文的内容相一致。

控阅的学位论文在解除后也遵守此规定。

☑公开　□控阅（____年）

作者签名：黄康杰　　导师签名：

签字日期：2026年5月27日　　签字日期：2026年5月27日

摘要

人-无人机协同系统通过融合人类的认知决策优势与无人机的自主执行能力，已成为执行复杂任务的重要技术路径。合理的人机信任是实现高效人机协同的重要前提。然而，由于无人机具有非线性动力学特性，且人类能够获取的感知信息有限，人类往往难以准确评估无人机的实际能力，从而导致信任失衡。信任失衡不仅会削弱人机协同效能，还可能增加任务风险。因此，有必要量化人机信任水平并刻画其动态演化规律，进而设计融合信任的人-无人机协同方法。

然而，现有研究在人机信任建模与协同方法设计方面仍存在不足。在信任建模层面，尚缺乏与机器性能紧密关联的动态信任模型；在协同方法层面，现有方法尚未将信任状态有效纳入规划与控制过程，难以缓解信任失衡引发的协作冲突。为此，本文面向人-无人机协同任务，开展了人机信任动态建模研究，并分别针对降落与搜索任务设计了融合信任的协同方法。主要研究工作如下：

（1）针对现有信任模型可解释性不足且难以刻画机器性能波动对人类信任影响的问题，提出了机器性能驱动的信任演化模型，实现了对人机信任动态的建模与跟踪，为后续融合信任的协同方法设计提供了模型基础。该模型明确区分了机器客观能力与人类主观认知，刻画了人机信任演化过程，并通过分析核心参数的物理意义与相关性质提升了模型的可解释性。

（2）针对人机协同降落任务中缺乏合理的控制权分配机制的问题，提出了融合信任的最小干预共享控制策略，在尽可能保留人类控制意图的同时，提高了任务成功率与降落精度。该策略首先基于意图推理、降落精度与任务安全性指标量化机器的客观能力与性能表现，并基于深度强化学习进行动作价值评估与候选动作集构造；随后，将信任状态引入仲裁过程，动态调节候选动作的价值损失阈值，并遵循最小干预原则，在可行域内选取与人类控制输入最接近的动作作为控制输出。

（3）针对人机协同搜索任务中轨迹规划对信任动态变化考虑不足的问题，提出了融合信任的人机协同轨迹规划方法，在降低人类工作负荷的同时，提升了飞行轨迹的平滑性。该方法首先基于飞行安全性与可视性指标量化机器的客观能力与性能表现；随后，在安全约束下根据人类输入扩展运动基元树，构造候选轨迹集；再结合信任状态进行轨迹选择，并对所选轨迹进行优化，从而生成信任感知的自适应飞行轨迹。

关键词：人机信任；人机协同；共享控制；轨迹规划

ABSTRACT

Human-UAV collaborative systems, which integrate human cognitive and decision-making advantages with the autonomous execution capabilities of unmanned aerial vehicles (UAVs), have become an important approach to performing complex tasks. Appropriate human-machine trust is an important prerequisite for achieving efficient human-UAV collaboration. However, due to the nonlinear dynamics of UAVs and the limited environmental and state information available to humans, humans often find it difficult to accurately assess the actual capabilities of UAVs, which may lead to trust imbalance. Trust imbalance not only degrades the effectiveness of human-UAV collaboration, but may also increase task risks. Therefore, it is necessary to quantify the level of human-machine trust, characterize its dynamic evolution, and further design trust-integrated human-UAV collaborative methods.

However, existing studies still have limitations in human-machine trust modeling and collaborative method design. At the trust-modeling level, there is still a lack of dynamic trust models that are closely associated with machine performance. At the collaborative-method level, existing approaches have not effectively incorporated trust states into planning and control processes, making it difficult to mitigate collaborative conflicts caused by trust imbalance. To address these issues, this thesis investigates dynamic human-machine trust modeling for human-UAV collaborative tasks and further designs trust-integrated collaborative methods for landing and search tasks. The main contributions are as follows:

(1) To address the limited interpretability of existing trust models and their difficulty in characterizing the influence of machine performance fluctuations on human trust, a machine performance-driven trust evolution model is proposed. The proposed model enables the modeling and tracking of human-machine trust dynamics, providing a foundation for the subsequent design of trust-integrated collaborative methods. It explicitly distinguishes the machine's objective capability from the human's subjective assessment, characterizes the evolution process of human-machine trust, and improves model interpretability by analyzing the physical meanings and related properties of the core parameters.

(2) To address the lack of a reasonable control authority allocation mechanism in human-UAV collaborative landing tasks, a trust-integrated minimal-intervention shared control strategy is proposed. This strategy improves the task success rate and landing

accuracy while preserving human intent as much as possible. Specifically, the machine's objective capability and performance are first quantified based on intent inference, landing accuracy, and task safety metrics, and action-value evaluation and candidate action set construction are then performed based on deep reinforcement learning. Subsequently, the trust state is incorporated into the arbitration process to dynamically adjust the value-loss threshold of candidate actions. Following the principle of minimal intervention, the action closest to human control input is selected from the feasible set as the control output.

(3) To address the insufficient consideration of dynamic trust variations in trajectory planning for human-UAV collaborative search tasks, a trust-integrated human-UAV collaborative trajectory planning method is proposed. This method improves flight trajectory smoothness while reducing human workload. Specifically, the machine's objective capability and performance are first quantified based on flight safety and visibility metrics. Then, under safety constraints, a motion primitive tree is expanded according to human control input to construct a candidate trajectory set. Trust is then incorporated into trajectory selection, and the selected trajectory is further optimized, thereby generating a trust-aware adaptive flight trajectory.

KEY WORDS: Human-Machine Trust, Human-Machine Collaboration, Shared Control, Trajectory Planning

目录

插图和附表清单

第 1 章　绪论

1.1　研究背景及意义

无人机（Unmanned Aerial Vehicles, UAVs）作为自主智能装备的典型代表[1]，凭借轻量化、高机动性及低成本等物理特性[2]，已广泛应用于货物运输[3]、农业作业[4]与监视侦察[5]等各类军民用任务中[6]。近年来，随着计算机视觉与群体智能等算法的融合，现代无人机进一步具备了目标识别[7]与协同作业[8]的自主决策能力，从而能够有效替代人类在高危与复杂环境中执行任务[9]。基于其广泛的应用前景与战略价值，无人机系统已成为全球主要国家和地区重点布局的方向。例如，我国多部委联合印发的《安全应急装备重点领域发展行动计划（2023-2025 年）》明确将无人机列为矿山、危化品事故处置及紧急生命救护等高危领域的核心装备[10]；美国国防部发布的《Unmanned Systems Integrated Roadmap FY 2017-2042》系统性规划了无人平台向全域自主协同演进的战略[11]；欧盟出台的《A Drone Strategy 2.0》致力于打破技术与监管壁垒，构建涵盖应急医疗救援等复杂场景的智能化无人机生态系统[12]。

然而，受限于硬件平台的物理约束、外部环境的干扰以及智能算法自身的局限性，现阶段的无人机在复杂高危场景下仍难以实现全自主作业。在硬件物理层面，无人机有限的搭载能力要求机载传感器与计算平台小型化，这直接限制了系统的算力与感知范围[13]；此外，旋翼等执行结构易受外部碰撞损坏，导致系统的抗碰撞能力有限。除了平台自身的物理局限，外部环境的不确定性也构成了严峻考验。无人机在实际飞行中常面临突发的气流[14]与电磁干扰[15]，这些干扰因素难以进行精确的数学建模，容易引发传感器的感知误差，进而降低决策的可靠性。更深层次的制约来自自主无人机所依赖的智能算法。当前主流的深度学习模型不仅具有黑箱特性、缺乏可解释性[16]，行为逻辑也具有一定的不可预测性；在面对动态目标和突发情境时，其决策过程还难以兼顾实时性与准确性，从而增加了系统运行的安全风险。

为弥补自主无人机的上述缺陷，将人类的感知与决策优势与无人机的自主执行能力深度融合，构建人-无人机协同系统，已成为当前重要的研究方向。在这一协同框架下，人类的优势首先体现在对环境信息的感知能力上，相较于深度相机与激光雷达等主流机载传感器，人类视觉对光照突变、水雾与烟雾等环境干扰具有更强的鲁棒性[13]；其次，在决策层面，专业操作员凭借前瞻性的预测与决策能力，能够提前约 1.5 秒对飞行状态进行判断[17]，其在复杂环境下的表现通常优于现有的自主导航算法[13]。相应地，自主无人机的优势主要体现在底层

运动控制能力上，其飞控系统通常以数百赫兹的频率进行姿态调节[18]，能够在毫秒级时间内对指令偏差进行精确补偿，从而保持飞行姿态稳定，其响应速度更快、控制精度也显著高于人类手动控制。

在人-无人机协同系统中，信任是影响协同效能与任务安全性的关键因素。人类需要根据对无人机能力的判断，动态调整自身的介入程度与协同策略，而信任水平在很大程度上反映了其对无人机能力的认识。因此，信任水平将直接影响人类对无人机的接受、依赖及协作行为[19]，是实现高效、安全协同的重要前提。然而，无人机具有复杂的高维非线性动力学特性，且人类通过遥控终端获取的环境与状态信息往往十分有限[20]。由此形成的信息不对称[21]使人类难以客观评估无人机的状态与性能，从而使其信任水平可能偏离与无人机实际能力相匹配的合理区间，进而引发信任失衡：过度信任可能使人类在无人机发生异常或决策失误时未能及时干预[22]，从而导致任务失败，甚至引发安全事故；而信任不足则会导致频繁且不必要的干预[23]，在增加操作负担的同时降低任务执行效率。

综上所述，人机信任失衡已成为制约人-无人机协同系统效能与安全性提升的关键问题。因此，针对人-无人机协同中的信任建模与方法设计开展系统性研究，不仅具有重要的理论价值，也具有显著的工程应用意义。从理论层面看，建立能够量化人机信任水平并刻画其动态演化规律的计算模型，有助于深化对人机交互机制的理解，为人机协同方法的设计提供理论支撑；从应用层面看，将信任状态纳入协同方法设计，能够缓解因信任失衡引发的协作冲突，提升人-无人机协同系统的整体效能。基于此，本文围绕人-无人机协同系统中的信任建模及其在协同任务中的应用展开研究，提出了机器性能驱动的信任演化模型，面向人-无人机协同降落任务设计了融合信任的最小干预共享控制策略，面向人-无人机协同搜索任务提出了融合信任的人机协同轨迹规划方法。

1.2　研究现状

1.2.1　人机信任研究现状

随着人工智能与机器人技术的发展，机器逐渐具备了自主执行复杂任务的能力。这使得机器不再仅仅作为工具存在，而是开始承担部分自主决策与执行任务，人机关系也由传统的“主-从”控制逐步演变为协作关系。在这一过程中，人类操作员需要将部分任务交由机器自主完成，而人类对机器的信任程度是影响协同效果的重要因素。在此背景下，学术界围绕人机信任的定义、测量方法、演化建模及应用开展了广泛研究。

目前，在人机交互领域，关于信任的定义尚未形成统一认识[24]。尽管对信任的定义具体表述各异，但现有研究普遍认同信任包含三个核心要素：信任者、

被信任者以及客观存在的风险或不确定性[25]。Lee 和 See 的观点被广泛引用，他们将信任描述为“在面临不确定性和脆弱性的情境下，认为代理将帮助实现个体目标的态度”[26]。Kok 和 Soh 对该定义进行了详细阐述，将信任进一步定义为“一个多维的潜在变量，在过往事件与信任者随后在不确定环境中选择是否依赖被信任者之间起中介作用”[27]。

由于信任具有潜在性和多面性，难以被直接观测。因此现有研究通常通过量化与信任相关的外在指标对信任进行测量。现有的测量方法主要分为以下三类：

(1) 自我报告。该方法主要通过问卷调查获取受试者的反馈，是目前应用较为广泛的信任测量手段。由于当前人机信任评估尚未建立统一标准，研究人员多采用定制量表或针对特定场景的验证量表，因而不同研究之间的信任水平对比存在困难。Jian 等人基于实验数据的聚类分析编制了包含 12 个条目的信任量表，是目前应用最广泛的量表之一[28]。Schaefer 等人构建并验证了包含 40 个条目的人机信任量表，用于测量人机交互前后的信任变化[29]。此外，学者们提出了涵盖能力信任与道德信任的多维信任量表[30]，以及基于信任形成理论构建并经可靠性验证的双因子自动化系统信任测试[31]。尽管自我报告法能够较直观地反映受试者的信任水平，但该方法高度依赖个体主观判断，且难以满足人机协同系统对实时、动态信任数据的获取需求。

(2) 生理测量。随着传感技术的发展，生理测量逐渐成为对自我报告方法的重要补充，为实时测量人机信任提供了新的技术手段。该方法主要通过分析心率变异性[32]、功能性磁共振成像[33]、功能性近红外光谱[34]以及脑电图[35-36]等生理信号或指标，评估人机信任状态。在具体研究中，Dong 等人通过调节机器的技术能力与拟人化线索，并结合脑电信号分析，研究了不同情境下人类对机器的信任变化，验证了利用生理信号测量信任的可行性[37]。Akash 等人利用脑电图和皮电反应数据构建了数据驱动的信任模型，不仅实现了信任与不信任状态的二元分类，还进一步探讨了通过动态调整系统透明度来优化人机交互过程[35]。此外，Gupta 等人在虚拟现实环境中结合生理与主观评价等多模态数据，研究了人类对虚拟智能体的信任，发现认知负荷和智能体准确率对信任形成具有重要影响[38]。这类方法客观性强且支持实时评估，但应用复杂，目前仍难以在生理信号与信任状态之间建立明确的映射关系。此外，生理传感器采集的信号通常具有非平稳特征，而许多传统机器学习分类算法均基于平稳性假设，这使得数据分析与特征提取面临较大挑战。

(3) 行为测量。该方法通过观察用户在实际交互过程中的行为反应与决策结果，来间接推断用户的信任状态。研究表明，信任的直接行为结果表现为人类对自主系统的依赖程度，具体表现为合理使用、因过度信任而产生的误用以及因缺乏信任而导致的弃用等典型行为[26]。在具体度量方式上，不同研究通常根据实

际交互场景选取不同的行为指标。例如，Hu 等人利用受试者在连续交互中对系统警报的二元决策响应，将其表示为概率分布，从而建立并验证了信任动态演化的计算模型[39]。Hudspeth 等人在共享物理空间的人机交互任务中，比较了不同界面形式与反馈机制对用户信任的影响，并结合自我报告与间接行为指标，评估了人类对机器人的信任[40]。这类方法能够较为直观地反映人机之间的实际协同效果与交互结果，但其局限性在于难以揭示信任演变背后的内在认知机理。

信任的离散测量数据为构建定量的信任演化模型提供了数据基础。根据信任的建模与推断方式不同，现有研究主要可分为以下三类：

(1) 时间序列模型。此类模型通过分析历史数据序列，刻画信任随时间的演化趋势以及外部事件对信任的影响。Lee 和 Moray 在半自动过程控制场景中，基于系统性能与机器故障状态构建了自回归滑动平均向量模型，分析机器故障对人类信任的影响[41]。Sadrfaridpour 等人在协作制造场景中，基于人类疲劳水平与人机速度差构建了时间序列信任模型，并结合手动、自主和协同三种机器人速度控制模式，量化并预测人类信任[42]。

(2) 基于微分/差分方程的动态系统模型。这类模型通常基于控制理论与系统辨识方法，将信任建模为系统中的状态变量。Hoogendoorn 等人在认知层面和神经层面分别建立了人类信任动态的微分模型，并通过仿真实验和形式化验证技术分析了其动态行为特征[43]。Gao 和 Lee 扩展了决策场理论，用于描述人类依赖自动化系统时的连续决策行为及信任演化过程[44]。Akash 等人利用灰箱系统辨识技术，基于大规模受试者行为数据建立了三阶线性模型，用于建模人类对障碍物检测传感器的信任[45]。Hu 等人构建了线性的人机信任计算模型，将信任水平描述为经验、累积信任和期望偏差的函数，并研究了机器误报与漏报对信任的不同影响[39]。

(3) 基于概率推断与机器学习的方法。该类方法侧重于处理交互环境和信任推断过程中的不确定性，并能够结合多模态数据进行在线估计。Xu 和 Dudek 利用动态贝叶斯网络构建了在线概率推断模型，并基于历史交互经验推断人类的潜在信任状态[46]。Akash 等人以部分可观测马尔可夫决策过程为框架，将信任与工作负荷建模为隐状态变量，描述了人机交互过程中状态的动态转移过程[47]。Hu 等人采用多种机器学习分类算法对连续心理生理信号进行分析，实现了对信任水平的分类[48]。Akash 等人则进一步提出了一种自适应概率分类算法，将马尔可夫决策过程与自适应贝叶斯二次判别器相结合，并利用脑电图与行为响应数据，实现了对信任状态的在线预测[35]。

构建信任演化模型有助于改善人机系统设计，从而降低信任失衡对协同效能的负面影响。当前将信任引入人机协同系统设计的研究主要集中在以下两个方向：

(1) 利用解释性信息进行信任校准。此类方法主要通过提升系统透明度、优化交互界面或提供解释性反馈，动态校准人类对系统的信任状态。Huang 等人研究表明，向人类展示任务的关键信息能够有效帮助其构建准确的心理模型并理解系统策略，从而将信任水平调节至合理区间[49]。Rezvani 等人在自动驾驶交互研究中发现，直观展示系统对环境异常的感知信息能提升驾驶员对系统的信任，而呈现系统自身的不确定性反而会损害交互体验和表现[50]。Okamura 等人提出了自适应信任校准框架，该框架通过实时监测用户的依赖行为，在识别到过度信任或信任不足时主动向用户提供相应提示信息，从而实现对信任状态的动态校准[51]。

(2) 利用信任状态调节人机协同策略。此类方法通常基于人机信任模型，对机器自主控制权限或人机协同策略进行动态调节。Zahedi 等人将人类的信任状态纳入机器人的任务规划过程中，使机器人能够根据当前的信任状态在行为可解释性与执行效率之间进行动态权衡[52]。Li 等人建立了人机双向信任的状态空间模型，并通过模型预测控制实现了移动机器人自动化程度的动态调节[53]。Hu 等人通过量化驾驶员期望与车辆实际性能之间的偏差来评估驾驶员的信任状态，并结合无模型自适应动态规划算法，实现了智能车辆转向控制权限的动态分配[54]。

综上，现有研究在人机信任的测量、建模与应用方面取得了一定进展，但面向无人机任务场景的信任研究仍存在不足。现有模型大多针对地面机器人、自动驾驶或一般自动化系统构建，对无人机任务中环境复杂、状态变化快以及人类感知受限等特点考虑不足，因而难以直接适用。因此，有必要进一步研究适用于无人机任务场景的信任建模方法。

1.2.2　人-无人机协同方法研究现状

针对复杂动态环境下的任务需求，学术界普遍认识到单一控制模式的局限性，进而转向探索人-无人机优势互补的协同解决方案。现有研究主要围绕如何将人类的感知与决策优势融入无人机控制与规划过程展开。根据协同方法作用层面的不同，相关研究主要可分为操作控制与轨迹规划两个层面。

在操作控制层面，人-无人机协同主要通过结合操作员输入以及无人机的自主控制、避障等能力，借助安全过滤、共享控制等机制提升操作安全性与稳定性。系统通常依据局部感知结果、人类意图和任务目标生成安全约束，并在此基础上对控制输入进行修正，以避免碰撞或偏离任务目标。Nieuwenhuisen 等人构建了结合多模态传感器与自中心栅格地图的全向障碍感知系统，并设计了快速反应式碰撞规避方法，以实现微型无人机在复杂环境中的安全飞行[55]。Odelga 等人针对遥操作多旋翼无人机的室内导航任务，结合局部障碍物检测与跟踪，提出了基于模型预测控制思想的避障方法，可通过预测障碍物状态实时修正操作

员指令，从而避免碰撞[56]。Ghaffari 提出了基于可达性分析与模型预测控制的碰撞规避辅助方法，通过构建安全凸区域并修正操作员的线加速度指令，实现安全遥操作[57]。Backman 等人在无人机着陆任务中结合跨模态变分自编码器与 TD3，通过隐式推断操作员的意图并生成辅助控制输入，辅助经验不足的操作员实现安全着陆[58]。

在轨迹规划层面，人-无人机协同主要关注将人类意图融入路径搜索与轨迹生成过程，并在满足动力学与环境约束的条件下规划安全、平滑的飞行轨迹。Yang 和 Michael 将人类意图建模为方向代价，并通过偏置增量动作采样生成与期望方向一致且满足避障要求的长时域平滑轨迹[59]。在此基础上，Yang 等人进一步提出了层次化遥操作框架，通过全局路径表示操作员在较长时间尺度上的期望飞行方向，并在局部生成兼顾安全性与动力学可行性的轨迹[60]。Wang 等人利用视线信息推断人类意图并构建拓扑路径，并进一步优化生成兼顾安全性、可视性与期望飞行速度的飞行轨迹[61]。Shi 等人则提出了人类引导的感知约束运动规划方法，通过初始路径搜索与轨迹优化，在兼顾可视性的同时保持与人类意图的一致性[62]。

综上，现有面向操作控制与轨迹规划的人-无人机协同方法，已在提升飞行安全性和减轻操作员负荷等方面取得了较好效果。然而，这类方法大多侧重于满足安全约束和响应操作意图，对人类信任状态的考虑相对有限，在设计上通常默认人类能够始终保持适度干预。因此，现有人-无人机协同方法在融合人类信任状态方面仍有待进一步完善。

1.2.3　研究现状总结

综上，现有研究已分别在人机信任建模和人-无人机协同方法设计方面取得一定进展，但二者之间的结合仍显不足，尚未形成将人机信任有效纳入人-无人机协同过程的系统框架。结合前述研究现状，当前主要存在以下三个方面的问题：

(1) 现有信任模型在无人机任务场景中的适用性与可解释性仍有待提升。现有研究大多围绕地面移动机器人、自动驾驶车辆等系统展开，而无人机任务场景通常具有更高的环境复杂性与更受限的感知条件，因此相关模型难以直接适用于人-无人机协同场景。进一步地，现有模型通常将影响信任演化的因素直接引入数学表达式中，但对模型参数的实际含义及其动态性质缺乏深入分析，因而难以清晰揭示信任变化的内在机理。这在一定程度上限制了模型的可解释性，也制约了其在人机协同机制设计中的应用。

(2) 人机控制权分配机制仍缺乏合理的动态调节方式。现有方法主要依据安全约束或人类意图进行人机控制权分配，以实现对人类控制输入的约束与修正。

尽管已有研究开始将信任因素引入共享控制设计，但大多仍采用线性映射、固定阈值或预设规则等方式对控制权进行调节，对信任的利用仍较为简单。这类方法难以根据信任状态对机器权限进行合理动态调节，从而难以保证机器干预的合理性。

(3) 现有轨迹规划方法对人类信任动态变化的考虑仍显不足。现有方法主要依据环境约束与人类意图生成轨迹，重点在于保证轨迹安全并响应人类意图，但通常未将任务过程中人类的信任状态有效纳入轨迹规划过程，因此难以根据信任状态对规划结果进行自适应调节。当人类对机器的信任偏离合理区间时，系统规划轨迹与其主观预期之间可能产生偏差，从而影响协同过程的一致性与稳定性。

1.3　研究内容与组织结构

1.3.1　研究内容

针对上述问题，本文提出了机器性能驱动的信任演化模型，并在此基础上分别面向降落与搜索任务设计了融合信任的人-无人机协同方法。总体研究框架如图1.1所示，主要包括以下三个方面：

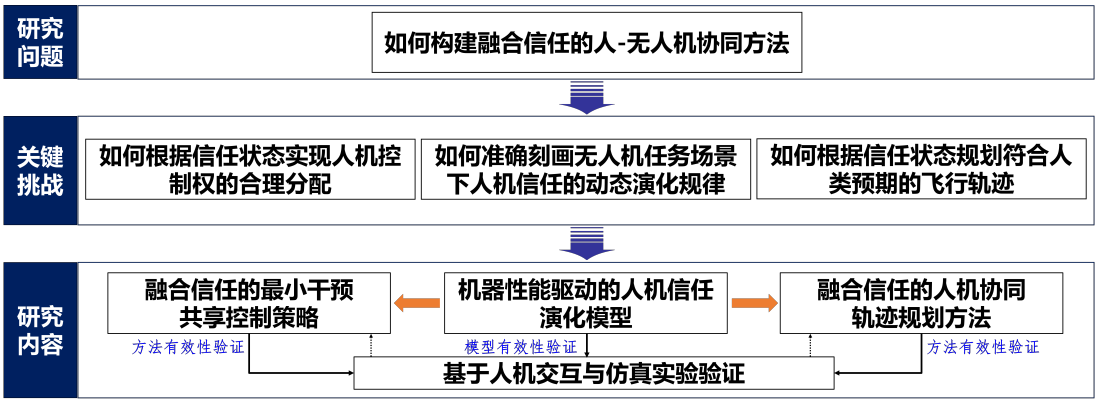


图 1.1　总体研究框架图

(1) 针对现有信任模型可解释性不足，未能充分考虑复杂环境中机器性能波动对人类信任影响的问题，提出了机器性能驱动的信任演化模型。该模型明确区分了机器的客观能力与人类的主观认知，分析了机器实时性能表现对信任动态演化的影响规律；在此基础上，基于机器性能指标实现信任的量化与动态更新，刻画了人机信任演化过程；同时，通过分析模型核心参数的物理意义及相关数学性质，提升了模型的可解释性。

(2) 针对人-无人机协同降落任务中控制权分配缺乏合理动态调节机制的问题，提出了融合信任的最小干预共享控制策略。该策略首先构建基于意图推理、降落精度与任务安全性的评价指标，用于量化机器的客观能力与性能表现；随

后，基于深度强化学习进行动作价值评估，并在连续动作空间中通过离散采样构造候选动作集，再依据价值约束筛选得到满足基础性能要求的可行动作集合；在此基础上，将人类信任状态引入共享控制仲裁过程，根据信任水平动态调节候选动作可接受的价值损失阈值，最终按照最小干预原则从满足约束的动作中选取与人类输入最接近的动作作为控制输出。

(3) 针对人-无人机协同搜索任务中轨迹规划对信任动态变化考虑不足的问题，提出了融合信任的人机协同轨迹规划方法。该方法首先以飞行安全性和可视性为评价指标，构建搜索任务下机器的性能表现与客观能力；其次，根据人类输入生成引导运动基元，并在安全约束下扩展运动基元树，构造候选轨迹集；然后，综合当前局部轨迹连续性、候选轨迹与人类意图的一致性以及候选轨迹对应的平均信任水平完成轨迹选择；最后，将信任状态引入轨迹优化过程，通过自适应调节优化权重生成信任感知的飞行轨迹。

1.3.2　组织结构

本文的组织结构如图1.2所示，具体安排如下：

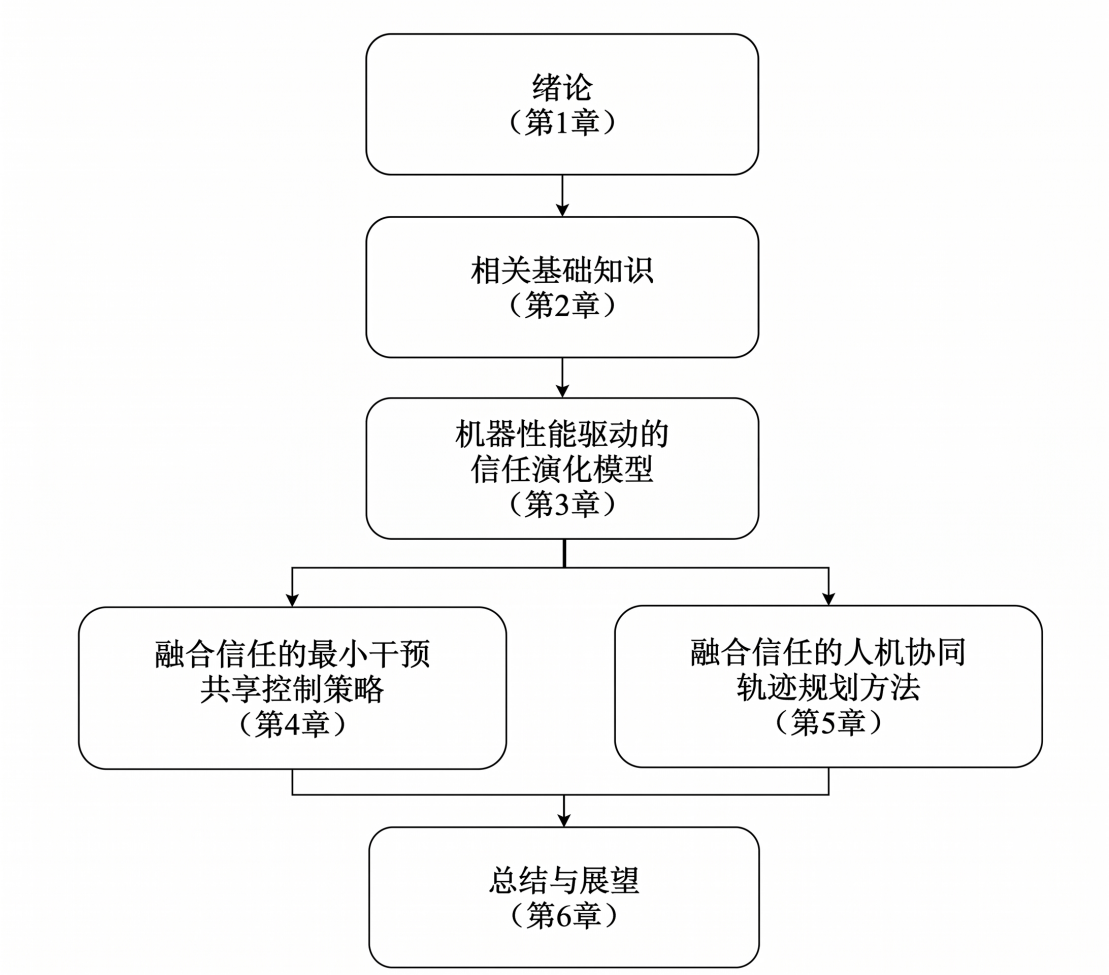


图 1.2　本文组织结构图

第一章：绪论。介绍了本文的研究背景，阐述了在人-无人机协同任务中引入人机信任的研究意义。同时总结了现有人机信任与人-无人机协同方法的研究现状，并在此基础上给出了本文的研究内容与组织结构。

第二章：相关基础知识。介绍了本文涉及的相关概念与基础知识，包括人机协同控制方式、强化学习基础理论以及意图推理方法。

第三章：机器性能驱动的信任演化模型。研究了在人机交互过程中，如何将机器性能表现与人类认知更新机制相结合，以刻画人类对机器能力的认知偏差及信任的动态演化规律。具体包括人机信任的定义、机器能力的刻画、人机信任演化模型的构建、关键性质分析以及人机交互实验验证等，为后续融合信任的协同方法设计提供理论依据与模型支撑。

第四章：融合信任的最小干预共享控制策略。研究了在人机协同降落任务中，如何结合人类信任状态与最小干预原则设计仲裁机制，以确保控制权分配的合理性。具体包括降落任务下机器能力构建、基于深度强化学习的动作价值评估与候选动作集构造、信任驱动的最小干预仲裁机制，以及人机协同降落仿真实验验证等。

第五章：融合信任的人机协同轨迹规划方法。研究了在人机协同搜索任务中，如何将人类意图与信任状态共同纳入轨迹生成、选择与优化过程，从而实现信任感知的自适应轨迹规划。具体包括搜索任务下机器能力构建、意图引导的候选轨迹生成与选择、基于信任状态自适应调节的轨迹优化策略，以及人机协同搜索仿真实验验证。

第六章：总结与展望。总结全文的研究成果，分析研究中存在的不足，并展望未来可以改进的方向。

第 2 章　相关基础知识

本章介绍了本文研究涉及的相关概念与基础知识。首先，介绍了人机协同中的两类典型控制方式，即共享控制与介入控制；其次，梳理了强化学习的基础理论，包括马尔可夫决策过程、基于价值与基于策略的方法以及深度强化学习；最后，介绍了人机协同系统中的两类意图推理方法，即目标意图推理和轨迹意图推理。

2.1　人机协同控制方式

随着机器自主决策能力的提升，传统的直接控制模式已难以满足复杂任务需求，因此人机协同控制受到广泛关注。人机协同控制的核心在于根据任务需求与环境变化，在人与机器之间合理分配控制权。尽管学术界对于具体协同模式尚未形成完全统一的分类标准，但从控制权的交互逻辑与分配机制出发，相关方法通常可分为共享控制与介入控制两类。

2.1.1　共享控制

共享控制是指人类与机器同时参与同一控制任务的人机协作模式[63]。如图2.1所示，共享控制通常处于人类直接控制与监督控制之间。在这种模式下，人机双方不通过离散切换来分配控制权，而是以连续协同的方式共同参与控制过程[64]。共享控制通过结合人类与机器各自的优势，能够有效弥补单一控制模式在复杂任务中的不足。

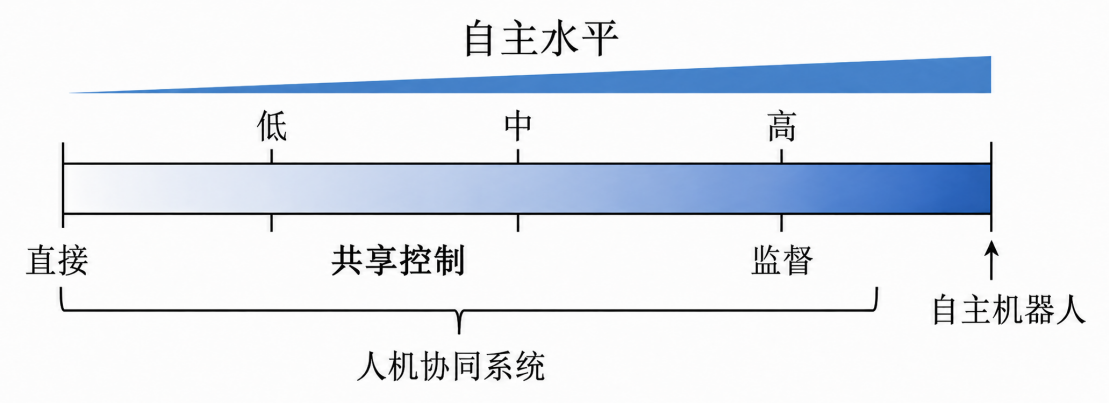


图 2.1　人机协同系统中的自主水平与控制模式示意图[65]

在共享控制系统设计中，核心问题在于建立合理的仲裁机制，以协调人类控制输入与机器控制指令之间的关系。仲裁机制是在共享控制过程中对人机指令

进行融合与冲突消解的动态机制，用于确定人类输入与机器控制之间的作用边界，以实现系统整体目标。在许多共享控制方法中，仲裁机制通过动态调节控制权重，将人类控制输入与机器控制指令进行加权融合。一个经典的仲裁形式可表示为：

$$u = \lambda u_m + (1-\lambda)u_h, \quad 0 \leqslant \lambda \leqslant 1 \tag{2.1}$$

其中，u 为系统最终输出的控制指令，u_m 与 u_h 分别表示机器控制指令与人类控制输入指令，λ 为控制权分配权重，用于表示机器在共享控制中的参与程度。通过合理设计权重 λ 的调节逻辑，系统可以在响应人类意图的同时提供必要的辅助。

2.1.2　介入控制

不同于共享控制中人机双方持续、同步参与控制过程，介入控制主要关注特定条件触发下的控制权转移与干预。如图2.2所示，在这种模式下，控制权通常由一方主导，并在满足触发条件时发生切换。根据干预发起主体的不同，介入控制可进一步划分为“机器介入人类”与“人类介入机器”两种典型模式。

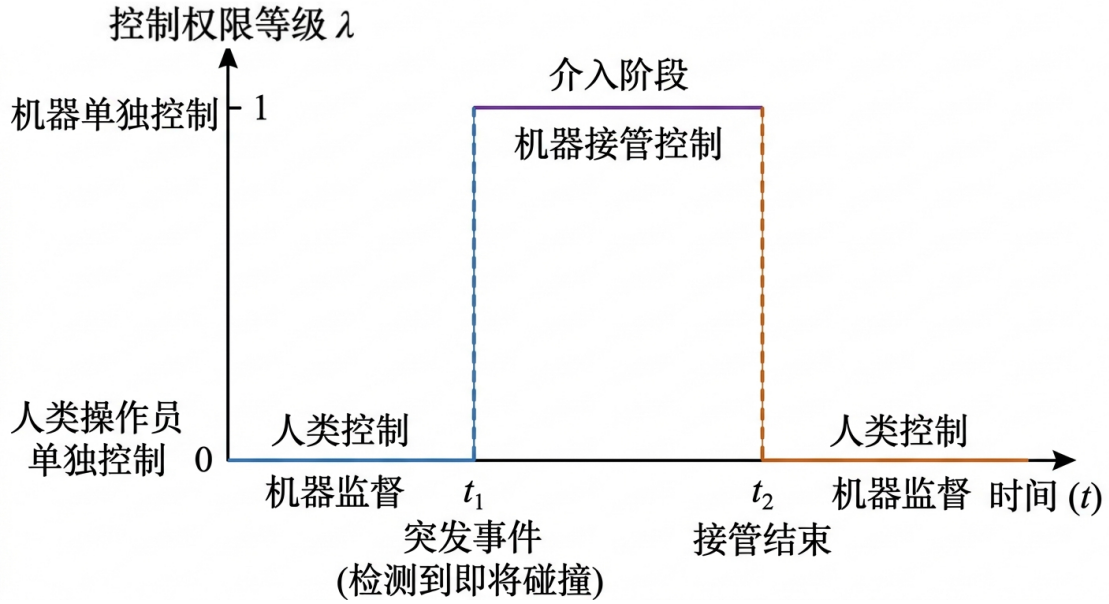


图 2.2　介入控制中的人机控制权转移示意图

一方面，在“机器介入人类”的模式中，系统持续监测人类输入与环境状态。当判定当前指令可能使系统逼近物理约束边界或引发安全风险时，系统会对人类发出的控制指令进行修正，并在必要时短暂接管控制权；在干预结束后，再将控制权交还给人类。通过在特定条件下对人类控制行为进行干预，系统能够在保障运行安全的同时尽可能保留人类的主导作用。

另一方面，“人类介入机器”的模式通常出现在自动化系统向人类发出接管请求的场景中。当人类将任务交由自主系统执行后，人类主要负责监视系统运

行。一旦环境变化超出系统的运行设计域，或系统因传感器、算法等问题无法有效应对突发情况，系统便会向人类发出接管请求，并由人类重新接管控制权。

无论由哪一方发起介入，介入时机的设计都是影响协同系统安全性的关键。这是因为，在控制权转移过程中，人类往往面临情境意识下降及其恢复的问题。特别是在“人类介入机器”的过程中，人类从监视系统运行转向重新理解环境并接管控制权，需要一定的认知与反应时间。如果介入时机过晚，人类可能在情境意识尚未充分恢复的情况下仓促接管，进而引发操作震荡，甚至导致系统失控；而介入时机过早，则可能造成人类注意力的不必要消耗。因此，介入时机的设计必须符合人类认知恢复的客观规律，才能确保控制权转移与人类认知状态调整之间的平稳衔接。

2.2　强化学习基础理论

强化学习主要研究智能体如何在与环境的持续交互过程中，根据奖励反馈不断调整策略，以最大化累积回报。在人机协同任务中，机器端的自主决策过程可由强化学习方法进行建模，其中环境由任务状态、外部约束以及交互对象等因素构成。由于能够利用交互反馈不断优化策略，强化学习已被广泛用于复杂非线性系统中的序贯决策与控制问题。

2.2.1　马尔可夫决策过程

强化学习通常在马尔可夫决策过程（Markov Decision Process, MDP）框架下进行建模。MDP 描述了一个序贯决策问题，其核心假设是系统的未来状态仅依赖于当前状态与所采取的动作，而与历史状态无关。一个标准的 MDP 由五元组 $(\boldsymbol{S}, \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{R}, \gamma)$ 表示：

- 状态空间（State Space, $\boldsymbol{S}$）：环境中所有可能状态构成的集合。在典型的物理控制系统中，状态通常包含位置、速度、姿态角等动态信息。
- 动作空间（Action Space, $\boldsymbol{A}$）：智能体在各状态下可执行动作的集合，可分为离散动作空间与连续动作空间。
- 状态转移概率（Transition Probability, $\boldsymbol{P}$）：描述环境状态转移规律的概率模型，记为 $\boldsymbol{P}(s'|s,a)$，即在当前状态 s 下执行动作 a 后转移至下一状态 s' 的概率。
- 奖励函数（Reward Function, $\boldsymbol{R}$）：智能体在状态 s 执行动作 a 时获得的即时反馈信号，记为 $\boldsymbol{R}(s,a)$，是引导策略优化的直接指标。
- 折扣因子（Discount Factor, γ）：取值范围为 $[0,1]$，用于权衡即时奖励与未来奖励的相对重要性。

在 MDP 框架下，强化学习的目标是寻找最优策略 π^*，使智能体获得的期望累积回报最大化。累积回报 G_t 定义为：

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{2.2}$$

2.2.2　基于价值与基于策略的强化学习方法

为了求解最优策略，强化学习方法通常可分为两类：基于价值的方法与基于策略的方法。

基于价值的方法通过评估状态或状态–动作对的长期期望回报来确定最优策略。动作价值函数 $Q_\pi(s,a)$ 表示在状态 s 下执行动作 a，并在此后遵循策略 π 时所能获得的期望累积回报，其定义为：

$$Q_\pi(s,a) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \,\middle|\, S_t = s, A_t = a \right] \tag{2.3}$$

在此基础上，最优动作价值函数 $Q^*(s,a)$ 满足贝尔曼最优方程：

$$Q^*(s,a) = \mathbb{E}_{s' \sim P(\cdot|s,a)} \left[R(s,a) + \gamma \max_{a'} Q^*(s',a') \right] \tag{2.4}$$

基于价值的方法通常通过迭代逼近上述最优方程来获得最优策略。然而，这类方法在输出控制动作时往往需要在动作空间中搜索使 Q 值最大的动作，因此难以直接应用于连续且高维的动作空间。

与此不同，基于策略的方法直接对策略函数 $\pi_\theta(a \mid s)$ 进行参数化建模，并通过优化参数 θ 最大化期望回报目标函数 $J(\theta)$。其核心思想是通过调整策略参数，提高高回报动作在相应状态下的选择概率。经典策略梯度定理可表示为：

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta} \left[\nabla_\theta \ln \pi_\theta(a \mid s)\, Q_{\pi_\theta}(s,a) \right] \tag{2.5}$$

基于策略的方法能够自然适应连续动作空间，因此常用于解决连续控制问题。但在实际训练过程中，策略梯度方法通常存在梯度估计方差较大、样本利用率较低等问题，从而限制了训练效率与稳定性。

2.2.3　深度强化学习

传统强化学习方法通常适用于低维离散决策问题，但在状态空间维数较高、动作空间连续且系统动力学复杂的场景下，这类方法往往难以有效处理高维状态表示与策略求解问题。为应对复杂控制任务，深度强化学习（Deep Reinforcement Learning, DRL）将深度神经网络引入强化学习框架，利用深度神经网络逼近价值函数或策略函数，以处理高维状态空间下的决策与控制问题。

在深度强化学习中，基于价值的方法通常以深度 Q 网络（Deep Q-Network, DQN）[66]为代表。DQN 使用神经网络逼近动作价值函数，并结合经验回放与目

标网络机制缓解训练样本相关性强和优化不稳定的问题。然而，DQN 仍需在离散动作集合中搜索使 Q 值最大的动作，因此难以直接扩展到连续动作控制场景。

相比之下，基于策略的深度强化学习方法直接利用深度神经网络表示策略函数 $\pi_\theta(a \mid s)$，能够更好地处理连续动作空间下的控制问题。但当任务环境较复杂、策略网络较深时，该类方法仍然面临梯度估计方差较大以及对超参数较为敏感等挑战。

Actor–Critic 框架将价值函数估计与策略梯度更新相结合，通过价值评估为策略更新提供指导，因此被广泛用于复杂连续控制问题的求解。在该框架中，Actor 负责根据当前状态输出控制动作或动作概率分布，对应参数化策略网络 π_θ；Critic 负责评估当前状态或状态–动作对的长期期望回报，对应价值网络 V_ϕ 或 Q_ϕ。借助 Critic 提供的评价信号，Actor 的策略更新能够获得更稳定的梯度估计，从而缓解策略梯度方法中方差过大的问题。

在常见的基于 Actor–Critic 架构的策略优化算法中，近端策略优化（Proximal Policy Optimization, PPO）[67]因具有较好的训练稳定性而得到广泛应用。PPO 通过引入裁剪机制限制新旧策略之间的相对变化幅度，以避免策略发生过大的单步更新。其裁剪代理目标函数为：

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min \left(r_t(\theta)\hat{A}_t,\ \mathrm{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t \right) \right] \tag{2.6}$$

其中，$r_t(\theta)$ 表示新旧策略输出动作概率的比值，$\hat{A}_t$ 为优势函数的估计值，ϵ 为控制裁剪范围的超参数。当概率比超出区间 $[1-\epsilon,\ 1+\epsilon]$ 时，目标函数将被截断，从而抑制过大的策略更新。

2.3　意图推理

在人机协同系统中，实现高效协同的前提在于机器能够准确推断人类的操作意图。然而，在实际运行中，机器通常无法直接获知人类的意图，这使得系统难以准确判断人类的真实目标或任务需求。尽管人类可以通过额外的交互接口直接表达目标或需求，但已有研究表明，与从人类操作行为中推断意图相比，这种方式不利于实现高效协作[58]，且还会增加额外的交互负担。

因此，在主流的人机协同架构中，机器通常需要依据连续的交互数据（如控制指令或运动学状态）对人类意图进行隐式推断。在这一过程中，意图推理的准确性直接影响协同策略的有效性与人机交互的流畅度。若机器频繁误判人类意图，其生成的辅助行为可能与人类预期不一致，迫使人类额外输入指令以修正机器行为。这不仅增加了人类的认知与物理负荷，还容易引发人机决策冲突，进而削弱人机系统的整体协作效能。

根据具体任务形式的不同，意图推理通常可分为目标意图推理与轨迹意图

推理两类。

2.3.1　目标意图推理

目标意图推理是指在离散的语义目标或空间目标集合中，推断人类希望到达的目标位置或所要执行的具体任务。递归贝叶斯滤波[68]是求解这类问题的经典方法。该方法将意图推理表述为后验概率估计问题，具有较强的可解释性，并适用于状态维度较低或目标集合有限的场景。

在这一理论框架下，通常假设环境中存在离散目标集合 $G = \{g_1, g_2, \cdots, g_N\}$，其中 $g_t \in G$ 表示 t 时刻的人类真实目标。机器通过观测人类连续的控制输入序列 $U_{0:t} = \{u_0, u_1, \cdots, u_t\}$ 来递归更新对各个目标的信念状态。信念状态 $b_t(g_t)$ 定义为在给定历史操作输入下，时刻 t 人类目标为 g_t 的后验概率：

$$b_t(g_t) \triangleq P(g_t \mid U_{0:t}), \qquad g_t \in G \tag{2.7}$$

根据贝叶斯定理，结合马尔可夫假设，后验概率可以展开为似然项与先验项的乘积。利用全概率公式对先验项进行递推，可得到信念状态的标准递归更新方程：

$$b_t(g_t) \propto P(u_t \mid g_t) \sum_{g_{t-1} \in G} P(g_t \mid g_{t-1})\, b_{t-1}(g_{t-1}) \tag{2.8}$$

其中，$P(u_t \mid g_t)$ 为似然函数，描述了在给定当前目标状态 g_t 时，人类采取当前操作指令 u_t 的生成概率；$P(g_t \mid g_{t-1})$ 为意图的转移概率，用于建模人类在执行任务过程中从上一时刻目标状态 g_{t-1} 转移到当前目标状态 g_t 的概率。

在完成概率更新后，系统可通过最大后验概率估计确定最可能的人类目标 g_t^*：

$$g_t^* = \arg\max_{g_t \in G} b_t(g_t) \tag{2.9}$$

为了衡量推理结果的可靠性，通常引入置信度指标，如最大与次大目标概率的差值或信息熵。当置信度指标满足设定阈值条件时，机器即可触发相应的辅助策略。

2.3.2　轨迹意图推理

轨迹意图推理不同于关注离散目标识别的目标意图推理，它主要根据系统当前及历史状态和操作输入，对未来一段时域内的运动状态序列进行预测，从而为机器的运动控制与协同规划提供前馈信息。

轨迹意图推理可表述为一个多变量时间序列预测问题。假设历史观测窗口长度为 H_p，预测时域长度为 H_f，则预测模型需要建立从历史状态 $X_{t-H_p+1:t}$ 与

操作输入 $U_{t-H_p+1:t}$ 到未来轨迹序列 $\hat{Y}_{t+1:t+H_f}$ 的映射关系 f_θ：

$$\hat{Y}_{t+1:t+H_f} = f_\theta(X_{t-H_p+1:t}, U_{t-H_p+1:t}) \tag{2.10}$$

由于人类控制行为具有较强的非线性特征，并伴随复杂的时序依赖关系，传统的线性运动学外推方法难以在较长预测时域内保持较高精度。因此，在实际研究中，映射函数 f_θ 通常建模为序列神经网络，如循环神经网络[69-70]或基于自注意力机制的模型[71]。这类方法通过提取时序特征，能够较好地学习人类控制行为中的非线性规律，从而生成平滑且符合运动学演化趋势的预测轨迹。

第 3 章　机器性能驱动的信任演化模型

在人机协同系统中，机器的性能表现是影响人类信任水平的重要因素[72]。然而，在实际任务执行过程中，机器的性能表现并不能直接对应人类的信任水平。无人机系统具有较强的非线性动力学特性，且人类操作员通过终端能够获取的环境与状态信息相对有限。因此，人类难以准确把握系统的实际运行状态以及机器在任务中的表现，从而使其对机器能力的判断容易出现偏差。

为了建模这一认知偏差以及信任随机器性能变化的动态演化过程，本章构建了一种机器性能驱动的人机信任演化模型，围绕人机信任的定义、客观机器能力与主观机器能力的量化方法、信任演化模型的构建、模型参数的物理意义与性质分析，以及仿真实验设计与模型有效性验证等内容展开研究。

3.1　人机信任定义

在人机协同系统中，人类对机器的信任来源于客观机器能力与主观机器能力之间的关系。客观机器能力反映机器真实具备的能力，决定了机器实际能够达到的水平；主观机器能力反映人类在交互过程中对机器能力形成的主观认知，决定了人类认为机器能够达到的水平。然而，在实际交互过程中，主观机器能力与客观机器能力之间往往存在偏差，使得人类容易高估或低估机器的真实能力，进而使人类处于过度信任或缺乏信任的状态。

定义 3.1 (人机信任)　为准确刻画人类对机器能力的认知偏差及其动态演化规律，将机器性能驱动的人机信任定义为人类的主观机器能力 $C_m^h(k)$ 与客观机器能力 $C_m^m(k)$ 的比值：

$$T_m^h(k) = \frac{C_m^h(k)}{C_m^m(k)}, \quad k \geqslant 0 \tag{3.1a}$$

$$T_m^h(t) = T_m^h(k), \quad t \in [t_k, t_{k+1}), \quad k \geqslant 0 \tag{3.1b}$$

其中，t_k 为人类第 k 次更新主观机器能力的时刻，$T_m^h(t)$ 为连续时间下 t 时刻的信任值；$T_m^h(k)$、$C_m^h(k)$、$C_m^m(k)$ 分别为第 k 次更新时刻对应的信任值、主观机器能力与客观机器能力。由式（3.1b） 可知，所定义的人机信任为分段常值函数，其值仅在特定离散时刻发生更新，而在相邻两次更新之间保持恒定。

进一步地，从上述定义出发，可以归纳出人机信任的几个主要性质：

- 信任是机器性能驱动的。机器性能已被证明是影响人机信任的重要因素[72]。在上述信任定义下，客观机器能力通过机器性能直接计算得出，主观机器能力则来源于人类对机器性能的观察与评估。因此，客观机器能力

与主观机器能力均以机器性能为基础，信任的建立与演化本质上由机器性能所驱动。

- 信任具有个体差异性。不同人类个体在感知与评估机器能力时存在认知差异，这一差异需要在信任建模中加以体现。通过引入主观机器能力，个体对机器能力的认知偏差可以直接反映在主观能力上，从而使信任定义具备个性化特征，能够针对不同个体进行差异化建模。
- 信任是动态变化的。由于机器性能会随环境与任务状态的改变而变化，因此机器性能驱动的信任同样会随任务与环境状态的演变而动态变化。
- 信任是可计算的。在上述信任定义下，主观机器能力虽具有一定主观性，但可通过具体方法进行量化。在确定个体认知偏差之后，可以实现对人机信任的计算，并通过信任演化模型预测信任的演化趋势，从而为人机系统设计提供依据。

基于上述对人机信任的定义及其可计算性质，可以进一步对人机交互中的信任状态进行明确的划分。在实际交互中，人类不恰当的信任主要表现为两种基本形式：一是过度信任，即人类高估了机器的实际能力；二是缺乏信任，即人类低估了机器的实际能力。在理想状态下，当主观机器能力与客观机器能力完全对齐（即 $T_m^h(t)=1$）时，信任处于适当的状态。然而，在真实的人机交互场景中，由于环境与任务状态动态变化，要求二者始终严格相等不仅难以实现，而且在工程应用中也缺乏实际必要性。事实上，只要主观认知与客观能力之间的偏差维持在较小的允许范围内，人类便不会产生不当的控制行为，系统依然能够维持良好的协同效能。

定义 3.2 (信任状态)　考虑上述情况，引入一个较小的容差阈值 $\delta \in (0,1)$，认为当信任处于区间 $[1-\delta, 1+\delta]$ 内时，均属于适当信任状态。在此基础上，连续时间 t 下的人机信任 $T_m^h(t)$ 可被划分为以下三种状态：

$$\text{信任状态}(t)=\begin{cases}\text{过度信任}, & T_m^h(t)>1+\delta \\ \text{适当信任}, & 1-\delta \leqslant T_m^h(t) \leqslant 1+\delta \\ \text{缺乏信任}, & T_m^h(t)<1-\delta\end{cases} \tag{3.2}$$

3.2　机器能力刻画

3.2.1　机器性能表现

人类对机器的信任并非直接来源于机器的设计指标或理论能力，而更多取决于人类在交互过程中对机器实际表现的持续观察与主观判断。因此，需要对机器的性能表现进行量化，作为信任更新的重要依据。

由于不同人机协作任务的目标形式、环境约束与人类关注重点存在差异，机器性能表现所包含的评价指标也不相同。以无人机降落任务为例，在具体降落位置事先未知的情况下，无人机需要根据人类的实时输入持续推理其意图，并最终完成降落任务。因此，该任务中的机器性能表现主要体现在意图推理结果、降落精度等方面。对于不同任务，只需根据具体任务目标和交互方式选取相应的评价指标即可。为描述不同任务下的机器性能表现，本节首先给出机器性能表现的一般形式。

设信任在离散时刻 t_k 进行更新，则相邻两次信任更新之间的时间区间记为：

$$\Delta k = [t_{k-1}, t_k), \qquad k \geqslant 1 \tag{3.3}$$

对于某一具体任务，设机器性能表现可由 M 个评价指标综合确定。记第 j 个评价指标在区间 Δk 内对应的评价结果为 $q_j(k)$，其中 $j \in \{1, 2, \cdots, M\}$，且 $q_j(k) \in [0, 1]$。$q_j(k)$ 的具体含义可根据任务特点确定。

考虑到人类在人机交互过程中对近期体验更为敏感，同时机器历史表现也会影响人类对机器的信任[26]，本文采用时间衰减机制对历史区间内各评价指标的评价结果进行动态加权。设时间窗口长度为 L，衰减因子为 $\gamma \in (0, 1)$，则第 j 个评价指标在第 k 个区间上的表现定义为：

$$P_j(k) = \frac{\sum_{i=0}^{\min(L-1,\,k-1)} \gamma^i q_j(k-i)}{\sum_{i=0}^{\min(L-1,\,k-1)} \gamma^i}, \qquad j = 1, 2, \cdots, M \tag{3.4}$$

记机器在区间 Δk 内的性能表现为 $P_m^m(k)$，其定义为：

$$P_m^m(k) = \sum_{j=1}^{M} \omega_j P_j(k) \tag{3.5}$$

其中，ω_j 为第 j 个评价指标的权重系数，满足：

$$\omega_j \geqslant 0, \qquad \sum_{j=1}^{M} \omega_j = 1 \tag{3.6}$$

结合式（3.4）–（3.6），即可得到机器性能表现的一般形式。后续结合具体任务给出各评价指标及其权重，即可得到相应场景下的机器性能表现。

3.2.2　客观机器能力

基于上一节对机器性能表现的定义，进一步给出客观机器能力 $C_m^m(k)$ 的定义。客观机器能力 $C_m^m(k)$ 用于表示机器在给定任务条件下实际可达到的能力水平，反映机器自身固有的、独立于人类主观认知的真实能力水平。尽管复杂任务

中的能力往往可从多个维度进行衡量，但受人类认知能力限制，在实际交互中，人类通常只能关注关键维度或某种综合性指标，因此本文将客观机器能力建模为标量形式。

客观机器能力的确定通常可结合两类信息：系统运行前的先验知识与运行过程中的性能观测。根据先验信息的充分程度，可考虑如下三类典型情形：

- 先验充分情形：若系统在上线前经过充分测试与验证，机器在特定任务下的能力已得到较为可靠的评估，则可将客观机器能力视为常数，建模为如下形式：

$$C_m^m(k) = C_m^m(0) = C_m^m, \quad \forall k \geqslant 0 \tag{3.7}$$

- 先验缺失情形：若缺乏足够的先验信息，机器能力无法在部署前得到可靠评估，则可利用观测到的机器性能对客观机器能力进行近似，建模为如下形式：

$$C_m^m(k) = \frac{1}{k}\sum_{i=1}^{k} P_m^m(i), \quad k \geqslant 1 \tag{3.8}$$

其中，$P_m^m(i)$ 表示第 i 个区间内观测到的机器性能表现。随着观测数据的积累，该近似将逐渐逼近机器的真实能力水平。

- 先验与后验融合情形：若先验存在但并不充分，初始评估具有一定参考价值，但仍需结合实际运行数据不断修正，则可采用加权递推融合的形式：

$$\begin{cases} C_m^m(0) = C_m^m \\ C_m^m(k) = (1-\alpha_m)C_m^m(k-1) + \alpha_m P_m^m(k), \quad k > 0 \end{cases} \tag{3.9}$$

其中，$\alpha_m \in (0,1)$ 为动态机器因子，用于平衡历史估计与当前观测之间的权重。

在实际任务中，机器的客观能力通常保持相对稳定，不会出现剧烈波动。基于这一事实，对客观机器能力提出如下假设：

假设 3.1 (客观机器能力的平稳性)　客观机器能力相对平稳，其相对变化满足：

$$\frac{|C_m^m(k) - C_m^m(k-1)|}{C_m^m(k-1)} < \epsilon, \quad \forall k > k_0 \tag{3.10}$$

其中，ϵ 为较小的正数，k_0 表示估计初期可能存在短暂波动的阶段。

需要指出的是，式（3.10） 中的 $C_m^m(k)$ 与 $C_m^m(k-1)$ 描述的是真实的客观机器能力，而式（3.8）–（3.9） 给出的则是其估计形式。假设 3.1 表明真实客观机器能力在稳定阶段内变化缓慢，因此其估计值也不应出现显著突变。因此，在采用动态更新时，动态机器因子 α_m 一般应取较小值，以避免当前观测性能的波动引起过大的估计偏移。

3.2.3　主观机器能力

主观机器能力 $C_m^h(k)$ 描述人类在交互过程中对机器实际能力所形成的主观认知与评估。与客观机器能力不同，主观机器能力是一种不可直接测量的内隐认知状态，其形成与演化受到个体经验、先验预期等因素的影响，因而具有显著的个体差异性。基于前文对客观机器能力与机器性能表现的定义，本文将主观机器能力的建模划分为两个阶段：先验认知阶段与动态演化阶段。前者反映人类在任务开始前对机器能力的初始预期，后者描述人类在交互过程中基于机器性能表现不断修正认知的过程。

第一阶段为先验认知的建立。在任务开始之前，人类通常已通过系统说明、训练或历史使用经验等途径获取与机器能力相关的先验信息。在此基础上，人类会结合自身经验、性格特质及专业技能水平等个体因素，对机器能力形成初始的主观判断。由于个体差异的存在，这一先验认知往往并非对客观能力的准确映射，而是伴随着一定程度的主观偏差。为此，引入初始人类因子 $\alpha_i > 0$ 以表示人类在该阶段的认知偏置，初始主观机器能力定义为：

$$C_m^h(0) = \alpha_i C_m^m(0) \tag{3.11}$$

第二阶段为动态演化阶段。在任务执行过程中，人类通过与机器的持续交互不断积累经验，并依据对机器实际表现的观察逐步修正已有认知。考虑到人类对机器性能的感知会受到认知偏置的影响，人类对机器性能的主观感知往往与机器的实际性能存在偏差。为此，引入动态观测因子 $\alpha_d > 0$ 以描述这一映射关系，则第 k 个区间上的主观感知性能 $P_m^h(k)$ 定义为：

$$P_m^h(k) = \alpha_d P_m^m(k), \quad k > 0 \tag{3.12}$$

其中，$P_m^m(k)$ 为机器在第 k 个区间上的实际性能表现。

在获得主观感知性能 $P_m^h(k)$ 后，人类会将其与上一更新时刻的认知状态 $C_m^h(k-1)$ 进行融合。考虑到主观机器能力的更新同时受历史认知与当前观测影响，本文采用线性加权递推结构对主观机器能力的更新过程进行建模。由此，主观机器能力 $C_m^h(k)$ 的更新过程可表示为：

$$C_m^h(k) = (1 - \alpha_h^*(k))C_m^h(k-1) + \alpha_h^*(k)P_m^h(k), \quad k > 0 \tag{3.13}$$

其中，$\alpha_h^*(k)$ 为人类认知更新因子，用于平衡历史认知与当前观测在主观机器能力更新中的相对权重。

由式（3.13） 可得主观机器能力的变化量为：

$$\begin{aligned}
\Delta C_m^h(k) &= C_m^h(k) - C_m^h(k-1) \\
&= \left[(1 - \alpha_h^*(k))C_m^h(k-1) + \alpha_h^*(k)P_m^h(k)\right] - C_m^h(k-1) \\
&= \alpha_h^*(k)\left(P_m^h(k) - C_m^h(k-1)\right)
\end{aligned}$$

令 $\Delta P_m^h(k)$ 表示当前主观感知性能与历史认知状态之间的偏差，即：

$$\Delta P_m^h(k) = P_m^h(k) - C_m^h(k-1) \tag{3.14}$$

则主观机器能力的变化量可等价写为：

$$\Delta C_m^h(k) = \alpha_h^*(k)\Delta P_m^h(k) \tag{3.15}$$

在确定人类认知更新因子 $\alpha_h^*(k)$ 的具体更新机制时，需考虑认知心理学中普遍存在的非对称敏感性现象：人类在面对正面信息与负面信息时，认知更新速率通常存在显著差异。当机器性能表现高于预期时，人类对机器能力的主观判断往往较为谨慎，调整过程相对缓慢；而当机器性能表现低于预期时，人类对机器能力的主观判断则更容易发生变化。这一规律已在人机交互研究中得到验证[73]。因此，主观机器能力的非对称更新主要表现为以下两种情况：

- 性能提升的保守性：当 $\Delta P_m^h(k) > 0$ 时，机器当前表现超过人类的历史认知。由于人类在提高对机器能力的主观评价时通常具有一定的保守性与迟滞性，单次较好表现往往不足以使人类对机器能力的判断发生明显提升。因此，正向更新因子 α_h^+ 应取较小的值，此时主观机器能力的单次变化量为：

$$\Delta C_m^h(k) = \alpha_h^+ \Delta P_m^h(k), \quad \Delta P_m^h(k) > 0 \tag{3.16}$$

- 性能下降的敏感性：当 $\Delta P_m^h(k) < 0$ 时，机器当前表现低于人类的历史认知。出于对潜在风险的警觉，人类对机器失误或性能下降的反应通常更为迅速，从而更快地降低对机器能力的主观评价。因此，负向更新因子 α_h^- 应取较大的值，此时主观机器能力的单次变化量为：

$$\Delta C_m^h(k) = \alpha_h^- \Delta P_m^h(k), \quad \Delta P_m^h(k) < 0 \tag{3.17}$$

根据上述两种非对称演化性质，正向与负向更新因子应满足 $0 < \alpha_h^+ < \alpha_h^- < 1$，以刻画主观机器能力在正负向信息作用下的非对称演化规律。

因此，综合式（3.11） 至式（3.17），主观机器能力的动态演化模型可表示为：

$$\begin{cases} C_m^h(k) = (1-\alpha_h^*(k))C_m^h(k-1) + \alpha_h^*(k)\alpha_d P_m^m(k), \quad \forall k > 0 \\ C_m^h(0) = \alpha_i C_m^m(0) \end{cases} \tag{3.18a}$$

$$\alpha_h^*(k) = \begin{cases} \alpha_h^+, & \Delta P_m^h(k) > 0 \\ 0, & \Delta P_m^h(k) = 0 \\ \alpha_h^-, & \Delta P_m^h(k) < 0 \end{cases} \tag{3.18b}$$

3.3　人机信任演化模型

3.3.1　信任演化模型

在前文定义信任、客观机器能力与主观机器能力的基础上，下面进一步推导信任演化模型具体形式。

根据式（3.1a），离散更新时刻 t_k 的信任为：

$$T_m^h(k) = \frac{C_m^h(k)}{C_m^m(k)}, \quad k \geqslant 0 \tag{3.19}$$

将式（3.12）代入主观机器能力递推式（3.13），可得：

$$C_m^h(k) = (1-\alpha_h^*(k))C_m^h(k-1) + \alpha_h^*(k)\alpha_d\, P_m^m(k), \quad k > 0 \tag{3.20}$$

将式（3.20）代入式（3.19），并利用 $C_m^h(k-1) = T_m^h(k-1)C_m^m(k-1)$，得到：

$$\begin{aligned} T_m^h(k) &= \frac{(1-\alpha_h^*(k))C_m^h(k-1) + \alpha_h^*(k)\alpha_d P_m^m(k)}{C_m^m(k)} \\ &= (1-\alpha_h^*(k))\frac{T_m^h(k-1)C_m^m(k-1)}{C_m^m(k)} + \alpha_h^*(k)\alpha_d \frac{P_m^m(k)}{C_m^m(k)} \end{aligned} \tag{3.21}$$

为简化记号，定义：

$$\rho_m(k) := \frac{C_m^m(k-1)}{C_m^m(k)}, \qquad \phi_m(k) := \frac{P_m^m(k)}{C_m^m(k)} \tag{3.22}$$

则式（3.21）可写为：

$$T_m^h(k) = (1-\alpha_h^*(k))\,\rho_m(k)\,T_m^h(k-1) + \alpha_h^*(k)\alpha_d\,\phi_m(k), \quad k > 0 \tag{3.23}$$

在假设 3.1 下，$C_m^m(k) \approx C_m^m(k-1)$，从而 $\rho_m(k) \approx 1$。进一步定义：

$$\psi_m(k) := \frac{P_m^m(k)}{C_m^m(k-1)} \tag{3.24}$$

其中，$\psi_m(k)$ 为机器性能相对系数，反映第 k 个区间上的实际机器性能相对于 $k-1$ 时刻客观机器能力的变化程度。

由此有 $\phi_m(k) \approx \psi_m(k)$，式（3.23）近似为：

$$T_m^h(k) \approx (1-\alpha_h^*(k))T_m^h(k-1) + \alpha_h^*(k)\alpha_d\,\psi_m(k), \quad k > 0 \tag{3.25}$$

由式（3.1b）可知：

$$T_m^h(t) = T_m^h(k), \quad t \in [t_k, t_{k+1}),\ k \geqslant 0 \tag{3.26}$$

同时，由式（3.11）与式（3.19）可得信任的初始值：

$$T_m^h(0) = \frac{C_m^h(0)}{C_m^m(0)} = \alpha_i \tag{3.27}$$

综上所述，可得到信任演化模型的一般形式为：

$$\begin{cases} T_m^h(k) = (1-\alpha_h^*(k))\,T_m^h(k-1) + \alpha_h^*(k)\alpha_d\,\psi_m(k), \quad k>0 \\ T_m^h(t) = T_m^h(k), \quad t\in[t_k, t_{k+1}),\ k\geqslant 0 \\ T_m^h(0) = \alpha_i \end{cases} \tag{3.28a}$$

其中：

$$\alpha_h^*(k) = \begin{cases} \alpha_h^+, & \Delta P_m^h(k) > 0 \\ 0, & \Delta P_m^h(k) = 0 \\ \alpha_h^-, & \Delta P_m^h(k) < 0 \end{cases} \tag{3.28b}$$

$$\Delta P_m^h(k) = \alpha_d P_m^m(k) - T_m^h(k-1) C_m^m(k-1) \tag{3.28c}$$

为便于后文表述，将式（3.28） 所示的信任演化模型命名为性能驱动的信任演化模型（Performance-Driven Trust Evolution Model, PDTEM），后续均以 PDTEM 指代该模型。

3.3.2　模型理论分析

本节对 PDTEM 的非负性、有界性、收敛性、扰动鲁棒性及一步更新非对称性进行分析，以说明模型的基本动态特性。

为便于表述，定义输入量为：

$$u_k := \alpha_d\,\psi_m(k) \tag{3.29}$$

则式（3.28a） 可写为：

$$T_m^h(k) = (1-\alpha_h^*(k))\,T_m^h(k-1) + \alpha_h^*(k)\,u_k, \quad k>0 \tag{3.30}$$

性质 3.1 (非负性保持)**.** *若 $T_m^h(0)\geqslant 0$ 且对所有 $k\geqslant 1$ 都有 $u_k\geqslant 0$，则对所有 $k\geqslant 0$ 都有：*

$$T_m^h(k) \geqslant 0 \tag{3.31}$$

证明　由式（3.28b） 以及 $0<\alpha_h^+<\alpha_h^-<1$ 可知，对任意 k 都有 $\alpha_h^*(k)\in[0,1)$。因此，由式（3.30） 可得：

$$T_m^h(k) = (1-\alpha_h^*(k))\,T_m^h(k-1) + \alpha_h^*(k)\,u_k \tag{3.32}$$

即 $T_m^h(k)$ 是 $T_m^h(k-1)$ 与 u_k 的非负线性组合。

下面用数学归纳法证明对所有 $k\geqslant 0$ 都有 $T_m^h(k)\geqslant 0$。

当 $k=0$ 时，由条件 $T_m^h(0)\geqslant 0$，命题成立。

假设当 $k=n-1$ 时有 $T_m^h(n-1)\geqslant 0$。又由于 $u_n\geqslant 0$ 且 $\alpha_h^*(n)\in[0,1)$，从而可得：

$$(1-\alpha_h^*(n))\,T_m^h(n-1)\geqslant 0, \qquad \alpha_h^*(n)\,u_n\geqslant 0 \tag{3.33}$$

将式（3.33） 代入式（3.32） 可得：

$$T_m^h(n) = (1 - \alpha_h^*(n))\, T_m^h(n-1) + \alpha_h^*(n)\, u_n \geqslant 0 \tag{3.34}$$

即当 $k = n$ 时命题也成立。

因此，由数学归纳法可知，对所有 $k \geqslant 0$ 都有 $T_m^h(k) \geqslant 0$，即式（3.31） 成立。 ∎

性质 3.2 (有界性). *若存在常数 $\underline{P} \leqslant \overline{P} < \infty$，使得 $\psi_m(k) \in [\underline{P}, \overline{P}]$，则信任序列 $\{T_m^h(k)\}$ 始终有界，并满足：*

$$T_m^h(k) \in \left[\min\{T_m^h(0), \alpha_d \underline{P}\},\ \max\{T_m^h(0), \alpha_d \overline{P}\} \right], \quad \forall k \geqslant 0 \tag{3.35}$$

证明　由定义 $u_k = \alpha_d \psi_m(k)$ 以及 $\psi_m(k) \in [\underline{P}, \overline{P}]$，可得：

$$u_k \in [\underline{u}, \overline{u}], \qquad \underline{u} := \alpha_d \underline{P}, \quad \overline{u} := \alpha_d \overline{P} \tag{3.36}$$

由式（3.30） 可知，$T_m^h(k)$ 为 $T_m^h(k-1)$ 与 u_k 的凸组合，因此有：

$$\min\{T_m^h(k-1), u_k\} \leqslant T_m^h(k) \leqslant \max\{T_m^h(k-1), u_k\} \tag{3.37}$$

再由 $u_k \in [\underline{u}, \overline{u}]$，可将式（3.37） 进一步放缩为：

$$\min\{T_m^h(k-1), \underline{u}\} \leqslant T_m^h(k) \leqslant \max\{T_m^h(k-1), \overline{u}\} \tag{3.38}$$

令：

$$L := \min\{T_m^h(0), \underline{u}\}, \qquad U := \max\{T_m^h(0), \overline{u}\} \tag{3.39}$$

显然，$T_m^h(0) \in [L, U]$。

下面用数学归纳法证明对所有 $k \geqslant 0$ 都有 $T_m^h(k) \in [L, U]$。

当 $k = 0$ 时，由式（3.39） 可知 $T_m^h(0) \in [L, U]$，结论成立。

假设当 $k = n-1$ 时有 $T_m^h(n-1) \in [L, U]$。又由于 $\underline{u}, \overline{u} \in [L, U]$，由式（3.38） 可得：$L \leqslant T_m^h(n) \leqslant U$，即 $T_m^h(n) \in [L, U]$。

因此，由数学归纳法可知，对所有 $k \geqslant 0$ 都有 $T_m^h(k) \in [L, U]$。结合式（3.39），即可得到式（3.35） 成立。

这说明在机器性能相对系数有界时，信任演化始终保持在一个有限区间内，不会出现无界发散。 ∎

性质 3.3 (收敛性). *若存在常数 $\bar{P} > 0$ 与整数 k_1，使得对所有 $k \geqslant k_1$ 都有：*

$$\psi_m(k) = \bar{P} \tag{3.40}$$

定义平衡点：

$$T^* := \alpha_d \bar{P} \tag{3.41}$$

以及误差：

$$e(k) := T_m^h(k) - T^* \tag{3.42}$$

则有：

$$\lim_{k\to\infty} e(k) = 0 \tag{3.43}$$

从而：

$$\lim_{k\to\infty} T_m^h(k) = T^* \tag{3.44}$$

证明　由式（3.29）与式（3.40）可得：

$$u_k = \alpha_d \psi_m(k) = \alpha_d \bar{P} = T^*, \qquad k \geqslant k_1 \tag{3.45}$$

因此，由式（3.30）可得：

$$T_m^h(k) = (1 - \alpha_h^*(k)) T_m^h(k-1) + \alpha_h^*(k) T^*, \qquad k \geqslant k_1 + 1 \tag{3.46}$$

两边同减 T^*，可得：

$$e(k) = (1 - \alpha_h^*(k)) e(k-1), \qquad k \geqslant k_1 + 1 \tag{3.47}$$

另一方面，由 $\psi_m(k) = P_m^m(k)/C_m^m(k-1)$ 及式（3.40）可得：

$$P_m^m(k) = \bar{P}\, C_m^m(k-1), \qquad k \geqslant k_1 \tag{3.48}$$

将式（3.48）代入式（3.28c），可得：

$$\Delta P_m^h(k) = C_m^m(k-1)\left(\alpha_d \bar{P} - T_m^h(k-1)\right) = -C_m^m(k-1) e(k-1), \qquad k \geqslant k_1+1 \tag{3.49}$$

由于 $C_m^m(k-1) > 0$，故 $\Delta P_m^h(k)$ 与 $-e(k-1)$ 同号。

下面分三种情况进行讨论。

情形一：若 $e(k_1) < 0$，则 $T_m^h(k_1) < T^*$。由式（3.49）可知：

$$\Delta P_m^h(k_1 + 1) > 0 \tag{3.50}$$

由式（3.28b）可得：

$$\alpha_h^*(k_1 + 1) = \alpha_h^+ \tag{3.51}$$

将式（3.51）代入式（3.47），可得：

$$e(k_1 + 1) = (1 - \alpha_h^+) e(k_1) \tag{3.52}$$

由于 $0 < 1 - \alpha_h^+ < 1$ 且 $e(k_1) < 0$，故 $e(k_1 + 1) < 0$。由递推关系可得，对所有 $k \geqslant k_1$ 都有：

$$e(k) = (1 - \alpha_h^+)^{k-k_1} e(k_1) < 0 \tag{3.53}$$

因此：

$$\lim_{k\to\infty} e(k) = 0 \tag{3.54}$$

情形二：若 $e(k_1)>0$，则 $T_m^h(k_1)>T^*$。由式（3.49） 可知：

$$\Delta P_m^h(k_1+1)<0 \tag{3.55}$$

由式（3.28b） 可得：

$$\alpha_h^*(k_1+1)=\alpha_h^- \tag{3.56}$$

将式（3.56） 代入式（3.47），可得：

$$e(k_1+1)=(1-\alpha_h^-)e(k_1) \tag{3.57}$$

由于 $0<1-\alpha_h^-<1$ 且 $e(k_1)>0$，故 $e(k_1+1)>0$。由递推关系可得，对所有 $k\geqslant k_1$ 都有：

$$e(k)=(1-\alpha_h^-)^{k-k_1}e(k_1)>0 \tag{3.58}$$

因此：

$$\lim_{k\to\infty}e(k)=0 \tag{3.59}$$

情形三：若 $e(k_1)=0$，即 $T_m^h(k_1)=T^*$。则由式（3.49） 可得：

$$\Delta P_m^h(k_1+1)=0 \tag{3.60}$$

因此，由式（3.28b） 可得：

$$\alpha_h^*(k_1+1)=0 \tag{3.61}$$

将式（3.61） 代入式（3.47），可得：

$$e(k_1+1)=0 \tag{3.62}$$

由此可得，对所有 $k\geqslant k_1$ 都有：

$$e(k)=0 \tag{3.63}$$

因此：

$$\lim_{k\to\infty}e(k)=0 \tag{3.64}$$

综合上述三种情况可得：

$$\lim_{k\to\infty}e(k)=0 \tag{3.65}$$

再结合式（3.42），可得：

$$\lim_{k\to\infty}T_m^h(k)=T^* \tag{3.66}$$

因此，信任序列 $\{T_m^h(k)\}$ 收敛到平衡点 T^*。

这一结果表明，当机器性能趋于稳定时，信任水平也会趋于稳定。 ∎

性质 3.4 (扰动鲁棒性). 若存在常数 $\bar{P}>0$、$\bar{\delta}\geqslant 0$ 与整数 k_1，使得对所有 $k\geqslant k_1$ 都有：

$$\psi_m(k)=\bar{P}+\delta_k,\qquad |\delta_k|\leqslant\bar{\delta} \tag{3.67}$$

且对所有 $k\geqslant k_1$ 都有：

$$\Delta P_m^h(k)\neq 0 \tag{3.68}$$

则信任相对平衡点 $T^*=\alpha_d\bar{P}$ 的稳态误差满足：

$$\limsup_{k\to\infty}\left|T_m^h(k)-\alpha_d\bar{P}\right|\leqslant\frac{\alpha_h^-\alpha_d\bar{\delta}}{\alpha_h^+} \tag{3.69}$$

证明　令：

$$T^*:=\alpha_d\bar{P},\qquad e(k):=T_m^h(k)-T^* \tag{3.70}$$

由式（3.29）与式（3.67）可得：

$$u_k=\alpha_d\psi_m(k)=\alpha_d(\bar{P}+\delta_k)=T^*+\alpha_d\delta_k \tag{3.71}$$

将式（3.71）代入式（3.30），并两边同减 T^*，得到：

$$\begin{aligned}e(k)&=(1-\alpha_h^*(k))\,T_m^h(k-1)+\alpha_h^*(k)\left(T^*+\alpha_d\delta_k\right)-T^*\\&=(1-\alpha_h^*(k))\left(T_m^h(k-1)-T^*\right)+\alpha_h^*(k)\alpha_d\delta_k\\&=(1-\alpha_h^*(k))\,e(k-1)+\alpha_h^*(k)\alpha_d\delta_k\end{aligned} \tag{3.72}$$

由式（3.68）、式（3.28b）以及 $0<\alpha_h^+<\alpha_h^-<1$ 可知，对所有 $k\geqslant k_1$ 都有：

$$\alpha_h^+\leqslant\alpha_h^*(k)\leqslant\alpha_h^- \tag{3.73}$$

对式（3.72）两边取绝对值，并结合 $|\delta_k|\leqslant\bar{\delta}$ 与式（3.73），可得：

$$|e(k)|\leqslant(1-\alpha_h^+)|e(k-1)|+\alpha_h^-\alpha_d\bar{\delta} \tag{3.74}$$

对式（3.74）从 k_1 起递推展开，得到：

$$\begin{aligned}|e(k)|&\leqslant(1-\alpha_h^+)^{k-k_1}|e(k_1)|+\sum_{j=k_1+1}^{k}(1-\alpha_h^+)^{k-j}\alpha_h^-\alpha_d\bar{\delta}\\&=(1-\alpha_h^+)^{k-k_1}|e(k_1)|+\alpha_h^-\alpha_d\bar{\delta}\sum_{r=0}^{k-k_1-1}(1-\alpha_h^+)^r\end{aligned} \tag{3.75}$$

由几何级数求和公式可得：

$$\sum_{r=0}^{k-k_1-1}(1-\alpha_h^+)^r=\frac{1-(1-\alpha_h^+)^{k-k_1}}{\alpha_h^+}\leqslant\frac{1}{\alpha_h^+} \tag{3.76}$$

将式（3.76）代入式（3.75），可得：

$$|e(k)|\leqslant(1-\alpha_h^+)^{k-k_1}|e(k_1)|+\frac{\alpha_h^-\alpha_d\bar{\delta}}{\alpha_h^+} \tag{3.77}$$

令 $k \to \infty$，由于第一项趋于 0，从而得到式（3.69）。

由此可见，当机器性能存在有界扰动时，信任相对平衡点的偏差仍保持有界，说明 PDTEM 对性能波动具有一定鲁棒性。 ∎

性质 3.5 (一步更新非对称性). 设 $\alpha_h^- > \alpha_h^+$。在相同初始信任状态 T_0 下，分别记正向输入与负向输入作用后的一步更新结果为 T^+ 与 T^-。若两次观测输入满足：

$$u^+ - T_0 = \Delta, \qquad u^- - T_0 = -\Delta \tag{3.78}$$

其中 $\Delta > 0$，且分别触发正向更新与负向更新，则一步更新幅度满足：

$$\left|T^- - T_0\right| > \left|T^+ - T_0\right| \tag{3.79}$$

即在相同幅值的正负偏差作用下，负向偏差引起的一步信任变化幅度更大。

证明　由式（3.30）可知，一步更新增量满足：

$$T_m^h(k) - T_m^h(k-1) = \alpha_h^*(k)\left(u_k - T_m^h(k-1)\right) \tag{3.80}$$

在正向情形中，由式（3.78）可知 $u^+ - T_0 = \Delta$，且此时 $\alpha_h^*(k) = \alpha_h^+$。因此：

$$T^+ - T_0 = \alpha_h^+ \Delta, \qquad |T^+ - T_0| = \alpha_h^+ \Delta \tag{3.81}$$

在负向情形中，由式（3.78）可知 $u^- - T_0 = -\Delta$，且此时 $\alpha_h^*(k) = \alpha_h^-$。因此：

$$T^- - T_0 = -\alpha_h^- \Delta, \qquad |T^- - T_0| = \alpha_h^- \Delta \tag{3.82}$$

由于 $\alpha_h^- > \alpha_h^+$ 且 $\Delta > 0$，可得：

$$\alpha_h^- \Delta > \alpha_h^+ \Delta \tag{3.83}$$

结合式（3.81）与式（3.82），可得 $|T^- - T_0| > |T^+ - T_0|$，故式（3.79）成立。

这体现了 PDTEM 对负面信息更敏感、对正面信息更保守的更新特征。 ∎

3.3.3　模型参数估计方法

PDTEM 的核心人因参数为 $\alpha_i, \alpha_d, \alpha_h^+, \alpha_h^-$。其中，$\alpha_i$ 为初始参数，可通过系统上线前的培训、说明与先验评估获得；$\alpha_d, \alpha_h^+, \alpha_h^-$ 为动态演化参数，需要基于人机交互数据进行估计。下面给出基于分段最小二乘的参数估计方法。

首先需要对样本进行分组。由于式（3.28c）中的偏差量 $\Delta P_m^h(k)$ 含有未知参数 α_d，若直接依据其符号对样本分组，会形成参数分组与参数估计之间的循环依赖。为避免这一问题，使用可直接观测的信任增量作为样本分组依据，定义：

$$\Delta T_m^h(k) := T_m^h(k) - T_m^h(k-1), \quad k > 0 \tag{3.84}$$

由式（3.30）可得：

$$\Delta T_m^h(k) = \alpha_h^*(k)\left(u_k - T_m^h(k-1)\right) \tag{3.85}$$

由于 $\alpha_h^*(k) \geqslant 0$，因此当 $\Delta T_m^h(k) \neq 0$ 时，$\Delta T_m^h(k)$ 与 $u_k - T_m^h(k-1)$ 具有相同符号。故可利用信任增量的符号区分正向更新与负向更新。因此，将样本划分为两类：

$$\mathcal{D}^+ = \{k : \Delta T_m^h(k) > 0\}, \qquad \mathcal{D}^- = \{k : \Delta T_m^h(k) < 0\} \tag{3.86}$$

其中，$\mathcal{D}^+$ 表示信任上升样本集，$\mathcal{D}^-$ 表示信任下降样本集。信任保持不变的样本由于对参数辨识贡献较弱，不将其纳入回归估计。

由上述符号关系可知，当 $k \in \mathcal{D}^+$ 时，当前样本对应正向更新情形，故有 $\alpha_h^*(k) = \alpha_h^+$；当 $k \in \mathcal{D}^-$ 时，当前样本对应负向更新情形，故有 $\alpha_h^*(k) = \alpha_h^-$。因此，式（3.28a） 在两类样本上可分别写为：

$$T_m^h(k) = (1 - \alpha_h^{\pm}) T_m^h(k-1) + \alpha_h^{\pm} \alpha_d \, \psi_m(k), \quad k \in \mathcal{D}^{\pm} \tag{3.87}$$

为将上式写成线性回归形式，对任意样本 $k \in \mathcal{D}^{\pm}$ 定义输入向量与输出为：

$$x(k) := \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} := \begin{bmatrix} T_m^h(k-1) \\ \psi_m(k) \end{bmatrix}, \qquad y(k) := T_m^h(k) \tag{3.88}$$

并定义回归参数向量：

$$\theta^{\pm} := \begin{bmatrix} \theta_1^{\pm} \\ \theta_2^{\pm} \end{bmatrix} := \begin{bmatrix} 1 - \alpha_h^{\pm} \\ \alpha_h^{\pm} \alpha_d \end{bmatrix} \tag{3.89}$$

则式（3.87） 可写为：

$$y(k) = \theta_1^{\pm} x_1(k) + \theta_2^{\pm} x_2(k), \qquad k \in \mathcal{D}^{\pm} \tag{3.90}$$

设正向样本数与负向样本数分别为 n^+ 与 n^-。对样本集 $\mathcal{D}^{\pm}$，构造设计矩阵与观测向量：

$$X^{\pm} := \begin{bmatrix} x_1^{\pm}(1) & x_2^{\pm}(1) \\ \vdots & \vdots \\ x_1^{\pm}(n^{\pm}) & x_2^{\pm}(n^{\pm}) \end{bmatrix}, \qquad Y^{\pm} := \begin{bmatrix} y^{\pm}(1) \\ \vdots \\ y^{\pm}(n^{\pm}) \end{bmatrix} \tag{3.91}$$

并采用最小二乘准则估计参数向量：

$$\hat{\theta}^{\pm} = \arg\min_{\theta^{\pm}} \| Y^{\pm} - X^{\pm} \theta^{\pm} \|_2^2 \tag{3.92}$$

其中 $\hat{\theta}^{\pm} = [\hat{\theta}_1^{\pm}, \hat{\theta}_2^{\pm}]^{\top}$。

由式（3.89） 可得，正向与负向更新率可分别恢复为：

$$\hat{\alpha}_h^{\pm} = 1 - \hat{\theta}_1^{\pm} \tag{3.93}$$

进一步，由 $\theta_2^{\pm} = \alpha_h^{\pm} \alpha_d$ 可得到动态观测因子 α_d 的两组估计值：

$$\widehat{\alpha_d}^{\pm} = \frac{\hat{\theta}_2^{\pm}}{\hat{\alpha}_h^{\pm}} \tag{3.94}$$

理论上，α_d 为与正负阶段无关的公共参数，因此可利用正向与负向估计结果进行一致性分析。定义一致性指标为：

$$\vartheta_d := \frac{\widehat{\alpha_d}^{+}}{\widehat{\alpha_d}^{-}} \tag{3.95}$$

当 $\vartheta_d \approx 1$ 时，说明正向与负向阶段对 α_d 的估计较为一致，模型在不同阶段具有较好的统一解释能力。为进一步获得单一的公共参数估计值，可对两段估计结果按样本量加权融合，得到：

$$\hat{\alpha}_d := \frac{n^{+}\widehat{\alpha_d}^{+} + n^{-}\widehat{\alpha_d}^{-}}{n^{+} + n^{-}} \tag{3.96}$$

3.4　人机信任模型验证

为验证 PDTEM 的有效性，本节围绕无人机自主穿越圆环任务设计了观察式人机信任实验。实验通过构建与任务特性相匹配的机器性能表现，并基于离线评估确定客观机器能力，同时收集受试者在持续观察机器行为过程中的主观反馈数据，用于模型参数估计与有效性验证。实验平台基于 PyFlyt 无人机仿真环境进行二次开发[74]，在此基础上构建了面向圆环穿越任务的实验场景。本文各仿真实验均在同一计算平台上完成，实验计算机搭载 Intel Core i9-14900KF 处理器（24 核，基础频率 3.2 GHz，最大睿频 6.0 GHz）和 NVIDIA GeForce RTX 4090 显卡（24 GB 显存），计算平台采用 Ubuntu 操作系统。

3.4.1　任务场景描述

无人机穿越圆环任务要求无人机在三维空间中自主规划飞行路径，并以合理的姿态从指定圆环内部通过，是无人机自主飞行研究中的一项典型测试任务。该任务对无人机的多项核心能力提出了较高要求：一方面，无人机需要具备准确的空间感知与目标定位能力，以识别圆环在三维空间中的位置与朝向；另一方面，无人机还需在飞行过程中完成实时轨迹跟踪与精细的机动控制，以确保最终能够无碰撞地穿越圆环。该任务目标明确、执行结果易于判定，适合作为人机信任建模的观察载体。

选取上述无人机穿越圆环任务作为信任建模的仿真载体。在所构建的 PyFlyt 仿真场景中，三维空间内设置 3 个位置各异的圆环，其中标定 1 个为目标圆环。无人机从给定初始区域内的随机位置起飞，在控制策略驱动下自主完成目标圆环穿越任务。图 3.1 给出了所构建的仿真环境示意图，展示了无人机与 3 个候选圆环的位置分布，以及目标圆环的标定方式。

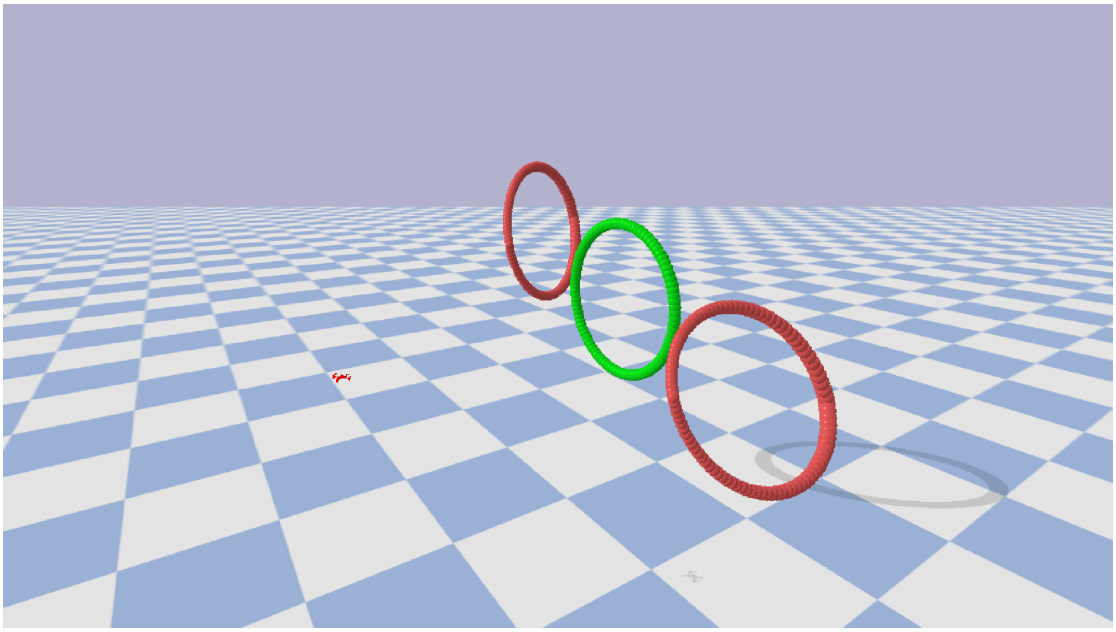


图 3.1　无人机自主穿越圆环任务仿真环境示意图

3.4.2　机器性能表现与客观机器能力构建

结合前文机器性能表现与客观机器能力的一般定义，下面给出圆环穿越任务下的具体构造方式。在该任务中，将第 k 轮任务对应为第 k 个信任更新区间，机器性能表现由任务成功指标和中心通过质量共同计算。

第 k 轮任务的成功指标定义为：

$$q_{\mathrm{s}}(k)=\begin{cases}1, & \text{任务成功}\\ 0, & \text{任务失败}\end{cases} \tag{3.97}$$

若第 k 轮任务成功，记无人机轨迹与目标圆环所在平面的交点到圆环中心的距离为 d_k，圆环半径为 R，则归一化中心偏差定义为：

$$\bar{d}_k=\frac{d_k}{R} \tag{3.98}$$

在任务成功的情况下，交点越接近圆环中心，说明穿越质量越高；若任务失败，则中心通过质量记为 0。第 k 轮任务的中心通过质量定义为：

$$q_{\mathrm{c}}(k)=\begin{cases}1-\bar{d}_k, & \text{任务成功}\\ 0, & \text{任务失败}\end{cases} \tag{3.99}$$

当任务成功时，$0\leqslant d_k\leqslant R$，因此 $q_{\mathrm{c}}(k)\in[0,1]$。

由式（3.4）和式（3.5），该任务下的机器性能表现可写为：

$$P_m^m(k)=w_s P_{\mathrm{s}}(k)+w_q P_{\mathrm{c}}(k),\qquad w_s+w_q=1 \tag{3.100}$$

其中，$P_{\mathrm{s}}(k)$ 和 $P_{\mathrm{c}}(k)$ 分别表示任务成功指标与中心通过质量在第 k 个更新区间上的表现值，w_s 和 w_q 为对应权重。实验中取 $w_s=0.7$、$w_q=0.3$，时间窗口长度取 $L=5$，衰减因子取 $\gamma=0.5$。

实验中，无人机执行穿越圆环任务的自主控制策略由 PPO 算法训练得到。对于客观机器能力，考虑到该控制策略在整个实验过程中保持不变，且任务设置与评价标准一致，因此可按式（3.7） 将其建模为常数。为此，在与正式实验一致的任务设置下，预先对无人机自主穿越策略进行了 N 轮离线评估。记任务成功指标与中心通过质量的平均值分别为：

$$\bar{q}_{\mathrm{s}} = \frac{1}{N}\sum_{k=1}^{N} q_{\mathrm{s}}(k), \qquad \bar{q}_{\mathrm{c}} = \frac{1}{N}\sum_{k=1}^{N} q_{\mathrm{c}}(k) \tag{3.101}$$

则该任务下的客观机器能力可表示为：

$$C_m^m(k) = C_m^m = w_s \bar{q}_{\mathrm{s}} + w_q \bar{q}_{\mathrm{c}}, \qquad k \geqslant 0 \tag{3.102}$$

离线评估结果表明，该策略在当前任务上的平均成功率为 84.60%，平均中心通过质量为 0.4756。在 $w_s = 0.7$、$w_q = 0.3$ 的设置下，得到客观机器能力 $C_m^m = 0.7349$。

3.4.3　实验设置

实验包括预实验和正式实验两个阶段。预实验用于个性化模型参数辨识，正式实验用于模型在线预测验证。实验共在校内招募 10 名受试者参与，受试者年龄分布为 22–28 岁，平均年龄约为 25 岁，均为在读硕士生或博士生。所有受试者在实验开始前均阅读统一的实验说明，了解任务规则、观察方式及评分要求。

实验场景中设置 3 个候选圆环，其中 1 个为目标圆环。圆环位置在整个实验过程中保持不变，每轮任务的目标圆环从 3 个候选圆环中随机选取。无人机自主完成目标圆环穿越任务，受试者仅观察无人机执行过程，不参与实际控制。实验中，受试者在每轮任务结束后给出对无人机主观机器能力的评价，评分采用 0.1 至 1.0 的十级离散量表。

在预实验阶段，受试者连续观察无人机自主穿越目标圆环的 100 轮任务，并在每轮任务结束后给出当前主观机器能力评分。结合各轮任务对应的区间机器性能表现，对每位受试者的个体化模型参数进行辨识。在正式实验阶段，受试者首先给出本阶段起始时刻的主观机器能力评分，作为状态递推的初始观测值，随后继续完成 100 轮观察任务，并在每轮任务结束后给出当前主观机器能力评分以及是否需要人工介入的主观意愿。

实验中记录的数据包括每轮任务是否成功、机器性能表现、主观机器能力评分以及干预意愿等。正式实验中，利用预实验阶段辨识得到的模型参数，并结合每轮任务对应的机器性能表现，对受试者信任状态进行预测，再将模型计算得到的信任值与由量表评分计算得到的信任值进行比较，以评估模型的在线预测能力。为验证所提模型中非对称更新机制的有效性，本节选取对称 PDTEM 作为对

比模型。该模型保留原模型的基本结构，但采用统一更新率，不再区分信任上升与下降阶段的更新差异。

3.4.4　实验结果与分析

表 3.1 给出了随机选取的四名受试者基于预实验数据得到的参数估计结果，以展示个性化模型的辨识情况。四名受试者均满足 $\alpha_h^- > \alpha_h^+$，这表明，当机器性能下降时，主观机器能力的调整幅度大于机器性能提升时的调整幅度。该现象与前文关于信任具有非对称更新特性的分析一致，说明 PDTEM 能够较好反映受试者在机器性能上升与下降两种情况下不同的认知变化规律。此外，四名受试者的一致性指标 ϑ_d 均接近于 1，表明主观机器能力上升与下降阶段对应的动态观测因子估计结果较为接近。这说明该模型在体现信任正负更新非对称特征的同时，参数估计也具有较好一致性。

表 3.1　四名受试者的 PDTEM 参数估计结果

受试者	α_d	α_h^+	α_h^-	ϑ_d
1	1.2412	0.7465	0.8722	0.9522
2	1.2624	0.4895	0.9195	0.9019
3	1.4905	0.5270	0.7050	0.9327
4	1.4127	0.4832	0.9900	0.9429

图 3.2 给出了受试者 4 在正式实验阶段的信任预测结果。从图中可见，两种模型均能够反映信任变化的整体趋势，但 PDTEM 与真实信任曲线更为接近，尤其在信任快速上升、下降以及局部波动较为明显的区间，对变化幅度和变化方向的刻画更为准确。相比之下，对称 PDTEM 虽然能够描述总体变化趋势，但由于未区分正向与负向更新过程，在信任变化方向发生切换以及局部峰值、低谷附近，预测结果与真实信任之间仍存在一定偏差。这说明在该受试者数据上，区分正向更新与负向更新过程有助于更准确地刻画信任的动态变化规律。

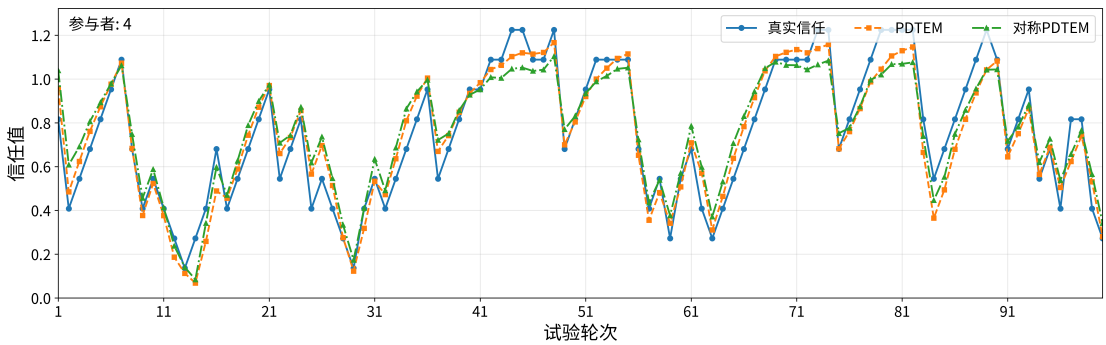


图 3.2　两种信任模型的预测结果对比图

表 3.2 给出了两种信任模型的预测效果对比。无论在受试者 4 的个体结果还

是总体结果上，PDTEM 的 RMSE 和 MAE 均更低，R^2 更高，表明其对信任随机器性能表现变化的动态演化过程具有更好的刻画能力。与此同时，PDTEM 的平均预测信任值也更接近平均真实信任值，表明其整体预测偏差较小。结合图 3.2 与表 3.2 可以看出，引入正向与负向非对称更新机制后，模型在整体趋势预测和局部变化响应两个方面均表现出更好的效果。

表 3.2　两种信任模型的预测效果对比

受试者	模型	RMSE	MAE	R^2	预测信任均值	真实信任均值
4	PDTEM	**0.0898**	**0.0729**	**0.9098**	**0.7471**	0.7552
4	对称 PDTEM	0.1066	0.0900	0.8728	0.7681	0.7552
总体	PDTEM	**0.0939**	**0.0771**	**0.8414**	**0.7514**	0.7501
总体	对称 PDTEM	0.0988	0.0812	0.8310	0.7549	0.7501

为了进一步分析信任状态与人类干预行为之间的关系，根据量表对应的 10 个等级对样本进行统计。记第 n 个量表等级对应的信任区间为 ΔT_n，该区间内的样本总数为 N_n，其中人类决定干预的次数为 N_n^i，则该区间内的人类干预频率定义为：

$$f_n = \frac{N_n^i}{N_n} \tag{3.103}$$

利用四分位距法剔除异常实验数据后，得到不同信任区间下的人类干预频率，如图 3.3 所示。

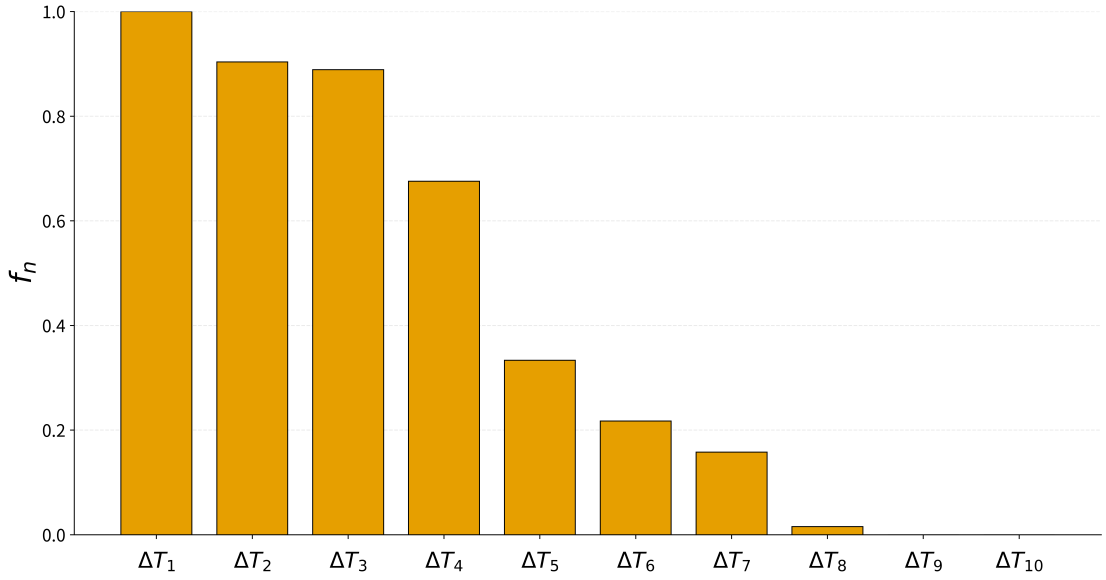


图 3.3　不同信任区间下的人类干预频率图

从图 3.3 可以发现，人类干预频率随信任水平提高整体呈下降趋势。低信任区间内的人类干预频率更高，说明当人类对无人机能力缺乏足够信任时，更倾向

于主动介入控制；随着信任水平提高，干预频率逐渐降低，此时人类更倾向于保持机器自主运行，人工干预相应减少。由此可见，信任状态会对人类的干预决策产生重要影响。若在协同控制过程中不考虑信任状态，人类可能因信任不足而频繁干预，从而削弱机器自主能力的有效发挥，并增加人机冲突发生的可能性。因此，在后续人-无人机协同方法设计中有必要引入信任状态。

3.5　本章小结

本章提出了机器性能驱动的信任演化模型（PDTEM），并从理论层面分析了模型的性质，同时通过人机交互实验验证了模型的有效性。实验结果表明，该模型能够较好地刻画人类信任变化的整体趋势及局部波动，且模型参数辨识结果体现出非对称更新特征，动态观测因子在不同更新阶段的估计结果较为接近。对人类干预行为的进一步分析发现，信任水平越低，人类干预频率越高。PDTEM 能够较好刻画无人机任务场景下的人机信任演化规律，为后续融合信任的人-无人机协同方法设计提供模型基础。

第 4 章　融合信任的最小干预共享控制策略

共享控制为人-无人机协同降落任务提供了一种有效的解决思路，其核心在于协调人类输入与机器辅助之间的关系。在任务过程中，系统辅助不足可能导致任务失败，而即使遵循最优策略，过度的辅助也会让人类操作员认为系统不可信[75]。因此，如何在保障任务安全的前提下实现适度干预，是共享控制方法设计中的关键。

为此，面向协同降落任务，本章提出一种融合信任的最小干预共享控制策略，将人类信任状态引入人机控制权的仲裁过程，并根据信任水平动态调节机器对人类输入的干预强度。本章围绕降落任务下机器能力与性能表现的构建、基于深度强化学习的动作价值评估与候选动作集构造、信任驱动的最小干预仲裁方法，以及仿真实验设计与结果验证等内容展开研究。

4.1　降落任务下机器能力构建

在第三章构建 PDTEM 的基础上，本节结合人-无人机协同降落任务特点构建任务评价指标，以计算客观机器能力 $\boldsymbol{C}_m^m$ 与机器性能表现 $\boldsymbol{P}_m^m(k)$。

4.1.1　任务评价指标

人机协同降落任务要求无人机不仅能够正确理解人类意图，还需精确降落至目标平台，并在整个执行过程中保证飞行安全。因此，本节从意图推理、降落精度和任务安全性三个方面构建协同降落任务的评价指标。设候选着陆平台集合为：

$$\mathcal{G} = \{g_1, g_2, \cdots, g_M\} \tag{4.1}$$

记第 k 次任务目标平台为 g_k^*。

1. 意图推理指标

意图推理指标 $q_{\mathrm{I}}(k)$ 用于衡量机器在任务结束时能否正确推理出目标平台。记第 k 次任务的完整观测历史为 $\boldsymbol{H}_k$，则任务结束时机器推理得到的目标平台为：

$$\hat{g}_k = \arg\max_{g_m \in \mathcal{G}} p(g_m \mid \boldsymbol{H}_k) \tag{4.2}$$

意图推理指标定义为：

$$q_{\mathrm{I}}(k) = \begin{cases} 1, & \hat{g}_k = g_k^* \\ 0, & \hat{g}_k \neq g_k^* \end{cases} \tag{4.3}$$

2. 降落精度指标

降落精度指标 $q_{\mathrm{L}}(k)$ 用于评价成功着陆后落点相对于目标平台中心的接近程度。记第 k 次任务的着陆点位置为 $\boldsymbol{p}_k$，目标平台中心位置为 $\boldsymbol{p}_{g_k^*}$，目标平台的有效半径为 $\boldsymbol{R}_g$，则位置偏差定义为：

$$d_{L,k} = \left\| \boldsymbol{p}_k - \boldsymbol{p}_{g_k^*} \right\|_2 \tag{4.4}$$

据此，降落精度指标定义为：

$$q_{\mathrm{L}}(k) = \begin{cases} 1 - \min\left(\dfrac{d_{L,k}}{R_g}, 1 \right), & \text{任务成功} \\ 0, & \text{任务失败} \end{cases} \tag{4.5}$$

3. 任务安全性指标

任务安全性指标 $q_{\mathrm{S}}(k)$ 用于评价第 k 次任务执行过程中是否发生碰撞。若任务过程中未发生碰撞，则记为安全；否则记为不安全。因此，任务安全性指标定义为：

$$q_{\mathrm{S}}(k) = \begin{cases} 1, & \text{未发生碰撞} \\ 0, & \text{发生碰撞} \end{cases} \tag{4.6}$$

4.1.2　机器性能表现

在协同降落任务中，机器性能表现从意图推理、降落精度和任务安全性三个方面进行评价，相应权重分别记为 ω_{I}、ω_{L} 和 ω_{S}，并满足：

$$\omega_{\mathrm{I}} + \omega_{\mathrm{L}} + \omega_{\mathrm{S}} = 1 \tag{4.7}$$

根据式（3.4），可分别得到第 k 次任务下意图推理、降落精度和任务安全性三项近期表现值 $P_{\mathrm{I}}(k)$、$P_{\mathrm{L}}(k)$ 和 $P_{\mathrm{S}}(k)$。将其代入式（3.5），可得第 k 次任务下的综合机器性能表现：

$$P_m^m(k) = \omega_{\mathrm{I}} P_{\mathrm{I}}(k) + \omega_{\mathrm{L}} P_{\mathrm{L}}(k) + \omega_{\mathrm{S}} P_{\mathrm{S}}(k) \tag{4.8}$$

4.1.3　客观机器能力

考虑到协同降落任务中的辅助策略可在部署前通过预先测试获得较为充分的能力评估，客观机器能力可按照式（3.7）进行计算。设预先测试任务总数为 K_0，则意图推理、降落精度和任务安全性三个评价指标在预先测试中的平均值分别为：

$$A_{\mathrm{I}} = \frac{1}{K_0} \sum_{r=1}^{K_0} q_{\mathrm{I}}(r), \qquad A_{\mathrm{L}} = \frac{1}{K_0} \sum_{r=1}^{K_0} q_{\mathrm{L}}(r), \qquad A_{\mathrm{S}} = \frac{1}{K_0} \sum_{r=1}^{K_0} q_{\mathrm{S}}(r) \tag{4.9}$$

由此，客观机器能力定义为：

$$C_m^m = \omega_{\mathrm{I}} A_{\mathrm{I}} + \omega_{\mathrm{L}} A_{\mathrm{L}} + \omega_{\mathrm{S}} A_{\mathrm{S}} \tag{4.10}$$

其中，各评价指标的权重系数与式（4.8）保持一致。

4.2　基于深度强化学习的动作价值评估与候选动作集构造

在协同降落任务中，控制输入是连续的，且当前动作会对后续飞行过程及最终着陆结果产生持续影响。因此，仅依据当前状态下的局部信息判断动作优劣并不充分，还需要对动作的长期回报进行评估。基于此，引入双延迟深度确定性策略梯度算法（Twin Delayed Deep Deterministic Policy Gradient, TD3）[76]对动作价值进行建模，从而为候选动作集构造提供依据。

4.2.1　动作价值评估网络

基于 TD3 框架，构建动作价值评估网络。该网络包含一个 Actor 网络 $\pi(s|\phi)$，用于输出确定性策略动作；同时包含两个 Critic 网络 $Q_1(s,\boldsymbol{u}|\theta_1)$ 和 $Q_2(s,\boldsymbol{u}|\theta_2)$，用于评估状态–动作对的价值。

在 TD3 中，智能体的训练目标是最大化累积期望回报。为此，Critic 网络通过最小化贝尔曼残差进行更新，其损失函数定义为：

$$L(\theta_i)=\mathbb{E}_{(s,\boldsymbol{u},r,s',d)\sim\mathcal{D}}\left[\left(y-Q_i(s,\boldsymbol{u}|\theta_i)\right)^2\right],\qquad i=1,2 \tag{4.11}$$

其中，$\mathcal{D}$ 表示经验回放池，$d\in\{0,1\}$ 表示终止标志。

为提高价值估计的稳定性，目标值 y 的计算中引入了目标策略平滑机制。首先，对目标策略输出叠加裁剪噪声：

$$\epsilon\sim\mathrm{clip}\left(\mathcal{N}(0,\sigma),-c,c\right) \tag{4.12}$$

然后将加噪后的目标动作裁剪到动作空间边界内：

$$\bar{\boldsymbol{u}}'=\mathrm{clip}\left(\pi'(s'|\phi')+\epsilon,\mathcal{U}\right) \tag{4.13}$$

目标值可表示为：

$$y=r+\gamma(1-d)\min_{i=1,2}Q_i'(s',\bar{\boldsymbol{u}}') \tag{4.14}$$

其中，Q_i' 和 π' 分别表示 Critic 网络与 Actor 网络的目标网络，γ 为折扣因子，ϵ 为加性平滑噪声。

在共享控制仲裁阶段，为与训练阶段的价值估计方式保持一致，采用两个 Critic 输出中的较小值作为动作价值评估量，即：

$$Q_{\mathrm{eval}}(s,\boldsymbol{u})=\min\{Q_1(s,\boldsymbol{u}),Q_2(s,\boldsymbol{u})\} \tag{4.15}$$

对于任意给定控制输入 $\boldsymbol{u}$，无论其来源于机器策略还是人类实时输入，$Q_{\mathrm{eval}}(s,\boldsymbol{u})$ 均可用于评价该动作在当前状态下对协同降落任务的长期预期回报。

4.2.2　候选动作集构造

为在连续动作空间中实施最小干预仲裁，需先构造有限的候选动作集合。借鉴均匀离散采样思想[77]，先在连续动作空间内进行离散采样，再结合动作价值

评估结果筛选满足基本性能要求的可行动作。

系统动作空间定义为：

$$\mathcal{U} = [u_1^{\min}, u_1^{\max}] \times [u_2^{\min}, u_2^{\max}] \times \cdots \times [u_d^{\min}, u_d^{\max}] \tag{4.16}$$

其中，d 表示动作维数。

对于第 ℓ 个动作分量，将其均匀离散为 n_ℓ 个采样点，且满足 $n_\ell \geqslant 2$。则该分量上的采样集合定义为：

$$\Xi_\ell = \left\{ u_\ell^{\min} + \frac{j-1}{n_\ell - 1}\left(u_\ell^{\max} - u_\ell^{\min}\right) \,\middle|\, j = 1, 2, \cdots, n_\ell \right\}, \qquad \ell = 1, 2, \cdots, d \tag{4.17}$$

由各分量采样点构成的笛卡尔积，可得到静态均匀网格采样动作集合：

$$\mathcal{U}_{\text{grid}} = \left\{ \boldsymbol{u} = (u_1, \cdots, u_d) \,\middle|\, u_\ell \in \Xi_\ell,\ \ell = 1, 2, \cdots, d \right\} \tag{4.18}$$

相应地，均匀网格采样动作集合中的总采样数为：

$$N = \prod_{\ell=1}^{d} n_\ell \tag{4.19}$$

记第 ℓ 个动作分量上的采样间距为：

$$h_\ell = \frac{u_\ell^{\max} - u_\ell^{\min}}{n_\ell - 1}, \qquad \ell = 1, 2, \cdots, d \tag{4.20}$$

对于任意连续动作 $\boldsymbol{u} \in \mathcal{U}$，在均匀网格中总存在一个采样动作 $\hat{\boldsymbol{u}} \in \mathcal{U}_{\text{grid}}$，使其分别满足：

$$\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_\infty \leqslant \frac{1}{2} \max_\ell h_\ell \tag{4.21}$$

$$\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2 \leqslant \frac{1}{2} \sqrt{\sum_{\ell=1}^{d} h_\ell^2} \tag{4.22}$$

由式（4.21）和式（4.22）可知，随着各分量采样间距减小，均匀网格采样对连续动作空间中任意动作的几何逼近误差也随之减小。因此，均匀离散采样能够为连续动作空间的离散近似提供基础，从而支持候选动作集的构造。

在共享控制过程中，还需进一步考虑人类输入 $\boldsymbol{u}_h(t)$ 的意图导向作用以及机器输入 $\boldsymbol{u}_m(t)$ 的参与作用。为此，将二者纳入均匀网格采样动作集合，可得时刻 t 的候选动作集合：

$$\mathcal{U}_{\text{cand}}(t) = \mathcal{U}_{\text{grid}} \cup \{\boldsymbol{u}_h(t), \boldsymbol{u}_m(t)\} \tag{4.23}$$

其中，$\boldsymbol{u}_m(t) = \pi(s(t))$ 表示当前状态下由策略网络输出的机器动作。

在候选动作集合上，先定义价值最高的参考动作为：

$$\boldsymbol{u}_{\text{ref}}(t) = \underset{\boldsymbol{u} \in \mathcal{U}_{\text{cand}}(t)}{\arg\max}\, Q_{\text{eval}}(s(t), \boldsymbol{u}) \tag{4.24}$$

进一步地，将候选动作 $\boldsymbol{u}$ 相对于参考动作 $\boldsymbol{u}_{\text{ref}}(t)$ 的价值损失定义为：

$$\Delta Q(t,\boldsymbol{u}) = Q_{\text{eval}}(s(t),\boldsymbol{u}_{\text{ref}}(t)) - Q_{\text{eval}}(s(t),\boldsymbol{u}) \tag{4.25}$$

保留所有价值损失不超过给定阈值的候选动作，可得到基础可行动作集合：

$$\mathcal{U}_Q(t) = \left\{ \boldsymbol{u} \in \mathcal{U}_{\text{cand}}(t) \,\middle|\, \Delta Q(t,\boldsymbol{u}) \leqslant \delta_Q \right\} \tag{4.26}$$

其中，$\delta_Q > 0$ 为价值损失容许阈值。该集合包含所有相对于参考动作价值损失不超过预设上界的候选动作，并作为后续仲裁的输入。

4.3　信任驱动的最小干预仲裁机制

在完成动作价值评估网络构建与候选动作集生成后，进一步将人类信任状态引入共享控制仲裁过程，根据信任水平动态调节候选动作可接受的价值损失上界，并在满足约束的动作中选取与人类输入最接近的动作作为控制输出，从而在保证任务性能的前提下尽可能保留人类的原始控制意图。该方法的整体架构如图 4.1 所示，完整算法流程见算法 4.1。

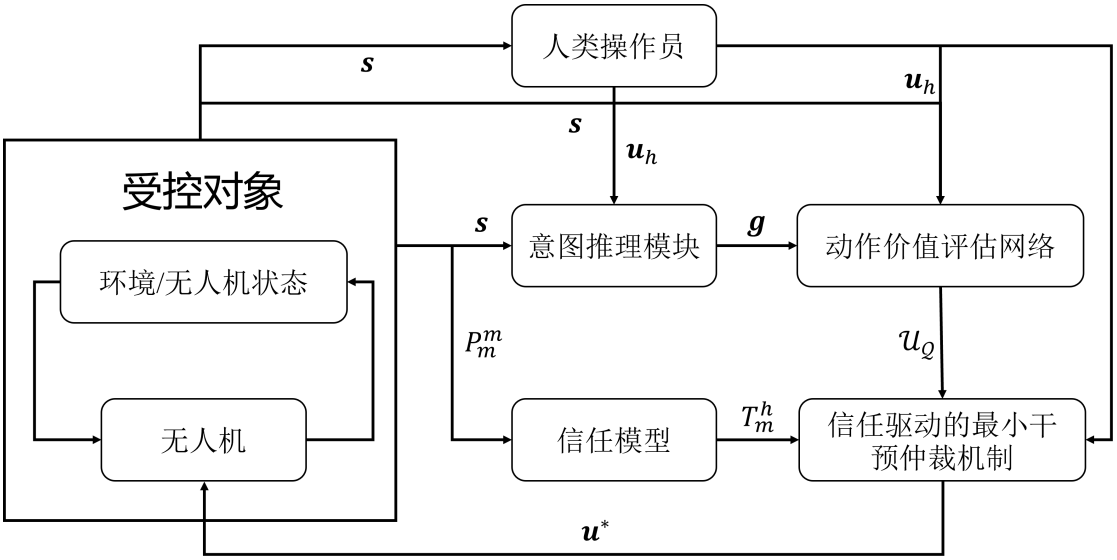


图 4.1　信任驱动的最小干预仲裁机制框架图

4.3.1　信任约束动作集

记第 k 次任务开始前人类对机器的信任为 $T_m^h(k)$。在任务执行过程中，该值保持不变，仅在任务间进行更新。为便于后续仲裁机制设计，先将其归一化到区间 $[0,1]$ 内：

$$\tilde{T}(k) = \begin{cases} 0, & T_m^h(k) \leqslant T_{\min} \\ \dfrac{T_m^h(k) - T_{\min}}{T_{\text{ref}} - T_{\min}}, & T_{\min} < T_m^h(k) < T_{\text{ref}} \\ 1, & T_m^h(k) \geqslant T_{\text{ref}} \end{cases} \tag{4.27}$$

其中，$T_{\min}$ 为信任归一化下界，T_{ref} 为归一化参考值。

为使系统能够根据信任水平调节对人类输入的干预强度，定义第 k 次任务的信任驱动价值损失阈值为：

$$\delta_T(k) = \delta_{\min} + (\delta_{\max} - \delta_{\min})(1 - \tilde{T}(k)) \tag{4.28}$$

其中，$\delta_{\min}$ 和 $\delta_{\max}$ 分别表示价值损失阈值的下界与上界，并满足：

$$0 < \delta_{\min} < \delta_{\max} \leqslant \delta_Q \tag{4.29}$$

由式（4.28）可知，信任水平越高，阈值 $\delta_T(k)$ 越小，系统对候选动作的价值约束越严格，满足约束的动作将更多集中在参考动作附近。此时，机器在仲裁过程中具有更强的干预作用。相反，信任水平越低，阈值 $\delta_T(k)$ 越大，系统允许保留更大范围的候选动作，人类输入附近的动作更容易进入仲裁集合，从而使人类对最终控制输出具有更大的影响。

在第 k 次任务执行过程中的时刻 t，在基础可行动作集合 $\mathcal{U}_Q(t)$ 上进一步施加信任约束，可得信任约束动作集合：

$$\mathcal{U}_T(t,k) = \left\{ \boldsymbol{u} \in \mathcal{U}_Q(t) \;\middle|\; \Delta Q(t,\boldsymbol{u}) \leqslant \delta_T(k) \right\} \tag{4.30}$$

该集合包含当前信任水平下同时满足基础价值约束与信任约束要求的候选动作，并作为后续最小干预仲裁的候选动作集合。

4.3.2　最小干预动作选择

在给定信任约束动作集合后，最终控制输出按照最小干预原则确定。由于参考动作 $\boldsymbol{u}_{\mathrm{ref}}(t)$ 属于候选动作集合，且满足：

$$\Delta Q\big(t, \boldsymbol{u}_{\mathrm{ref}}(t)\big) = 0 \tag{4.31}$$

因此，在 $\delta_T(k) > 0$ 的条件下，信任约束动作集合 $\mathcal{U}_T(t,k)$ 始终非空。

基于此，从 $\mathcal{U}_T(t,k)$ 中选取与当前人类输入 $\boldsymbol{u}_h(t)$ 加权距离最小的动作作为控制输出：

$$\boldsymbol{u}^*(t) = \underset{\boldsymbol{u} \in \mathcal{U}_T(t,k)}{\arg\min} \big(\boldsymbol{u} - \boldsymbol{u}_h(t)\big)^{\mathsf{T}} \mathbf{W} \big(\boldsymbol{u} - \boldsymbol{u}_h(t)\big) \tag{4.32}$$

其中，$\mathbf{W}$ 为正定对角权重矩阵，用于计算各控制分量上的干预代价。

4.4　仿真实验与结果分析

为验证所提融合信任的最小干预共享控制策略在人-无人机协同降落任务中的有效性，实验分为两部分：一部分为模拟人类实验，用于在可控条件下比较不同技能和不同控制方法下的系统表现；另一部分为真实受试者实验，用于进一步验证该方法在真实人机交互过程中的效果。两部分实验均在基于 PyFlyt 构建的仿真环境中进行，仿真环境如图 4.2 所示。

算法 4.1　信任驱动的最小干预仲裁算法

Input: 机器策略 $\pi(\cdot)$，动作价值评估网络 $Q_{\text{eval}}(\cdot,\cdot)$，信任值 $T_m^h(k)$，权重矩阵 $\mathbf{W}$，动作空间边界 $\mathcal{U}$，各分量采样点数 $\{n_\ell\}_{\ell=1}^{d}$，阈值参数 $T_{\min}$，T_{ref}，δ_Q，$\delta_{\min}$，$\delta_{\max}$

Output: 控制序列 $\{\boldsymbol{u}^*(t)\}$

根据 $\mathcal{U}$ 与 $\{n_\ell\}_{\ell=1}^{d}$ 构造静态均匀网格采样动作集合 $\mathcal{U}_{\text{grid}}$;
根据式（4.27） 计算归一化信任值 $\tilde{T}(k)$;
根据式（4.28） 计算第 k 次任务的信任驱动价值损失阈值 $\delta_T(k)$;
while 当前任务未结束 **do**
    获取当前环境状态 $s(t)$ 和人类输入 $\boldsymbol{u}_h(t)$;
    根据机器策略计算机器输入 $\boldsymbol{u}_m(t) \leftarrow \pi(s(t))$;
    构造候选动作集合 $\mathcal{U}_{\text{cand}}(t) \leftarrow \mathcal{U}_{\text{grid}} \cup \{\boldsymbol{u}_h(t), \boldsymbol{u}_m(t)\}$;
    计算参考动作 $\boldsymbol{u}_{\text{ref}}(t) \leftarrow \arg\max_{\boldsymbol{u}\in\mathcal{U}_{\text{cand}}(t)} Q_{\text{eval}}(s(t),\boldsymbol{u})$;
    foreach $\boldsymbol{u} \in \mathcal{U}_{\text{cand}}(t)$ **do**
        根据式（4.25） 计算价值损失 $\Delta Q(t,\boldsymbol{u})$;
    end
    构造基础可行动作集合 $\mathcal{U}_Q(t) \leftarrow \{\boldsymbol{u} \in \mathcal{U}_{\text{cand}}(t) \mid \Delta Q(t,\boldsymbol{u}) \leqslant \delta_Q\}$;
    构造信任约束动作集合 $\mathcal{U}_T(t,k) \leftarrow \{\boldsymbol{u} \in \mathcal{U}_Q(t) \mid \Delta Q(t,\boldsymbol{u}) \leqslant \delta_T(k)\}$;
    根据式（4.32） 选择控制输出 $\boldsymbol{u}^*(t)$;
end
return $\{\boldsymbol{u}^*(t)\}$;

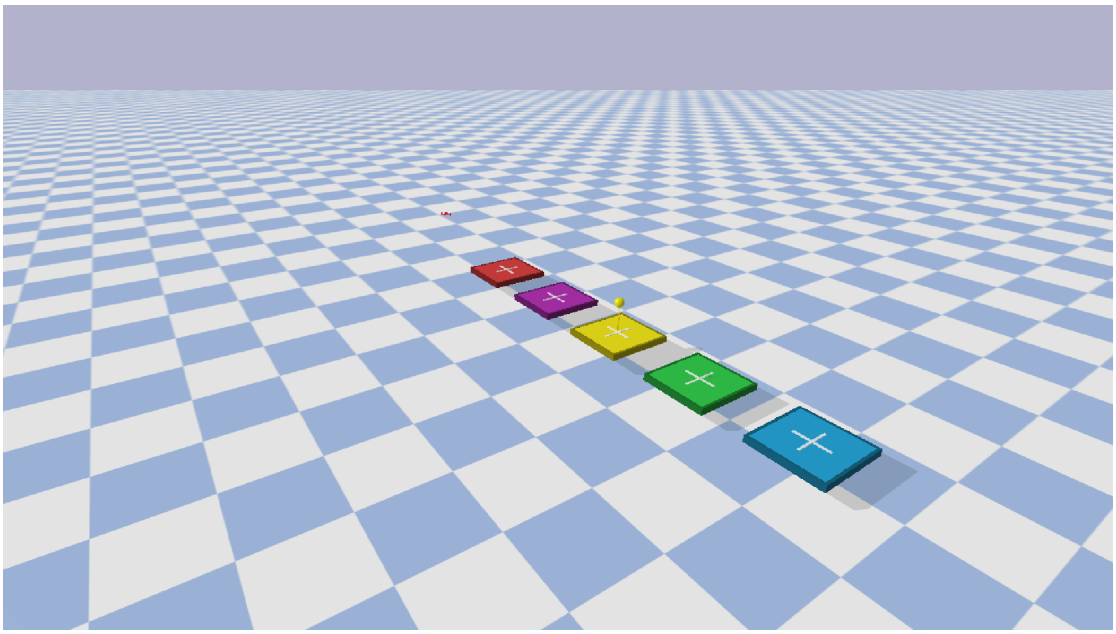


图 4.2　无人机降落任务仿真环境示意图

4.4.1　实验一：模拟人类实验

4.4.1.1　模拟人类建模

为模拟人类操作员在协同降落过程中的控制行为，构建模拟人类模型。该模型主要包括目标平台、内部目标估计、技能参数、接受度参数和动作生成机制。在本节中，将一次完整的协同降落任务执行记为一个实验回合。

设第 k 个实验回合的目标平台为 g_k^*，其中心位置记为 $\boldsymbol{p}_{g_k^*}$。模拟人类已知当前回合的目标平台，但模拟人类对无人机与目标平台相对位置的判断存在偏差，其动作生成并非直接依据真实目标位置，而是依据时刻 t 对目标位置的内部估计 $\hat{\boldsymbol{p}}_g(t)$ 进行。

在每个实验回合开始时，目标估计围绕真实目标位置进行初始化，其初始偏差由技能参数 $\beta_{\mathrm{s}} \in [0,1]$ 控制：技能水平越低，初始偏差越大；技能水平越高，初始估计越接近真实目标。对于部分低技能设置，目标估计的初始偏差可能沿相邻平台方向产生较大偏移，以反映回合开始阶段对目标相对方位判断不准确的情况。

在动作生成阶段，模拟人类根据当前目标估计和无人机状态产生原始控制动作。设时刻 t 无人机的位置和速度分别为 $\boldsymbol{p}(t)$ 和 $\boldsymbol{v}(t)$，则无人机相对于目标估计的位置误差为：

$$\boldsymbol{e}(t) = \hat{\boldsymbol{p}}_g(t) - \boldsymbol{p}(t) \tag{4.33}$$

基于误差 $\boldsymbol{e}(t)$ 和速度 $\boldsymbol{v}(t)$，采用 PD 控制生成模拟人类的原始动作，并叠加与技能水平相关的零均值高斯噪声，以描述不同技能水平的模拟人类在操作精度和稳定性上的差异。技能水平越低，噪声强度越大。

在回合执行过程中，模拟人类的目标估计会随任务进展动态更新，其更新形式为：

$$\hat{\boldsymbol{p}}_g(t+1) = \hat{\boldsymbol{p}}_g(t) + \frac{\beta_{\mathrm{s}}}{\boldsymbol{K}_\beta}\left(\boldsymbol{p}_{g_k^*} - \hat{\boldsymbol{p}}_g(t)\right) + \frac{\alpha_{\mathrm{acc}}(k)}{\boldsymbol{K}_\alpha}\boldsymbol{r}_m(t) \tag{4.34}$$

其中，$\boldsymbol{K}_\beta$ 和 $\boldsymbol{K}_\alpha$ 分别为技能修正项缩放系数和机器辅助项缩放系数，用于调节对应项对目标估计更新的影响强度。式中，第一项为历史项，用于保留上一时刻的目标估计；第二项表示模拟人类依据自身目标判断对目标估计的修正，其修正幅度由技能参数 β_{s} 决定；第三项表示机器辅助在接受度调节下对目标估计的修正作用。$\alpha_{\mathrm{acc}}(k)$ 表示模拟人类对机器辅助的接受度，由回合开始前的信任状态 $T_m^h(k)$ 线性映射得到，并在当前回合内保持不变。信任水平越高，接受度越高，模拟人类越倾向于依据机器辅助修正目标估计；反之，接受度越低，模拟人类越倾向于依赖自身判断，而较少依据机器辅助修正目标估计。

机器辅助对应的目标估计修正项 $\boldsymbol{r}_m(t)$ 定义为：

$$\boldsymbol{r}_m(t) = \mathcal{M}\left(\boldsymbol{u}^*(t) - \boldsymbol{u}_h(t)\right) \tag{4.35}$$

其中，$\boldsymbol{u}_h(t)$ 为模拟人类的原始动作，$\boldsymbol{u}^*(t)$ 为无人机的最终执行动作，$\mathcal{M}(\cdot)$ 表示将动作偏移映射为目标估计修正方向的变换。

4.4.1.2　实验设置

在本组实验中，共比较四种控制方法：模拟人类控制（Pure Human Control under Simulated User, PHC）、固定权重共享控制（Fixed-Weight Shared Control, FWSC）、最小干预共享控制（Minimal Intervention Shared Control, MISC）以及融合信任的最小干预共享控制（Trust-Aware Minimal Intervention Shared Control, TA-MISC）。其中，PHC 由模拟人类独立完成控制；FWSC 按固定权重融合模拟人类动作与机器动作，其中机器权重设为 0.2；MISC 在动作价值约束下执行最小干预仲裁，并采用固定价值损失阈值 δ_{fix} 筛选候选动作；TA-MISC 则在最小干预仲裁基础上根据信任状态动态调节价值损失阈值。

实验中的意图推理模块采用高斯朴素贝叶斯分类器，用于根据模拟人类输入与无人机状态信息在线推断目标平台。为分析不同技能水平条件下各方法的表现，将模拟人类技能参数 β_{s} 设置为 0.05 至 0.95，步长为 0.05，并在每种方法的各技能水平设置下重复运行 100 个实验回合。

在正式实验开始前，先采用中等技能水平的模拟人类对机器辅助能力进行预评估，分别统计机器在意图推理、降落精度和任务安全性三个方面的表现，并据此计算客观机器能力常数 C_m^m。模拟人类实验中的关键参数设置如表 4.1 所示。

4.4.1.3　实验结果与分析

表 4.2 给出了四种控制方法在模拟人类实验中的总体结果。可以看出，TA-MISC 在四项指标上均优于其余三种方法，表明在最小干预仲裁中引入信任状态后，人机系统有更好的综合表现。

从任务完成情况和执行效率来看，PHC 的平均成功率仅为 24.89%，说明在存在目标判断误差和控制噪声的条件下，仅依赖模拟人类自身控制较难稳定完成协同降落任务。FWSC 的成功率提升至 51.79%，表明固定权重共享控制能够在一定程度上改善任务完成效果，但其平均回合步数达到 510.29，为四种方法中最高，说明固定权重融合方式虽然能够提供辅助，但在控制权分配上缺乏针对性，往往需要更长的持续控制与反复修正才能完成任务。相比之下，MISC 和 TA-MISC 的成功率分别达到 66.74% 和 73.79%，同时平均回合步数分别降至 424.25 和 409.61。这表明基于最小干预原则的仲裁方式能够在提高任务成功率的同时减少任务完成所需的控制步数。

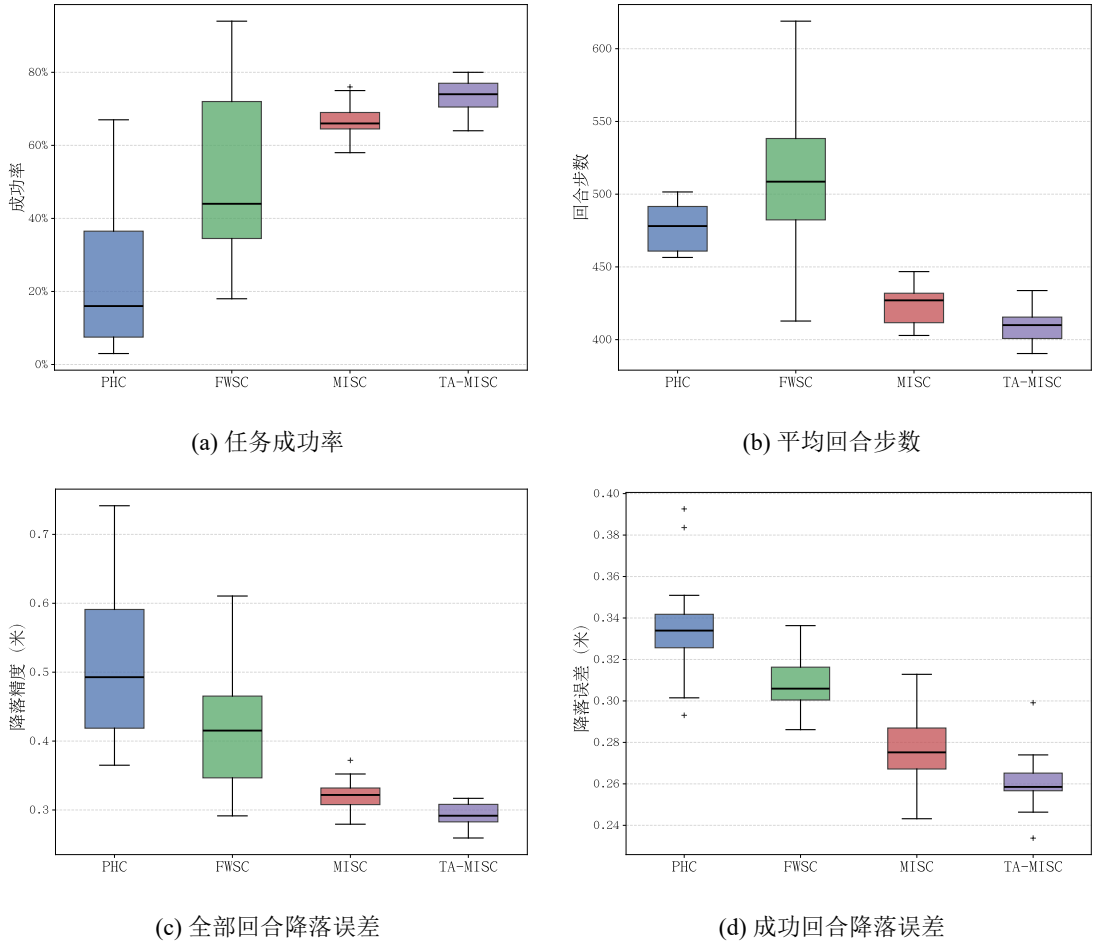
表 4.1　模拟人类实验关键参数设置

参数	值
初始人类因子 α_i	0.6
客观机器能力 C_m^m	0.55
正向认知更新因子 α_h^+	0.18
负向认知更新因子 α_h^-	0.32
动态观测因子 α_d	1.0
时间窗口长度 L	5
衰减因子 γ	0.5
基础价值损失阈值 δ_Q	2.0
固定价值损失阈值 δ_{fix}	1.0
动态阈值下界 $\delta_{\min}$	0.3
动态阈值上界 $\delta_{\max}$	2.0
技能修正项缩放系数 K_β	20.0
机器辅助项缩放系数 K_α	15.0

表 4.2　模拟人类实验结果对比

方法	成功率（%）	回合步数	降落误差（m）	成功回合降落误差（m）
PHC	24.89 ± 20.32	477.72 ± 15.91	0.51 ± 0.11	0.34 ± 0.04
FWSC	51.79 ± 23.01	510.29 ± 47.94	0.41 ± 0.08	0.31 ± 0.04
MISC	66.74 ± 4.55	424.25 ± 12.77	0.32 ± 0.04	0.28 ± 0.03
TA-MISC	$\mathbf{73.79 \pm 4.48}$	$\mathbf{409.61 \pm 10.93}$	$\mathbf{0.29 \pm 0.04}$	$\mathbf{0.26 \pm 0.02}$

从着陆质量来看，PHC 的全部回合平均降落误差最高，说明模拟人类单独控制下的整体着陆质量相对较差。FWSC 将该指标降低至 0.41 m，说明固定权重共享控制在一定程度上改善了整体着陆表现。MISC 和 TA-MISC 的全部回合平均降落误差进一步下降至 0.32 m 和 0.29 m。在成功回合中，四种方法的平均降落误差分别为 0.34 m、0.31 m、0.28 m 和 0.26 m，其中 TA-MISC 最接近目标平台中心，表明其在保证任务完成能力的同时仍能保持较好的着陆精度。



(a) 任务成功率　　(b) 平均回合步数

(c) 全部回合降落误差　　(d) 成功回合降落误差

图 4.3　模拟人类实验结果图

图 4.3 给出了四项指标在不同技能水平下的分布情况，更能反映不同方法对模拟人类技能差异的适应能力。PHC 的结果离散程度较大，说明在技能水平较低或控制能力波动较明显的条件下，仅依赖模拟人类控制难以保证任务执行的稳定性，因此有必要引入共享控制机制。FWSC 虽能够提供一定辅助，但整体离散程度仍然较大，反映出固定权重融合方式对不同技能水平的适应性有限。相比之下，MISC 与 TA-MISC 的结果更为集中，说明最小干预仲裁能够在保留人类控制参与的同时，更有效地适应不同技能水平条件下的任务需求。进一步地，TA-MISC 较 MISC 波动更小，表明在最小干预仲裁中引入信任状态，有助于提高协同控制策略的稳定性。

4.4.2　实验二：真实受试者实验

4.4.2.1　实验设置

在本组实验中，共比较三种控制方法：人类单独控制（PHC）、最小干预共享控制（MISC）以及融合信任的最小干预共享控制（TA-MISC）。其中，PHC 由人类受试者单独控制；MISC 采用固定价值损失阈值进行候选动作筛选；TA-MISC 则在此基础上根据信任状态对价值损失阈值进行动态调节。MISC 和 TA-MISC 均采用人类与机器共享控制的方式。

为验证所提方法在真实人机交互过程中的有效性，邀请 10 名受试者参与实验。正式实验开始前，受试者先进行若干回合练习，以熟悉无人机运动特性、环境视角与交互方式。随后，每名受试者分别在三种控制方法下完成 20 次降落任务。实验环境设置、信任模型及共享控制等相关参数与实验一保持一致，意图推理模块采用高斯朴素贝叶斯分类器。

实验从任务完成情况、执行效率、着陆质量和人机协同状态四个方面进行分析，统计指标包括成功率、回合步数、降落误差、平均信任水平和干预强度。

4.4.2.2　实验结果与分析

表 4.3 给出了三种控制方法在真实人机交互实验中的总体结果。整体来看，TA-MISC 在成功率、回合步数、降落误差和平均信任水平等指标上均表现最优，但在最小干预仲裁中引入信任状态后，对人类的干预强度也相对更高。下面结合表 4.3 和图 4.4 对实验结果进行具体分析。

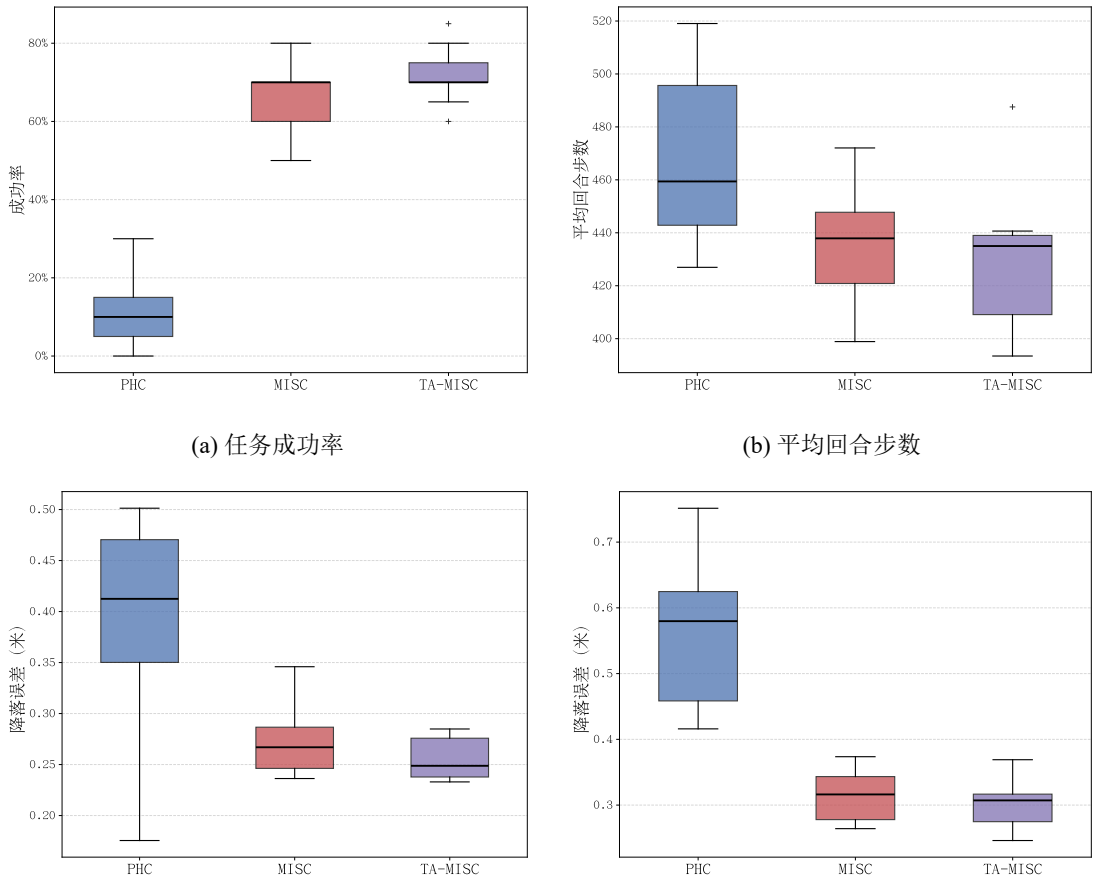
表 4.3　真实人类实验结果对比

方法	成功率（%）	回合步数	降落误差（m）	平均信任	干预强度
PHC	10.56 ± 9.26	471.26 ± 32.93	0.56 ± 0.10	–	–
MISC	65.56 ± 8.64	437.84 ± 22.28	0.32 ± 0.04	0.74 ± 0.06	0.36 ± 0.01
TA-MISC	$\mathbf{72.22 \pm 7.11}$	$\mathbf{428.22 \pm 27.44}$	$\mathbf{0.26 \pm 0.02}$	$\mathbf{0.79 \pm 0.03}$	$\mathbf{0.42 \pm 0.02}$

从任务完成情况与执行效率来看，PHC 的平均成功率仅为 10.56%，平均回合步数达到 471.26，说明单独依靠人类操作较难稳定完成降落任务。相比之下，MISC 与 TA-MISC 的平均成功率分别提高至 65.56% 和 72.22%，平均回合步数分别下降至 437.84 和 428.22。

从着陆质量来看，PHC 的平均降落误差最高，为 0.56 m，反映出人类单独控制下着陆精度相对较低。MISC 将该指标降低至 0.32 m，TA-MISC 进一步降低至 0.26 m。综合来看，在无人机降落任务中引入共享控制能够改善任务表现与着陆质量，其中 TA-MISC 的整体表现最好。

从人机协同状态来看，PHC 不涉及机器辅助，因此不统计平均信任水平与干预强度。对于两种共享控制方法，MISC 与 TA-MISC 的平均信任水平分别为 0.74 和 0.79，说明引入信任调节机制后，人类在任务过程中的整体信任水平更高。与此同时，TA-MISC 的平均干预强度为 0.42，高于 MISC 的 0.36。虽然 TA-MISC 会对人类进行更高程度的干预，但由于其在仲裁机制中考虑了人类信任状态，这种干预是与人类信任水平相匹配的合理干预。其更高的平均信任水平也表明，其对系统辅助行为的接受程度也相对更高，有助于改善人机协同效果。



(a) 任务成功率　　(b) 平均回合步数

(c) 全部回合降落误差　　(d) 成功回合降落误差

图 4.4　真实受试者实验结果图

图 4.4 进一步展示了三种控制方法在核心任务指标上的分布情况。整体趋势与表 4.3 中的结果一致：PHC 在各项任务指标上的表现相对较弱，MISC 已能够明显改善任务结果，TA-MISC 则在成功率、回合步数和降落误差方面进一步优于 MISC。由此可见，融合信任的最小干预共享控制策略在真实人机交互条件下仍表现出更好的整体效果。

4.5　本章小结

本章针对人机协同降落任务中缺乏合理控制权分配机制的问题，提出了融合信任的最小干预共享控制策略。围绕协同降落任务，构建了任务评价指标，并据此计算机器性能表现与客观机器能力；进一步构建了动作价值评估网络，设计了候选动作集构造与筛选机制，并在此基础上实现了信任驱动的最小干预仲裁。实验结果表明，所提方法能够提高任务成功率，改善人机协同过程中的控制表现，为人机协同降落任务中的控制权分配提供了有效方法。

第 5 章　融合信任的人机协同轨迹规划方法

在人-无人机协同搜索任务中，人类信任状态是影响轨迹规划效果的重要因素。由于人类操作员通常只能依赖无人机回传的图像获取局部环境信息，而无人机避障与轨迹决策过程又缺乏足够透明性，操作员难以准确理解机器能力及当前规划依据，因此人类对系统的信任更容易发生变化。当规划轨迹与操作员预期不一致时，操作员可能产生不必要的修正输入，不仅会加重操作负担，也可能会导致人机冲突。因此，需要在轨迹规划过程中考虑人类信任的动态变化。

为此，面向协同搜索任务，本章提出一种融合信任的人机协同轨迹规划方法，将人类信任状态纳入轨迹规划过程，以平衡飞行质量、人类意图响应与信任状态的要求。本章围绕协同搜索任务下机器性能表现与客观机器能力的计算、意图引导的候选轨迹生成与选择机制、信任感知的轨迹优化策略，以及仿真实验设计与结果验证等内容展开研究。

5.1　问题描述

设无人机在时刻 t 的状态向量为 $\mathbf{x}(t) \in \mathcal{X} \subset \mathbb{R}^n$，其中 $\mathcal{X}$ 表示状态空间；控制输入记为 $\mathbf{u}(t) \in \mathcal{U} \subset \mathbb{R}^l$，其中 $\mathcal{U}$ 表示允许的控制输入集合，n 和 l 分别表示状态向量与控制输入的维数。在协同搜索任务中，无人机需要根据当前状态、操作员输入以及环境信息规划飞行轨迹。

记 $\xi(\tau)$ 为规划时域 $\tau \in [0, T]$ 上的状态轨迹，其位置分量记为 $\mathbf{p}(\tau)$。为保证飞行过程中的安全性与物理可行性，轨迹需要满足动力学可行性约束，即：

$$\xi^{(i)}(\tau) \leqslant \mathbf{x}_{\max}^{(i)}, \quad \forall \tau \in [0, T] \tag{5.1}$$

其中，$\xi^{(i)}(\tau)$ 表示轨迹的 i 阶导数，$\mathbf{x}_{\max}^{(i)}$ 表示无人机对应的物理极限，例如最大速度、最大加速度等。同时，轨迹需要在整个规划过程中始终位于安全空间内，即:

$$\mathbf{p}(\tau) \in \mathcal{X}_{\text{safe}}, \quad \forall \tau \in [0, T] \tag{5.2}$$

其中，$\mathcal{X}_{\text{safe}}$ 表示空间中的安全区域。

在协同搜索任务场景中，往往不存在一个已知的固定目标，操作员更多是通过实时控制指令 $\mathbf{u}_h(t)$ 持续引导无人机的运动方向。因此，本章不将操作员意图表示为状态空间中的显式目标，而将其建模为方向性意图[78]。在此基础上，定义意图代价函数 $C_{\mathbf{u}_h}(\xi)$，用于衡量轨迹 $\xi(\tau)$ 与人类意图之间的一致程度；同时，定义飞行质量代价函数 $C_q(\xi)$，用于反映轨迹的平滑性等飞行质量。

考虑到人类当前信任状态 $T_m^h(t)$ 会影响系统在意图响应与飞行质量之间的权衡，可将人机协同轨迹规划问题形式化为：

$$\begin{aligned} \min_{\xi(\tau)} \quad & C_{\mathbf{u}_h}(\xi) + \lambda\left(T_m^h(t)\right) C_q(\xi) \\ \text{s.t.} \quad & \xi^{(i)}(\tau) \leqslant \mathbf{x}_{\max}^{(i)}, \quad \forall \tau \in [0, T] \\ & \mathbf{p}(\tau) \in \mathcal{X}_{\text{safe}}, \quad \forall \tau \in [0, T] \end{aligned} \tag{5.3}$$

上述优化问题从整体上描述了融合信任的人机协同轨迹规划目标。基于这一问题描述，后续将通过候选轨迹生成与选择以及轨迹优化，具体实现融合信任的人机协同轨迹规划方法。

5.2　搜索任务下机器能力构建

在第三章构建 PDTEM 的基础上，本节结合人-无人机协同搜索任务特点构建任务评价指标，以计算客观机器能力 $C_m^m(k)$ 与机器性能表现 $P_m^m(k)$。

5.2.1　任务评价指标

人机协同搜索任务要求无人机在任务执行过程中既能够保证飞行安全，又能够维持对环境和目标区域的有效观察。因此，本节选取安全性和可视性作为协同搜索任务的评价指标。

1. 安全性指标

安全性指标用于衡量无人机在飞行过程中的碰撞风险。记时刻 t 无人机指向最近障碍物的相对位置矢量为 $\boldsymbol{d}(t) = [d_x(t), d_y(t), d_z(t)]$，则到最近障碍物的欧氏距离为：

$$\Delta d(t) = \|\boldsymbol{d}(t)\| = \sqrt{d_x^2(t) + d_y^2(t) + d_z^2(t)} \tag{5.4}$$

设无人机的飞行速度矢量为 $\boldsymbol{v}(t) = [v_x(t), v_y(t), v_z(t)]$，则沿障碍物方向的接近速度分量定义为：

$$v_d(t) = \begin{cases} \dfrac{\boldsymbol{v}(t) \cdot \boldsymbol{d}(t)}{\Delta d(t)}, & \boldsymbol{v}(t) \cdot \boldsymbol{d}(t) > 0 \\ 0, & \boldsymbol{v}(t) \cdot \boldsymbol{d}(t) \leqslant 0 \end{cases} \tag{5.5}$$

其中，$\boldsymbol{v}(t) \cdot \boldsymbol{d}(t)$ 表示速度矢量在障碍物方向上的投影。

在此基础上，定义时刻 t 的安全性函数 $S(t)$ 为：

$$S(t) = e^{-\frac{\gamma_s v_d^2(t)}{2\Delta d(t)}} \tag{5.6}$$

其中，$\gamma_s \in \mathbb{R}^+$ 为风险敏感度调节系数，$\frac{v_d^2(t)}{2\Delta d(t)}$ 表示沿障碍物方向为避免碰撞所需的最小恒定减速度。

2. 可视性指标

可视性指标用于量化无人机视场（field of view, FOV）内的遮挡程度。视场内遮挡越严重，操作员能够获取的环境信息越少，从而影响其决策。

如图 5.1 所示，设时刻 t 无人机位置为 $\boldsymbol{p}(t)=[p_x(t),p_y(t),p_z(t)]$，当前引导目标位置为 $\boldsymbol{c}(t)=[c_x(t),c_y(t),c_z(t)]$，并假设无人机始终朝向该引导目标。为降低计算复杂度，采用一组球形区域 $\{\mathcal{B}_1(t),\mathcal{B}_2(t),\cdots,\mathcal{B}_N(t)\}$ 对视场内的可视情况进行近似评估[61]。各球形区域的中心 $\boldsymbol{c}_i(t)$ 与半径 $r_i(t)$ 定义为：

$$\begin{aligned}\boldsymbol{c}_i(t)&=\boldsymbol{p}(t)+\psi_i\big(\boldsymbol{c}(t)-\boldsymbol{p}(t)\big)\\ r_i(t)&=\rho\,\psi_i\,\|\boldsymbol{c}(t)-\boldsymbol{p}(t)\|\end{aligned}\tag{5.7}$$

其中，$\psi_i=\frac{i}{N}$，$i\in\{1,2,\cdots,N\}$，N 表示球形区域总数，ρ 为与 FOV 大小相关的参数。

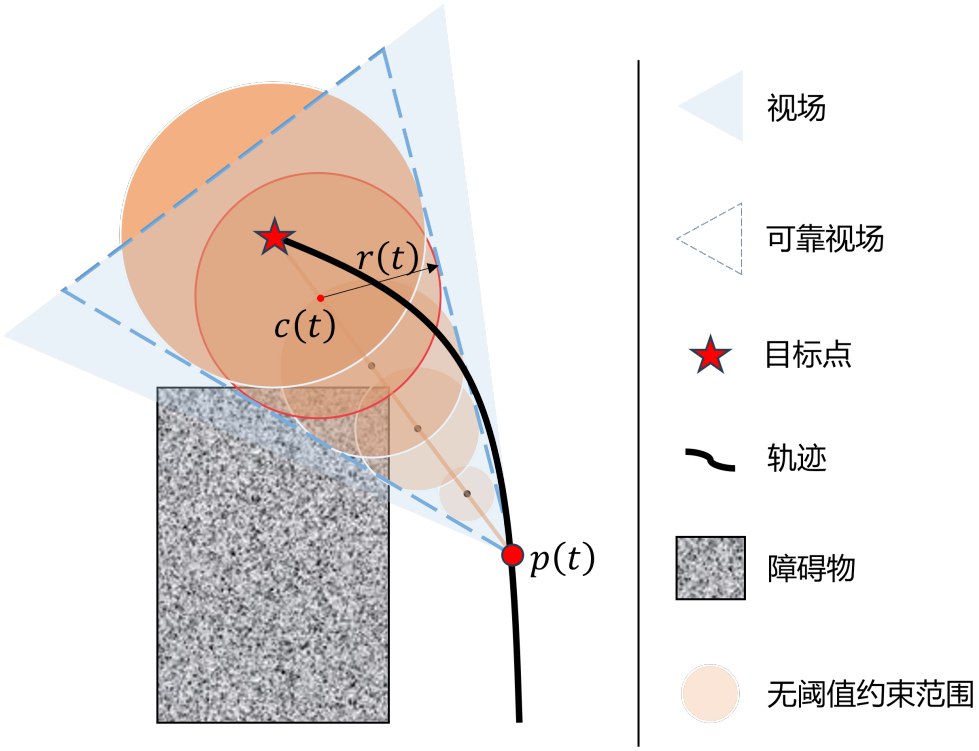


图 5.1　可视性指标示意图

对于每个区域 $\mathcal{B}_i(t)$，通过比较中心点 $\boldsymbol{c}_i(t)$ 到最近障碍物的最小距离 $\Xi(\boldsymbol{c}_i(t))$ 与半径 $r_i(t)$ 判断其遮挡情况。当满足 $\Xi(\boldsymbol{c}_i(t))\geqslant r_i(t)$ 时，可认为该区域未被遮挡。由此，定义时刻 t 的可视性函数 $F(t)$ 为：

$$F(t)=1-\frac{\sum_{i=1}^{N}\max\big(0,r_i^2(t)-\Xi^2(\boldsymbol{c}_i(t))\big)}{\sum_{i=1}^{N}r_i^2(t)}\tag{5.8}$$

5.2.2　机器性能表现

机器性能表现由安全性与可视性两项评价指标共同计算。对区间 $\Delta k = [t_{k-1}, t_k)$ 按固定采样间隔 δt 进行采样，共得到 N_k 个采样时刻，记第 i 个采样时刻为：

$$\tau_{k,i} = t_{k-1} + (i-1)\delta t, \qquad i = 1, 2, \cdots, N_k \tag{5.9}$$

由此，第 k 个区间上的安全性与可视性评价结果分别为：

$$q_{\mathrm{S}}(k) = \frac{1}{N_k}\sum_{i=1}^{N_k} S(\tau_{k,i}), \qquad q_{\mathrm{F}}(k) = \frac{1}{N_k}\sum_{i=1}^{N_k} F(\tau_{k,i}) \tag{5.10}$$

设安全性与可视性的权重分别为 ω_{S} 和 ω_{F}，并满足：

$$\omega_{\mathrm{S}} + \omega_{\mathrm{F}} = 1 \tag{5.11}$$

根据式（3.4），可分别得到第 k 个区间上的安全性表现 $P_{\mathrm{S}}(k)$ 和可视性表现 $P_{\mathrm{F}}(k)$。再由式（3.5），搜索任务下的综合机器性能表现为：

$$P_m^m(k) = \omega_{\mathrm{S}} P_{\mathrm{S}}(k) + \omega_{\mathrm{F}} P_{\mathrm{F}}(k) \tag{5.12}$$

5.2.3　客观机器能力

在人-无人机协同搜索任务执行过程中，随着无人机位置和状态的变化，其周围障碍物及视场遮挡情况会不断变化，机器的性能表现也会随之波动。因此，仅依靠离线评估得到的初始客观机器能力，难以对任务过程中的机器能力作出准确的评估。为此，采用前文先验与后验融合的动态更新形式，结合任务过程中的机器性能表现对客观机器能力进行持续修正。

记离线评估得到的初始客观机器能力为 C_m^m，则由式（3.9），搜索任务下的客观机器能力更新为：

$$\begin{cases} C_m^m(k) = (1-\alpha_m)C_m^m(k-1) + \alpha_m P_m^m(k), & k > 0 \\ C_m^m(0) = C_m^m, & k = 0 \end{cases} \tag{5.13}$$

5.3　意图引导的候选轨迹生成与选择

在前文构建搜索任务下机器性能表现与客观机器能力的基础上，进一步结合人类信任状态，开展意图引导的候选轨迹生成与选择。具体地，以操作员输入对应的运动基元为引导，在安全约束下构建运动基元树并生成候选轨迹集；随后结合当前局部轨迹连续性、人类意图一致性以及候选轨迹对应的平均信任水平，从候选轨迹集中选出参考轨迹。整体框架如图 5.2 所示，具体流程见算法 5.1。

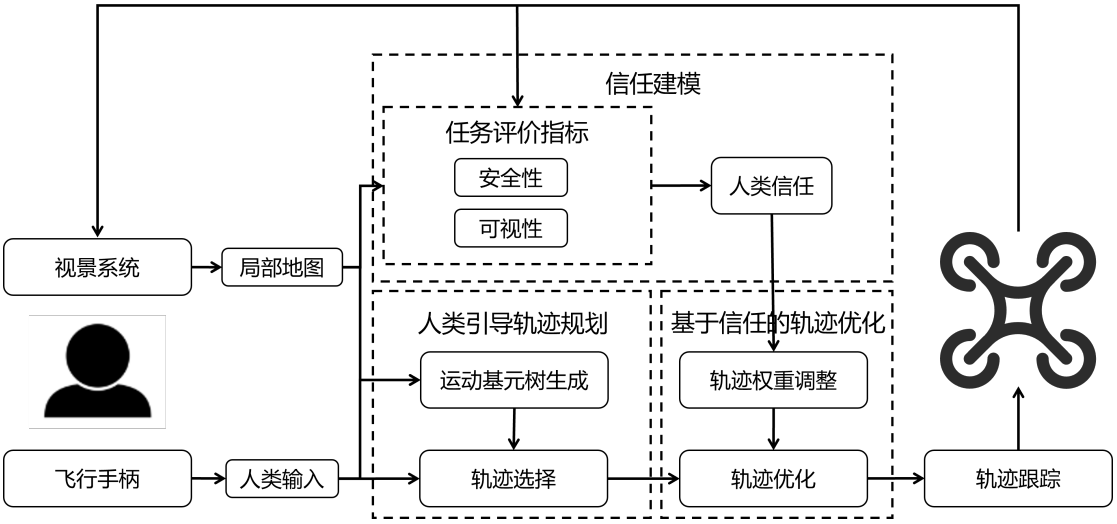


图 5.2　融合信任的人机协同轨迹规划框架图

算法 5.1　融合信任的人机协同轨迹生成与选择算法

Input: 当前状态 $\mathbf{x}_c$，当前局部轨迹 ξ_L，操作员输入 $\mathbf{u}_h$，信任状态 $T_m^h(t)$，安全空间 $\mathcal{X}_{\text{safe}}$，节点动作空间 $\mathcal{A}_n$，每轮扩展节点数 J，候选子节点保留数 K

Output: 最优轨迹 ξ^*

初始化根节点 η_c，采样集 $\mathcal{S} \leftarrow \{\eta_c\}$，运动基元树 $\Gamma \leftarrow \{\eta_c\}$;
根据式（5.16）、式（5.17）和式（5.18）生成引导基元 $\gamma_{\mathbf{u}_h}$;
while $\mathcal{S} \neq \varnothing$ 且未满足停止条件 **do**
    $J_s \leftarrow \min(J, |\mathcal{S}|)$;
    从 $\mathcal{S}$ 中选取 J_s 个采样节点 $\{\eta_n\}_{n=1}^{J_s}$;
    for $n = 1, \cdots, J_s$ **do**
        从 $\mathcal{S}$ 中移除节点 η_n;
        初始化候选子节点集 $\mathcal{C}_n \leftarrow \varnothing$;
        for 动作空间 $\mathcal{A}_n$ 中的每个采样动作 **do**
            生成对应候选基元 $\gamma_{\mathbf{u}_n}(t)$ 及其子节点;
            if $\gamma_{\mathbf{u}_n}(t) \subset \mathcal{X}_{\text{safe}}$ **then**
                将对应子节点加入 $\mathcal{C}_n$;
            end
        end
        if $\mathcal{C}_n \neq \varnothing$ **then**
            根据式（5.19）计算各候选子节点的方向代价;
            选取方向代价最小的 $\min(K, |\mathcal{C}_n|)$ 个子节点;
            将选中的子节点及其对应运动基元加入运动基元树 Γ;
            将选中的子节点加入采样集 $\mathcal{S}$;
        end
    end
end
根据运动基元树 Γ 生成候选轨迹集 $\{\xi_i\}_{i=1}^{M}$;
根据式（5.22）计算各候选轨迹的选择代价;
从 $\{\xi_i\}_{i=1}^{M}$ 中选择最优轨迹 ξ^*;
return ξ^*;

5.3.1　运动基元树

四旋翼无人机具有微分平坦特性[79]，其状态与控制输入可由位置 x、y、z 及偏航角 θ 及其导数唯一表示，因此可基于这些平坦输出构造运动基元轨迹。

记无人机系统状态为：

$$\mathbf{x} = [x, y, z, \theta]^{\top} \in \mathcal{X} \tag{5.14}$$

其中，状态空间满足 $\mathcal{X} \subset \mathbb{R}^4$。操作员在水平坐标系 $\mathcal{C}$ 下给出的控制输入为：

$$\mathbf{u}_h = [v_x, \omega, v_z]^{\top} \tag{5.15}$$

其中，水平坐标系 $\mathcal{C}$ 与无人机原点固连，且其 z 轴与世界坐标系 $\mathcal{W}$ 的 z 轴保持对齐。

采用独轮车模型描述运动基元的状态演化。设运动基元持续时间为 T，则对于当前操作员输入，其对应的运动基元满足：

$$\dot{\mathbf{f}}_{\mathbf{u}_h}(t) = [v_x \cos(\omega t),\ v_x \sin(\omega t),\ v_z,\ \omega]^{\top}, \qquad t \in [0, T] \tag{5.16}$$

然而，该模型无法直接提供高阶参考。考虑到姿态参考可由平坦输出的高阶导数直接计算，因此进一步采用八阶多项式对运动基元进行参数化。将操作员输入对应的运动基元记为引导基元 $\gamma_{\mathbf{u}_h}(t)$，其持续时间记为 T_h。在世界坐标系 $\mathcal{W}$ 下，引导基元可进一步表示为：

$$\gamma_{\mathbf{u}_h}(t) = \sum_{i=0}^{8} \mathbf{c}_i t^i \tag{5.17}$$

给定当前状态 $\mathbf{x}_c$ 及其四阶导数，则引导基元的多项式系数由下列边界条件唯一确定：

$$\begin{cases} \gamma^{(j)}(0) = \mathbf{x}_c^{(j)}, & j = 0, 1, 2, 3, 4 \\ \dot{\gamma}(T_h) = R_{t,\mathcal{C}}^{\mathcal{W}} \dot{\mathbf{f}}_{\mathbf{u}_h}(T_h) & \\ \gamma^{(j)}(T_h) = 0, & j = 2, 3, 4 \end{cases} \tag{5.18}$$

其中，$R_{t,\mathcal{C}}^{\mathcal{W}}$ 表示时刻 t 从水平坐标系 $\mathcal{C}$ 到世界坐标系 $\mathcal{W}$ 的坐标变换矩阵。由于 $\mathcal{C}$ 与 $\mathcal{W}$ 的 z 轴保持对齐，偏航角速度 ω 保持不变。引导基元用于表示当前时刻由操作员输入确定的期望运动方向，并为后续搜索提供引导。

在构建运动基元树时，将前向速度 v_x 固定为操作员输入值，并对角速度 ω、垂直速度 v_z 以及基元持续时间 T 进行离散采样。进一步地，在节点 η_n 对应的离散动作空间 $\mathcal{A}_n$ 中采用增量式动作采样方法[59]进行迭代采样，从而扩展得到候选子节点集 $\mathcal{C}_n$。对于每个采样动作，均采用与引导基元相同的多项式生成方式构造对应的候选基元 $\gamma_{\mathbf{u}_n}(t)$。若该基元在整个持续时间内始终位于安全空间 $\mathcal{X}_{\text{safe}}$ 内，则将其保留为候选基元。

为衡量候选基元与引导基元之间的方向一致性，定义方向代价函数为：

$$f_c^n = 1 - \mathbf{d}_n \cdot \mathbf{d}_h$$
$$\mathbf{d}_n = \frac{p(\gamma_{\mathbf{u}_n}(T_n)) - p(\mathbf{x}_n)}{\|p(\gamma_{\mathbf{u}_n}(T_n)) - p(\mathbf{x}_n)\|}, \quad \mathbf{d}_h = \frac{p(\gamma_{\mathbf{u}_h}(T_h)) - p(\mathbf{x}_c)}{\|p(\gamma_{\mathbf{u}_h}(T_h)) - p(\mathbf{x}_c)\|} \tag{5.19}$$

其中，$p(\cdot)$ 表示状态到空间位置的映射，T_n 表示节点 η_n 对应运动基元的持续时间，$\mathbf{x}_n$ 表示节点 η_n 对应的初始状态。根据代价 f_c^n 对候选子节点集 $\mathcal{C}_n$ 进行排序，并选取若干低代价子节点加入采样集 $\mathcal{S}$，用于后续轮次的运动基元扩展。

当 $\|p(\mathbf{x}_n) - p(\mathbf{x}_c)\|$ 超出局部感知范围，或节点数量达到预设上限时，停止运动基元树的扩展。通过连接运动基元树中不同深度的运动基元，最终得到候选轨迹集 $\{\xi_i\}_{i=1}^{M}$。

5.3.2　轨迹选择机制

在得到候选轨迹集 $\{\xi_i\}_{i=1}^{M}$ 后，需要从中选出最优轨迹 ξ^*。轨迹选择同时考虑三方面因素：一是候选轨迹与当前局部轨迹之间的差异，以保证轨迹切换的连续性；二是候选轨迹与操作员输入对应引导基元之间的差异，以保持轨迹与人类意图的一致性；三是候选轨迹对应的平均信任水平，以使轨迹选择倾向于信任水平较高的候选轨迹。

为度量轨迹之间的相似性，这里采用离散 Fréchet 距离（Discrete Fréchet Distance, DFD）[80]。设两条连续轨迹 f 和 g 经过离散采样后分别形成点序列 $\mathcal{P} = \{f_1, f_2, \cdots, f_n\}$ 和 $\mathcal{Q} = \{g_1, g_2, \cdots, g_m\}$，则二者的 DFD 定义为：

$$\delta_{dF}(f, g) := \min_{(\alpha,\beta)} \max_{i\in\{1,\cdots,k\}} d\left(f_{\alpha(i)}, g_{\beta(i)}\right) \tag{5.20}$$

其中，(α, β) 表示保持顺序的完整对应关系。进一步地，对于时间参数化轨迹 $\boldsymbol{\Gamma}$ 和 $\boldsymbol{\Phi}$，其 DFD 可写为：

$$\delta_{dF}(\boldsymbol{\Gamma}, \boldsymbol{\Phi}) := \min_{(\alpha,\beta)} \max_{i\in\{1,\cdots,k\}} d\left(\boldsymbol{\Gamma}_{\alpha(i)}, \boldsymbol{\Phi}_{\beta(i)}\right) \tag{5.21}$$

其中，$\boldsymbol{\Gamma}_i = \boldsymbol{\Gamma}(i \cdot \Delta t)$ 表示以固定采样间隔 Δt 对轨迹 $\boldsymbol{\Gamma}$ 进行离散采样。离散 Fréchet 距离的计算方式如图 5.3 所示。

记当前局部轨迹为 ξ_L，操作员输入对应的引导基元为 $\gamma_{\mathbf{u}_h}$，则最优轨迹由下式确定：

$$\xi^* = \arg\min_{\xi\in\{\xi_i\}_{i=1}^{M}} \left[w_L\, \delta_{dF}(\xi, \xi_L) + w_U\, \delta_{dF}(\xi, \gamma_{\mathbf{u}_h}) - w_T\, \bar{T}_m^h(\xi) \right] \tag{5.22}$$

其中，$\bar{T}_m^h(\xi)$ 表示候选轨迹 ξ 在规划区间内对应的信任均值，w_L、w_U 和 w_T 分别为轨迹连续性、意图一致性和信任水平对应的权重系数。由式（5.22） 可知，前两项分别约束候选轨迹相对于当前局部轨迹和操作员输入引导基元的偏离程

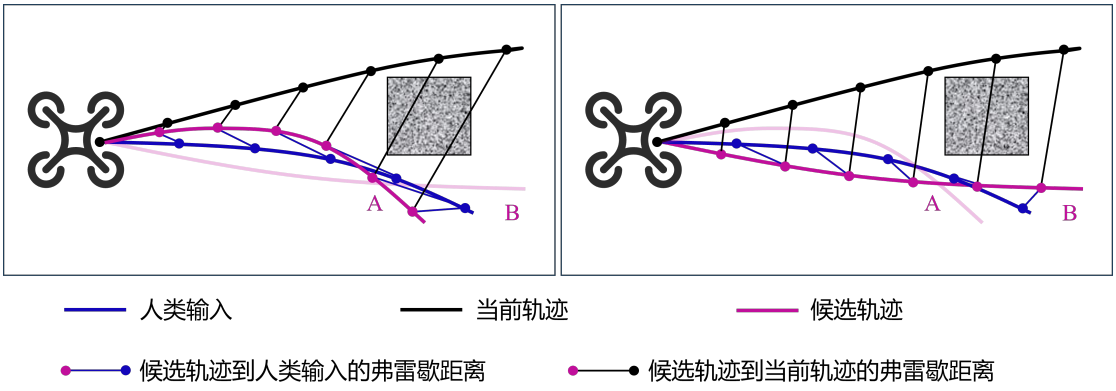


图 5.3　离散 Fréchet 距离计算示意图

度，第三项则倾向于选择对应信任水平较高的轨迹。通过上述机制，可在保证轨迹切换平滑性的同时兼顾人类意图与信任状态对轨迹选择的影响。

5.4　信任驱动的轨迹优化策略

5.4.1　轨迹参数化表示

从候选轨迹集选出的轨迹 ξ^* 作为轨迹优化的参考轨迹。为便于连续优化，采用 MINCO（Minimum Control）对其进行重参数化[81]。优化对象取平坦输出中的位置分量：

$$\mathbf{p}(t) = [x(t), y(t), z(t)]^{\mathsf{T}} \tag{5.23}$$

航向角不作为此处的主优化变量，仍由参考轨迹给定。

MINCO 通过线性复杂度映射对分段多项式轨迹进行时空解耦，从而将轨迹的空间参数与时间分配分离。设轨迹由 M 段多项式构成，则 MINCO 轨迹类可表示为：

$$\mathfrak{T}_{\mathrm{MINCO}} = \left\{ \mathbf{p}(t) : [0, T_\Sigma] \mapsto \mathbb{R}^m \;\middle|\; \mathbf{c} = \mathcal{M}(\mathbf{q}, \mathbf{T}), \quad \mathbf{q} \in \mathbb{R}^{m(M-1)}, \quad \mathbf{T} \in \mathbb{R}^M_{>0} \right\} \tag{5.24}$$

其中，$m = 3$，$\mathbf{q} = (\mathbf{q}_1, \cdots, \mathbf{q}_{M-1})$ 为中间点，$\mathbf{T} = [T_1, \cdots, T_M]^{\mathsf{T}}$ 为各段时间分配，$\mathbf{c} = (\mathbf{c}_1^{\mathsf{T}}, \cdots, \mathbf{c}_M^{\mathsf{T}})^{\mathsf{T}}$ 为各段多项式系数，$T_\Sigma = \sum_{i=1}^{M} T_i$ 为总轨迹时间。

这里采用五次多项式轨迹表示。第 i 段轨迹写为：

$$\mathbf{p}_i(t) = \mathbf{c}_i^{\mathsf{T}} \beta(t), \qquad t \in [0, T_i] \tag{5.25}$$

其中，$T_i = t_i - t_{i-1}$，$\mathbf{c}_i \in \mathbb{R}^{6\times m}$，$\beta(t) = [1, t, \cdots, t^5]^{\mathsf{T}}$ 为自然基。整条轨迹可表示为：

$$\mathbf{p}(t) = \mathbf{p}_i(t - t_{i-1}), \qquad t \in [t_{i-1}, t_i) \tag{5.26}$$

在求解过程中，以参考轨迹 ξ^* 的中间点与分段时间分别作为 $\mathbf{q}$ 和 $\mathbf{T}$ 的初值。

5.4.2　代价函数设计

为生成平滑、安全且满足动力学约束的轨迹，将各项约束以惩罚项的形式纳入目标函数，并采用无约束优化方法求解。轨迹优化问题可表述为：

$$\min_{\mathbf{q},\mathbf{T}} \left[J_t, J_o, J_e, J_d\right] \cdot \lambda \tag{5.27}$$

其中，$\lambda = \left[\lambda_t, \lambda_o, \lambda_e, \lambda_d\right]^T$ 为权重向量，分别对应执行时间惩罚、安全性惩罚、控制量惩罚和动力学可行性惩罚的权重。下面先给出各项代价对多项式系数 $\mathbf{c}$ 与时间分配 $\mathbf{T}$ 的梯度形式，再结合 $\mathbf{c} = \mathcal{M}(\mathbf{q}, \mathbf{T})$ 利用链式法则求得关于优化变量 $(\mathbf{q}, \mathbf{T})$ 的梯度。

对第 i 段轨迹，定义约束点为：

$$\mathring{\mathbf{p}}_{i,j} = \mathbf{p}_i\left(\frac{j}{\kappa_i}T_i\right), \qquad j \in \{0, 1, \cdots, \kappa_i\} \tag{5.28}$$

其中，κ_i 为第 i 段轨迹上的采样点数。

1. 执行时间惩罚 J_t

为提高轨迹执行效率，对总执行时间施加惩罚：

$$J_t = \sum_{i=1}^{M} T_i \tag{5.29}$$

其梯度为：

$$\frac{\partial J_t}{\partial \mathbf{c}} = \mathbf{0}, \qquad \frac{\partial J_t}{\partial \mathbf{T}} = \mathbf{1} \tag{5.30}$$

2. 安全性惩罚 J_o

基于欧几里得符号距离场（Euclidean Signed Distance Field，ESDF）构建安全性惩罚。定义采样点处的距离约束函数为：

$$\psi_o\left(\mathring{\mathbf{p}}_{i,j}\right) = \Xi_{\text{thr}} - \Xi\left(\mathring{\mathbf{p}}_{i,j}\right) \tag{5.31}$$

其中，Ξ_{thr} 为安全距离阈值，$\Xi\left(\mathring{\mathbf{p}}_{i,j}\right)$ 为采样点到障碍物的 ESDF 距离。则安全性惩罚定义为：

$$J_o = \sum_{i=1}^{M} \frac{T_i}{\kappa_i} \sum_{j=0}^{\kappa_i} \bar{\omega}_j \max\left(\psi_o\left(\mathring{\mathbf{p}}_{i,j}\right), 0\right)^3 \tag{5.32}$$

其中，$\bar{\omega}_j$ 为梯形求积系数，满足：

$$(\bar{\omega}_0, \bar{\omega}_1, \cdots, \bar{\omega}_{\kappa_i}) = \left(\frac{1}{2}, 1, \cdots, 1, \frac{1}{2}\right) \tag{5.33}$$

其梯度形式为：

$$\begin{aligned} \frac{\partial J_o}{\partial \mathbf{c}_i} &= \sum_{j=0}^{\kappa_i} \frac{\partial J_o}{\partial \psi_o}\frac{\partial \psi_o}{\partial \mathbf{c}_i} \\ \frac{\partial J_o}{\partial T_i} &= \sum_{j=0}^{\kappa_i} \frac{\partial J_o}{\partial \psi_o}\frac{\partial \psi_o}{\partial t}\frac{\partial t}{\partial T_i} + \sum_{j=0}^{\kappa_i} \frac{\bar{\omega}_j}{\kappa_i} \max\left(\psi_o\left(\mathring{\mathbf{p}}_{i,j}\right), 0\right)^3 \end{aligned} \tag{5.34}$$

其中：

$$\frac{\partial J_o}{\partial \psi_o} = 3\frac{T_i}{\kappa_i}\bar{\omega}_j \max\left(\psi_o\left(\mathring{\mathbf{p}}_{i,j}\right),0\right)^2$$

$$\frac{\partial \psi_o}{\partial \mathbf{c}_i} = -\beta(t)\nabla\Xi\left(\mathring{\mathbf{p}}_{i,j}\right)^T, \qquad \frac{\partial \psi_o}{\partial t} = -\nabla\Xi\left(\mathring{\mathbf{p}}_{i,j}\right)^T \mathbf{p}_i^{(1)}(t) \tag{5.35}$$

$$t = \frac{j}{\kappa_i}T_i, \qquad \frac{\partial t}{\partial T_i} = \frac{j}{\kappa_i}$$

3. 控制量惩罚 J_e

为保证轨迹平滑性并抑制高阶控制量，对轨迹的三阶导数施加惩罚：

$$J_e = \sum_{i=1}^{M}\int_0^{T_i}\left\|\mathbf{p}_i^{(3)}(t)\right\|^2 dt \tag{5.36}$$

其梯度为：

$$\frac{\partial J_e}{\partial \mathbf{c}_i} = 2\left(\int_0^{T_i}\beta^{(3)}(t)\beta^{(3)}(t)^T dt\right)\mathbf{c}_i$$

$$\frac{\partial J_e}{\partial T_i} = \beta^{(3)}(T_i)^T\mathbf{c}_i\mathbf{c}_i^T\beta^{(3)}(T_i) \tag{5.37}$$

4. 动力学可行性惩罚 J_d

为保证轨迹满足物理可行性要求，对最大速度、加速度和加加速度施加约束。定义约束函数为：

$$\mathcal{G}_* = \left\|\mathbf{p}_i^{(n)}(t)\right\|^2 - \max_*^2, \qquad * \in \{v,a,j\} \tag{5.38}$$

其中，当 $*=v,a,j$ 时，n 分别取 $1,2,3$，$\max_v$、$\max_a$ 和 $\max_j$ 分别表示最大允许速度、最大允许加速度和最大允许加加速度。则惩罚项表示为：

$$(J_d)_* = \sum_{i=1}^{M}\frac{T_i}{\kappa_i}\sum_{j=0}^{\kappa_i}\bar{\omega}_j \max\left(\mathcal{G}_*\left(\mathring{\mathbf{p}}_{i,j}\right),0\right)^3, \qquad * \in \{v,a,j\} \tag{5.39}$$

总动力学可行性惩罚为：

$$J_d = (J_d)_v + (J_d)_a + (J_d)_j \tag{5.40}$$

其梯度形式为：

$$\frac{\partial (J_d)_*}{\partial \mathbf{c}_i} = \sum_{j=0}^{\kappa_i}\frac{\partial (J_d)_*}{\partial \mathcal{G}_*}\frac{\partial \mathcal{G}_*}{\partial \mathbf{c}_i}$$

$$\frac{\partial (J_d)_*}{\partial T_i} = \sum_{j=0}^{\kappa_i}\frac{\partial (J_d)_*}{\partial \mathcal{G}_*}\frac{\partial \mathcal{G}_*}{\partial t}\frac{\partial t}{\partial T_i} + \sum_{j=0}^{\kappa_i}\frac{\bar{\omega}_j}{\kappa_i}\max\left(\mathcal{G}_*\left(\mathring{\mathbf{p}}_{i,j}\right),0\right)^3 \tag{5.41}$$

其中：

$$\frac{\partial (J_d)_*}{\partial \mathcal{G}_*} = 3\frac{T_i}{\kappa_i}\bar{\omega}_j \max\left(\mathcal{G}_*\left(\mathring{\mathbf{p}}_{i,j}\right),0\right)^2$$

$$\frac{\partial \mathcal{G}_*}{\partial \mathbf{c}_i} = 2\beta^{(n)}(t)\mathbf{p}_i^{(n)}(t)^T, \qquad \frac{\partial \mathcal{G}_*}{\partial t} = 2\beta^{(n+1)}(t)^T\mathbf{c}_i\mathbf{p}_i^{(n)}(t) \tag{5.42}$$

$$t = \frac{j}{\kappa_i}T_i, \qquad \frac{\partial t}{\partial T_i} = \frac{j}{\kappa_i}$$

5.4.3　自适应权重调节

人类信任状态会影响操作员对轨迹规划结果的接受程度及干预意愿。当信任水平较低时，操作员容易进行不必要的干预，因此轨迹规划应更加侧重安全性，以生成相对保守的飞行轨迹；当信任水平较高时，操作员更容易接受轨迹规划结果，轨迹规划应更加关注任务执行效率。基于此，优化目标中的时间权重 λ_t 和安全权重 λ_o 根据人类信任水平 $T_m^h(t)$ 进行动态调整，其形式为：

$$\begin{aligned} \lambda_t &= \lambda_{t,0} + \Delta\lambda_t\left(T_m^h(t)\right) \\ \lambda_o &= \lambda_{o,0} + \Delta\lambda_o\left(T_m^h(t)\right) \end{aligned} \tag{5.43}$$

根据信任水平的不同，可将其划分为低信任区间 $[0, T_L)$、中等信任区间 $[T_L, T_H)$ 和高信任区间 $[T_H, \infty)$。在不同信任区间内采用如下权重调节策略：当信任水平处于低信任区间时，适当增大安全权重 $\Delta\lambda_o$，并减小时间权重 $\Delta\lambda_t$；当信任水平处于中等信任区间时，逐步增加效率相关权重并降低安全性相关权重；当信任水平处于高信任区间时，适当增大时间权重 $\Delta\lambda_t$，以提升搜索效率。

为实现权重在不同信任区间之间的平滑过渡，调节函数采用 Sigmoid 形式构造，如图 5.4 所示：

$$\begin{aligned} \Delta\lambda_t\left(T_m^h(t)\right) &= \frac{\Lambda_t}{1 + e^{-k_1\left(T_m^h(t) - T_H\right)}} \\ \Delta\lambda_o\left(T_m^h(t)\right) &= \frac{\Lambda_o}{1 + e^{k_2\left(T_m^h(t) - T_L\right)}} \end{aligned} \tag{5.44}$$

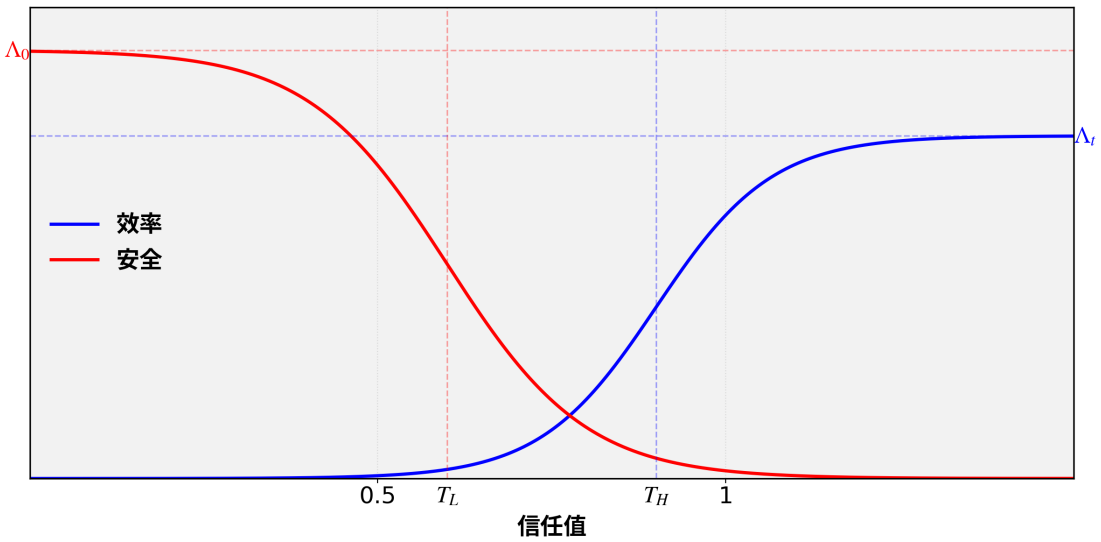


图 5.4　轨迹优化权重调节函数示意图

5.5　仿真实验与结果分析

为验证所提方法的有效性，基于 ROS 机器人软件框架搭建了随机森林仿真场景，并在该场景下开展人机协同仿真实验，同时与未考虑信任因素的偏置增量动作采样（Biased Incremental Action Sampling, BIAS）[59] 轨迹规划基准方法进行对比。实验重点分析两类方法在任务效率、操作员负荷、轨迹质量以及信任水平等方面的差异。

5.5.1　实验设置

为评估所提方法在不同环境复杂度下的适用性，实验设置了三种树木密度不同的随机森林场景。仿真区域为长 70 m、宽 20 m、高 3 m 的矩形空间，其中分别布置 50 棵、100 棵和 200 棵树木，以对应稀疏、中等和密集三类环境。

实验共招募 15 名年龄在 22–28 岁的志愿者参与，受试者均无无人机操控经验。正式实验开始前，每位受试者接受统一说明，并进行 5 分钟熟悉练习。实验任务为：在森林环境中控制四旋翼无人机从区域一端飞行至另一端，当无人机到达目标区域时视为任务完成。受试者通过第一人称视角观察环境，并通过操纵杆输入控制指令。

为评估两类方法的性能，实验从以下两个方面进行比较：

1. 任务执行效率：包括飞行距离、飞行时长以及平均速度；
2. 人机协同表现：包括操作员输入次数、Jerk 积分以及平均信任水平。

其中，操作员输入次数用于反映人工修正频率及操作负担；Jerk 积分用于反映轨迹变化的剧烈程度和平滑性；平均信任水平用于反映操作员在任务过程中的整体信任程度。一般而言，操作员输入次数越高，说明人工修正越频繁、操作负担越重；Jerk 积分越高，则说明轨迹变化越剧烈、平滑性越差。两项指标可在一定程度上反映人机协同过程的稳定性与协调性。

实验中的运动基元采样参数设置如下：基元持续时间 $\boldsymbol{T}$ 的取值范围为 [0.3, 1.5] s，离散化数量为 5；角速度 ω 的取值范围为 $[-0.75, 0.75]$ rad/s，离散化数量为 15；垂直速度 v_z 的取值范围为 $[-0.75, 0.75]$ m/s，离散化数量为 3。其余参数设置如表 5.1 所示。

5.5.2　实验结果与分析

表 5.2 给出了两种方法在不同密度的随机森林环境下的实验结果。为更直观地展示两类方法的轨迹差异，图 5.5 和图 5.6 分别给出了对应的二维与三维飞行轨迹对比。

从任务执行效率来看，两种方法在飞行距离上总体接近，说明本章方法并未明显增加完成任务所需的路径长度。飞行时长方面，本章方法在三类环境中均低

表 5.1　协同搜索实验关键参数设置

参数	值
操作员输入最大线速度	2 m/s
每轮扩展节点数 J	2
候选子节点保留数 K	5
球形区域总数 N	20
FOV 大小参数 ρ	0.8
风险敏感度调节系数 γ_s	2
初始人类因子 α_i	0.5
动态观测因子 α_d	0.92
正向认知更新因子 α_h^+	0.09
负向认知更新因子 α_h^-	0.42
初始客观机器能力 C_m^m	0.85

表 5.2　随机森林环境下的实验结果对比

方法	距离 (m)	时长 (s)	平均速度 (m/s)	Jerk 积分 (m^2/s^3)	输入次数	平均信任水平
稀疏环境 (50 棵树)						
基准方法	71.15 ± 0.46	46.62 ± 1.49	1.53 ± 0.05	27.49 ± 1.63	39 ± 4	0.744 ± 0.013
本章方法	**70.04 ± 0.63**	**45.53 ± 0.47**	**1.54 ± 0.02**	**21.30 ± 3.39**	**34 ± 3**	**0.813 ± 0.016**
中等环境 (100 棵树)						
基准方法	71.33 ± 1.16	48.88 ± 1.71	1.46 ± 0.06	44.88 ± 10.38	45 ± 7	0.657 ± 0.015
本章方法	**70.29 ± 2.44**	**45.84 ± 2.39**	**1.52 ± 0.03**	**31.48 ± 10.60**	**34 ± 5**	**0.750 ± 0.013**
密集环境 (200 棵树)						
基准方法	71.73 ± 1.71	54.89 ± 5.29	1.31 ± 0.14	84.69 ± 9.68	56 ± 8	0.611 ± 0.017
本章方法	**71.38 ± 2.00**	**49.86 ± 1.05**	**1.43 ± 0.08**	**48.11 ± 9.44**	**43 ± 5**	**0.693 ± 0.023**

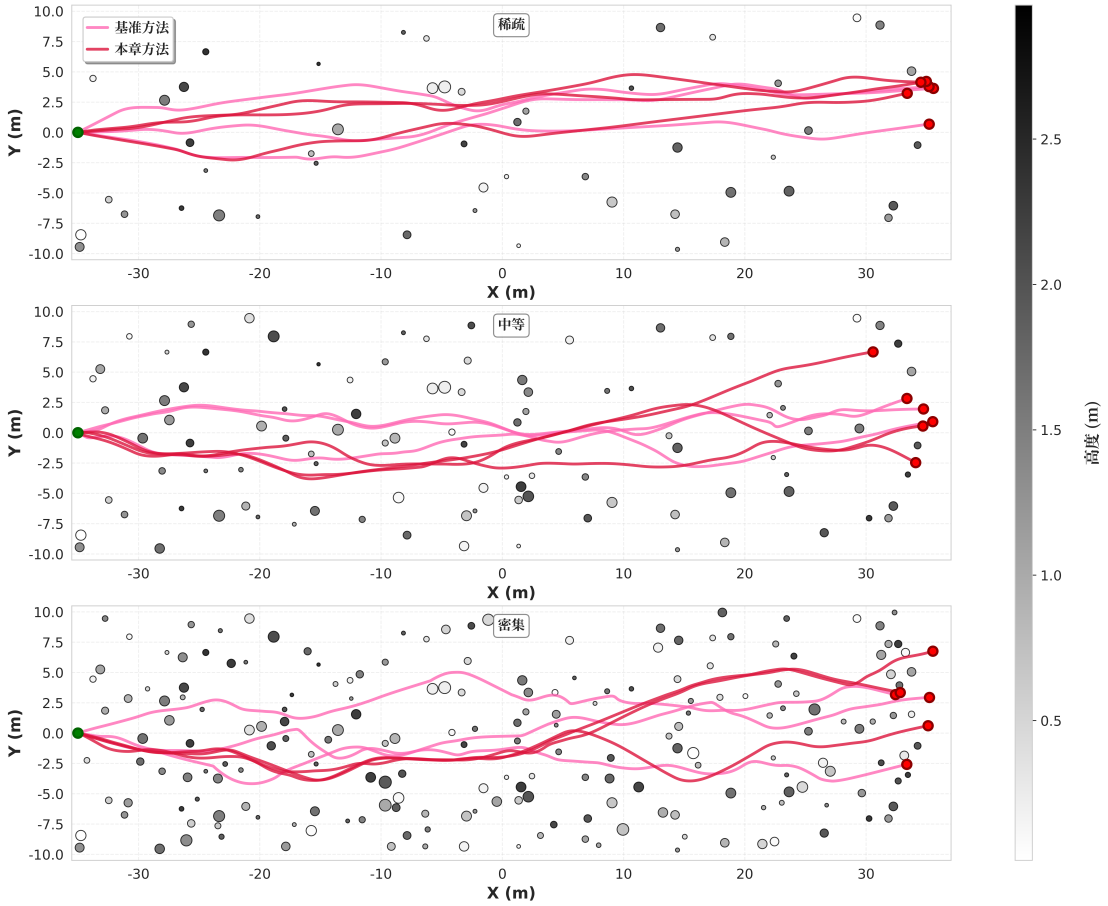


图 5.5　不同密度随机森林环境下两种轨迹规划方法的二维飞行轨迹对比图

于基准方法，尤其在中等和密集环境中完成时间更短；相应地，其平均速度也略高于基准方法。这表明，在飞行距离总体接近的情况下，本章方法能够在一定程度上提升任务执行效率。

从操作员负荷来看，本章方法在所有环境中的输入次数均低于基准方法。在稀疏环境中，输入次数由 39 次降至 34 次；在中等环境中，由 45 次降至 34 次；在密集环境中，则由 56 次降至 43 次。随着障碍物密度增加，两类方法的输入次数均有所增加，但本章方法的增长幅度相对较小。由此可见，考虑信任的轨迹规划方法能够在一定程度上减少操作员的人工修正次数，从而减轻操作员的负担。

从轨迹质量来看，本章方法在三类环境下的 Jerk 积分均低于基准方法，表现出更好的轨迹平滑性。具体而言，在稀疏、中等和密集环境中，本章方法的 Jerk 积分分别为 21.30、31.48 和 48.11，均低于基准方法对应的 27.49、44.88 和 84.69。结合图 5.5 与图 5.6 可以看出，基准方法在障碍物较多时更容易出现局部绕行和频繁方向调整等现象，而本章方法生成的轨迹整体更平滑，路径过渡也更为连贯。

从信任水平来看，本章方法在三类环境中的平均信任水平均高于基准方法。在稀疏环境中，平均信任水平由 0.744 提升至 0.813；在中等环境中，由 0.657 提

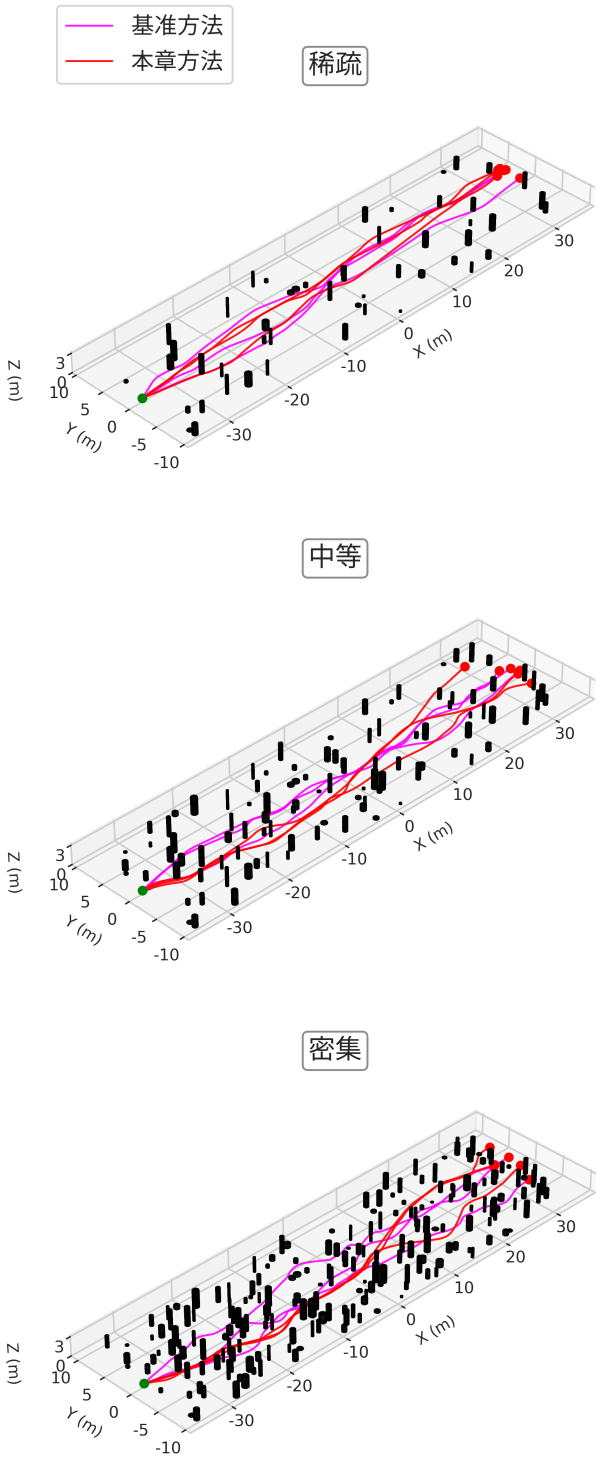


图 5.6　不同密度随机森林环境下两种轨迹规划方法的三维飞行轨迹对比图

升至 0.750；在密集环境中，由 0.611 提升至 0.693。随着环境复杂度增加，两类方法的信任水平均有所下降，但本章方法始终高于基准方法，且下降幅度略小。这一结果说明，在障碍物密度增大、环境更复杂的情况下，在轨迹规划中考虑信任仍有助于维持较高的操作员信任水平。

综合上述结果可以看出，在飞行距离总体接近的情况下，本章方法在任务执行效率、操作员输入次数、轨迹平滑性和平均信任水平等方面均优于基准方法，且在中等和密集环境中的优势更为明显。当人类信任水平较高时，操作员对规划轨迹的干预意愿降低，因而更容易接受轨迹规划结果，进而减少了人工修正次数并减轻了操作负荷。这进一步表明，将人类信任状态融入轨迹规划方法设计后，能够使规划结果更符合操作员预期，从而提升人机协同过程的稳定性。

5.6　本章小结

本章针对人机协同搜索任务中轨迹规划对信任动态变化考虑不足的问题，提出了融合信任的人机协同轨迹规划方法。围绕协同搜索任务，构建了安全性与可视性评价指标，并据此计算机器性能表现与客观机器能力；基于人类意图生成引导运动基元，扩展运动基元树并构造候选轨迹集；进一步将操作员信任状态引入轨迹选择与优化过程，实现对规划轨迹的自适应调节。仿真实验结果表明，所提方法在保持任务执行效率稳定的同时，能够降低操作员负担，提升飞行轨迹平滑性，为人机协同搜索任务提供了有效的轨迹规划方法。

第 6 章　总结与展望

6.1　总结

本文围绕人-无人机协同系统中的人机信任问题开展研究，以人-无人机协同搜索与协同降落两类典型任务作为研究对象，围绕信任建模、控制权分配与轨迹规划三个关键问题，提出了机器性能驱动的信任演化模型以及融合信任的协同方法，从而提升人-无人机系统的协同效能。本文的主要研究工作和创新点可总结如下：

（1）针对现有人机信任模型可解释性不足、难以反映机器实时性能波动对人类信任动态影响的问题，提出了一种机器性能驱动的信任演化模型。该模型从机器客观能力与人类主观认知之间可能存在的偏差出发，将机器性能表现与人类认知更新过程相结合，构建了适用于人-无人机协同任务的动态信任演化模型。通过对模型核心参数物理意义及其数学性质的分析，增强了模型的可解释性与合理性。人机交互实验结果表明，所提出的模型能够较好刻画人机交互过程中信任水平的动态变化，为后续融合信任的协同方法设计提供了模型基础。

（2）针对人机协同降落任务中控制权分配缺乏合理调节机制的问题，提出了一种融合信任的最小干预共享控制策略。该策略结合协同降落任务特点，构建了意图推理、降落精度和任务安全性三类任务评价指标，并据此计算机器能力与机器性能表现。在此基础上，利用深度强化学习建立动作价值评估网络，构造候选动作集，并根据信任状态动态调节候选动作的价值损失阈值，最终在满足约束的前提下，依据最小干预原则选择与人类输入偏差最小的控制动作。实验结果表明，该策略能够在尽可能保留人类控制意图的同时，提高协同降落任务的成功率与任务质量。

（3）针对人机协同搜索任务中轨迹规划方法忽略人类信任动态变化的问题，提出了一种融合信任的人机协同轨迹规划方法。该方法结合协同搜索任务特点，构建了安全性与可视性两类任务评价指标，并据此计算机器能力与机器性能表现。在此基础上，根据人类输入生成引导基元，并在安全约束下扩展运动基元树，构造候选轨迹集；进一步将人类信任状态融入轨迹选择与优化过程，从而在兼顾人类意图的同时，实现融合信任的人机协同轨迹规划。仿真实验结果表明，该方法能够在降低操作员负荷的同时，提升飞行轨迹平滑性，从而改善人机协作的整体表现。

6.2　展望

尽管本文围绕人-无人机协同中的信任建模与应用开展了较为系统的研究，并取得了一定成果，但仍有若干问题值得进一步探讨，主要体现在以下几个方面：

（1）提升信任模型的泛化能力与在线适应能力。本文构建的机器性能驱动信任演化模型主要基于任务执行过程中可观测的性能指标对人类信任状态进行刻画，能够较好反映特定任务场景下的信任动态。然而，在更复杂的实际应用环境中，影响信任变化的因素往往更加多样，除机器性能外，还可能受到任务风险、环境复杂度、人类经验水平以及交互方式等因素的共同影响。未来可引入多维影响因素，构建更具通用性的信任建模框架，并结合在线参数辨识与自适应学习方法，提高模型在不同任务场景与不同个体条件下的适应能力。

（2）增强融合信任的人机协同方法在真实场景中的适用性。本文针对协同降落与协同搜索任务分别设计了融合信任的共享控制方法与轨迹规划方法，并在仿真环境中验证了其有效性。然而，相较于仿真场景，真实任务环境通常具有更强的不确定性，感知误差、通信时延以及外部扰动等因素都可能影响信任估计的准确性和协同方法的稳定性。未来可进一步研究融合时延补偿、鲁棒优化的人机协同方法，并结合真实无人机平台开展实验验证，以推动相关研究向工程应用转化。

（3）拓展融合信任方法的应用场景与系统规模。本文主要面向单操作员与单无人机条件下的协同任务展开研究。未来在应急救援、巡检监测等复杂任务中，往往涉及多无人机协同、多任务耦合甚至多操作员协同决策。随着系统规模的扩大，人机之间以及机群内部的交互关系将更加复杂，也将面临更高的建模与设计难度。未来可在此基础上进一步探索多无人机协同场景下的群体信任建模、基于信任的任务分配与协同决策方法，以提升复杂系统中的整体协作效率与鲁棒性。

总体而言，本文的研究表明，信任是连接人类认知决策与无人机自主执行能力的重要桥梁。未来可围绕信任动态建模、协同方法优化以及真实系统验证等方面持续开展研究，进一步推动融合信任的人-无人机协同控制理论与方法向智能化、实用化和面向复杂场景应用的方向发展。

参考文献

[1] Ramezani M, Atashgah M, Sanchez-Lopez J L, et al. Human-centric aware UAV trajectory planning in search and rescue missions employing multi-objective reinforcement learning with AHP and similarity-based experience replay[C]//Proceedings of the 2024 International Conference on Unmanned Aircraft Systems. 2024: 177-184.

[2] Abdelmaksoud S I, Mailah M, Abdallah A M. Control strategies and novel techniques for autonomous rotorcraft unmanned aerial vehicles: a review[J]. IEEE Access, 2020, 8: 195142-195169.

[3] Villa D K, Brandao A S, Sarcinelli-Filho M. A survey on load transportation using multirotor UAVs[J]. Journal of Intelligent & Robotic Systems, 2020, 98(2): 267-296.

[4] Boursianis A D, Papadopoulou M S, Diamantoulakis P, et al. Internet of things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: a comprehensive review[J]. Internet of Things, 2022, 18: 100187.

[5] Li X, Savkin A V. Networked unmanned aerial vehicles for surveillance and monitoring: a survey[J]. Future Internet, 2021, 13(7): 174.

[6] Saeed A S, Bani Younes A, Cai C, et al. A survey of hybrid unmanned aerial vehicles[J]. Progress in Aerospace Sciences, 2018, 98: 91-105.

[7] 江波, 屈若锟, 李彦冬, 等. 基于深度学习的无人机航拍目标检测研究综述[J]. 航空学报, 2021, 42(4): 137-151.

[8] 刘森, 姜雪松, 张宇晨, 等. 林火搜救多无人机协同任务分配方法[J]. 中国新技术新产品, 2024(6): 133-136.

[9] Murphy R R, Tadokoro S, Nardi D, et al. Search and rescue robotics[M]//Siciliano B, Khatib O. Springer Handbook of Robotics. Berlin, Heidelberg: Springer, 2008: 1151-1173.

[10] 工业和信息化部, 国家发展和改革委员会, 科学技术部, 等. 安全应急装备重点领域发展行动计划（2023–2025 年）: 工信部联安全〔2023〕166 号[R]. 工业和信息化部, 2023.

[11] U.S. Department of Defense. Unmanned systems integrated roadmap FY 2017–2042[R]. Washington, DC: Office of the Secretary of Defense, 2018.

[12] Scott B I, Andritsos K I. A drone strategy 2.0 for a smart and sustainable unmanned aircraft eco-system in Europe[J]. Air and Space Law, 2023, 48(3): 273-296.

[13] Zhang Z, Chen H, Lye S W, et al. Human-guided safe and efficient trajectory replanning for unmanned aerial vehicles[C]//Proceedings of the 2022 International Conference on Control, Automation and Diagnosis. 2022: 1-6.

[14] Luo Y, Yu H, Zhang H, et al. A novel Newton–Euler method-based nonlinear anti-swing control for a quadrotor UAV carrying a slung load[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2024, 54(4): 2266-2275.

[15] He B, Ji X, Li G, et al. Key technologies and applications of UAVs in underground space: a review[J]. IEEE Transactions on Cognitive Communications and Networking, 2024, 10(3): 1026-1049.

[16] 赵云波, 康宇, 朱进. 人机混合智能系统自主性理论和方法[M]. 北京: 科学出版社, 2021.

[17] Pfeiffer C, Scaramuzza D. Human-piloted drone racing: visual processing and control[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 3467-3474.

[18] Eschmann J, Albani D, Loianno G. Learning to fly in seconds[J]. IEEE Robotics and Automation Letters, 2024, 9(7): 6336-6343.

[19] Gunia A. The role of trust in human-machine interaction: cognitive science perspective[M]// Artificial Intelligence, Management and Trust. London, U.K.: Routledge, 2024: 85-126.

[20] Ibrahim B, Elhajj I H, Asmar D. 3D autocomplete: enhancing UAV teleoperation with AI in the loop[C]//Proceedings of the 2024 IEEE International Conference on Robotics and Automation. 2024: 17829-17835.

[21] Steyvers M, Tejeda H, Kerrigan G, et al. Bayesian modeling of human–AI complementarity [J]. Proceedings of the National Academy of Sciences, 2022, 119(11): e2111547119.

[22] Lyons J B, Stokes C K. Human–human reliance in the context of automation[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2012, 54(1): 112-121.

[23] Nam C, Walker P, Li H, et al. Models of trust in human control of swarms with varied levels of autonomy[J]. IEEE Transactions on Human-Machine Systems, 2020, 50(3): 194-204.

[24] Gebru B, Zeleke L, Blankson D, et al. A review on human–machine trust evaluation: human-centric and machine-centric perspectives[J]. IEEE Transactions on Human-Machine Systems, 2022, 52(5): 952-962.

[25] Hoff K A, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2015, 57(3): 407-434.

[26] Lee J D, See K A. Trust in automation: designing for appropriate reliance[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2004, 46(1): 50-80.

[27] Kok B C, Soh H. Trust in robots: challenges and opportunities[J]. Current Robotics Reports, 2020, 1(4): 297-309.

[28] Jian J Y, Bisantz A M, Drury C G. Foundations for an empirically determined scale of trust in automated systems[J]. International Journal of Cognitive Ergonomics, 2000, 4(1): 53-71.

[29] Schaefer K E. Measuring trust in human robot interactions: development of the trust percep-

tion scale-HRI[M]. Boston, MA: Springer, 2016.

[30] Ullman D, Malle B F. What does it mean to trust a robot? steps toward a multidimensional measure of trust[C]//Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. 2018: 263-264.

[31] Wojton H M, Porter D, Lane S T, et al. Initial validation of the trust of automated systems test (TOAST)[J]. The Journal of Social Psychology, 2020, 160(6): 735-750.

[32] Khalid H M, Liew W S, Nooralishahi P, et al. Exploring psycho-physiological correlates to trust: implications for human-robot-human interaction[C]//Proceedings of the Human Factors and Ergonomics Society Annual Meeting: Vol. 60. SAGE Publications, 2016: 697-701.

[33] Goodyear K, Parasuraman R, Chernyak S, et al. An fMRI and effective connectivity study investigating miss errors during advice utilization from human and machine agents[J]. Social Neuroscience, 2017, 12(5): 570-581.

[34] Palmer S, Richards D, Shelton-Rayner G, et al. Human-agent teaming - an evolving interaction paradigm: an innovative measure of trust[C]//20th International Symposium on Aviation Psychology. 2019: 438-443.

[35] Akash K, Hu W L, Jain N, et al. A classification model for sensing human trust in machines using EEG and GSR[J]. ACM Transactions on Interactive Intelligent Systems, 2018, 8(4): 1-20.

[36] Park C, Shahrdar S, Nojoumian M. EEG-based classification of emotional state using an autonomous vehicle simulator[C]//Proceedings of the 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop. 2018: 297-300.

[37] Dong S Y, Kim B K, Lee K, et al. A preliminary study on human trust measurements by EEG for human-machine interactions[C]//Proceedings of the 3rd International Conference on Human-Agent Interaction. 2015: 265-268.

[38] Gupta K, Hajika R, Pai Y S, et al. Measuring human trust in a virtual assistant using physiological sensing in virtual reality[C]//Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces. 2020: 756-765.

[39] Hu W L, Akash K, Reid T, et al. Computational modeling of the dynamics of human trust during human–machine interactions[J]. IEEE Transactions on Human-Machine Systems, 2019, 49(6): 485-497.

[40] Hudspeth M, Balali S, Grimm C, et al. Effects of interfaces on human-robot trust: specifying and visualizing physical zones[C]//Proceedings of the 2022 International Conference on Robotics and Automation. IEEE, 2022: 11265-11271.

[41] Lee J, Moray N. Trust, control strategies and allocation of function in human-machine systems [J]. Ergonomics, 1992, 35(10): 1243-1270.

[42] Sadrfaridpour B, Saeidi H, Burke J, et al. Modeling and control of trust in human-robot collaborative manufacturing[M]//Robust Intelligence and Trust in Autonomous Systems. Boston, MA: Springer, 2016: 115-141.

[43] Hoogendoorn M, Jaffry S W, Treur J. Cognitive and neural modeling of dynamics of trust in competitive trustees[J]. Cognitive Systems Research, 2012, 14(1): 60-83.

[44] Gao J, Lee J D. Extending the decision field theory to model operators' reliance on automation in supervisory control situations[J]. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2006, 36(5): 943-959.

[45] Akash K, Hu W L, Reid T, et al. Dynamic modeling of trust in human-machine interactions [C]//Proceedings of the 2017 American Control Conference. 2017: 1542-1548.

[46] Xu A, Dudek G. OPTIMo: online probabilistic trust inference model for asymmetric human-robot collaborations[C]//Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. ACM, 2015: 221-228.

[47] Akash K, Polson K, Reid T, et al. Improving human-machine collaboration through transparency-based feedback–part I: human trust and workload model[J]. IFAC-PapersOnLine, 2019, 51(34): 315-321.

[48] Hu W L, Akash K, Jain N, et al. Real-time sensing of trust in human-machine interactions[J]. IFAC-PapersOnLine, 2016, 49(32): 48-53.

[49] Huang S H, Bhatia K, Abbeel P, et al. Establishing appropriate trust via critical states[C]// Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2018: 3929-3936.

[50] Rezvani T, Driggs-Campbell K, Sadigh D, et al. Towards trustworthy automation: user interfaces that convey internal and external awareness[C]//Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems. IEEE, 2016: 682-688.

[51] Okamura K, Yamada S. Empirical evaluations of framework for adaptive trust calibration in human-AI cooperation[J]. IEEE Access, 2020, 8: 220335-220351.

[52] Zahedi Z, Verma M, Sreedharan S, et al. Trust-aware planning: modeling trust evolution in iterated human-robot interaction[C]//Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction. ACM, 2023: 281-289.

[53] Li Y, Cui R, Yan W, et al. Reconciling conflicting intents: bidirectional trust-based variable autonomy for mobile robots[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5615-5622.

[54] Hu C, Shi Y, Ge S, et al. Trust-based shared control of human-vehicle system using model free adaptive dynamic programming[J]. IEEE Transactions on Intelligent Vehicles, 2025, 10 (7): 4103-4115.

[55] Nieuwenhuisen M, Droeschel D, Schneider J, et al. Multimodal obstacle detection and collision avoidance for micro aerial vehicles[C]//Proceedings of the 2013 European Conference on Mobile Robots. 2013: 7-12.

[56] Odelga M, Stegagno P, Bülthoff H H. Obstacle detection, tracking and avoidance for a teleoperated UAV[C]//Proceedings of the 2016 IEEE International Conference on Robotics and Automation. 2016: 2984-2990.

[57] Ghaffari S. Collision avoidance assistance in UAV teleoperation[D]. Hamilton, ON, Canada: McMaster University, 2021.

[58] Backman K, Kulić D, Chung H. Learning to assist drone landings[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 3192-3199.

[59] Yang X, Michael N. Assisted mobile robot teleoperation with intent-aligned trajectories via biased incremental action sampling[C]//Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2020: 10998-11003.

[60] Yang X, Cheng J, Michael N. An intention guided hierarchical framework for trajectory-based teleoperation of mobile robots[C]//Proceedings of the 2021 IEEE International Conference on Robotics and Automation. 2021: 482-488.

[61] Wang Q, Liu D, Carmichael M G, et al. Computational model of robot trust in human co-worker for physical human-robot collaboration[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 3146-3153.

[62] Shi H, Luo L, Gao S, et al. Human-guided motion planner with perception awareness for assistive aerial teleoperation[J]. Advanced Robotics, 2024, 38: 152-167.

[63] Inagaki T. Adaptive automation: sharing and trading of control[M]//Handbook of Cognitive Task Design. Boca Raton, FL, USA: CRC Press, 2003: 171-194.

[64] Abbink D A, Carlson T, Mulder M, et al. A topology of shared control systems—finding common ground in diversity[J]. IEEE Transactions on Human-Machine Systems, 2018, 48 (5): 509-525.

[65] Li G, Li Q, Yang C, et al. The classification and new trends of shared control strategies in telerobotic systems: a survey[J]. IEEE Transactions on Haptics, 2023, 16(2): 118-133.

[66] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.

[67] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[A/OL]. arXiv (2017). https://arxiv.org/abs/1707.06347.

[68] Jain S, Argall B. Recursive bayesian human intent recognition in shared-control robotics [C]//Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2018: 3905-3912.

[69] Zhang Q, Xu Z, Wang Y, et al. Predicted trajectory guidance control framework of teleoperated ground vehicles compensating for delays[J]. IEEE Transactions on Vehicular Technology, 2023, 72(9): 11264-11274.

[70] Zhang Q, Yang L, Huang Z, et al. A delay compensation framework based on eye-movement for teleoperated ground vehicles[J]. IEEE Transactions on Vehicular Technology, 2025, 74 (1): 390-401.

[71] Giuliari F, Hasan I, Cristani M, et al. Transformer networks for trajectory forecasting[C]// 2020 25th International Conference on Pattern Recognition. 2021: 10335-10342.

[72] Hancock P A, Billings D R, Schaefer K E, et al. A meta-analysis of factors affecting trust in human-robot interaction[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2011, 53(5): 517-527.

[73] Xia R, Zhao Y B, Lu J, et al. Trust-modulated authority allocation in human-guided goal recognition tasks[C]//Proceedings of the 2nd International Conference on Artificial Intelligence and Human-Computer Interaction. IOS Press, 2024: 516-522.

[74] Tai J J, Wong J, Innocente M, et al. PyFlyt–UAV simulation environments for reinforcement learning research[A/OL]. arXiv (2023). https://arxiv.org/abs/2304.01305.

[75] Nikolaidis S, Hsu D, Srinivasa S. Human-robot mutual adaptation in collaborative tasks: models and experiments[J]. The International Journal of Robotics Research, 2017, 36(5-7): 618-634.

[76] Fujimoto S, van Hoof H, Meger D. Addressing function approximation error in actor-critic methods[C]//Proceedings of the 35th International Conference on Machine Learning. PMLR, 2018: 1587-1596.

[77] Broad A, Murphey T, Argall B. Highly parallelized data-driven MPC for minimal intervention shared control[C]//Proceedings of Robotics: Science and Systems XV. Robotics: Science and Systems Foundation, 2019: 1-10.

[78] Yang X, Agrawal A, Sreenath K, et al. Online adaptive teleoperation via motion primitives for mobile robots[J]. Autonomous Robots, 2019, 43(6): 1357-1373.

[79] Mellinger D, Kumar V. Minimum snap trajectory generation and control for quadrotors[C]// 2011 IEEE International Conference on Robotics and Automation. 2011: 2520-2525.

[80] Eiter T, Mannila H. Computing discrete Fréchet distance: CD-TR 94/64[R]. Christian Doppler Laboratory for Expert Systems, 1994.

[81] Wang Z, Zhou X, Xu C, et al. Geometrically constrained trajectory optimization for multicopters[J]. IEEE Transactions on Robotics, 2022, 38(5): 3259-3278.

致谢

三年的研究生生活转瞬即逝，离家求学的日子也即将告一段落。回首在科大的这段时光，充实而又难忘。我不仅收获了扎实的专业知识，更在心智上得到了极大的成长与蜕变。在此，我想对所有关心、支持和帮助过我的人表达最真挚的感谢。

首先，我要特别感谢我的导师赵云波老师。感谢赵老师当年给予我进入科大求学的宝贵机会。您是我科研道路上的引路人，不仅以开阔的学术视野指引我前行，更以身作则，培养了我严谨求实的学术态度。从教导我如何去正确地看待科学研究及其背后的逻辑，到面对具体研究问题时的一丝不苟，再到悉心指导我如何将复杂的学术问题深入浅出地讲清楚，您的每一次点拨都让我受益良多。这份深厚的师恩，我定当铭记于心。

其次，我要诚挚地感谢李鹏飞老师。这三年里，李老师无论是在学术上还是在生活上，都对我有着无微不至的关照。从最初寻找研究方向、梳理复杂的逻辑思路，到后来一遍遍地指导论文写作，李老师都倾注了大量的心血。每当我遇到科研上的挫折和困难，甚至陷入焦虑时，您总能耐心地开导我、鼓励我，帮我重新找回方向和信心。能遇到您这样良师益友般的长辈，是我莫大的幸运。

此外，我也要感谢一路同行的朋友们和实验室的同门。感谢大家在科研遇到瓶颈时的热烈讨论与无私帮助，让枯燥的代码和晦涩的公式变得不再那么令人头疼；也感谢大家在生活中的陪伴与倾听，分担了我的压力，见证了彼此的成长。正是有了你们的鼓励与支持，这段或许有些单调的科研岁月才变得如此丰富多彩、充满温度。

我还要将最深的谢意献给我的父母。二十多年来，你们始终如一地做我最坚强的后盾。无论是在求学路上的重重考验，还是在生活中的方方面面，你们坚定不移的支持与无条件的关爱，永远是我不断前行、克服困难的最大动力。如今我已学有所成，即将走向社会，未来的岁月里，一定会换我来好好守护你们。

纸短情长，言不尽意。祝愿所有的师长、亲友及同窗，身体健康，工作顺利。愿自己永葆热忱，不忘初心，砥砺前行。

在读期间取得的科研成果

会议论文：

[1] **Huang Kangjie**, Zhang Wen, Wang Ruoshan, Jin Xiaoran, Zhao Yun-Bo, Li Pengfei. Trust-Aware Minimal-Intervention Shared Control via Deep Reinforcement Learning[C]. 2026 6th International Conference on Neural Networks, Information and Communication Engineering, Hefei, China, 2026.