

分类号 _____

U D C _____

密 级 _____

编 号 10486

武汉大学

博 士 学 位 论 文

基于模仿学习的机械臂运动规划与柔顺
控制研究

研 究 生 姓 名: 高 霄

指导教师姓名、职称: 肖晓晖 教授

学 科、专 业 名 称: 机械电子工程

研 究 方 向: 机器人学

二〇二一年五月

Research of Motion Planning and Compliant Control for Robot Manipulators Based on Imitation Learning

Candidate: GAO XIAO

Supervisor: PROF. XIAO XIAOHUI

Major: Mechatronic Engineering

Speciality: Robotics



School of Power and Mechanical Engineering

WUHAN UNIVERSITY

May, 2021

论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的研究成果。除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

学位论文作者（签名）：

年 月 日

论文创新点

1. 提出一种融合位置和姿态的微分同胚映射算法，实现了模仿学习中复杂示教轨迹可逆且连续可微的位姿变换。该算法分析了映射的基本约束条件，提出了融合位置和姿态的映射框架，可保证位置和姿态的同步变换。在该框架下，设计了两种实施方式：基于径向基函数迭代的映射算法，和基于可逆神经网络的映射算法，并在仿真和实验中验证其有效性。
2. 提出一种基于模仿学习和微分同胚映射的运动规划算法，实现了动态非结构化环境下的机械臂实时运动规划。该算法建立模仿学习中的人工示教轨迹和隐空间中简单轨迹之间的微分同胚映射，并将隐空间中全局渐进稳定的动态系统映射至示教空间，从而构建出基于动态系统的运动规划算法。该算法实现了融合位置和姿态且仅基于系统状态的运动规划，且具有内在的在线适应性和全局稳定性。
3. 提出一种基于模仿学习的多模态任务表达和学习方法，实现了复杂接触环境下机械臂的自主拟人化柔顺操作。通过采集人执行任务过程中的数据（位置、速度、交互力和交互刚度），建立表达人类多模态操作技能的概率统计模型，用于规划机械臂的运动、力和刚度，并基于该映射模型，设计了典型接触式任务中的自适应位置/力控制策略。结合上述规划和控制策略实现接触环境下机械臂的自主拟人化柔顺操作。

摘 要

机器人广泛应用于人们的生产生活中,将人们从单调重复性的工作中解放出来,大大提高了生产效率和质量。然而,现阶段的机器人灵巧操作水平和智能决策能力仍远低于人类,仅能应用于结构化环境下重复性的任务操作。针对动态非结构化环境下机器人的操作,难以设计具有高适应性和稳定性的运动规划算法。针对复杂接触环境下的任务操作,其位置和力的规划与控制仍需针对特定的任务进行调试,难以推广应用于不同场景。因此,本文提出基于模仿学习的机械臂运动规划与柔顺控制方法,用以提高机器人的易用性和智能性,实现复杂场景下机器人自主拟人化操作。

为提高机器人的易用性和对环境的适应性,可采用基于动态系统的机械臂运动规划方法,但传统的动态系统模型设计过程较为复杂。本文提出了基于微分同胚映射的动态系统设计方法,可以实现单步快速训练。从动态系统的约束条件出发,分析微分同胚映射的约束,并提出了统一的融合位置和姿态的微分同胚映射框架。在此框架基础上提出了两种映射算法:基于径向基函数的迭代算法,和基于可逆神经网络的映射算法。随后在仿真和实验中进行了对比和分析,验证该映射算法框架的有效性。

其次,现有的时不变动态系统仅考虑位置的规划,姿态的规划仍采用时域插值,无法实现全过程位置和姿态同步性,本文提出了融合位置和姿态的动态系统设计方法。该方法通过建立示教数据空间和隐空间之间的微分同胚映射,并在隐空间中构建一个全局渐近稳定的动态系统,从而得到示教空间中全局渐进稳定且能精确表征示教数据的动态系统,并设计了 Lyapunov 函数证明了其稳定性。采用二维空间仿真和三维空间中机器人实验进行验证,结果表明该方法构造的运动规划算法对时间和空间扰动的适应性好,在动态变化的环境中可以在线规划运动轨迹,证明了算法的有效性。

最后,考虑接触式的任务操作,需要对接触力进行规划并基于力和位置反馈进行柔顺控制器设计,本文提出了接触式任务操作的技能表达模型。通过采集人执行接触式任务操作过程中的数据,构建接触式任务中位置、速度、交互力和交互刚度之间的关系,将其用于指导机器人自主执行时的运动规划和柔顺控制器设计。并提出了基于模仿学习的接触式任务柔顺操作技能学习框架,针对典型的接触式任务设计了基于系统状态的自适应力/位混合控制器和阻抗控制器。为验证该框架,设计了三类机器人实验:机械臂打磨实验,轴孔装配实验,变刚度交互实验。实验中均能实现机械臂的拟人化柔顺操作,并成功的完成了任务,验证了该柔顺操作技能学习框架的有效性。

本文系统地研究了机械臂的运动规划与柔顺控制方法,并提出了基于模仿学习的规划和控制算法,可以快速适应不同的操作场景,并能适应动态变化的环境和未知干

扰，对机器人的运动规划和柔顺控制研究和快速应用具有一定的价值和意义。

关键词: 机械臂，模仿学习，微分同胚映射，动态系统，运动规划，柔顺控制

ABSTRACT

Robots have been widely used in human's daily life, which can liberate human from monotonous and repetitive work, and also improve production efficiency and quality. However, the capabilities for dexterous manipulation and intelligent decision-making are still far lower than humans. Robots can only be used for repetitive tasks in a structured environment. It is non-trivial to design a motion planning algorithm with high adaptability and stability for robots in a dynamic unstructured environment. For tasks in a complex contact-based environment, the planning and control of the force and position profiles still need to be designed carefully for specific tasks, and it is difficult to extend application to different scenarios. Therefore, this paper proposes a motion planning and compliant control method for robot manipulators based on imitation learning to improve the ease of use and intelligence of the robot, and to make robots accomplish tasks like a human in complex scenarios.

In order to improve the ease of use and adaptability of the robot to the environment, a promising way is to use a dynamical system for motion planning. However, traditional dynamical system methods often use time-based planning or are complicated to design. This paper proposes a dynamical system design method based on a diffeomorphism, which can realize single-step training. Based on the constraints of the dynamical system, the constraints of the diffeomorphic mapping are analyzed. Then a unified framework of the diffeomorphic mapping that couples the position and orientation is proposed. Based on this framework, two mapping algorithms are proposed: an iterative algorithm based on radial basis functions, and a mapping algorithm based on invertible neural networks. Subsequently, comparison and analysis were carried out in simulation and experiment to verify the effectiveness of the mapping algorithm framework.

Secondly, current time-invariant dynamical systems only consider the position planning, and the orientation planning still adopts time-domain interpolation, which cannot realize the synchronization between position and orientation planning. This paper proposes a dynamical system design method that integrates position and orientation. This method establishes a diffeomorphic mapping between the teaching space and the latent space, and constructs a globally asymptotically stable system in the latent space, thereby obtaining a globally asymptotically stable system that can accurately represent the demonstration data. The stability of the dynamical system is proved by Lyapunov function. It is verified by two-dimensional space simulation and real robot experiments in three-dimensional space. The results show that the online motion planning method has good adaptability to time and space disturbances and can

replan the motion on-the-fly in dynamical changing environment, which proves the effectiveness of the algorithm.

Finally, for contact-based tasks, it is necessary to plan force and position profiles and design a compliant controller based on force and position feedback. This paper proposes a skill representation model for contact-based task manipulation. By collecting demonstration data in the process of human performing contact-based tasks, the relationship between position, speed, interaction force and interaction stiffness in contact tasks is constructed, which is used for the planning and control of the robot manipulator during autonomous execution. Then a skill learning framework for compliant manipulation is proposed based on imitation learning for contact-based tasks, and an adaptive hybrid force/position controller and an adaptive impedance controller are designed based on system state for typical contact tasks. In order to verify the framework, three types of robot experiments are designed: robot polishing experiments, peg-in-hole assembly experiments, and variable stiffness interactive experiments. In these experiments, the robot manipulators can realize compliant manipulation like a human to accomplish contact-based tasks, which verifies the effectiveness of the skill learning framework for compliant manipulation.

This paper systematically studies the motion planning and compliant control methods of robot manipulators, and proposes a motion planning and control algorithm based on imitation learning, which can quickly adapt to different operating scenarios, and also adapt to dynamic environment and unknown disturbances. The researches in this paper are of great value and significance to the research on motion planning and compliant control, and also the application of robots.

Key words: robot manipulators, imitation learning, diffeomorphism, dynamical system, motion planning, compliant control

目 录

摘要	I
ABSTRACT	III
1 绪论	1
1.1 课题来源	1
1.2 研究背景与意义	1
1.3 国内外研究综述	2
1.4 本文研究内容与结构	16
2 基于模仿学习的机械臂操作基础	19
2.1 微分同胚映射	19
2.2 动态系统及其稳定性	20
2.3 模仿学习	22
2.4 机械臂运动规划与控制	23
2.5 本章小结	28
3 融合位置和姿态的微分同胚映射算法	29
3.1 引言	29
3.2 约束分析	30
3.3 映射算法框架	31
3.4 基于径向基函数的迭代算法	32
3.5 基于可逆神经网络的映射算法	35
3.6 仿真对比与分析	37
3.7 实验研究	39
3.8 本章小结	45
4 基于模仿学习和微分同胚映射的运动规划	46
4.1 引言	46
4.2 运动规划算法及其稳定性	46
4.3 融合位置和姿态的运动规划	48
4.4 运动规划仿真	53
4.5 动态环境运动规划实验	58
4.6 本章小结	64

5	基于模仿学习的柔顺操作技能学习	65
5.1	引言	65
5.2	技能表达方法	65
5.3	接触式任务柔顺操作技能学习	66
5.4	机打磨实验研究	70
5.5	轴孔装配实验研究	74
5.6	变刚度操作实验研究	80
5.7	本章小结	83
6	总结与展望	85
6.1	总结	85
6.2	展望	86
	参考文献	87
	攻读博士期间发表的科研成果	101
	致谢	102

1 绪论

1.1 课题来源

本学位论文来源于国家自然科学基金“基于模拟学习的软操作柔顺交互控制研究（项目编号：51705371）”，人因工程国防科技重点实验室开放基金资助项目“基于多传感融合的人机交互规划与柔性控制”（项目编号：6142222180311），国家留学基金委国家建设高水平大学公派研究生项目（资助编号：201906270139），欧洲 LEARN-REAL 项目（CHIST-ERA, <https://www.chistera.eu/projects/learn-real>），以及瑞士国家科学基金会。

1.2 研究背景与意义

智能机器人广泛应用于自动化产线、医疗、物流、服务等领域，将人们从单调重复性的工作中解放出来，大大提高了生产效率和质量。以汽车产线为例，其采用大量的工业机器人实现自动化焊接^[1]、喷涂^[2]和装配^[3]等操作。此类工业机器人一般采用定点示教编程，根据事先设定的作业任务进行重复运动，仅能工作于结构化环境，若需进行产线改造则需要专家对其进行重新编程调试。社会的高速发展使得此类机器人无法胜任多样化、个性化的生产任务。基于柔性制造^[4]的生产模式逐渐成为趋势，同一套生产体系可以生产多种产品，并能快速响应市场需求。因此，对机器人进行快速编程，提高其易用性和智能性是柔性智能制造中的关键。

随着协作机器人^[5,6]的兴起，机器人开始走进人类的生产生活，然而，现阶段机器人的灵巧操作水平和智能决策能力仍远低于人类。实际柔性制造产线中常采用人和机器人共享操作空间的方式协同作业，实现优势互补，提高生产效率；服务行业中开始逐渐采用机器人服务人类生活。但人类操作不确定性强且运动复杂多变，机器人与人共享工作空间给机器人操作提出了更多的需求：交互过程中首先需保证人类和自身不受到伤害，保证安全性；其次，人机协作^[7]的场景复杂多变，需要机器人能适应动态变化的场景；当机器人运动过程中受到外界的时间或空间扰动时，仍能稳定地完成任任务。传统基于离线运动编程的方式已经无法满足需求，迫切需要新的机器人运动规划方法，提高机器人对动态环境的适应性。

另一方面，机器人普遍采用高增益位置控制，无法直接应用于接触式任务的操作。针对接触式的操作任务，需考虑交互安全性和接触过程的约束关系，接触力的规划与柔顺交互控制对任务完成质量的影响很大。如打磨抛光^[8]，去毛刺，装配^[9]等任务，需要基于任务约束进行力和位置的规划和控制。若仅采用传统的运动规划和控制方法，由于建模误差和环境的不确定性极易导致过大的接触力，损伤机器人或工件。基于柔顺控制的方法可以让机器人具有柔顺性交互行为，避免上述缺陷。但目前此类机器人仍

严重依赖针对特定任务的手工编程，难以快速适应其他的接触式操作场景。

综上所述，目前的机器人运动规划和控制算法主要存在以下问题：1) 编程调试复杂，需要专家知识，易用性和可迁移性差；2) 针对动态非结构化环境，难以设计具有高适应性和稳定性的运动规划算法；3) 接触式任务操作中位置和力的规划与控制复杂，且难以推广应用于不同场景。为提高机器人的易用性和智能性，急需研究解决上述问题的方法。因此，本文提出基于模仿学习的机械臂运动规划与柔顺控制算法，研究机械臂的快速轨迹规划算法，并保证算法的高适应性和稳定性，研究接触式任务下的人类柔顺操作技能表达模型，并将其迁移至机械臂实现自主执行。本文提出的算法可提高机械臂的易用性和智能性，对机械臂技术的发展与应用具有重要价值与意义。

1.3 国内外研究综述

1.3.1 传统的运动规划方法

使用机器人中最基础的任务是规划其运动，使得机器人从起始位置无碰撞地运动至目标位置。机器人运动规划包含路径规划和轨迹规划两方面，路径规划指的是生成连接起点和终点无碰撞的几何路径。轨迹规划需要在路径的基础上赋予时间信息，生成运动过程的速度、加速度等信息。传统的运动规划方法可以分为以下五种：基于插值的规划，网格搜索法，基于采样的规划，人工势场法，以及启发式方法。

基于插值的规划方法可分为关节空间规划和操作空间（一般为笛卡尔空间）规划两大类。关节空间规划一般给定起点、终点和中间点的信息，在关节限制和轨迹连续性平滑性等约束条件下，基于梯形速度规划（带有抛物线过渡的线性插值）^[10]，S 形速度规划，高阶多项式插值^[11]等方法计算整个过程的关节位置、速度和加速度时间曲线。由于仅关注于关节空间插值，受正运动学非线性的影响，机械臂末端执行器的运动轨迹不易预测。操作空间规划已知末端执行器的起点和终点位姿（位置和姿态），以及中间关键点的位姿，基于圆弧或直线路径插值等方式，求解末端执行器的运动轨迹。此类方法需通过逆运动学和雅可比矩阵实时求解关节位置和速度，用于对末端运动轨迹有要求的场合。基于运动学的规划方法使用方便，计算简单，一般用于结构化且工作流程固定的工业场景。针对新的任务场景需要重新指定终点和中间点并生成轨迹，适应性差。操作场景中存在障碍物时难以直观地规划出无碰撞轨迹。

网格搜索法首先将空间离散成网格点，在每个网格点，可向相邻的网格点移动，同时保证路径是无碰撞的，不断重复搜索求解连接起点和终点的无碰撞最优路径。代表性的方法有 Dijkstra^[12] 和 A*^[13] 算法。A* 算法效率较高，但仅仅适用于静态场景。在其基础上，D*^[14] 和 AD*^[15] 算法将其扩展到动态场景的路径规划。然而，此类方法都存在计算量随维度增大而指数增大的问题。

基于采样的规划方法中，最常用的有快速搜索随机树 (Rapidly-exploring Random Trees, RRT)^[16] 和概率路径图 (Probabilistic Roadmap Method, PRM)^[17-19]。如图1.1(a)所示，RRT的基本思想是从起点（左下角）出发向各个方向随机生长得到子节点，并以子节点为起始点继续生长得到随机树，直至抵达目标点（右上角），从而构建起点到终点的路径（红色曲线）。PRM首先随机采样空间中的点（图1.1(b)蓝色圆点），并采用k近邻算法或基于距离阈值连接某些相邻点，并添加到路径图中，重复上述步骤直至路线图足够稠密为止，最后采用D*找出连接起点和终点的最短路径。PRM中采样点的数量增加，搜索到的路径会逐渐接近最短路径，但搜索效率也会下降。此类方法搜索空间大，计算量大，导致规划所需时间太长，难以进行实时规划。基于采样的规划方法在机器人研究中应用最为广泛，开放式运动规划库 (Open Motion Planning Library, OMPL)^[20] 将常用的基于采样的运动规划方法进行整合并集成到机器人操作系统 (Robot Operating System, ROS)^[21] 中，便于研究者进行算法研究和快速应用。然而，此类方法针对相同的起点和终点，每个规划的运动可能不一致。

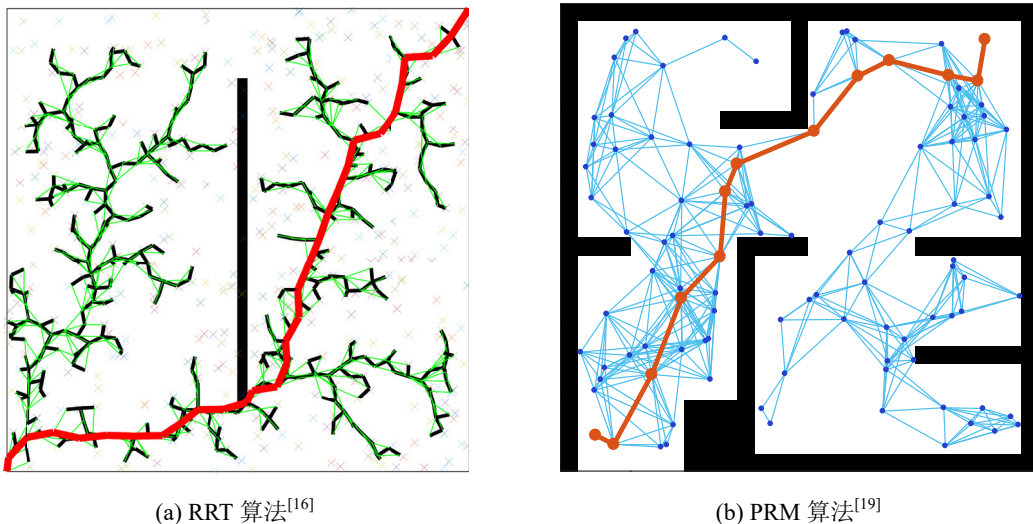


图 1.1 采用 RRT 和 PRM 的路径规划二维仿真

人工势场法^[22]将机器人视为势场中的一个点，其受到来自障碍物的排斥力和来自终点的吸引力，从而生成无碰撞的从起点到终点的路径。这种方法计算量小，添加新的障碍物十分容易。然而该方法可能会陷入势场的局部极小值中，而不是朝着全局最小值运动，易找到非最佳路径甚至无法找到路径，因此通常与其他逃离局部极小值的方法结合使用。

启发式方法包括神经网络法 (Neural Network, NN)，模糊逻辑算法 (Fuzzy Logic, FL)，受自然界启发的方法，以及多种算法的结合。NN^[23-25]因强大的非线性拟合和并行计算能力，广泛应用于机器人路径规划，但其训练耗时长，且有可能难以收敛到最优

解。FL 算法^[26, 27] 基于 if-then 逻辑模仿人脑的不确定性概念判断、推理思维方式构建机器人的运动路径, 但难以选取合适的规则和成员函数, 通常与基于 NN 的方法结合使用^[28, 29]。受自然界启发的算法有遗传算法 (Genetic Algorithm, GA)^[30-32], 粒子群优算法 (Particle Swarm Optimization, PSO)^[33-35], 蚁群算法 (Ant Colony Optimization, ACO)^[36-38]。相较于传统方法, 此类方法性能较好, 但一般需要较长时间求解轨迹, 常用于离线运动规划。

1.3.2 基于优化的运动规划方法

机器人的运动规划期望在运动过程中与环境无碰撞, 且运动平稳、快速和消耗能量小。传统的运动规划方法 (如基于采样的运动规划方法等), 获得的无碰撞路径往往是分段线性连接的, 难以满足机器人运动的约束条件 (如关节速度、关节力矩、关节奇异、末端执行器姿态等) 的路径, 无法直接用于机器人的运动。基于优化的运动规划方法 (Optimization-based Motion Planning, OMP) 在机器人运动规划中有两个方面应用: 1) 优化其他方法生成轨迹的长度和光滑性等, 以满足约束条件, 2) 直接计算满足各种约束条件的优化轨迹。

基于传统的运动规划方法获得的无碰撞路径进行后处理, 可以缩短路径和提高路径平滑性。Shortcut 算法^[39] 采用局部规划器对子路径进行优化以缩短路径长度, 作为后端优化的预处理部分。还可以利用该方法, 反复选取路径上的两点, 在满足碰撞约束前提下用一个较短的无碰撞段代替中间轨迹^[40]。对于无碰撞路径的平滑优化, 可以采用曲线对两段直线路径进行优化, 如 Boissonnat 等人^[41] 使用 line-circle-line 方法平滑移动机器人的路径, 能够提升移动机器人在路径跟踪时的灵活性, 但得到的曲率不连续变化的路径势必导致机器人的加速度发生突变。此外, 还有采用其他平滑曲线平滑的方法, 如螺旋线和贝塞尔曲线^[42], 三次样条曲线^[43], 极多项式曲线^[44], 五次非均匀 B 样条平滑轨迹^[45] 等。以弹性带为基础的基于轨迹优化的运动规划方法^[46], 是针对全局规划器生成的初始轨迹相对于最短路径进行变换, 借助内部收缩力和外部排斥力来改善路径形状, 同时避开障碍物。在此基础上发展而来的时间弹性带^[47, 48], 生成的局部轨迹由一系列带有时间信息的离散位姿组成, 通过算法优化使得离散位姿组成的轨迹达到时间最短、距离最短和避开障碍物等目标, 同时限制速度与加速度使轨迹满足机器人运动学约束。

此外, 还有基于多约束非线性优化的方法求解目标轨迹。先设计目标函数, 并利用梯度下降法、图优化等优化算法直接生成满足避障和其他约束条件的目标轨迹。通过设计连续的目标函数, 并求解受等式或不等式约束的非线性规划 (Nonlinear Program, NLP) 问题, 若其导数可直接求解或者近似, 可计算一阶导数和二阶导数并基于牛顿法求解不受约束的 NLPs。针对复杂约束条件下的 NLPs, 可采用增广拉格朗日函数 (Aug-

mented Lagrangian) 法, 内点法 (Interior Point) 或序列线性二次规划 (Sequential Linear Quadratic, SLQ) 法^[49]。例如基于差分动态规划^[50] 的迭代线性二次规划 (iterative Linear Quadratic Regulator, iLQR)^[51], 可用于生成能耗最优的运动轨迹。

求解 NLPs 主要的不同点在于如何使用目标函数的导数和环境信息。Ratliff 等人提出了 CHMOP 算法^[52](Covariant Hamiltonian Optimization for Motion Planning), 利用函数协变梯度下降技术, 对前进项和障碍项组成的目标函数优化, 以最小化代价函数为目标, 求取最优路径, 但易陷入局部最小。在此算法基础上发展而来的 T-CHOMP(Time-Covariant Hamiltonian Optimization for Motion Planning) 算法, 将梯度优化部分扩展到时间域, 可一定程度上避免陷入局部最小。Kalakrishnan 等人提出随机轨迹优化算法 (Stochastic Trajectory Optimization for Motion Planning, STOMP)^[53], 该算法采用类似的优化目标函数, 但没有采用梯度优化法, 而是对可能路径的概率做积分, 利用期望最大化 (Expectation Maximum, EM) 估计计算下一步代价更低的轨迹。其中, CHOMP 和 STOMP 对机器人本体进行球体近似, 因此需要对环境和机器人进行离线或在线预处理。为降低碰撞检测的计算消耗, Pavlichenko 等人^[54] 提出自适应碰撞检测度的方法, 然而, 目标函数中的避障项仍然容易导致陷入局部极小值。

为解决同时存在静止障碍物和运动障碍物场景下轨迹优化的问题, Park 等人提出增量式轨迹优化算法 (Incremental Trajectory Optimization for Motion Planning, ITOMP)^[55], 在短时间内计算障碍物位置和轨迹边界, 使用最近障碍物距离计算的约束或成本来及时避障。Schulman 等人提出了 TrajOpt 算法^[56], 利用顺序凸优化约束凸形和非凸形扫描体之间的最小符号距离来考虑连续时间安全性。Deits 和 Tedrake^[57] 没有直接使用障碍物形状, 而是将贪婪凸分割和混合整数优化应用于动态四旋翼路径规划, 确保规划的运动位于无障碍空间的凸子区域内。

实际使用中, 基于优化的方法通常采用零初始化^[58], 或者连接起点和终点的直线插值路径作为初值。然而, 这些初始化极有可能无法求解出最优轨迹, 或者需要很长时间的随机搜索才能求解出结果。因此, 通常会先采用其他方法 (如 RRT, PRM) 求出合适的初始轨迹进行初始化, 再基于优化的方法求解最优轨迹。基于优化的方法一般需设计出明确的目标函数, 并结合机器人约束和环境约束设置边界条件来求解受约束的非线性优化问题, 此类方法仍旧需要专家经验知识, 难以适用于动态的任务场景。

1.3.3 基于学习的运动规划方法

基于学习的机器人运动规划方法主要分为两大类: 强化学习^[59] 和模仿学习^[60-62]。强化学习通过智能体与环境进行试错学习来建立策略, 对环境适应性和泛化能力好, 但大多数算法训练效率低。模仿学习基于人的操作过程数据建立技能模型, 并将其迁移至机器人, 以模仿人的操作, 实现简单方便。

(1) 强化学习

基于强化学习 (Reinforcement Learning, RL) 的运动规划方法先定义一个奖励函数, 机器人与环境交互过程中根据环境反馈得到奖励, 不断调整运动策略, 以最大化奖励为目标, 得到最优的运动轨迹。

基于强化学习的运动规划可以提高机器人自主决策与学习能力, 对不确定环境有更好的适应性。Peters^[63] 提出将动作基元与随机策略梯度学习相结合, 并评估不同策略梯度算法在参数化运动基元上的适用性, 在机械臂打棒球任务中验证了算法的高效性。Theodorou^[64, 65] 等人提出一种基于具有路径积分的随机优化控制框架, 并设计概率强化学习策略 PI^2 , 该方法中用于学习的更新式无需矩阵的逆解和梯度学习率, 对于高维问题的数值鲁棒性好, 且性能优于基于梯度优化的强化学习方法, 并在 12 自由度的机械狗上进行了仿真和实验验证。Schenck 等人^[66] 采用优化后的卷积神经网络建立挖取和倾倒颗粒介质的预测模型, 实现了控制 KUKA 机器人挖取和堆放颗粒的操作技能学习。

随着视觉图像技术的发展, 基于视觉信息反馈和强化学习的机器人运动规划研究也越来越多。Asada 等人^[67] 基于视觉信息构件得到机器人工作环境中目标物体的几何尺寸及方位信息, 采用 Q-learning 算法完成了机器人将球击打到指定位置的操作任务。Kroemer 等人^[68] 的混合控制策略, 成功实现了机器人抓取不同种类物体的任务目标。

近年来, 众多学者们将庞大且深层的神经网络与强化学习结合, 即深度强化学习。许多深度强化学习算法被陆续提出, 如信赖域策略优化 (Trust Region Policy Optimization, TRPO)^[69], 深度确定性策略梯度 (Deep Deterministic Policy Gradient, DDPG)^[70], 归一化优势函数 (Normalized Advantage Functions, NAF)^[71] 等, 基于这些算法, 通过端到端的学习实现从原始输入到输出的直接映射, 可有效解决高维复杂环境下的决策问题, 实现较为复杂的运动轨迹规划, 且具有很强的泛化能力。在此基础上, Levine 等人^[72] 提出利用一个 7 层、92000 个参数的深度卷积神经网络进行策略学习, 该过程中不再需要全部的状态信息即可掌握未知的位姿信息, 并可使用随机梯度下降等常规方法进行训练。而 Gu 等人^[73] 指出, 由于高采样复杂度, 深度强化学习的应用仅限于仿真与简单的实际任务, 因此他们提出一个适用于复杂 3D 机器人任务的基于深度 Q 网络的离线策略训练方法。实验证明该方法可以在不需要先验示教数据或人为设计的情况下学习一系列复杂的 3D 机器人运动技能。李广创等人^[74] 则以点焊机器人为研究对象, 将机器人在三维空间的自动点焊转化为机器人的避障路径规划问题, 使用一个三层的 DNN 网络, 经过离线训练后, 自行训练出一条接近于最优的运动轨迹, 成功地避开障碍物到达目标点。

将物理仿真器中的基于深度神经网络训练机器人获取操作技能, 并将其迁移至实际机器人系统中, 被称为 Sim2Real^[75], 受到了学者们的广泛研究。其中, 以 DeepMind

公司和 OpenAI 研究所为代表的相关研究中, 为提高强化学习训练效率, 采用多个机器人在仿真器中^[76-78] 进行强化学习训练, 可以更快地求解出运动规划策略, 并解决高维度的规划和控制问题。直接让机器人在真实环境中进行强化学习探索十分危险, 且效率低下, 成本较高。如 DeepMind 公司^[79] 采用 6 至 14 台机器人在现实环境中收集了百万次抓取数据, 耗时两个月, 采用大型卷积神经网络实现了无手眼标定下的抓取技能学习, 并将其迁移至其他机器人实现抓取操作。Sim2Real 的方法通用性好, 泛化能力强, 然而需要大量的计算资源和长时间的网路训练, 且仿真环境和真实的机器人环境存在较大差异, 直接迁移策略效果一般很差, 目前难以大规模地应用于实际生产生活中。

此外, 还有在已有轨迹基础上采用强化学习进行优化的方法, 原理类似于基于优化的运动规划方法, 以提高学习效率。Kalakrishnan 等人^[80] 使用动觉示教初始化机器人的初始位置控制策略, 且利用 PI^2 算法通过优化目标函数来学习末端感知器所需要的力、转矩等信息, 从而在转动把手开门和拾取桌上的笔两个环境的实验中取得了良好的效果。Ota 等人^[81] 采用 RRT 算法生成的路径作为强化学习的初始轨迹, 实现快速收敛。Chiang 等人^[82] 提出 RL-RRT 算法将强化学习作为局部优化器, 实现运动规划中的避障。Faust^[83] 提出 PRM-RL 算法使用强化学习来处理小范围内点到点的运动规划。Ha 等人^[84] 针对多机械臂系统训练一种分散策略来控制机器人达到目标末端执行器姿态, 该系统以基于采样的运动规划算法作为初值, 提高强化学习的学习效率, 同时保证神经网络的快速推理。

基于深度强化学习的运动规划, 需要利用庞大且复杂的网络, 计算资源消耗大, 具有很强的不稳定性, 一般难以直接在现实环境中应用。

(2) 模仿学习

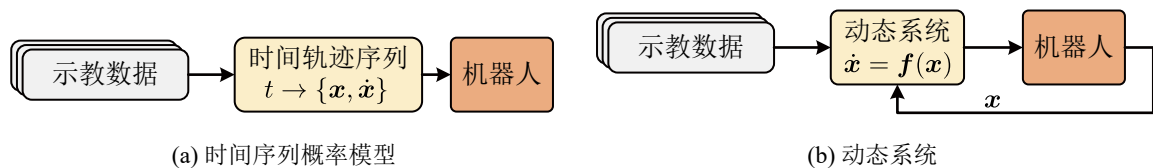


图 1.2 基于模仿学习的运动规划方法

模仿学习 (Imitation Learning, IL), 又被称为示教学习 (Learning from Demonstration, LfD)。基于模仿学习的机器人运动规划从人执行任务过程中的数据提取规律, 并构建运动规划模型, 主要可分为基于时间序列的概率模型, 和基于状态反馈的动态系统 (Dynamical System, DS) 模型两种方法 (图1.2)。

图1.2(a)中, 时间序列模型将人的示教时间轨迹序列看作随机过程, 基于最大似然估计来求解选定的统计模型参数值, 并用该模型来描述时间到轨迹的映射关系, 又被

称为行为克隆 (Behavior Clone, BC)。一般采用机器学习中的回归算法来建立模型, 如高斯混合回归 (Gaussian Mixture Regression, GMR)^[85], 高斯过程回归 (Gaussian Process Regression, GPR)^[86], 支持向量回归 (Support Vector Regression, SVR)^[87], 局部加权回归 (Locally Weighted Regression, LWR)^[61], 隐式马尔科夫模型 (Hidden Markov Model, HMM)^[88] 等。由于生成的轨迹精确依赖时间, 为开环的运动规划方法, 其对时空扰动十分敏感, 难以在线适应动态环境。

高斯混合模型 (Gaussian Mixture Model, GMM) 是一种有限混合模型, 在模仿学习中, 利用多个高斯过程函数来表征多次示教采集的数据, 每个高斯核函数表示一个阶段, 并使用高斯分布的协方差来描述不同阶段示教轨迹的不确定性。图1.3展示了采用 GMM 学习多次手写字母 G 的轨迹数据 (图1.3(a)), 对 5 个高斯模型 (图1.3中的椭圆) 进行线性混合, 并求解其 GMR 的轨迹 (图1.3(b)), 从而建立时间到二维坐标的映射模型 (图1.3(c)和图1.3(d))。

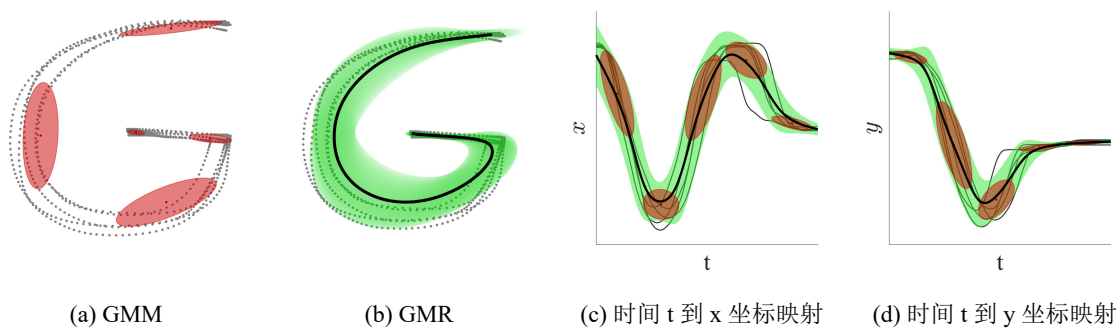


图 1.3 GMM 和 GMR 编码二维示教数据^[89]

Calinon 等人^[90] 最早在仿人机器人上采用 GMM 实现了对基本的抓取动作的学习, 随后采用 GMM 的协方差对运动过程不同阶段任务空间和关节空间轨迹跟踪精度进行建模^[91], 应用到不同连杆长度的两自由度机械臂模型中。此外, Calinon 还提出了一种任务参数化的混合模型 (Task-Parameterized Gaussian Mixture Model, TP-GMM)^[92, 93], 采用相对轨迹来学习两次 GMM 模型, 并求 GMM 乘积作为轨迹模型, 以提高其泛化能力。Osa 等人^[94] 使用 GMM 对机械臂末端轨迹进行建模, 并使用学习到的协方差构建有障碍物的场景下满足不同跟踪精度要求的目标函数。GMR 需要存储所有训练数据用于回归计算, 计算复杂度和存储空间均较大。Schneider 等人^[95] 通过聚类将输入空间划分为小的子空间, 只利用测试输入附近子空间的训练样本进行预测, 减少了计算量, 提高计算实时性。然而, 采用 GMM/GMR、GPR 等方法生成的轨迹无法保证轨迹稳定收敛至目标点^[96, 97]。

HMM 对笛卡尔坐标系下的空间轨迹进行编码, 并采集训练集中具有最大相似性的轨迹复现技能。Akgun 等人^[98] 利用 HMM 对动作数据和环境感知数据进行建模。Cali-

non 等人^[99] 采用一条带有明确时间周期的 HMM 模型来描述运动中的时空约束, 在任务执行阶段显示出了一定的抗扰动性。Kulic 等人^[100, 101] 利用动作捕捉系统采集人体的运动数据, 并采用 HMM 进行编码, 基于模型的相似性用分层树的结构进行表示, 通过构造原始运动图来学习各基本动作之间的时间关系, 利用基本运动图来组织相关运动。Kruger 等^[102] 从连续观察中学习基本动作, 将每个基本动作及其对周边环境中表示相应的参数化的 HMM。为解决基于 HMM 模型生成的路径不连续的问题, Lee 等人^[103, 104] 利用 HMM 对一系列运动轨迹进行编码, 并对其中不连续的部分进行插值处理。Aletti 等人^[105] 在获取轨迹关键点后, 采用非均匀有理 B 样条曲线在关键点插值生成新的轨迹。

机器人在执行轨迹过程中可能受到外界的干扰, 且其起始点可能并不在示教轨迹的区域内。在这种情况下, 采用时间序列概率模型的方法难以抵抗干扰, 且无法将轨迹较好地泛化至整个工作空间, 同时保证其稳定收敛至目标点。如图1.2(b), 基于动态系统的机器人轨迹学习方法存在内在的抗扰动能力和稳定性, 其将人的示教轨迹描述为一类微分方程:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad (1.1)$$

其中 $\mathbf{x}$ 为系统的状态 (一般为末端位姿), 基于示教数据求解该方程参数。对任意空间中起始点, 对式(1.1)积分可得到包含人示教特征的运动轨迹。在机器人运动过程中, 动态系统基于当前的机器人位姿反馈, 实时计算出期望的位置 $\mathbf{x}_d$ 速度 $\dot{\mathbf{x}}_d$, 将其作为参考轨迹发送给机器人。由于存在状态反馈环, 基于动态系统的运动规划方法具有内在的对未知扰动和对动态环境的在线适应能力。

为构建动态系统, Ijspeert 等人提出动态运动基元 (Dynamical Movement Primitives, DMP)^[106] 的方法, 由基于单次示教数据的非线性前馈项和弹簧阻尼的线性反馈项两部分叠加, 可实现全局稳定地收敛至轨迹终点。DMP 已成为基于动态系统的机器人运动规划方法中最常用的一种, 并有多种基于 DMP 的改进方法。为编码多次示教轨迹以提高系统的泛化能力, Paraschos 等人基于 DMP 提出采用轨迹概率表达的概率运动基元 (Probabilistic Movement Primitives, ProMP)^[107]。Ude^[108] 将 DMP 扩展到包含姿态的示教轨迹学习和建模。核运动基元 (Kernelized Movement Primitive, KMP)^[109] 提高了对高维示教轨迹的适应性和泛化能力, 以弥补 DMP 和 ProMP 的不足。此类系统通常用于时间轨迹序列的规划, 不关心存在扰动的情况下的规划。针对超出示教区域的起点, 泛化能力仍然不足。

为提高动态系统对时空扰动的快速适应性, 需要构建时间无关的动态系统模型, 即式(1.1)中, $\mathbf{f}(\mathbf{x})$ 仅为状态量的函数。其中, 最主要的难点是如何建立全局渐进稳定的动态系统。Hersch^[110] 提出了可精确表征示教数据的动态系统, 但仅能保证在示教区域的稳定性, 无法保证全局收敛至目标点。为提高系统对时空扰动的适应性, 同时保证全

局稳定性,瑞士洛桑联邦理工的 Khansari^[96] 等人提出了动态系统的稳定估计器 (Stable Estimator of Dynamical Systems, SEDS), 基于 GMM^[60] 将示教数据编码为一阶自治系统 (定常微分方程), 并推导了保证该系统全局渐近稳定的充分条件, 采用二次规划求解稳定性约束下的最大似然问题。然而, SEDS 采用二次型的 Lyapunov 函数, 该模型生成的轨迹始终是收缩型 (轨迹到终点距离单调递减), 如图 1.4(a) 所示, 每条轨迹线 (带箭头的蓝色曲线) 逐渐收缩至终点, 若存在时或空间的扰动, 轨迹线也能收敛至终点。由于 SEDS 的二次型 Lyapunov 函数过于严格, 无法精确表征非收缩型的示教轨迹, 即对复杂的示教轨迹学习效果较差。SEDS 以牺牲部分表征精度为代价, 保证了全局渐进稳定性, 被称为动态系统学习中精度和稳定性的两难困境^[96]。

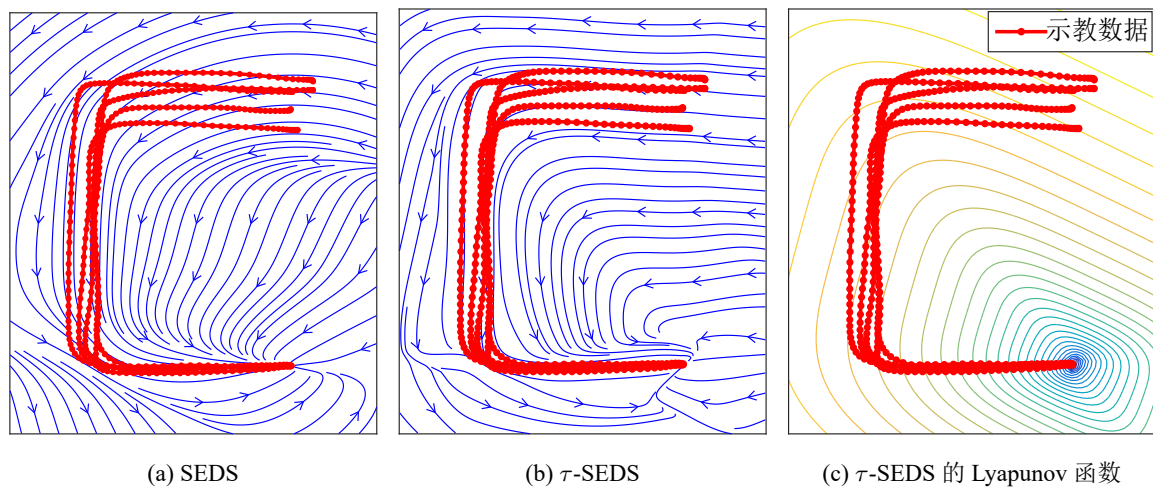


图 1.4 SEDS^[96] 和 τ -SEDS^[111] 编码二维示教数据

为改进 SEDS, 众多学者开始为构建全局稳定且可表征复杂轨迹的 DS 展开研究, 一般可分为以下两步: 1) 基于示教数据灵活构建 Lyapunov 函数; 2) 基于 Lyapunov 稳定性约束求解 DS。Khansari 为改进 SEDS 分别执行步骤 1) 和 2), 在 2) 中采用常见的回归算法将示教轨迹建模为 DS, 同时保证其 Lyapunov 稳定性, 该方法被命名为 SEDS-II^[112]。Neumann 等人指出 SEDS 因二次型 Lyapunov 函数导致存在稳定性和精度的权衡, 提出了 τ -SEDS^[111], 该方法在 1) 和 2) 之间构建一个微分同胚映射 (Diffeomorphic Mapping, DM)¹ 将示教数据映射到隐空间, 在隐空间中采用 SEDS 学习全局稳定的 DS, 再用其逆映射生成原空间的 DS。图 1.4(b) 为 τ -SEDS 在“C”型二维示教数据下的仿真效果, 蓝色带箭头的曲线为空间中的轨迹线, 对比 SEDS 的轨迹线 (图 1.4(a)), 可以看出, 其对示教数据的表征效果更好。图 1.4(c) 为 τ -SEDS 的 Lyapunov 函数, 相较于二次型的 Lyapunov 函数 (以终点为圆心的同心圆), 可表征更复杂的运动。浙江工业大学的金聪聪^[113] 等人使用狄利克雷过程高斯混合模型代替 SEDS 中的 GMM, 采用变分

¹微分同胚映射是一个可逆、连续且连续可微的映射, 其逆映射也是连续且连续可微的。

推断训练模型，可以基于示教数据自动确定合适的 GMM 混合分量个数，并设计参数化的 Lyapunov 函数，以提高学习轨迹精度，同时保证稳定性。文献^[114]提出了一种基于能量稳定器的动态系统 (Energy-based Stabilizer of Dynamical Systems, ESDS)，在二次型 Lyapunov 函数的基础上叠加虚拟能量，可实现快速高精度的非自洽 DS（时变系统）生成。但其仍精确依赖时间，难以应对存在扰动的环境。

采用两步学习的方法通常耗时较长，且需调整的参数较多。为解决该问题，Perrin 等人在文献^[115]指出上述 SEDS 及其变种方法只能建立与 $\dot{\mathbf{x}} = -\mathbf{x}$ 微分同胚的 DS，因此可学习的轨迹仍受限。Perrin 提出了一种基于迭代法求解 DM，并建立示教数据和某个已知全局稳定系统 $\dot{\mathbf{x}} = -\gamma(\mathbf{x})\mathbf{A}\mathbf{x}$ 之间的 DM，由此得到表征示教数据的 DS。该方法训练速度极快，但初始 DS 的选择尤为关键，直接决定了生成 DS 的质量。且只能处理单条轨迹，或是基于多条轨迹的平均。Urain 提出一种基于标准化流生成稳定的 DS：模仿流 (imitation Flow, iFlow)^[116]，提出了统一的框架同时训练 DM 和 DS 的参数，可学习点到点运动和极限环稳定系统，提高了模型表达精度，但基于神经网络的方法模型复杂度高，可解释性差，训练耗时较长。

表 1.1 常用的动态系统方法对比

	时不变系统	位置	姿态	单步学习	训练耗时
ProMP ^[107] , KMP ^[109]	-	✓	-	✓	短
DMP ^[108]	-	✓	✓	✓	短
SEDS ^[96]	✓	✓	-	-	长
SEDS-II ^[112]	✓	✓	-	-	中等
τ -SEDS ^[111] , i-SEDS ^[113]	✓	✓	-	-	长
FDM ^[115]	✓	✓	-	✓	短
ESDS ^[114]	-	✓	-	✓	短
iFlow ^[116]	✓	✓	-	✓	长

表1.1总结了上述基于动态系统的运动规划方法的主要特征。基于时间序列的概率模型采用开环的方式进行轨迹规划，因此难以在线适应动态环境或是抵抗时空扰动。SEDS 及其改进方法训练耗时较长，需对 Lyapunov 函数和 DS 分别学习，需调整的参数较多，实现较为复杂。基于神经网络的方法^[116, 117]一般训练耗时长，且调参困难，模型可解释性差。其次，真实的机器人操作空间是包含三维位置和三维姿态的 6 维空间，目前大部分研究仅针对位置的运动规划，考虑姿态的动态系统相关研究较少，通常采用恒定姿态进行仿真或实验，或是设计姿态的时间插值算法，因此生成的轨迹仍为时

间序列。为保证在位置和姿态的扰动下仍能在线生成收敛至终点的轨迹，且适应动态变化的终点位姿，需设计融合位置和姿态的时不变系统来进行运动规划。

1.3.4 柔顺控制

机器人的柔顺控制主要是指控制机器人与环境或人之间的交互力，其目的是防止交互力过大，损坏机器人本体或接触对象，一般需基于交互力反馈不断调整机器人与环境的相对位姿进行任务操作。柔顺控制算法一般可分为被动柔顺控制和主动柔顺控制两种。被动柔顺控制基于关节、连杆或末端执行器本身的固有的材料柔性特性。被动柔顺装置一般结构简单，价格低廉，对机器人精度要求较低，比主动柔顺控制更容易实现，并且具有更高的性价比。如远端中心弹性 (Remote Center of Compliant, RCC) 装置^[118, 119]，如图1.5所示，即微偏心柔顺装置。可将 RCC 装置安装于机械臂末端用于轴孔装配操作^[120]，一定程度上补偿装配中的位置和姿态误差。但是 RCC 装置的柔顺中心一旦确定，就很难调整，并且装置的刚度也是固定不可调的，不能适应不同刚度的作业需求，通用性较差。此外，有直接基于柔性关节实现人机交互中的被动柔性，如基于串联弹性驱动器 (Series Elastic Actuator, SEA) 关节的 Baxter 机械臂^[121]。

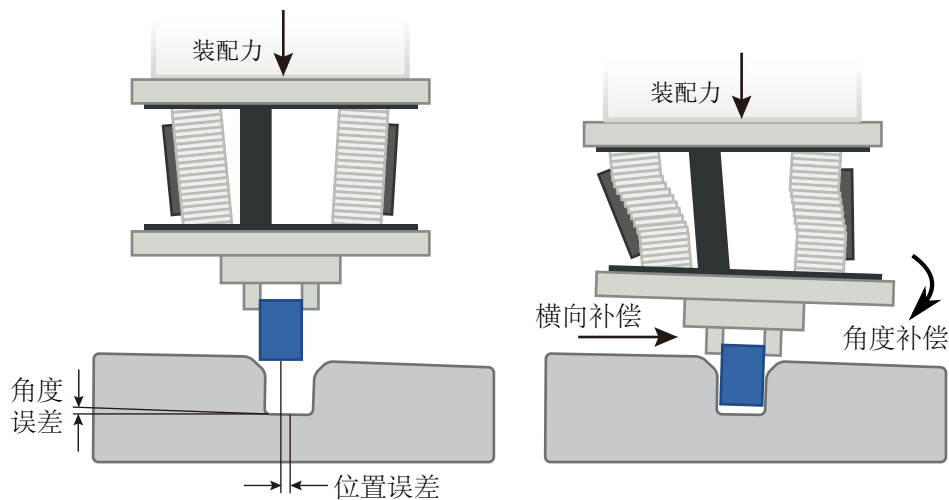


图 1.5 基于 RCC 装置的轴孔装配

主动柔顺控制需要根据实时检测的接触力调整机器人的运动，常用柔顺控制方法有力/位混合控制和阻抗控制等。1981 年 Raibert 提出了力/位混合控制^[122]，主要通过柔顺选择矩阵把任务空间分为力和位置的子空间，然后分别根据对应子空间参考信号的误差设计控制器，实现了力和位置的同时跟踪控制。Cheah 等人^[123]针对机器人本身运动学和动力学的不确定性提出了一种反馈控制方法，保证了机器人末端运动的稳定性。Mazur 等人^[124]提出了一种针对末端受约束的力/位混合控制方法，使用改进的 Arimoto 算法，在关节空间对力和位置进行解耦，并证明了力和位置的跟踪是收敛的。

常健^[125]提出了一种 7 自由度机械臂的力/位混合控制方法,包括位置环(位置、速度、加速度)和力环,针对环境约束,提出冗余机械臂适从坐标系建立方法,能够同时控制机器人末端的位置和与环境的接触力。喻洋^[126]提出了一种基于速度的力/位混合控制的装配方法,借鉴双曲正切函数设计了机械臂末端速度-力曲线函数,保证了柔顺装配的平稳性。力/位混合控制引入了柔顺选择矩阵,可以在相互正交的力子空间和位置子空间进行独立控制,但是实际操作环境复杂,难以确定对应的柔顺选择矩阵。此外,还需要在自由空间和接触空间进行控制器的切换,难以保证控制的稳定性。

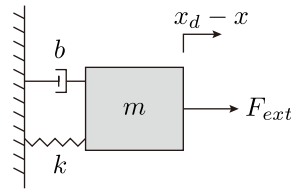


图 1.6 弹簧质量阻尼系统

1985 年 Hogan^[127-129]提出了阻抗控制的概念,通过调节机器人末端的位置和力之间的动态特性来实现柔顺性,把机器人与环境的交互建模为弹簧质量阻尼系统(图1.6),其表达式为:

$$m(\ddot{x}_d - \ddot{x}) + b(\dot{x}_d - \dot{x}) + k(x_d - x) = F_{ext}, \quad (1.2)$$

其中, x_d 为期望位置, x 为机器人当前位置, k, m, b 分别为刚度、质量和阻尼系数, F_{ext} 为外力。阻抗控制 (Impedance control) 通过测量目标位置和当前位置的误差,通过调节各关节力矩来调整末端输出力,以实现式(1.2)的二阶闭环系统。导纳控制 (Admittance Control) 通过感知外部的接触力 F_{ext} , 基于式(1.2)调整机器人运动位置和速度,通过修正参考轨迹实现与环境的柔顺接触,其闭环系统也为式(1.2)。

众多学者在 Hogan 研究的基础上进行改进,并应用于各种机器人场景。Bussy 等人^[130]使用导纳控制和有限状态机实现人机协作中人和机器人主被动角色的切换,以实现基于位置控制的机器人与环境之间的柔顺交互。Agravante^[131, 132]结合视觉和阻抗控制实现被搬运的物体水平抬升。Oh 等人^[133]针对人机交互过程安全性和轨迹跟踪的精确性,提出了一种频率型阻抗控制方法 (Frequency-Shaped Impedance Control, FSIC),使得人机交互力占主导时柔顺,摩擦力占主导时跟踪精度高。Kouris 等人^[134]针对人机协作中对安全性和交互的流畅性,提出了一种区分意外碰撞和预期协作交互接触的方法,对交互力进行基于快速傅里叶变换发现意外碰撞包含的频率成分更高,以此结合导纳控制达到了较好的交互效果。2020 年 Hamedani 等人^[135]针对于未知多变环境的交互,提出一种结合模糊动态增益的变阻抗控制策略,保证了在较大干扰下控制系统的稳定性。

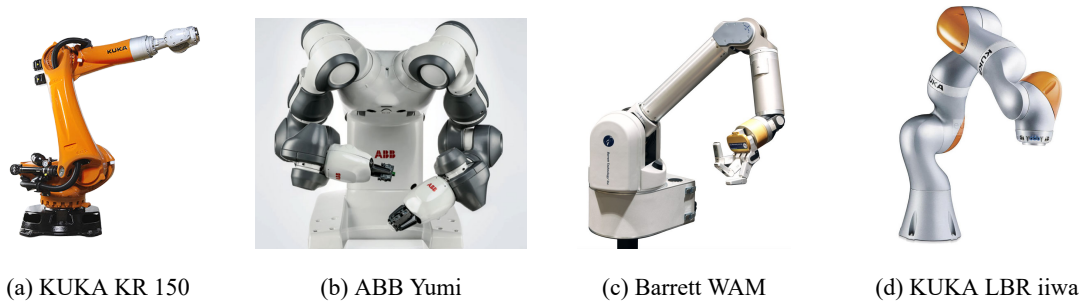


图 1.7 常见的机械臂

为实现主动柔顺控制，一般先需要实现精确的接触力感知，并根据力反馈进行控制器设计。由于传统的工业机器人（图1.7(a)）一般采用高增益位置控制，并采用大减速比齿轮箱，以实现大负载下高精度的轨迹跟踪。机器人自身刚度较大，微小的位置偏差将产生极大的接触力。此类机器人一般在其末端安装力传感器以实现力的精确感知。ABB 公司的 Yumi 双臂机器人（图1.7(b)）采用低负载 (0.5 kg) 设计，从而可使用低减速比齿轮箱，降低摩擦力影响，通过建立各关节摩擦模型并基于关节电流来估计可实现精确的交互力感知^[136]。Barrett WAM 机械臂（图1.7(c)）采用直驱电机的方式以大大降低摩擦力对力估计的干扰。KUKA LBR iiwa 机械臂（图1.7(d)）的 7 个关节均安装有力矩传感器，可实现连杆端力矩的精确测量^[137]，并结合其动力学模型实现外部接触力感知。

传统的柔顺操作任务中，需规划参考信号（位置、速度/力等），设计控制器并调整控制器参数。其规划和控制器参数固定，一般难以适应不同的机器人应用场景。基于模仿学习的方法从人的柔顺操作数据中进行学习，通常用于指导控制器的参数设计或参考位置/力的生成。Medina 等人^[138]使用导纳控制获取协同搬运的位置和力信息并分割为基元，最后使用 HMM 模型编码，实现简单预测交互力，但在模型学习阶段中分割算法需要根据信号维度手动调整。Nemec 等人^[139]使用 DMP 学习人机协作搬运的轨迹和速度，机器人的控制则采用阻抗控制以保证交互的安全性，并能泛化轨迹到不同起点和终点，但有较多参数需要仔细调节（包括速度阈值、权重因子等）。Rozo 等人^[140-143]提出了基于 TP-GMM 模型的人机协作搬运变刚度控制策略，能够根据新任务生成对应的轨迹和刚度，但机械臂可能运动到期望区域外，且难以保证终点位置精度。为提取人在任务执行过程中的交互刚度信息，彭涛^[144]提出了一种基于表面肌电信号的变阻抗控制策略，建立了手臂肌电信号到人手末端的刚度映射，并把人手的刚度纳入机器人阻抗控制器。武文^[145]对冗余机械臂操作空间和零空间运动引起的驱动力进行解耦控制，通过构造机械臂和环境的阻抗关系，并规划零空间速率，实现了力/位混合的柔顺控制。段江晔^[146]设计了神经网络对示教数据进行编码，同时学习轨迹和柔顺行为，并且实现了在线学习，提高了学习的效率。

然而，此类接触式操作任务中，需要针对每种任务设定期望的学习模型，仍缺乏统一的学习指导框架。

表 1.2 柔顺控制方法对比

方法	适用场景	优点	缺点
被动柔顺	固定作业任务	简单，成本低， 响应最快	对环境适应性差， 只能应对小的位姿偏差
阻抗控制	力矩控制的机器人	响应速度较快	需精确机器人动力学模型
导纳控制	基于位置/速度 控制的机器人	无需力矩控制接口， 适用范围广	须安装外置力/力矩传感器， 响应速度比阻抗控制慢
力/位置 混合控制	结构化场景	同时跟踪轨迹和力	需力/力矩传感器 需精确定义环境约束

表1.2总结了现有的柔顺控制方法适用场景及其优缺点。被动柔顺控制常用于固定场景的作业任务，是一种简单可靠的方法，其响应速度最快。主动柔顺控制方法中通过主动调节机器人的运动或控制力矩以实现与环境/人的安全交互，而且可以调整控制算法的参数实现以适应不同作业环境。实际应用中通常将被动柔顺和主动柔顺相结合，实现对过大交互力的快速响应和对变化环境的适应性。基于模仿学习的柔顺控制方法采用模仿人在完成交互任务中的调整策略来适应不同的环境，无需专家经验知识和复杂的控制器调参过程，应用前景较好。

1.3.5 研究现状小结

由以上国内外研究现状可以看出，在机械臂的运动规划方面：1) 大多数传统规划方法可以实现固定场景下的离线运动规划，一般需较长的时间进行离线求解；基于优化的方法效果较好，但仍旧需要专家经验知识，难以适应经常变化的任务场景；2) 基于强化学习的方法可以提高机器人对环境适应性，但其可解释性较差，网络结构复杂，训练耗时长，且极容易不稳定，难以直接用于实际机器人的规划；3) 基于模仿学习的方法可以提高机器人的易用性，实现简单，其中基于时间概率模型的方法可以很好地模仿人类的运动轨迹，但由于其严格地依赖于时间，难以适应动态环境，难以抵抗时空扰动，基于 DS 的运动规划方法可以提高其在线适应能力，但现有的时不变 DS 方法仅针对于位置信息，没有考虑姿态的规划问题。在柔顺控制方面：1) 基于被动柔顺控制的方法应用广泛，近年来主要集中在柔顺结构的设计研究，对动态环境适应性差；

2) 基于主动柔顺控制方法可以适应不同的场景,并能调节期望的交互特性,但调参较为复杂;3) 基于模仿学习的柔顺控制方法沿用主动柔顺控制的思路,其控制参数由人的示教数据中提取,场景适应性好。

因此,机械臂的运动规划中,由于应用场景复杂多变,且时常会受到人或动态环境的时空扰动影响,如何设计使用简便,且能适应动态变化的环境,同时具有全局稳定性的规划方法是一个极大的挑战。另一方面,针对接触式任务操作的柔顺控制,如何适应变化的接触场景,并设计主动柔顺控制算法也是一个亟需解决的问题。

1.4 本文研究内容与结构

为提高机器人的易用性和智能性,本文基于模仿学习构建了机械臂运动规划与柔顺控制算法。研究并提出了具有高度适应性和稳定性的运动规划算法,且可基于人工示教数据来快速构建,在仿真和实验中验证了所提方法的有效性。其次,针对接触式的操作任务,为保证接触的安全性和任务执行效果,提出基于模仿学习的接触式任务柔顺操作学习框架,并进行了机器人的实验验证。本论文的主要研究思路以及个章节的关系如图1.8所示,首先采集人完成任务过程的示教数据,一方面学习出运动规划模型,另一方面进行柔顺操作任务的技能表达模型的学习,将其用于指导力的规划和控制器的设计,最后进行实验验证。

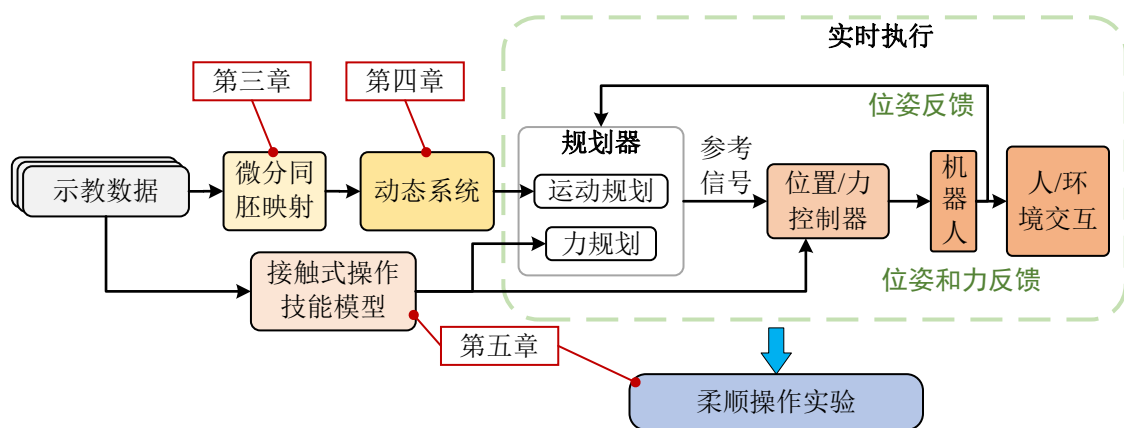


图 1.8 全文研究思路

全文共分为 6 章,各章节主要研究内容如图1.9。

第1章阐述了机器人运动规划与柔顺控制技术在实际应用中的研究背景与意义,并从运动规划方法和柔顺控制方法两大类综述了国内外的研究现状,并给出了本文的研究内容和结构。

第2章中概述了本文中常用基础理论和方法,包括微分同胚映射算法的基本概率

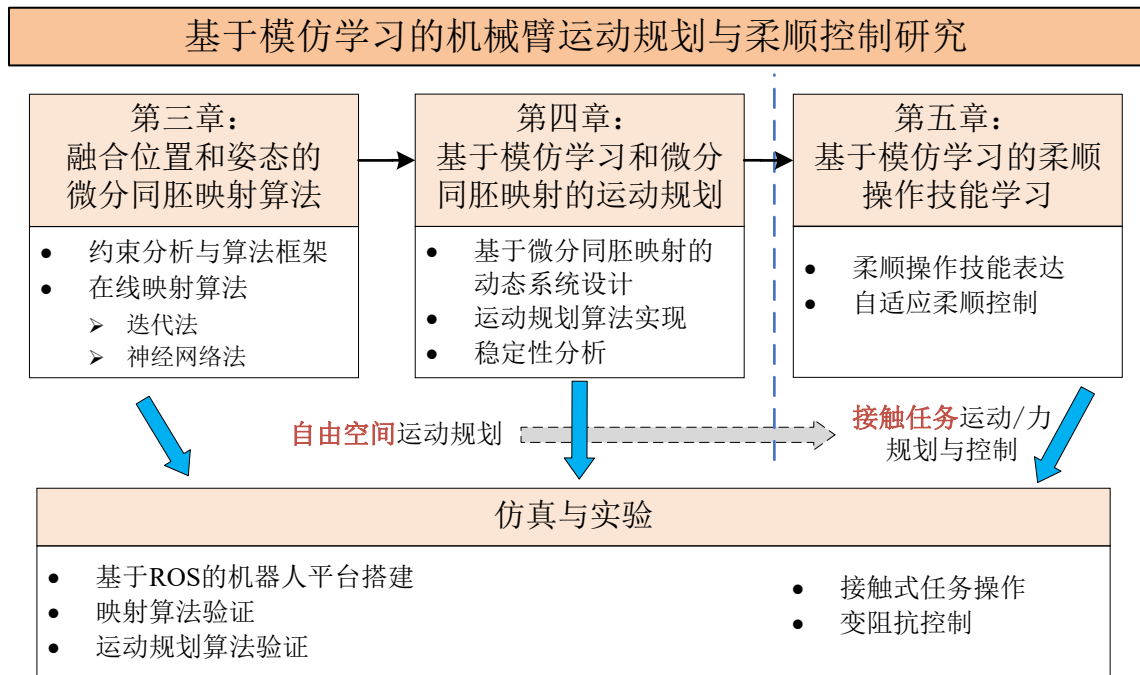


图 1.9 全文研究内容

和性质，基于 DS 的运动规划方法及其稳定性，GMM，以及机器人学的相关知识。

第3章针对目前基于模仿学习的规划方法存在的不足，设计了融合位置和姿态的微分同胚映射算法框架，实现了复杂示教轨迹可逆且连续可微的位姿变换，用于后续基于模仿学习的运动规划算法设计。在该映射算法框架的基础上提出了两种实现形式，基于径向基函数的迭代算法，和基于可逆神经网络的映射算法。并进行了二维空间的仿真和三维空间中机器人的遥操作实验，验证了这两种微分同胚映射算法的有效性。

第4章提出了基于模仿学习和微分同胚映射的运动规划算法。首先提出了基于微分同胚映射的 DS 构造思路，然后提出了仅含位置的 DS 构造算法，并采用类似的原理提出了融合位置和姿态的 DS 模型构造算法，并证明其全局稳定性。结合第3章的微分同胚映射算法，进行了二维空间的运动规划仿真，验证其稳定性和有效性。最后设计了动态场景下的交互实验，包含物体递交和动态抓取两组实验，验证该运动规划算法对动态环境的适应性和有效性。

第5章针对接触式的任务操作，提出了基于模仿学习的柔顺操作模型和技能表达方法，并进行实验验证。考虑人执行接触式任务中位置、速度、交互力和交互刚度之间的关系，提出了基于模仿学习的接触式任务操作技能学习框架，并研究了典型的接触式任务中控制器设计方法，包含与系统状态相关的自适应力/位混合控制器和自适应阻抗控制器。最后，设计了三个实验场景，用于验证第4章中的运动规划方法和前述柔顺操作技能模型。1) 设计机械臂打磨平台，并研究基于模仿学习的自主打磨操作，从

人工示教数据中学习技能模型进行力的规划和控制。2) 设计轴孔装配实验平台和示教数据采集平台, 从人工示教数据中分析并学习操作规律, 构建姿态调节控制器并用于机器人的自主执行。3) 设计变刚度柔顺擦白板实验, 研究人完成该任务时位置与手臂刚度之间的映射关系, 并将其用于指导阻抗控制器设计, 以实现变刚度柔顺控制, 提高作业效率。

第6章总结了全文的工作, 并对今后的研究方向进行了展望。

2 基于模仿学习的机械臂操作基础

本文的研究基于现有的基础理论和方法，如微分同胚映射的基本概念和性质；模仿学习中的动态系统基础理论及其 Lyapunov 稳定性分析方法；模仿学习的基本思路，以及其中常用的非线性回归算法；机械臂运动学、动力学以及常用的控制策略。本章简要概述这些基础理论和方法，后续章节将以此为基础做进一步的研究。

2.1 微分同胚映射

本小结主要介绍微分同胚映射的基本概念和性质，并简述如何通过微分同胚映射进行可逆的轨迹变换。

定义 2.1.1 (微分同胚) 对于给定的两个流形 M 和 N ，若存在一个可逆且连续可微的函数 $f: M \rightarrow N$ ，其逆映射 $f^{-1}: N \rightarrow M$ 也是连续可微的，则 f 为微分同胚映射。若 f 为 r 阶连续可微，则称 f 为 C^r 光滑的微分同胚映射。

定理 2.1.1 (可逆函数定理^[147]) 对于连续可微的函数 $g(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$ ，若其雅可比矩阵 $J_g \in \mathbb{R}^{n \times n}$ 可逆，则 $g(x)$ 可逆。

定理 2.1.2 (矩阵可逆性) 矩阵 $A \in \mathbb{R}^{n \times n}$ 可逆的充要条件是其行列式 $\det(A) \neq 0$ 。

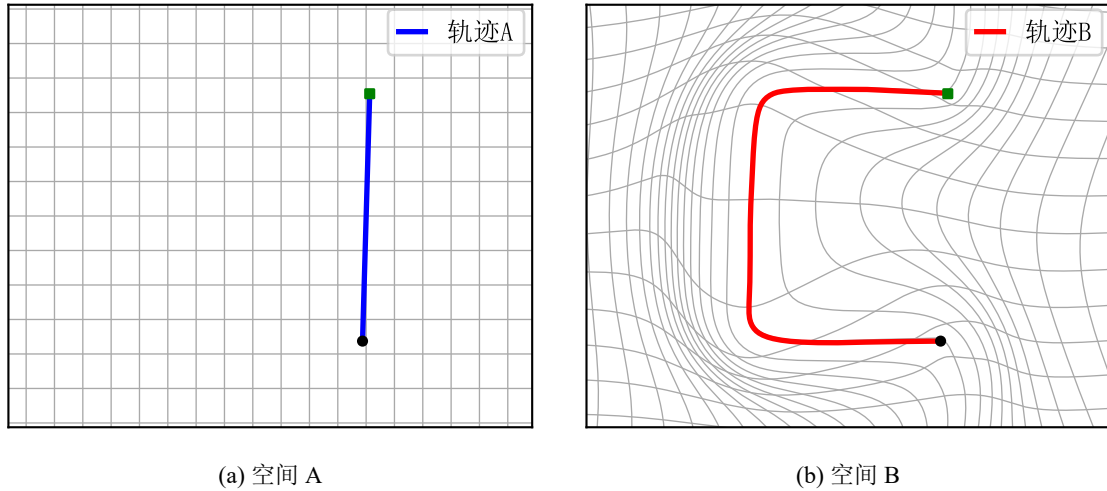


图 2.1 二维微分同胚映射示例

为便于可视化，考虑二维空间 A（图2.1(a)）和空间 B（图2.1(b)），设其中分别有时间轨迹 A 为 $x_a(t)$ 和轨迹 B 为 $x_b(t)$ 。若空间 A 和空间 B 存在微分同胚映射 $f(x)$ ，使得在任意时刻 t ，有：

$$x_b(t) = f(x_a(t)), \quad (2.1)$$

即每个时刻轨迹 B 与轨迹 A 一一对应，则有 $f^{-1}(x_b(t)) = x_a(t)$ 。图2.1展示了空间 A 中等间隔采样的网格线在微分同胚映射函数 $f(x)$ 的作用下被平滑地扭曲变形，形成

空间 B 所示的网格线。

对式(2.1)求时间的导数, 可得轨迹 B 和轨迹 A 的速度关系式:

$$\dot{\mathbf{x}}_b = \mathbf{J}_f(\mathbf{x}_a)\dot{\mathbf{x}}_a. \quad (2.2)$$

因此, 可以通过微分同胚映射对轨迹进行可逆的变换。若轨迹为 C^r 光滑的, 且 $\mathbf{f}(\mathbf{x})$ 为 C^r 光滑的微分同胚映射, 则映射后的轨迹 $\mathbf{f}(\mathbf{x}_a)$ 也是 C^r 光滑的。

2.2 动态系统及其稳定性

考虑状态变量 $\boldsymbol{\xi} \in \mathbb{R}^D$ (关节角度或末端位姿), 一个连续非线性的 DS 将运动轨迹描述为下述非线性微分方程:

$$\dot{\boldsymbol{\xi}} = \mathbf{f}(\boldsymbol{\xi}). \quad (2.3)$$

若已知初值状态点 $\boldsymbol{\xi}(0)$, 将上式对时间积分可以得到系统的运动轨迹 $\boldsymbol{\xi}(t)$, $t \geq 0$ 。

定义 2.2.1 (平衡点) 若一旦 $\boldsymbol{\xi}(t_0) = \boldsymbol{\xi}^*$, $\boldsymbol{\xi}(t) \equiv \boldsymbol{\xi}^*$, $t \geq t_0$, 则状态 $\boldsymbol{\xi}^*$ 称为系统的一个平衡点。

设感兴趣的平衡点为 $\boldsymbol{\xi}^*$, 为了记号和分析的方便, 构建变量 $\boldsymbol{\xi}' = \boldsymbol{\xi} - \boldsymbol{\xi}^*$, 得到新的方程:

$$\dot{\boldsymbol{\xi}}' = \mathbf{f}(\boldsymbol{\xi}' + \boldsymbol{\xi}^*). \quad (2.4)$$

因此若要研究方程(2.3)在平衡点 $\boldsymbol{\xi}^*$ 的性质, 只需要研究方程(2.4)在零点处的性质。为描述的简便, 本文设定方程(2.3)的平衡点为零点。

定义 2.2.2 (局部稳定) 若 $\forall R > 0$, $\exists r > 0$, 使得当 $\|\boldsymbol{\xi}(0)\| < r$ 时, $\|\boldsymbol{\xi}(t)\| < R$, $t > 0$, 则平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 是局部稳定的。

定义 2.2.3 (渐进稳定) 若平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 是局部稳定的, 且 $\exists r > 0$, 使得当 $\|\boldsymbol{\xi}(0)\| < r$ 时, $\lim_{t \rightarrow \infty} \boldsymbol{\xi}(t) = \mathbf{0}$, 则平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 是局部渐进稳定的。

定义 2.2.4 (全局渐进稳定) 对 $\forall \boldsymbol{\xi}(0) \in \mathbb{R}^D$, 平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 都是局部渐进稳定的, 则平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 是全局渐进稳定的。

若式(2.3)为线性时不变系统:

$$\dot{\boldsymbol{\xi}} = \mathbf{A}\boldsymbol{\xi}, \quad (2.5)$$

即 $\mathbf{A}$ 为常矩阵, 当且仅当矩阵 $\mathbf{A}$ 的所有特征值实部为负数时, 该系统全局渐进稳定, 轨迹线为 $\boldsymbol{\xi}(t) = e^{\mathbf{A}t}\boldsymbol{\xi}(0)$, 其指数收敛至唯一平衡点 $\boldsymbol{\xi} = \mathbf{0}$ 。

若式(2.3)为非线性系统, 一般基于 Lyapunov 稳定判据^[148] 来分析其稳定性, 包含 Lyapunov 间接法 (又称为线性化方法) 和 Lyapunov 直接法。式(2.3)在零点的处可线性

化为 $\dot{\xi} = A\xi$, 其中 A 为其雅可比矩阵:

$$A = \left. \frac{\partial f}{\partial \xi} \right|_{\xi=0}. \quad (2.6)$$

定理 2.2.1 (Lyapunov 间接法^[148]) 若零点为 $\dot{\xi} = f(\xi)$ 的平衡点, 且矩阵 A 的所有特征值实部为负数, 则其在零点处局部渐进稳定。

定理 2.2.2 (Lyapunov 直接法^[148]) 若存在一个标量的连续可微函数 $V(\xi)$ 满足:

$$V(\xi) \geq 0, \forall \xi \in \mathbb{R}^D \text{ 且 } V(\xi) = 0 \Leftrightarrow \xi = 0, \quad (2.7a)$$

$$\dot{V}(\xi) \leq 0, \forall \xi \in \mathbb{R}^D \text{ 且 } \dot{V}(\xi) = 0 \Leftrightarrow \xi = 0, \quad (2.7b)$$

$$\text{当 } \|\xi\| \rightarrow \infty, V(\xi) \rightarrow \infty \text{ (径向无界)}, \quad (2.7c)$$

则该系统为全局渐进稳定的, 平衡点为 0 。满足式(2.7a)–(2.7b)的函数 $V(\xi)$ 称为 Lyapunov 函数。

定理 2.2.3 若零点为 $\dot{\xi} = f(\xi)$ 的平衡点, 设 $A(\xi)$ 为其雅可比矩阵, 若对 $\forall \xi \in \Omega$, A 所有特征值都为负, 则零点为其渐近稳定的平衡点, 该系统的 Lyapunov 函数可选为

$$V(\xi) = f^T(\xi)f(\xi), \quad (2.8)$$

其中 Ω 为原点的一个邻域, 若 Ω 为整个空间 $\mathbb{R}^D$, 且当 $\|\xi\| \rightarrow \infty, V(\xi) \rightarrow \infty$, 则该平衡点为全局渐近稳定的。

证明 矩阵 $A(\xi)$ 特征值都为负, 则其为负定矩阵, 对 $\forall \xi \in \Omega (\xi \neq 0)$, 有:

$$\xi^T A \xi < 0. \quad (2.9)$$

对上式转置可得

$$(\xi^T A \xi)^T = \xi^T A^T \xi < 0. \quad (2.10)$$

两式叠加得:

$$\xi^T A \xi + \xi^T A^T \xi = \xi^T (A + A^T) \xi < 0. \quad (2.11)$$

所以矩阵 $A + A^T$ 也是负定的。

由于零点为其平衡点, 则有 $f(0) = 0$ 。假设存在 $\xi_1 \in \Omega (\xi_1 \neq 0)$, 使得 $f(\xi_1) = 0$ 。因 $f(x)$ 在 Ω 内可导, 由中值定理可知, 必定存在介于 ξ_1 和 0 之间一点 ξ_2 , 使得 $\left. \frac{\partial f(\xi)}{\partial \xi} \right|_{\xi=\xi_2} = 0$, 即 $A(\xi_2) = 0$, 这与矩阵 $A(\xi)$ 的负定性矛盾。所以原假设不成立, 即当 $\xi \neq 0$ 时, $f(\xi_1) \neq 0$ 。Lyapunov 函数 $V(\xi) = f^T(\xi)f(\xi)$ 满足式(2.7a)。

将式 $\dot{\xi} = f(\xi)$ 两边对时间求导, 可得:

$$\ddot{\xi} = \dot{f} = A\dot{\xi} = Af. \quad (2.12)$$

将 Lyapunov 函数式(2.8)对时间求导：

$$\dot{V}(\boldsymbol{\xi}) = \dot{\mathbf{f}}^T \mathbf{f} + \mathbf{f}^T \dot{\mathbf{f}} = \mathbf{f}^T \mathbf{A}^T \mathbf{f} + \mathbf{f}^T \mathbf{A} \mathbf{f} = \mathbf{f}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{f}. \quad (2.13)$$

由于矩阵 $\mathbf{A} + \mathbf{A}^T$ 是负定的，所以 $\dot{V}(\boldsymbol{\xi})$ 也是负定的，即满足式(2.7b)。因此，零点为其渐进稳定的平衡点。若 Ω 为整个空间 $\mathbb{R}^D$ ，且当 $\|\boldsymbol{\xi}\| \rightarrow \infty, V(\boldsymbol{\xi}) \rightarrow \infty$ ，则该平衡点（零点）为全局渐近稳定的。□

本论文主要基于式(2.2)构建全局渐进稳定的时不变系统，用其描述机械臂末端的运动轨迹。若其平衡点（即轨迹的终点）在机械臂运动的过程中发生变化，轨迹也能动态的适应并收敛至平衡点，体现了基于 DS 模型的运动规划方法的在线适应性。若机械臂运动过程中受到位姿扰动或者时间扰动，由于机械臂运行速度仅与当前状态（即当前位姿）相关，运动轨迹也能收敛至平衡点。式(2.2)可以将系统状态取机械臂的位置和速度项，则可进行加速度的规划。本文中基于动态系统的运动规划仅考虑位置和速度的规划，并保证位置和速度的连续可微。

2.3 模仿学习

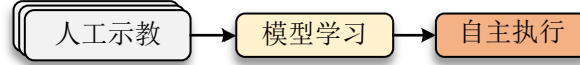


图 2.2 模仿学习框架

模仿学习可以分为三个部分，包含示教、模型学习、执行，如图2.2所示。

其中，模型学习部分一般采用非线性回归算法来编码人的示教数据。非线性回归算法用于构建输入数据 $\boldsymbol{\xi}^i \in \mathbb{R}^m$ 和输出数据 $\boldsymbol{\xi}^o \in \mathbb{R}^n$ 之间的连续映射函数 $\mathbf{f} : \boldsymbol{\xi}^i \rightarrow \boldsymbol{\xi}^o$ 。映射函数 $\mathbf{f}$ 的参数 $\boldsymbol{\theta}$ 可通过优化算法进行求解。机器学习中常用的非线性回归方法有 GMM, GPR, SVR, LWR 等。本节中介绍 GMM 算法^[85]，其将示教数据编码为概率模型，并采用 GMR 进行数据回归。

考虑示教数据 $\boldsymbol{\xi} = [\boldsymbol{\xi}^i, \boldsymbol{\xi}^o]$ ，其中 $(\cdot)^i$ 表示映射输入， $(\cdot)^o$ 表示其输出。采用含有 K 个高斯模型的混合来编码数据，数据点 $\boldsymbol{\xi}_0 \in \mathbb{R}^D$ 在数据集 $\boldsymbol{\xi} = [\boldsymbol{\xi}^i, \boldsymbol{\xi}^o]$ 下的概率为：

$$\begin{aligned} p(\boldsymbol{\xi}_0) &= \sum_{k=1}^K \pi_k \mathcal{N}(\boldsymbol{\xi}_0; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \\ &= \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^D |\boldsymbol{\Sigma}_k|}} e^{-\frac{1}{2}((\boldsymbol{\xi}_0 - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{\xi}_0 - \boldsymbol{\mu}_k))}, \end{aligned} \quad (2.14)$$

其中 $\pi_k \in [0, 1]$ 为先验概率, $\sum_{k=1}^K \pi_k = 1$, $\boldsymbol{\mu}_k \in \mathbb{R}^D$, $\boldsymbol{\Sigma}_k \in \mathbb{R}^{D_i \times D_o}$ 分别表示 GMM 中高斯分布的均值和方差矩阵。其次, 定义输入和输出量的均值和方差:

$$\boldsymbol{\mu}_k = \begin{bmatrix} \boldsymbol{\mu}_k^i \\ \boldsymbol{\mu}_k^o \end{bmatrix}, \boldsymbol{\Sigma}_k = \begin{bmatrix} \boldsymbol{\Sigma}_k^i & \boldsymbol{\Sigma}_k^{io} \\ \boldsymbol{\Sigma}_k^{oi} & \boldsymbol{\Sigma}_k^o \end{bmatrix}. \quad (2.15)$$

对给定的输入变量 $\boldsymbol{\xi}^i$, 其输出 $\boldsymbol{\xi}^o$ 的条件概率为:

$$p(\boldsymbol{\xi}^o | \boldsymbol{\xi}^i) = \sum_{k=1}^K p(k | \boldsymbol{\xi}^i) \mathcal{N}(\hat{\boldsymbol{\xi}}_k, \hat{\boldsymbol{\Sigma}}_k), \quad (2.16a)$$

其中:

$$\hat{\boldsymbol{\xi}}_k = \boldsymbol{\mu}_k^o + \boldsymbol{\Sigma}_k^{oi} (\boldsymbol{\Sigma}_k^i)^{-1} (\boldsymbol{\xi}^i - \boldsymbol{\mu}_k^i), \quad (2.16b)$$

$$\hat{\boldsymbol{\Sigma}}_k = \boldsymbol{\Sigma}_k^o - \boldsymbol{\Sigma}_k^{oi} (\boldsymbol{\Sigma}_k^i)^{-1} \boldsymbol{\Sigma}_k^{io}. \quad (2.16c)$$

权重函数 $p(k | \boldsymbol{\xi}^i)$ 表示 $\boldsymbol{\xi}^i$ 在第 k 个高斯模型的概率, 且有:

$$p(k | \boldsymbol{\xi}^i) = \frac{p(k) p(\boldsymbol{\xi}^i | k)}{\sum_{j=1}^K p(j) p(\boldsymbol{\xi}^i | j)} = \frac{\pi_k \mathcal{N}(\boldsymbol{\xi}^i; \boldsymbol{\mu}_k^i, \boldsymbol{\Sigma}_k^i)}{\sum_{j=1}^K \pi_j \mathcal{N}(\boldsymbol{\xi}^i; \boldsymbol{\mu}_j^i, \boldsymbol{\Sigma}_j^i)}. \quad (2.17)$$

基于式2.16a, GMR 通过计算给定 $\boldsymbol{\xi}^i$ 下 $\boldsymbol{\xi}^o$ 的条件期望:

$$E\{p(\boldsymbol{\xi}^o | \boldsymbol{\xi}^i)\} = \sum_{k=1}^K p(k | \boldsymbol{\xi}^i) \hat{\boldsymbol{\xi}}_k. \quad (2.18)$$

因此, GMM/GMR 基于参数 $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$, 且一般由 EM 算法来迭代求解。其中, 超参数 K 基于贝叶斯信息准则 (Bayesian Information Criterion, BIC) 来选取。

相较于其他非线性回归算法, GMR 可以高效地建立多输入多输出的概率密度模型, 训练效率高, 在机械臂运行过程中也可在线计算当前状态下的后验概率, 可用于机械臂的实时规划和控制。

2.4 机械臂运动规划与控制

为控制机械臂执行期望的动作, 需要规划机械臂的运动轨迹并控制机械臂各个关节的运动以跟踪该轨迹。本节首先介绍机械臂的位姿描述方式, 主要是基于四元数的姿态表达, 其次介绍机械臂的运动学和动力学, 最后介绍常用的几种机械臂运动控制方法。

2.4.1 位姿描述

刚体在三维空间中自由度为 6 个，包含 3 个位置和 3 个姿态自由度。位置的描述可以直接采用笛卡尔坐标系下的三维坐标。姿态的表示方式有欧拉角，旋转矩阵，四元数和轴角四种方式。欧拉角由 3 个绕 x , y , z 轴转动的角度组成，自身存在万向节锁死^[149]、插值困难等问题。旋转矩阵为 3×3 的单位正交矩阵，存在表达冗余，计算量大等问题。基于四元数^[150] 的姿态表示方式因其插值简单，无万向节锁死，旋转的复合简单，应用最为广泛。本节简述四元数的基本表达和性质。

单位四元数为定义在 4 维空间的 3 维超球面，因此被称作为李群 $\mathcal{S}^3$ 。考虑单位四元数 $\mathbf{q} \in \mathcal{S}^3$:

$$\mathbf{q} = [q_w, \mathbf{q}_v^T]^T, \|\mathbf{q}\| = 1, \quad (2.19)$$

其中 $q_w \in \mathbb{R}$, $\mathbf{q}_v \in \mathbb{R}^3$ 分别为四元数的实部和虚部。本文中的四元数均默认为单位四元数。四元数的逆:

$$\mathbf{q}^{-1} = \bar{\mathbf{q}} = [q_w, -\mathbf{q}_v^T]^T, \quad (2.20)$$

其中 $\bar{\mathbf{q}}$ 为四元数的共轭。考虑两个单位四元数 $\mathbf{p}$ 和 $\mathbf{q}$ ，四元数乘法定义为:

$$\mathbf{p} * \mathbf{q} = [p_w q_w - \mathbf{p}_v \cdot \mathbf{q}_v, p_w \mathbf{q}_v^T + q_w \mathbf{p}_v^T + \mathbf{p}_v^T \times \mathbf{q}_v^T]^T, \quad (2.21)$$

即为在四元数 $\mathbf{q}$ 的基础上旋转 $\mathbf{p}$ 得的四元数。本文用 “*” 来表示两个四元数之间的乘法，以便于和矩阵乘法作区分。四元数乘法满足结合律，但不满足交换律，即:

$$\mathbf{q}_0 * (\mathbf{q}_1 * \mathbf{q}_2) = (\mathbf{q}_0 * \mathbf{q}_1) * \mathbf{q}_2. \quad (2.22)$$

一般来说 $\mathbf{q}_0 * \mathbf{q}_1 \neq \mathbf{q}_1 * \mathbf{q}_0$ 。

两个四元数复合的导数为:

$$\frac{d(\mathbf{p} * \mathbf{q})}{dt} = \dot{\mathbf{p}} * \mathbf{q} + \mathbf{p} * \dot{\mathbf{q}}. \quad (2.23)$$

三维空间中的特殊正交 (Special Orthogonal) 欧式群被称为 $SO(3)$ ，为姿态旋转矩阵的集合。设旋转矩阵 $\mathbf{R} = (r_{i,j})_{i=1,2,3}^{j=1,2,3}$ ，可将旋转矩阵转化为单位四元数:

$$\begin{cases} q_w = \frac{1}{2} \sqrt{1 + r_{11} + r_{22} + r_{33}}, \\ \mathbf{q}_v = \left[\frac{r_{32} - r_{23}}{4q_w}, \frac{r_{13} - r_{31}}{4q_w}, \frac{r_{21} - r_{12}}{4q_w} \right]^T. \end{cases} \quad (2.24)$$

设 $\mathbf{q} = [a, b, c, d]^T$ ，可将四元数转化为旋转矩阵:

$$\mathbf{R} = \begin{bmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2bd + 2ac \\ 2bc + 2ad & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2cd + 2ab & a^2 - b^2 - c^2 + d^2 \end{bmatrix}. \quad (2.25)$$

由于 $\mathbf{q}$ 和 $-\mathbf{q}$ 表示相同的旋转矩阵 $\mathbf{R}$, 因此四元数流形 $\mathcal{S}^3$ 是 $SO(3)$ 的双倍覆盖。

为实现姿态的优化和距离度量, 需将四元数表达为三维笛卡尔空间的向量。将四元数 $\mathbf{q}$ 在无旋转四元数 $\mathbf{q}_I = [1, 0, 0, 0]^T$ (identity quaternion) 的切平面投影定义为四元数的对数 $\log: \mathcal{S}^3 \mapsto \mathbb{R}^3$:

$$\mathbf{r} = \log(\mathbf{q}) = \begin{cases} \text{acos}(q_w) \frac{\mathbf{q}_v}{\|\mathbf{q}_v\|}, & \|\mathbf{q}_v\| \neq 0, \\ [0, 0, 0]^T, & \|\mathbf{q}_v\| = 0, \end{cases} \quad (2.26)$$

其中 $\mathbf{r} \in \mathbb{R}^3$, acos 为修正后的反余弦函数^[151]:

$$\text{acos}(\theta) = \begin{cases} \arccos(\theta) - \pi, & -1 \leq \theta < 0, \\ \arccos(\theta), & 0 \leq \theta \leq 1. \end{cases} \quad (2.27)$$

因此 $\text{acos}^*(\theta) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, 定义指数函数 $\exp: \mathbb{R}^3 \mapsto \mathcal{S}^3$, 将其恢复为单位四元数:

$$\mathbf{q} = \exp(\mathbf{r}) = \begin{cases} \left[\cos(\|\mathbf{r}\|), \sin(\|\mathbf{r}\|) \frac{\mathbf{r}^T}{\|\mathbf{r}\|} \right]^T, & \text{if } \|\mathbf{r}\| \neq 0, \\ [1, 0, 0, 0]^T, & \|\mathbf{r}\| = 0. \end{cases} \quad (2.28)$$

旋转角速度 $\boldsymbol{\omega} \in \mathbb{R}^3$ 和四元数导数之间的关系为:

$$[0, \boldsymbol{\omega}^T]^T = 2\dot{\mathbf{q}} * \bar{\mathbf{q}}, \quad (2.29)$$

即有 $\dot{\mathbf{q}} = \frac{1}{2} [0, \boldsymbol{\omega}^T]^T * \mathbf{q}$ 。若考虑离散过程, 设将四元数 $\mathbf{q}_0$ 旋转得到四元数 $\mathbf{q}_1$, 过程耗时为 t , 则该旋转平均角速度为:

$$\boldsymbol{\omega} = 2 \log(\mathbf{q}_1 * \bar{\mathbf{q}}_0) / t. \quad (2.30)$$

四元数可以实现简单平滑的旋转插值, 考虑两个四元数 $\mathbf{q}_0$ 和 $\mathbf{q}_1$, 期望设计四元数 $\mathbf{q}(t), 0 \leq t \leq 1$, 基于以下表达可实现球面线性插值:

$$\mathbf{q}(t) = \mathbf{q}_0 * \exp(t \log(\bar{\mathbf{q}}_0 * \mathbf{q}_1)) = \mathbf{q}_0 * (\bar{\mathbf{q}}_0 * \mathbf{q}_1)^t. \quad (2.31)$$

当 $t = 0$ 时, $\mathbf{q} = \mathbf{q}_0$; 当 $t = 1$ 时, $\mathbf{q} = \mathbf{q}_1$ 。若 $\mathbf{q}_0 = \mathbf{q}_I$, 则 $\mathbf{q}(t) = \mathbf{q}_1^t$ 。

在优化任务中, 一般基于四元数之间的距离设计目标函数, 两个四元数 $\mathbf{q}_0$ 和 $\mathbf{q}_1$ 之间的距离定义为:

$$d(\mathbf{q}_0, \mathbf{q}_1) = 2\|\log(\mathbf{q}_0 * \bar{\mathbf{q}}_1)\|. \quad (2.32)$$

$d(\mathbf{q}_0, \mathbf{q}_1) \in [0, \pi]$ 为两个四元数之间的测地线角度, 当且仅当 $\mathbf{q}_0 = \pm \mathbf{q}_1$ 时, $d(\mathbf{q}_0, \mathbf{q}_1) = 0$ 。

2.4.2 运动学

机械臂运动学主要研究各个关节运动和末端运动之间的关系，在运动规划中，最关心的问题是求解关节角度序列以使得机械臂末端执行器达到期望的位姿。

考虑由 n 个旋转关节串联的机械臂，一般采用 DH 法^[152] 来建立关节空间和操作空间的齐次变换矩阵。设末端执行器的位姿为 $\mathbf{p} \in \mathbb{R}^6$ ，关节角度序列为 $\boldsymbol{\theta} \in \mathbb{R}^n$ ，正运动学模型为已知关节角度求解末端位姿：

$$\mathbf{p} = \mathbf{f}(\boldsymbol{\theta}). \quad (2.33)$$

对上式求时间的导数，可得末端速度与关节角速度的关系：

$$\dot{\mathbf{p}} = \mathbf{J}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}}, \quad (2.34)$$

其中， $\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \in \mathbb{R}^{6 \times n}$ 为正运动学的雅可比矩阵。

逆运动学为已知末端位姿求解对应的关节角度序列 $\boldsymbol{\theta}$ ，一般有解析法和迭代法两种方式。解析法直接求解式(2.33)，一般商用的机械臂采用此方式，求解速度快，且能输出多个解。迭代法中，设参考轨迹为 $\mathbf{p}_d$ ，对位姿误差 $\mathbf{r} = \mathbf{f}(\boldsymbol{\theta}) - \mathbf{p}_d$ 在 $\boldsymbol{\theta} = \boldsymbol{\theta}_k$ 处进行 Taylor 展开：

$$\mathbf{r} \approx \mathbf{r}(\boldsymbol{\theta}_k) + \mathbf{J}(\boldsymbol{\theta}_k)(\boldsymbol{\theta} - \boldsymbol{\theta}_k). \quad (2.35)$$

由牛顿法求 $\mathbf{r} = \mathbf{0}$ 的解，可取迭代式：

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \alpha \mathbf{J}^\dagger(\boldsymbol{\theta}_k)(\mathbf{f}(\boldsymbol{\theta}_k) - \mathbf{p}_d), \quad (2.36)$$

其中， $\mathbf{J}^\dagger = \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}$ 为 $\mathbf{J}$ 的伪逆， $\alpha > 0$ 为迭代步长。当误差 $\mathbf{r}$ 足够小或迭代超时，停止迭代并输出结果。

迭代法的通用性较好，但一般仅能输出一个解，且计算效率比解析法低。本文的机械臂实验采用 KDL 运动学库¹构建机械臂的运动学模型，并基于 TRAC-IK^[153] 逆运动学求解器²计算末端参考轨迹对应的关节角度序列。TRAC-IK 的求解采用迭代法，相较于其他迭代算法效率较高，稳定性好，且已集成至 ROS 中，便于开发和使用。

2.4.3 动力学

机械臂的动力学主要研究其运动和作用力之间的关系，一般采用拉格朗日 (Lagrange) 法或牛顿-欧拉 (Newton-Euler) 法建模。 n 自由度串联机械臂的刚体动力学模型^[150] 可描述为：

$$\boldsymbol{\tau} = \mathbf{M}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \mathbf{c}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \mathbf{g}(\boldsymbol{\theta}) + \mathbf{J}(\boldsymbol{\theta})^T \mathbf{f}_{ext}, \quad (2.37)$$

¹https://github.com/orocos/orocos_kinematics_dynamics

²https://bitbucket.org/tracilabs/trac_ik/src/master/

其中 $\theta, \dot{\theta}, \ddot{\theta} \in \mathbb{R}^n$ 分别为关节位置、速度和加速度。质量矩阵 $M(\theta) \in \mathbb{R}^{n \times n}$ ，科里奥利力 (Coriolis)、离心力矩阵 $c(\theta, \dot{\theta}) \in \mathbb{R}^n$ ，以及重力项 $g(\theta) \in \mathbb{R}^n$ 为仅依赖于状态的动力学模型参数。 $f_{ext} \in \mathbb{R}^6$ 为笛卡尔空间的接触外力/力矩，其左乘雅可比矩阵的转置可将其转换至关节空间。 $\tau \in \mathbb{R}^n$ 为电机控制力矩。

建立精确的机械臂动力学模型，可以设计基于模型的控制器的，在给定参考轨迹（操作空间或关节空间）下，计算所需的控制力矩 τ ，以实现更好的轨迹跟踪效果和动态性能。

2.4.4 运动控制

一般来说，机械臂运动控制的目的是控制机械臂各关节运动以跟踪参考轨迹，参考轨迹可以是关节空间或末端执行器的操作空间（笛卡尔空间）。由于本文中的操作任务都是采用末端执行器与环境交互，规划的轨迹在笛卡尔空间中，采用逆运动学求解关节角度序列，并使用关节空间的控制器的跟踪期望的运动。针对力矩控制的机械臂，可以设计关节空间或笛卡尔空间的控制器的，直接求解关节控制力矩。

传统的工业机械臂一般仅开放位置/速度控制接口，其内部采用采用纯比例-积分-微分 (Proportional-Integral-Derivative, PID) 反馈的控制器的实现高精度轨迹跟踪：

$$\tau = K_p e + K_i \int e dt + K_d \dot{e}, \quad (2.38)$$

其中 $e = \theta_d - \theta$ 为关节角度误差， $K_p, K_i, K_d \in \mathbb{R}^{n \times n}$ 分别为比例积分微分系数，均为对角正定矩阵。

为实现高精度轨迹跟踪， K_p 一般取值较大，即机械臂的刚度很大。若需要应用此类机械臂进行接触式任务的操作，需在其末端安装力传感器，以检测机械臂末端与环境之间的交互力，并根据交互力来修正机械臂的运动，以实现柔顺的交互操作。为建立交互力和位置修正量之间的关系，常采用式(1.2)所述的弹簧质量阻尼系统，即导纳控制，其控制系统框图2.3如图所示。

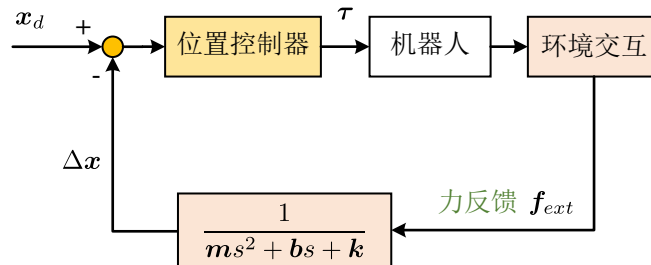


图 2.3 导纳控制器框图

导纳控制器根据力传感器测量的机械臂末端与环境之间接触力 f_{ext} 计算所需的修正位姿，用于修正参考轨迹 x_d ，并通过位置跟踪控制器来控制机械臂运动。由于实际

机械臂的控制力矩不可能无限大，跟踪修正后的位姿 $\mathbf{x}_d - \Delta \mathbf{x}$ 需要一定的时间，即导纳控制器存在的一定的时间滞后，响应速度较慢。实际中通常与被动柔顺装置相结合，以提高响应速度，避免接触力过大。

对于机械臂动力学模型已知且力矩控制接口开放的机械臂，为实现参考轨迹的快速精密跟踪，可设计基于计算力矩前馈和 PD 反馈的控制率^[150]：

$$\boldsymbol{\tau} = \mathbf{M}(\boldsymbol{\theta})(\ddot{\boldsymbol{\theta}}_d + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e}) + \mathbf{c}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \mathbf{g}(\boldsymbol{\theta}). \quad (2.39)$$

将上式带入式(2.37)，若无外力影响，得其闭环系统为：

$$\ddot{\mathbf{e}} + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e} = \mathbf{0}. \quad (2.40)$$

为保证系统的稳定性，取 $\mathbf{K}_p$ 和 $\mathbf{K}_v$ 为对角正定矩阵，则跟踪误差 $\mathbf{e}$ 将收敛至零。

在接触式任务操作中，期望机械臂能与环境柔顺接触，针对动力学模型已知的机械臂，可设计如下阻抗控制器^[154]：

$$\boldsymbol{\tau} = \mathbf{K} \mathbf{e} + \mathbf{D} \dot{\mathbf{e}} + \mathbf{M}(\boldsymbol{\theta}) \ddot{\boldsymbol{\theta}} + \mathbf{c}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \mathbf{g}(\boldsymbol{\theta}). \quad (2.41)$$

其中 $\mathbf{K}, \mathbf{D} \in \mathbb{R}^{n \times n}$ 分别为对角正定的刚度阻尼矩阵。将上式带入式(2.37)，可得其闭环系统：

$$\mathbf{K} \mathbf{e} + \mathbf{D} \dot{\mathbf{e}} + \mathbf{M} \ddot{\mathbf{e}} = \mathbf{J}(\boldsymbol{\theta})^T \mathbf{f}_{ext}. \quad (2.42)$$

即机械臂与环境之间交互为弹簧质量阻尼系统模型。

2.5 本章小结

本章概述了本文所需的基础理论和工具。首先引入了微分同胚映射的基本概念和性质，随后阐述了基于 DS 的运动规划算法，以及 Lyapunov 稳定判据。其次，简述了模仿学习中的 GMM 方法，用于构建输入和输出之间的映射关系。为实现机械臂的运动规划和控制，采用四元数描述机械臂的姿态，并描述其基本性质和方法。最后，分析了机械臂的运动学和动力学，以及常用的运动控制方法。

3 融合位置和姿态的微分同胚映射算法

3.1 引言

在运动规划中，基于模仿学习的方法因其实现简单方便，无需专家经验知识，可以快速生成轨迹，应用十分广泛。其中，基于行为克隆的方法首先选取某种描述模型，如 GMR, GPR, SVR, LWR 或 NN 等，将示教数据看作为时间到轨迹的映射观测点，并采用最大似然估计来求解模型的参数。这种方法可以生成非常精确地编码多维非线的示教数据。然而，由于此类方法精确依赖于时间且为开环（无状态反馈）的规划算法，在有外界时间或空间的扰动时，机械臂极易失稳。一般需要基于时间做重新采样和在线优化，以生成新的轨迹，但对位置动态环境和干扰的适应性仍然很差，也难以设计出通用的重采样算法。

基于 DS 的运动规划方法可以实现基于状态反馈的轨迹在线生成。该方法本身便具有内在的对动态环境和时空扰动的在线适应性。用 DS 来直接描述人的示教轨迹数据，可以直接用于机械臂运动轨迹的生成。其包含了整个空间中所有可能的运动至终点的轨迹。最常用的 DS 方法为 DMP，可以实现全局稳定的收敛，但如表1.1所示，此类方法及其改进方法仍旧严格依赖时间量。SEDS^[96] 及其后续改进方法虽能构建出全局渐进稳定的时不变 DS，但一般依赖于两步学习：构建 Lyapunov 函数和求解 DS，其训练耗时较长，且由于稳定性的约束导致其对示教数据的表征效果较差。此外，此类方法仅设计了二维或三维位置的 DS，针对姿态的规划没有做相关研究。当控制机械臂与环境或人在真实三维空间交互时，姿态的规划和控制也十分重要。

为快速的构造 DS 以规划机械臂的运动，可基于微分同胚映射来对复杂的示教数据进行简化，并利用简化后的轨迹建立 DS 模型^[111, 115]。然而，现有的微分同胚映射只考虑位置的可逆且连续可微的变换，没有考虑姿态信息，因此无法针对含有姿态的轨迹数据进行学习和建模。

为快速构建融合位置和姿态的时不变 DS 来进行运动规划，本章设计一个融合位置和姿态的微分同胚映射框架，可实现示复杂教数据可逆的简化。通过构建用于表征简化后数据的 DS 模型，并将其映射回原空间，以实现原空间中基于 DS 的运动规划。第4章中将建立示教数据空间的 DS 和一个简化后空间中全局稳定的 DS 之间的映射关系，从而快速求解出示教空间中全局稳定的 DS。

本章首先分析了该融合融合位置和姿态的微分同胚映射的约束条件，并提出了统一的设计框架，随后在此框架下提出了两种实施方法：基于径向基函数的迭代法和可逆神经网络映射算法。在二维仿真中验证了两种算法的有效性并对比了性能。最后采用双机械臂实时遥操作实验验证两个映射算法的有效性。

3.2 约束分析

本节首先提出了采用微分同胚映射来构建 DS 的基本流程，并分析了映射算法中的约束条件。

为快速地生成满足任务约束条件的运动轨迹，一般基于人工示教数据来学习 DS。考虑包含 N 个数据点的一组示教数据 $B = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N$ ，其中 $\mathbf{x}' \in \mathbb{R}^{D_0}$, ($D_0 = 2$ 或 3), $\mathbf{q}' \in \mathcal{S}^3$ 分别为位置笛卡尔坐标和姿态四元数。采用 $B = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N := \{\mathbf{x}'_i - \mathbf{x}'_N, \bar{\mathbf{q}}'_N * \mathbf{q}'_i\}_{i=1}^N$ 将数据进行平移和旋转，可得终点位置为 $\mathbf{0}$ 且终点姿态为无旋转四元数 $\mathbf{q}_I$ 的示教数据。下面结合二维空间中一个实例来说明具体过程，如图3.1(b)所示红色轨迹线为“J”形的示教数据，目的是基于该示教数据构建式2.3中的 DS，并满足如下约束：

- 1) 全局渐进稳定，且只有唯一平衡点 $\xi^* = [\mathbf{0}, \mathbf{q}_I]$ ；
- 2) 该 DS 应该尽可能精确地表征示教数据 B ；

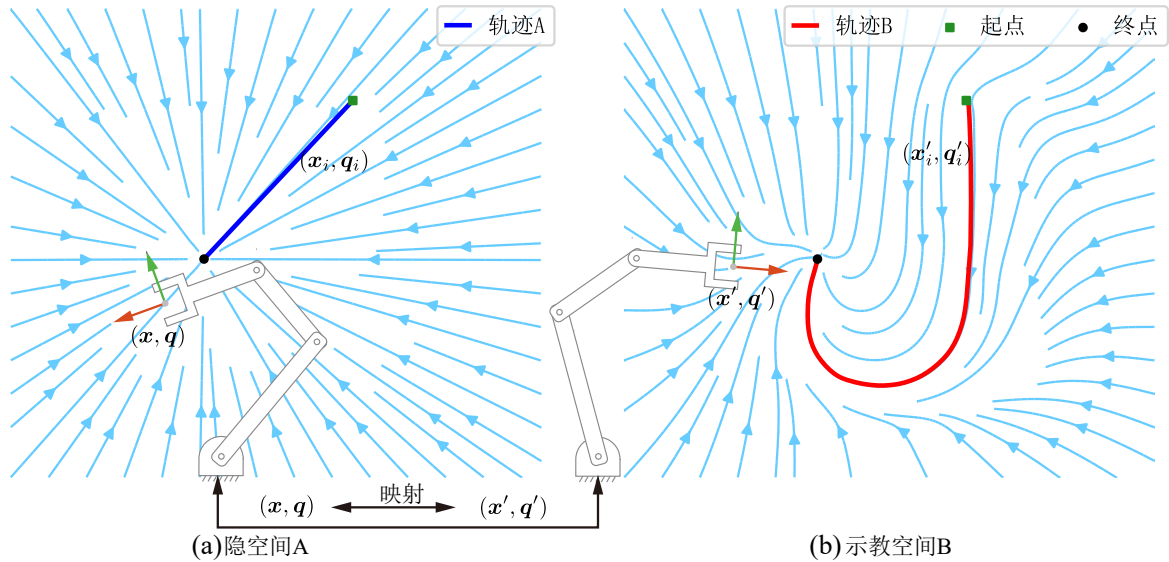


图 3.1 采用映射的方法生成 DS 的原理

如图3.1所示，为快速构建全局稳定的 DS，本文首先在隐空间 A 中构建出全局渐进稳定收敛至零点的 DS: $\xi = f(\xi)$ ，然后建立两个空间之间的微分同胚映射 Φ ，以使得轨迹 A 可按式(2.1)映射至轨迹 B，从而得到示教空间中所需的 DS。若在隐空间 A 中设计出收敛至零点且能精确表征轨迹 A 的 DS，则可生成空间 B 中满足要求的 DS。(具体证明见章节4.2。)

设示教数据 B 在隐空间中对应的数据为 $A = \{\mathbf{x}_i, \mathbf{q}_i\}_{i=1}^N$ 。本章的目的是已知 B 和 A 的条件下，求解两个空间之间的微分同胚映射关系 Φ ，以使得 $\Phi: A \rightarrow B$ 。设图3.1中，隐空间 A 中机械臂 A 末端位姿为 $\xi = (\mathbf{x}, \mathbf{q})$ ，空间 B 中机械臂 B 末端位姿为 $\xi' = (\mathbf{x}', \mathbf{q}')$ ，映射 $\Phi: (\mathbf{x}, \mathbf{q}) \Leftrightarrow (\mathbf{x}', \mathbf{q}')$ ，若隐空间 A 中机械臂可以全局稳定的收敛至

终点 $\mathbf{A}_N$ ，则机械臂 B 也是全局稳定的收敛至终点 $\mathbf{B}_N$ 。其中，映射 Φ 需满足如下约束条件：

- 1) 当左侧机械臂靠近第 i 个点时，即 $\mathbf{x} \rightarrow \mathbf{x}_i$ ，右侧机械臂也应该靠近相应的点，即 $\mathbf{x}' \rightarrow \mathbf{x}'_i$ 。
- 2) 当机械臂靠近某个点时，两侧相对该点的姿态应该相同，即若 $\mathbf{x} = \mathbf{x}_i$, $\Delta \mathbf{q}_i = \bar{\mathbf{q}}_i * \mathbf{q} = \bar{\mathbf{q}}'_i * \mathbf{q}'$ ，其中 $\Delta \mathbf{q}_i$ 为机械臂末端执行器相对于第 i 个点的姿态。
- 3) 位置和姿态的映射函数均应该为可逆、连续且连续可微的。
- 4) 位置和姿态映射的梯度不应距离单位矩阵太远，以保证机械臂的运动在安全的速度限制之内。

第一和第二约束保证两侧的机械臂运动趋势相同，且当贴近示教轨迹时，相对的姿态相同。第三条约束保证机械臂的运动是平滑的。首先，位姿映射的连续性和连续可微性保证了若机械臂 A 的轨迹是连续可微的，则机械臂 B 位置和速度也是连续的。映射的可逆性保证了两侧机械臂的运动轨迹时严格一一对应的，不存在多解的情况。最后一条约束设定了位置位姿的梯度应该在安全允许的范围之内。

3.3 映射算法框架

根据章节3.2中的约束条件分析，本节提出了融合位置和姿态的映射算法框架，并定义了该映射算法的误差。

3.3.1 基本框架

根据映射的约束条件1) 和2)，位置和姿态是紧密联系的。姿态的约束建立在位置之上，仅当机械臂的位置 $\mathbf{x}$ 靠近数据 $\mathbf{A}$ 时，才对姿态有一定的约束。本节首先定义位置的可逆且连续可微的映射函数，即图3.1中两个机械臂末端位置 $\mathbf{x}$ 和 $\mathbf{x}'$ 之间的映射，定义正向映射 Φ_x 和反向映射 Φ_x^{-1} ：

$$\mathbf{x}' = \Phi_x(\mathbf{x}), \quad \mathbf{x} = \Phi_x^{-1}(\mathbf{x}'). \quad (3.1)$$

基于约束2)，可以看出，两个机械臂的姿态偏差仅仅取决于当前的机械臂位置 $\mathbf{x}$ ，因此，本节设计姿态映射为线性函数：

$$\mathbf{q}' = \Phi_q(\mathbf{x}, \mathbf{q}) = \mathbf{g}(\mathbf{x}) * \mathbf{q}, \quad \mathbf{q} = \bar{\mathbf{g}}(\mathbf{x}) * \mathbf{q}'. \quad (3.2)$$

其中 $\mathbf{g}(\mathbf{x}) \in \mathcal{S}^3$ 为两个机械臂末端姿态偏差，且为单位四元数，表示为位置 $\mathbf{x}$ 的函数。

根据式(3.1)的位置映射和式(3.2)姿态映射，两边对时间求导数，可得其速度映射关系：

$$\dot{\mathbf{x}}' = \frac{\partial \Phi_x(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}}, \quad \dot{\mathbf{x}} = \left(\frac{\partial \Phi_x(\mathbf{x})}{\partial \mathbf{x}} \right)^{-1} \dot{\mathbf{x}}', \quad (3.3a)$$

$$\dot{\mathbf{q}}' = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} * \mathbf{q} + \mathbf{g}(\mathbf{x}) * \dot{\mathbf{q}}, \quad (3.3b)$$

$$\dot{\mathbf{q}} = \frac{\partial \bar{\mathbf{g}}(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} * \mathbf{q}' + \bar{\mathbf{g}}(\mathbf{x}) * \dot{\mathbf{q}}', \quad (3.3c)$$

其中,角速度可以由式(2.29)计算得出。如约束4)所述,速度映射可以由位姿映射的梯度直接计算出来。可以看出,位姿映射的雅可比矩阵即为速度映射的缩放矩阵。

由定理2.1.1和定理2.1.2可知,针对连续可微的映射函数 $\Phi(\mathbf{x})$,若 $\forall \mathbf{x} \in \mathbb{R}^D$,有 $\det\left(\frac{\partial \Phi(\mathbf{x})}{\partial \mathbf{x}}\right) > 0$,则函数 $\Phi(\mathbf{x})$ 为可逆函数。

公式(3.1)–(3.3)为位置和姿态的微分同胚映射的基础框架。映射算法的连续性和连续可微性将由算法本身特性来保证。位置映射的可逆性可由映射函数的雅可比矩阵的行列式值来判别。针对姿态映射(3.2),由于姿态 $\mathbf{q}'$ 和 $\mathbf{q}$ 之间为线性关系,可逆性已满足,只需设计连续可微的函数 $\mathbf{g}(\mathbf{x})$ 即可。

3.3.2 映射误差

式(2.1)所描述的绝对理想的微分同胚映射一般难以实现。实际场景中,基于约束1)和2),只需映射算法的误差足够小即可,以保证机械臂可以以良好的位置精度完成操作任务。映射算法 Φ 的位姿和姿态误差定义为:

$$F_p(\Phi_x(\mathbf{A}), \mathbf{B}) = \frac{1}{N} \sum_{i=1}^N \|\Phi_x(\mathbf{x}_i) - \mathbf{x}_i'\|, \quad (3.4a)$$

$$F_q(\Phi_q(\mathbf{A}), \mathbf{B}) = \frac{1}{N} \sum_{i=1}^N d(\Phi_q(\mathbf{q}_i), \mathbf{q}_i'), \quad (3.4b)$$

其中位置误差函数 F_p 为笛卡尔空间的距离的均值,姿态误差函数 F_q 基于公式(2.32)求解两组四元数之间距离的均值。

可根据上述误差定义设计目标函数,并采用基于优化的方法来求解最优的微分同胚映射。

3.4 基于径向基函数的迭代算法

Perrin 在其论文^[115]中提出了一种基于迭代法的快速微分同胚映射算法。本节以该方法为基础,将其扩展到姿态的微分同胚映射,采用径向基函数(Radial Basis Function, RBF)进行迭代求解映射函数的参数,分别复合 K 个局部加权平移 $\phi_1(\mathbf{x}), \dots, \phi_K(\mathbf{x})$ 和 K 个局部加权旋转 $\varphi_1(\mathbf{x}), \dots, \varphi_K(\mathbf{x})$ 来构建位置映射 $\Phi_x(\mathbf{x})$ 和姿态映射 $\Phi_q(\mathbf{x})$,并分析如何求出最优的参数以满足约束条件1)–4)。

3.4.1 局部加权平移和旋转

考虑如下局部加权平移变换：

$$\phi(\mathbf{x}) = \mathbf{x} + k_\phi(\mathbf{x})\mathbf{v}_1, \quad (3.5a)$$

$$\text{且有 } k_\phi(\mathbf{x}) = e^{-\rho_1^2 \|\mathbf{x} - \mathbf{c}_1\|^2}, \quad (3.5b)$$

其中 $\mathbf{v}_1 \in \mathbb{R}^D$ 为平移变换的方向， $k_\phi(\cdot)$ 为中心为 $\mathbf{c}_1$ ，宽度为 $\rho_1 > 0$ 的 RBF。易知 $\phi(\mathbf{x})$ 是 C^∞ 光滑的函数。求式(3.5a)的雅可比矩阵，有：

$$\mathbf{J}_\phi = \mathbf{I}_2 - 2\rho^2 e^{-\rho^2 \|\mathbf{x} - \mathbf{c}_1\|} \mathbf{v}_1 (\mathbf{x} - \mathbf{c}_1)^T, \quad (3.6)$$

其中， $\mathbf{I}_2$ 为 2 阶单位矩阵。对上式求行列式值，可得：

$$\det(\mathbf{J}_\phi) \geq 1 - 2\rho^2 e^{-\rho^2 \|\mathbf{x} - \mathbf{c}_1\|} \|\mathbf{v}_1\| \cdot \|\mathbf{x} - \mathbf{c}_1\|. \quad (3.7)$$

对于函数 $f(x) = xe^{-\rho^2 x^2}$, $x \geq 0$ ，当 $x = \frac{1}{\sqrt{2}\rho}$ 时，其有最大值，因此：

$$\det(\mathbf{J}_\phi) \geq 1 - \sqrt{2} \|\mathbf{v}_1\| \rho e^{-\frac{1}{2}}. \quad (3.8)$$

由小结3.3.1的分析可知，若 $\forall(\mathbf{x}, \mathbf{v}_1) \in \mathbb{R}^D \times \mathbb{R}^D$, $\det\left(\frac{\partial \phi(\mathbf{x})}{\partial \mathbf{x}}\right) > 0$ ，则 $\phi(\mathbf{x})$ 为一个微分同胚映射。满足该要求的一个充分条件为：

$$1 - \sqrt{2} \|\mathbf{v}_1\| \rho e^{-\frac{1}{2}} > 0, \quad (3.9)$$

即 $0 < \rho_1 < \rho_{\max}(\mathbf{v}_1) = \frac{1}{\sqrt{2}\|\mathbf{v}_1\|} e^{\frac{1}{2}}$ 时， $\phi(\mathbf{x})$ 为一个微分同胚映射。

其次，考虑如下局部加权旋转变换：

$$\varphi(\mathbf{q}) = \mathbf{v}_2^{k_\varphi(\mathbf{x})} * \mathbf{q}, \quad (3.10a)$$

$$\text{且有 } k_\varphi(\mathbf{x}) = e^{-\rho_2^2 \|\mathbf{x} - \mathbf{c}_2\|^2}, \quad (3.10b)$$

其中， $\mathbf{v}_2 \in \mathcal{S}^3$ 为单位四元数，用于表示旋转方向。 $\mathbf{v}_2^{k_\varphi(\mathbf{x})}$ 根据式(2.31)计算球面线性插值， $k_\varphi(\mathbf{x}) \in (0, 1]$ 为插值权重。 $k_\varphi(\cdot)$ 是以 $\mathbf{c}_2$ 为中心， ρ_2 为宽度的 RBF 核函数。姿态四元数旋转 $\mathbf{v}_2^{k_\varphi(\mathbf{x})}$ 类似于局部加权平移变换中的 $k_\phi(\mathbf{x})\mathbf{v}_1$ 平移项。

3.4.2 迭代算法描述

本小节基于式(3.5)和式(3.10)迭代，目标是使得映射误差表达式(3.4)足够小，算法伪代码见算法1。其包含三个超参数 $\{\mu, \beta_1, \beta_2\}$ ，用于调节映射的空间变形程度。算法输出 $\rho, \mathbf{C}, \mathbf{V}$ 为局部加权平移和旋转变换的参数，并构成最终的位姿映射 Φ_x 和姿态映射 Φ_q 。为求解该参数，首先用 $\mathbf{A}$ 初始化变量 $\mathbf{Z} = \{\mathbf{Z}_x, \mathbf{Z}_q\} = \{\mathbf{z}_i, \hat{\mathbf{q}}_i\}_{i=1}^N$ ，并进行迭代直至映射误差(3.4)满足要求，每次迭代步骤如下：

Algorithm 1: 位姿的微分同胚映射迭代算法

输入: $A = \{\mathbf{x}_i, \mathbf{q}_i\}_{i=1}^N$, $B = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N$,

参数: $K \in \mathbb{N}^+$, $0 < \mu < 1$, $0 < \beta_1, \beta_2 \leq 1$,

输出: $\rho = \{\rho_{1,j}, \rho_{2,j}\}_{j=1}^K$, $C = \{\mathbf{c}_{1,j}, \mathbf{c}_{2,j}\}_{j=1}^K$, $V = \{\mathbf{v}_{1,j}, \mathbf{v}_{2,j}\}_{j=1}^K$,

1 初始化 $Z = \{Z_x, Z_q\} = \{z_i, \hat{q}_i\}_{i=1}^N$, $Z := A$,

while 式(3.4)中的映射误差足够小 **do**

2 $m := \arg \max_{i \in [1, \dots, N]} (\|z_i - \mathbf{x}'_i\|)$

3 $\mathbf{c}_{1,j} := z_m$

4 $\mathbf{v}_{1,j} := \beta_1(\mathbf{x}'_m - \mathbf{c}_{1,j})$

5 $\rho_{1,j} := \arg \min_{\rho_1 \in (0, \mu \rho_{\max}(\mathbf{v}_{1,j}))} (F_p(\phi(Z_x), B))$

6 $Z_x := \phi_j(Z_x)$

7 $n := \arg \max_{i \in [1, \dots, N]} (d(\hat{q}_i, \mathbf{q}'_i))$

8 $\mathbf{c}_{2,j} := z_n$

9 $\mathbf{v}_{2,j} := (\hat{q}_n * \bar{q}'_n)^{\beta_2}$

10 $\rho_{2,j} := \arg \min_{\rho_2 > 0} (F_q(\varphi(Z), B))$

11 $Z_q := \varphi_j(Z_q)$

12 **end**

- 1) 找出 $(z_i)_{i=1}^N$ 与 $(b_i)_{i=1}^N$ 之间距离最大的点, 计该点的索引为 m , 并保存式(3.5b)中的中心 $\mathbf{c}_{1,j}$ 和式(3.5a)中的方向 $\mathbf{v}_{1,j}$;
- 2) 在边界约束 $\rho_1 \in (0, \mu \rho_{\max}(\mathbf{v}_{1,j}))$ 的条件下求使得 $F_p(\phi(Z_x), B)$ 取最小值时的参数值 ρ_1 , 并保存 $\rho_{1,j}$;
- 3) 根据式(3.5)更新位置 Z_x ;
- 4) 找出姿态距离函数 $d(\hat{q}_i, \mathbf{q}'_i)$ 取最大值的点, 计该点的索引为 n , 并保存式(3.10b)中的旋转变换中心 $\mathbf{c}_{2,j}$;
- 5) 将姿态偏差 $\hat{q}_n * \bar{q}'_n$ 按权重 β_2 插值, 得 $\mathbf{v}_{2,j}$;
- 6) 在单边约束 $\rho_2 > 0$ 的条件下求使得 $F_q(\varphi(Z), B)$ 取最小值的 ρ_2 , 并保存为 $\rho_{2,j}$;
- 7) 根据式(3.10)更新位置 Z_q 。

在第 j 个迭代步中, 有更新式:

$$Z_x := \phi_j(Z_x), \quad (3.11a)$$

$$\mathbf{Z}_q := \varphi_j(\mathbf{Z}) = \mathbf{v}_{2,j}^{k_{\varphi,j}(\mathbf{Z}_x)} * \mathbf{Z}_q. \quad (3.11b)$$

若迭代 K 次, 最终的映射函数为:

$$\Phi_x(\mathbf{x}) = \phi_K \circ \cdots \circ \phi_2 \circ \phi_1(\mathbf{x}), \quad (3.12a)$$

$$\Phi_q(\mathbf{x}, \mathbf{q}) = \varphi_K \circ \cdots \circ \varphi_2 \circ \varphi_1(\mathbf{x}, \mathbf{q}) = \mathbf{g}(\mathbf{x}) * \mathbf{q}, \quad (3.12b)$$

其中, $\circ$ 表示函数复合, 姿态映射 Φ_q 的表达式基于四元数乘法结合律(2.22)进行了化简。因此, 正向映射与映射算法框架(3.1)–(3.2)形式相同。反向映射 Φ^{-1} 基于牛顿迭代法求非线性方程(3.12)的根。由于本节采用无限光滑(定义域内无穷阶数连续可导)的 RBF 函数作为核函数进行平移和旋转, 因此最终的位姿映射 Φ 也是无限光滑的, 从而映射函数 Φ 也是 C^∞ 光滑的微分同胚映射。

算法1中第 5 行和第 10 行需求解边界约束下的最小值问题, 本节采用序列二次规划方法 (Sequential Quadratic Programming, SQP)^[155], 优化计算过程快。可以通过给 ρ_1 和 ρ_2 添加上界, 以使得位置和姿态的映射变得更为平滑, 从而保证位姿映射的雅可比矩阵在合适的范围, 以满足约束4), 但可能会牺牲精度, 导致需要更多的迭代次数 K 以达到满足任务要求的映射误差。 β_1 和 β_2 分别用于调节位置和姿态迭代的步长, 步长过小可能导致迭代效率低下, 可随着迭代的过程不断增大 β_1 和 β_2 , 以提高效率。

3.5 基于可逆神经网络的映射算法

由于迭代算法存在一定的缺陷(将在仿真3.6和实验部分3.7具体分析), 本节提出基于可逆神经网络的方法实现式(3.1)和(3.2)的位置姿态映射, 求解出最优的映射函数 $\Phi_x(\mathbf{x})$ 和 $\Phi_q(\mathbf{x})$ 。本节基于标准化流模型构建位置的微分同胚映射 $\Phi_x(\mathbf{x})$, 并采用全连接神经网络来构建姿态映射中的比例项 $\mathbf{g}(\mathbf{x})$, 最后结合两个神经网络实现融合位置和姿态的映射算法。

3.5.1 位置映射网络

本节采用标准化流模型^[156]的基本思想, 将一系列双射(即可逆)变换组合起来, 构建观测空间和隐空间之间的映射关系。本小节通过构建示教数据 $\mathbf{B}$ 和 $\mathbf{A}$ 之间的映射, 并采用类似于式(3.12a)的 K 个变换的复合。

基于 Real NVP^[157]的结构, 本小节构建了 4 个耦合层, 且每一层为一个简单的仿射变换。考虑 D_0 维的位置输入 $\mathbf{x}$, 某一层的输入 $\mathbf{x}'$ 为:

$$\begin{cases} \mathbf{x}'_{1:d} = \mathbf{x}_{1:d}, \\ \mathbf{x}'_{d+1:D_0} = \mathbf{x}_{d+1:D} \odot e^{s(\mathbf{x}_{1:d})} + \mathbf{t}(\mathbf{x}_{1:d}), \end{cases} \quad (3.13)$$

其中 $\mathbf{s}$ 和 $\mathbf{t}$ 分别为缩放因子和平移变换因子，且为两个卷积网络， $1 \leq d \leq D_0 - 1$ 。 $\odot$ 表示对应元数相乘。其反向映射为：

$$\begin{cases} \mathbf{x}_{1:d} = \mathbf{x}'_{1:d}, \\ \mathbf{x}_{d+1:D_0} = (\mathbf{x}'_{d+1:D_0} - \mathbf{t}(\mathbf{x}_{1:d})) \odot e^{-\mathbf{s}(\mathbf{x}_{1:d})}. \end{cases} \quad (3.14)$$

正向映射（式(3.13)）的雅可比矩阵为：

$$\mathbf{J}(\mathbf{x}) = \frac{\partial \Phi_{\mathbf{x}}(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \frac{\partial \mathbf{x}'_{d+1:D_0}}{\partial \mathbf{x}_{1:d}} & \text{diag}(e^{\mathbf{s}(\mathbf{x}_{1:d})}) \end{bmatrix}, \quad (3.15)$$

其中， $\mathbf{J}(\mathbf{x}) \in \mathbb{R}^{D_0 \times D_0}$ 为下三角矩阵，且对角线元素均大于 0（ $\mathbf{I}_d$ 是 d 维的单位矩阵）。因此， $\forall \mathbf{x} \in \mathbb{R}^{D_0}$ ， $\det\left(\frac{\partial \Phi_{\mathbf{x}}(\mathbf{x})}{\partial \mathbf{x}}\right) > 0$ ，即雅可比矩阵 $\mathbf{J}(\mathbf{x})$ 可逆。由定理 2.1.1 可知，每一层变换都是可逆的。此外，每层变换都是无限光滑的，因此最终的位置映射也是无限光滑的，即最终位置映射为无限光滑的微分同胚映射，满足小节 3.3.1 中的约束要求。

为训练该神经网络，定义其代价函数：

$$Q_1 = F_p(\Phi_{\mathbf{x}}(\mathbf{A}), \mathbf{B}) + \lambda_1 \frac{1}{N_1} \sum_{i=1}^{N_1} |\mathbf{J}(\mathbf{x}_i^s) - \mathbf{I}_{D_0}|, \quad (3.16)$$

其中，第一项为基于式(3.4a)的位置误差，第二项为考虑约束 4) 的惩罚项。采样点 $\mathbf{x}_i^s \sim \mathcal{N}(\boldsymbol{\mu}_a, 3\boldsymbol{\Sigma}_a)$ 为 N_1 个基于高斯分布的采样，其中 $\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a$ 为 $\mathbf{A}$ 中位置 $\{\mathbf{x}_i\}_{i=1}^N$ 高斯分布均值和方差， $\lambda_1 > 0$ 为该惩罚项的权重。该惩罚项计算雅可比矩阵 $\mathbf{J}(\mathbf{x}_i^s)$ 和单位矩阵 $\mathbf{I}_D$ 之间的差值，从而在优化过程中保证位置映射的雅可比矩阵 $\mathbf{J}(\mathbf{x})$ 靠近单位矩阵，以使得速度的缩放矩阵靠近单位矩阵。

3.5.2 姿态映射网络

本小节设计全连接神经网络来计算式(3.2)中的 $\mathbf{g}(\mathbf{x})$ ，采用 4 层神经网络，每层 24 个单元，并用 $\tanh$ 作为激活函数，其函数式为：

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (3.17)$$

因此， $\tanh(x) \in (-1, 1)$ 。采用式(2.26)求四元数的对数，从而在笛卡尔空间训练姿态映射。该神经网络用于构建位置到姿态的映射，即 $\mathbf{x} \mapsto \mathbf{r}$ ，其中 $\mathbf{r} = \frac{2}{\pi} \log(\mathbf{q}' * \bar{\mathbf{q}})$ ，然后采用四元数指数恢复出四元数 $\mathbf{g}(\mathbf{x}) = \exp(\frac{\pi}{2}\mathbf{r})$ 。

基于映射约束 2)，设计用于映射训练的观测数据 $\{\mathbf{x}_i \mapsto \frac{2}{\pi} \log(\mathbf{q}'_i * \bar{\mathbf{q}}_i)\}_{i=1}^N$ 。姿态映射的代价函数为：

$$Q_2 = \frac{1}{N} \sum_{i=1}^N d(\exp(\mathbf{r}_i) * \mathbf{q}_i, \mathbf{q}'_i) + \lambda_2 \frac{1}{N_1} \sum_{j=1}^{N_1} F_q(\hat{\mathbf{g}}(\mathbf{x}_j^s), \mathbf{q}_I) + \lambda_3 \sum_n \|\mathbf{w}_n\|^2, \quad (3.18)$$

其中, 第一项为采用式(2.32)计算的四元数距离, $\mathbf{r}$ 为该神经网络的输出, 当 $\mathbf{x} = \mathbf{x}_i$ 时, $\mathbf{r} = \mathbf{r}_i$ 。第二项类似于式(3.16)中的惩罚项, 用于保持两个空间的姿态偏差向无旋转四元数 $\mathbf{q}_I$ 靠近。第三项为 L2 正则化项, 计算网络中权重的平方和, 用于避免网络过拟合。正实数 λ_2, λ_3 为代价函数后两项的权重。由于全连接层为线性映射, $\tanh$ 为无限光滑的函数, 因此最终的映射 $\mathbf{g}(\mathbf{x})$ 也是无限光滑的, 即无限阶连续可微, 满足小节3.3.1中的约束要求。

3.5.3 位姿映射网络

将上述的两个映射网络结合在一起, 定义联合代价函数:

$$Q = Q_1 + Q_2, \quad (3.19)$$

并用 Adam 优化器^[158] 进行训练。训练数据为 $\{\mathbf{x}_i \mapsto \mathbf{x}'_i, \frac{2}{\pi} \log(\mathbf{q}'_i * \bar{\mathbf{q}}_i)\}_{i=1}^N$, 用以表示网络输入输出的观测值。代价函数 Q 中, $\boldsymbol{\lambda} = \{\lambda_i\}_{i=1}^3$ 为基于经验选择的权重因子。参数 λ_1 用于平衡位置映射的精度和平滑性, 较大的 λ_1 将使得位置映射更为平缓, 从而满足约束4)。 λ_2 和 λ_3 共同决定的姿态映射的平滑性, 数值越大平滑性越好, 但精度越差。针对速度映射公式(3.3), 可以基于神经网络的梯度快速求得。

3.6 仿真对比与分析

本节中, 在二维仿真场景下对比3.4节的迭代映射算法和3.5节的神经网络法。如图3.2所示, 设定数据 $\mathbf{A}$ 为第一行第二列子图中的五个物体, 并设定物体的轮廓轨迹的采样点数量为 $N_s = 1, 10, 100$, $\mathbf{B}$ 为第二行和第三行彩色点/轮廓所示的位置和姿态, 其中 $\mathbf{B}$ 与 $\mathbf{A}$ 中的姿态存在较大的差异。本节分别用迭代映射算法和可逆神经网络法构建 $\mathbf{A}$ 和 $\mathbf{B}$ 之间的可逆映射。迭代算法中, 设定迭代参数为 $\mu = 0.6, \beta_1 = 0.95, \beta_2 = 0.9$ 。可逆神经网络算法训练中, 设定 $\boldsymbol{\lambda} = [0.002, 0.01, 0.3]$, $N_1 = 200$, $K = 4$ 。设定算法终止条件为式(3.4)中的位置误差和姿态误差分别小于 1 mm 和 0.025 rad。本章的仿真¹和实验均运行在 4 核心的 Intel i7-7700HQ CPU 上, 电脑内存为 20 GB。

二维度映射的结果如图3.2所示, 第一行为原空间 $\mathbf{A}$ 中的物体, 并绘制由等间距的网格线, 第二行和第三行为被映射算法扭曲的空间, 其中灰色网格线的位置和姿态由正向映射计算得到。根据式(3.2)可知, 只需关注两侧的姿态偏差, 即 $\mathbf{g}(\mathbf{x})$, 因此, 为更清晰地展示姿态映射, 设定原空间中数据 $\mathbf{A}$ 的姿态均为无旋转四元数 $\mathbf{q}_I$ 。物体或网格点的姿态均用箭头来表示其 x 轴方向。从图中可以看出, 两个映射算法都将网格线平滑地拉扯变形了。贴近于数据 $\mathbf{A}$ 的网格线, 在变形空间中也贴近数据 $\mathbf{B}$, 且其姿态

¹代码链接<https://github.com/happygaoxiao/Diffeormorphism>

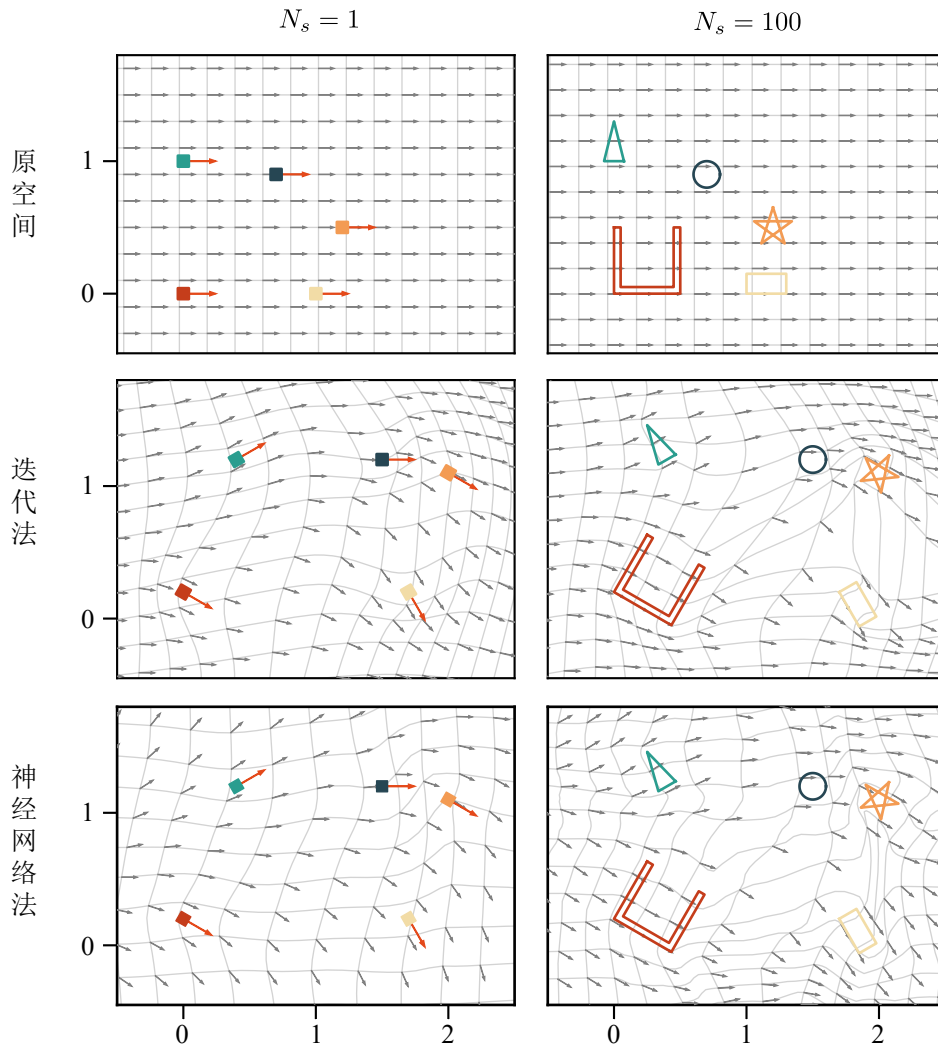


图 3.2 二维场景下迭代法和神经网络法映射仿真结果

方向保持几乎一致。从直观上可以看出两个映射算法的平滑性较好，且映射误差在允许的范围內。

表格3.1了两个算法的训练时间和正向、反向、速度计算所需的时间，可逆神经网络法简记为 iNN。相比于神经网络法，迭代法训练耗时很短，但映射耗时较长，且随着采样点的数量增加而增大，主要原因为迭代点数量增加导致达到相同精度所需的迭代次数 K 增加，从而使得需要更多的函数复合来构造位姿映射，较长的映射输出时间可能会导致机械臂控制频率降低甚至机械臂失稳。在可逆神经网络算法中，尽管需要较长的时间用于训练，其映射计算的速度非常快，实时性好，便于机械臂的实时控制，且映射效率不受采样点数量的影响。

表 3.1 迭代法和可逆神经网络法映射仿真时间对比

	训练耗时 (s)			正向映射 (ms)			反向映射 (ms)			速度映射 (ms)		
N_s	1	10	100	1	10	100	1	10	100	1	10	100
迭代	0.051	0.61	9.5	0.44	4.4	28	1.2	9.7	68	5.1	45	320
iNN	1500	2500	4600	0.54	0.58	0.59	0.55	0.56	0.58	1.0	1.1	1.0

3.7 实验研究

为验证本章提出的映射算法的有效性, 本节设计了双边遥操作机械臂实验。以遥操作中的本地和远端操作空间中被操作对象的位置和姿态建立微分同胚映射, 实现连续双边遥操作, 用于验证所提出的迭代算法和可逆神经网络映射均为光滑的微分同胚映射, 并对比其性能。

3.7.1 实验平台

本节中设计两个实验平台, 分别为遥操作物体收集实验平台(图3.3)和遥操作拧阀门实验平台(图3.4)。对于每个实验平台, 左侧空间被视为本地空间, 右侧为远端空间, 中间用木板隔开以阻止操作员直接观察远端, 从而模拟现实场景下的遥操作, 该场景与图3.1相对应。采用一个操作杆的按钮来同时控制两个机械臂的夹爪开合。桌面上固定了一个 Intel Realsense D435 相机, 并采用 Aruco 标记码²来定位两个黑色方块, 从而实时获取各个待操作物体的位姿。其中, 物体相对于 Aruco 标记码之间的位姿变换, 以及相机和两个机械臂之间的变换, 已事先标定好。控制系统基于 ROS 搭建。左侧的机械臂由操纵员动觉拖动示教, 右侧的机械臂为远端机械臂。实验中采用两个 Franka Emika 公司的 Panda 机械臂, 均为基于力矩控制的 7 自由度机械臂。

物体收集实验平台中, 本地空间有四个不同颜色的小球和一个圆柱桶, 计其相对于本地机械臂基坐标系的位置为 $\mathbf{A}$, 右侧也有对应的小球和一个圆柱桶, 计其相对于远端机械臂基坐标系的位置为 $\mathbf{B}$ 。采用章节3.4中的迭代算法构建映射函数 $\Phi: \mathbf{A} \rightarrow \mathbf{B}$, 目标是当拖动本地机械臂连续地收集四个小球至圆柱桶内, 远端机械臂也能收集远端空间中的小球至圆柱桶。

在拧阀门遥操作实验平台中, 两侧空间各有四个阀门, 阀门按椭圆标记的颜色一一对应, 固定于黑色方块上, 设阀门相对机械臂基座的位姿为 $\mathbf{A}$ 和 $\mathbf{B}$, 注意其位置和姿态存在较大差异。将 $\mathbf{A}$ 和 $\mathbf{B}$ 其用于求解章节3.4和章节3.5的映射函数, 目标是通过拖动本地机械臂实现连续性地拧动远端阀门。

²源码链接https://github.com/pal-robotics/aruco_ros

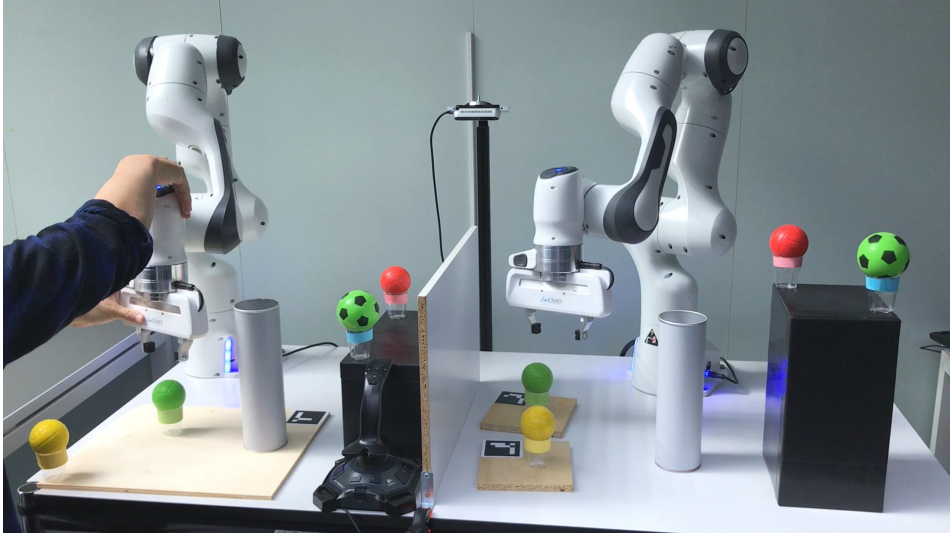


图 3.3 物体收集实验平台

3.7.2 阻抗控制

两个机械臂的控制算法均采用阻抗控制器^[154]，从而操作员可以通过示教本地机械臂来控制远端机械臂运动，并获取远端机械臂与环境之间的交互触觉感知，笛卡尔空间阻抗控制器的控制率为：

$$\begin{aligned} \tau &= \hat{M}(\theta)\ddot{\theta} + \hat{c}(\theta, \dot{\theta}) + \hat{g}(\theta) + J(\theta)^T f, \\ \text{其中 } f &= \begin{bmatrix} K_p(x_d - x) + K_v(\dot{x}_d - \dot{x}) \\ K_{pr} \log(q_d * \bar{q}) + K_{vr}(\omega_d - \omega) \end{bmatrix}, \end{aligned} \quad (3.20)$$

其中， $\hat{(\cdot)}$ 表示对式(2.37)中动力学模型参数的估计值。 $x, \dot{x} \in \mathbb{R}^3$ 为机械臂末端执行器的位置和速度， $q \in \mathcal{S}^3$ 为末端姿态四元数， $w \in \mathbb{R}^3$ 为角速度，由正运动学计算得出，速度项由正运动学对时间的导数 $[\dot{x}^T, \omega^T]^T = J(\theta)\dot{\theta}$ 计算。 $(\cdot)_d$ 表示期望值（即参考轨迹）。 $K_{(\cdot)} \in \mathbb{R}^{3 \times 3}$ 表示阻抗控制器的刚度和阻尼矩阵，均为对角矩阵，且对角线元素均为正。若忽略建模误差，将式(3.20)带入(2.37)得系统的闭环系统方程：

$$\begin{bmatrix} K_p(x_d - x) + K_v(\dot{x}_d - \dot{x}) \\ K_{pr} \log(q_d * \bar{q}) + K_{vr}(\omega_d - \omega) \end{bmatrix} = f_{ext}, \quad (3.21)$$

即机械臂与外界交互关系为弹簧-阻尼系统模型，若无外力将收敛至期望轨迹。该控制率与式(2.41)之间的不同在于，这里考虑参考姿态和当前姿态的最短距离（四元数的测地线距离），相较于欧拉角直接作差效果更好。且此处直接对关节惯量进行补偿，因此闭环系统无质量项。

针对本节的遥操作场景，某个机械臂的参考轨迹为另一个机械臂末端位姿经微分

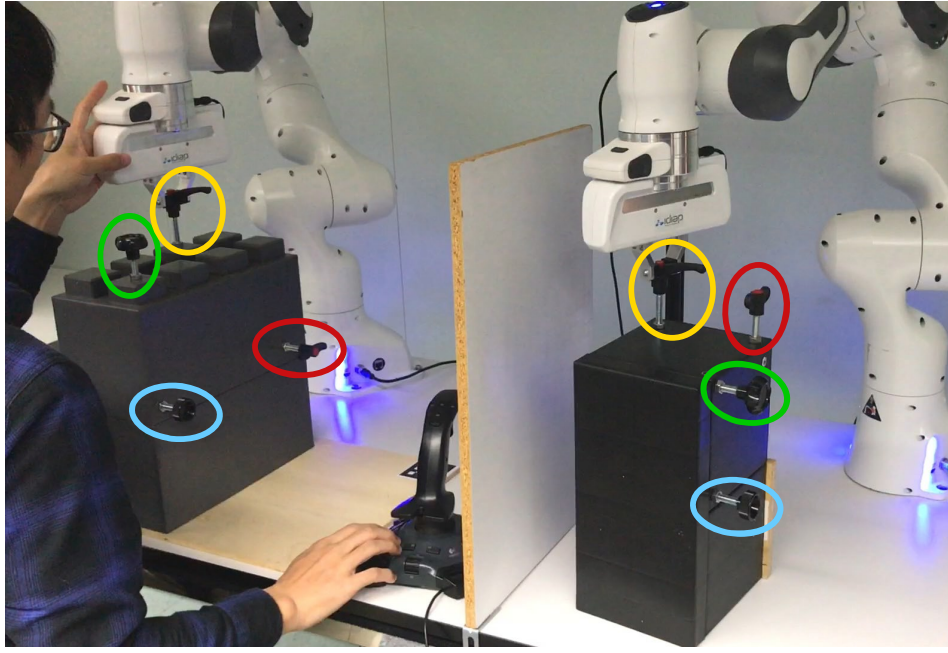


图 3.4 拧阀门遥操作实验平台

同胚映射的输出，即按式(3.1)–(3.2)：

$$\begin{aligned} \mathbf{x}_d^r &= \Phi_x(\mathbf{x}^l), & \mathbf{q}_d^r &= \mathbf{g}(\mathbf{x}^l) * \mathbf{q}^l, \\ \mathbf{x}_d^l &= \Phi_x^{-1}(\mathbf{x}^r), & \mathbf{q}_d^l &= \bar{\mathbf{g}}(\Phi_x^{-1}(\mathbf{x}^r)) * \mathbf{q}^r, \end{aligned} \quad (3.22)$$

其中， $(\cdot)^l$ ， $(\cdot)^r$ 分别表示本地 (local/left) 机械臂和远端 (remote/right) 机械臂末端夹爪的位置/姿态。在物体收集实验中，球体的姿态对抓取操作没有影响，因此这里取 $\mathbf{g}(\mathbf{x}^l) = \mathbf{q}_l$ 为无旋转四元数。式(3.20)中，远端机械臂的阻抗增益参数设置为 $\mathbf{K}_p = 600\mathbf{I}_3$ ， $\mathbf{K}_v = 20\mathbf{I}_3$ ， $\mathbf{K}_{pr} = 30\mathbf{I}_3$ ， $\mathbf{K}_{vr} = 2\mathbf{I}_3$ ，本地机械臂的参数设置为该参数的 0.1 倍，以便于操作人员可以轻松地拖动机械臂，并减小所感知的远端交互力。电机控制指令发送频率为 1 kHz，并采用低通滤波器以平滑控制力矩。参考轨迹更新频率为 100 Hz。迭代法和可逆神经网络法的模型参数与3.6节仿真的参数相同，其中位置数据由二维变为三维，映射算法中，取 $N_s = 1$ 。

3.7.3 实验结果分析

(1) 物体收集实验

实验过程如图3.5所示，依次通过遥操作收集绿色、黄色、花色和红色小球至圆柱形容容器中，实验视频见³。针对每一个彩色小球，操作人员拖动本地端的机械臂，使得其末二指夹爪在合适的抓取位置，并按下中间的操作杆上的按钮，以发送夹爪闭合指令。

³实验视频<https://www.bilibili.com/video/BV1ey4y1s7eB/>

二指夹爪闭合并抓取小球后，操作员拖动机械臂并携带移动至圆柱形容器的口，再次按下按钮释放二指夹爪，从而将小球放入容器中。该过程中本地端和远端的机械臂运动轨迹如图3.6所示，黑色曲线为机械臂末端夹爪在其基坐标系下的轨迹线，彩色原点代表对应的待抓取和收集的小球。可以看出，机械臂运行轨迹平滑，在微分同胚映射算法的帮助下，操作员可以专注于本地端的遥操作任务，远端也会同步地执行相应的任务，即便小球的位置存在差异。该实验证明了微分同胚映射算法可以用于机械臂的实时轨迹规划。

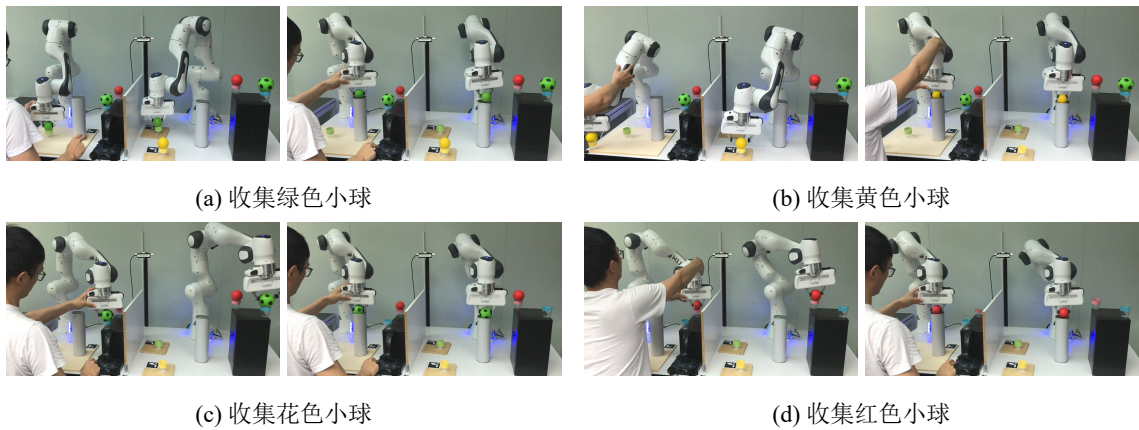


图 3.5 遥操作收集小球实验过程

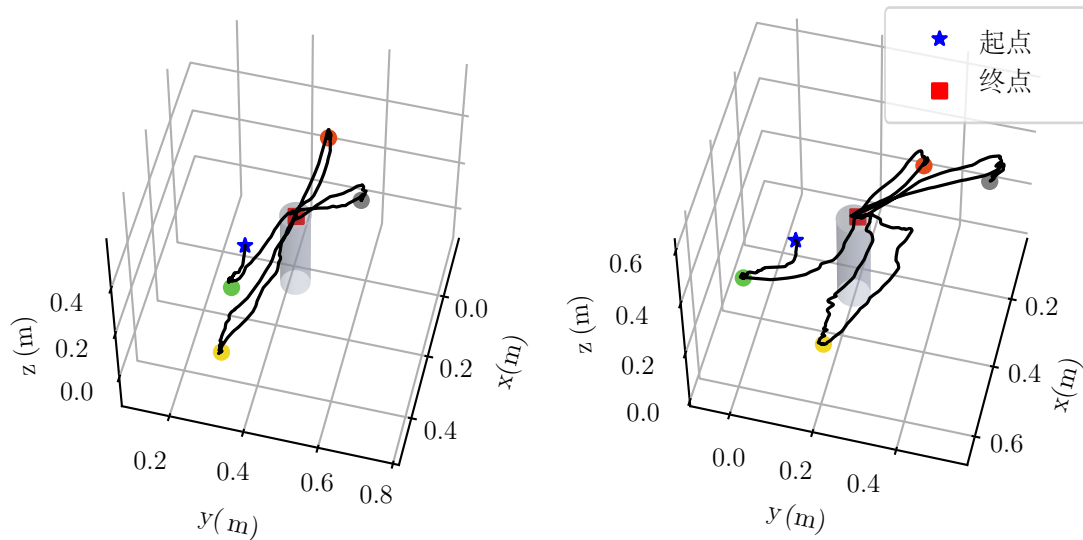


图 3.6 收集小球过程中机械臂的运动轨迹

(2) 拧阀门实验

本实验中，双边遥操作的目标是按顺序顺时针旋转四个阀门 90° 。某次实验过程如

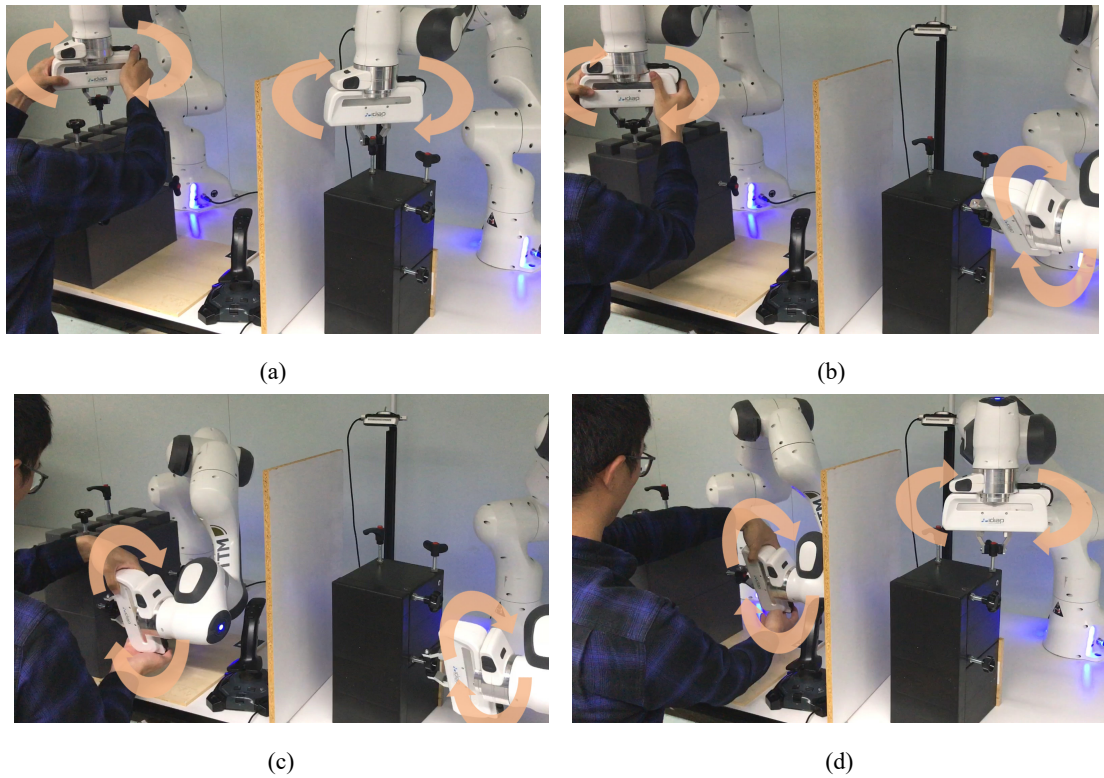


图 3.7 遥操作拧阀门实验截图

图3.7所示，完整的遥操作视频见⁴。在本地端（每幅图的左侧），操作员旋转阀门，远端的机械臂同时执行相同的任务。采用迭代法建立运动映射的遥操作实验中本地机械臂和远端机械臂的运动轨迹如图3.8(a)所示，图3.8(b)为采用神可逆神经网络法实验中的机械臂轨迹图。本地空间中采用三维等间距网格线，经微分同胚映射算法得到远端空间中扭曲的网格线。尽管两侧空间中阀门的位姿不相同 ($\mathbf{A} \neq \mathbf{B}$)，两种映射算法都能满足约束条件1)–2)，即本地机械臂在拧动某个阀门的时候，远端机械臂也正操作对应的阀门，且算法生成的轨迹平滑，机械臂可以平稳执行。

表 3.2 三种映射算法结果对比

	迭代法	可逆神经网络	直接法
任务耗时 (s)	101 ± 18	96 ± 16	195 ± 26.5
交互力 (N)	1.12 ± 0.21	0.81 ± 0.14	1.05 ± 0.15

在本节的实验中，采用了第三种运动映射算法，基于直接位置映射^[159]的方法作为本章提出的两种算法的测试基准。采用直接位置映射的实验中，将图3.4中的竖直方

⁴实验视频<https://www.bilibili.com/video/BV1ZU4y1a7vX>

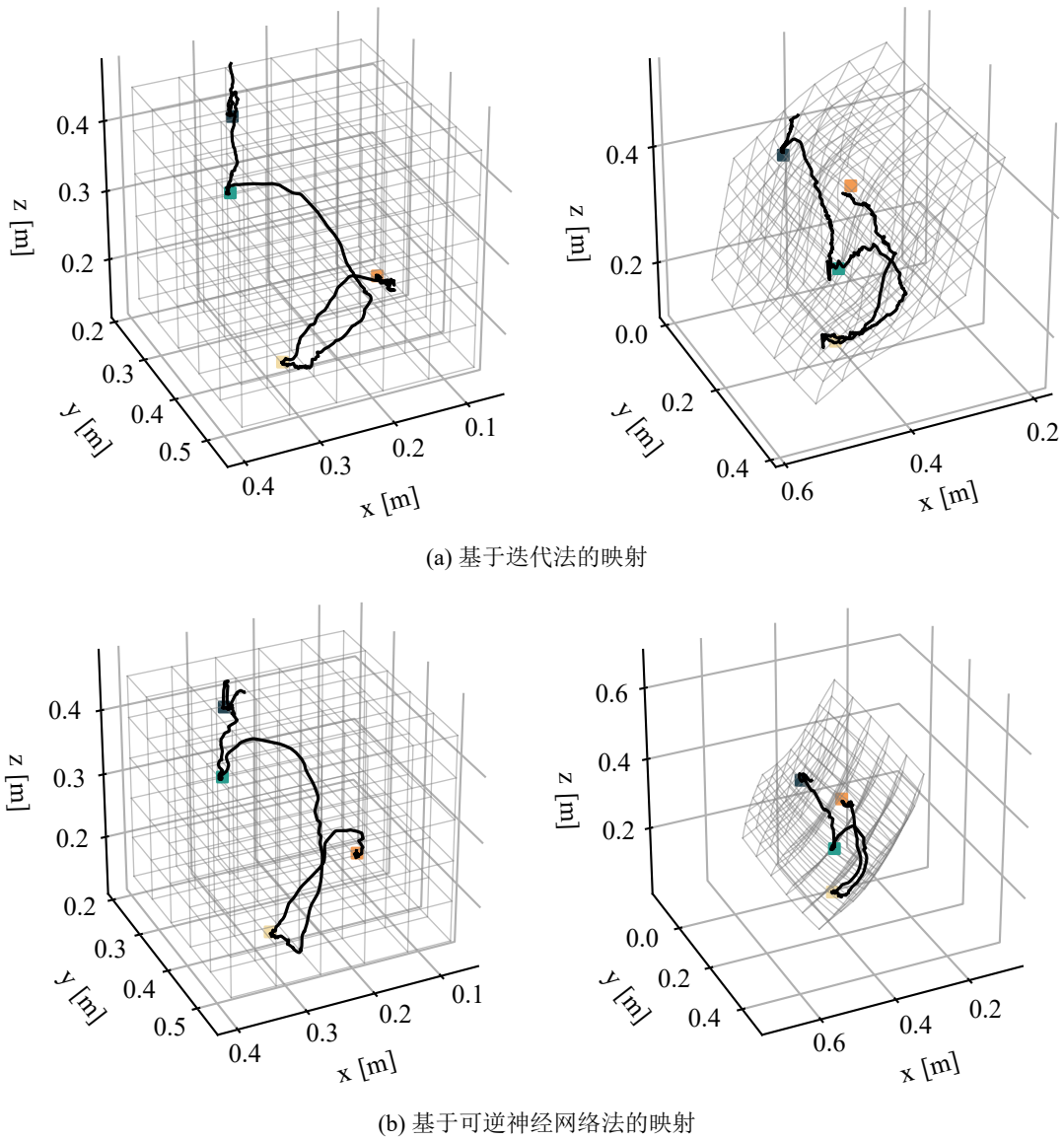


图 3.8 工作空间映射以及机械臂轨迹

向的隔板移除，示教者直接观看远端工作空间，从而模拟基于视频流反馈的遥操作。

三种映射算法各执行 10 次，表 3.2 对比了三种方法的实验结果。采用迭代法拧四个阀门耗费时间稍多于神经网络法，且交互力较大，主要原因可能是其正向、反向和速度映射耗时较长 (见表 3.1)，导致实时的参考位姿更新频率降低，操作员需要较大的力来拖动本地机械臂。迭代法和可逆神经网络法性能都优于直接法。直接法受限于人/机械臂的运动空间限制，需要频繁暂停遥操作通信，移动本地机械臂至合适的位置再重新进行任务，因此耗费时间较长。相较于迭代法，可逆神经网络法由于需要较长的时间用于模型训练，难以快速迁移至新的场景。

3.8 本章小结

本章分析了采用微分同胚映射构建 DS 的基本流程，并分析了位置和姿态的映射算法的约束条件，提出了一种融合位置和姿态的微分同胚映射算法框架。在该框架的基础上，设计了两种实现方式：基于径向基函数的迭代算法，以及基于可逆神经网络的方法。为验证算法的有效性，首先设计了二维仿真场景，用以展示映射算法对空间的平滑扭曲效果。仿真结果表明，两种方法都能成功实现平滑的位置和姿态映射。迭代法训练耗时较短，但其位姿正向、反向和速度映射耗时较长（几十毫秒），在实际机械臂执行时可能影响实时性。神经网络法训练所需时间长（1500 s 至 4600 s），但其正向反向映射效率极高（2ms），可以满足机械臂实时控制需求。最后，本章设计了双边遥操作场景，将两种映射算法用于建立遥操作中的本地空间和远端空间之间的映射。当两侧空间中被操作对象的位置和姿态有差异时，两种映射算法都能实现连续的双边遥操作，且性能均优于传统的直接位置映射法。实验证明两种微分同胚映射算法均能实现在线的机械臂运动规划，且生成的轨迹较为平滑。

4 基于模仿学习和微分同胚映射的运动规划

4.1 引言

如表1.1所示, 现有的时不变 DS 方法一般仅研究位置的 DS 及其收敛性。在基于模仿学习的运动规划方法中, 同时考虑位置和姿态的规划一般采用行为克隆的方法, 基于示教数据建立时间到位姿的映射概率模型。本章提出一种基于模仿学习和微分同胚映射的 DS 设计框架, 并证明其稳定性。该 DS 为时不变系统, 具有内在的对时空扰动的适应性, 此外该方法为基于位姿反馈的规划方法, 可以适应动态变化的场景。基于该框架, 本文结合第3章的映射算法, 设计了 DS 的具体形式。最后, 在 2D 仿真中进行了运动规划的测试, 并在 Panda 机械臂上进行了自由空间的运动规划实验, 验证该运动规划算法对动态环境的适应性和稳定性。

4.2 运动规划算法及其稳定性

本节分析了基于模仿学习和微分同胚映射构建运动规划算法的基本思路, 并证明其稳定性。

4.2.1 构建方法

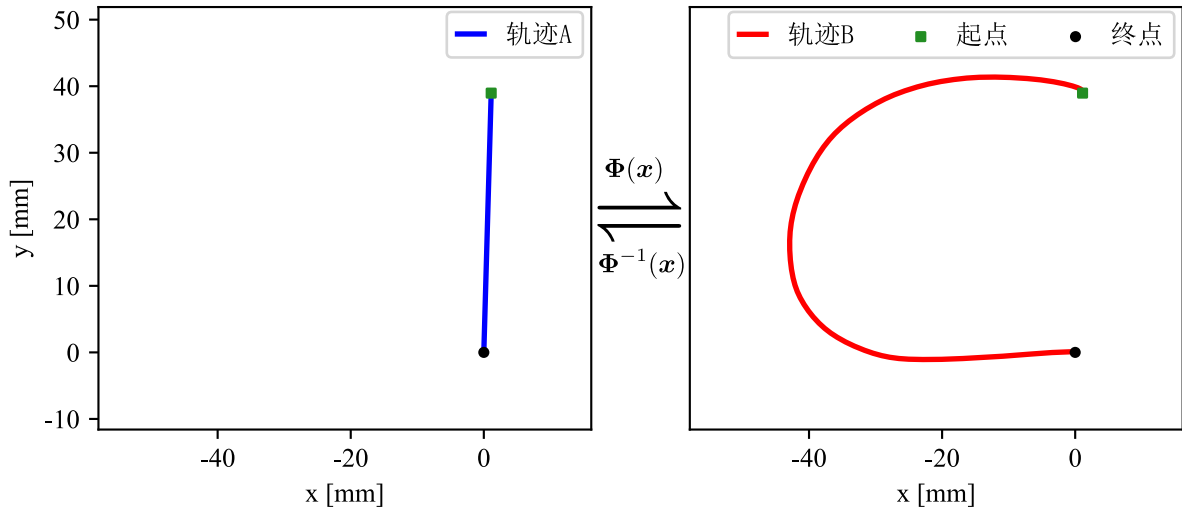


图 4.1 基于微分同胚映射的初始表达

如图4.1所示, 设有初始示教数据 $B = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N$, 且已将数据进行如章节3.2中的归零化处理, 即 $\mathbf{x}'_N = \mathbf{0}, \mathbf{q}'_N = \mathbf{q}_I$ 。示教数据 B 的起点为绿色方块, 终点为黑色原点。取轨迹 B 的起点和终点位姿, 并经线性插值得到轨迹 A , 其中:

$$\mathbf{A} = \{\mathbf{x}_i, \mathbf{q}_i\}_{i=1}^N = \{t\mathbf{x}_1, \mathbf{q}_1^t\}, t = \frac{N-i}{N-1} (i = 1, 2, \dots, N). \quad (4.1)$$

由章节3.1中的微分同胚映射表达, 以 $\mathbf{A} \mapsto \mathbf{B}$ 作为映射的观测值, 用于训练两个空间的映射函数 $\Phi(\mathbf{x})$ 。若左侧空间 $\mathbf{A}$ 中存在一个全局渐进稳定的 DS 收敛于零点, 则以空间内任意一点为起点, 其轨迹线最终将收束到零点。若在空间 $\mathbf{B}$ 内任取一点为起始点 $\mathbf{x}$, 在空间 $\mathbf{A}$ 中有 $\Phi^{-1}(\mathbf{x})$ 与其对应, 且空间 $\mathbf{A}$ 中 $\Phi^{-1}(\mathbf{x})$ 的轨迹线将收束至零点, 则将该轨迹线映射至空间 $\mathbf{B}$ 中, 可作为空间 $\mathbf{B}$ 中起始点 $\mathbf{x}$ 的轨迹线, 其轨迹也是收束至零点的。接下来将给出严格的稳定性证明。

4.2.2 稳定性分析

定义 4.2.1 设有 DS $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 和 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$, 若存在一个微分同胚映射 $\Phi: \mathbb{R}^D \mapsto \mathbb{R}^D$, 使得

$$\forall \mathbf{x} \in \mathbb{R}^D, \mathbf{f}_2(\Phi(\mathbf{x})) = \mathbf{J}_\Phi(\mathbf{x})\mathbf{f}_1(\mathbf{x}), \quad (4.2)$$

则这两个 DS 被称为微分同胚的, 其中 $\mathbf{J}_\Phi(\mathbf{x}) = \frac{\partial \Phi(\mathbf{x})}{\partial \mathbf{x}}$ 。

因此, 若 DS $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 和 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 满足定义4.2.1, 则有:

$$\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x}) = \mathbf{J}_\Phi(\Phi^{-1}(\mathbf{x}))\mathbf{f}_1(\Phi^{-1}(\mathbf{x})). \quad (4.3)$$

定理 4.2.1 若两个 DS $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 和 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 是微分同胚的, 其映射函数为 Φ 。若 $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 是全局渐进稳定的, 平衡点为 $\mathbf{x}^*$, 则 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 也是全局渐进稳定的, 其平衡点为 $\tilde{\mathbf{x}}^* = \Phi(\mathbf{x}^*)$ 。

证明 设 DS $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 的 Lyapunov 函数为 $\mathcal{V}(\mathbf{x})$, 则其满足定理2.2.2。针对 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$, 取候选 Lyapunov 函数 $\mathcal{V}_2(\mathbf{x}) = \mathcal{V}(\Phi^{-1}(\mathbf{x}))$, 两边对时间求导数, 则有:

$$\begin{aligned} \dot{\mathcal{V}}_2(\mathbf{x}) &= \left. \frac{\partial \mathcal{V}(\mathbf{y})}{\partial \mathbf{y}} \right|_{\mathbf{y}=\Phi^{-1}(\mathbf{x})} \frac{\partial \Phi^{-1}(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} \\ &= \left. \frac{\partial \mathcal{V}(\mathbf{y})}{\partial \mathbf{y}} \right|_{\mathbf{y}=\Phi^{-1}(\mathbf{x})} [\mathbf{J}_\Phi(\Phi^{-1}(\mathbf{x}))]^{-1} \mathbf{f}_2(\mathbf{x}) \\ &= \left. \frac{\partial \mathcal{V}(\mathbf{y})}{\partial \mathbf{y}} \right|_{\mathbf{y}=\Phi^{-1}(\mathbf{x})} [\mathbf{J}_\Phi(\Phi^{-1}(\mathbf{x}))]^{-1} \mathbf{J}_\Phi(\Phi^{-1}(\mathbf{x}))\mathbf{f}_1(\Phi^{-1}(\mathbf{x})) \\ &= \left. \frac{\partial \mathcal{V}(\mathbf{y})}{\partial \mathbf{y}} \right|_{\mathbf{y}=\Phi^{-1}(\mathbf{x})} \mathbf{f}_1(\Phi^{-1}(\mathbf{x})) = \left. \frac{\partial \mathcal{V}(\mathbf{y})}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial t} \right|_{\mathbf{y}=\Phi^{-1}(\mathbf{x})} \\ &= \dot{\mathcal{V}}(\mathbf{y}) = \dot{\mathcal{V}}(\Phi^{-1}(\mathbf{x})). \end{aligned} \quad (4.4)$$

当 $\Phi^{-1}(\mathbf{x}) = \mathbf{x}^*$ 时, 即 $\mathbf{x} = \Phi(\mathbf{x}^*)$, 由定理2.2.2可知:

$$\begin{cases} \mathcal{V}(\Phi^{-1}(\mathbf{x})) = 0 \\ \dot{\mathcal{V}}(\Phi^{-1}(\mathbf{x})) = 0 \end{cases} \implies \begin{cases} \mathcal{V}_2(\mathbf{x}) = 0 \\ \dot{\mathcal{V}}_2(\mathbf{x}) = 0 \end{cases}. \quad (4.5)$$

当 $\Phi^{-1}(\mathbf{x}) \neq \mathbf{x}^*$ 时, 即 $\mathbf{x} \neq \Phi(\mathbf{x}^*)$, 则有:

$$\begin{cases} \mathcal{V}(\Phi^{-1}(\mathbf{x})) > 0 \\ \dot{\mathcal{V}}(\Phi^{-1}(\mathbf{x})) < 0 \end{cases} \implies \begin{cases} \mathcal{V}_2(\mathbf{x}) > 0 \\ \dot{\mathcal{V}}_2(\mathbf{x}) < 0 \end{cases}. \quad (4.6)$$

因此, $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 是全局渐近稳定的, 且平衡点为 $\tilde{\mathbf{x}}^* = \Phi(\mathbf{x}^*)$ 。若 $\text{DS}\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 的 Lyapunov 函数为 $\mathcal{V}(\mathbf{x})$, 则 $\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 的 Lyapunov 函数可取 $\mathcal{V}_2(\mathbf{x}) = \mathcal{V}(\Phi^{-1}(\mathbf{x}))$ 。□

4.3 融合位置和姿态的运动规划

本节首先提出基于微分同胚映射, 且只有考虑位置的 DS 模型设计方法, 用该 DS 模型来编码人的示教轨迹数据。随后提出了融合位置和姿态的 DS 构建方法。

4.3.1 位置的动态系统

首先分析二维空间 ($D = 2$) 中的 DS 构建方法。如小节4.2.1所述, 设有人工示教数据 $\mathbf{B} = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N$, 其轨迹 B 如图4.2(b)所示红色曲线。根据章节3.2中对 DS 的约束条件1)–2), 本节的目的是求解出空间 B 中全局渐进稳定的 DS (本小节只考虑位置), 且该系统尽可能精确地表征示教轨迹 B。

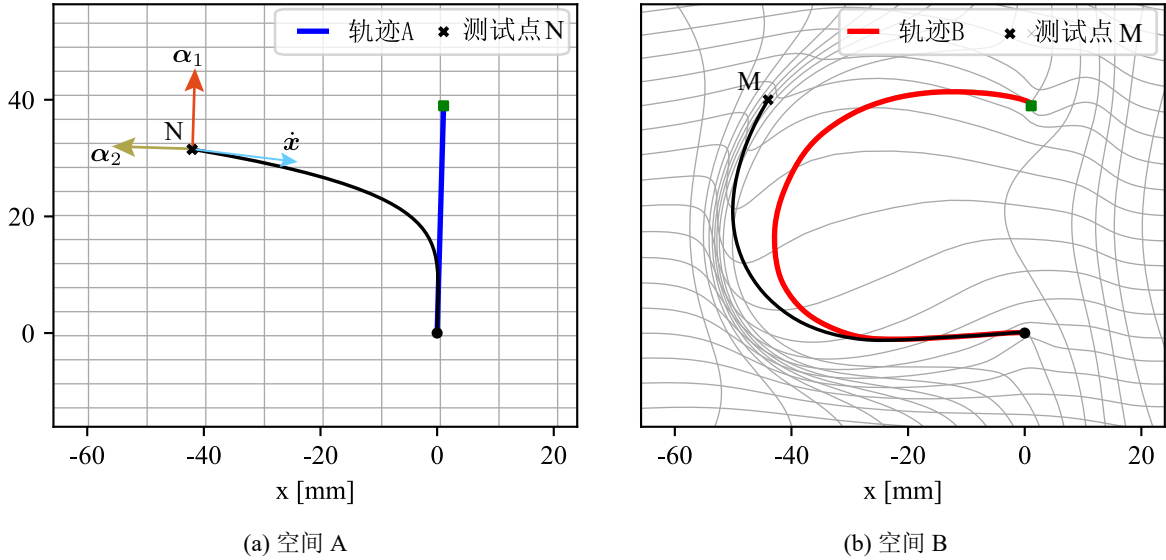


图 4.2 二维空间中微分同胚映射

由式(4.1)可基于示教数据 $\mathbf{B}$ 生成 $\mathbf{A} = \{\mathbf{x}_i, \mathbf{q}_i\}_{i=1}^N$, 得到空间 A 中蓝色轨迹线 A (图4.2(a)), 设轨迹 B 总耗时为 T 。基于章节3.1中的微分同胚映射算法框架, 以式(3.4a)为代价函数, 求解出最优的映射函数 $\Phi(\mathbf{x})$, 使得位置误差保持在允许的范围。例如, 可采用章节3.4的迭代算法或章节3.5的可逆神经网络法构造 $\Phi(\mathbf{x})$ 。图4.2中, 空间 A 中绘

制了等间距的网格线,该网格线在映射函数 $\Phi(\mathbf{x})$ 的作用下可得空间 B 中的网格线,其被平滑地扭曲变形,且轨迹 A 被映射至轨迹 B ,且逆映射也成立。

设空间 A 中有任意一测试点 N ,其位置为 $\mathbf{x}$ 。 N 点在空间 B 中的映射点为 M ,其位置为 $\Phi(\mathbf{x})$ 。本节期望测试点 M 的运动轨迹收敛至零点(图中黑色原点),且其应贴近示教轨迹 B 。在可逆映射 $\Phi(\mathbf{x})$ 作用下,即期望空间 A 中,测试点 N 的运动轨迹收敛至零点,且尽可能贴近轨迹 A 。为此,本节首先在空间 A 中设计全局渐进稳定的 DS,然后基于式(4.3)求解空间 B 中满足要求的 DS。

轨迹 A 起点为 $\mathbf{x}_1$, 终点为 $\mathbf{0}$, 取如下两个特征向量:

$$\begin{cases} \alpha_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}, \\ \alpha_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}, \end{cases} \quad (4.7)$$

其中, α_1 与向量 $\mathbf{x}_1$ 平行, α_2 由 α_1 逆时针旋转 90° 得到。设矩阵 P 有特征向量 α_1 和 α_2 , 对应的特征值分别为 λ_1, λ_2 。构建矩阵 $R = [\alpha_1 \ \alpha_2]$, 则有:

$$P = R \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R^{-1}, \quad (4.8)$$

其中, R 为正交矩阵, 因此有 $R^{-1} = R^T$, 因此 $P^T = P$, P 为对称矩阵。对空间 A , 构建如下 DS:

$$\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x}) = \gamma(\mathbf{x})P\mathbf{x}, \gamma(\mathbf{x}) > 0. \quad (4.9)$$

由式(4.3), 可得空间 B 中位置的 DS:

$$\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x}) = J_\Phi(\Phi^{-1}(\mathbf{x}))\mathbf{f}_1(\Phi^{-1}(\mathbf{x})) = \gamma(\Phi^{-1}(\mathbf{x}))J_\Phi(\Phi^{-1}(\mathbf{x}))P\Phi^{-1}(\mathbf{x}). \quad (4.10)$$

针对三维空间 ($D = 3$), 可采用类似的构造方法。取旋转轴 ϕ^* 为:

$$\phi^* = [0, 0, 1]^T \times \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}, \quad (4.11)$$

旋转角度为 $\theta^* = \arccos([0, 0, 1] \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|})$, 该旋转将三维空间中 z 轴旋转至与 $\mathbf{x}_1$ 向量共线, 从而将其等价为二维问题。该旋转的四元数为 $\mathbf{q}^* = [\cos(\frac{\theta^*}{2}), \sin(\frac{\theta^*}{2})\phi^*]^T$, 根据式(2.25)将该四元数转化为旋转矩阵 R , 可取 P 矩阵为:

$$P = R \begin{bmatrix} \lambda_2 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix} R^{-1}. \quad (4.12)$$

同样可构建式(4.9)中的 DS, 三维空间中的 DS, 与二维空间中的推导相同。

4.3.2 位置和姿态的动态系统

如图4.3(c)所示, 仍有空间 B 中原始示教数据 $B = \{x'_i, q'_i\}_{i=1}^N$ 和空间 A 中经过线性插值的轨迹 (含姿态) $A = \{x_i, q_i\}_{i=1}^N$, 并在图4.1的基础上构造空间 D (图4.3(a)) 中轨迹 $D = \{x_i, q_I\}_{i=1}^N$ 。其中, 位置坐标与 A 相同, 姿态四元数全为无旋转四元数 (水平向右的箭头), 图中箭头指代该点的姿态的 x 轴方向。

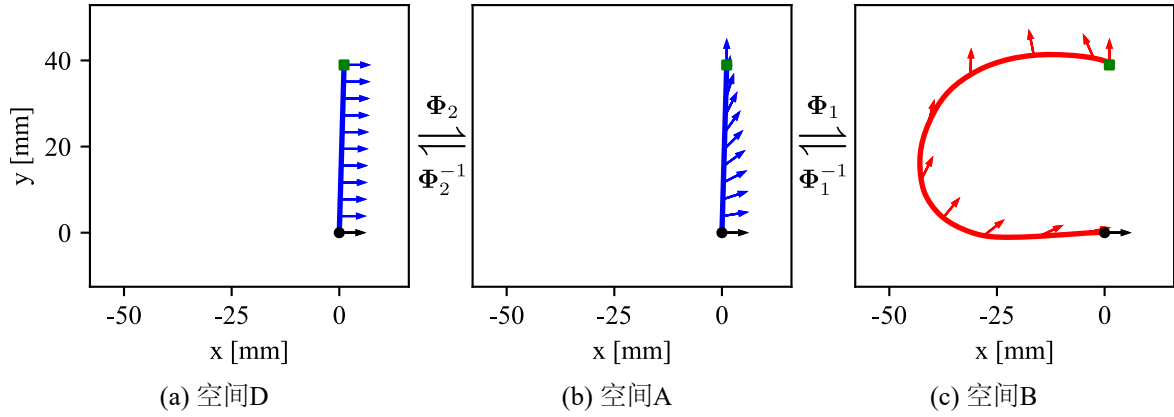


图 4.3 融合和位置和姿态的映射

首先, 构建基于式(3.1)–(3.2)的微分同胚映射 Φ_1 , 使得轨迹 A 映射至轨迹 B, 并保证其映射误差 (式(3.4)) 满足要求, 即空间 A 至空间 B 的映射:

$$\Phi_1 : \begin{cases} x' = h_1(x), \\ q' = g_1(x) * q, \end{cases} \quad (4.13)$$

其中 (x, q) 为空间 A 中任意一点, (x', q') 为其在空间 B 中的映射对应点。

同样地, 构建微分同胚映射 Φ_2 , 将轨迹 D 映射至空间 A 中的轨迹 A, 有:

$$\Phi_2 : \begin{cases} x = h_2(x_D) = x_D, \\ q = g_2(x_D) * q_I = g_2(x), \end{cases} \quad (4.14)$$

其中 (x_D, q_I) 为空间 A 中点 (x, q) 在逆映射 Φ_2^{-1} 作用下, 在空间 D 中的对应点, $g_2(x) \in S^3$ 为单位四元数, 且为位置 x 的函数。由于轨迹 D 和轨迹 A 中位置完全相同, 因此可直接取映射函数 $h_2(x_D) = x_D$ 。

对于位置的 DS, 可同样地在空间 A 中使用式(4.9)构建。

针对空间 A 中姿态的 DS, 设计如下函数:

$$\omega = \gamma_2(q)\beta \log(q * \bar{g}_2(x)) + \omega_r, \gamma_2(q) > 0, \beta < 0 \quad (4.15a)$$

$$\text{且有, } \omega_r = -2q * \frac{\partial \bar{g}_2(x)}{\partial x} \dot{x} * g_2(x) * \bar{q} \quad (4.15b)$$

其中 $\omega \in \mathbb{R}^3$ 为角速度, $\log$ 为采用式(2.26)的四元数对数。 ω_r 为角速度前馈项, 即式(4.15a)为前馈加反馈的角速度控制, 目的是调节姿态四元数 q , 以跟踪参考姿态 $g_2(x)$ 。直观上看, 由于在位置的 DS 中, x 最终将收敛至原点, 则 $g_2(x)$ 也将收敛至无旋转四元数 q_I , 因此 q 将最终收敛至 q_I 。

因此, 空间 A 中融合位置和姿态的 DS 可写为:

$$\dot{\xi} = \begin{bmatrix} \dot{x} \\ \omega \end{bmatrix} = f_A(x, q) = \begin{bmatrix} f_{A,x}(x) \\ f_{A,q}(x, q) \end{bmatrix} = \begin{bmatrix} \gamma_1(x)Px \\ \gamma_2(q)\beta \log(q * \bar{g}_2(x)) + \omega_r \end{bmatrix}, \quad (4.16)$$

其中, $\xi = [x^T; q^T]^T$ 为系统的状态量, $\gamma_1(x) > 0$, $\gamma_2(q) > 0$, $\beta < 0$, P 为对称负定矩阵。由式(4.3), 可写出空间 B 中位置的 DS 为:

$$\dot{x} = f_{B,x}(x) = \gamma_1(h_1^{-1}(x))J_{h_1}(h_1^{-1}(x))Ph_1^{-1}(x). \quad (4.17)$$

基于式(3.3b), 可知:

$$\dot{q}' = \frac{\partial g_1(x)}{\partial x} \dot{x} * q + g_1(x) * \dot{q} = J_{g_1}(x) \dot{x} * q + g_1(x) * \dot{q}. \quad (4.18)$$

角速度可以由式(2.29)计算, $\omega_B = 2\dot{q}' * \bar{q}'$, 并结合式(4.13)的反函数, 因此最终空间 B 中姿态的 DS 为 $\omega = f_{B,q}(x, q)$:

$$\begin{aligned} \omega &= 2(J_{g_1}(x_A)\dot{x}_A * q_A + g_1(x_A) * \dot{q}_A) * \bar{q} \\ &= 2\left(J_{g_1}(x_A)f_{A,x}(x_A) * \bar{g}_1(x_A) * q + g_1(x_A) * \frac{1}{2}f_{A,q}(x_A, q_A) * q_A\right) * \bar{q} \\ &= 2J_{g_1}(x_A)f_{A,x}(x_A) * \bar{g}_1(x_A) + g_1(x_A) * f_{A,q}(x_A, q_A) * \bar{g}_1(x_A), \end{aligned} \quad (4.19)$$

其中:

$$\begin{cases} x_A = h_1^{-1}(x), \\ q_A = \bar{g}_1(x_A) * q = \bar{g}_1(h_1^{-1}(x)) * q. \end{cases} \quad (4.20)$$

将式(4.16)和(4.20)带入式(4.19)即可得到空间 B 中姿态的 DS。其中, $x \in \mathbb{R}^D$, $D = 2$ 、 $D = 3$ 分别表示二维空间和三维空间。单位四元数 $q \in \mathcal{S}^3$ 且始终包含 4 个元素。因此, 空间 B 中融合位置和姿态的 DS 为:

$$\dot{\xi} = \begin{bmatrix} \dot{x} \\ \omega \end{bmatrix} = f_B(x, q) = \begin{bmatrix} f_{B,x}(x) \\ f_{B,q}(x, q) \end{bmatrix}. \quad (4.21)$$

4.3.3 动态系统稳定性分析

(1) 位置的稳定性

为保证式(4.9)的 DS 稳定性, 取特征值 $\lambda_1 < 0, \lambda_2 < 0$ 。若 $\gamma(\mathbf{x}) = \gamma_0 > 0$ 为常数, 则式(4.9)为线性时不变系统, 且有特征值 $\gamma_0\lambda_1 < 0, \gamma_0\lambda_2 < 0$, 则其为全局渐进稳定的, 且有唯一平衡点 $\mathbf{0}$ 。

若式(4.9)中 $\gamma(\mathbf{x})$ 为状态量 $\mathbf{x}$ 的函数, 取 Lyapunov 函数 $\mathcal{V}_1(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$, 有 $\mathcal{V}_1(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \mathbb{R}^D$, 当且仅当 $\mathbf{x} = \mathbf{0}$ 时, $\mathcal{V}_1(\mathbf{x}) = 0$, 当 $\|\mathbf{x}\| \rightarrow \infty, \mathcal{V}_1(\mathbf{x}) \rightarrow \infty$, 即满足定理2.2.2中式(2.7a)和式(2.7c)。对 $\mathcal{V}_1(\mathbf{x})$ 其求时间的导数:

$$\dot{\mathcal{V}}_1(\mathbf{x}) = 2\mathbf{x}^T \dot{\mathbf{x}} = 2\mathbf{x}^T \gamma(\mathbf{x}) \mathbf{P} \mathbf{x} = 2\gamma(\mathbf{x}) \mathbf{x}^T \mathbf{P} \mathbf{x}, \quad (4.22)$$

对称矩阵 $\mathbf{P}$ 的特征值全为负数, 所以 $\mathbf{P}$ 负定, 对任意的 $\mathbf{x} \neq \mathbf{0}$, 有其负定二次型 $\mathbf{x}^T \mathbf{P} \mathbf{x} < 0$, 又因 $\gamma(\mathbf{x}) > 0$ 且为标量, 即 $\dot{\mathcal{V}}_1(\mathbf{x}) < 0$, 当且仅当 $\mathbf{x} = \mathbf{0}$ 时, $\dot{\mathcal{V}}(\mathbf{x}) = 0$, 即满足式(2.7b)。因此 $\dot{\mathbf{x}} = \mathbf{f}_1(\mathbf{x})$ 为全局渐进稳定的。

由定理4.2.1知, 空间 B 中 $\text{DS}\dot{\mathbf{x}} = \mathbf{f}_2(\mathbf{x})$ 也是全局渐进稳定的, 且有唯一平衡点 $\mathbf{0}$, 其 Lyapunov 函数可取为:

$$\mathcal{V}_2(\mathbf{x}) = \mathcal{V}_1(\Phi^{-1}(\mathbf{x})) = [\Phi^{-1}(\mathbf{x})]^T \Phi^{-1}(\mathbf{x}). \quad (4.23)$$

(2) 姿态的稳定性

针对式(4.15)所示姿态的 DS, 构建如下 Lyapunov 函数:

$$\mathcal{V}_2 = \mathbf{e}^T \mathbf{e}, \quad (4.24a)$$

$$\text{其中: } \mathbf{e} = \log(\mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x})), \quad (4.24b)$$

从而有 $\mathcal{V}_2 \geq 0, \forall \mathbf{q} \in \mathcal{S}$, 当且仅当 $\mathbf{q} = \pm \mathbf{g}_2(\mathbf{x})$ 时, $\mathbf{e} = \mathbf{0}, \mathcal{V}_2 = 0$ 。

由四元数 $\mathbf{q}_e = \mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x})$ 的对数表达式(2.26), 对 $\log(\mathbf{q}_e)$ 求对 $\mathbf{q}_e$ 的偏导数, 则有:

$$\frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} = \begin{bmatrix} -\frac{q_2}{1-q_1^2} & \frac{(q_3^2+q_4^2)\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & -\frac{q_2q_3\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & -\frac{q_2q_4\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} \\ -\frac{q_3}{1-q_1^2} & -\frac{q_2q_3\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & \frac{(q_2^2+q_4^2)\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & -\frac{q_3q_4\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} \\ -\frac{q_4}{1-q_1^2} & -\frac{q_2q_4\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & -\frac{q_3q_4\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} & \frac{(q_2^2+q_3^2)\cos(q_1)}{(1-q_1^2)^{\frac{3}{2}}} \end{bmatrix}, \quad (4.25)$$

其中, $[q_1, q_2, q_3, q_4]^T = \mathbf{q}_e, \mathbf{e} = \log(\mathbf{q}_e)$ 。

对式(4.24a)求对时间的导数, 可得:

$$\dot{\mathcal{V}}_2 = 2\mathbf{e}^T \dot{\mathbf{e}} = 2\mathbf{e}^T \frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} \left(\dot{\mathbf{q}} * \bar{\mathbf{g}}_2(\mathbf{x}) + \mathbf{q} * \frac{\partial \bar{\mathbf{g}}_2(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} \right), \quad (4.26)$$

由于 $\dot{\mathbf{q}} = \frac{1}{2}\boldsymbol{\omega} * \mathbf{q}$ ，并将式(4.15a)和(4.15b)带入上式，得：

$$\begin{aligned}\dot{\mathcal{V}}_2 &= 2\mathbf{e}^T \frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} \left(\frac{1}{2} \boldsymbol{\omega} * \mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x}) + \mathbf{q} * \frac{\partial \bar{\mathbf{g}}_2(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} \right) \\ &= 2\mathbf{e}^T \frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} \left(\frac{1}{2} (\gamma_2(\mathbf{q})\beta \log(\mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x})) + \boldsymbol{\omega}_r) * \mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x}) + \mathbf{q} * \frac{\partial \bar{\mathbf{g}}_2(\mathbf{x})}{\partial \mathbf{x}} \dot{\mathbf{x}} \right) \\ &= \mathbf{e}^T \frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} \gamma_2(\mathbf{q})\beta \log(\mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x})) * \mathbf{q} * \bar{\mathbf{g}}_2(\mathbf{x}) \\ &= \beta \gamma_2(\mathbf{q}) \log(\mathbf{q}_e)^T \frac{\partial \log(\mathbf{q}_e)}{\partial \mathbf{q}_e} \log(\mathbf{q}_e) * \mathbf{q}_e.\end{aligned}$$

将式(4.25)带入上式，并采用 $[q_1, q_2, q_3, q_4]^T = \mathbf{q}_e$ 将上式展开，且有 $q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1$ ，可得：

$$\dot{\mathcal{V}}_2 = \beta \gamma_2(\mathbf{q}) \operatorname{acos}^2 q_1,$$

其中， $\beta < 0, \gamma_2(\mathbf{q}) > 0$ 。因此 $\dot{\mathcal{V}}_2 \leq 0$ ，当且仅当 $\operatorname{acos} q_1 = 0$ 时， $\dot{\mathcal{V}}_2 = 0$ ，则有 $q_1 = \pm 1$ ， $\mathbf{q}_e = [\pm 1, 0, 0, 0]^T$ ， $\mathbf{q} = \pm \mathbf{g}_2(\mathbf{x})$ ， $\mathbf{q}$ 和 $-\mathbf{q}$ 表示相同的姿态。因此，式(4.15)所示姿态的 DS 是全局渐进稳定的，且收敛至姿态四元数 $\mathbf{g}_2(\mathbf{x})$ 。由于位置 $\mathbf{x}$ 将在有限时间内收敛至零点， $\mathbf{g}_2(\mathbf{x})$ 也将在有限时间内收敛至无旋转四元数 $\mathbf{q}_I = [1, 0, 0, 0]^T$ ，即该姿态的 DS 将最终稳定于 $\mathbf{q}_I$ 。

(3) 位姿的稳定性

可取 Lyapunov 函数 $\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2$ ，易知其满足式(2.7)。因此，式(4.16)中位姿的 DS 是全局渐进稳定的，其位置将收敛至零点，姿态将跟踪参考姿态 $\mathbf{g}_2(\mathbf{x})$ ，并最终稳定至 $\mathbf{q}_I$ 。

针对二维空间或者三维空间中，式(4.16)所描述的位姿的 DS，矩阵 $\mathbf{P}$ 的特征值 $\lambda_1 < 0$ 代表着沿 $\mathbf{x}_1$ 朝着原点方向的运动速度比例系数， $\lambda_2 < 0$ 表示朝着垂直于 $\mathbf{x}_1$ 方向的运动速度比例系数。为保证约束条件2)，即精确地表征示教数据，一般取 $|\lambda_2| > |\lambda_1|$ ， $|\lambda_2|$ 越大，空间 A 中测试轨迹越靠近轨迹 A，则空间 B 中轨迹越靠近示教轨迹 B。 $\gamma(\mathbf{x}) > 0$ 用于调节系统的运动速度， $\gamma(\mathbf{x})$ 越大，收敛至原点耗时越短。 $\beta \gamma_2(\mathbf{q})$ 为角速度的比例系数，其绝对值越大则收敛至 $\mathbf{g}_2(\mathbf{x})$ 的速度越快。

由定理4.2.1可知，可得示教数据空间中的，式(4.21)所述的融合位置和姿态的 DS 也是全局渐进稳定的，且将稳定于示教轨迹的终点位姿 $\mathbf{x}'_N, \mathbf{q}'_N$ 。

4.4 运动规划仿真

基于上一节提出的 DS 框架，本节在二维空间中进行轨迹生成仿真。仍采用上一节中的方法基于示教数据 $\mathbf{B} = \{\mathbf{x}'_i, \mathbf{q}'_i\}_{i=1}^N$ 构建空间 A 中的数据 A。采用第3章中3.4小节的迭代算法分别构建空间 A 至空间 B 的映射 Φ_1 （式(4.13)）和 Φ_2 （式(4.14)）。如

图4.4所示, 采用瑞士洛桑联邦理工大学学习算法与系统 (Learning Algorithms and Systems Laboratory, LASA) 实验室的手写字母数据集¹进行仿真测试。

空间 B 中有“C”形示教数据 (红色曲线), 在空间 A 中为蓝色轨迹 A。图4.5表示形状为 τ 的手写字母仿真轨迹线。图4.6为汤勺形示教数据的 DS 仿真。图4.7为示教轨迹为 sin 曲线的 DS 仿真结果。左侧为隐空间 A, 右侧为原示教数据空间, 可以看出, 本章提出的 DS 所生成的轨迹线均能稳定收敛至零点, 且能较好的描述示教数据。图4.8中, 左侧绘制了 LASA 数据集的 7 次手绘 sin 曲线的路径。红色曲线为以各条路径的起点为 DS 的初始点进行仿真的轨迹, 右侧图片为其对应的时间速度曲线, 可以看出, DS 复现的速度与示教数据吻合。

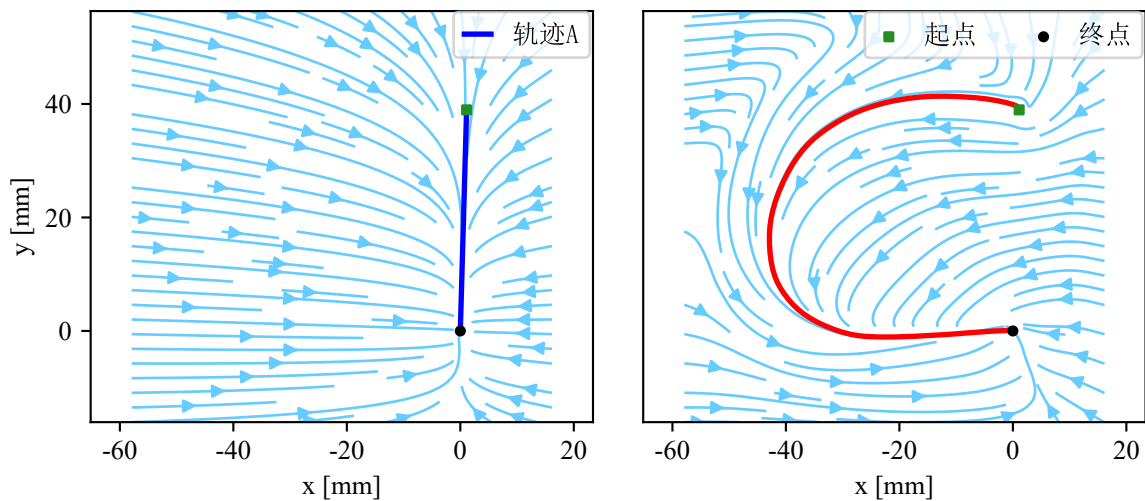


图 4.4 C 形曲线的 DS 仿真

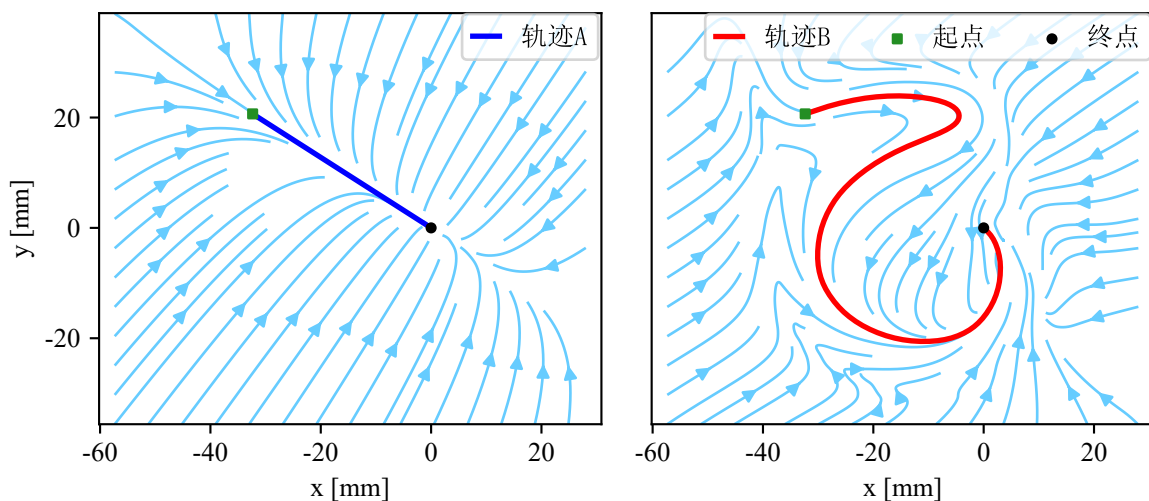


图 4.5 τ 形曲线的 DS 仿真

¹<http://bitbucket.org/khansari/lasahandwritingdataset>

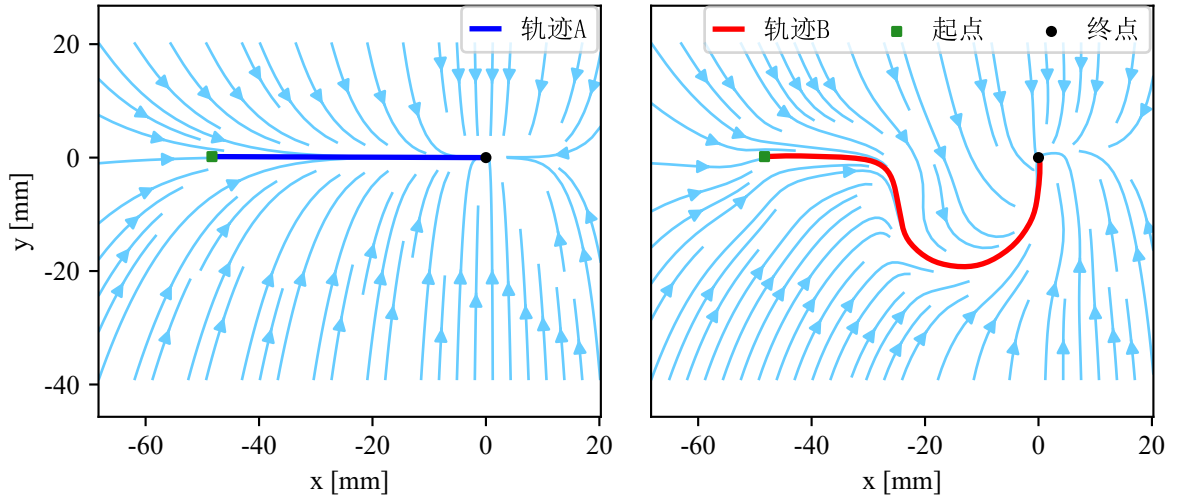


图 4.6 汤勺形曲线的 DS 仿真

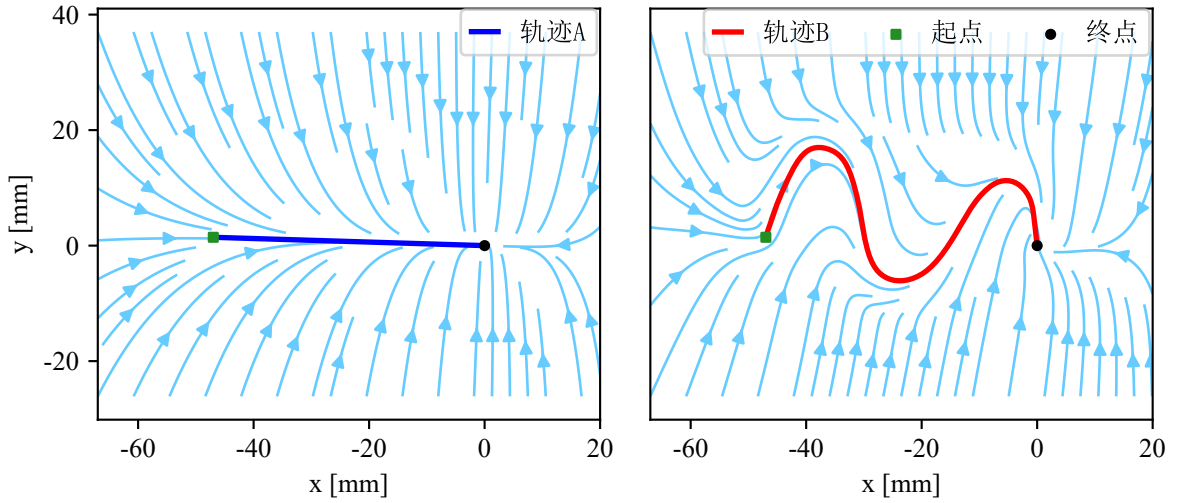


图 4.7 sin 曲线的 DS 仿真

设示教数据 $\mathbf{B}$ 采样间隔为 Δt , 则有 $t = (i - 1)\Delta t$ 时, $(\mathbf{x}, \mathbf{q}') = \mathbf{x}'_i, \mathbf{q}'_i$, 该条示教数据总时长为 $T = (N - 1)\Delta t$. 为保证空间 A 中, 当 $t = (i - 1)\Delta t, i = 1, 2, \dots, N$ 时, 以 $\mathbf{x} = \mathbf{x}_1$ 为起始点的运行轨迹刚好到达 $\frac{N-i}{N-1}\mathbf{x}_1$ 点, 取 $\gamma_1(\mathbf{x})$ 为:

$$\gamma_1(\mathbf{x}) = \begin{cases} \frac{\|\mathbf{x}_1\|}{N\Delta t\|\mathbf{x}\|}, & \|\mathbf{x}\| \geq \frac{\|\mathbf{x}_1\|}{N} \\ \frac{1}{\Delta t}, & \text{其他.} \end{cases}, \quad (4.27)$$

此外, 本节取 $\lambda_1 = -1, \lambda_2 = -5$. 注意到, $\gamma_1(\mathbf{x})$ 为位置 $\mathbf{x}$ 的函数且连续.

图4.4中, DS 运动轨迹线为所示的浅蓝色曲线, 由于左右两侧的 DS 都是全局渐进稳定的, 且有唯一平衡点 $\mathbf{0}$, 因此轨迹线都收敛至原点. 空间 A 中, 轨迹线在收敛至原点的同时, 也向轨迹 A 靠近. 在微分同胚映射 Φ_1 的作用下, 空间 B 中 DS 的运动

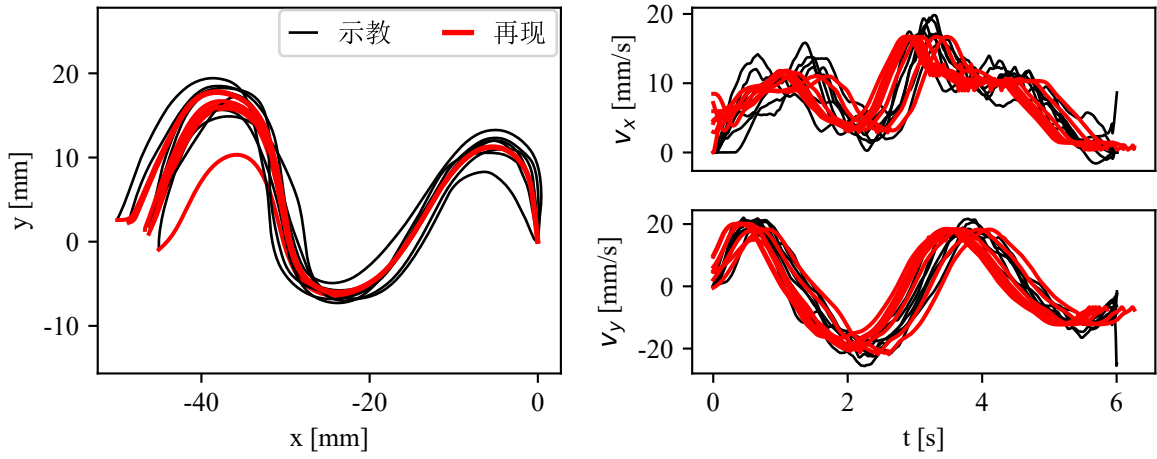


图 4.8 轨迹复现效果

轨迹线也在收敛至原点的同时向红色示教轨迹靠近。

类似于式(4.28)，为保证空间 A 中，当 $t = (i-1)\Delta t, i = 1, 2, \dots, N$ 时，以 $(\mathbf{x}, \mathbf{q}) = \mathbf{x}_1, \mathbf{q}_1$ 为起始姿态的运行轨迹刚好到达 $(k\mathbf{x}_1, \mathbf{q}^k), k = \frac{N-i}{N-1} (i = 1, 2, \dots, N)$ 点。这样可以保证 B 空间生成的 DS 满足如下性质：以 $(\mathbf{x}'_1, \mathbf{q}'_1)$ 为起始位姿的轨迹线与示教轨迹 B 相同，且其附近的轨迹线与轨迹 B 有相似的运动速度。因此，取 $\gamma_2(\mathbf{q})$ 为：

$$\gamma_2(\mathbf{q}) = \begin{cases} \frac{\theta_1}{N\Delta t\theta(\mathbf{q})}, & \theta(\mathbf{q}) \geq \frac{\theta_1}{N}, \\ \frac{\theta_1}{N}, & \text{其他.} \end{cases} \quad (4.28a)$$

$$\text{其中: } \theta_1 = d(\mathbf{q}_1, \mathbf{q}_I), \quad (4.28b)$$

$$\theta(\mathbf{q}) = d(\mathbf{q}, \mathbf{q}_r(\mathbf{x})), \quad (4.28c)$$

$$\mathbf{q}_r(\mathbf{x}) = \mathbf{g}_1(\mathbf{h}_1^{-1}(\mathbf{x})) * \mathbf{g}_2(\mathbf{h}_1^{-1}(\mathbf{x})). \quad (4.28d)$$

θ_1 为采用式(2.32)求示教初始点的姿态 $\mathbf{q}_1$ 和无旋转四元数之间的距离。为防止 θ_1 较小时， $\gamma_2(\mathbf{q})$ 过小导致姿态收敛速度过慢，若 $\theta_1 < \frac{\pi}{2}$ ，则取 $\theta_1 = \frac{\pi}{2}$ 。 $\theta(\mathbf{q})$ 为四元数 $\mathbf{q}$ 与参考姿态四元数 $\mathbf{q}_r(\mathbf{x})$ 之间的距离，其中 $\mathbf{q}_r(\mathbf{x})$ 为仅含位置变量 $\mathbf{x}$ 的函数。

图4.9绘制了采用迭代法（章节3.4）构建微分同胚映射 Φ_1 和 Φ_2 的效果。在空间 D （图4.9(a））中绘制等间距网格线，对每个网格点设定姿态为无旋转四元数 $\mathbf{q}_I$ 。映射 Φ_2 将空间 D 映射至空间 A （图4.9(b）），可以看出，其网格点位置不变，姿态发生较大变化：当网格点靠近轨迹 A 时，其姿态方向（灰色箭头）也靠近数据 A 所示的方向（蓝色箭头）。映射 Φ_1 将空间 A 映射至空间 B （图4.9(c）），网格线在其作用下发生扭曲，同时靠近轨迹 B 的网格点姿态方向也与轨迹 B 上的姿态（红色箭头）相近。

二维空间中，一个刚体的自由度为 3，即只需 3 个独立变量来完整描述其位姿，这里取其二维笛卡尔坐标和角度 $\theta(\mathbf{q})$ 。针对二维空间中姿态的 DS ， $\theta(\mathbf{q})$ 为其与 x 轴正

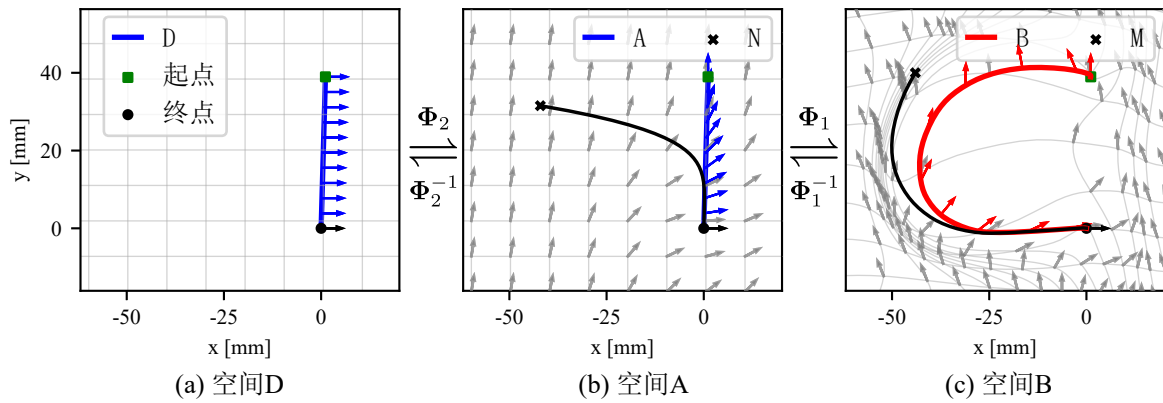


图 4.9 两个姿态微分同胚映射二维仿真

方向的夹角。

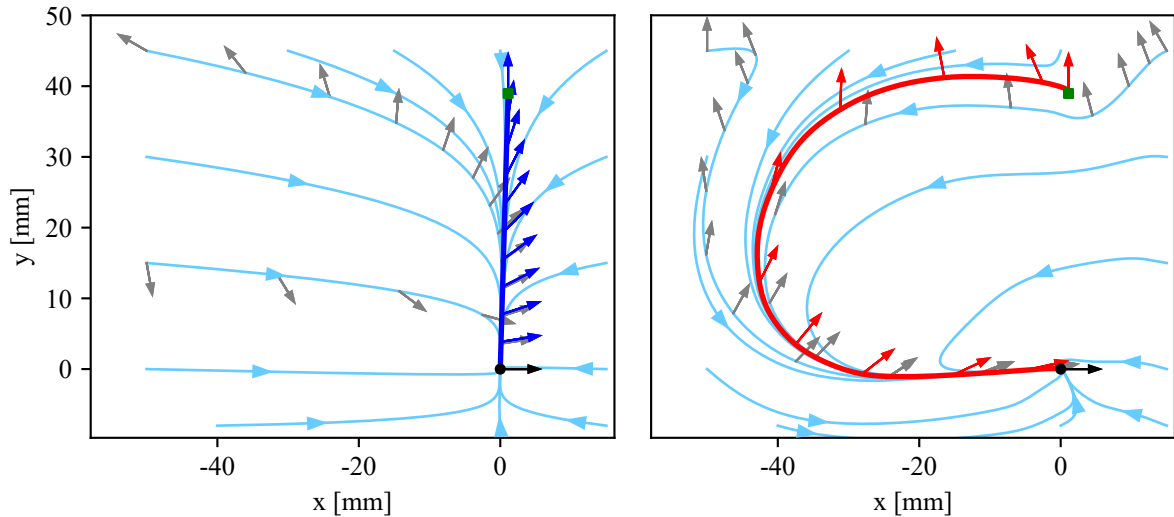


图 4.10 融合位置和姿态的 DS

如图4.10所示，浅蓝色曲线仍为 DS 的运动轨迹线。图中箭头表示该点的姿态方向。仍然采用 C 形的示教轨迹，并在每个点赋予其姿态角，姿态方向如图中红色箭头所示，初始姿态角为 90° ，随着运动过程先增加，随后减小至 0 。空间 A 中左侧以某个点为起始点，任意生成该点的姿态（灰色箭头），可以看出其姿态方向将逐渐靠近 x 方向（水平向右），当其位置靠近轨迹 A 时，其姿态将靠近 A 的姿态。在空间 B 中也体现出相同的特性，即该 DS 可以用于表征示教数据。

为可视化空间 B 中姿态的 DS 收敛情况，取某一点为起点 $\mathbf{x}_{\text{test}}$ ，并给定三组随机的初始姿态，其姿态角的时间轨迹如图4.11所示。可以看出，测试点的姿态角缓慢跟踪参考姿态角，并随参考轨迹一起收敛至 0 。结果表明，姿态的 DS 也是稳定的，且能准确描述示教数据的姿态。

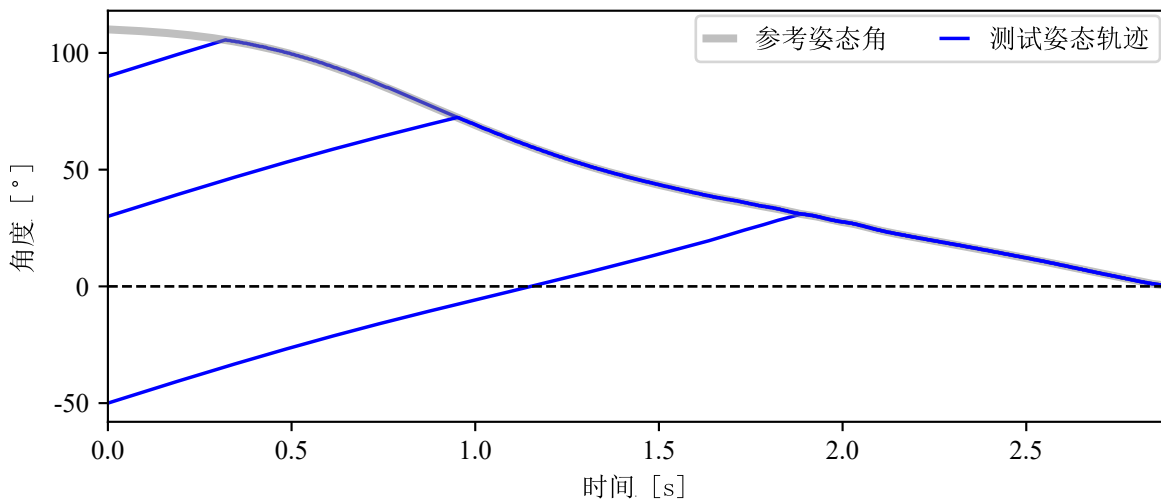
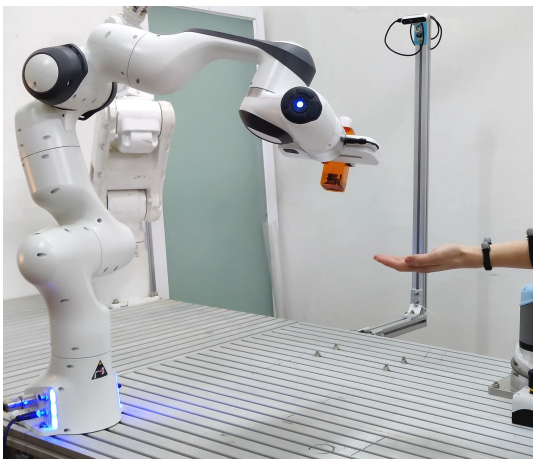


图 4.11 姿态角随时间变化曲线

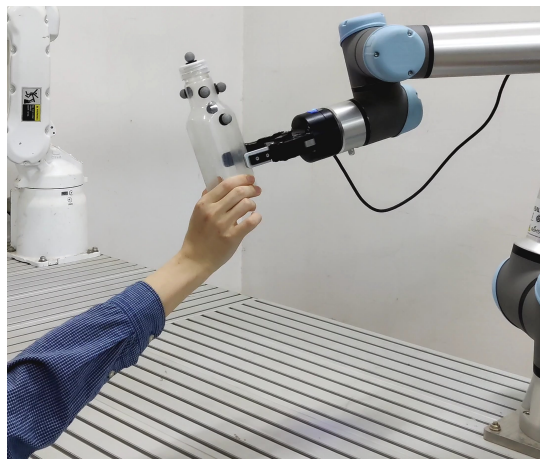
4.5 动态环境运动规划实验

本节采用上述的 DS 用于机械臂的在线运动规划，设计了人机交互递交实验平台和动态抓取实验平台，用于验证算法对未知动态环境和干扰的在线适应性。

4.5.1 实验平台



(a) 人机交互递交实验平台



(b) 动态抓取实验平台

图 4.12 基于 DS 的运动规划实验平台

本节搭建了两个机械臂实验平台，如图4.12所示，均采用 Vicon 8 X Vero2.2 光学三维运动捕捉系统进行刚体的实时定位，该系统与机械臂基座之间的位姿变换矩阵已事先标定好。控制系统基于 ROS 进行搭建。

在人机交互递交实验平台（图4.12(a)）中，Panda 机械臂抓取棕色药瓶并递交给人

手，人手腕处贴有光学三维运动捕捉系统的标记点，用于实时获取人手心位姿。DS 的构建以人手心的位姿为基准坐标系，实验过程中，人手可随意移动，当机械臂运动轨迹收敛至零点时，则其恰好处于人手心位姿处，以完成递交任务。

在动态抓取实验（图4.12(b)）中，人手持药瓶，需控制机械臂来抓取药瓶，该过程中人手可随意变换位姿，以测试该运动规划算法在未知动态环境下的性能。药瓶上贴有 6 个动作捕捉系统的标记点，用于实时测量其位姿。在该实验中采用优傲 (Universal Robot 5 e-series, UR5e) 的 6 轴机械臂，其末端安装有 Robotiq 85 二指手，以实现抓取操作。其中 DS 模型的基准坐标系为药瓶的待抓取坐标系。

4.5.2 任务示教与运动规划模型建立

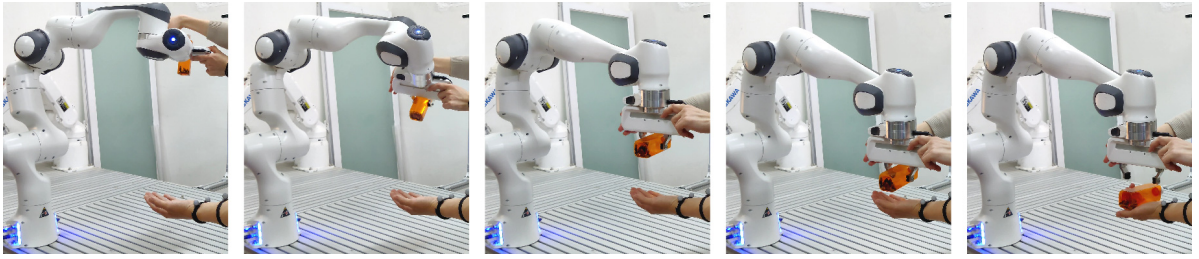


图 4.13 递交示教过程

本节采用拖动示教的方法采集数据，示教过程基于零力控制，考虑机械臂的动力学方程(2.37)，设计如下控制器：

$$\tau = \hat{M}(\theta)\ddot{\theta} + \hat{c}(\theta, \dot{\theta}) + \hat{g}(\theta). \quad (4.29)$$

这种方式需要精确地建模机械臂动力学参数 $\hat{M}, \hat{c}, \hat{g}$ ，并进行力矩补偿，将上式带入机械臂动力学方程(2.37)，若忽略建模误差，则闭环系统为 $f_{ext} = 0$ ，即机械臂与外界的接触力始终维持在零，从而操作者可拖动机械臂去执行所需的任务，并由机械臂自身的位置/力传感器来记录示教数据。

实验一中，示教将药瓶放置至人手腕的轨迹进行示教，如图4.13所示。1 号操作人员手腕贴有运动捕捉系统的位置标记点，2 号操作人员双手持 Panda 机械臂末端并拖动机械臂向人手腕运动，并放下药瓶。记录该过程中的位姿和人手位姿数据。将该数据表达至人手坐标系中，作为示教轨迹 B（图4.14(b)中红色轨迹），起点为绿色方形点，示教轨迹向左下方运动，终点为黑色圆点。采用式(4.1)生成隐空间 A 中的轨迹 A（图4.14(a)中蓝色轨迹），基于章节4.3中的方法构建隐空间 A 中的全局渐进稳定的 DS，并采用章节3.4中的迭代算法构建微分同胚映射 Φ ，从而生成空间 B 中融合位置和姿态的全局渐进稳定的 DS。图4.14(a)中绘制有等间距的灰色网格线，可以看出在映射 Φ 的作用下，

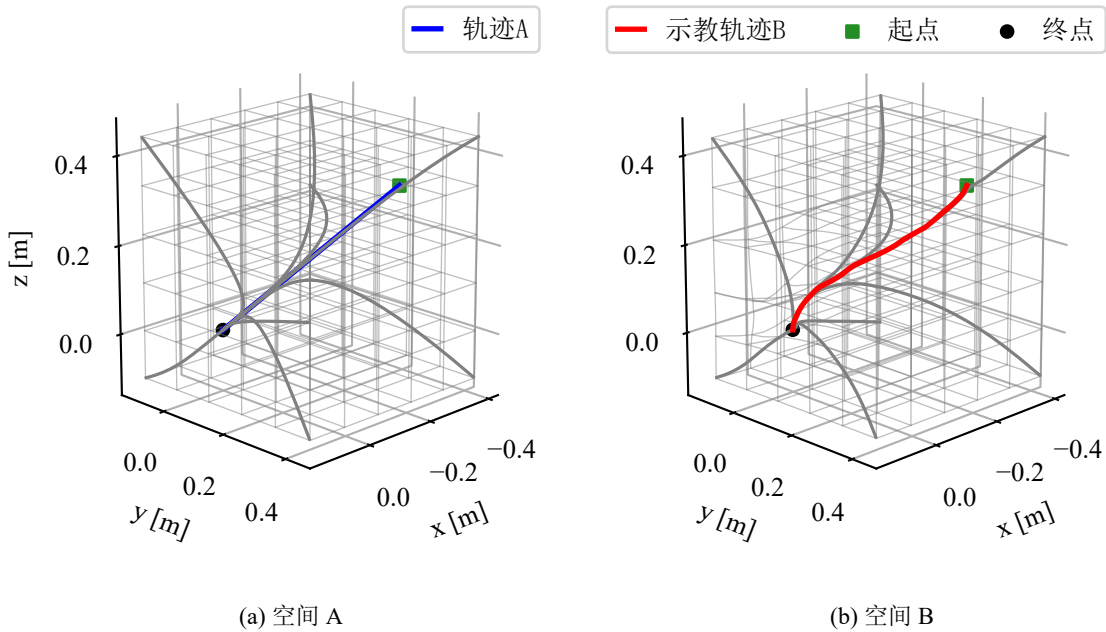


图 4.14 递交示教轨迹和其 DS 结果

空间被平滑地扭曲。取该网格的最外侧的 8 个顶点为机械臂的初始点，对 DS 进行积分得到其轨迹线。可以看出，其轨迹平滑，往示教轨迹靠近的同时往终点运动。

实验二中的示教过程类似于实验一，且采用相同的方式构建 DS 模型，用于后续机械臂实验的动态运动规划。

4.5.3 实验过程和结果分析

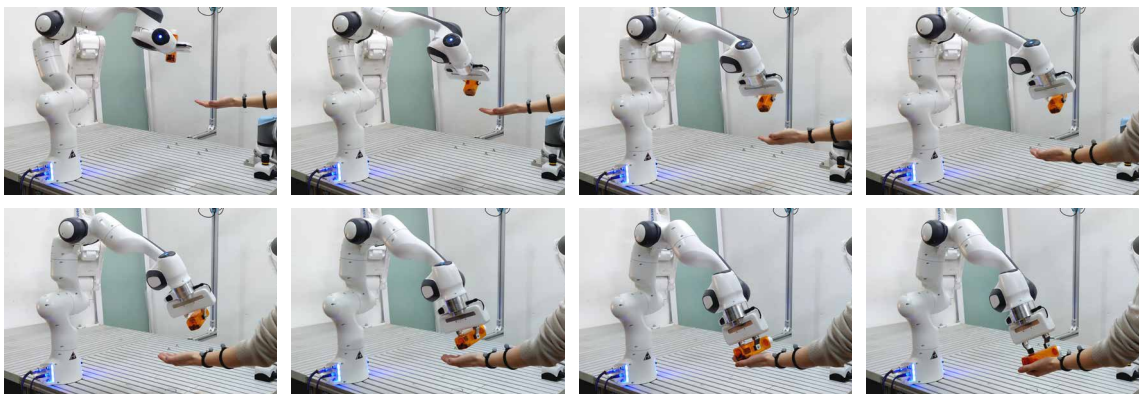


图 4.15 动态递交过程

(1) 动态递交实验

Panda 机械臂控制器采用 3.7.2 中的式 3.20 所描述的阻抗控制器，可实现机械臂操作过程中与环境/人之间柔性的交互，其参考轨迹由 DS(4.21) 经离散积分得到，设 t 时刻

机械臂末端位姿为 $(\boldsymbol{x}, \boldsymbol{q})$ ，则 $t + \Delta t$ 时刻其参考位置和姿态为：

$$\begin{aligned} \boldsymbol{x}_d &= \boldsymbol{x} + \dot{\boldsymbol{x}}\Delta t, \\ \boldsymbol{q}_d &= \exp\left(\frac{1}{2}\boldsymbol{\omega}\Delta t\right) * \boldsymbol{q}, \end{aligned} \quad (4.30)$$

其中 $\dot{\boldsymbol{x}}$ 和 $\boldsymbol{\omega}$ 由融合位置和姿态的 DS 表达式(4.21)计算， $\Delta t = 20 \text{ ms}$ 为控制周期。阻抗控制器的参数设置为 $\boldsymbol{K}_p = 800\boldsymbol{I}_3$, $\boldsymbol{K}_v = 30\boldsymbol{I}_3$, $\boldsymbol{K}_{pr} = 250\boldsymbol{I}_3$, $\boldsymbol{K}_{vr} = 2\boldsymbol{I}_3$ 。

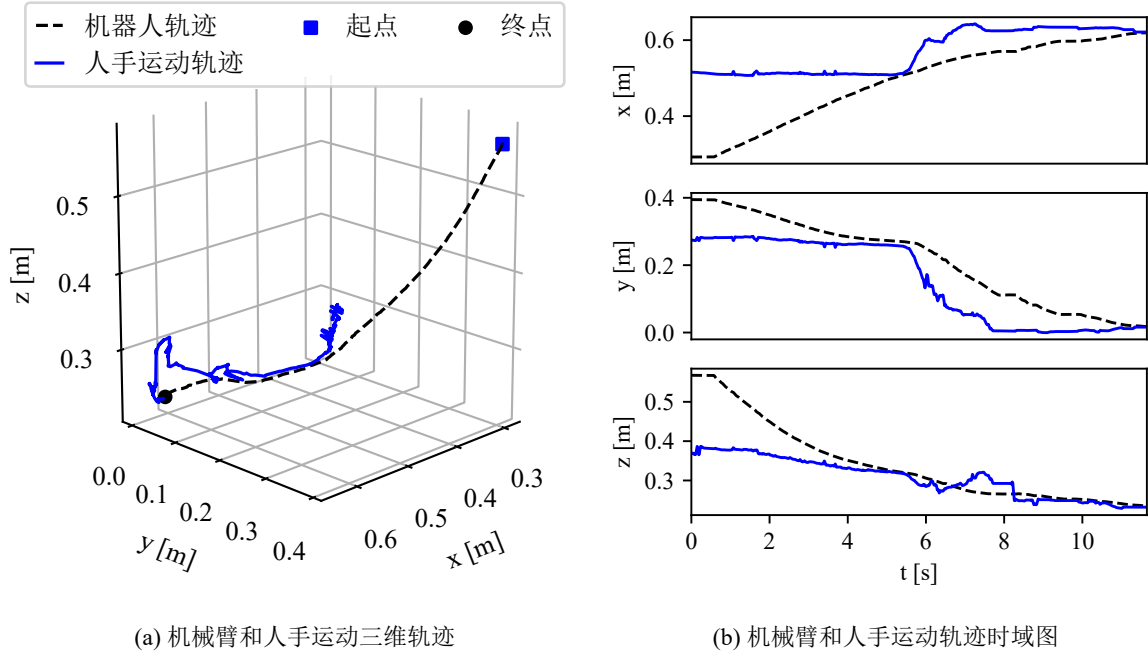


图 4.16 递交过程机械臂和人手运动轨迹

动态递交过程如图4.15所示，机械臂夹取药瓶并向人手心运动，运行途中人手忽然移动至新的位置，机械臂也能适新的目标点并重新向新的终点运动，从而顺利的实现物体的递交过程。实验视频见²。

机械臂和人手运动轨迹如图4.16所示，蓝色实线为人手运动轨迹，黑色虚线为机械臂运动轨迹。图4.16(a)展示了其三维空间中的运动情况，图4.16(b)绘制了其 x 、 y 和 z 坐标（均相对于 Panda 机械臂基座）随时间变化曲线。可以看出 0-5 s 内，人手 x 和 y 坐标几乎不变，但 z 方向的位置缓慢减小，该过程中，机械臂末端位置不断向人手靠近。在 $t \approx 4.8 \text{ s}$ 时人手开始向远处移动，可以看出，机械臂的运动轨迹也能适应其运动并最终到达手心上方，实现递交任务。

(2) 动态抓取实验

UR5e 机械臂支持位置和速度控制模式，控制频率为 500 Hz。UR5e 机械臂末端集成有六维力/力矩传感器，可实现末端接触力的实时测量。为实现 UR5e 机械臂与人/环

²实验视频<https://www.bilibili.com/video/BV1QQ4y1Z7iM/>

境之间的柔顺交互，采用图2.3所示的导纳控制器。首先由 DS 模型计算出其线速度和角速度，并由式(4.30)计算参考轨迹，随后采用式(2.34)计算其关节角速度。

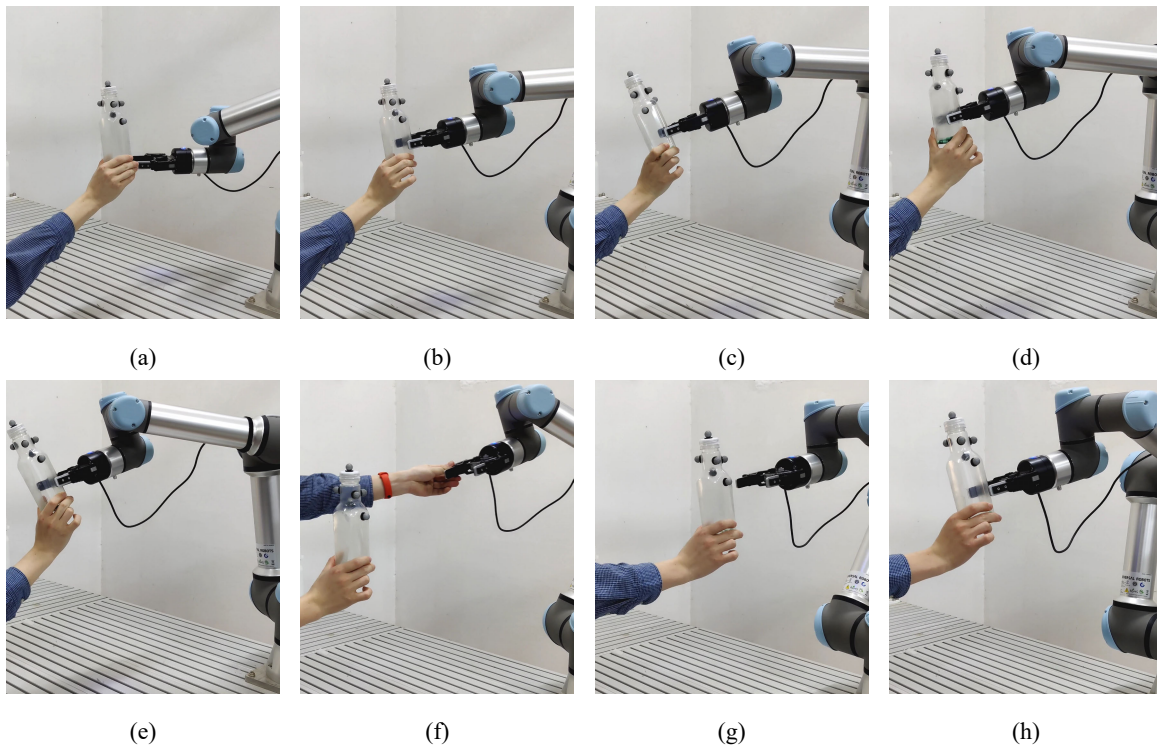


图 4.17 动态抓取过程

实验过程如图4.17所示，并记录了实验过程中机械臂末端位姿（二指手末端）和药瓶位姿（药瓶的期望抓取位姿）变化曲线。所有位姿均相对于 UR5e 的基座坐标系。实验视频见³。轨迹曲线包含 x 、 y 和 z 坐标的时间曲线（图4.18）和其四元数变化曲线（图4.20），并根据式(2.32)求机械臂末端姿态和药瓶姿态之间的距离，可得其测地线角度时间曲线（图4.19）。在图4.17(a-e)中，操作员右手持药瓶并不断调整其位置和姿态，可以看出，机械臂末端夹爪可以动态适应人手的移动。从机械臂和药瓶的轨迹线时间曲线图4.18和图4.20可以看出，药瓶的位置和姿态不断发生变化，机械臂依然能不断向药瓶运动并靠近。在 $t \approx 28$ s 时（图4.18、图4.19和图4.20中红色竖直线段处开始），操作员用左手推动机械臂，以改变其末端位姿（图4.17(f)）。约 3 s 后停止推动，从图4.18、图4.19和图4.20中灰色区域可以看出，机械臂末端位置和姿态发生了较大变化。当外力撤去后，在基于 DS 的运动规划算法的作用下，机械臂又能重新跟踪药瓶位姿。机械臂与药瓶之间的相对位置和姿态不断减小，并随药瓶的运动而运动。机械臂轨迹和药瓶轨迹的最小距离小于 1 mm，姿态误差小于 1° 。

(3) 实验结果分析

³实验视频<https://www.bilibili.com/video/BV1Fq4y1J7hf/>和<https://www.bilibili.com/video/BV1Gp4y147Y9/>

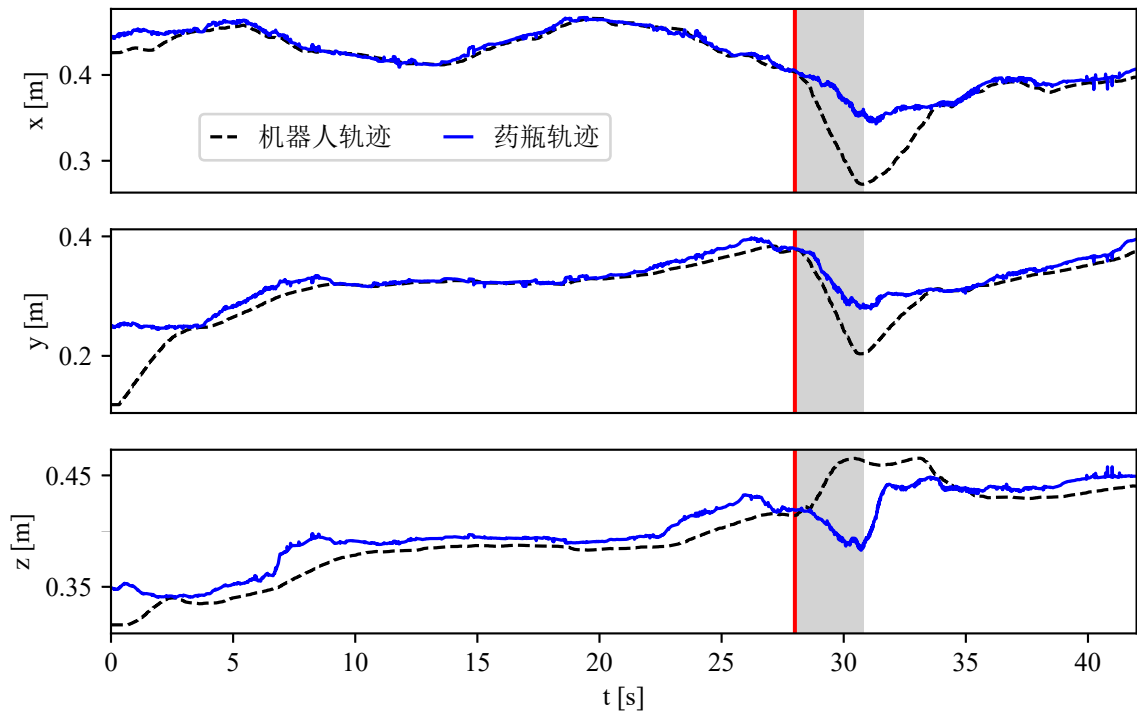


图 4.18 动态抓取过程机械臂和药瓶轨迹

本节进行了两个机械臂实验：物体递交实验和动态抓取实验。从实验结果可以看出，基于模仿学习和微分同胚映射构造的 DS 模型，可以用于动态环境下机械臂的实时运动规划，该方法对在线变化的终点位姿有较好的适应性，并且能在外界干扰的情况下在线规划新的轨迹。机械臂的运动轨迹表明，该模型能较好的表征示教数据，并能保证实时性，以实现机械臂的在线运动规划。

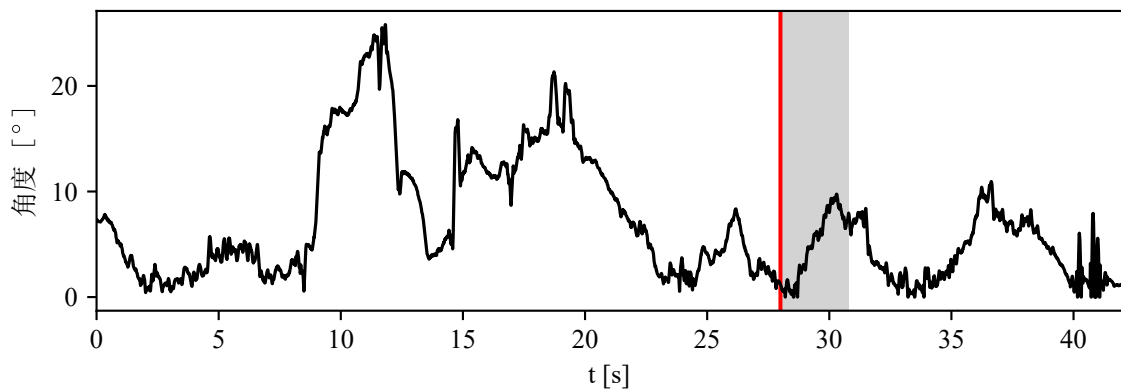


图 4.19 机械臂和药瓶相对姿态角时间曲线

在机械臂的实验中，本节没有过多地研究机械臂的运动控制算法，仅关注于其运动规划算法。采用 DS 进行 Panda 机械臂的在线运动规划过程中，由于动力学模型的不

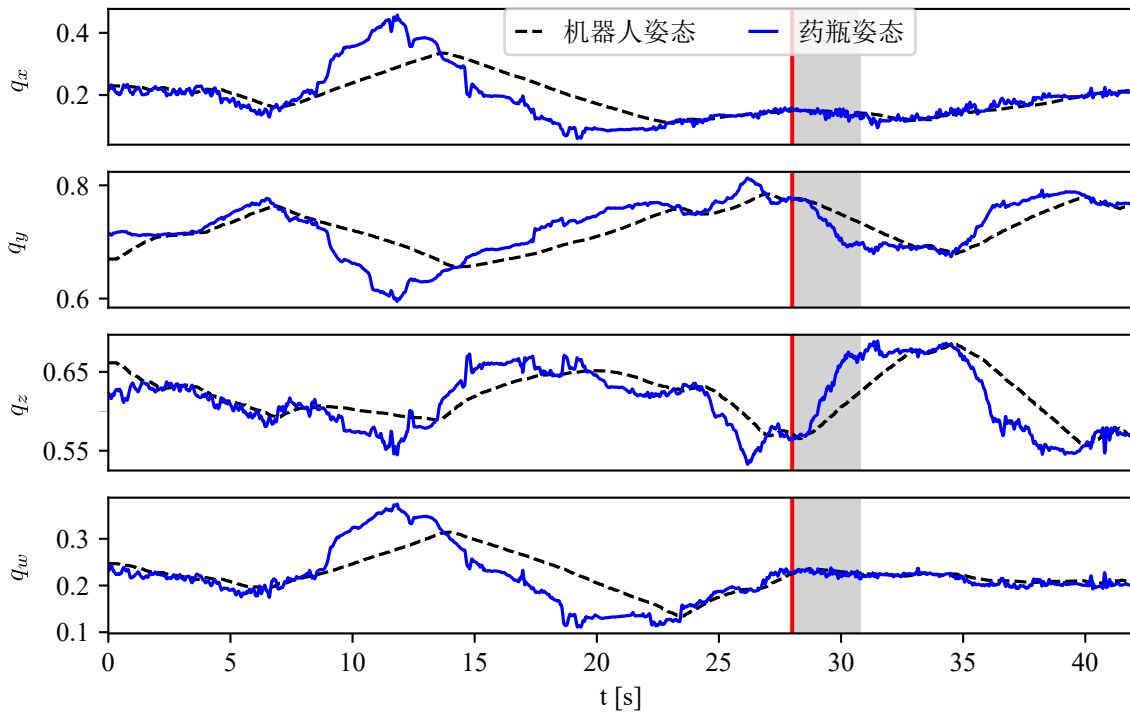


图 4.20 动态抓取过程机械臂和药瓶姿态

精确，导致实际中基于阻抗控制的轨迹跟踪误差无法收敛至零，且较小的参考位姿变动下其运动较慢，需增大其阻抗控制器参数或增大式(4.27)中的 γ_1 和式(4.28)中的 γ_2 ，以增大参考速度。UR5e 机械臂采用位置/速度控制，直接发送关节运动指令，其下位机采用式(2.38)所示的 PID 控制器可实现轨迹的精确跟踪，无需对参考速度进行放大。

4.6 本章小结

本章首先提出了基于模仿学习和微分同胚映射构造 DS 模型的基本思路，并证明了一个全局渐进稳定的 DS 经微分同胚的映射生成的 DS 也是全局渐进稳定的。然后提出了二维空间中位置的 DS 构造方法，首先在隐空间构造一个简单的全局渐进稳定的 DS，然后采用微分同胚映射建立示教数据空间和隐空间之间的映射，从而得到示教空间中全局渐进稳定的 DS。采用类似的思想，提出了融合位置和姿态的 DS 设计框架，并采用 Lyapunov 稳定判据证明了该系统是全局渐进稳定的。基于该框架，在二维空间中进行了运动规划仿真，结果表明该运动规划算法可以很好地表征示教数据，并保证空间内任意起始点的运动轨迹线都收敛至零点。最后，设计了物体递交和动态抓取实验，实验结果表明，采用该 DS 进行运动规划的机械臂可以较好地抵抗外界干扰，并能在在线适应动态的操作环境，验证了本章提出的基于模仿学习和微分同胚映射的运动规划算法，并具有对动态环境的适应性，且能直接用于机械臂的在线运动规划。

5 基于模仿学习的柔顺操作技能学习

5.1 引言

自由空间中仅需考虑机械臂运动规划，针对接触式的任务，仅仅考虑运动规划是不够的。传统的机械臂规划和控制方式中，先规划离线运动轨迹，再根据该运动轨迹设计高增益 PID 控制器实现位置的精确跟踪。然而接触式任务中，微小的位置偏差可能导致过大的接触力损伤机械臂或工件，且极易导致交互过程中机械臂失稳，尤其是在刚性接触环境中。为了让机械臂能像人类一样灵活且安全可靠地在接触式任务中进行操作，并在非结构化环境下具有一定的对变化环境适应能力，一般需要能实时感知接触力并设计力控制策略，以保证机械臂与环境的柔顺接触，并满足任务需求。然而针对不同的接触式任务操作，仍需重新设计和规划位置轨迹，力轨迹和力控制器。

针对上述问题，本章提出基于模仿学习的柔顺操作技能表达模型，建立了统一的接触式任务操作学习框架，用于指导典型的接触式任务操作中的力规划与力控制器设计。通过采集人在完成接触式任务操作中的位置、速度、力、肌电信号等数据，从而提取人完成该任务中的特征，并可将其迁移至机械臂实现仿人的智能操作。最后本章设计了三个实验进行验证，包含机械臂打磨实验，轴孔装配实验和变刚度擦白板实验。

5.2 技能表达方法

接触式的任务操作需要机械臂具有柔顺的操作行为，包括如何调整交互力和末端位姿之间的关系，可描述为如下表达：

$$\Omega : \{\mathbf{X}, \dot{\mathbf{X}}, \mathbf{F}, \mathbf{k}, S\}, \quad (5.1)$$

其中 Ω 指代某个任务， $\mathbf{X}, \dot{\mathbf{X}}$ 分别为任务空间中的位姿和速度， $\mathbf{F}$ 为交互力和力矩， $\mathbf{k}$ 为人手与环境之间的交互刚度， S 为任务约束。技能建模与表达目的是发现这些数据中的内在关系。任务 Ω 的技能模型可表示为：

$$\Omega(\mathbf{X}, \dot{\mathbf{X}}, \mathbf{F}, \mathbf{k}, S) = 0. \quad (5.2)$$

机械臂典型的接触式任务操作场景有打磨抛光，装配，抓取等等，在这些任务中交互力对任务完成质量影响极大。在打磨和抛光任务中，要求末端安装有加工头的机械臂沿着工件表面运动，通常采用力/位混合控制来同时控制接触力和运动，在工件表面接触点处的法线方向控制接触力，其他方向跟踪位置。任务约束 S 用于约束打磨点的位置并要求打磨力始终在合适的范围内。接触力的控制效果直接影响加工质量，如表面粗糙度，尺寸公差，形状公差和位置公差。此类技能可表示为映射函数：

$$\mathbf{X} \xrightarrow{S} \{\dot{\mathbf{X}}, \mathbf{F}, \mathbf{k}\}. \quad (5.3)$$

对于特定的打磨接触点，构建从位置到速度和接触力的映射函数。当有新的打磨任务时，可在线生成期望的打磨速度和打磨力。

针对抓取和装配任务，力和力矩的数据隐含了接触状态的信息。人类可以在无视觉反馈的情况下仅基于触觉反馈完成这类任务，在示教过程中，人基于感知的交互力来调整末端执行器的位姿和速度。例如，在轴孔装配中，约束 S 用于指定 x 和 y 方向的接触力和力矩为 0。此类任务可建模为：

$$\mathbf{F} \xrightarrow{S} \{\mathbf{X}, \dot{\mathbf{X}}, \mathbf{k}\}. \quad (5.4)$$

即基于力传感器反馈的接触力信息来指导位姿和速度。

其中，为采用操作员在执行接触式任务过程中与环境之间的交互刚度 $\mathbf{k}$ ，一般可以基于笛卡尔空间中的弹簧质量阻尼模型（式(1.2)）建立 x 、 y 和 z 方向的交互刚度，并采用最小二乘法估计该二阶模型参数值。然而，此类基于位置偏差和接触力的测量方法，在机械臂与人交互过程中难以实现高实时性和稳定性的参数估计。人体表面肌电信号用于检测手臂肌肉的激活程度，一定程度上反映了手臂的刚度信息。因此，可基于人体表面肌电信号实现人手臂的刚度估计，通过离线测量并标定其映射关系，随后用于接触式任务中示教过程刚度数据的采集。

5.3 接触式任务柔顺操作技能学习

5.3.1 模仿学习框架

人类或动物的学习过程通常是结合模仿和试错进行的。模仿学习通过示教者给出的示例进行学习。通过模仿，使学习者避免了盲目地全局搜索，而能专注于更有意义地局部优化。一般工业机械臂部署需专业编程经验的人进行标定、编程和测试，但实际现场工人无相关的专业知识，因此工业机械臂仅仅用于单一性的重复化任务。模仿学习通过采集人执行任务时的数据，并结合机器学习的方法进行数据处理与模型训练，得到人完成某种任务时的技能，最后将该技能赋予给机械臂，实现机械臂拟人化的任务执行。

具体来说，如图5.1所示。首先采集多组人工示教数据，一方面基于该示教数据构建第4章中的 DS，用于运动规划；另一方面，基于本章的接触式操作技能模型规划接触力信号。运动规划和力的规划器均基于当前的位姿反馈更新参考信号。此外，接触式操作技能可用于指导位置/力的控制器设计，避免了复杂的调参过程。机械臂接收控制信号（控制力矩或者位置/速度）与人或环境进行交互，以完成期望的操作任务。

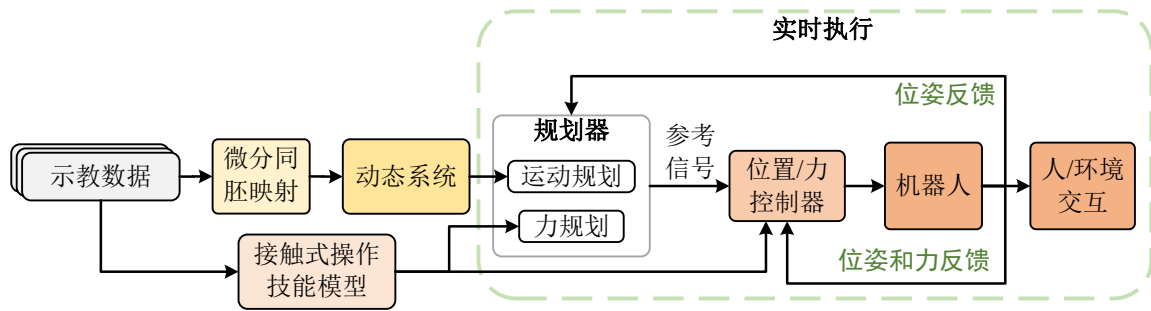


图 5.1 基于模仿学习的规划与控制框架

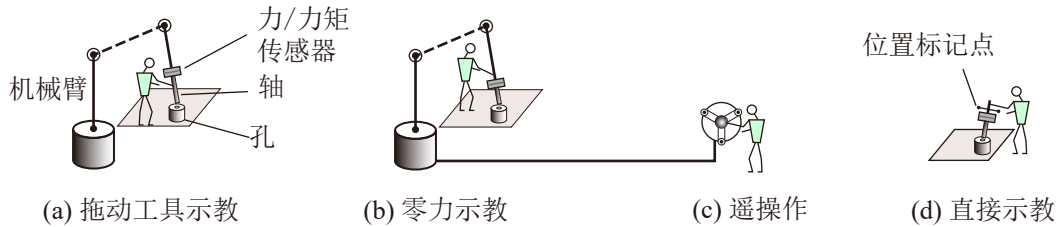


图 5.2 常用的示教方式

5.3.2 人工操作数据采集

常用的示教方式有三种：动觉示教（图5.2(a-b)），远程遥控示教（图5.2(c)），和手工直接示教（图5.2(d)）。针对接触式的任务，尤其是接触力对任务完成质量影响很大的任务，都需要在示教时采集人完成该任务过程的力和位置的数据，因此一般需要位置和力的传感器。动觉示教可分为两类，其中拖动工具进行示教可基于力传感器反馈来拖动机械臂运动，可适用于任何机械臂；直接拖动机械臂进行动觉示教的方式一般采用式(4.29)零力控制来实现拖动示教。动觉示教的方式可以直接记录机械臂的关节角度或是由正运动学记录末端执行器的位姿。遥操作的示教方式中操作员通过控制遥操作杆的运动来控制远端的机械臂执行任务，双边遥操作^[160]中还具有力的反馈从而更好地进行接触式任务的示教。采用手工直接示教的方法最为方便，但需要外部安装位置传感器以获取工具位姿，如惯性测量单元 (Inertial Measurement, IMU)、视觉相机或运动捕捉系统等，此外，还会存在一致性问题 (correspondence problem)^[161]，即人的操作数据迁移至机械臂上可能受限于机械臂/环境自身约束，导致需要对数据进行预处理。

5.3.3 技能学习

5.3.4 技能学习与自主执行控制

考虑一系列人工示教数据，一般采用机器学习的方法进行数据编码和模型表达。如 GMM, GPR, HMM, SVR 等，用于建立式(5.3)或式(5.4)的映射模型。

如图5.1所示，基于示教数据可训练得到运动和力的规划器，并由此生成参考信号。主动柔顺控制器的类型可根据任务的不同灵活选取，控制器的参数可以直接从示教数据中学习得到，并基于手臂表面肌电信号进行交互刚度估计。本章采用 GMM 构建位置到末端操作刚度的映射关系，将其作为机械臂控制器的阻抗参数。

(1) 力/位混合控制

打磨和装配任务中，一般在某些方向上需要控制位置，其他方向上需控制接触力，目前普遍采用的方法为力/位混合控制。力/位混合控制的原理是指在机械臂末端沿着约束方向上进行力控制，与约束方向相正交的方向上进行位置控制，通过设置力和位置控制律对力和位置分别来控制。

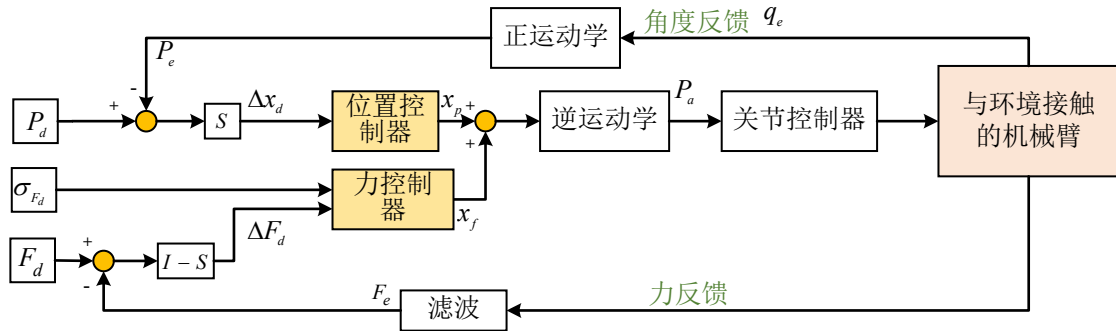


图 5.3 自适应力/位混合控制

如图5.3所示，考虑基于位置控制的机械臂，控制率 $\mathbf{X}_c$ 为包含位置控制器的输出 $\mathbf{x}_p$ 和力控制器的输出 $\mathbf{x}_f$ 。本章中位置控制器采用 PID 控制器，力控制器采用自适应 PI 控制器。所有的坐标系标记于图5.5(a)中，包括机械臂的基坐标系 $\{B\}$ ，力传感器的坐标系 $\{F\}$ ，机械臂末端执行器的坐标系 $\{E\}$ ，以及工件坐标系 $\{W\}$ 。力和力矩的反馈信号由 $\{F\}$ 变换至机械臂末端接触点 $\{E\}$ ，该变换包含旋转矩阵 ${}^E_F \mathbf{R} \in \mathbb{R}^{3 \times 3}$ 和平移向量 ${}^E_F \mathbf{P} \in \mathbb{R}^3$ 。忽略惯性力，末端坐标系 $\{E\}$ 下的接触力为：

$$\mathbf{F}_e = {}^E_F \mathbf{T} \mathbf{F}_f, \quad (5.5a)$$

且有

$${}^E_F\mathbf{T} = \begin{bmatrix} {}^E_F\mathbf{R} & \mathbf{0}_{3 \times 3} \\ \mathbf{P}_t {}^E_F\mathbf{R} & {}^E_F\mathbf{R} \end{bmatrix}, \quad (5.5b)$$

其中 ${}^E_F\mathbf{T} \in \mathbb{R}^{6 \times 6}$ 为力的变换矩阵. $\mathbf{P}_t$ 为向量 ${}^E_F\mathbf{P}$ 的反斜对称矩阵。

因此，位置控制量 $\mathbf{X}_c$ 为：

$$\mathbf{X}_c = \mathbf{X}_p + \mathbf{X}_f, \quad (5.6a)$$

$$\mathbf{X}_f = \mathbf{K}_p^f \Delta \mathbf{F}_d + \mathbf{K}_i^f \int \Delta \mathbf{F}_d dt, \quad \mathbf{K}_p^f = \mathbf{K}_p^{f_0} / \sigma_{F_d}, \quad (5.6b)$$

$$\Delta \mathbf{F}_d = \mathbf{S}(\mathbf{F}_d - \mathbf{F}_e), \quad (5.6c)$$

其中 $\mathbf{X}_f \in \mathbb{R}^6$ 为跟踪期望力 $\mathbf{F}_d \in \mathbb{R}^6$ 的位置调整量。 $\mathbf{K}_p^{f_0}, \mathbf{K}_i^f \in \mathbb{R}^{6 \times 6}$ 为 PI 控制器的参数。 σ_{F_d} 表示由小结5.3.3计算出的力分布的标准差。 $\mathbf{S}$ 为基于任务约束 k 生成的方向选择矩阵。实际中精确的力跟踪难以实现，也没有必要，从人的示教数据中可发现，只需要保证接触力在一定的范围即可。本节设计了自适应 PI 控制器用于力的跟踪，较小的 σ_{F_d} 意味着对力的要求较高，因此可采用式(5.6b)计算出较大的比例系数 $\mathbf{K}_p^f$ ，以实现期望力的快速跟踪。

(2) 自适应阻抗控制

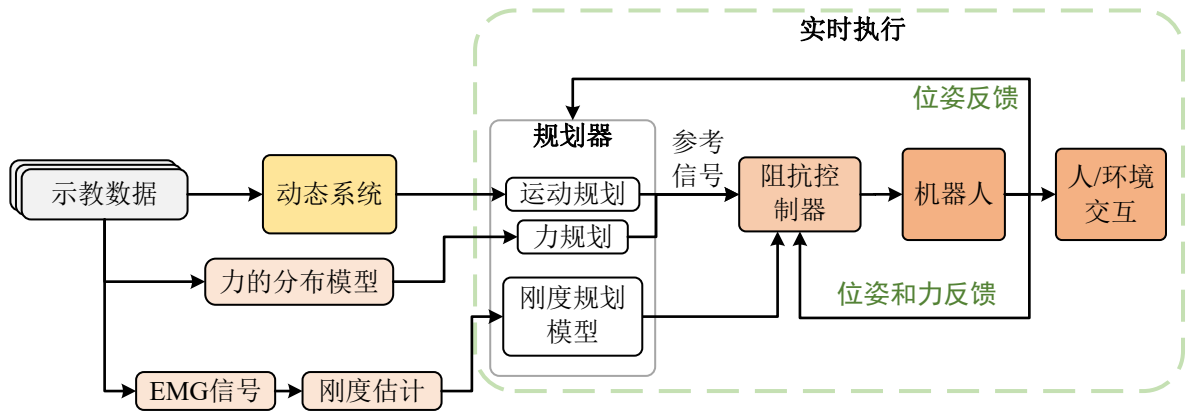


图 5.4 自适应阻抗控制

实际生活化场景中，大多是接触任务并不需要精确地进行位置和力的跟踪，如擦桌子，开门，开抽屉，推箱子，擦玻璃等等。只需要保持接触过程中的柔顺性并规划合适的轨迹进行运动即可，本节提出如图5.4所示的自适应阻抗控制器，用于从人的经验中学习交互刚度参数，并指导机械臂控制器的设计。

其中，运动规划方法采用第4章中的 DS，可采用 GMM 对示教数据建模，构建位姿到速度的映射模型。力规划通过 GMM 构建位姿到力的概率分布模型。刚度规划模

型也可将其描述为位姿到刚度的映射关系。从而，整个规划器均为仅基于机械臂状态量的反馈规划方法，自身便具有对扰动和对动态环境的适应性。

5.4 机打磨实验研究

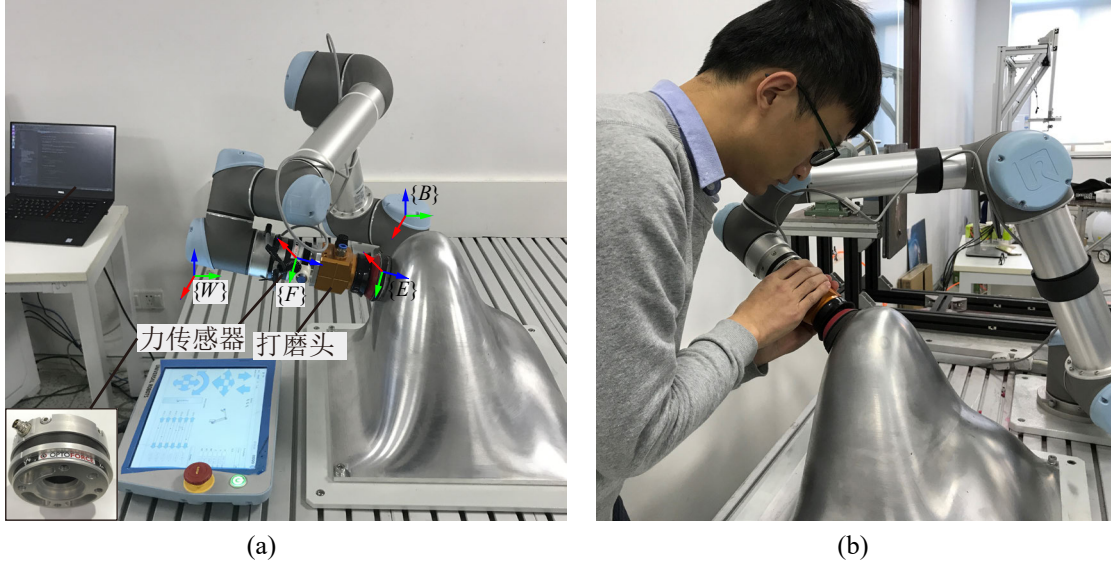


图 5.5 打磨实验平台

本节中，设计了典型的工业场景下的机械臂打磨任务，以验证提出的基于模仿学习的接触式任务框架的有效性。实验采用 UR5 机械臂和 Optoforce 的 HEX-70-XE-200N 六维力/力矩传感器。基于 ROS 进行机械臂的控制，UR 机械臂的 ROS 驱动采用 Andersen 的驱动程序^[162]。本节进行打磨任务的实验，实验平台如图 5.5(a) 所示。其中待打磨构件为 6061 铝合金材质，整体尺寸为 400 mm × 400 mm × 300 mm。六维力/力矩传感器安装于 UR 机械臂末端，打磨头安装于力传感器上。

5.4.1 打磨任务示教和数据采集

针对打磨任务，采用动觉示教的方式实现数据采集，示教者可以直接拖动打磨头以便于采集数据。尽管 UR 机械臂自身具有力控模式，其关节力矩由双编码器差值来估计，精度较低，一般需较大的力进行拖动示教，难以实现柔顺示教。为保持示教过程中打磨头能在示教者的控制下打磨工件表面，本小结设计了基于 PI 跟踪控制器和导纳控制器的示教算法，如图 5.6 所示。

UR 机械臂采用速度控制模式，其计算式为：

$$\dot{q} = J^{-1} S_P V_c, \quad (5.7a)$$

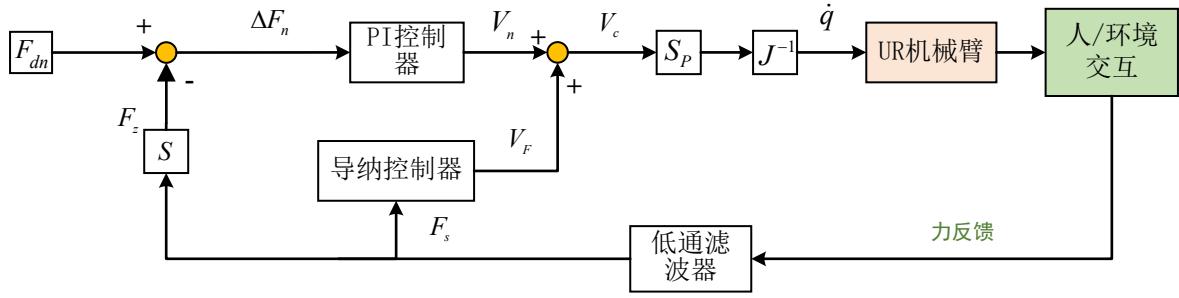


图 5.6 打磨动觉示教算法

$$\mathbf{V}_c = \begin{bmatrix} 0 & 0 & \mathbf{V}_n & 0 & 0 & 0 \end{bmatrix}^T + \mathbf{V}_F, \quad (5.7b)$$

其中:

$$\mathbf{V}_n = \mathbf{K}_p^D \Delta F_n + \mathbf{K}_i^D \int \Delta F_n dt, \quad (5.7c)$$

$$\mathbf{M}_a^D \dot{\mathbf{V}}_F + \mathbf{D}_a^D \mathbf{V}_F = \mathbf{F}_s, \quad (5.7d)$$

其中 $\dot{\mathbf{q}} \in \mathbb{R}^6$ 为机械臂的关节角速度, $\mathbf{J} \in \mathbb{R}^{6 \times 6}$ 为机械臂正运动学 ($\{E\}$ 相对于 $\{B\}$) 的雅可比矩阵。 $\mathbf{S}_P \in \mathbb{R}^{6 \times 6}$ 为打磨投影矩阵。 $\mathbf{V}_c \in \mathbb{R}^6$ 为期望的笛卡尔空间速度, $\mathbf{F}_s \in \mathbb{R}^6$ 为坐标系 $\{E\}$ 下的接触力。采用 Butterworth 低通滤波器对力传感器的反馈数据进行滤波。 $\mathbf{V}_n \in \mathbb{R}$ 为使得打磨头与曲线恒力接触的速度调整量。 $\mathbf{V}_F \in \mathbb{R}^6$ 由导纳控制器计算得出。 $\mathbf{K}_p^D, \mathbf{K}_i^D \in \mathbb{R}$ 为 PI 控制器中的比例和积分系数。 $\mathbf{M}_a^D, \mathbf{D}_a^D \in \mathbb{R}^{6 \times 6}$ 为导纳控制器的质量阻尼参数。

由于机械臂运动平缓,可忽略惯性力。对打磨头末端进行受力分析可得 $\mathbf{F}_e = \mathbf{F}_s + \mathbf{F}_h$, 其中 $\mathbf{F}_h \in \mathbb{R}^6$ 为示教者施加于末端执行器得力, $\mathbf{F}_e \in \mathbb{R}^6$ 为工件的接触反力。机械臂的运动数据以时间序列的形式记录下来,并在每次示教之后让机械臂回放之前的运动,从而有 $\hat{\mathbf{F}}_e = \hat{\mathbf{F}}_s = \mathbf{F}_e$, 并将具有人工经验的接触力数据和位置数据记录下来。

采用式5.7和图5.6所示的控制率进行动觉示教。控制器参数为 $\mathbf{K}_p^D = 2.5 \times 10^{-4}$, $\mathbf{K}_i^D = 1.9 \times 10^{-4}$, $\mathbf{M}_a^D = \mathbf{I}_{6 \times 6} \times 10^4$, $\mathbf{D}_a^D = 5\mathbf{I}_{6 \times 6} \times 10^3$ 。为学习打磨图5.5(a)所示的工件曲面力的分布,选择 18 个平面对工件进行 20° 等间隔分割,从而生成投影矩阵 $\mathbf{S}_P$ 。人工拖动示教打磨过程如图5.7(a)所示,记录机械臂的关节运动数据,并重放该运动(图5.7(b)),接触力数据 $F_p \in \mathbb{R}$, 打磨接触中心点位置 $\mathbf{x}_p \in \mathbb{R}^3$ 和速度 $\dot{\mathbf{x}}_p \in \mathbb{R}^3$ 以 125 Hz 进行记录,每条轨迹示教 5 次。

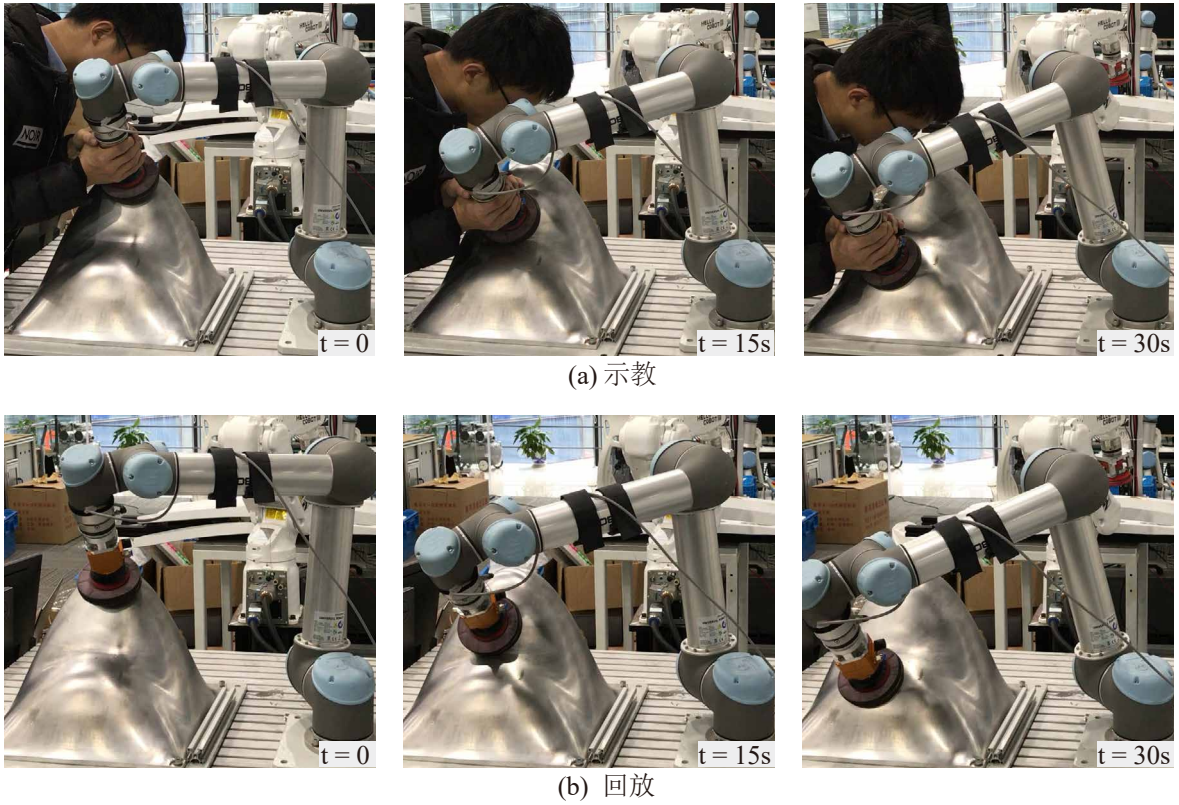


图 5.7 打磨示教过程

5.4.2 打磨技能学习

首先，需对示教数据进行预处理，采用动态时间规整 (Dynamical Time Wrapping, DTW) 对齐某个打磨路径的 5 次示教数据。采用包含 N 个数据点的位置、速度和力的训练数据 $\{x_p^i, \dot{x}_p^i, F_p^i\}_{i=1}^N$ 。并定义任务完成度 $\alpha_i = i/N, \alpha_i \in [0, 1]$ 。取 $\{\alpha_i\}_{i=1}^N$ 为 GMM 的输出分量， $\{F_p^i\}_{i=1}^N$ 为其输出，并用 GMR 对力的分布进行回归，以学习其内在关联。单条打磨路径的映射结果如图 5.8 所示，首先将同一打磨任务重复 5 次，然后采用 5 个高斯模型来近似其联合概率分布 $p(\alpha, \hat{F})$ ，采用 EM 训练 GMM，图中绿色标记点表示每个高斯模型的均值，最后基于该 GMM 模型计算 GMR，其中蓝色区域表示接触力分布的 3σ 置信区域。接触力 $\sigma_{\hat{F}}$ 的标准差也为 α 的函数。由于针对每条示教路径， α_i 与 x_p 一一对应，因此实际上建立了映射函数 $f(x) = [\hat{v}, \hat{F}, \sigma_{\hat{F}}]$ ， $\hat{v} \in \mathbb{R}^N$ 为打磨速度。因此，共包含 18 组映射函数 $F() : x \rightarrow \{\hat{v}_a, \hat{m}F_a, \sigma_{\hat{F}_a}\}_{n=1}^{18}$ ，其中位置到力的映射策略 $x \rightarrow \hat{F}_a$ 如图 5.9(a) 所示，曲线上颜色表示在该点的接触力大小。最后，采用高斯过程回归算法得到整个曲面的接触力分布，如图 5.9(b) 所示。

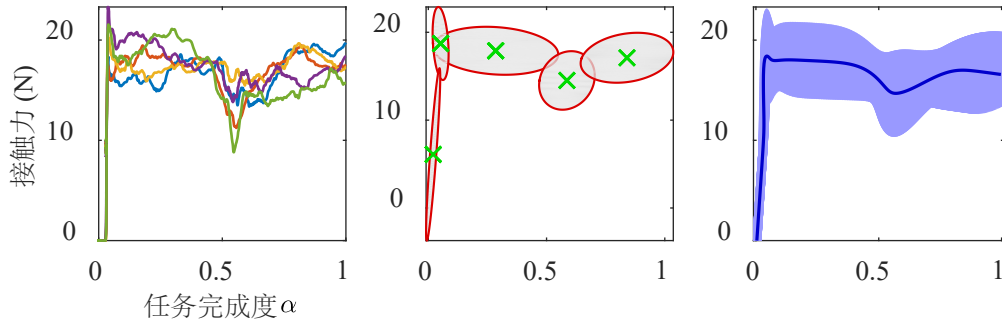


图 5.8 单条打磨路径的 GMM 和 GMR 学习

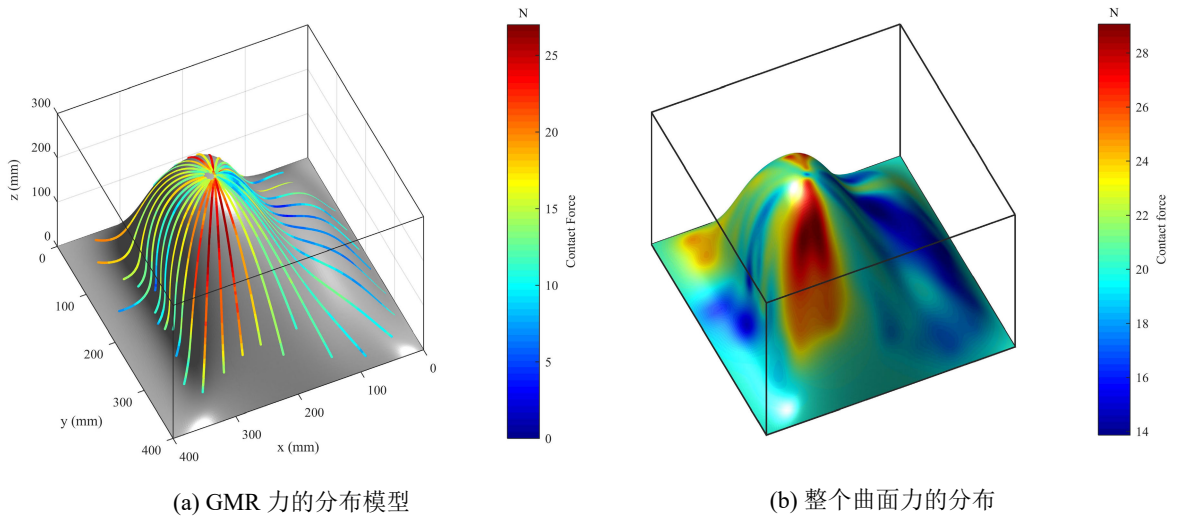


图 5.9 打磨曲面上位置到接触力的映射分布模型

5.4.3 打磨操作自主执行

在任务执行阶段，取如图5.10(a)所示的打磨路径，该路径可由人工拖动示教完成，也可由人工手动输入关键点来生成。基于前述的模型，计算这条路径上的位置点对应的速度 $\hat{v}_e$ ，接触力 $\hat{F}_d$ ，力的方差值 $\sigma_{\hat{F}_d}$ 。采用图5.3所示的自适应力/位混合控制器实现轨迹和接触力的跟踪。力控制器的参数选为 $\mathbf{K}_p^{f_0} = \text{diag}(1.6\mathbf{I}_{3 \times 3} \times 10^{-4}, \mathbf{0}_{3 \times 3})$, $\mathbf{K}_i^f = \text{diag}(3\mathbf{I}_{3 \times 3} \times 10^{-4}, \mathbf{0}_{3 \times 3})$ ，选择矩阵 $\mathbf{S}^p = \text{diag}(0, 0, 1, 0, 0, 0)$ 表示该点所在工件表面的法线方向。打磨过程如图5.11所示，每隔 5 s 一张截图。其中，某次打磨过程力控跟踪结果如图5.10(b)所示，蓝色区域为接触力概率分布的 3σ 区间，黑色曲线为参考力，其他彩色曲线表示打磨过程的接触力时间变化过程，在 90.3 % 的时间内接触力都在满足要求的范围内，整个过程力控跟踪的误差为 1.58 N，满足打磨要求。从人的示教数据中可以看出，打磨中只需保证接触力在一定的范围内即可，无需实现高速、高精密的力控跟踪。实验视频见¹。

¹实验视频<https://www.bilibili.com/video/BV165411g74K/>

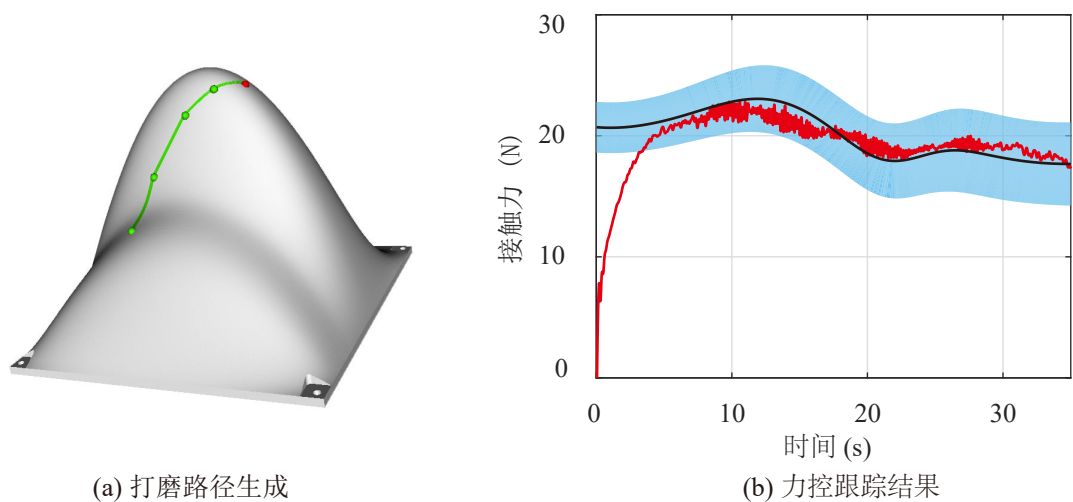


图 5.10 打磨路径生成和力控跟踪曲线

5.5 轴孔装配实验研究

5.5.1 装配实验平台

表 5.1 不同尺寸轴

	轴 1	轴 2	轴 3
公差	h6	h6	h6
尺寸 mm	$\phi 10_{-0.009}^0$	$\phi 8_{-0.009}^0$	$\phi 12_{-0.011}^0$

表 5.2 不同尺寸/公差的孔

	孔 1	孔 2	孔 3	孔 4	孔 5	孔 6
公差	H7	F7	E7	D7	H7	H7
尺寸 mm	$\phi 10_0^{+0.015}$	$\phi 10_{+0.013}^{+0.028}$	$\phi 10_{+0.025}^{+0.04}$	$\phi 10_{+0.04}^{+0.055}$	$\phi 8_0^{+0.015}$	$\phi 12_0^{+0.018}$

为评估所提出的基于模仿学习的接触式任务操作技能学习框架，设计了多组轴孔装配实验。本节设计了一系列不锈钢的轴和孔，每组具有不同的尺寸和公差配合精度，如图 Fig. 5.12(a) 所示，轴的尺寸见表5.1，孔的尺寸如表5.2所示，孔深均为 30 mm。实验平台见图5.12(b)，本节设计了一个简单的 RCC 装置，采用弹簧和直线轴承实现轴线方向的被动柔性。力传感器固定于 UR3 机械臂末端，并依次连接 RCC 装置和轴，孔固定于工作台上。

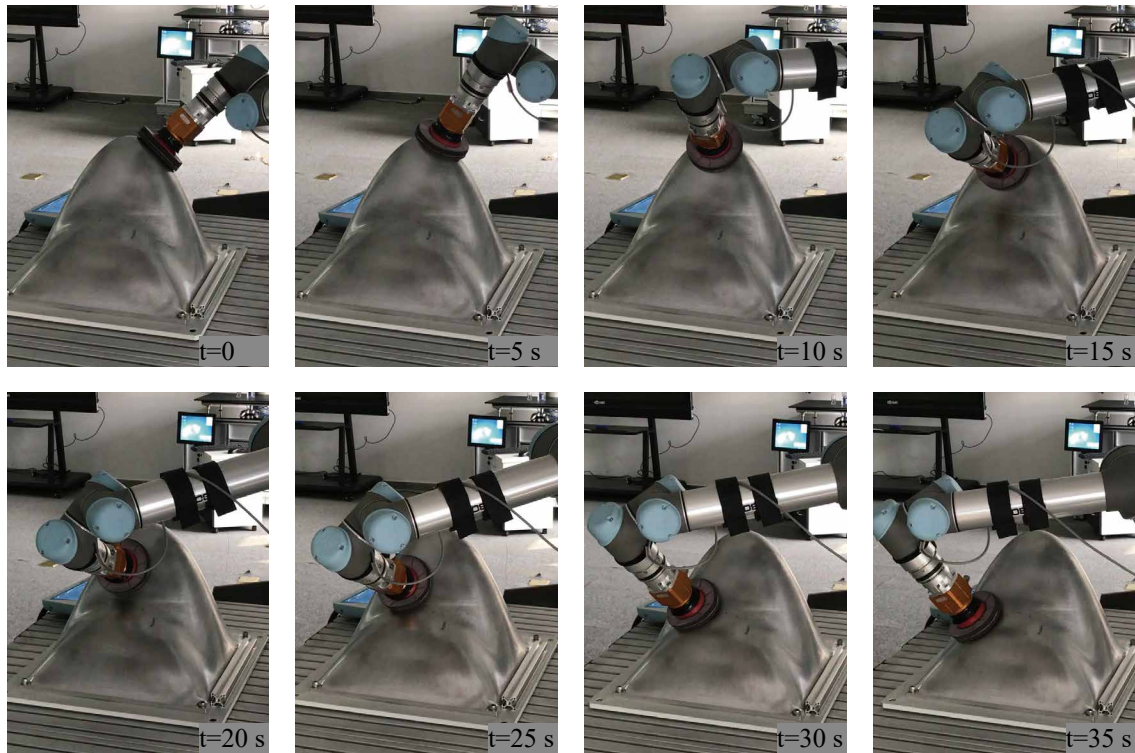


图 5.11 自主力控打磨过程

5.5.2 轴孔装配数据采集

取轴 1 和孔 1 用于示教数据采集，并搭建如图 5.13 所示的装配数据采集平台。设计示教手柄并粘贴 5 个光学运动捕捉系统的标记点，并将力传感器安装于示教手柄末端，轴固定于力传感器末端，此外，将孔固定于安川 MH5F 机械臂末端。在每次示教开始前，安川机械臂将在 x 方向和 y 方向生成随机角度偏差值 $\delta\theta_x, \delta\theta_y \in [-20^\circ, 20^\circ]$ ，其均一分布。根据 5 个标记点的实时位置可计算出轴的位置，并将其表达于孔所在的坐标系中，同时，将末端的重力进行实时补偿，以使得力传感器测量的接触力/力矩为轴孔之间的接触力/力矩。

示教过程中采集轴相对于孔的位姿，接触力，并计算角速度值，某次示教的数据如图 5.14 所示。分析人完成该任务时的数据，可初步将其分为以下 5 个阶段：

- (a) 轴接近孔的平面，z 方向相对位置不断减小，直至 z 方向力突变为非零值；
- (b) 轴与孔平面接触，z 方向距离基本不变，采用螺旋形扫描寻孔，z 方向依然保持一定的接触力；
- (c) z 方向的力突然减小，代表寻孔完成，并开始调整其姿态，使得角度误差减小，该阶段角速度值波动最大；
- (d) 缓慢调节轴的位姿，同时保持接触力/力矩不至于过大，使轴能顺利向下运动，z 方向的位置偏差继续减小；

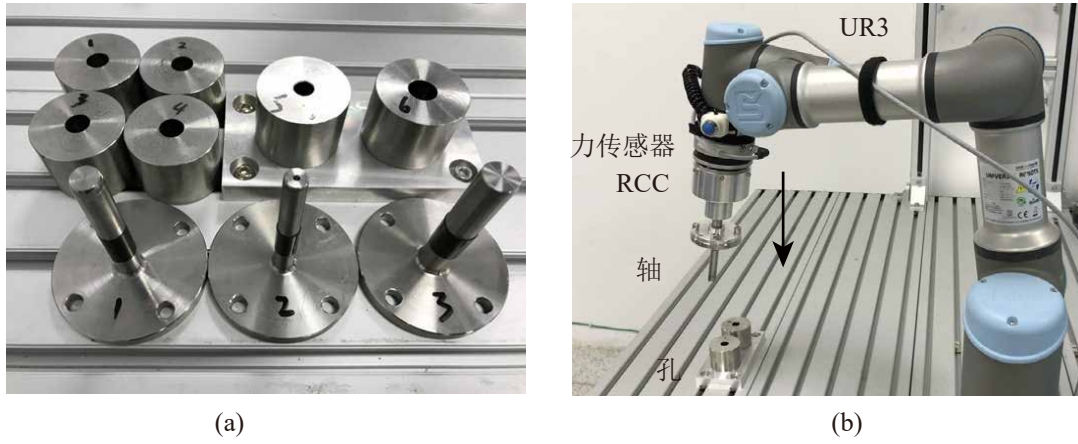


图 5.12 轴孔装配实验平台

(e) 完成装配任务。

可以看出，步骤 (b)-(e) 中，始终存在一定的接触力，且无需为恒定值，步骤 (c) 和 (d) 中姿态的调整策略最为关键。人在无视觉反馈的条件下可轻松完成该装配任务，可能的策略在于人能根据感知到的接触力信息来分析当前所处的接触姿态，并进行姿态调整。本节采集 20 组示教数据为后续装配技能模型做准备。

5.5.3 装配调整策略学习

采用 GMM 构建交互力矩 $\mathbf{M} \in \mathbb{R}^{2 \times N}$ 到末端角速度 $\boldsymbol{\omega} \in \mathbb{R}^{2 \times N}$ 之间的映射关系。图5.15(a) 为单次示教中 x 方向力矩和 x 方向角速度的关系。图5.15(b) 展示了 GMM 的分布模型，为 20 次示教数据经 GMM 建模的 x 方向力矩与其角速度的映射模型，绿色标记点为其均值点，红色椭圆表示 GMM 的 3σ 置信区间。根据式(2.16a)和该 GMM 模型，角速度的条件概率 $\hat{\boldsymbol{\omega}} \sim p(\boldsymbol{\omega}|\mathbf{M})$ ，结合式(2.18)，可计算角速度期望值。

5.5.4 轴孔装配自主执行

位置和姿态的调整策略是轴孔装配成功的关键，采用图5.3中的力/位混合控制进行控制率的分解。此外，基于前述的 GMM 设计姿态的调整策略。轴的期望姿态由角速度 $\boldsymbol{\omega}_a = [\omega_x, \omega_y, \omega_z]^T$ 计算，且有：

$$\mathbf{R}(t + \Delta t) = (\Delta t \mathbf{S}(\boldsymbol{\omega}_a) + \mathbf{I}_{3 \times 3}) \mathbf{R}(t), \quad (5.8a)$$

其中 $\mathbf{S}(\boldsymbol{\omega}_a)$ 为反斜对称矩阵：

$$\mathbf{S}(\boldsymbol{\omega}_a) = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}, \quad (5.8b)$$

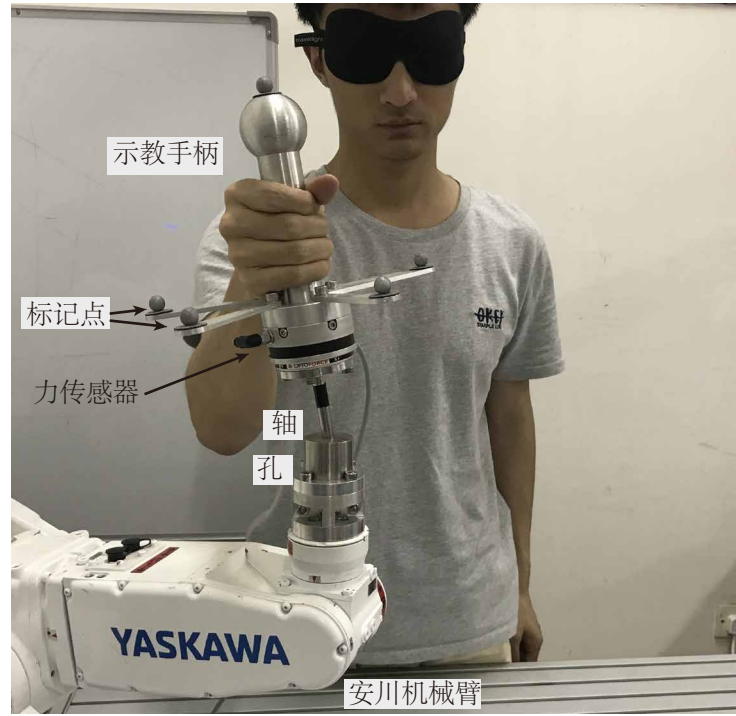


图 5.13 轴孔装配数据采集平台

其中, Δt 为控制周期, $I_{3 \times 3}$ 为单位矩阵, 为对比本节的方法和其他方法的不同, 设计了四种控制器:

- A: 角速度服从均一分布 $\hat{\omega} \sim \mathcal{U}(\mathbf{a}, \mathbf{b})$, 其中 $\mathbf{a} = \mathbf{0}$, $\mathbf{b} = [\sigma_x, \sigma_y]^T$ 由 GMM 的方差矩阵确定。
- B: 角速度服从高斯分布 $\hat{\omega} \sim \mathcal{N}(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B)$, 且 $\boldsymbol{\mu}_B = [0, 0]^T$, $\boldsymbol{\Sigma}_B = [0.1, 0.1]^T$ 为常数。
- C: 角速度 $\hat{\omega} \in p(\omega|\mathbf{M})$ 为基于实时感知力矩的 GMM 后验概率。
- D: 角速度 $\hat{\omega} = E\{p(\omega|\mathbf{M})\}$, 即 GMR。

自主执行操作包含以下四个步骤:

- 移动至孔平面, 如图5.16(a)所示;
- 采用阿基米德螺旋线搜孔, 同时保持恒力接触;
- 大范围姿态调整。设计力矩控制器 $\hat{\omega} = -3\boldsymbol{\Sigma}_B \text{sign}(\mathbf{M})$ 以减小法向力带来的力矩, 直到接触力矩小于阈值 $M_0 = 0.1\text{Nm}$, 对应图5.16(c)过程;
- 基于前述的四种控制器, 小范围调整姿态, 如图5.16(e-g), 其中根据任务完成度 α 设计衰减函数以减小接触力。因此, 参考角速度为 $\hat{\omega} = |\gamma(\alpha)\hat{\omega}| \text{sign}(\mathbf{M})$, 并根据式(5.8)计算期望姿态。

其中, 力矩控制器在第三和第四步中实现柔顺接触。其中位置和姿态均基于当前的力和力矩反馈进行计算。设计期望的接触力 $F_d = [0, 0, F_z^0]^T$ 对轴施加一定的装配力, 并

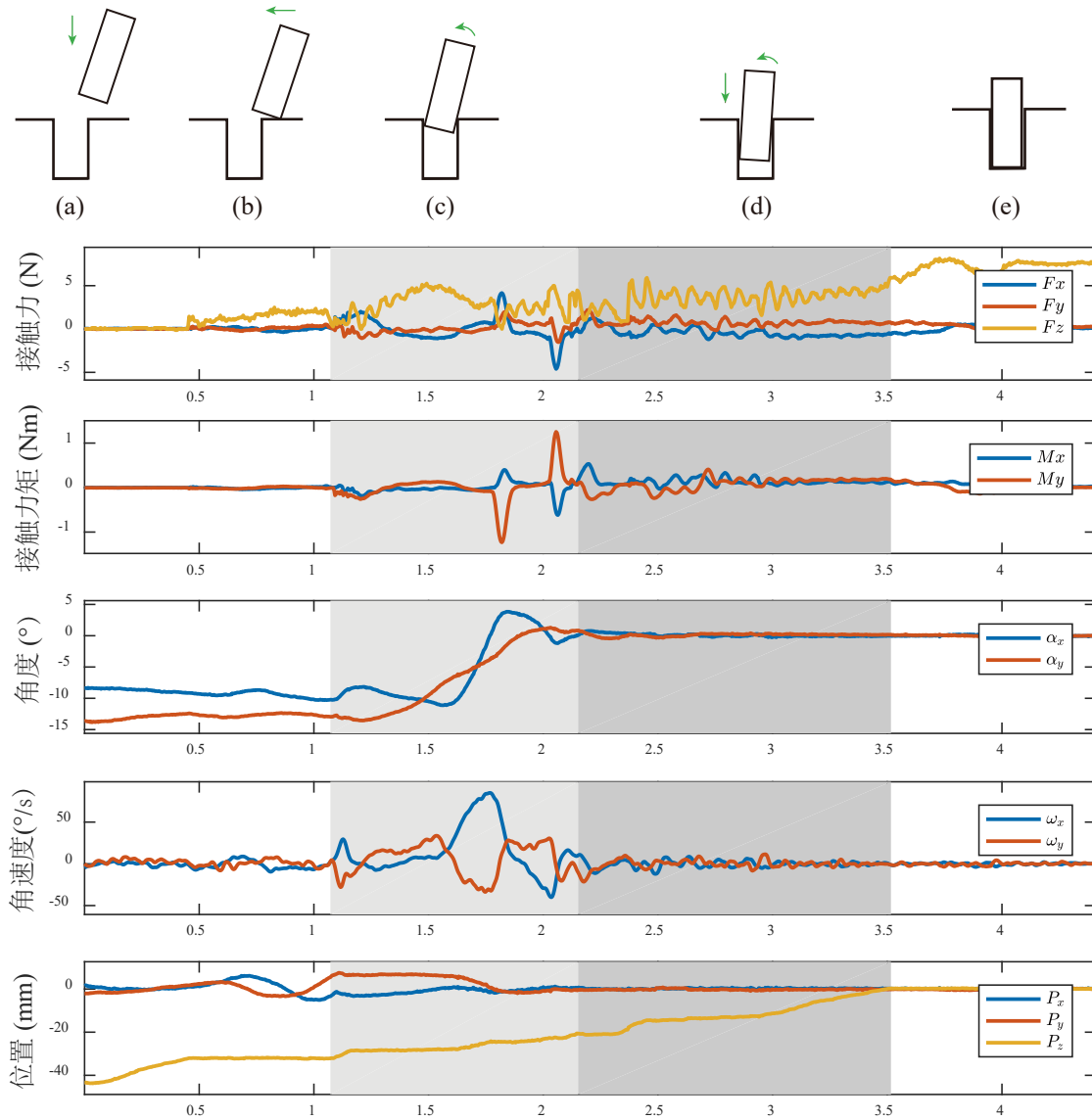


图 5.14 某次轴孔装配示教数据

采用 PI 控制器实现接触力的跟踪, 该控制器将减小 x 和 y 方向的接触力并保持 z 方向接触力的恒定, 同时, 力矩控制器由上述的四种控制器来实现, 用于调整轴的姿态。PI 控制器的参数为 $\mathbf{K}_p^a = \text{diag}(5, 5, 8) \times 10^{-5}$, $\mathbf{k}_i^a = \text{diag}(2, 2, 3) \times 10^{-5}$, $\mathbf{S} = \mathbf{0}_{6 \times 6}$, 即所有方向均采用力控制。

本节设计了两组轴孔装配实验, 第一组测试上述四种控制器在相同的轴和孔下的性能, 第二组实验测试本节提出的基于 GMM 的控制器 C 在不同的场景下的泛化能力, 包括不同尺寸的装配任务, 以及多种公差配合场景的装配。每组实验均测试 25 次, 并设计等间距采用的姿态误差 $[-20^\circ, 20^\circ]$ 和位置误差 $[-6 \text{ mm}, 6 \text{ mm}]$ 。注意表 5.2 中, 孔 1-4 具有相同的基准尺寸和不同的公差, 且轴 1 与孔 1, 轴 2 与孔 5, 轴 3 与孔 6 为具有相同公差等级但不同基准尺寸。

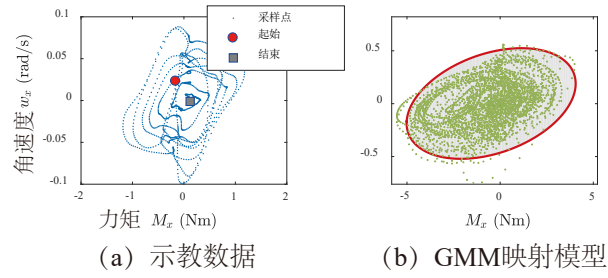


图 5.15 轴孔装配技能模型

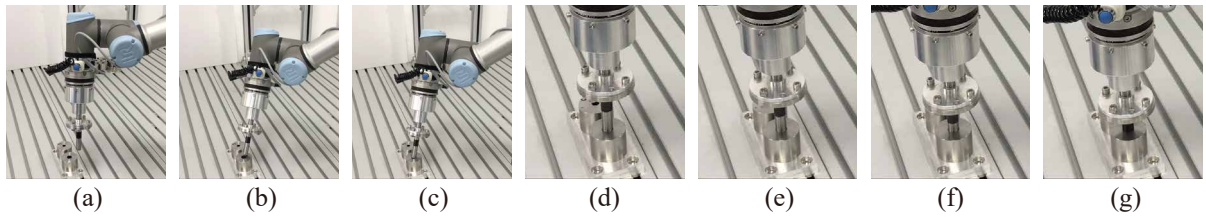


图 5.16 轴孔装配过程

轴孔装配的性能可由多个方面进行评价,包括成功率,平均任务完成时间,以及接触力的大小,平均能量消耗等。装配过程中接触力越小,时间越短,成功率越高,说明柔顺性越好。第一组的实验结果如表5.3所示,采用轴1和孔1进行实验,某次实验的装配过程如图5.16所示,实验视频见²。控制器C基于所学习的模型和实时力矩反馈,计算条件概率分布 $\hat{\omega} \in p(\omega|M)$ 。控制器A和B从静态的分布中采样角速度。从表5.3可以看出,控制器C实现了更高的成功率和较低的任务完成时间。其中,能量消耗值由 $W = \sum_{i=1}^n \int_0^T U_i I_i dt$ 计算, n 表示关节数量, U_i, I_i 分别为第 i 个关节的驱动电压和电流, T 为任务执行总耗时。控制器C实现了较低的能量消耗,这可能主要是由于较短的完成时间带来的收益。此外,控制器C的接触力均值和方差较大。控制器D采用GMR计算角速度,实验结果表面其成功率最低且任务完成时间最长。主要原因是轴和孔的配合误差仅为0.01mm左右,极容易导致装配卡死。由于控制器A、B、C中角速度由随机采样生成,因此可能更容易克服装配卡死的问题,随机的扰动有助于精密的轴孔装配。其他原因可能是力反馈的不精确或是学习模型中的建模误差。

对于第二组实验,结果如表5.4所示,包含6组实验,装配性能均比第一组实验中好,实验结果表明,轴1和孔1的装配技能模型可以迁移至不同尺寸和不同公差配合的轴孔装配中,证明了本章提出的基于模仿学习的接触式任务操作技能表达模型具有一定的泛化能力。

²实验视频<https://www.bilibili.com/video/BV165411g74K/>

表 5.3 不同控制器下 H7/h6 轴孔装配结果对比

	成功率 (%)	平均完成时间 (s)	接触力均值 (N) (标准差)	平均能量消耗 (J)
A	64	20.1	8.5 (4.5)	1847
B	56	30.1	7.0 (4.6)	2433
C	88	15.2	11.1 (6.2)	1483
D	40	38.5	4.1 (3.5)	2204

表 5.4 不同尺寸/公差配合轴孔装配结果对比

孔编号	成功率 (%)	平均完成时间 (s)	接触力均值 (N) (标准差)	平均能量消耗 (J)
1	88	15.2	11.0 (6.2)	1483
2	76	14.5	11.2 (6.3)	1413
3	80	15.4	10.8 (6.5)	1568
4	88	12.5	11.7 (6.4)	1336
5	88	14.1	11.1 (5.7)	1464
6	76	14.5	11.7 (6.1)	1531

5.6 变刚度操作实验研究

5.6.1 实验平台

本节进行变刚度擦白板实验，实验平台如图5.17所示，白板固定于桌面，用黑笔在白板上写字母，并控制机械臂夹持白板擦去擦除笔迹。

实验目标是一次擦拭清除所有的字母痕迹。可采用图(5.3)所示的力/位混合控制完成该任务，基本思路是并预先设定好参考轨迹，控制机械臂与白板法线方向的接触力在一定范围。然而，力/位混合控制实现该任务需要分为两步：1) 在自由空间中，需采用纯位置控制控制机械臂与白板接触；2) 接触后，采用力/位混合控制实现位置和力的跟踪。由于存在分段过程和控制率的切换，可能耗费更多时间完成该任务。



图 5.17 擦白板实验平台

第二种方法为采用自适应阻抗控制（图5.4）实现，可以实现单步操作，并基于章节4.3中基于 DS 的运动规划算法规划机械臂的运动。

5.6.2 实验结果分析

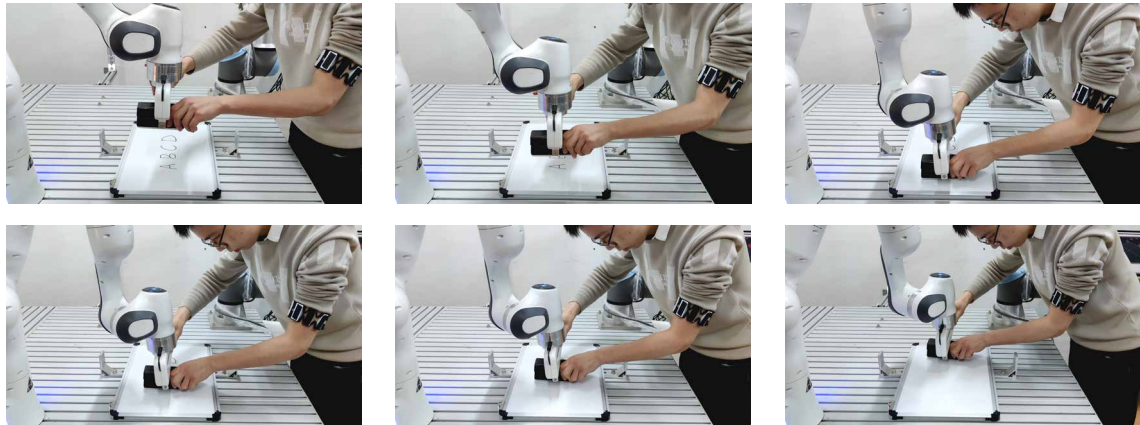


图 5.18 擦白板示教数据采集

采用式(4.29)对 Panda 机械臂进行零力控制，从而示教者可轻松地拖动机械臂完成擦白板任务。如图5.18所示，示教者拖动机械臂向下接触白板，然后从左往右擦除“ABCD”字迹。示教者在其大臂佩戴 MYO 手环用于测量表面肌电信号，并与机械臂末端轨迹同步记录数据，共记录三组。根据弹簧质量阻尼模型估计刚度值，绘制位置和刚度之间的变化曲线（图5.19中黑色曲线）。

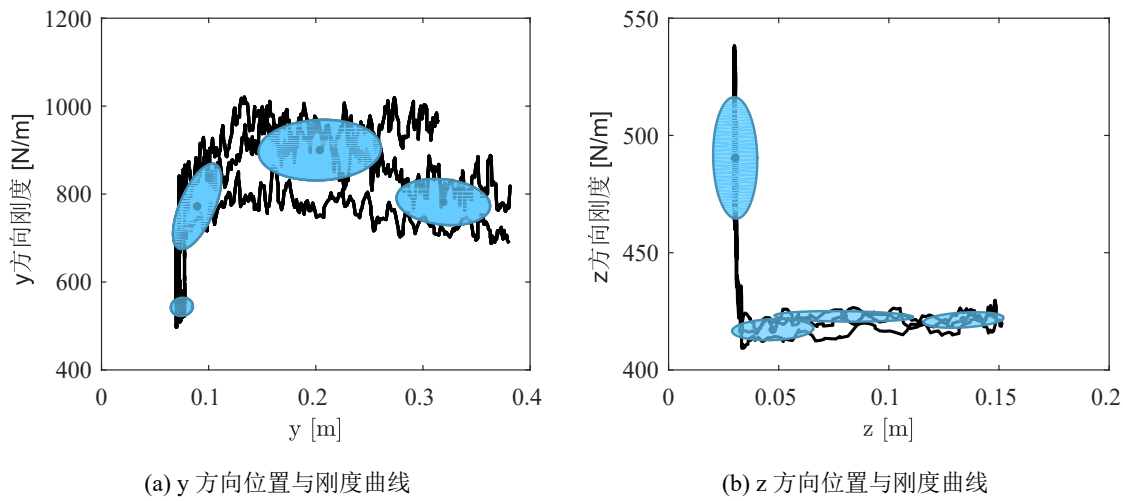


图 5.19 y 和 z 方向位置和刚度映射模型

图5.19绘制了 y 和 z 方向位置和刚度的关系。其中，z 方向为竖直方向（垂直于白板），y 方向为字母书写方向（示教者视角从左往右）。可以看出，在自由空间中 z 方向的刚度较小，当板擦与白板接触后，刚度开始急剧增大，y 方向的刚度也增大，当擦拭即将完成时，y 方向的刚度又会减小。采用 4 个高斯模型建立位置到刚度的映射模型，图5.19中的椭圆表示高斯模型的 σ 置信区间。

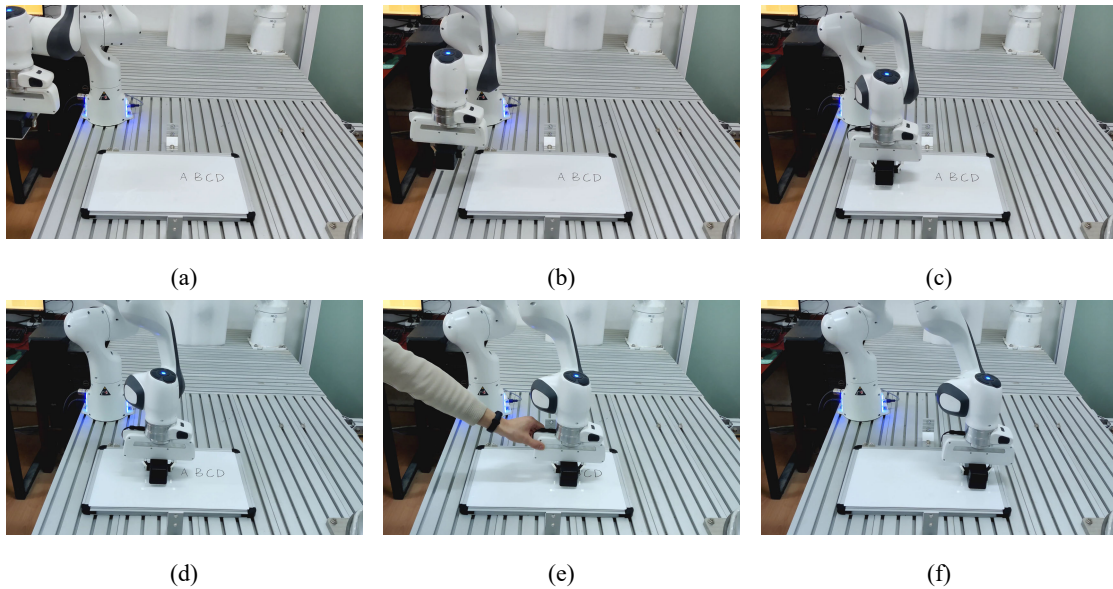


图 5.20 机械臂柔顺擦拭操作

将示教轨迹进行平均，并采用章节4.3中的方法构造 DS，将其用于机械臂的运动规划。根据图5.19所示的位置和刚度之间的映射关系，采用 GMR 获取其位置-刚度曲线，将其用于阻抗控制器的刚度在线调节。从而实现在非接触状态以较小的刚度运动，

接触并擦拭状态下增大刚度，以实现柔顺擦拭操作。

实验过程如图5.20所示，可以看出机械臂可以顺利擦除白板上的字母。实验视频见³。机械臂的运动轨迹如图5.21，可以看出，其运动轨迹平滑， z 方向的坐标逐渐减小，与白板接触后基本不变化。 y 方向的坐标不断增大，代表了擦拭方向。 x 方向基本无变化。图5.20(e)处，给机械臂施加扰动，仍能正常完成擦除任务，证明第4章提出的DS不仅适用于自由空间中的运动规划，也适用于接触式任务操作中的规划，且具有一定在线适应性。此外，验证了本章中提出的接触式任务柔顺操作框架，可在机械臂上实现变刚度的接触式柔顺操作。

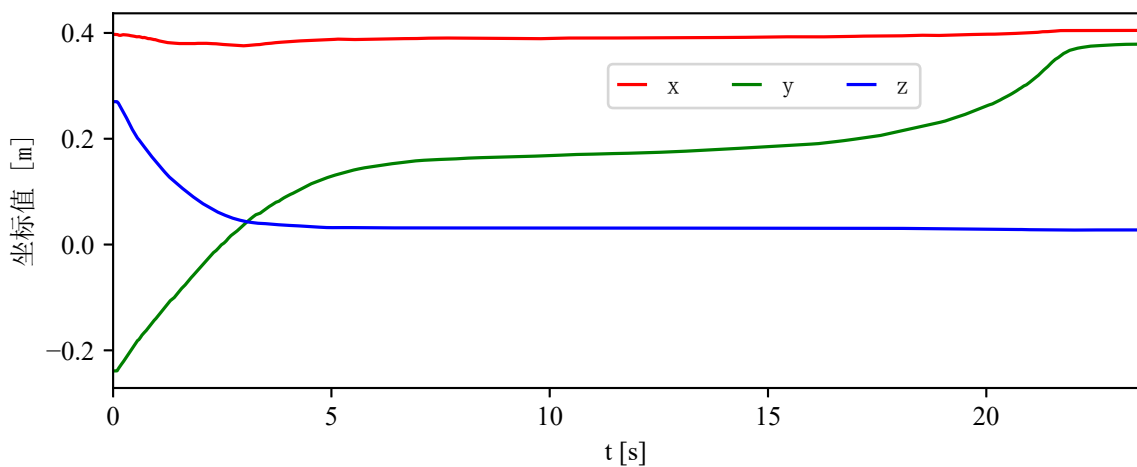


图 5.21 擦白板过程机械臂运动轨迹

最后，本节采用图5.3所示的力/位混合控制完成同样的擦拭任务，擦拭过程中法向接触力设置为 20 N。两种控制方法分别实验 10 次，并记录其任务完成所用时间。实验结果表明，两种方法均能完成该擦拭任务。变刚度阻抗控制器耗时 22 ± 4.8 s，力/位混合控制耗时 30 ± 5.3 s。可以看出，该任务采用变刚度阻抗控制效率更高。力/位混合控制由于需要进行精确的力和位置的跟踪，因此需将该任务分解为自由空间中的纯位置控制和接触操作过程的力/位置跟踪，耗时较长。

5.7 本章小结

本章研究柔顺操作技能表达和学习方法，提出了柔顺操作的技能表达模型，构建了接触式任务中位置、速度、交互力和交互刚度之间的关系。其次，提出了基于模仿学习的接触式任务操作技能学习框架，可用于指导常见的接触式交互任务，并设计了自适应力/位混合控制和自适应阻抗控制。随后，结合第4章所提出的DS和该柔顺操作

³实验视频<https://www.bilibili.com/video/BV1J64y117Cc/>

技能表达模型，设计了三个实验进行验证。1) 设计了机械臂打磨操作平台，设计融合导纳控制器和 PI 控制器的拖动示教算法，实现曲面上的拖动示教打磨数据采集，并采用 GMM 算法构建整个打磨曲面上位置到力的映射模型，最后进行打磨任务的自主执行，力轨迹跟踪误差仅为 1.58 N，且 90.3 % 的时间内接触力都在满足要求的范围内。2) 搭建了轴孔装配实验平台和人工示教数据采集平台，通过分析人完成该轴孔装配操作过程分析器内在规律，并基于 GMM 构建示教者的姿态调整策略模型，同时设计了三种其他的控制器进行对比，结果表明该姿态调整策略成功率最高，耗时短，且能适应不同尺寸和公差配合的轴孔装配，具有一定的泛化能力。3) 搭建了白板字迹擦拭平台，通过示教的方式采集人的轨迹和肌电信号数据，并建立位置到刚度的映射关系，并将其用于机械臂执行过程中的阻抗控制器设计，实现变刚度柔顺控制，实验结果表明，采用 DS 的轨迹规划算法效果良好，使用基于状态的变刚度控制器具有对时间、空间扰动和对动态环境的在线适应性。

6 总结与展望

6.1 总结

本文针对机械臂运动规划与柔顺控制算法展开研究,为提高机器人的易用性和智能性,提出基于模仿学习的运动规划与柔顺控制框架,应用于典型场景下机器人的自主操作,主要工作如下:

第3章中,提出了采用微分同胚映射构建 DS 的方法,并提出了满足约束条件的微分同胚映射算法框架。首先分析了所需的 DS 的约束要求,从而推导出微分同胚映射的约束条件,并设计了统一的微分同胚映射框架,包含位置姿态及其速度的映射关系。在该框架的基础上,提出了两种映射算法的实现形式,包括基于径向基函数的迭代算法,和基于可逆神经网络的映射算法。随后进行了二维空间仿真对比和三维空间中的遥操作实验,验证所提出的微分同胚算法的有效性。其中迭代算法训练速度快,但正反向映射估计耗时较长,可逆神经网络法训练耗时较长,但正反向映射计算效率高。

第4章中,提出了一种基于模仿学习和微分同胚映射的运动规划算法。首先,证明了一个全局渐进稳定的 DS 的微分同胚映射系统也是全局渐进稳定的。针对位置的 DS,在隐空间中设计简单的 DS 并将其映射至示教数据空间,针对姿态的 DS 设计了其角速度表达,随后提出了融合位置和姿态的 DS,并构造了 Lyapunov 函数证明其稳定性。为验证其稳定性和有效性,采用第3章中的映射算法,设计了二维空间中 DS 的仿真,仿真中,从任意一点以任意的初始姿态出发,都能收敛至零点。最后,在三维空间中设计了物体递交实验和动态抓取实验,测试随机扰动和动态环境对运动规划算法的影响,结果表明该方法构造的 DS 对时间和空间扰动的适应性好,且能适应动态变化的环境,证明了算法的有效性。

第5章中,针对接触式任务操作,提出了统一的技能表达方法,并进行实验验证。其中,构建了接触式任务中位置、速度、交互力和交互刚度之间的关系。为指导机器人自主执行时的柔顺控制器设计,采用将人的操作刚度迁移至机器人的方法。其次,提出了基于模仿学习的接触式任务操作技能学习框架,可用于指导常见的接触式交互任务。针对两种不同场景设计了两个自适应控制器:力/位混合控制和自适应阻抗控制。力/位混合控制中根据位置到力标准差的映射动态调节力控制器的增益参数,阻抗控制器中用建立人在执行该任务过程中位置到刚度的映射关系,用于指导其刚度参数设计。最后,设计了三个实验¹用于验证第4章和前述的方法。1) 机器人打磨实验,设计了打磨实验平台,并提出了一种接触式任务示教方法,采集了人执行打磨任务中的位置、速度和接触力,基于 GMM 构建位置到速度和接触力的映射概率模型。基于该模型实现机

¹ 本文所有的实验视频<https://space.bilibili.com/2157139/channel/detail?cid=177206>

机器人的自主打磨实验，多次打磨过程力控跟踪的均方根误差为 1.58 N，满足打磨要求，且在 90.3 % 的时间内接触力都在力分布的置信区间内。2) 轴孔装配实验，设计了装配实验平台和示教数据采集平台，通过采集并分析人执行该装配任务过程的位置、姿态、力/力矩数据，对状态进行划分并提出了轴孔装配过程中的姿态调整策略，在机器人自主执行阶段，该策略相比于其他算法成功率更高，且耗时短，随后进行了不同尺寸和公差配合的轴孔装配实验，结果表明该姿态调整策略可应用于不同的场景。3) 设计变刚度擦白板实验，通过采集人执行该任务过程的位姿、速度和表面肌电信号，建立位姿到速度和位姿到手臂刚度的映射，从而设计了基于 DS 的运动规划算法和变刚度阻抗控制器，实验结果表明该算法具有对扰动的适应性，且能适应不同的任务起点，效率较高。

6.2 展望

本文提出了基于模仿学习的机械臂运动规划与柔顺控制方法，在全局渐进稳定的 DS 设计和柔顺控制方面取得了一定的成果，为进一步提高机器人易用性和智能性，仍有许多工作有待进一步研究，主要包括：

(1) 在基于 DS 的运动规划方法中，本文只考虑了如何设计出全局渐进稳定且能精确表征示教数据的 DS，没有考虑示教空间中存在障碍物的情况，如何保证该 DS 全局渐进稳定的同时避开障碍物，甚至是动态的障碍物，是一项十分具有挑战性的工作。

(2) 本文虽然提出了基于微分同胚映射的 DS 设计方法，然而，尚未找到一种既能在短时间内完成训练，又能在正反向映射计算时保持高效的方法。高效的训练和预测有助于其快速迁移至不同的应用场景，并能保证能实时提供机器人的参考轨迹。

(3) 本文提出的基于模仿学习的接触式操作技能表达，仅仅是针对子单元任务的规划和控制，针对复杂的零件加工或装配等操作，需要首先进行任务规划，将复杂的任务序列拆分成多个小单元任务结构，并设计其组合方法。此类多任务长时间序列的表达和学习方法仍有待进一步研究。

参考文献

- [1] Erkaya S. Investigation of joint clearance effects on welding robot manipulators[J]. Robotics and Computer-Integrated Manufacturing, 2012, 28(4): 449-457.
- [2] Xie F G, Liu X J, Wu C, et al. A novel spray painting robotic device for the coating process in automotive industry[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2015, 229(11): 2081-2093.
- [3] Duque D A, Prieto F A, Hoyos J G. Trajectory generation for robotic assembly operations using learning by demonstration[J]. Robotics and Computer-Integrated Manufacturing, 2019, 57: 292-302.
- [4] Zheng C, Qin X, Eynard B, et al. SME-oriented flexible design approach for robotic manufacturing systems[J]. Journal of Manufacturing Systems, 2019, 53: 62-74.
- [5] Djuric A M, Urbanic R, Rickli J. A framework for collaborative robot (CoBot) integration in advanced manufacturing systems[J]. SAE International Journal of Materials and Manufacturing, 2016, 9(2): 457-464.
- [6] Bahrin M A K, Othman M F, Azli N H N, et al. Industry 4.0: A review on industrial automation and robotic[J]. Jurnal Teknologi, 2016, 78(6-13): 137-143.
- [7] Matheson E, Minto R, Zampieri E G, et al. Human-robot collaboration in manufacturing applications: a review[J]. Robotics, 2019, 8(4): 100.
- [8] Tian F J, Li Z G, Lv C, et al. Polishing pressure investigations of robot automatic polishing on curved surfaces[J]. The International Journal of Advanced Manufacturing Technology, 2016, 87(1): 639-646.
- [9] Chutima P. A comprehensive review of robotic assembly line balancing problem[J]. Journal of Intelligent Manufacturing, 2020: 1-34.
- [10] Siciliano B, Sciavicco L, Villani L, et al. Robotics: modelling, planning and control[M]. Springer Science & Business Media, 2010.
- [11] Chand S, Doty K L. On-line polynomial trajectories for robot manipulators[J]. The International Journal of Robotics Research (IJRR), 1985, 4(2): 38-48.
- [12] Dijkstra E W, et al. A note on two problems in connexion with graphs[J]. Numerische mathematik, 1959, 1(1): 269-271.

- [13] Hart P E, Nilsson N J, Raphael B. A formal basis for the heuristic determination of minimum cost paths[J]. IEEE transactions on Systems Science and Cybernetics, 1968, 4(2): 100-107.
- [14] Stentz A. Optimal and efficient path planning for unknown and dynamic environments[R]. CARNEGIE-MELLON UNIV PITTSBURGH PA ROBOTICS INST, 1993.
- [15] Likhachev M, Ferguson D, Gordon G, et al. Anytime search in dynamic graphs[J]. Artificial Intelligence, 2008, 172(14): 1613-1643.
- [16] LaValle S M, et al. Rapidly-exploring random trees: A new tool for path planning[J]. Iowa State University, 1998.
- [17] Kavraki L, Latombe J C. Randomized preprocessing of configuration space for path planning: Articulated robots[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). Vol. 3. 1994: 1764-1771.
- [18] Amato N M, Wu Y. A randomized roadmap method for path and manipulation planning[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). Vol. 1. 1996: 113-120.
- [19] Kavraki L E, Svestka P, Latombe J C, et al. Probabilistic roadmaps for path planning in high-dimensional configuration spaces[J]. IEEE transactions on Robotics and Automation, 1996, 12(4): 566-580.
- [20] Sucan I A, Moll M, Kavraki L E. The open motion planning library[J]. IEEE Robotics & Automation Magazine, 2012, 19(4): 72-82.
- [21] Quigley M, Conley K, Gerkey B, et al. ROS: an open-source Robot Operating System[C]// ICRA workshop on open source software. 2009.
- [22] Khatib O. Real-time obstacle avoidance for manipulators and mobile robots[G]// Autonomous robot vehicles. Springer, 1986: 396-404.
- [23] Yang S X, Meng M. An efficient neural network approach to dynamic robot motion planning[J]. Neural networks, 2000, 13(2): 143-148.
- [24] Parhi D R, Singh M K. Real-time navigational control of mobile robots using an artificial neural network[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2009, 223(7): 1713-1725.
- [25] Janglová D. Neural networks in mobile robot motion[J]. International Journal of Advanced Robotic Systems, 2004, 1(1): 2.

- [26] Wang M, et al. Fuzzy logic based robot path planning in unknown environment[C]// Proc. International Conference on Machine Learning and Cybernetics. Vol. 2. 2005: 813-818.
- [27] Pandey A, Sonkar R K, Pandey K K, et al. Path planning navigation of mobile robot with obstacles avoidance using fuzzy logic controller[C]// Proc. IEEE International Conference on Intelligent Systems and Control (ISCO). 2014: 39-41.
- [28] Parhi D R. Neuro-fuzzy navigation technique for control of mobile Robots[J]. INTECH-Mobile robot motion planning, 2008: 409-432.
- [29] Pradhan S K, Parhi D R, Panda A K. Neuro-fuzzy technique for navigation of multiple mobile robots[J]. Fuzzy Optimization and Decision Making, 2006, 5(3): 255-288.
- [30] Alajlan M, Koubaa A, Chaari I, et al. Global path planning for mobile robots in large-scale grid environments using genetic algorithms[C]// Proc. International Conference on Individual and Collective Behaviors in Robotics (ICBR). 2013: 1-8.
- [31] Hu Y, Yang S X. A knowledge based genetic algorithm for path planning of a mobile robot[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). Vol. 5. 2004: 4350-4355.
- [32] Tuncer A, Yildirim M. Dynamic path planning of mobile robots with improved genetic algorithm[J]. Computers & Electrical Engineering, 2012, 38(6): 1564-1572.
- [33] Zhang Y, Gong D w, Zhang J h. Robot path planning in uncertain environment using multi-objective particle swarm optimization[J]. Neurocomputing, 2013, 103: 172-185.
- [34] Roberge V, Tarbouchi M, Labonté G. Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning[J]. IEEE Transactions on industrial informatics, 2012, 9(1): 132-141.
- [35] Qin Y Q, Sun D B, Li N, et al. Path planning for mobile robot using the particle swarm optimization with mutation operator[C]// Proc. the International Conference on Machine Learning and Cybernetics. Vol. 4. 2004: 2473-2478.
- [36] Englot B, Hover F. Multi-goal feasible path planning using ant colony optimization[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2011: 2255-2260.
- [37] Iser R, Wahl F M. AntSLAM: global map optimization using swarm intelligence[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2010: 265-272.

- [38] Garcia M P, Montiel O, Castillo O, et al. Path planning for autonomous mobile robot navigation with ant colony optimization and fuzzy cost function evaluation[J]. Applied Soft Computing, 2009, 9(3): 1102-1110.
- [39] Geraerts R, Overmars M H. Creating high-quality paths for motion planning[J]. The International Journal of Robotics Research (IJRR), 2007, 26(8): 845-863.
- [40] Hauser K, Ng-Thow-Hing V. Fast smoothing of manipulator trajectories using optimal bounded-acceleration shortcuts[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2010: 2493-2498.
- [41] Boissonnat J D, Cérézo A, Leblond J. Shortest paths of bounded curvature in the plane[J]. Journal of Intelligent and Robotic Systems, 1994, 11(1): 5-20.
- [42] Segovia A, Rombaut M, Preciado A, et al. Comparative study of the different methods of path generation for a mobile robot in a free environment[C]// Proc. the International Conference on Advanced Robotics' Robots in Unstructured Environments. 1991: 1667-1670.
- [43] Saska M, Macas M, Preucil L, et al. Robot path planning using particle swarm optimization of Ferguson splines[C]// Proc. the IEEE Conference on Emerging Technologies and Factory Automation. 2006: 833-839.
- [44] 单伟, 孟正大. 基于改进 A* 算法的平滑路径设计[J]. 东南大学学报 (自然科学版), 2010, 40(S1): 155-161.
- [45] 邹细勇, 徐德, 李子印. 非完整移动机器人路径跟踪的模糊控制[D]. 2008.
- [46] Quinlan S, Khatib O. Elastic bands: Connecting path planning and control[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 1993: 802-807.
- [47] Rösmann C, Feiten W, Wösch T, et al. Trajectory modification considering dynamic constraints of autonomous robots[C]// Proc. German Conference on Robotics(ROBOTIK). 2012: 1-6.
- [48] Rösmann C, Feiten W, Wösch T, et al. Efficient trajectory optimization using a sparse model[C]// Proc. European Conference on Mobile Robots. 2013: 138-143.
- [49] Sideris A, Bobrow J E. An efficient sequential linear quadratic algorithm for solving nonlinear optimal control problems[C]// Proc. the American Control Conference. 2005: 2275-2280.

- [50] MAYNE B D. A Second-order Gradient Method for Determining Optimal Trajectories of Non-linear Discrete-time Systems[J]. International Journal of Control, 1966, 3(1): 85-95.
- [51] Todorov E, Li W. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems[C]// Proc. the American Control Conference. 2005: 300-306.
- [52] Ratliff N, Zucker M, Bagnell J A, et al. CHOMP: Gradient optimization techniques for efficient motion planning[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2009: 489-494.
- [53] Kalakrishnan M, Chitta S, Theodorou E, et al. STOMP: Stochastic trajectory optimization for motion planning[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2011: 4569-4574.
- [54] Pavlichenko D, Behnke S. Efficient stochastic multicriteria arm trajectory optimization[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2017: 4018-4025.
- [55] Park C, Pan J, Manocha D. ITOMP: Incremental trajectory optimization for real-time replanning in dynamic environments[C]// Proc. the International Conference on Automated Planning and Scheduling. Vol. 22. 2012.
- [56] Schulman J, Duan Y, Ho J, et al. Motion planning with sequential convex optimization and convex collision checking[J]. The International Journal of Robotics Research (IJRR), 2014, 33(9): 1251-1270.
- [57] Deits R, Tedrake R. Efficient mixed-integer planning for UAVs in cluttered environments[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2015: 42-49.
- [58] Ratliff N, Toussaint M, Schaal S. Understanding the geometry of workspace obstacles in motion optimization[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2015: 4202-4209.
- [59] Sutton R S, Barto A G. Reinforcement learning: An introduction[M]. MIT press, 2018.
- [60] Calinon S. Robot programming by demonstration[M]. EPFL Press, 2009.
- [61] Argall B D, Chernova S, Veloso M, et al. A survey of robot learning from demonstration[J]. Robotics and autonomous systems, 2009, 57(5): 469-483.

- [62] Ravichandar H, Polydoros A S, Chernova S, et al. Recent Advances in Robot Learning from Demonstration[J]. Annual Review of Control, Robotics, and Autonomous Systems, 2020, 3(1): 297-330.
- [63] Peters J, Schaal S. Reinforcement learning of motor skills with policy gradients[J]. Neural networks, 2008, 21(4): 682-697.
- [64] Theodorou E, Buchli J, Schaal S. Reinforcement learning of motor skills in high dimensions: A path integral approach[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2010: 2397-2403.
- [65] Theodorou E, Buchli J, Schaal S. A generalized path integral control approach to reinforcement learning[J]. The Journal of Machine Learning Research, 2010, 11: 3137-3181.
- [66] Schenck C, Tompson J, Levine S, et al. Learning robotic manipulation of granular media[C]// Proc. Conference on Robot Learning. 2017: 239-248.
- [67] Asada M, Noda S, Tawaratsumida S, et al. Purposive behavior acquisition for a real robot by vision-based reinforcement learning[J]. Machine learning, 1996, 23(2): 279-303.
- [68] Kroemer O B, Detry R, Piater J, et al. Combining active learning and reactive control for robot grasping[J]. Robotics and Autonomous systems, 2010, 58(9): 1105-1116.
- [69] Schulman J, Levine S, Abbeel P, et al. Trust region policy optimization[C]// Proc. Intl Conf. on Machine Learning (ICML). 2015: 1889-1897.
- [70] Lillicrap T P, Hunt J J, Pritzel A, et al. Continuous control with deep reinforcement learning[J]. ArXiv preprint arXiv:1509.02971, 2015.
- [71] Gu S, Lillicrap T, Sutskever I, et al. Continuous deep q-learning with model-based acceleration[C]// Proc. Intl Conf. on Machine Learning (ICML). 2016: 2829-2838.
- [72] Levine S, Finn C, Darrell T, et al. End-to-end training of deep visuomotor policies[J]. The Journal of Machine Learning Research, 2016, 17(1): 1334-1373.
- [73] Gu S, Holly E, Lillicrap T, et al. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2017: 3389-3396.
- [74] 李广创, 程良伦. 基于深度强化学习的机械臂避障路径规划研究[J]. 软件工程, 2019, 3(22): 12-15.

- [75] Höfer S, Bekris K, Handa A, et al. Perspectives on Sim2Real Transfer for Robotics: A Summary of the R:SS 2020 Workshop[Z]. 2020arXiv: 2012.03806 [cs.R0].
- [76] OpenAI, Akkaya I, Andrychowicz M, et al. Solving Rubik's Cube with a Robot Hand[J]., 2019arXiv: 1910.07113 [cs.LG].
- [77] Jeong R, Aytar Y, Khosid D, et al. Self-Supervised Sim-to-Real Adaptation for Visual Robotic Manipulation[Z]. 2019arXiv: 1910.09470 [cs.R0].
- [78] OpenAI, Andrychowicz M, Baker B, et al. Learning Dexterous In-Hand Manipulation[J]., 2019arXiv: 1808.00177 [cs.LG].
- [79] Levine S, Pastor P, Krizhevsky A, et al. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection[J]. The International Journal of Robotics Research (IJRR), 2018, 37(4-5): 421-436.
- [80] Kalakrishnan M, Righetti L, Pastor P, et al. Learning force control policies for compliant manipulation[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2011: 4639-4644.
- [81] Ota K, Jha D K, Oiki T, et al. Trajectory Optimization for Unknown Constrained Systems using Reinforcement Learning[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2019: 3487-3494. DOI: 10.1109/IROS40897.2019.8968010.
- [82] Chiang H T L, Hsu J, Fiser M, et al. RL-RRT: Kinodynamic motion planning via learning reachability estimators from RL policies[J]. IEEE Robotics and Automation Letters (RA-L), 2019, 4(4): 4298-4305.
- [83] Faust A, Oslund K, Ramirez O, et al. PRM-RL: Long-range robotic navigation tasks by combining reinforcement learning and sampling-based planning[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2018: 5113-5120.
- [84] Ha H, Xu J, Song S. Learning a Decentralized Multi-arm Motion Planner[J]., 2020arXiv: 2011.02608 [cs.R0].
- [85] Calinon S. A tutorial on task-parameterized movement learning and retrieval[J]. Intelligent Service Robotics, 2016, 9(1): 1-29.
- [86] Schneider M, Ertel W. Robot learning by demonstration with local gaussian process regression[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2010: 255-260.

- [87] Misimi E, Olofsson A, Eilertsen A, et al. Robotic handling of compliant food objects by robust learning from demonstration[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2018: 6972-6979.
- [88] Calinon S, Billard A. Recognition and reproduction of gestures using a probabilistic framework combining PCA, ICA and HMM[C]// Proc. Intl Conf. on Machine Learning (ICML). 2005: 105-112.
- [89] Calinon S. Mixture Models for the Analysis, Edition, and Synthesis of Continuous Time Series[M]// Mixture Models and Applications. Cham: Springer International Publishing, 2020: 39-57.
- [90] Calinon S, Guenter F, Billard A. On Learning, Representing, and Generalizing a Task in a Humanoid Robot[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2007, 37(2): 286-298. DOI: 10.1109/TSMCB.2006.886952.
- [91] Calinon S, Billard A. A probabilistic programming by demonstration framework handling constraints in joint space and task space[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2008: 367-372.
- [92] Calinon S, Alizadeh T, Caldwell D G. On improving the extrapolation capability of task-parameterized movement models[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2013: 610-616.
- [93] Alizadeh T, Calinon S, Caldwell D G. Learning from demonstrations with partially observable task parameters[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2014: 3309-3314.
- [94] Osa T, Esfahani A M G, Stolkin R, et al. Guiding trajectory optimization by demonstrated distributions[J]. IEEE Robotics and Automation Letters (RA-L), 2017, 2(2): 819-826.
- [95] Schneider M, Ertel W. Robot learning by demonstration with local gaussian process regression[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2010: 255-260.
- [96] Khansari-Zadeh S M, Billard A. Learning stable nonlinear dynamical systems with gaussian mixture models[J]. IEEE Transactions on Robotics, 2011, 27(5): 943-957.
- [97] Khansari-Zadeh S M. A dynamical system-based approach to modeling stable robot control policies via imitation learning[R]. Epfl, 2012.

- [98] Akgun B, Cakmak M, Jiang K, et al. Keyframe-based learning from demonstration[J]. International Journal of Social Robotics, 2012, 4(4): 343-355.
- [99] Calinon S, Pistillo A, Caldwell D G. Encoding the time and space constraints of a task in explicit-duration hidden Markov model[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2011: 3413-3418.
- [100] Kulić D, Takano W, Nakamura Y. Incremental learning, clustering and hierarchy formation of whole body motion patterns using adaptive hidden markov chains[J]. The International Journal of Robotics Research (IJRR), 2008, 27(7): 761-784.
- [101] Kulić D, Ott C, Lee D, et al. Incremental learning of full body motion primitives and their sequencing through human motion observation[J]. The International Journal of Robotics Research (IJRR), 2012, 31(3): 330-345.
- [102] Kruger V, Herzog D L, Baby S, et al. Learning actions from observations[J]. IEEE robotics & automation magazine, 2010, 17(2): 30-43.
- [103] Lee D, Nakamura Y. Stochastic model of imitating a new observed motion based on the acquired motion primitives[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2006: 4994-5000.
- [104] Lee D, Nakamura Y. Mimesis scheme using a monocular vision system on a humanoid robot[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2007: 2162-2168.
- [105] Aleotti J, Caselli S. Robust trajectory learning and approximation for robot programming by demonstration[J]. Robotics and Autonomous Systems, 2006, 54(5): 409-413.
- [106] Ijspeert A J, Nakanishi J, Shibata T, et al. Nonlinear dynamical systems for imitation with humanoid robots[C]// Proc. IEEE Intl Conf. on Humanoid Robots (Humanoids). 2001: 219-226.
- [107] Paraschos A, Daniel C, Peters J R, et al. Probabilistic Movement Primitives[C]// Burges C J C, Bottou L, Welling M, et al. Proc. Advances in Neural Information Processing Systems (NIPS). Vol. 26. Curran Associates, Inc., 2013: 2616-2624.
- [108] Ude A, Nemec B, Petrić T, et al. Orientation in cartesian space dynamic movement primitives[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2014: 2997-3004.

- [109] Huang Y L, Rozo L, Silvério J, et al. Kernelized movement primitives[J]. The International Journal of Robotics Research (IJRR), 2019, 38(7): 833-852.
- [110] Hersch M, Guenter F, Calinon S, et al. Dynamical system modulation for robot learning via kinesthetic demonstrations[J]. IEEE Transactions on Robotics, 2008, 24(6): 1463-1467.
- [111] Neumann K, Steil J J. Learning robot motions with stable dynamical systems under diffeomorphic transformations[J]. Robotics and Autonomous Systems, 2015, 70: 1-15.
- [112] Khansari-Zadeh S M, Billard A. Learning control Lyapunov function to ensure stability of dynamical system-based robot reaching motions[J]. Robotics and Autonomous Systems, 2014, 62(6): 752-765.
- [113] 金聪聪, 刘安东, Steven L, 等. 基于改进动态系统稳定估计器的机器人技能学习方法[J]. 自动化学报, 2020, 46(x): 1-11.
- [114] Saveriano M. An energy-based approach to ensure the stability of learned dynamical systems[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2020: 4407-4413.
- [115] Perrin N, Schlehuber-Caissier P. Fast diffeomorphic matching to learn globally asymptotically stable nonlinear dynamical systems[J]. Systems & Control Letters, 2016, 96: 51-59.
- [116] Uraïn J, Ginesi M, Tateo D, et al. ImitationFlow: Learning Deep Stable Stochastic Dynamic Systems by Normalizing Flows[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2020: 5231-5237.
- [117] Lemme A, Neumann K, Reinhart R F, et al. Neural learning of vector fields for encoding stable dynamical systems[J]. Neurocomputing, 2014, 141: 3-14.
- [118] Whitney D E, Rourke J M. Mechanical Behavior and Design Equations for Elastomer Shear Pad Remote Center Compliances[J]. Journal of Dynamic Systems, Measurement, and Control, 1986, 108(3): 223-232.
- [119] Gandhi P S, Bobade R S, Chen C. On novel compliant mechanisms for remote center motion[J]. Advances in Mechanical Engineering, 2018, 10(4).
- [120] Lee S. Development of a new variable remote center compliance (VRCC) with modified elastomer shear pad (ESP) for robot assembly[J]. IEEE Transactions on Automation Science and Engineering, 2005, 2(2): 193-197.

- [121] Fitzgerald C. Developing baxter[C]// Proc. IEEE Conference on Technologies for Practical Robot Applications (TePRA). 2013: 1-6.
- [122] Raibert M H, Craig J J. Hybrid Position/Force Control of Manipulators[J]. Journal of Dynamic Systems, Measurement, and Control, 1981, 103(2): 126-133.
- [123] Cheah C C, Kawamura S, Arimoto S. Stability of hybrid position and force control for robotic manipulator with kinematics and dynamics uncertainties[J]. Automatica, 2003, 39(5): 847-855.
- [124] Mazur A, Kaczmarek M, Ratajczak J, et al. Modified position-force control for a manipulator geometrically constrained by round obstacles[J]. Journal of Automation Mobile Robotics and Intelligent Systems, 2019, 13(2): 60-67.
- [125] 常健, 王亚珍, 李斌, 等. 基于力/位混合算法的 7 自由度机械臂精细操控方法[J]. 机器人, 2016, 38(5): 531-539.
- [126] 喻洋, 王耀兵, 魏世民, 等. 基于柔顺控制的机器人装配技术研究[J]. 北京邮电大学学报, 2020, 43(4): 1-6.
- [127] Hogan N. Impedance Control: An Approach to Manipulation: Part I—Theory[J]. Journal of Dynamic Systems, Measurement, and Control, 1985, 107(1): 1-7.
- [128] Hogan N. Impedance Control: An Approach to Manipulation: Part II—Implementation[J]. Journal of Dynamic Systems, Measurement, and Control, 1985, 107(1): 8-16.
- [129] Hogan N. Impedance Control: An Approach to Manipulation: Part III—Applications[J]. Journal of Dynamic Systems, Measurement, and Control, 1985, 107(1): 17-24.
- [130] Bussy A, Gergondet P, Kheddar A, et al. Proactive behavior of a humanoid robot in a haptic transportation task with a human partner[C]// Proc. IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). 2012: 962-967.
- [131] Agravante D J, Cherubini A, Bussy A, et al. Collaborative human-humanoid carrying using vision and haptic sensing[C]// Proc. IEEE Intl Conf. on Robotics and Automation (ICRA). 2014: 607-612.
- [132] Agravante D J, Cherubini A, Bussy A, et al. Human-humanoid joint haptic table carrying task with height stabilization using vision[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2013: 4609-4614.

- [133] Oh S, Woo H, Kong K. Frequency-shaped impedance control for safe human-robot interaction in reference tracking application[J]. IEEE/ASME Transactions On Mechatronics, 2014, 19(6): 1907-1916.
- [134] Kouris A, Dimeas F, Aspragathos N. Contact distinction in human-robot cooperation with admittance control[C]// Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2016: 1951-1956.
- [135] Hamedani M H, Zekri M, Sheikholeslam F, et al. Recurrent fuzzy wavelet neural network variable impedance control of robotic manipulators with fuzzy gain dynamic surface in an unknown varied environment[J]. Fuzzy Sets and Systems, 2020.
- [136] Wahrburg A, Morara E, Cesari G, et al. Cartesian contact force estimation for robotic manipulators using Kalman filters and the generalized momentum[C]// Proc. IEEE International Conference on Automation Science and Engineering (CASE). 2015: 1230-1235.
- [137] Albu-Schäffer A, Haddadin S, Ott C, et al. The DLR lightweight robot: design and control concepts for robots in human environments[J]. Industrial Robot: an international journal, 2007.
- [138] Medina J R, Lawitzky M, Mörtl A, et al. An experience-driven robotic assistant acquiring human knowledge to improve haptic cooperation[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2011: 2416-2422.
- [139] Nemec B, Likar N, Gams A, et al. Human robot cooperation with compliance adaptation along the motion trajectory[J]. Autonomous robots, 2018, 42(5): 1023-1035.
- [140] Rozo L, Calinon S, Caldwell D G, et al. Learning physical collaborative robot behaviors from human demonstrations[J]. IEEE Transactions on Robotics, 2016, 32(3): 513-527.
- [141] Rozo L, Bruno D, Calinon S, et al. Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints[C]// Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS). 2015: 1024-1030.
- [142] Rozo L, Calinon S, Caldwell D G. Learning force and position constraints in human-robot cooperative transportation[C]// Proc. IEEE International Symposium on Robot and Human Interactive Communication. 2014: 619-624.

- [143] Calinon S. A tutorial on task-parameterized movement learning and retrieval[J]. Intelligent service robotics, 2016, 9(1): 1-29.
- [144] 彭涛. 基于 sEMG 的人机协作机械臂变阻抗示教技术研究[D]. 哈尔滨工业大学, 2018.
- [145] 武文. 冗余自由度机器人操作臂主动柔顺控制研究[D]. 哈尔滨工业大学, 2019.
- [146] 段江哗. 面向机器人操作的示教学习与智能控制方法研究[D]. 中国科学院大学 (中国科学院深圳先进技术研究院), 2020.
- [147] Allendoerfer C B. Calculus of several variables and differentiable manifolds[M]. Macmillan, 1974.
- [148] Slotine J J E, Li W, et al. Applied nonlinear control[M]. Vol. 199. Prentice hall Englewood Cliffs, NJ, 1991.
- [149] Hoag D. Apollo guidance and navigation: Considerations of apollo imu gimbal lock[J]. Cambridge: MIT Instrumentation Laboratory, 1963: 1-64.
- [150] Siciliano B, Khatib O. Springer handbook of robotics[M]. springer, 2016.
- [151] Zeestraten M J A, Havoutis I, Silvério J, et al. An Approach for Imitation Learning on Riemannian Manifolds[J]. IEEE Robotics and Automation Letters (RA-L), 2017, 2(3): 1240-1247.
- [152] Denavit J, Hartenberg R S. A kinematic notation for lower-pair mechanisms based on matrices[J]. Trans. ASME E, Journal of Applied Mechanics, 1955, 22: 215-221.
- [153] Beeson P, Ames B. TRAC-IK: An open-source library for improved solving of generic inverse kinematics[C]// Proc. IEEE Intl Conf. on Humanoid Robots (Humanoids). 2015: 928-935.
- [154] Cho H C, Park J H. Stable bilateral teleoperation under a time delay using a robust impedance control[J]. Mechatronics, 2005, 15(5): 611-625.
- [155] Nocedal J, Wright S J. Sequential quadratic programming[J]. Numerical optimization, 2006: 529-562.
- [156] Kobyzev I, Prince S, Brubaker M. Normalizing Flows: An Introduction and Review of Current Methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.
- [157] Dinh L, Sohl-Dickstein J, Bengio S. Density estimation using Real NVP[C]// Proc. International Conference on Learning Representations. 2017.

- [158] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization[Z]. 2017arXiv: 1412.6980 [cs.LG].
- [159] Conti F, Khatib O. Spanning large workspaces using small haptic devices[C]// Proc. World Haptics Conference. 2005: 183-188.
- [160] Hokayem P F, Spong M W. Bilateral teleoperation: An historical survey[J]. Automatica, 2006, 42(12): 2035-2057.
- [161] Nehaniv C L, Dautenhahn K, et al. The correspondence problem[J]. Imitation in animals and artifacts, 2002, 41.
- [162] Andersen T T. Optimizing the Universal Robots ROS driver[J]. Technical University of Denmark, 2015.

攻读博士期间发表的科研成果

期刊论文

- [1] **Gao Xiao**, Silvério João, Pignat Emmanuel, Calinon Sylvain, Li Miao, Xiao Xiaohui. Motion Mappings for Continuous Bilateral Teleoperation[J]. *IEEE Robotics and Automation Letters (RA-L)*, 2021, 6(3): 5048-5055. (IF=3.608, Q1, 中科院二区)
- [2] **Gao Xiao**, Ling Jie, Xiao Xiaohui, Li Miao. Learning Force-Relevant Skills from Human Demonstration[J]. *Complexity*, 2019, DOI: <https://doi.org/10.1155/2019/5262859>. (IF=2.462, Q2, 中科院二区)
- [3] **Gao Xiao**, Silvério João, Calinon Sylvain, Li Miao, Xiao Xiaohui. Bilateral teleoperation with object-adaptive mapping [J]. *Complex & Intelligent Systems*, 2021, Under review. (IF=3.791, Q2, 中科院二区)
- [4] 汤自林, **高霄**, 肖晓晖. 基于模仿学习的变刚度人机协作搬运控制 [J]. *浙江大学学报 (工学版)*, 2021. (EI 源刊, 已接收)
- [5] 解五一, **高霄**, 何思宇, 肖晓晖. 面向复合材料自动铺放设备的输带速度与张力协同解耦控制 [J]. *浙江大学学报 (工学版)*, 2019, 53(03): 455-462. (EI 源刊)
- [6] 谢啸, 张涵, 汤自林, **高霄**, 肖晓晖. 面向复合材料自动铺放设备的输带速度与张力协同解耦控制 [J]. *中南大学学报*, 2021. (EI 源刊, 已接收)

会议论文

- [1] **Gao Xiao**, Li Miao, Xiao Xiaohui. Learning Dynamical System for Grasping Motion[C]. *ICRA Workshop*, 2021, accepted.

发明专利（已授权）

- [1] **高霄**, 李淼, 简磊, 肖晓晖. 一种基于示教学习的机器人轴孔装配系统及方法. (ZL 201811275792.8).
- [2] **高霄**, 李淼, 简磊, 肖晓晖. 一种基于示教学习的机器人打磨系统及方法. (ZL 201811275785.8).
- [3] **高霄**, 赵慧, 肖晓晖, 刘胜. 一种基于六轴机器人式的自动铺带装置及方法. (ZL 201510050759.5)

致 谢

时光飞逝，从 2011 年进入武大，十年了。从懵懵懂懂的少年，到现在经过严格科研训练的博士，成长了不少。感谢武大为我提供良好的学习和生活环境，让我有机会遇见很多优秀的老师和朋友。

本文是在导师肖晓晖教授的悉心指导下完成的，从论文选题、讨论提纲、撰写初稿、修改与定稿的全过程，都离不开导师的辛勤指导与帮助。在此，谨向肖老师表达我最深的谢意！生活上老师无微不至地关心我，科研上给予我足够的自由，去探索自己喜欢的研究方向，感谢肖老师这些年来对我如同子女般的呵护，让我可以尽情地追逐梦想。

感谢李淼老师对我在机器人和模仿学习方面的指导，在我科研最迷茫的时候给我指明了方向。李老师分析问题一针见血，思路清晰，每次与李老师讨论我都受益匪浅。感谢实验室郭朝老师这些年来对我的悉心指导，郭老师科研认真负责，对学生有耐心。感谢实验室的朱逸颖老师，让 A9506 更加朝气蓬勃。

读博期间的成长离不开实验室博士们的帮助。感谢实验室的王杨师兄、姚道金师兄、孙剑韬师兄、丁加涛、吴垚在我读博期间的陪伴与相互支持。吴垚博士总能为自己制定好周密的计划并如期完成，这一点让我由衷敬佩。特别感谢凌杰师兄、冯朝师兄一直以来细致入微的帮助和指导，他们对我小论文的撰写、修改和定稿等方面给出了很多建议，对我大论文的框架和思路也提了很多改进意见，他们优秀的科研习惯和脚踏实地的做事态度一直是我学习的榜样。感谢明敏师妹对我的帮助和鼓励，与她讨论问题总能给我新的灵感。感谢何思宇、曹力科、杨尚尚、张晓晖在实验室的陪伴，营造出良好的科研氛围。喜欢在 A9506 与大家分析讨论问题，从天文到地理，从历史到时事热点，从人文到最新科技进展，一群博士在一起讨论总能碰撞出思维的火花，也让我收获不少。读博是一个人孤独的旅程，但有你们在一起总能感到温暖与快乐。与一群优秀的人一起学习，是我最怀念的时光。

此外，感谢实验室的硕士们一直以来对我的支持。感谢“机械臂”课题组的张勇、汤自林、谢啸等人对我的帮助。从自己孤身一人，到队伍逐渐发展壮大，机械臂相关问题的讨论让我们共同成长。特别感谢与我同级的伙伴赵慧、寇重光、陈潇、张匆的陪伴与帮助，一起进入实验室，一起成长的时光让我终生难忘。

衷心地感谢国家留学基金委（CSC）对我的资助，让我有幸去 Idiap 研究所进行联合培养。感谢外导 Calinon Sylvain 对我的帮助和耐心指导，以及论文投稿和修改的建议。特别感谢博后 Silvério João 一次次耐心的讨论和帮助，提高了我的科研素养，João 在科研上给予了我很多建议，生活上也乐于帮助我。感谢 Troussard Martin 对我的生活上的帮助，Martin 的幽默风趣总能给大家带来欢乐。感谢组里的 Girgin Hakan, Shetty

Suhan, Lembono Teguh, Pignat Emmanuel 对我的帮助, 让我在瑞士的一年里过得很开心, 科研上也收获不少。

感谢父母一直对我的爱与支持, 父母虽然不懂我的研究, 依然无条件的支持我, 尽可能地为我创造良好的学习条件。

最后, 衷心感谢百忙之中参与论文评阅与答辩的各位老师, 让我的论文更加完善。