

# 机器人自动轴孔装配研究进展

徐德, 秦方博

(中国科学院自动化研究所, 北京 100190)

**摘要:** 轴孔装配是加工制造业常见的一类操作任务。基于工业机器人研究轴孔自动装配, 对于机器人在装配领域的应用具有重要价值。对于高精密和形状复杂的零件, 高效可靠的轴孔装配仍然具有很大挑战性。基于此, 从控制的角度对机器人自动轴孔装配进行了全面梳理。首先, 介绍了机器人自动轴孔装配过程。然后, 在对基于传统模型的装配控制进行论述的基础上, 对新兴的基于学习的智能装配控制进行了讨论, 重点阐述了模仿学习和强化学习在机器人自动装配中的应用。传统方法与人工智能方法的结合, 将为机器人自动轴孔装配注入新的活力, 将成为未来的重要发展趋势。

**关键词:** 轴孔装配; 智能机器人; 模仿学习; 强化学习; 技能学习

**中图分类号:** TP242.6

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-6652.202223

## Research development on automated robotic peg-in-hole assembly

XU De, QIN Fangbo

Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

**Abstract:** Peg-in-hole assembly is a typical operation task in manufactory. The research of peg-in-hole assembly based on industrial robots is valuable for the application of robots in the automated assembly area. For the peg-in-hole components with high precision or complex shapes, the efficient and reliable assembly is still very challenging. The development of automated robotic peg-in-hole assembly of peg-in-hole was reviewed from the view of control. First, the process of robotic peg-in-hole assembly was introduced. Secondly, the assembly control methods based on the traditional models were described. The newly emerged intelligent assembly methods based on learning mechanism were discussed, especially the applications of imitation learning and reinforcement learning in the automated robotic assembly. The combination of the traditional methods and the artificial intelligent methods will provide new energy for the automated robotic assembly, which will be one of the important developing tendencies in future.

**Key words:** peg-in-hole assembly, intelligent robot, imitation learning, reinforcement learning, skill learning

### 0 引言

轴孔装配是自动装配任务中比较典型且常见的一类工业任务。在过去数十年来, 自动装配是具有挑战性的研究领域<sup>[1-2]</sup>, 高性能的轴孔装配更是具有挑战性的热点问题。例如, 2019年, 清华大学

Xu J 等人<sup>[3]</sup>指出, 由于传统的控制策略需要对接触模型进行复杂分析, 利用机器人进行多轴孔自动装配依然具有挑战性。2020年, 大阪大学 Beltran-Hernandez C C 等人<sup>[4]</sup>认为, 在非结构化环境中的复杂高精度自动装配依然是一个有待解决的问题。2021年, 北京邮电大学 Song J Z 等人<sup>[5]</sup>

收稿日期: 2022-01-14; 修回日期: 2022-04-14

通信作者: 徐德, de.xu@ia.ac.cn

基金项目: 科技创新 2030—“新一代人工智能”重大项目 (No.2018AAA0103005); 国家自然科学基金资助项目 (No.61873266, No.62103413)

**Foundation items:** The National Key Research and Development Program of China (No.2018AAA0103005), The National Natural Science Foundation of China (No.61873266, No.62103413)

指出, 长期以来, 利用机器人装配零件一直是研究热点之一。

目前, 工业机器人已被广泛用于自动执行多种工业任务, 例如电焊、喷涂、包装、码垛等。这些任务的共同特点是机器人的运动采用位置控制。但零件之间的装配常常涉及装配力, 难以依靠位置控制实现对零件的自动装配。此外, 待装配零件的不确定性也增加了机器人自动装配的难度, 例如末端执行工具、轴、孔等具有的不确定性。因此, 基于工业机器人研究轴孔自动装配, 对于拓展机器人在装配领域的应用具有重要价值<sup>[6-7]</sup>。

## 1 装配过程

按照轴与孔之间的接触情况, 可以将单轴孔装配过程分为多种状态。例如, 2014 年, 卢森堡大学 Jasim I F 等人<sup>[8]</sup>针对单轴孔间隙配合, 将装配过程分为 5 种接触状态。状态 1, 自由状态; 状态 2, 轴接近孔; 状态 3, 轴初步入孔, 姿态需要调整; 状态 4, 轴姿态调整完毕, 轴对准孔; 状态 5, 完成轴孔装配。在轴初步入孔状态下, 插入过程往往具有卡阻问题。2021 年, 浙江工业大学潘柏松等人<sup>[9]</sup>对卡阻问题进一步细化, 将单轴孔的卡阻状态分为两类, 一类为径向卡阻状态, 包括孔外单点、孔外两点、孔内单点的接触状态; 另一类为径轴双向卡阻状态, 包括孔内两点、孔内三点的接触状态。针对这两类卡阻状态, 建立了单轴孔在各接触状态下的几何与力学模型。

轴孔装配过程通常可以分为 3 个阶段, 分别为搜索阶段、对准阶段和装入阶段。也有部分文献将搜索阶段和对准阶段统称为搜索对准孔阶段。搜索阶段和对准阶段可以采用视觉伺服, 装入阶段需要根据受力/力矩调整轴/孔的位姿。对于单轴孔装配, 在轴孔之间具有较大间隙的情况下, 只要轴孔对准具有较高的精度, 就能够将轴装入孔中; 在轴孔之间具有较小间隙或过渡配合、过盈配合的情况下, 轴孔对准存在的微小位姿偏差都会给装入带来较大的困难, 需要对装配过程中的接触状态和力进行感知, 增加装配过程的适应性<sup>[2]</sup>。对于多轴孔装配, 由于零件间接触情况更加复杂, 存在接触刚度的各向异性, 并且多个轴间存在不确定的形变与制造误差, 仅仅依靠轴孔对准, 难以将多轴装入多孔中。例如, 卫星燃料加注过程中燃料枪的插入, 属于双轴孔装配; 航空插头等接插件的对接, 电源模块等

4 脚、6 脚元件在电路板上的插入, 属于多轴孔装配。这些操作具有共同特点: 仅仅依靠对准不能保证插入成功, 插入过程存在轴向卡阻问题。

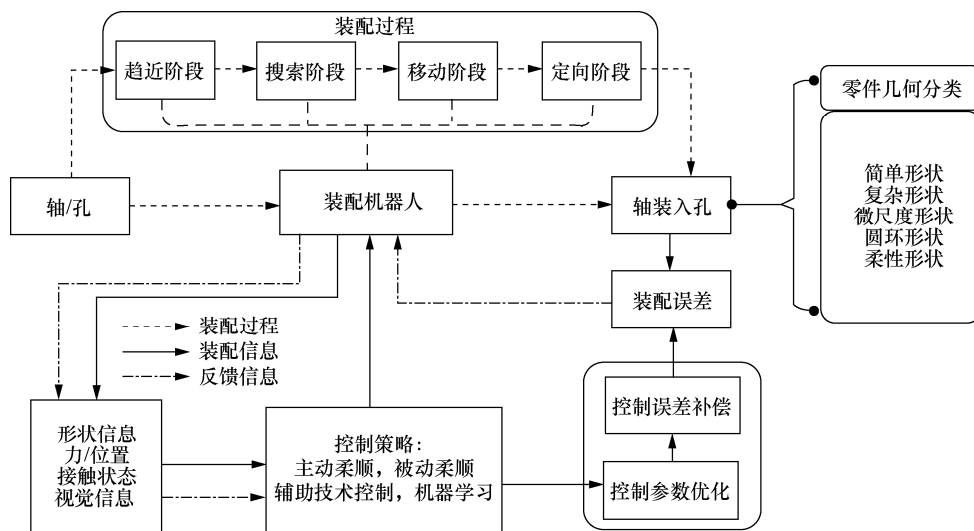
针对装配精度和零件变形问题, 2016 年, 名古屋大学 Takahashi J 等人<sup>[10]</sup>提出了一种基于被动对准原理 (passive alignment principle, PAP) 的接合技术, 共有 4 步。①趋近, 机器人抓取薄环, 接近杆件, 末端绕 X 轴有倾角。②放松, 手指放松以降低夹持力, 三手指外移, 保持薄环与杆件接触。③被动对准, 去掉夹持力, 使得薄环可以在水平方向自由移动, 完成被动对准。④接合装入, 末端绕 X 轴旋转, 使绕 X 轴的倾角为 0, 下推薄环, 装入杆件上。虽然轴孔装配中刚性零件的变形很小, 但被动对准的思路具有很好的借鉴意义。被动对准的核心是充分利用轴孔之间的导向性。

根据装配过程中部件的运动情况, 2020 年, 科罗拉多大学 Watson J 等人<sup>[11]</sup>将轴孔装配分为轴装入孔 (peg-in-hole, 也简称轴孔) 和孔装在轴 (hole-on-peg) 任务, 例如, 将圆柱滚子插入轴承属于轴装入孔, 将螺帽套在螺钉上属于孔装在轴。针对装配过程的孔 (或轴) 搜索和装入两个阶段, Watson J 等人<sup>[11]</sup>提出了两种规范算法, 分别为基于螺旋的搜索算法和倾斜装入算法。这两种算法利用在力/力矩域中人工设定的阈值, 检测装配过程中的临界点 (critical point)。这些算法对于轴初步入孔具有很好的参考价值。

2020 年, 哈尔滨理工大学 Jiang J G 等人<sup>[1]</sup>对机器人轴孔装配方法进行了梳理, 将机器人装配过程分为 4 个阶段: 趋近阶段 (approaching phase)、搜索阶段 (searching phase)、移动阶段 (moving phase)、定向阶段 (orientating phase)。机器人轴孔装配过程与装配零件分类如图 1 所示。轴孔装配面临的困难取决于轴孔的形状, 轴孔形状可分为简单形状、复杂形状、微尺度形状、圆环形状、柔性形状等。

2020 年, 南丹麦大学 Sloth C 等人<sup>[12]</sup>建议对装配任务演示编程, 而不是直接对整个过程或行为进行示教, 将演示与装配基元合并以增强鲁棒性。与其他演示编程方法不同, 作者执行的不是一次演示, 而是 4 个子演示的序列<sup>[12]</sup>, 用于提取装配基元参数和期望的机器人轨迹。实际上, 这 4 个子演示对应装配的不同阶段, 也可以认为是轴孔装配的 4 种行为。

对于多轴孔装配, 其接触状态和装配过程更加

图1 机器人轴孔装配过程与装配零件分类<sup>[1]</sup>

复杂。其接触状态形式多样，并具有刚度的各向异性，难以进行精确建模。其装配过程需要更多的试探性，需要更加复杂的装配流程和装配技巧，这使得模仿学习、强化学习等人工智能方法在装配领域具有良好的应用前景<sup>[13]</sup>。

## 2 基于传统模型的装配控制

基于传统模型的装配控制需要在对准阶段建立待装配零件的模型或者对零件进行定位，在装入阶段需要建立装配过程的受力模型。例如，2003年，中国工程物理研究院李侃等人<sup>[14]</sup>提出双轴孔装配方法，利用相机和激光传感器测量零件1的位置和姿态，利用机器人抓取零件2，将零件2对准零件1后，沿轴线运动实现装入。零件1与零件2为间隙配合，不涉及装配力控制，属于早期的装配方式。2015年，西安理工大学魏明明等人<sup>[15]</sup>将建模与试探结合，实现了间隙配合的轴孔装配。在趋近阶段，分别建立趋近装配目标点和避障的行为动力学模型，在视觉伺服引导下，使装配机器人末端执行器从起始点无碰撞地到达装配目标点。在装配阶段，利用视觉获取孔轴线的空间姿态，基于吸引强度利用宏微运动平台调整X轴和Z轴的夹角，使得轴孔和轴线平行；然后，利用建立的受力与位姿偏差模型，根据力传感器的数据计算末端执行器侧向误差，调整X轴、Y轴方向位置，利用微动平台在Z轴方向进行插入，完成了轴孔装配。2019年，东南大学徐远等人<sup>[16]</sup>提出了基于视觉和力反馈的自动装配方法。在趋近阶段基于视觉进行工件定位，将轴运动

到孔处。在装入阶段，采用直线运动与螺旋线运动两种装配轨迹规划策略，基于反馈的力与力矩将轴装入孔。相比之下，直线运动的效率较高，螺旋线运动的成功率在小间隙时较高。2020年，陕西科技大学陈婵娟等人<sup>[17]</sup>建立了受力与位姿偏差的模型，利用位移传感器检测孔零件的位姿，实现轴孔粗对准；利用受力与位姿偏差的模型将力/力矩传感器信号转换为位姿偏差，实现轴孔精对准，从而实现间隙配合零件的装入。

位置控制具有很高的刚度，装入阶段采用位置控制容易导致零件损伤甚至损坏。主动柔顺控制可以控制零件之间的接触力，使得装配具有一定的柔性，有利于在装配过程中保护零件。常见的主动柔顺控制方法包括阻抗控制和力位控制等。例如，2019年，天津大学Zeng F等人<sup>[18]</sup>将环境的阻抗模型与机器人带有摩擦的动态模型联合在一起，提出了虚拟接触表面，并将其用于优化装配。作者设计的柔顺控制装配系统为双闭环结构，将基于PD（proportional-derivative）的位置控制作为内环，将阻抗控制作为外环。2017年，首尔大学Park H等人<sup>[19]</sup>设计了一种控制方案，机器人能够在力位混合控制下产生接触柔顺行为。在该方案中，Kinect设备被安装在仿人机器人头部，用于识别孔的位置和朝向；系统中未安装力传感器，将轴孔位置偏差乘以系数作为力。Park H等人<sup>[19]</sup>针对圆柱状单轴孔装配的装入阶段，仿照人的装配过程设计了4项基本运动，如图2所示，分别为推（pushing）、揉（rubbing）、摆（wiggling）、旋（screwing）。经过沿



轴向的“推”运动,轴接触孔;经过绕轴向螺旋的“揉”运动,轴端部进入孔;经过以轴端部为中心变轴向的“摆”运动,轴前部进入孔;经过绕轴向旋转的“旋”运动,轴装入孔。

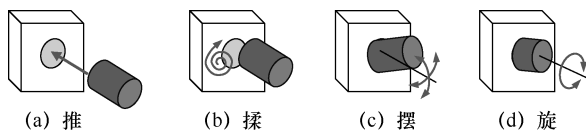


图2 仿照人的装配过程的4项基本运动<sup>[19]</sup>

2017年,北京邮电大学陈钢等人<sup>[20]</sup>提出了一种双臂协调操作的轴孔装配力位混合控制方法,轴和孔采用相向运动模式,轴和孔相对惯性系各自运动0.1 m,实现轴对孔0.2 m深度的装入。2018年,中国科学院沈阳自动化研究所吴炳龙等人<sup>[21]</sup>设计了基于力位混合控制的工业机器人精密轴孔装配系统。在孔搜索阶段,采用力位混合控制。其中,在 $X$ 轴、 $Y$ 轴方向进行位置控制,沿螺旋线轨迹以接触点为起始点由内向外进行孔搜索;在 $Z$ 轴方向采用力控制,保持一个恒定的力值,以便在找到孔的时候能够将轴插入孔。2019年,北京卫星环境工程研究所董恂等人<sup>[22]</sup>提出了基于动力学前馈的空间机器人多销孔装配力柔顺控制方法,在阻抗控制的基础上,基于动力学前馈直接对关节进行力矩补偿,再基于机器人关节的数学模型,将前馈算法的补偿量进一步映射到电机电流环,从而进一步加快关节响应速度,保证力柔顺装配效果。

中国科学院自动化研究所针对微零件的自动装配开展了大量研究,建立了基于多路显微视觉的微零件的位姿测量模型,分析了相对位姿偏差与微零件受力分布机理,实现了直径mm级过盈配合柱状微零件之间的装配,实现了直径 $10\mu\text{m}$ 级轴孔的高精度装配,实现了多零件的协调装配<sup>[23-33]</sup>。

基于模型的装配控制在装入阶段对接触模型的建模精度依赖性强,对不同任务的适应性较差,并且需要进行精确的人工标定与调试。对于多轴孔零件,其复杂的接触状态更加难以精确建模和在线感知,基于模型的装配控制难以取得令人满意的效果。

### 3 基于学习的智能装配控制

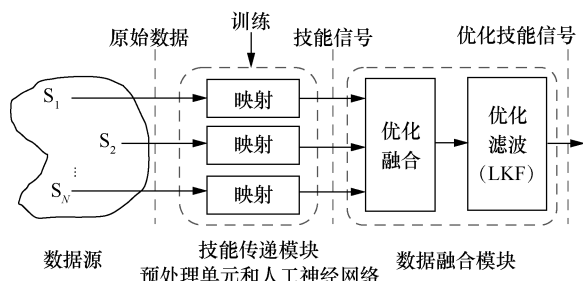
为了提高装配控制的适应性,降低对模型的依赖性,研究人员开发出一系列用于装配的智能控制方法。不同于基于模型的装配控制,智能装配控制

不需要建立准确的模型。模糊控制、自适应控制等智能控制方法是装配领域比较常见的智能控制方法。例如,2015年,卢森堡大学Jasim I F等人<sup>[34]</sup>针对动力学未知且具有任意开关未知约束的机器人提出了一种自适应滑模模糊控制策略。其中,未知参数在线自适应,闭环信号保证是有界的。将上述自适应滑模模糊控制策略应用于轻量级库卡(KUKA)机器人的轴孔装配任务,取得了良好的效果。为了提高动态环境中机器人执行装配任务的灵活性,2021年,德国宇航中心Nottensteiner K等人<sup>[35]</sup>提出了一种机器人自主装配的方法。该方法采用视觉和内部触觉,在贝叶斯框架下连续跟踪零件。即使在视觉系统受到遮挡时,机器人也可以在零件估计姿态的引导下,实现面向目标的装配。采用模糊控制、自适应控制等智能控制方法,虽然能够提高机器人装配的适应能力,但装配精度较低。

近年来,随着人工智能的快速发展,模仿学习、强化学习等学习方法被引入装配领域,并得到研究人员的普遍重视。

#### 3.1 基于模仿学习的装配技能获取

模仿学习从示教装配的过程中学习技能,通过操作员演示形成一系列数据,通过对数据的拟合生成运动行为,进而获得装配技能。早在2004年,科英布拉大学Cortesao R等人<sup>[36]</sup>利用神经网络建立感觉-运动(sensory-motor)函数,根据传感器信号控制电机运动。Cortesao R等人<sup>[36]</sup>设计了一种基于人类感觉-运动技能的机器人装配数据融合框架,通过几何感知信号向动态感知信号的映射,将人的技能传递给机器人。基于感觉-运动的技能获取框架如图3所示,该框架含有两个模块,分别为技能传递模块(skill transfer modular)和数据融合模块(data fusion modular)。技能传递模块包含预处理单元和人工神经网络,用于将原始传感数据转换为技能信号。数据融合模块具有两个子模块,分别为优化融合和优化滤波模块,用于对技能信号进行优化。技能信号由柔顺运动信号(即力、速度)构成,可从视觉和姿态数据中获得。优化融合采用加权平均,优化滤波采用线性卡尔曼滤波(linear Kalman filter, LKF)。操作员示教装配过程,形成传感器和技能信号的数据对,利用这些数据对对神经网络进行训练。该框架中的映射虽然比较直观,但表达能力和适应性都较弱,其模仿学习的能力远远不能满足轴孔装配的需求。

图3 基于感觉-运动的技能获取框架<sup>[36]</sup>

近年来，动态运动基元（dynamic movement primitive, DMP）、高斯混合模型（Gaussian mixture model, GMM）和高斯混合回归（Gaussian mixture regression, GMR）等成为模仿学习中比较常用的模型。DMP 是非线性二阶动态系统，在二阶阻尼系统上施加了非线性扰动，从而实现任意轨迹形状的非线性运动。利用 DMP 可以对示教的运动轨迹进行编码，进而生成具有人类行为特点的运动轨迹，具有良好的泛化能力。GMM 和 GMR 用于拟合示教数据，即从统计意义上对示教轨迹进行整合。其中利用 GMM 将多次示教的轨迹数据编码为概率模型，利用 GMR 进行时间分布预测。

DMP 在模仿学习中被广泛应用，近年来也不断得到改进。例如，2018 年，瑞士洛桑联邦理工学院 Petrič T 等人<sup>[37]</sup>提出了一种基于柔顺运动基元（compliant movement primitive, CMP）的自主扩展数据库方法，用于生成柔顺精密运动。受人的感觉运动（sensorimotor）学习能力的启发，该方法结合了特定任务的运动学和相应的前馈-反馈动力学轨迹，设计了一种自主学习与给定运动轨迹关联的 CMP。换言之，CMP 是带有力矩的运动轨迹，包含 DMP 和力矩基元（torque primitive, TP）。CMP 的获取分为两步：第 1 步利用演示或者预定义得到笛卡儿空间的运动轨迹，并将其编码为 DMP，其权重利用增量式局部加权回归（locally weighted regression, LWR）学习获得；第 2 步利用基于误差学习的递归回归（recursive regression）获得相应的力矩，并将其编码为 TP。利用上述方法，加快了机器人轴孔装配过程中 TP 的学习速度。2020 年，中国科学院沈阳自动化研究所 Liu N L 等人<sup>[38]</sup>提出了一种在笛卡儿空间具有力位混合反馈的改进 DMP 模型，用于基于演示学习的机器人轴孔装配。为了保证轴孔装配过程中的柔顺交互，采用笛卡儿空间的阻抗控制方法跟踪改进 DMP 模型生成的轨迹。

对于多次示教的数据，可以利用 GMM 将其编

码为概率模型，然后利用 GMR 进行状态的时间分布预测。例如，2014 年，卢森堡大学 Jasim I F 等人<sup>[8]</sup>为对轴孔装配的接触状态建模，采集的信号包括力、力矩、位置、姿态，并采用基于期望最大的 GMM 对信号数据进行编码。2016 年，高丽大学 Song H C 等人<sup>[39]</sup>提出了一种基于导引算法的复杂形状零件轴孔装配策略。该策略受到人的运动模式的启发，即通过观察人执行类似动作确定装配方向和最大力阈值。为了执行装配任务，装配方向根据螺旋排列和复杂零件的几何信息进行选择，通过肌肉运动示教由 GMM 获得所需要的力。采用具有导纳滤波器的阻抗控制器控制基于位置的工业机器人，执行稳定接触运动。2019 年，中国科学院自动化研究所 Qin F B 等人<sup>[40]</sup>提出了一种基于 GMM-GMR 的微装配技能学习方法，实现了柱状微零件的自动装配。该方法采用 GMM 对示教的行动数据进行编码，采用 GMR 对状态的分布进行预测，不仅使运动具有非线性，也保证了运动的全局稳定性和泛化能力。设计的对准、装入可以在不同的技能中得到复用，分别用于显微视觉伺服控制器和微力柔顺控制器。该方法仅需 10 余次示教就可以完成学习，运动稳定性可保障，对于起始条件变化具有良好的适应性，对于过程干扰具有鲁棒性。

2020 年，哈尔滨工业大学 Ding G W 等人<sup>[41]</sup>指出，传统的任务预编程方法限制了人与机器人技能传递的效率。Ding G W 等人<sup>[41]</sup>建立了无标记的视觉采集系统，可以连续采集人手的运动；提出了基于阈值的启发式分割算法，可以将完整的运动分割为不同的运动基元，将人手的运动编码为面向任务的模型。但参考文献[41]中的轴与孔为间隙配合，并且间隙较大，装入过程不涉及力。参考文献[41]注重将含轴零件拾取后运动到孔的上方对准孔，释放后实现轴孔装配。2019 年，西班牙马拉加大学 Perez-Del-Pulgar C J 等人<sup>[42]</sup>将演示学习方法扩展到遥操作，将任务分解为一系列需要探测的手势。通过演示学习，对每个手势进行 GMM 编码，并将其存储在手势库中。利用这些编码手势对 KUKA 机器人进行遥操作，实现了远程控制下的机器人轴孔装配。

针对装配过程中与力相关的模仿学习问题，2019 年，武汉大学 Gao X 等人<sup>[43]</sup>将与力相关的技能表达为统计模型，其关键参数从运动和力的演示数据中学习得到。基于学习到的技能模型，任务规划器根据给定的操作任务指定运动与力包络（force

profile)。利用 GMM 将演示数据编码为概率模型, 利用 GMR 进行时间分布预测。在装配过程中, 将学习到的技能模型与自适应控制器结合, 根据力/力矩调整末端执行器的位置、姿态、速度, 实现与轴孔装配任务相符的力适应性。2019 年, Ti B Y 等人<sup>[44]</sup>指出, 技能学习过程中人的介入可以极大地改善机器人的性能。但是, 通过演示学习在变化环境中生成具有人类行为特点的行为具有挑战性。演示中力包络的形状和运动轨迹都反映出人的意图与装配需求, 在与力相关的装配技能模仿学习中, 如何综合学习力包络和运动轨迹有待进一步研究。

将模仿学习与基于模型的控制结合, 有利于有效集成两者的优势。2021 年, 北京邮电大学 Song J Z 等人<sup>[5]</sup>提出了一种新的装配策略, 从示教装配的过程中学习技能, 以应对几何轮廓参数和材料弹性参数的不准确问题。Song J Z 等人<sup>[5]</sup>将装配过程分为趋近、接触、调整、对准和装入阶段。在趋近和装入阶段, 采用 PID(proportional-integral-derivative) 位置控制。在调整阶段, 在  $X$  轴、 $Y$  轴方向采用阻抗控制保持轴与孔的接触力, 在  $Z$  轴方向采用位置控制。在对准阶段, 采用基于 GMM/GMR 的柔顺控制, 如图 4 所示(其中  $X_d$  是给定位姿,  $q_r$  是参考关节角,  $q$  是关节角,  $\dot{q}$  是关节速度,  $F$  是力,  $T$  是力矩)。下面以  $X$  轴方向对准为例, 说明其控制原理。人工拖拽机器人末端, 将轴对准孔, 形成  $X$  轴方向受力与角度偏差的  $F_x$ - $\theta$  示教数据。利用 GMM 对  $F_x$ - $\theta$  示教数据的概率分布进行学习, 然后利用 GMR 按照 68% 的置信度估计出  $F_x$  随  $\theta$  变化的平滑曲线带。在对准阶段, 根据测量到的  $F_x$ , 利用  $F_x$ - $\theta$  的平滑曲线带给出不同角度偏差  $\theta$  的概率, 经逆向运动学将  $\theta$  转化为关节角, 作为反馈量参与柔顺控制。实验结果表明, 间隙 0.18 mm 零件的装配成功率为 100%。该方法对机器人运动学模型的依赖性较强, 而且  $F_x$ - $\theta$  曲线带的学习精度会影响装配运动控制效果。

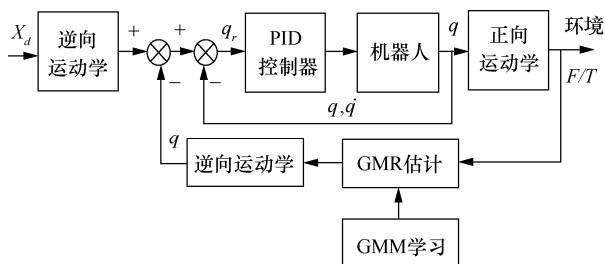


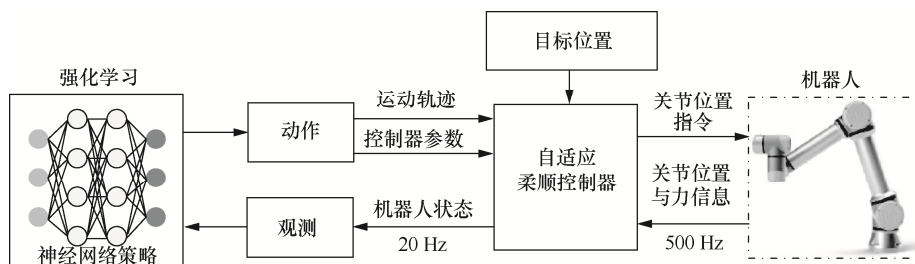
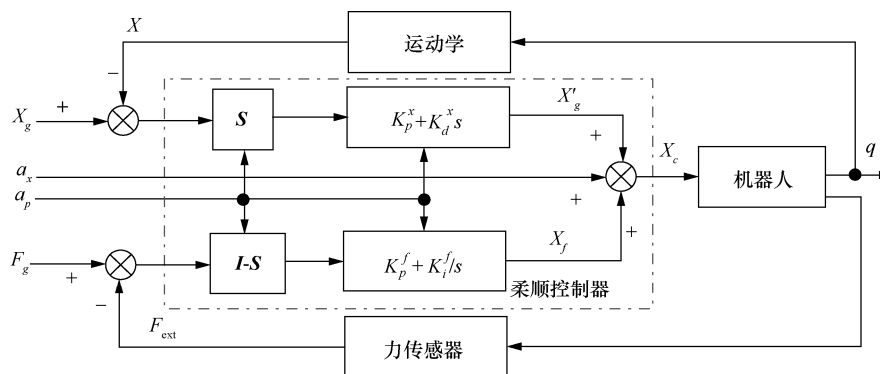
图4 基于 GMM/GMR 的柔顺控制<sup>[5]</sup>

由于人工演示的装配性能受限于人机交互技术和演示者的技能水平, 用于模仿学习的演示数据可能不是最优的, 甚至对于多轴孔复杂装配任务, 基于常规人机交互设备的反馈信号有限, 人工演示难以实现成功装配。

### 3.2 基于强化学习的技能优化

强化学习使机器人以试错机制与环境进行交互, 通过最大化累计奖赏方式学习到最优技能策略。演员-评价者(actor-critic)是连续动作空间下强化学习算法常采用的网络结构。近年来, 研究人员将强化学习引入轴孔装配, 对装配技能进行优化。例如, 2014 年, 京都大学 Senda K 等人<sup>[45]</sup>提出了一种利用强化学习生成技能的方法。该方法考虑了对象变化导致的模型差异, 在对象位姿上添加扰动, 利用强化学习获得动作策略。其中, 在轴孔装配过程中, 孔的位姿由立体视觉系统测量, 并将孔的位置和姿态误差建模为正态分布。为了避免传统控制策略对接触模型的依赖性, 2019 年, 清华大学 Xu J 等人<sup>[3]</sup>将强化学习引入多轴孔的机器人自动装配。Xu J 等人<sup>[3]</sup>将轴孔装配任务作为马尔可夫决策过程, 采用 actor-critic 强化学习网络, 提出了一种模型驱动的深度确定策略梯度(model-driven deep deterministic policy gradient, MDDPG)算法, 在不显式分析接触状态的情况下, 利用强化学习在虚拟环境下训练获得策略, 并将其迁移至实际环境, 实现了双轴孔的机器人自动装配。

将强化学习与传统方法结合, 有利于发挥传统方法的优势, 并弥补强化学习效率低的不足。例如, 2019 年, 德国博世中央研究院 Hoppe S 等人<sup>[13]</sup>将无模型强化学习与规划结合, 生成操作任务的信息数据。2020 年, 大阪大学 Beltran-Hernandez C C 等人<sup>[4]</sup>将强化学习与自适应控制结合, 利用基于学习的方法, 解决具有孔位不确定性的轴孔装配问题。基于强化学习的自适应柔顺控制系统如图 5 所示。利用强化学习获得策略后, 以机器人状态为输入, 神经网络输出动作, 包括机器人运动轨迹和自适应控制器的参数。自适应控制器是基于力位混合控制的柔顺控制器, 如图 6 所示。图 6 中, 虚线框表示柔顺控制器;  $X_g$  是机器人的目标位姿,  $X$  是机器人位姿,  $X'_g$  是位置通道的输出位姿,  $X'_f$  是力通道的输出位姿,  $X_c$  是综合后发送给机器人的期望位姿;  $F_g$  是机器人的目标力,  $F_{ext}$  是机器受到的外力;  $q$  是关节角;  $a_p$  是神经网络输出动作中的控制器参数,  $a_x$  是神经

图5 基于强化学习的自适应柔顺控制系统<sup>[4]</sup>图6 基于力位混合控制的柔顺控制器<sup>[4]</sup>

网络输出动作中的位姿； $S$  是位置通道的选择矩阵， $I-S$  是力通道的选择矩阵； $K_p^x + K_d^x s$  是位置通道 PD 控制器的传递函数， $K_p^x$  是位置通道 PD 控制器的比例系数， $K_d^x$  是位置通道 PD 控制器的微分系数； $K_p^f + K_i^f/s$  是力通道 PI (proportional- integral) 控制器的传递函数， $K_p^f$  是力通道 PI 控制器的比例系数， $K_i^f$  是力通道 PI 控制器的积分系数。 $a_p$  是图 6 中 PD 控制器和 PI 控制器的参数。神经网络输出动作中的位姿  $a_x$  与位置通道的输出位姿  $X_g'$  和力通道的输出位姿  $X_f$  相加，构成机器人的目标位姿。2020 年，高丽大学 Kim Y L 等人<sup>[46]</sup>认为将强化学习用于接触式任务时，采用已有的力控制器解决接触问题是必要的。Kim Y L 等人<sup>[46]</sup>提出了一种基于神经网络的运动基元 (neural-network based movement primitive, NNMP)，设计了适合 NNMP 的模仿学习算法，生成与演示轨迹类似的初始运动轨迹。然后，利用强化学习对生成的初始运动轨迹进行优化。在控制系统中，将 NNMP 产生的运动轨迹输入力控制器，由力控制器控制机器人的运动，实现了方形轴孔的自动装配。2021 年，Khader S A 等人<sup>[47]</sup>将变阻抗控制 (variable impedance control, VIC) 引入强化学习，并利用交叉熵 (cross entropy) 方法优化采样分布，解决强化学习的全时稳定性问题。上述将强化学习与传统方法结合的方式，对于多轴孔的机

器人自动装配具有重要的参考价值。

近年来，研究人员提出了一些新的强化学习方法，并用于轴孔自动装配。例如，2020 年，京都大学 Senda K 等人<sup>[48]</sup>提出了一种基于近似参数模型的贝叶斯强化学习方法，基于贝叶斯估计模型进行在线规划。该方法用少量的实例即可学习新的机器人任务，对于模型不确定性具有鲁棒性，并且需要的计算资源在可行的范围内。上述新的强化学习方法将为智能化的轴孔装配提供重要参考。

### 3.3 多种方式结合的技能学习

先验知识有利于提高轴孔装配的质量和效率。因此，先验知识对于轴孔装配任务具有十分重要的作用。如何将先验知识有机地与学习结合起来，一直是研究人员努力的重要方向之一。例如，2007 年，伍伦贡大学 Dong S 等人<sup>[49]</sup>研发出一种新的技能学习系统，在虚拟环境中结合任务的先验知识生成位置、接触力、力矩以及姿态，用于识别和学习新的演示任务中的技能。记录的数据分为不同的接触状态，利用隐马尔可夫模型 (hidden Markov model, HMM) 识别人的技能。在触觉渲染虚拟环境中，通过评估每一种接触状态下控制器的性能，机器人形成仿人的装配运动轨迹。

将多种学习方式结合有利于增强装配技能学习的能力。目前，研究人员开始尝试将模仿学习与

强化学习结合,在技能学习的不同阶段发挥模仿学习和强化学习各自的优势。2020年,韩国汉阳大学 Cho N J 等人<sup>[50]</sup>提出了一种基于模仿学习和自学习的技能学习框架,机器人能够学习、改进和泛化运动技能,如图7所示。机器人共有两种运动技能,分别为孔搜索和轴装入。机器人从人的演示中学习初始运动技能,并通过强化学习改进或泛化这些技能。初始运动技能表示为 HMM 和 DMP 的级联,对输入信号分类并产生运动轨迹。HMM 用于动作选择,DMP 用于运动生成。将已建模的类作为熟悉类,对初始运动技能进行改进,以完善机器人对熟悉类的反应。将未建模的类作为非熟悉类,对技能进行泛化,以解决机器人对非熟悉类的反应。利用该框架,机器人实现了对多种形状轴孔的自动装配。

为了提高对装配对象和环境变化的自适应能力,在2018年,东北大学王斐等人<sup>[51]</sup>提出了一种多深度确定性策略梯度(multiple deep deterministic policy gradient, M-DDPG)算法修正装配参数,在演示编程的基础上,通过强化学习,机器人可以稳定执行任务。王斐等人<sup>[51]</sup>提出了一种改进的并行卷积神经网络(parallel convolution neural network, PCNN),即通过一维的卷积与池化过程自动提取惯性信息与肌电信息特征,增强手势识别的泛化性和准确率;采用 GMM 对演示数据进行统计编码,利用 GMR 方法实现机器人轨迹运动再现,消除噪声点。2021年,中国科学院自动化研究所 Ma Y Q 等人<sup>[52]</sup>提出了结合初始策略和剩余策略的微零件装配技能学习方法,利用模仿学习从演示数据获得初始策略,利用强化

学习获得剩余策略,有效提高了训练效率,实现了高精度的多轴孔微器件的装配。

虽然粗略模型具有较大误差,但也含有十分有用的知识,能够有效减少学习过程中漫无目的的探索,提高技能学习的效率。因此,将建模与学习结合,也是一条十分重要的途径,受到研究人员重视。2021年,浙江工业大学潘柏松等人<sup>[9]</sup>提出一种基于几何约束物理模型与 HMM 相结合的机器人轴孔装配策略。首先,通过建立轴孔的各接触状态下的几何与力学模型,分析零件装配过程中几何约束与受力特点,获得轴孔装配策略的理论装配轨迹。然后,采用少样本示教学习方法获得 HMM-GMR 模型,将机器人实际位姿输入提前训练好的 HMM-GMR 模型,得到装配过程中的期望力,通过导纳控制器对理论轨迹进行补偿,实现对期望接触力的跟踪,从而保证实际装配过程中接触运动的准确性和鲁棒性。针对最小间隙为 0.16 mm 的轴孔零件进行了装配实验,取得了良好的效果。针对微零件装入过程中接触状态复杂并难以准确建模的问题,2021年,中国科学院自动化研究所 Li Y 等人<sup>[53]</sup>提出了一种结合建模与强化学习的微零件装入技能学习方法。首先,将装入动作分解为专家动作和改进动作两部分,分别用于保证安全性和效率。其次,设计了专家动作获取方法,分别利用多次、单次示教数据或力的雅可比矩阵,获取专家动作。然后,通过强化学习获取改进动作。该方法有效提高了强化学习的训练效率,增强了装配的适应性和灵活性。

### 3.4 技能迁移与复用

技能迁移是技能在不同载体、环境、任务中的

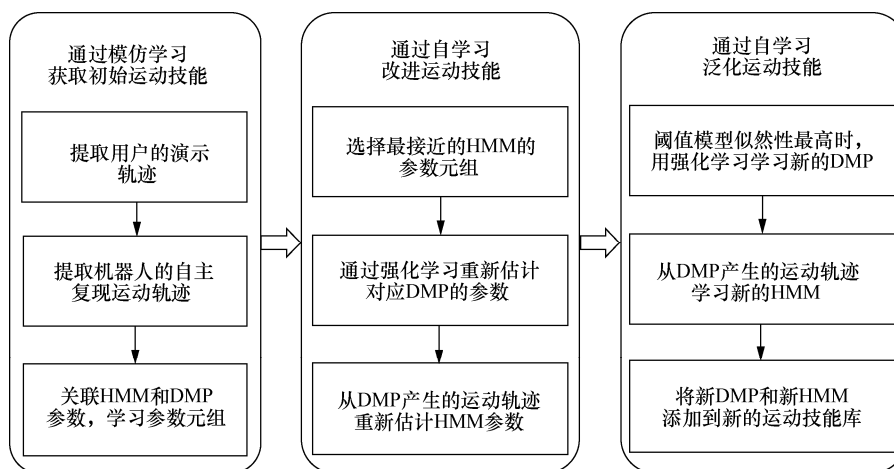


图7 基于模仿学习和自学习的技能学习框架<sup>[50]</sup>



传递,包括不同机器人之间技能的传递、不同环境间技能的传递、不同任务间技能的传递等。例如,针对不同装配任务,2014年,斯洛文尼亚约瑟夫·斯特凡研究所 Abu-Dakka F J 等人<sup>[2]</sup>提出了一种基于力的技能学习方法,利用装配过程中的运动与力/力矩包络学习装配技能,并利用强化学习来优化学习到的力/力矩包络以及最优的刚度参数。在新的任务中,利用视觉获得新零件的位姿,利用迭代控制学习对笛卡儿空间轨迹进行自适应调整,使其与学习到的力/力矩包络匹配。2019年,中国科学院沈阳自动化研究所刘乃龙等人<sup>[54]</sup>在3D图形仿真环境下基于深度强化学习训练机器人学习轴孔装配任务,使得机器人能处理更多环境不确定性因素下的轴孔装配问题。2020年,大阪大学 Beltran-Hernandez C C 等人<sup>[4]</sup>采用离线策略和无模型强化学习方法,利用多种迁移学习技术和域随机化,提升了训练速度。技能传递对于完成新任务、学习新技能具有重要作用,也是实现技能复用的重要方法,但目前处在起步阶段,尚有许多问题有待探索。

#### 4 未来发展趋势

在当前的机器人轴孔自动装配中,模仿学习在对准阶段主要用于模仿演示操作的零件轨迹,在装入阶段主要模仿演示操作形成的零件之间的力/力矩包络。这种对轨迹或者受力的模仿过于注重局部,未将装配过程作为整体考虑。为了改变模仿学习“知其然而不知其所以然”的现状,有必要改变模仿学习的目标,由只学习轨迹改为学习运动意图与行为流程,从而注重整体与关联性。模仿人的技能学习过程,研究结合先验知识的装配运动与行为的模仿学习方法,包括在模仿过程中如何运用已有经验,从演示操作和输入知识中获取技能的关键因素等,有望实现从简单演示到技能获取的飞跃;研究从学习到的技能到行为的生成方法,包括根据当前状态和所学到的技能实时生成运动序列的方法,结合行为与状态进行带约束的动态路径规划方法等,能够跨越操作技能到具体操作运动指令的鸿沟。因此,仿照人的技能学习过程,充分运用先验知识,基于模仿学习通过简单演示获取技能,是技能获取未来重要的发展方向之一。对于已经获取的装配技能,脱离具体执行机构的技能提纯以及结合具体执行机构的实例化再现,是操作技能生成运动指令面临的挑战,也是技能再现未来的重要发展方

向之一。

在当前的机器人轴孔自动装配中,强化学习在装入阶段主要用于学习装入策略。学习到的装入策略以参数形式存储于神经网络内,形成神经网络控制器。如果将其用作唯一控制器或者主控制器,则难以保证控制系统的性能。为了更好地发挥强化学习的优势,有必要将基于强化学习的神经网络控制器与其他智能控制器结合起来。将基于强化学习的神经网络控制器用作辅助控制器,以智能控制器为主控制器,以多轴孔机器人装配为背景,研究基于强化学习的运动控制方法,包括对智能控制器的参数进行学习优化,将模仿学习和强化学习两者优势结合的方法等,有利于实现多轴孔的高效、安全装配。因此,将基于强化学习的神经网络控制器与其他智能控制器结合,是基于强化学习的轴孔装配运动控制系统未来的重要发展方向。将基于模型的方法与基于人工智能的方法结合,有利于取长补短,是多轴孔装配控制未来发展的重要方向之一。

机器人学习到的技能普适性较差,并且不同的技能之间复用性差,难以将已有技能向新任务迁移,远远没有达到融会贯通、触类旁通的程度。为了弥补上述不足,有必要赋予机器人有效积累知识并不断进化完善技能的能力。仿照人对新任务的学习过程,研究技能增量式学习方法,包括新技能的学习方法、知识积累算法等,实现对技能的不断学习与完善,是提高技能适应性的重要途径。研究面向任务的知识与技能实例化方法,包括知识与技能的关联算法、技能与任务的匹配算法、不同技能片段的融合复用方法等,实现对已有知识与技能的有效运用,是进行技能复用与再生的核心,也是装配技能学习的重要发展趋势之一。

#### 5 结束语

随着人工智能技术的快速发展,将模仿学习、强化学习等学习机制引入机器人轴孔装配系统,已成为近年来轴孔装配的重要发展趋势,并逐渐成为轴孔装配研究的热点。但机器人自动轴孔装配与人相比还有很大差距,从技能学习的角度来看,一方面模仿学习过于关注运动轨迹,对运动意图和环境约束关注较少;另一方面强化学习只在局部步骤上应用,远远没有达到优化整个装配过程的程度。从技能迁移与复用的角度来看,一方面机器人学习得到的技能仍然过于单一,对任务的针对性强,对应

用条件的要求比较严格,难以形成经验的有效积累与技能的不断进化完善;另一方面学习到的技能的灵活性和普适性较差,不同的技能之间复用性差,难以将已有技能向新任务迁移,远远没有达到融会贯通、触类旁通的程度。从知识使用的角度来看,人类的先验知识难以加入学习机制中,导致学习效率偏低。上述问题的解决必将加快装配智能化前进的步伐。

与单轴孔装配相比,多轴孔装配更加复杂。针对多轴孔的机器人自动装配研究尚处于起步阶段,相关研究较少。多轴孔装配过程中接触状态的复杂性和装配动作的技巧性,决定了多轴孔自动装配是很好的研究对象。以多轴孔装配为背景,仿照人的装配过程,能够系统深入地探讨技能学习机制、先验知识运用等问题。因此,以多轴孔装配为研究对象,将基于模型的方法与基于学习的人工智能方法结合,将先验知识加入学习机制中,将是轴孔装配智能化未来重要的发展趋势。

## 参考文献:

- [1] JIANG J G, HUANG Z Y, BI Z M, et al. State-of-the-art control strategies for robotic PIH assembly[J]. *Robotics and Computer-Integrated Manufacturing*, 2020(65): 101894.
- [2] ABU-DAKKA F J, NEMEC B, KRAMBERGER A, et al. Solving peg-in-hole tasks by human demonstration and exception strategies[J]. *Industrial Robot: an International Journal*, 2014, 41(6): 575-584.
- [3] XU J, HOU Z M, WANG W, et al. Feedback deep deterministic policy gradient with fuzzy reward for robotic multiple peg-in-hole assembly tasks[J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(3): 1658-1667.
- [4] BELTRAN-HERNANDEZ C C, PETIT D, RAMIREZ-ALPIZAR G I, et al. Variable compliance control for robotic peg-in-hole assembly: a deep-reinforcement-learning approach[J]. *Applied Sciences-Basel*, 2020, 10(19): 6923.
- [5] SONG J Z, CHEN Q L, LI Z D. A peg-in-hole robot assembly system based on Gauss mixture model[J]. *Robotics and Computer-Integrated Manufacturing*, 2021, 67: 101996.
- [6] 朱博文, 马立, 陈韦男, 等. 自动轴孔装配系统的现状与关键技术[J]. *现代制造工程*, 2019(5): 156-161.  
ZHU B W, MA L, CHEN W N, et al. Present situation and key technology of the automatic peg-in-hole assembly system[J]. *Modern Manufacturing Engineering*, 2019(5): 156-161.
- [7] 王鹏, 吴成东, 张凤. 机器人轴孔装配控制技术研究进展[J]. *科技广场*, 2007(3): 6-10.  
WANG P, WU C D, ZHANG F. The research progress on robot peg-in-hole assembly control technology[J]. *Science Mosaic*, 2007(3): 6-10.
- [8] JASIM I F, PLAPPER P W. Contact-state monitoring of force-guided robotic assembly tasks using expectation maximization-based Gaussian mixtures models[J]. *The International Journal of Advanced Manufacturing Technology*, 2014, 73(5/6/7/8): 623-633.
- [9] 潘柏松, 颜天野, 胡鑫达, 等. 基于几何约束与隐马尔可夫链模型的轴孔装配策略[J]. *计算机集成制造系统*, 2021: 已录用.  
PAN B S, YAN T Y, HU X D, et al. Peg-in-hole assembly strategy based on geometric constraint and hidden Markov model[J]. *Computer Integrated Manufacturing Systems*, 2021: accepted.
- [10] TAKAHASHI J, FUKUKAWA T, FUKUDA T. Passive alignment principle for robotic assembly between a ring and a shaft with extremely narrow clearance[J]. *IEEE/ASME Transactions on Mechatronics*, 2016, 21(1): 196-204.
- [11] WATSON J, MILLER A, CORRELL N. Autonomous industrial assembly using force, torque, and RGB-D sensing[J]. *Advanced Robotics*, 2020, 34(7/8): 546-559.
- [12] SLOTH C, KRAMBERGER A, ITURRATE I. Towards easy setup of robotic assembly tasks[J]. *Advanced Robotics*, 2020, 34(7/8): 499-513.
- [13] HOPPE S, LOU Z Y, HENNES D, et al. Planning approximate exploration trajectories for model-free reinforcement learning in contact-rich manipulation[J]. *IEEE Robotics and Automation Letters*, 2019, 4(4): 4042-4047.
- [14] 李侃, 汪法根, 廖正菊, 等. 一种用于多轴孔装配的机器人系统研究[J]. *机械设计与研究*, 2003, 19(3): 25-26.  
LI K, WANG F G, LIAO Z J, et al. Study on a robotic system for multi-axial bore assembling[J]. *Machine Design and Research*, 2003, 19(3): 25-26.
- [15] 魏明明, 傅卫平, 蒋家婷, 等. 操作机器人轴孔装配的行为动力学控制策略[J]. *机械工程学报*, 2015, 51(5): 14-21.  
WEI M M, FU W P, JIANG J T, et al. Dynamics of behavior control strategy in peg-in-hole assembly task of manipulator[J]. *Journal of Mechanical Engineering*, 2015, 51(5): 14-21.
- [16] 徐远, 宋爱国, 李会军. 基于机器视觉和力反馈的自动装配技术研究[J]. *测控技术*, 2019, 38(4): 11-16.  
XU Y, SONG A G, LI H J. Research on automatic assembly technology based on machine vision and force feedback[J]. *Measurement & Control Technology*, 2019, 38(4): 11-16.
- [17] 陈婵娟, 赵飞飞, 李承, 等. 多传感器协助机器人精确装配[J]. *机械设计与制造*, 2020(3): 281-284.  
CHEN C J, ZHAO F F, LI C, et al. Multi-sensor assisted robotic accurate assembly[J]. *Machinery Design & Manufacture*, 2020(3): 281-284.
- [18] ZENG F, XIAO J L, LIU H T. Force/torque sensorless compliant control strategy for assembly tasks using a 6-DOF collaborative robot[J]. *IEEE Access*, 2019, 7: 108795-108805.
- [19] PARK H, PARK J, LEE D H, et al. Compliance-based robotic peg-in-hole assembly strategy without force feedback[J]. *IEEE Trans-*

- actions on Industrial Electronics, 2017, 64(8): 6299-6309.
- [20] 陈钢, 王玉琦, 贾庆轩, 等. 机器航天员轴孔装配过程中的力位混合控制方法[J]. 宇航学报, 2017, 38(4): 410-419.  
CHEN G, WANG Y Q, JIA Q X, et al. Hybrid force and position control strategy of robonaut performing peg-in-hole assembly task[J]. Journal of Astronautics, 2017, 38(4): 410-419.
- [21] 吴炳龙, 曲道奎, 徐方. 基于力/位混合控制的工业机器人精密轴孔装配[J]. 浙江大学学报(工学版), 2018, 52(2): 379-386.  
WU B L, QU D K, XU F. Industrial robot high precision peg-in-hole assembly based on hybrid force/position control[J]. Journal of Zhejiang University (Engineering Science), 2018, 52(2): 379-386.
- [22] 董恣, 张立建, 易旺民, 等. 基于动力学前馈的空间机器人多销孔装配力柔顺控制[J]. 机械工程学报, 2019, 55(4): 207-217.  
DONG Q, ZHANG L J, YI W M, et al. Force compliance control of multi-peg-in-hole assembling by space robot based on dynamic feed-forward[J]. Journal of Mechanical Engineering, 2019, 55(4): 207-217.
- [23] XING D P, LIU F F, XU D, et al. Simultaneous control in belief space for circular insertion in precision assembly[J]. IEEE Transactions on Industrial Informatics, 2021, 17(3): 1842-1851.
- [24] XING D P, LIU X W, LIU F F, et al. Efficient insertion strategy for precision assembly with uncertainties using a passive mechanism[J]. IEEE Transactions on Industrial Informatics, 2021, 17(2): 1263-1273.
- [25] XING D P, LIU F F, XU D. Joint alignment and simultaneous insertion of multiple objects in precision assembly[J]. IEEE Transactions on Industrial Informatics, 2021, 17(1): 230-239.
- [26] XING D P, LIU F F, XU D. Efficient coordinated control strategy to handle randomized inclination in precision assembly[J]. IEEE Transactions on Industrial Informatics, 2020, 16(9): 5814-5824.
- [27] LIU S, XING D P, LI Y F, et al. Robust insertion control for precision assembly with passive compliance combining vision and force information[J]. IEEE/ASME Transactions on Mechatronics, 2019, 24(5): 1974-1985.
- [28] 曲吉旺, 许家忠, 张大朋, 等. 基于显微视觉的微球微管精密装配[J]. 高技术通讯, 2019, 29(9): 914-924.  
QU J W, XU J Z, ZHANG D P, et al. Precision assembly for micro-sphere and micro-tube based on microscopic vision[J]. Chinese High Technology Letters, 2019, 29(9): 914-924.
- [29] XING D P, LYU Y, LIU S, et al. Efficient insertion of multiple objects parallel connected by passive compliant mechanisms in precision assembly[J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 4878-4887.
- [30] XING D P, LIU F F, LIU S, et al. Efficient insertion of partially flexible objects in precision assembly[J]. IEEE Transactions on Automation Science and Engineering, 2019, 16(2): 706-715.
- [31] XING D P, LIU F F, LIU S, et al. Efficient collision detection and detach control for convex prisms in precision manipulation[J]. IEEE Transactions on Industrial Informatics, 2018, 14(12): 5316-5326.
- [32] LIU S, XU D, LIU F F, et al. Relative pose estimation for alignment of long cylindrical components based on microscopic vision[J]. IEEE/ASME Transactions on Mechatronics, 2016, 21(3): 1388-1398.
- [33] XING D P, LIU F F, QIN F B, et al. Coordinated insertion control for inclined precision assembly[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 2990-2999.
- [34] JASIM I F, PLAPPER P W, VOOS H. Adaptive sliding mode fuzzy control for unknown robots with arbitrarily-switched constraints[J]. Mechatronics, 2015, 30: 174-186.
- [35] NOTTENSTEINER K, SACHTLER A, ALBU-SCHAEFFER A. Towards autonomous robotic assembly: using combined visual and tactile sensing for adaptive task execution[J]. Journal of Intelligent and Robotic Systems, 2021, 101(3): 49.
- [36] CORTESAO R, KOEPE R, NUNES U, et al. Data fusion for robotic assembly tasks based on human skills[J]. IEEE Transactions on Robotics, 2004, 20(6): 941-952.
- [37] PETRIČ T, GAMS A, COLASANTO L, et al. Accelerated sensorimotor learning of compliant movement primitives[J]. IEEE Transactions on Robotics, 2018, 34(6): 1636-1642.
- [38] LIU N L, ZHOU X D, LIU Z M, et al. Learning peg-in-hole assembly using Cartesian DMPs with feedback mechanism[J]. Assembly Automation, 2020, 40(6): 895-904.
- [39] SONG H C, KIM Y L, SONG J B. Guidance algorithm for complex-shape peg-in-hole strategy based on geometrical information and force control[J]. Advanced Robotics, 2016, 30(8): 552-563.
- [40] QIN F B, XU D, ZHANG D P, et al. Robotic skill learning for precision assembly with microscopic vision and force feedback[J]. IEEE/ASME Transactions on Mechatronics, 2019, 24(3): 1117-1128.
- [41] DING G W, LIU Y B, ZANG X Z, et al. A task-learning strategy for robotic assembly tasks from human demonstrations[J]. Sensors, 2020, 20(19): 5505.
- [42] PEREZ-DEL-PULGAR C J, SMISEK J, RIVAS-BLANCO I, et al. Using Gaussian mixture models for gesture recognition during haptically guided telemanipulation[J]. Electronics, 2019, 8(7): 772.
- [43] GAO X, LING J, XIAO X H, et al. Learning force-relevant skills from human demonstration[J]. Complexity, 2019: 5262859.
- [44] TI B Y, GAO Y S, LI Q, et al. Human intention understanding from multiple demonstrations and behavior generalization in dynamic movement primitives framework[J]. IEEE Access, 2019, 7: 36186-36194.
- [45] SENDA K, TANI Y. Autonomous robust skill generation using reinforcement learning with plant variation[J]. Advances in Mechanical Engineering, 2014(6): 276264.
- [46] KIM Y L, AHN K H, SONG J B. Reinforcement learning based on movement primitives for contact tasks[J]. Robotics and Computer-Integrated Manufacturing, 2020, 62: 101863.
- [47] KHADER S A, YIN H, FALCO P, et al. Stability-guaranteed reinforcement learning for contact-rich manipulation[J]. IEEE Robotics and Automation Letters, 2021, 6(1): 1-8.
- [48] SENDA K, HISHINUMA T, TANI Y. Approximate Bayesian reinforcement learning based on estimation of plant[J]. Autonomous Robots, 2020, 44(5): 845-857.

- [49] DONG S, NAGHDY F. Application of hidden Markov model to acquisition of manipulation skills from haptic rendered virtual environment[J]. Robotics and Computer-Integrated Manufacturing, 2007, 23(3): 351-360.
- [50] CHO N J, LEE S H, KIM J B, et al. Learning, improving, and generalizing motor skills for the peg-in-hole tasks based on imitation learning and self-learning[J]. Applied Sciences-Basel, 2020, 10(8): 2719.
- [51] 王斐, 齐欢, 周星群, 等. 基于多源信息融合的协作机器人演示编程及优化方法[J]. 机器人, 2018, 40(4): 551-559.
- WANG F, QI H, ZHOU X Q, et al. Demonstration programming and optimization method of cooperative robot based on multi-source information fusion[J]. Robot, 2018, 40(4): 551-559.
- [52] MA Y Q, XU D, QIN F B. Efficient insertion control for precision assembly based on demonstration learning and reinforcement learning[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 4492-4502.
- [53] LI Y, XU D. Skill learning for robotic insertion based on one-shot demonstration and reinforcement learning[J]. International Journal of Automation and Computing, 2021, 18(3): 457-467.
- [54] 刘乃龙, 刘钊铭, 崔龙. 基于深度强化学习的仿真机器人轴孔装配

研究[J]. 计算机仿真, 2019, 36(12): 296-301.

LIU N L, LIU Z M, CUI L. Deep reinforcement learning based robotic assembly in simulation[J]. Computer Simulation, 2019, 36(12): 296-301.

#### 作者简介:



徐德 (1965-), 男, 博士, 中国科学院自动化研究所研究员, 主要研究方向为机器人视觉测量、视觉控制、智能控制、视觉定位、显微视觉、微装配、技能学习。



秦方博 (1992-), 男, 博士, 中国科学院自动化研究所副研究员, 主要研究方向为机器人视觉感知、精密装配机器人、植入手术机器人、深度学习。