

# 机器人操作技能模型综述

秦方博<sup>1,2</sup> 徐德<sup>1,2</sup>

**摘要** 机器人技能学习是人工智能与机器人学的交叉领域,目的是使机器人通过与环境 and 用户的交互得到经验数据,基于示教学习或强化学习,从经验数据中自主获取和优化技能,并应用于以后的相关任务中.技能学习使机器人的任务部署更加灵活快捷和用户友好,而且可以让机器人具有自我优化的能力.技能模型是技能学习的基础和前提,决定了技能效果的上限.日益复杂和多样的机器人操作任务,对技能操作模型的设计实现带来了许多挑战.本文给出了技能操作模型的概念与性质,阐述了流程、运动、策略和效果预测四种技能表达模式,并对其典型应用和未来趋势做出了概括.

**关键词** 机器人操作, 智能机器人, 机器人学习, 技能学习, 自主系统

**引用格式** 秦方博, 徐德. 机器人操作技能模型综述. 自动化学报, 2019, 45(8): 1401–1418

**DOI** 10.16383/j.aas.c180836

## Review of Robot Manipulation Skill Models

QIN Fang-Bo<sup>1,2</sup> XU De<sup>1,2</sup>

**Abstract** Robot skill learning is an interdisciplinary field of artificial intelligence and robotics. The aim is to enable that a robot autonomously obtains a certain manipulation skill via imitation learning or reinforcement learning, based on the experience data generated by the robot's interaction with the environment or human users. Skill learning can make the robot task deployments more flexible, efficient and user-friendly, even realize the sustainable self-improvement of robot. Robot skill model is the foundation of skill learning, which determines upper limit of the skill outcome. The complexity and variety of the modern robot manipulation tasks pose many challenges to the robot skill modeling. This paper describes the definition and characteristics of robot manipulation skill model, introduces the four skill representation methods considering procedure, motion, policy and outcome, respectively. Finally, the typical application and future tendency of robot manipulation skill models are presented.

**Key words** Robot manipulation, intelligent robot, robot learning, skill learning, autonomous system

**Citation** Qin Fang-Bo, Xu De. Review of robot manipulation skill models. *Acta Automatica Sinica*, 2019, 45(8): 1401–1418

机器人操作技能 (Robot manipulation skill) 是指机器人基于自身的传感、感知、决策、规划与控制能力,在有限时间内操作环境中特定物体,使物体由初始状态达到目标状态.技能在生活和生产中无处不在,例如物流中的拆垛与码垛、生产线上的零部件装配、餐厅中的端茶递水、体育运动中的打乒乓球等,机器人被寄希望于协助甚至代替人类完成越来越多、越来越复杂的技能.早在 1985 和 1988 年

分别已有德国和日本学者对机器人技能学习进行研究<sup>[1-2]</sup>.目前机器人技能大多通过人工预定义的规则实现,而人类主要是通过模仿和与环境交互获取技能,在多模态感知、自主决策、以及运动的灵活性和自适应性上,机器人与人类相比还有较大差距<sup>[3]</sup>.随着机器人软硬件技术的不断提高,以及近几年来人工智能技术的迅速发展,为机器人操作技能建模与学习带来新的机遇.

2015 年 5 月,国务院印发《中国制造 2025》,明确提出“以加快新一代信息技术与制造业深度融合为主线,以推进智能制造为主攻方向”.2017 年 7 月,国务院发布了《新一代人工智能发展规划》,将人工智能定位为国家战略,提出了明确的发展目标,即到 2020 年“人工智能技术应用成为改善民生的新途径”,到 2030 年“人工智能理论、技术与应用总体达到世界领先水平”.文献 [4] 中认为智能机器人行业是未来“脑科学研究”和“脑认知与类脑计算”研究成果的重要产出方向.机器人已经在大中型生产厂商中广泛应用,例如汽车行业的生产线,但是机器

收稿日期 2018-12-17 录用日期 2019-03-19  
Manuscript received December 17, 2018; accepted March 19, 2019

国家重点研究发展计划 (2018YFD0400902), 国家自然科学基金 (61873266, 61733004) 资助

Supported by National Key Research and Development Program of China (2018YFD0400902) and National Natural Science Foundation of China (61873266, 61733004)

本文责任编辑 贺威

Recommended by Associate Editor HE Wei

1. 中国科学院自动化研究所精密感知与控制研究中心 北京 100190  
2. 中国科学院大学人工智能学院 北京 101408

1. Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190  
2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 101408

人在小型生产厂商的应用中受限于技能编程和工具定制的高成本,而且其生产环境经常是非结构化的,例如各车间的生产工具和工作台布局存在不一致.在全球化市场竞争当中,产品需求变化很快,制造系统的功能也要随之改变,所以带来更短的迭代周期、更多样化的种类和小批量定制化的生产.近年来越来越多的协作机器人出现在市场中,例如 Rethink Robotics Baxter、Universal Robots UR10、ABB YuMi、KUKA IIWA 等,由于协作任务的动态性和多样性更强,所以对机器人协作技能的灵活部署能力和适应能力提出更高的要求.服务机器人在人类环境中工作往往要面对物体的空间排布不规则,例如物体可能会被放置在盒子、货架或柜子中,对机器人技能执行造成挑战<sup>[5]</sup>.综上所述,机器人操作技能的实现具有迫切的需求与重大的意义.

如图 1 所示,机器人系统从下到上的层级分别是机器人本体、感知与控制模块、技能模型与技能学习.机器人本体与物理环境直接交互,具有一系列执行器与传感器.感知与控制模块是机器人本体与技能模型之间的中介层,传感器对物理环境的原始观测信号,经过感知模块被加工为感兴趣的状态变量,技能模型所产生的运动指令,经过控制模块来产生机器人本体的运动.技能模型进行决策、规划和预测的高层信息处理,其输入输出往往与机器人本体不直接联系.特殊地,策略模型中的端到端模型可以实现传感器原始信号到底层驱动信号的直接映射,即将感知与控制模块也包含在技能模型中.技能模型被认为是在一定范围内通用的,一定的参数配置可将技能模型实例化为具体的技能,例如对于拾放 (Pick-and-place) 技能模型,通过改变其参数如物体种类、运动轨迹、目标位置等,可获得摆放茶杯、零件传送等不同的技能实例.最后,技能参数的获取可通过技能学习实现,技能学习类型包括示教学习 (Learning from demonstration)<sup>[6]</sup>、强化学习 (Reinforcement learning)<sup>[7]</sup> 等.技能模型显式或隐式地决定了技能效果的上限,技能学习解决的问题是如何基于经验数据获取模型参数,使技能效果趋近该上限.

机器人操作技能模型应具有如下特点: 1) 紧凑性.技能模型应具有紧凑的数学表达,对输入、输出和参数具有明确定义.换言之,并不是所有的机器人技能实现都可以称之为技能模型,通过繁冗的系统集成与调试工作同样可以实现技能,但是不具备紧凑性,难以实现模型的分析、学习和泛化应用. 2) 综合性.机器人技能最大的特点是需要与环境进行物理交互,而非仅仅运行于虚拟数据空间.所以技能模型需要考虑与环境感知和运动控制模块的紧密融合. 3) 稳定性.机器人技能执行是一个持续的动态过程,

会受到环境不确定性、自身观测和过程噪声的影响,最终应实现被操作物体的状态收敛于目标状态. 4) 实时性.模型计算导致的延迟,会对机器人性能产生不良的影响,甚至造成系统的不稳定,所以需要技能模型具有足够的计算实时性. 5) 可学习性.硬编码的技能往往需要大量的专家开发调试工作,且没有跨任务复用的能力,所以部署效率低、成本高.具有可学习性的技能模型,可以通过示教学习、强化学习或用户交互,实现技能的获取、适应和优化,从而使机器人的跨任务部署更加灵活和低成本. 6) 安全性.机器人与环境的安全交互至关重要,环境中的人和物体不可受到损害,其次也要避免机器人自身受到损害.技能模型应引导机器人在安全范围内工作,技能执行效果应受到实时的监督,对异常或危险情况做出及时的反应. 7) 复杂性.为了表示复杂技能,例如长时间多步骤的技能,或对不规则物体的灵巧操作技能,需要更高的模型复杂度,而高复杂度模型的设计和实现具有很大的挑战性.因为不仅要实现复杂功能,同时也要考虑技能的紧凑性、综合性、稳定性、安全性以及可学习性的保证所带来的制约.为满足复杂技能的表达,同时满足以上性质,技能模型的设计与实现仍面对很大挑战,属于开放的研究问题.

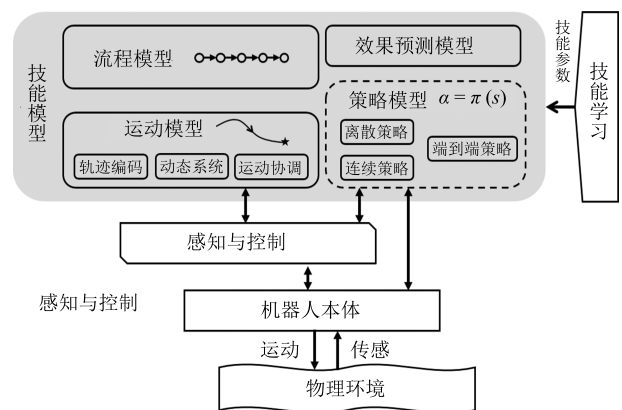


图 1 机器人操作技能模型框图

Fig. 1 Diagram of robot manipulation skill model

本文内容安排如下:第 1 节介绍技能流程模型;第 2 节介绍运动模型,包括轨迹编码、动态系统以及运动协调三种类型的运动模型;第 3 节介绍策略模型,包括离散策略、连续策略以及端到端策略;第 4 节介绍效果预测模型;第 5 节阐述了机器人操作技能模型的典型应用和未来趋势;最后对全文进行了总结.

## 1 流程模型

技能表示具有层次性,最底层的技能表示是机

器人和环境物体的具体状态的一段连续轨迹, 通过由下而上的抽象, 可以将技能表示为与具体硬件无关的抽象语义级的流程. 例如, 机器人轴孔装配可以表示为机械手和零件位姿的一段连续轨迹, 通过分割和抽象可表示为“趋近轴-接触轴-抓握轴-趋近孔-位姿对准-柔顺插入-撤离复位”的流程, 进一步地抽象表示为“抓取轴-插入轴孔-撤离复位”, 甚至可以进一步抽象为“装配轴孔”, 而多个“装配轴孔”步骤可以组成更复杂的装配技能流程. 所以, 技能具有层次性, 对层次的划分目前并没有统一标准, 例如文献 [8] 中仅用一个策略即可表达开门技能, 不需要再将开门过程划分为更细的流程步骤.

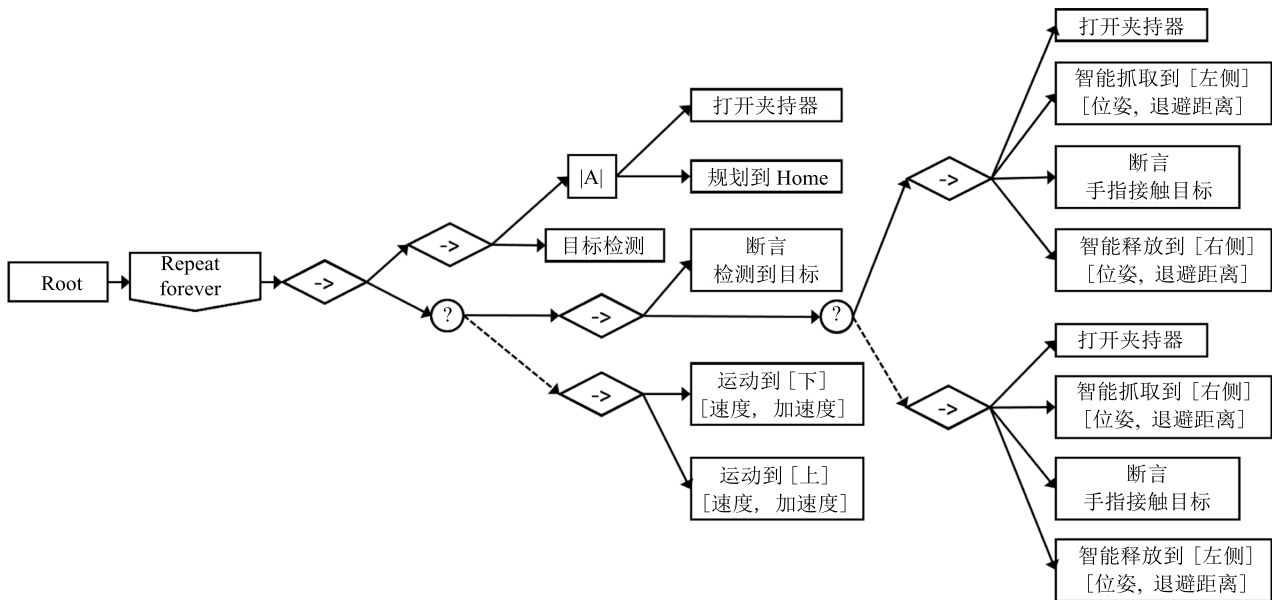
基于流程的技能表示的优点在于: 1) 实现复杂技能表达. 将一个多步骤的复杂技能, 仅用一个运动或策略模型来表达是非常困难的, 若将复杂技能划分为流程, 则流程中的每一个步骤均可以被表示为更简单紧凑的模型. 2) 使技能更容易维护和修改. 由于流程模型将整体技能划分为模块, 增删和修改其中某一个模块时, 对其他步骤的模块不产生影响, 否则, 对技能的局部改变, 可能影响到整个技能. 3) 抽象的语义级流程, 使机器人技能表示易于被人类理解, 且因为不依赖于具体硬件, 所以可应用于不同的机器人系统.

流程模型首先需要定义一系列行动基元 (Action primitive 或 Atomic action), 记为  $\mathcal{A}$ , 然后利用图模型或知识推理等方法, 对行动基元的种类选择和先后顺序进行表示, 技能执行时流程模型会生成  $N$  个行动基元的序列  $\{\langle a_1, \theta_1 \rangle, \langle a_2, \theta_2 \rangle, \dots, \langle a_N, \theta_N \rangle\}$ , 其中,  $a_i \in \mathcal{A}$  且  $\theta_i$  是行动基元的参数. 有限状态机 (Finite state machine, FSM) 是一种简洁易用的方法, 其假设系统具有有限个状态, 每个时刻只能处于一种状态, 在一定条件下可进行状态转移<sup>[9]</sup>. 为表达机器人技能, FSM 的状态节点可以是机器人的行动、查询或断言, FSM 的拓扑结构以及状态转移条件表达了技能的流程语义. FSM 在状态数量较少时具有直观易懂的优点, 但是在状态数量很多时, FSM 的节点多、节点间连接复杂, 导致其难以理解、维护和扩展, 尤其在技能流程并非简单顺序式, 而是具有选择执行、失败返回等流程. 层次式有限状态机 (Hierarchical finite state machine, HFSM) 假设一个超状态 (Super state) 可以包含多个状态, 超状态之间的转移被称为广义转移<sup>[10]</sup>. HFSM 相对于 FSM 具有了可模块化的优势, 即修改超状态内部的结构, 对超状态之间的连接结构没有影响.

行为树 (Behavior tree) 是一种具有可扩展性和层次性的流程表达方式, 而且具有可读性的优点, 即行为树的含义可以通过文字或可视化图形来展

示给用户. 行为树由根节点 (Root)、控制流节点 (Control flow node) 和执行节点 (Execution node) 构成<sup>[11]</sup>. 在每一个执行周期, 先由根节点发出执行信号, 通过访问子节点逐层传播, 经过控制流节点并终止于某一执行节点, 该执行节点向机器人发出当前周期的行为指令. 每一个子节点都具有三种状态信号, 即运行、成功、失败, 父节点可以访问子节点的状态信号. 控制流节点根据子节点的状态信号决定下一个访问节点, 包括选择节点、序列节点、并行节点、重复节点以及装饰节点, 分别用于表达流程中的选择执行、顺序执行、并行执行、重复执行以及自定义策略. 执行节点包括条件节点以及行动节点, 分别用于条件查询以及行动执行. 直观上理解, 行为树是一种基于规则的系统, 描述了在某一阶段和条件下应该执行哪一种行动. 文献 [12] 中将机器人的系统功能 (System capability) 表示为一个三元组  $\langle$ 操作集合, 输入集合, 输出集合 $\rangle$ , 输入集合和输出集合取决于机器人硬件, 操作集合是语义级的技能表达, 包括行动、查询以及断言. 在系统功能的定义基础上, 可继续定义技能规范 (Skill specification), 同样表达为一个三元组  $\langle$ 行为树, 系统功能, 规范参数 $\rangle$ . 在一定的系统功能下, 可以通过行为树来定义技能的语义级表达. 在不改变行为树与系统功能的情况下, 针对具体场景实例化规范参数, 可以使机器人的技能适应场景中的条件变化, 例如待抓取零件种类变化后可以仅更改“零件图像模板”来适应抓取技能. 如图 2 所示, Paxton 等<sup>[13]</sup> 开发了 CoSTAR (Collaborative system for task automation and recognition) 系统, 基于行为树实现了使用户可以理解技能和编辑技能的人机交互界面, 系统从传感器中提取抽象的感知信息, 使技能表达不受限于环境或机器人本体的细节. 例如, CoSTAR 定义了 SmartMoves, 可以抽象表达“从桌子右边抓起物品 A”这种技能目标, 类似的抽象表达符合人类对技能的思考方式, 而不是利用机器人的具体传感信息例如固定的位姿, 从而使系统应用对用户更加友好<sup>[14]</sup>. 文献 [15] 基于行为树开发了程序库 BART (Behaviour architecture for robotic tasks), 用于组织和管理有限能力的行动基元以表达复杂的任务技能. 文献 [16] 提出随机行为树 (Stochastic behavior tree) 来评估行为树的成功率以及执行时间等指标.

隐马尔科夫模型 (Hidden Markov model, HMM) 是一种概率图模型, 具有有限个隐含状态, 且隐含状态之间可以进行随机状态转移, 隐含状态无法被直接观测, 而是输出相应的观测变量序列. Akgun 等<sup>[17]</sup> 利用 HMM 表示技能的执行模型和目标模型, 两个模型互相独立, HMM 的观测变量主要来自运动学位姿和 RGBD 相机中提取到的目标状

图 2 基于行为树的技能流程表示<sup>[14]</sup>Fig. 2 Behavior tree based skill procedure representation<sup>[14]</sup>

态, 隐含状态可以视为技能的流程节点, 执行技能时通过 Dijkstra 算法产生隐含状态节点间的最短路径. 目标模型基于 HMM 计算观测序列的似然度, 如果似然度大于阈值, 说明技能执行成功. 文献 [18] 将上述方法拓展到机器人技能自我提升, 即将目标模型的输出用于优化执行模型. Kroemer 等<sup>[19]</sup> 利用基于状态的转移自回归隐马尔科夫模型 (State-based transitions autoregressive hidden Markov model, STARHMM) 对多阶段技能进行建模, 最佳阶段数量基于谱聚类 and 试凑法得到. HMM 的隐含变量是技能阶段, 观测变量是机器人和环境的状态, 状态转移通过行动基元来实现. 通过将状态转移概率作为相应阶段的回报函数, 该模型可用于指导行动基元的强化学习. 文献 [20] 将 HMM 与线性参数变化系统 (Linear parameter varying, LPV) 系统结合为 HMM-LPV 模型, 用于表达稳定的多步骤技能, HMM 的节点表示行动基元, 且行动基元是否完成由一个二值终止状态描述, 终止状态判断保证了行动基元之间的正确切换, 基于该模型可以实现重复执行的多步骤技能, 例如具有“接触-切削-后撤”三个循环重复步骤的水果削皮技能.

Pardowitz 等<sup>[21]</sup> 提出了一种任务优先级图 (Task precedence graph), 以操作为节点形成有向图, 有向图的边用于表示技能中各个操作的优先级关系. 行为网络 (Behavior network) 利用有向无环图来表示技能中各个行为间的拓扑关系, 包括永久前提条件、使能前提条件和顺序前提条件<sup>[22]</sup>. 文献 [23] 提出一种团/链层次式任务网络 (Cliques/chain

hierarchical task network), 其拓扑性质可以从任务图中自动提取获得. Ahmadzadeh 等<sup>[24]</sup> 提出视空间技能 (Visuospatial skill), 主要关注视觉图像空间中的目标位置和方向, 并实现学习和复现平面上抓取-放置的技能. 该方法从示教中记录一个 (抓取图像观测, 放置图像观测) 的序列, 然后在复现技能时通过图像匹配方法将观测序列转换为 (抓取位置, 抓取方向, 放置位置, 放置方向) 的序列, 通过循环执行序列中每一个元素定义的抓取-放置位姿来复现技能. 图像经过图像空间到工作平面的单应性矩阵 (Homography matrix) 矫正, 而单应性矩阵在不同视角下可以重新标定, 所以该视空间技能具有视角不变性. 文献 [25] 将视空间技能学习得到的每个操作的前提条件和效果进一步用于符号级规划 PDDL (Planning domain definition language). 文献 [26] 基于一种符号级任务规划器, 即时域快速向下模块 (Temporal fast downward modules) 来实现机器人的高层控制, 该规划器工作于观测-监视-执行的循环中, 并且可以利用语义附件来实现断言. Beetz 等<sup>[27]</sup> 提出一种认知机器人抽象机 (Cognitive robot abstract machine), 使机器人可以通过感知信息和知识库自动推断技能流程, 且可以实现并行行动、失败情况处理以及对环境变化的响应. KnowRob 是一种机器人知识处理 (Knowledge processing) 系统, 用于桥接模糊的技能描述知识与机器人执行技能时需要的具体信息, 通过知识库 (Knowledge base) 的应答来实现技能推断<sup>[28]</sup>, 且可用于云机器人之间的知识交换<sup>[29]</sup>.

## 2 运动模型

运动是直接产生技能效果的必要途径, 运动模型通过离线或在线生成一段连续的机器人空间状态表示技能, 具有与物理系统相关的明确含义, 例如位置、姿态和接触力. 一般地, 假设运动是非线性的. 运动模型主要包括轨迹编码、运动系统以及运动协调三种模式.

### 2.1 轨迹编码

轨迹编码是用紧凑的数学模型对轨迹的形状、约束等信息进行表示. 如果示教并存储一条或多条具体的固定轨迹, 然后执行技能时准确地复现该轨迹, 那么运动技能没有泛化能力, 而且轨迹存储占用很大存储空间. 轨迹编码可以用少量的参数和具有泛化能力的模型来表达技能轨迹, 主要通过建立时间-空间关联来实现, 即在给定时间  $t$  时, 空间状态的概率分布为

$$\mathbf{x} \sim \mathcal{T}(t; \Theta) \quad (1)$$

其中,  $\mathbf{x} \in \mathbf{R}^d$  是机器人状态, 例如机械手末端位置、姿态或作用力,  $\Theta$  是参数集合.  $\mathbf{x}$  的期望描述了轨迹的中心位置, 协方差描述了轨迹的不确定性, 如图 3 所示. 因为在很多情况下, 机器人并不需要严格沿着轨迹中心运动, 有一定范围内的偏离并不影响技能成功执行. 所以协方差信息所表示的不确定性, 可以

用于表示运动约束信息. 例如, 搬运物体时, 起始抓取位置和终止放置位置需要比较准确, 而中间过程中允许轨迹具有较大的不确定性; 而对于焊接机器人来说, 运动轨迹的不确定性始终很小, 使焊枪沿着焊缝精确运动.

Calinon 等<sup>[30]</sup> 用多元变量高斯混合模型 (Gaussian mixture model, GMM) 实现轨迹编码. 假设 GMM 有  $K$  个高斯分量, 多元变量由机器人空间状态和时间组成, 即  $\boldsymbol{\xi} = (\mathbf{x}, t)^T$ , GMM 参数为各高斯分量的均值与方差, 即  $\Theta = \{\theta_k\}_{k=1}^K = \{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$ , 则多元变量的概率密度函数为

$$p(\mathbf{x}, t) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}, t; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (2)$$

其中,  $\pi_k$  是高斯分量  $k$  的权重,  $\mathcal{N}$  是高斯概率密度函数. 在执行技能时, 为了重建随时间变化的轨迹, 用高斯混合回归 (Gaussian mixture regression, GMR) 来计算给定已知时间时状态的条件分布  $\mathbf{x} \sim p(\mathbf{x}|t; \Theta)$ , 其条件期望和条件协方差显式表达为

$$\begin{cases} \bar{\mathbf{x}} = \sum_{k=1}^K \beta_k \left[ \bar{\mathbf{x}}_k + \boldsymbol{\Sigma}_{xt,k} (\boldsymbol{\Sigma}_{t,k})^{-1} (t - \bar{t}_k) \right] \\ \boldsymbol{\Sigma}_x = \sum_{k=1}^K \beta_k \left[ \boldsymbol{\Sigma}_{x,k} - \boldsymbol{\Sigma}_{xt,k} (\boldsymbol{\Sigma}_{t,k})^{-1} \boldsymbol{\Sigma}_{tx,k} \right] \end{cases} \quad (3)$$

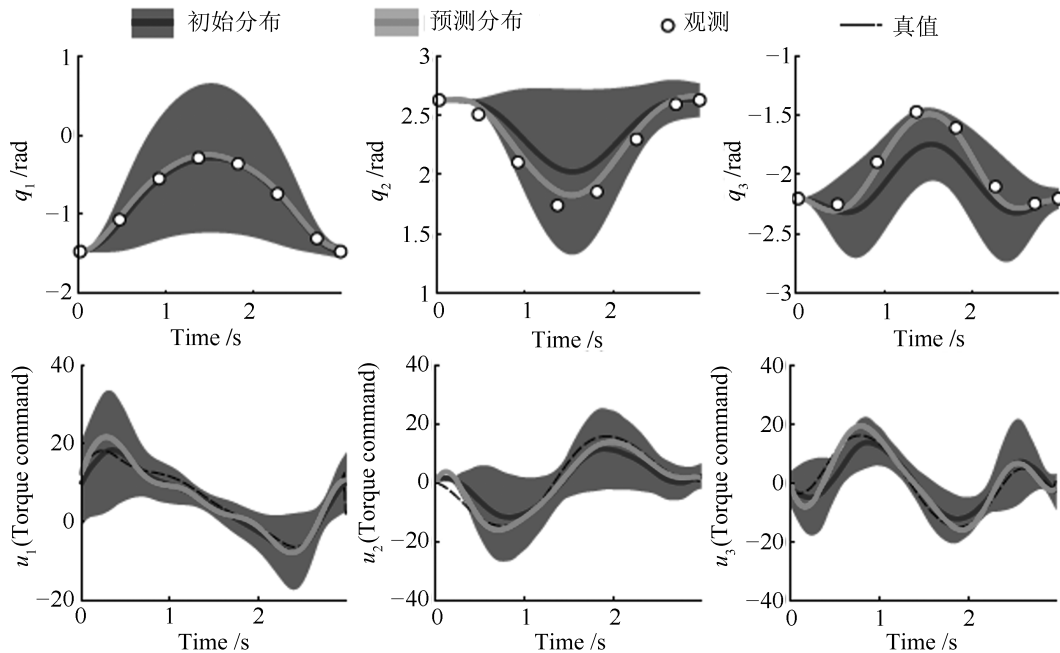


图 3 基于概率运动基元的轨迹编码<sup>[31]</sup>

Fig. 3 ProMP based trajectory encoding<sup>[31]</sup>

其中,  $\beta_k = \mathcal{N}(t|\bar{t}_k, \Sigma_{t,k}) / \sum_{i=1}^K \mathcal{N}(t|\bar{t}_i, \Sigma_{t,i})$ ,  $\bar{\mathbf{x}}_k$  和  $\bar{t}_k$  是期望  $\boldsymbol{\mu}_k$  的分量,  $\Sigma_{t,k}$ ,  $\Sigma_{x,k}$ ,  $\Sigma_{tx,k}$  和  $\Sigma_{xt,k}$  是协方差  $\Sigma_k$  的分量. 上述方法是在固定的坐标系下表达轨迹, 但是轨迹可以在不同参考坐标系中表达, 且参考坐标系的变换关系也可能发生变化, 所以 Calinon 等<sup>[32]</sup> 又提出任务参数化的高斯混合模型 (Task parameterized GMM, TP-GMM), 将坐标系的原点和旋转变换矩阵作为模型中的任务参数, 从而允许在不同的参考坐标系下观察和复现运动轨迹. 由于 TP-GMM 需要预先定义若干参考坐标系, 而并没有对各参考坐标系对于相关技能的重要性做区分, 文献 [33] 设计了基于置信度权重的框架, 允许用户将技能与各参考坐标系相关性的先验置信度知识用于 TP-GMM 的示教学习, 使用迭代特征选择方法确定重要参考坐标系, 并删除冗余参考坐标系. 任务参数化的半捆绑 GMM (TP-semi-tied GMM, TP-SGMM), 将具有共同协同基向量的高斯分量的协方差矩阵进行捆绑, 使技能中具有协调性的部分得到复用, 从而提高了模型对于新情景的适应能力<sup>[34]</sup>. 文献 [35] 中基于 TP-GMM 对位置和姿态轨迹同时进行编码, 并基于特殊正交旋转群实现姿态运动的复现.

轨迹编码模型生成的轨迹分布, 可以和最优控制结合, 优化目标函数是使复现轨迹与期望轨迹中心的偏差最小, 轨迹状态包括笛卡尔空间和关节空间的位置, 协方差矩阵的逆用于作为状态各分量偏差的权重, 同时加入机器人运动学模型作为约束, 通过迭代计算获取关节角的最优轨迹<sup>[30]</sup>. 在文献 [36] 中利用线性二次型调节器 (Linear quadratic regulator) 实现了最小干预控制器 (Minimal intervention controller), 其中轨迹编码模型输出的期望作为参考轨迹, 协方差矩阵用于调节控制器的刚度和阻尼系数, 同时最小化运动指令来实现最小干预控制. 为了学习人机协作技能, Rozo 等<sup>[37]</sup> 基于 TP-GMM 对位置和力的约束进行编码, 通过最优反馈控制器来最小化机器人的输出力矩和人的干预, 提高了人机交互的安全程度, 并利用了轨迹中位置、速度和力的不确定性来调节柔顺程度. 当状态变量的维度较高, 会对技能学习带来维度灾难, 主成分分析 (Principal component analysis, PCA) 等方法可将原始状态空间的高维状态变量投影到低维隐空间中, 然后在隐空间中学习轨迹, 有利于简化轨迹编码模型, 避免维度灾难<sup>[30]</sup>. 例如, 在平面运动例如擦桌子和直线运动例如焊接, 其三维运动轨迹可以分别投影到二维和一维空间进行表达. 在执行阶段, 隐空间中生成的轨迹应该恢复到原始空间后再发送给运动控制器.

Paraschos 等<sup>[38-39]</sup> 提出一种概率运动基元 (Probabilistic movement primitives, ProMPs) 模

型, 对轨迹  $\boldsymbol{\tau}$  的概率分布进行建模

$$p(\boldsymbol{\tau}|\boldsymbol{w}) = \prod_t \mathcal{N}(\mathbf{y}_t | \Phi_t^T \boldsymbol{w}, \Sigma_y) \quad (4)$$

其中, 轨迹  $\boldsymbol{\tau} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T\}$ , 轨迹上的一点由位置和速度组成, 即  $\mathbf{y}_t = [\mathbf{x}_t, \dot{\mathbf{x}}_t] = \Phi_t^T \boldsymbol{w} + \boldsymbol{\varepsilon}_y$ ,  $\Phi_t$  是时间依赖的基矩阵 (Basis matrix),  $\boldsymbol{w}$  是权重向量, 决定了轨迹的形状. 由于同一技能的轨迹形状可在一定范围内变化, 而不同的轨迹形状需要由不同的  $\boldsymbol{w}$  表示, 所以将  $\boldsymbol{w}$  的变化表示以  $\boldsymbol{\theta}$  为参数的概率分布  $p(\boldsymbol{w}; \boldsymbol{\theta})$ , 则轨迹的概率分布可以表示为层次式贝叶斯模型, 即

$$p(\boldsymbol{\tau}; \boldsymbol{\theta}) = \int_{\boldsymbol{w}} p(\boldsymbol{\tau}|\boldsymbol{w}) p(\boldsymbol{w}; \boldsymbol{\theta}) d\boldsymbol{w} \quad (5)$$

将基矩阵的基函数定义为高斯形式和 Von-Mises 形式, 可使 ProMP 模型分别表达单次运动和有节奏运动. 为了表达轨迹模型对任务条件例如目标状态  $\mathbf{y}^*$  的适应性, 可引入新的条件分布  $p(\boldsymbol{w}|\mathbf{y}^*)$ , 根据改变  $\boldsymbol{w}$  的分布从而调节轨迹形状. 通过轨迹概率分布相乘, 多个 ProMPs 可以被结合为一个新的轨迹. 基于 ProMP 的轨迹生成基于随机反馈控制器 (Stochastic feedback controller) 实现, 并生成每个时刻  $\mathbf{y}_t$  的轨迹均值与协方差. 文献 [40] 中用 ProMP 模型同时学习轨迹分布和相应的底层控制信号, 从而在系统动力学未知情况下应用 ProMP 模型. 期望最大化的降维方法可用于提高 ProMP 模型在强化学习的策略搜索框架下的学习效率<sup>[41]</sup>. 用单个复杂的 ProMP 模型来表示技能, 往往会降低模型的通用能力, 故 Lioutikov 等<sup>[42]</sup> 提出了概率分割 (Probabilistic segmentation, Probs) 方法, 从多次示教轨迹中学习若干 ProMPs 构成的库, 使技能的各基元都可以被独立地优化、替换或复用. 为了实现人机协作技能, 文献 [31] 提出交互概率运动基元 (Interactive ProMP), 并用于对机器人和人的运动相关性建立概率模型, 在执行协作技能时, 计算当人的运动状态已知时机器人运动轨迹的条件分布. 混合概率交互基元 (Mixture of interactive ProMPs) 可表示多模态的人机协作技能, 在执行多模态技能时, 根据人的运动状态来选择混合模型中概率最大的分量, 例如, 机械手需要避开圆柱并触碰人手中的物体, 多模态技能可以使机器人自动根据观测的人手和圆柱位置, 来选择从左边或右边绕过圆柱这两种不同的模式完成触碰.

高斯过程回归 (Gaussian process regression, GPR) 也可以用于轨迹编码, 并输出轨迹的期望与方差. 但是与 GMR 不同的是, GPR 是一种被动学习 (Lazy learning). 已知训练样本集  $\{(t_i, \mathbf{x}_i)\}_{i=1}^N$  和测试样本  $(t, \mathbf{x})$ , GPR 假设  $\mathbf{x}$  和  $\{\mathbf{x}_i\}_{i=1}^N$  遵从一

个多元高斯分布, 用协方差函数衡量样本之间的相似度, 则已知训练样本集时, 测试样本的期望和协方差可以通过计算条件概率分布获得. GPR 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. 文献 [43] 中采用 local GPR, 即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, 从而提高计算实时性. 文献 [44] 中利用 HMM 来生成轨迹的关键点, 然后用样条插值将离散的关键点转化为连续轨迹, 文中指出在没有精确机器人逆运动学模型时, 轨迹编码可在关节空间进行. 针对非刚体目标操作技能, 可将示教的场景与当前场景进行非刚体点云配准, 得到弯曲函数 (Warping function) 用于将示教轨迹变换到当前场景, 通过平均滤波来获得一条光滑的参考轨迹, 如果要考虑避障问题, 需要将轨迹优化与参考轨迹结合<sup>[45-46]</sup>.

## 2.2 动态系统

动态系统是状态按照一定规律自动演化的系统. 动态系统与轨迹编码相比, 主要有两点不同: 首先动态系统不显式地依赖于时间变量, 仅考虑空间状态与其时间导数之间的关系, 例如, 一阶动态系统可以用一阶微分方程表示为

$$\dot{\mathbf{x}} = \mathcal{D}(\mathbf{x}) \quad (6)$$

其中,  $\mathbf{x} \in \mathbf{R}^d$  是机器人系统状态,  $\mathcal{D}$  是演化函数, 可根据当前状态决定当前的状态变化率, 故动态系统的迭代演化可以表示为

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \tau \times \mathcal{D}(\mathbf{x}_t) \quad (7)$$

其中,  $\tau$  是时间系数, 决定了状态演化的快慢. 运动的终止可以用动态系统的收敛性表达, 即当  $t \rightarrow \infty$  时,  $\dot{\mathbf{x}} \rightarrow 0$  且  $\mathbf{x} \rightarrow \mathbf{x}^*$ , 其中  $\mathbf{x}^*$  为吸引子.

其次, 动态系统可以在线地生成轨迹, 对于扰动具有在线自适应能力. 扰动可以划分为空间和时间扰动, 外部和内部扰动, 以及瞬时和持续扰动. 例如, 在试图抓取物体时, 物体被突然移动, 则该扰动属于外部的瞬时空间扰动. 动态系统受到扰动后, 可以根据扰动后的状态继续按照一定规律迭代演化. 而轨迹编码模型是显式地生成一条固定的轨迹, 当机器人受到扰动并偏离该轨迹后, 无法实时地调整该轨迹, 需要重新生成新轨迹.

动态运动基元 (Dynamic motion/movement primitives, DMP) 于 2002 年由 Ijspeert 等<sup>[47-49]</sup> 提出, 是一种可产生任意形状运动轨迹的动态系统模型. DMP 的基本思想是用一个正则系统 (Canonical system) 驱动一个变换系统 (Transform system), 假设正则系统是一阶系统  $\dot{s} = -\tau\alpha_s s$ , 其变量

$s$  表示隐含的相位,  $\tau$  和  $\alpha_s$  分别是时间系数和保证系统收敛的常数,  $\mathbf{x}$  是机器人状态,  $\mathbf{x}^*$  和  $\mathbf{x}_0$  分别是目标状态和起始状态, 则变换系统表示为

$$\ddot{\mathbf{x}} = \tau\alpha_x(\beta_x(\mathbf{x}^* - \mathbf{x}) - \dot{\mathbf{x}}) + f(s) \quad (8)$$

其中,  $\alpha_x$  和  $\beta_x$  分别是二阶动态系统的阻尼和弹性系数,  $f(s)$  相当于在二阶线性系统上叠加了非线性扰动, 从而实现任意轨迹形状的非线性运动, 其表达式为

$$f(s) = \frac{\sum_{i=1}^N \Psi_i(s) \times w_i}{\sum_{i=1}^N \Psi_i(s)} \times s \times (\mathbf{x}^* - \mathbf{x}_0) \quad (9)$$

其中,  $N$  为基函数的个数,  $N$  的大小影响轨迹拟合精度.  $w_i$  为可学习的权重系数, 非线性基函数为高斯形式  $\Psi_i(s) = \exp(-(s - c_i)^2 / (2\sigma_i^2))$  且常数系数为  $c_i$  和  $\sigma_i$ . 为实现多自由度运动, 可用同一个正则系统驱动多个变换系统, 则每个变换系统的变量是同步改变的. 上述 DMP 可以表示以任意非线性轨迹从起始状态到目标状态的运动, 例如网球挥拍击球的动作. 通过在基函数中使用余弦项, 可将 DMP 用于表示周期性运动, 例如敲鼓动作. DMP 主要取决于目标状态  $\mathbf{x}^*$  和权重系数  $w_i$ , 前者由环境或设置确定, 后者可基于局部权重回归 (Linear weighted regression, LWR) 方法从示教轨迹中学习, 并通过强化学习继续优化<sup>[50-51]</sup>. LWR 方法仅能从单次示教中学习 DMP 模型参数, 文献 [52] 中基于高斯混合模型来估计 DMP 的参数, 从而实现从多次示教轨迹中学习 DMP. 针对需要对目标施加一定作用力的技能, 例如熨烫衣物, 可用两个 DMP 分别表示位置轮廓 (Profile) 和力轮廓, 然后两个 DMP 由同一个正则系统驱动<sup>[53]</sup>. DMP 具有参数数量少的优点, 可以用基于策略搜索 (Policy search) 的强化学习方法, 通过较少次数的试探进行参数优化<sup>[54-55]</sup>. 多个 DMP 的序列可表示更复杂的操作技能<sup>[56]</sup>. 在一些任务中, 执行相同的技能需要根据当前情景选择使用不同的运动模式, 例如打乒乓球, 需要根据来球速度、击球点等状态来选择击球方式, 所以文献 [57] 提出一种混合动态运动基元 (Mixture of motor primitives, MoMP), 用选通网络 (Gaiting network) 计算对当前状态下每个 DMP 的权重  $\gamma_i(\mathbf{s}) = \eta \exp[\theta_i^T \phi_i(\mathbf{s})]$ , 其中,  $\eta$  是归一化系数,  $\phi_i$  表示特征向量,  $\theta_i$  表示参数, 理想情况下选通网络应只选择一种 DMP 执行, 但是现实情况较为复杂, 为了取得良好的泛化性能, MoMP 基于选通网络输出的权重, 对所有 DMP 的输出进行加权平均得到最终输出. 文献 [58] 利用协作矩阵 (Coordination

matrix) 来表示多自由度对应的多 DMPs 间的耦合关系, 同时利用迭代降维方法来减少协作矩阵中不必要的自由度, 更少的自由度有利于高效地强化学习. Denisa 等提出了柔顺运动基元 (Compliant movement primitives, CMP) 模型, 且对于柔顺运动的学习不依赖于显式的环境动力学模型<sup>[59]</sup>. CMP 的学习分为两步: 1) 从示教轨迹中学习 DMP; 2) 记录实际的关节力矩轨迹并利用径向基函数编码为力矩基元 (Torque primitive, TP). DMP 和 TP 共同组成 CMP, 在执行运动技能时, 力矩项被用于前馈控制, 位置项被用于低增益的反馈控制.

Gribovskaya 等<sup>[60]</sup> 指出, DMP 方法虽然不依赖显式时间变量, 但是依赖于隐含的相位变量, 而且没有对多维状态变量的分量之间的关系进行描述, 并在实验中发现 DMP 受到扰动后可能会产生不理想的运动模式, 尤其是起始位置变化造成的扰动. DMP 依赖于隐含的相位变量和外部线性稳定器 (Stabilizer), 外部稳定器会扭曲运动的模式. 所以 Gribovskaya 等基于 GMM 直接对多元变量动态系统 (Multivariate dynamical system) 进行建模, 将动态系统建模视为多元回归问题, 直接建立一个从状态到状态变化率的非线性映射  $\mathcal{D}$ . GMM 用于表达状态与状态变化率的联合概率分布

$$p(\dot{\mathbf{x}}, \mathbf{x}) = \sum_{k=1}^K \pi_k \mathcal{N}(\dot{\mathbf{x}}, \mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (10)$$

在执行技能时, 利用感知模块获取当前状态, 通过 GMR 计算状态变化率的条件概率分布. 基于多元变量动态系统的运动模型形式非常简洁, 但是其如何保证一个非线性动态系统的稳定性是关键问题. Khansari-Zadeh 等<sup>[61]</sup> 提出了一种动态系统的稳定估计 (Stable estimator of dynamical system, SEDS) 方法用于学习 GMM 参数. SEDS 方法根据李雅普诺夫稳定性定理得到 GMM 参数约束并用于优化问题, 使从示教数据中学习到的动态系统具有全局渐进稳定性质 (Global asymptotical stability), 即从任意的起始状态开始运动都可以逐渐收敛到目标状态

$$\lim_{t \rightarrow \infty} \mathbf{x}_t = \mathbf{x}^*, \quad \forall \mathbf{x}_{t=0} \in \mathbf{R}^d \quad (11)$$

如图 4 所示, 在执行技能时, 动态系统的输入状态是感知模块获得的实际状态反馈, 输出是期望运动状态, 并将其发送给底层控制器作为参考信号, 然后通过执行器来改变实际状态, 底层控制器的跟踪误差可被视为对动态系统的扰动. 在一些技能中, 运动有多个可选目标状态, 例如物体上有多个抓取位置, 每一个抓取位置对应一个位置吸引子和一种趋近轨迹, 需要从中选择最优的抓取位置和其对应的动态系统. 如果通过简单的状态空间划分, 并在每个子空间采用一个动态系统, 会导致运动到子空间边界上时来回在两个子空间切换, 从而产生震荡. 所以

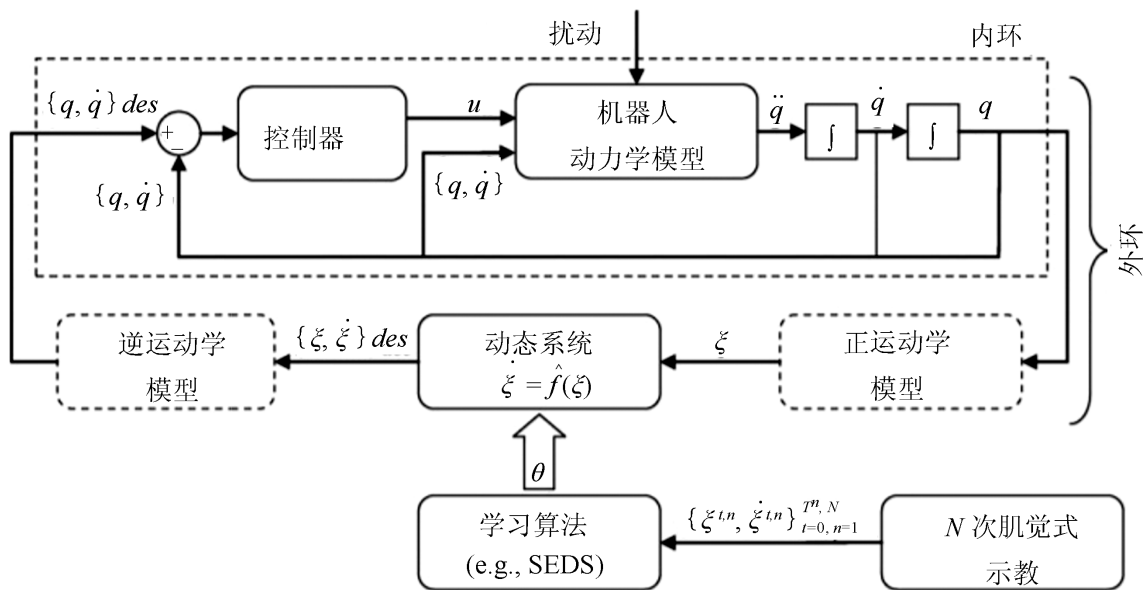


图 4 基于多元变量动态系统的运动技能执行框架, 其中,  $q$ ,  $u$  和  $\xi$  分别表示机器人的关节角度、运动指令和动态系统的状态变量 (此处为笛卡尔空间中的末端位置)<sup>[61]</sup>

Fig. 4 Multivariate dynamical system based motion skill,  $q$ ,  $u$  and  $\xi$  label the robot's joint angle, motor command and dynamical system's state variable (end-effector position in Cartesian space)<sup>[61]</sup>



强化支持向量机 (Augmented support vector machines, A-SVM) 被用于划分状态空间, 并利用分类函数的梯度用于调节动态系统, 从而使机器人运动始终被限制在一个子空间中, 而不会越过其边界<sup>[62]</sup>. Khansari-Zadeh 等<sup>[61]</sup> 指出, 基于 SEDS 的动态系统学习方法对于保证全局稳定性的严格条件, 可能会影响复杂运动建模的准确性, 即运动的稳定性和准确性间具有权衡. 文献 [63] 中利用微分同胚变换 (Diffeomorphic transformation) 将 SEDS 改进为  $\tau$ -SEDS, 以提高模型表达准确性, 但是模型学习更加耗时. Duan 等<sup>[64]</sup> 基于极限学习机 (Extreme learning machines, ELMs) 提出一种快速稳定的动态系统建模 (Fast and stable modeling for dynamical systems, FSM-DS) 方法, 利用了 ELMs 良好的学习效率以及泛化能力, 同时用参数化的李雅普诺夫函数来提高运动建模的准确性, FSM-DS 在实验中展示出同时具有良好的学习效率、全局稳定性和准确性的优点.

### 2.3 运动协调

运动协调是指运动中不同变量之间的耦合关系和联动执行. 例如, 在抓取技能中, 手指运动与手臂运动是互相协调的, 当手部趋近物体时, 手指应张开, 当手部接触物体时, 手指应逐渐握紧. Shukla 等<sup>[65]</sup> 提出耦合动态系统 (Coupled dynamical systems, CDS) 来表达手部移动与手指抓握的运动协调, 即利用两个多元变量动态系统来建模手部移动和手指抓握的运动, 变量分别是手部位置  $\mathbf{x}$  和手指角度  $\theta$ , 两个动态系统是主从关系, 即在  $t$  时刻, 主动态系统独立演化并更新手部位置  $\mathbf{x}_t$ , 而从动态系统的状态  $\theta_t$  由主动态系统的状态  $\mathbf{x}_t$  推断得到, 然后从动态系统输出手指的增量运动. 手部位置和手指角度的联合概率分布基于 GMM 表示, 并用 GMR 实现状态推断. 耦合动态系统既实现了两个不同变量的协调, 同时也保留了各变量的运动模式. Ureche 等<sup>[66]</sup> 在将耦合动态系统用于运动中位置与姿态的协调, 同时采用 GMM 对阻抗控制器的刚度因子和末端力随机机械手末端位置的变化轮廓进行表达, 在执行技能时, 通过 GMR 根据机械手末端位置, 调节阻抗控制器的刚度和期望力. 当机器人与环境或其他机器人产生交互时, 需要考虑交互对运动的影响, 例如避障或者双机协作, 可在 DMP 中引入调整项 (Modulation term), 例如根据机器人与障碍物间的距离来调节运动速度. 同样, 双手之间的相对位置也可以用于各手 DMP 的调整项, 从而保证双手操作同一个物体时, 双手间的作用力或相对位置符合期望. 所以 Gams 等<sup>[67]</sup> 提出协作动态运动基元 (Cooperative DMPs), 假设双手的期望距离和实际

距离分别是  $d_d$  和  $d_a$ , 则  $C_1 = c[k(d_d - d_a)]l_1$  和  $C_2 = c[-k(d_d - d_a)]l_2$  分别引入机械手 1 和 2 的 DMP 公式作为调整项, 且  $l_1$  和  $l_2$  决定了双手间的主从关系, 例如, 当  $l_1 = 1$  且  $l_2 = 0$  时, 机械手 2 将完全跟随机手 1 的运动. 文献 [68] 利用 GMM 编码软体手术机器人套管针距离和末端笛卡尔位置的关系, 即随着套管针插入时, 末端位置会进行与之协调的运动, 从而实现手术中灵活而不伤害重要器官的针头移动技能.

## 3 策略模型

策略  $\pi$  是状态  $s$  到行动  $a$  的一个映射, 即  $a = \pi(s)$ , 可以直观解释为机器人不断观察当前状态并做出反应行动, 从而以闭环的形式执行技能. 策略可分为离散策略和连续策略, 其输出分别是离散的行动选择和连续的行动指令. 特殊地, 端到端策略将感知模块和控制模块融合进策略模型中, 故可以直接将传感器原始信号端作为输入, 并将底层执行器的驱动信号端作为输出.

### 3.1 离散策略

离散策略的输出是离散量, 决定了行动基元  $a$  的选择, 且  $a$  属于行动基元集合  $\mathcal{A}$ . 相比于流程模型, 离散策略并没有显式地表达出技能中行动基元的种类数量与拓扑关系, 而是需要根据当前状态  $s$  来在线地生成状态-行动基元的序列  $\{\langle s_1, a_1 \rangle, \langle s_2, a_2 \rangle, \dots, \langle s_T, a_T \rangle\}$ .

在非结构化环境中, 环境状态会具有不可预料的变化. 例如机器人取桌子上的书, 需要考虑的状态包括手到书的距离、相对方向、书的厚度、书上是否有其他物体; 如果书很薄, 需要将书先推到桌边再抓取; 如果书上放着杯子, 则需要先将杯子移开. Sung 等<sup>[69]</sup> 提出一种动态规划策略 (Dynamic planning strategy), 假设  $t$  时刻环境状态为  $s_t$ , 为了能随着环境状态变化动态地从  $\mathcal{A}$  中选择最佳行动基元, 定义评分函数

$$S(s_t, a, \theta_t) = \mathbf{w}^T \phi(s_t, a, \theta_t), \quad a \in \mathcal{A} \quad (12)$$

其中,  $\theta_t$  是行动基元参数, 例如目标种类等,  $\mathbf{w}$  是权重向量,  $\phi$  表示特征提取. 当前状态下, 分数最高的行动基元被选择和执行

$$a_t = \arg \max_a S(s_t, a, \theta_t), \quad a \in \mathcal{A} \quad (13)$$

为了避免特征向量维数过高, 评分函数被分解为马尔科夫随机场 (Markov random field, MRF) 形式, 并允许特征中包括历史状态, 权重向量的学习可形式化为结构化支持向量机的优化问题. 文献 [70] 提出基于置信度的策略表达, 用 GMM 来表示当前

观测状态下各个行为基元的置信度, 并选择置信度最大的行为基元用于执行. 在执行过程中, 机器人根据置信度调节自主程度 (Adjustable autonomy). 对于当前观测, 模型返回一个推荐的行为和其置信度. 如果该置信度高于一定阈值, 则机器人选择自主执行该行为; 否则, 机器人请求并等待用户新的示教, 获取新的示教数据后更新模型. 可调节的自主程度减小了用户的在示教上的工作强度, 已经具有较高置信度的情况无需重复示教, 也可以向用户提供反馈, 使用户注意到未经示教的情况. Edmonds 等<sup>[71]</sup> 结合由顶至下 (Top-bottom) 和由底至上 (Bottom-top) 两种模型来表示行动基元选择策略, 前者基于一种随机语法模型与或图 (And-or graph, AOG) 表示, 后者用 softmax 分类器表示. 该方法利用自编码器将机器人的原始观测状态编码为低维空间的隐含状态 (Fluent), 由顶至下的输出表示已知过去的行动基元序列时下一个行动基元的概率分布, 由下至上的模型输出已知当前隐含状态和行为基元类别时下一个行动基元的概率分布, 实验中机器人基于该方法实现了打开具有安全开关的药瓶的技能. 文中指出, 自顶至下的模型和自底至上的模型具有互补性, 前者是非马尔科夫模型, 对技能的先验语义结构进行编码, 避免执行无关的行动基元; 后者是马尔科夫模型, 使机器人对实时的反馈信号做出反应, 实现精细的交互操作.

在强化学习中的 Q 学习 (Q-learning) 框架中, 函数  $Q(s_t, a_t)$  表示  $t$  时刻状态是  $s_t$  时, 选择行动  $a_t$  以后所能得到的回报期望值. 机器人控制策略即选择当前状态下使期望回报最大的行动, 即当状态和行动均为离散时,  $Q$  函数可由一个  $Q$  表格表示 (Q-table), 每一行代表一种状态, 每一列代表一种行动, 机器人反复执行试探行动并优化  $Q$  表格, 直到  $Q$  表格收敛. 当状态空间维度较高时,  $Q$  表格占用内存很大, 且因为维度灾难而难以学习, 故可采用神经网络等模型来表示  $Q$  函数, 只需要学习模型参数, 然后通过模型的泛化能力来近似在整个状态空间中的  $Q$  值分布. 文献 [72] 中采取长短时记忆 (Long short term memory, LSTM) 神经网络表示紧配合的轴孔装配策略的  $Q$  函数, 如图 5 所示, 输入状态包括三维力、二维力矩、以及轴的二维水平位置, 行动是若干个离散的控制量, 由三维力和二维旋转构成. 模型包含两个节点数分别为 20 和 15 的 LSTM 层, 用 Q 学习的框架训练, 通过 50 和 80 分钟的试探交互分别学习搜索和插入, 并达到 100% 的成功率.

### 3.2 连续策略

连续策略不依赖于预定义行动基元集合, 可以

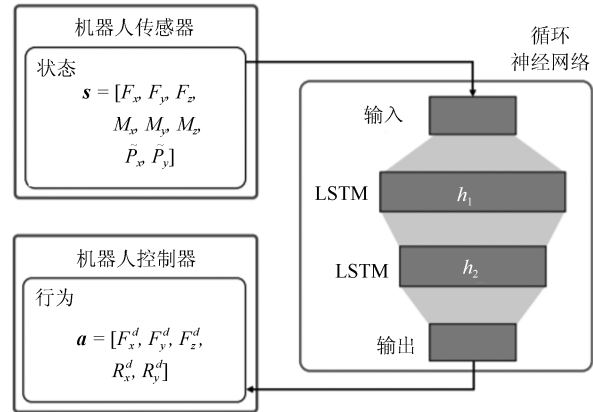


图 5 基于 LSTM 的装配策略模型<sup>[72]</sup>

Fig. 5 LSTM based assembly policy model<sup>[72]</sup>

直接根据当前状态  $s_t$  输出具体的行动指令  $a \in \mathbf{R}^d$ . 文献 [73] 采用简单的线性策略  $a_t = \mathbf{A}s_t + \mathbf{b}$  来实现低成本机械手搭积木的技能, 状态  $s_t$  是 RGBD 相机测量得到的被夹持积木的三维位置, 行动  $a_t$  是四个关节控制信号. 该机械手不具有提供位姿反馈的内部传感器, 所以文献 [74] 基于高斯过程回归从数据中学习系统模型, 然后基于强化学习方法来优化线性策略.

神经网络 (Neural networks) 是一种具有高维非线性表达能力的模型, 由输入层、输出层和若干隐含层组成. Levine 等<sup>[75]</sup> 利用具有两层隐含层的前馈神经网络表示机器人操作物体的技能, 状态包括机器人关节的角度、速度以及物体位置, 策略直接输出关节力矩信号. 神经网络的两个隐含层均具有 40 个单元, 激活函数为软整流单元, 输出层为线性连接. 由于神经网络参数维度高, 其学习需要大量样本, 故首先用试探运动的轨迹高效地学习线性高斯控制器 (Linear Gaussian controller), 线性高斯控制器对于任务条件变化的适应能力有限, 所以在不同任务条件下训练多个线性高斯控制器, 基于引导的策略搜索 (Guided policy search, GPS) 方法, 利用多个线性高斯控制器和轨迹优化生成的轨迹, 有监督地训练一个泛化能力更强的非线性神经网络策略. 为表示复合的多阶段技能, 针对每个阶段训练一个正向策略和相应的复位策略, 复位策略在该阶段失败时执行并复位该阶段的执行环境, 学习到的多个基于线性高斯控制器的正向策略和复位策略同样可以基于 GPS 方法训练为一个神经网络策略<sup>[76]</sup>. Finn 等<sup>[77]</sup> 同样基于 GPS 方法训练神经网络控制策略, 策略输入中引入了图像特征点信息, 并指出策略学习中的重要问题是, 从原始传感数据中提取并建立合适的状态空间表达, 人工设计的状态表达在复杂的非结构化环境下应用范围有限. 故基于深度空间自编码器 (Deep spatial auto-encoder) 实现

无监督的特征坐标提取自学习, 并通过特征裁剪和卡尔曼滤波实现获取更加稳定和有利于控制的图像特征. 文献 [78] 中设计了一种操作网络 (Manipulation networks) 模型, 将图像中的机器人手部目标位置、当前手部位置以及当前机器人的七维关节角度作为输入, 经过 7 层全卷积层 (Fully connected layer) 后输出机器人的 7 维关节力矩, 该操作网络模型的作用于逆运动学模型类似, 但是可以生成更加自然的手臂运动. Gu 等<sup>[79]</sup> 采用双隐层前馈神经网络, 隐含节点数为 100, 且采用了整流线性单元 (Rectified linear unit) 作为激活函数, 输出层激活函数为 tangent 型, 以保证输出行动是有界的. 实验中将非线性神经网络与更简单的线性策略对比, 实验结果表明线性策略可以成功用于捡起-放置技能, 但是无法成功实现旋转门把手并开门的技能, 故更加复杂的非线性策略模型对于表达复杂技能是必要的.

### 3.3 端到端策略

非端到端的技能模型是基于预定义的感知模块与底层控制模块, 仅对中间层的策略进行优化. 当感知模块和控制模块不是最优的时候, 会对技能的实现效果带来限制. 端到端模型将感知和控制功能均融合于技能表示, 使机器人在优化策略的同时, 也优化感知和控制模块, 使其与技能更好地匹配.

深度神经网络 (Deep neural networks, DNNs) 已经在很多有监督学习视觉任务中达到了最先进水平, 并实现了端到端的学习. 卷积神经网络 (Convolutional neural networks) 是 DNNs 中常用的一种模型, 尤其适合用于视觉信息的特征提取. Levine 等<sup>[80]</sup> 提出针对机器人操作任务的端到端深度视觉运动策略 (Visuomotor policy), 直接将原始的观测状态, 包括关节角、关节速度、末端位姿、末端速度、和 RGB 图像作为策略输入, 并输出 7 维关节力矩. 如图 6 所示, 该模型基于 CNN 实现, 其结构共有 7 层, 原始图像经过 3 个卷积层得到响应图像, 并基于空间 softmax 函数获取图像特征的二维坐标. 这些图像特征的二维坐标与机器人本体状态串联并输入两层全连接层, 最终通过线性连接输出关节力矩.

该方法不仅实现了较为复杂的操作技能, 而且不需要相机标定, 且不需要人工设计图像特征提取与控制的算法模块, 但是需要通过机器人提供的三维位置反馈来学习特定技能的图像特征. 文献 [8] 中所设计的端到端深度神经网络模型, 采用了 6 个卷积层进行图像特征提取, 最大值池化 (max-pooling) 被用于增加感受野和降低分辨率, 6 个卷积层中的后 5 层的输出特征图被上采样到相同尺寸并相加, 在经过卷积和空间 softmax 获得 32 个特征点坐标向量. 当前帧特征点坐标向量、起始帧特征点坐标向量以及 33 维的机器人状态经过两层全卷积层, 最后输出 7 维的关节力矩. 该端到端模型的卷积层模块首先在位姿预测任务上进行预训练, 然后再基于异步分布式引导策略搜索 (Asynchronous distributed guided policy search, ADGPS) 方法对模型进行端到端训练. 实验中基于该模型实现了 4 个机器人通过集体策略学习 (Collective policy learning) 获取了扭转门把手并开门的技能.

循环神经网络 (Recurrent neural networks, RNN) 是一种具有记忆的神经网络, 即其输入不仅是当前状态也包含了隐含层的历史状态, 故 RNN 适合于处理时间序列数据和非马尔科夫策略. 文献 [81] 中利用端到端的 RNN 模型表示绘画技能, 该方法不需要显式的图像特征提取或形状基元定义, 而是直接将绘画过程表示为  $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{h}_{t-1})$ , 其中状态  $\mathbf{x}_t$  包含当前的已绘图像和绘画动作, 绘画动作由二维位置和笔尖接触状态构成,  $\mathbf{h}_{t-1}$  是上一时刻 RNN 的隐含状态, 可视为记忆状态, 则策略输出是下一时刻预测的绘制图像和绘画动作. 模型中利用卷积层对图像进行编码得到降维后的低维图像表达, 利用单层全连接层对动作进行编码, 然后将图像和动作的编码一起输入 LSTM 模块进行序列状态预测, LSTM 的输出进一步通过转置卷积和全连接层分别解码为图像和动作. Kase 等<sup>[82]</sup> 提出一种基于卷积自编码器 (Convolutional autoencoder, CAE) 和多时间尺度循环神经网络 (Multiple timescale RNN, MTRNN) 的端到端模型, 其中 CAE 用于提取图像特征, MTRNN 用于生成运动. 该方法中的 MTRNN 采用了两个具有不同时间常数的上下文

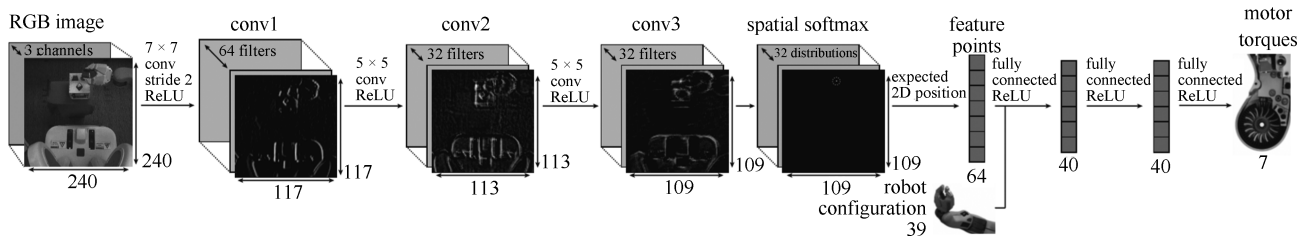


图 6 基于深度神经网络的端到端策略模型<sup>[80]</sup>

Fig. 6 DNN based end-to-end policy model<sup>[80]</sup>

层 (Context layer), 即快速上下文层和慢速上下文层, 分别从当前上下文和过去上下文中提取更多的信息. 通过在损失函数中加入对起始时刻上下文和终止时刻上下文相似的约束, 实现了模型自动在多个阶段间切换, 例如, 当机器人打开盒子后, 实验人员帮其完成了把物品放入盒子的拾放阶段, 然后机器人会自动跳过拾放阶段, 直接关闭盒子.

受到神经科学启发, 熟练的技能包括预测式控制、反应式控制和生物力学控制. 反应式控制是针对每一个传感信号产生一个运动信号, 会受到传感延迟的影响. 预测式控制给定一个观察, 而系统连续地输出一个不间断的运动信号, 快速实现基本的动作<sup>[83]</sup>. 受此启发, Ghadirzadeh 等<sup>[84]</sup> 提出一种深度预测策略 (Deep predictive policy), 可以将 1 幅无标定的原始图像作为输入, 并预测底层运动控制器的一段连续运动轨迹. 策略模型由感知、策略和行为 3 个模块级联构成, 感知模块基于空间自编码器将输入图像编码为图像特征点, 图像特征点组成状态向量  $\mathbf{s}_t$ ; 策略模块基于变分自编码器由状态向量生成一个 5 维的行动正态分布, 并从中采样出  $\mathbf{a}_t$ ; 最后行为模块将  $\mathbf{a}_t$  映射为底层控制信号轨迹  $\mathbf{u}_{t:t+T}$ . 为了提高训练的样本效率, 感知和行为模块的参数从合成和仿真数据中学习, 策略模块的少量参数基于实体机器人数据学习.

#### 4 效果预测模型

技能模型中的核心是技能的执行和参数化, 但是为了使技能具有可预测性, 使技能的执行可以受到监督, 需要应用效果预测模型. 效果预测模型对技能的前提条件 (Pre-condition)、后置条件 (Post-condition) 以及执行过程进行评估<sup>[85]</sup>, 所以效果预测模型的输出可以是成功率、异常判断、效果预判等. 人类在执行技能时, 会对失败情况做出预判, 例如人在行走时, 会通过自身姿态和地面反作用力来预判异常地形是否会导致摔倒. 文献 [55] 中提出失败条件检测, 可以在技能失败发生之前提前停止执行, 应用多种传感器信号包括关节位置、速度、加速度, 夹持器加速度, 和手指的压力传感器, 预测方法主要基于 z-test, 判断当前状态与成功经验数据的分布是否一致. 如果不一致, 则认为当前状态下可能导致失败. 触觉反馈可用于预测抓握稳定性<sup>[86]</sup>, 文献 [87] 中使用词袋 (Bag-of-words) 模型来表示接触的三维空间分布, 从而降低特征空间的维数, 并用 SVM 分类器学习并预测抓握的稳定与不稳定. Levine 等<sup>[88]</sup> 提出了一种大数据驱动的手眼协调模型, 将 CNN 用于表达抓取成功率的预测函数  $g(I_t, \mathbf{v}_t) \in [0, 1]$ , 其中,  $I_t$  和  $\mathbf{v}_t$  分别是  $t$  时刻原始输入图像和候选运动指令, 所以该函数表示了给定图像

观测时, 执行某个运动指令导致抓取成果的概率. 该模型由 16 层卷积层和 2 层全连接层构成, 输出单元为 sigmoid 型. 在执行抓取技能时, 机器人需要基于观测图像估计最优的运动指令, 所以采用交叉熵方法 (Cross-entropy method), 通过随机采样和迭代优化获得最优运动指令  $\mathbf{v}_t^*$ . 机械手的升降以及夹持器的打开闭合取决于  $p = g(I_t, \mathbf{0})/g(I_t, \mathbf{v}_t^*)$ , 即机械手停止和进一步运动的成功概率之比, 如果大于高阈值则夹持, 如果小于低阈值则升高夹持器, 其他情况则执行  $\mathbf{v}_t^*$ . 该方法不需要相机标定和手眼标定, 也不需要目标上增加人工标记, 直接从单目图像中提取空间信息与物体种类信息. 在实验中, 14 个机械臂在两个月中收集了超过 800 000 次抓取尝试的大量数据, 并利用这些数据训练模型, 结果执行抓取的成功率达到 82.5%. 这些机械臂上安装的相机相对于基座的姿态并不完全一致, 抓取的目标不同, 而且夹持器手指在长时间的运行过程中产生了不同的磨损. 从大量的抓取数据中, 可以发现一些抓取模式, 例如针对柔软目标和刚体目标的抓取方式不同. 文献 [78] 中提出一种卷积未来回归 (Convolutional future regression) 模型, 基于当前帧图像  $I_t$  经过卷积编码器获得的当前隐含场景编码  $F_t$ , 预测未来的隐含场景编码  $F_{t+1}$ , 然后未来的隐含场景编码可经过解码器生成重建特征图并输入目标检测网络, 得到预测到的未来目标位置  $Y_{t+1}$ . 深度视频预测技术可用于无监督的物理交互效果学习, 该方法提出基于卷积 LSTM 模型的卷积动态神经流 (Convolutional dynamic neural advection, CDNA) 模型, 可根据当前帧、机器人状态与机器人行动, 预测下一帧的像素运动分布<sup>[89]</sup>. 深度视频预测模型可与模型预测控制 (Model predictive control, MPC) 进一步结合实现技能执行, 实验中机器人基于该模型学习到将物体移动到指定位置的技能<sup>[90]</sup>.

### 5 典型应用与未来趋势

#### 5.1 典型应用

在现实的工业与科研应用中, 机器人操作技能模型的应用非常广泛, 典型的例子如图 7 所示. 凡是需要完成操作任务的场合, 都需要相应的技能, 而通过设计具有紧凑性、综合性、稳定性、安全性、可学习性和复杂性的机器人操作技能模型, 是使机器人获取、学习和优化这些技能的关键.

装配机器人所需要的技能主要是通过末端执行器操作零件, 从而使零件之间具有期望的位姿关系. 例如, 基于柔顺运动基元实现轴孔装配<sup>[91]</sup>、基于合成技能实现家具组装<sup>[92]</sup>、微器件精密装配<sup>[93]</sup> 以及焊缝焊接<sup>[94]</sup>. 基于行为树的交互式技能编辑系统 Co-

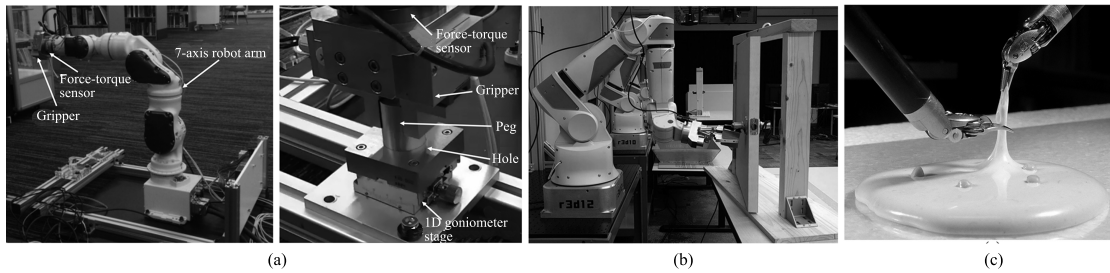


图7 机器人操作模型的典型应用 ((a) 轴孔装配技能<sup>[72]</sup>; (b) 开门技能<sup>[8]</sup>; (c) 手术切除技能<sup>[95]</sup>)  
 Fig. 7 Typical application of robot manipulation skill model ((a) peg-in-hole assembly<sup>[72]</sup>;  
 (b) door opening<sup>[8]</sup>; (c) resection surgery<sup>[95]</sup>)

STAR 系统可以跨平台地应用于多种机器人本体, 例如 UR5 机器人和 KUKA LBR iiwa 机器人, 在弯折钢丝、打磨、抛光、装配等多种任务中展现出可行性与易用性<sup>[13]</sup>. 协作机器人近年来也得到越来越多的应用, 机器人需要和人协作来提高任务工作效率, 所以需要更加安全和灵活的协作技能模型, 例如机械臂与工人协作搬运物体<sup>[37]</sup>, KUKA 机器人观察人的运动并以不同模式传递工具<sup>[96]</sup>.

服务机器人通常工作于人类环境中, 人类环境中物体种类繁多、排列不规则, 而且往往会发生动态变化, 使机器人可以灵活操作日常生活中的物体具有很大意义. 例如, 双机械手摇晃杯子<sup>[97]</sup>、机械臂擀面团<sup>[98]</sup>、熨烫衣服和递盘子<sup>[99]</sup>. 基于肌觉式示教学习, 6 自由度机器人 Katana 6M 获取抓取杯子、放在咖啡机、并按下咖啡机按钮的技能<sup>[43]</sup>, KUKA 机械臂可学习到蔬菜刨丝和抠取电池的技能<sup>[66]</sup>, 以及搅拌液体和挖取果肉的技能<sup>[100]</sup>. ASIMO 人形机器人可通过模仿学习获取倒水技能<sup>[101]</sup>. 基于强化学习, PR2 机器人可获取如下技能, 玩具积木拼接、旋紧瓶盖、挂晾衣架、把米袋放入碗中中等生活技能, 但是执行场景与训练场景差别较大时, 学习到的技能会无法适用<sup>[77, 80]</sup>. 文献 [102] 中的实验实现了软体机械手拨算盘、转动阀门和抓取瓶子的技能. 在运动方面, SARCOS 仿人机械臂可通过强化学习获取击打棒球技能<sup>[103]</sup>, PR2 机器人可使用球杆击打台球<sup>[55]</sup>, Barrett WAM 机械臂可实现打乒乓球技能<sup>[57]</sup>.

手术机器人大多通过遥操作, 即在人类医生的控制下, 完成精细灵巧而流程复杂的手术. 手术机器人自主完成完整的技能具有很大的难度, 但是使机器人学习局部的手术技能并在监督下自主执行该局部技能, 具有很好的前景. 例如, 管状器械的柔顺插入与黏膜剥离<sup>[104]</sup>、达芬奇手术机器人的清创与切割技能<sup>[105]</sup>、以及脑部肿瘤切除技能<sup>[95]</sup>.

## 5.2 未来趋势

机器人操作技能模型一直是机器人学与人工智

能的研究者关注的话题, 以瑞士联邦理工学院学习算法与系统实验室 (Learning Algorithms and Systems Laboratory) 为代表的研究团队, 重视机器人动力学与创新学习算法的结合, 使机械臂可以完成动态性很强的复杂技能任务, 例如实现非对称的双机械手协作<sup>[106]</sup> 和高速抓取抛在空中的瓶子<sup>[107]</sup>. 而以加州大学伯克利分校机器人人工智能与学习实验室 (Robotic Artificial Intelligence and Learning Laboratory) 为代表的研究团队, 重视用复杂的深度神经网络来实现通用机器人操作, 希望通过神经网络的强大表达能力来避免技能学习对相机标定、机器人动力学模型、人工定位标记的依赖. 该团队近年来致力于基于深度强化学习 (Deep reinforcement learning, DRL) 来学习操作技能, 展示出了很强的通用泛化能力, 但是目前实现的操作技能相对简单, 操作对象是摆放在平面上的小件物品<sup>[77, 108]</sup> 和对操作精度要求不高的简易物品<sup>[75, 80]</sup>, 因为目前基于深度神经网络的技能模型的输入主要包括单目图像和机械臂运动学信息, 基于单目图像对于立体目标的感知精度有限, 而且技能模型并未考虑交互力和长期记忆. 未来工作中, 机器人操作技能建模可参考如下方向:

1) 高效学习. 日常生活中的很多技能, 人类只需要观察一次演示过程, 甚至根据一段技能流程描述的语句, 即可获取该技能. 而机器人需要从若干次演示中才能获取一个较为简单的技能, 例如学习刨削蔬菜需要 18 次示教<sup>[66]</sup>. 在文献 [88] 的实验中, 研究者利用 14 个机械臂耗时两个月收集了超过 800 000 次物体抓取技能的大量数据, 并利用这些数据训练模型, 使结果执行抓取的成功率达到 82.5%. 在文献 [77] 中机器人需要进行十几分钟的试探交互操作才能获取推动积木等简单技能.

2) 模型解释. 机器人技能模型如果具有可解释性, 则可以让人类更好地理解机器人的行动, 确保机器人的行动是安全而合理的, 当机器人运行出现问题时, 人类可以基于技能模型寻找内部原因. 相反, 如果技能模型是“黑箱式”的, 例如, 文献 [80] 中的

深度神经网络策略模型有 92 000 个参数, 没有有显式的物理意义, 所以人类无法理解机器人学到了什么, 也无法预知机器人技能是否是安全的, 只能通过实验来验证, 这种验证工作在实际应用中往往是不可行的. 以行为树为代表的流程模型具有一定的可解释性, 可以被人类直观理解, 运动模型也可以通过轨迹生成来直观展示出技能效果. 以神经网络为代表的策略模型, 虽然具有很强的表达能力和通用性, 但是无法被用户直观理解, 也难以通过数学分析来评估该模型.

3) 统一框架. 人工智能中的视觉、语音等任务, 往往具有统一的任务框架, 不依赖于具体的硬件平台, 故可以实现“众人拾柴火焰高”, 典型例子是视觉任务中具有若干的大型公共数据集如 ImageNet<sup>[109]</sup>, 使世界各地研究者具有了统一的数据资源和评价指标, 解决同一个核心问题并互相对比竞争, 显著促进了视觉技术的进步. 机器人技能学习任务, 面临着机器人平台和任务指标的差异性, 研究者不仅要考虑技能模型设计, 还要考虑场景设计、数据收集、感知与控制模块开发的问题, 从而对相关研究的规模和速度带来一定限制.

4) 云机器人. 互联网技术在最近十年内得到了爆发式增长, 对机器人领域也产生了积极影响. 云技术为机器人提供了互联网的数据和计算资源, 可以支持和提高机器人的技能执行, 而不是仅仅依赖于机器人自身有限的计算能力和存储容量<sup>[110]</sup>. 云技术可以提供大数据、云计算、集体机器人学习 (Collective robot learning) 和人力计算 (Human computation), 从而实现机器人共享大量的图像、轨迹、地图、控制策略数据资源, 并在执行技能时将一部分复杂计算在云端执行, 而且人工参与可实现众包 (Crowd-sourcing) 数据分析与标注<sup>[111]</sup>. 近几年推出的云机器人的项目有云机器人 (CloudRobot)<sup>[112]</sup>、机器人大脑 (RoboBrain)<sup>[113]</sup> 和机器人地球 (RoboEarth)<sup>[114]</sup>. 平行机器人的概念于 2015 年首次提出<sup>[115]</sup>, 文献 [116] 中指出“平行机器人是物理机器人、软件机器人、仿真系统、物联网、数据库、广义的人工智能技术等相结合而成的机器人控制与管理系统”. 其中物理机器人在物理、网络空间中执行具体的任务技能, 而软体机器人在网络、交互社会空间中执行搜索、推理、决策、优化等知识自动化的任务<sup>[117]</sup>.

5) 类脑智能. 类脑智能是以计算建模为手段, 受脑神经和人类认知行为机制启发, 并通过软硬件协同实现的机器智能<sup>[4]</sup>. 人脑能够使人类在新的环境和新的任务中自动获取技能, 并在复杂的动态的环境中执行技能并展现出长时间稳定和低功耗的优点, 人类甚至能在自身受到损伤时仍然保持技能的鲁棒

性. 近年来通过多学科交叉实验研究, 在脑区、神经簇、神经元等不同尺度下研究人脑的工作机制, 可以对机器人技能建模具有更多的启发. 机器人所展现的技能水平仍然与人类相差甚远, 在流程层面, 人类可以掌握外科手术、烹饪食物等具有很大不确定性和复杂流程的技能, 在物理运动层面, 人类既可以精细地操作缝衣针, 也可以用足够的准确度和适当的力量击打乒乓球. 如何使机器人操作技能更接近人类水平, 将是一个具有极大挑战的长期目标.

## 6 结论

机器人操作技能模型在生产和生活中具有巨大的应用价值, 在科研上仍然是开放的热点问题. 本文首先介绍了机器人操作技能模型的定义、意义与基本性质, 并对基于流程、运动、策略和效果预测四种模式的技能模型分别进行阐述. 最后, 本文展示了机器人操作技能模型在装配、协作、服务和手术机器人领域的典型应用, 并指出了一些未来可能的发展趋势. 机器人操作技能建模与学习是人工智能与机器人学的交叉领域, 从不同学科视角对该问题进行审视、研究和融合, 对于科研人员既是挑战又是机遇. 使机器人操作技能不断趋近甚至超越人类水平, 对于科研人员是长期而富有挑战性的目标.

## References

- Hirzinger G, Landzettel K. Sensory feedback structures for robots with supervised learning. In: Proceedings of the 1985 IEEE International Conference on Robotics and Automation. St. Louis, MO, USA: IEEE, 1985. 627–635
- Asada H, Asari Y. The direct teaching of tool manipulation skills via the impedance identification of human motions. In: Proceedings of the 1988 IEEE International Conference on Robotics and Automation. Philadelphia, PA, USA: IEEE, 1988. 1269–1274
- Zeng Yi, Liu Cheng-Lin, Tan Tie-Niu. Retrospect and outlook of brain-inspired intelligence research. *Chinese Journal of Computers*, 2016, **39**(1): 212–223 (曾毅, 刘成林, 谭铁牛. 类脑智能研究的回顾与展望. *计算机学报*, 2016, **39**(1): 212–223)
- Tao Jian-Hua, Chen Yun-Ji. Current status and consideration on brain-like computing chip and brain-like intelligent robot. *Bulletin of Chinese Academy of Sciences*, 2016, **31**(7): 803–811 (陶建华, 陈云霁. 类脑计算芯片与类脑智能机器人发展现状与思考. *中国科学院院刊*, 2016, **31**(7): 803–811)
- Ersen M, Oztop E, Sariel S. Cognition-enabled robot manipulation in human environments: requirements, recent work, and open problems. *IEEE Robotics and Automation Magazine*, 2017, **24**(3): 108–122
- Argall B D, Chernova S, Veloso M, Browning B. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 2009, **57**(5): 469–483
- Kober J, Bagnell J A, Peters J. Reinforcement learning in robotics: a survey. *The International Journal of Robotics Research*, 2013, **32**(11): 1238–1274

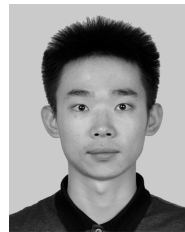
- 8 Yahya A, Li A, Kalakrishnan M, Chebotar Y, Levine S. Collective robot reinforcement learning with distributed asynchronous guided policy search. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, BC, Canada: IEEE, 2017. 79–86
- 9 Foukarakis M, Leonidis A, Antona M, Stephanidis C. Combining finite state machine and decision-making tools for adaptable robot behavior. In: Proceedings of the 8th International Conference on Universal Access in Human-Computer Interaction. Heraklion, Crete, Greece: Springer, 2014. 625–635
- 10 Zhou H T, Min H S, Lin Y H, Zhang S N. A robot architecture of hierarchical finite state machine for autonomous mobile manipulator. In: Proceedings of the 10th International Conference on Intelligent Robotics and Applications. Wuhan, China: Springer, 2017. 425–436
- 11 Colledanchise M, Parasuraman R, Ögren P. Learning of behavior trees for autonomous agents. *IEEE Transactions on Games*, 2019, **11**(2): 183–189
- 12 Guerin K R, Lea C, Paxton C, Hager G D. A framework for end-user instruction of a robot assistant for manufacturing. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 6167–6174
- 13 Paxton C, Hundt A, Jonathan F, Guerin K, Hager G D. CoSTAR: instructing collaborative robots with behavior trees and vision. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation. Singapore, Singapore: IEEE, 2017. 564–571
- 14 Paxton C, Jonathan F, Hundt A, Mutlu B, Hager G D. Evaluating methods for end-user creation of robot task plans. In: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. Madrid, Spain: IEEE, 2018. 6086–6092
- 15 Bagnell J A, Cavalcanti F, Cui L, Galluzzo T, Hebert M, Kazemi M, et al. An integrated system for autonomous robotics manipulation. In: Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura, Portugal: IEEE, 2012. 2955–2962
- 16 Colledanchise M, Marzinotto A, Ögren P. Performance analysis of stochastic behavior trees. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation. Hong Kong, China: IEEE, 2014: 3265–3272
- 17 Akgun B, Thomaz A. Simultaneously learning actions and goals from demonstration. *Autonomous Robots*, 2016, **40**(2): 211–227
- 18 Akgun B, Thomaz A L. Self-improvement of learned action models with learned goal models. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015. 5259–5264
- 19 Kroemer O, Daniel C, Neumann G, van Hoof H, Peters J. Towards learning hierarchical skills for multi-phase manipulation tasks. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 1503–1510
- 20 Medina J R, Billard A. Learning stable task sequences from demonstration with linear parameter varying systems and hidden Markov models. In: Proceedings of the 2017 Conference on Robot Learning. Mountain View, California, USA, 2017: 175–184
- 21 Pardowitz M, Knoop S, Dillmann R, Zollner R D. Incremental learning of tasks from user demonstrations, past experiences, and vocal comments. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2007, **37**(2): 322–332
- 22 Nicolescu M N, Mataric M J. Natural methods for robot task learning: instructive demonstrations, generalization and practice. In: Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems. Melbourne, Australia: ACM, 2003. 241–248
- 23 Hayes B, Scassellati B. Autonomously constructing hierarchical task networks for planning and human-robot collaboration. In: Proceedings of the 2016 IEEE International Conference on Robotics and Automation. Stockholm, Sweden: IEEE, 2016. 5469–5476
- 24 Ahmadzadeh S R, Kormushev P, Caldwell D G. Interactive robot learning of visuospatial skills. In: Proceedings of the 2013 International Conference on Advanced Robotics. Montevideo, Uruguay: IEEE, 2013: 1–8
- 25 Ahmadzadeh S R, Paikan A, Mastrogiovanni F, Natale L, Kormushev P, Caldwell D G, et al. Learning symbolic representations of actions from human demonstrations. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 3801–3808
- 26 Dornhege C, Hertle A. Integrated symbolic planning in the tidyup-robot project. In: Proceedings of the 2013 Designing Intelligent Robots: Reintegrating AI: Papers Form the AAAI Spring Symposium. Palo Alto, California, USA: AAAI, 2013.
- 27 Beetz M, Mösenlechner L, Tenorth M. CRAM — a cognitive robot abstract machine for everyday manipulation in human environments. In: Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, China: IEEE, 2010. 1012–1017
- 28 Tenorth M, Beetz M. KnowRob: a knowledge processing infrastructure for cognition-enabled robots. *The International Journal of Robotics Research*, 2013, **32**(5): 566–590
- 29 Bozcuoğlu A K, Kazhoyan G, Furuta Y, Stelter S, Michael B, Kei O, et al. The exchange of knowledge using cloud robotics. *IEEE Robotics and Automation Letters*, 2018, **3**(2): 1072–1079
- 30 Calinon S, Guenter F, Billard A. On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2007, **37**(2): 286–298
- 31 Maeda G J, Neumann G, Ewerton M, Lioutikov R, Kroemer O, Peters J. Probabilistic movement primitives for coordination of multiple human-robot collaborative tasks. *Autonomous Robots*, 2017, **41**(3): 593–612
- 32 Calinon S, Li Z B, Alizadeh T, Tsagarakis N G, Caldwell D G. Statistical dynamical systems for skills acquisition in humanoids. In: Proceedings of the 12th IEEE-RAS International Conference on Humanoid Robots. Osaka, Japan: IEEE, 2012. 323–329
- 33 Huang Y L, Silvério J, Roza L, Caldwell D G. Generalized task-parameterized skill learning. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation. Brisbane, QLD, Australia: IEEE, 2018. 5667–5674
- 34 Tanwani A K, Calinon S. Learning robot manipulation tasks with task-parameterized semitied hidden semi-Markov model. *IEEE Robotics and Automation Letters*, 2016, **1**(1): 235–242

- 35 Silvério J, Rozo L, Calinon S, Caldwell D G. Learning bimanual end-effector poses from demonstrations using task-parameterized dynamical systems. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015. 464–470
- 36 Calinon S, Bruno D, Caldwell D G. A task-parameterized probabilistic model with minimal intervention control. In: Proceedings of the 2014 IEEE International Conference on Robotics and Automation. Hong Kong, China: IEEE, 2014. 3339–3344
- 37 Rozo L, Bruno D, Calinon S, Caldwell D G. Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015. 1024–1030
- 38 Paraschos A, Daniel C, Peters J, Neumann G. Probabilistic movement primitives. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada: ACM, 2013. 2616–2624
- 39 Paraschos A, Daniel C, Peters J, Neumann G. Using probabilistic movement primitives in robotics. *Autonomous Robots*, 2018, **42**(3): 529–551
- 40 Paraschos A, Rueckert E, Peters J, Neumann G. Probabilistic movement primitives under unknown system dynamics. *Advanced Robotics*, 2018, **32**(6): 297–310
- 41 Colomé A, Neumann G, Peters J, Torras C. Dimensionality reduction for probabilistic movement primitives. In: Proceedings of the 2014 IEEE-RAS International Conference on Humanoid Robots. Madrid, Spain: IEEE, 2014. 794–800
- 42 Lioutikov R, Neumann G, Maeda G, Peters J. Learning movement primitive libraries through probabilistic segmentation. *The International Journal of Robotics Research*, 2017, **36**(8): 879–894
- 43 Schneider M, Ertel W. Robot learning by demonstration with local Gaussian process regression. In: Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, China: IEEE, 2010: 255–260
- 44 Garrido J, Yu W, Soria A. Human behavior learning for robot in joint space. *Neurocomputing*, 2015, **155**: 22–31
- 45 Schulman J, Ho J, Lee C, Abbeel P. Learning from demonstrations through the use of non-rigid registration. *Robotics Research*. Cham: Springer International Publishing, 2016. 339–354
- 46 Lee A X, Lu H, Gupta A, Levine S, Abbeel P. Learning force-based manipulation of deformable objects from multiple demonstrations. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 177–184
- 47 Ijspeert A J, Nakanishi J, Schaal S. Learning attractor landscapes for learning motor primitives. In: Proceedings of the 15th International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2002. 1547–1554
- 48 Ijspeert A J, Nakanishi J, Schaal S. Movement imitation with nonlinear dynamical systems in humanoid robots. In: Proceedings of the 2002 IEEE International Conference on Robotics and Automation. Washington, DC, USA: IEEE, 2002. 1398–1403
- 49 Ijspeert A J, Nakanishi J, Hoffmann H, Pastor P, Schaal S. Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Computation*, 2013, **25**(2): 328–373
- 50 Kober J, Peters J. Policy search for motor primitives in robotics. *Machine Learning*, 2011, **84**(1–2): 171–203
- 51 Kober J, Peters J. Learning motor primitives for robotics. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan: IEEE, 2009. 2112–2118
- 52 Yang C G, Chen C Z, He W, Cui R X, Li Z J. Robot learning system based on adaptive neural control and dynamic movement primitives. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, **30**(3): 777–787
- 53 Kormushev P, Calinon S, Caldwell D G. Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input. *Advanced Robotics*, 2011, **25**(5): 581–603
- 54 Kupcsik A, Deisenroth M P, Peters J, Loh A P, Vadakkepatt P. Model-based contextual policy search for data-efficient generalization of robot skills. *Artificial Intelligence*, 2017, **247**: 415–439
- 55 Pastor P, Kalakrishnan M, Chitta S, Theodorou E, Schaal S. Skill learning and task outcome prediction for manipulation. In: Proceedings of the 2011 IEEE International Conference on Robotics and Automation. Shanghai, China: IEEE, 2011. 3828–3834
- 56 Stulp F, Theodorou E A, Schaal S. Reinforcement learning with sequences of motion primitives for robust manipulation. *IEEE Transactions on Robotics*, 2012, **28**(6): 1360–1370
- 57 Mülling K, Kober J, Kroemer O, Peters J. Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, 2013, **32**(3): 263–279
- 58 Colomé A, Torras C. Dimensionality reduction for dynamic movement primitives and application to bimanual manipulation of clothes. *IEEE Transactions on Robotics*, 2018, **34**(3): 602–615
- 59 Deniša M, Gams A, Ude A, Petrič T. Learning compliant movement primitives through demonstration and statistical generalization. *IEEE/ASME Transactions on Mechatronics*, 2016, **21**(5): 2581–2594
- 60 Gribovskaya E, Khansari-Zadeh S M, Billard A. Learning non-linear multivariate dynamics of motion in robotic manipulators. *The International Journal of Robotics Research*, 2011, **30**(1): 80–117
- 61 Khansari-Zadeh S M, Billard A. Learning stable nonlinear dynamical systems with Gaussian mixture models. *IEEE Transactions on Robotics*, 2011, **27**(5): 943–957
- 62 Shukla A, Billard A. Augmented-SVM for gradient observations with application to learning multiple-attractor dynamics. *Support Vector Machines Applications*. Cham: Springer International Publishing, 2014. 1–21
- 63 Neumann K, Steil J J. Learning robot motions with stable dynamical systems under diffeomorphic transformations. *Robotics and Autonomous Systems*, 2015, **70**: 1–15
- 64 Duan J H, Ou Y S, Hu J B, Wang Z Y, Jin S K, Xu C. Fast and stable learning of dynamical systems based on extreme learning machine. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, **49**(6): 1175–1185
- 65 Shukla A, Billard A. Coupled dynamical system based arm-hand grasping model for learning fast adaptation strategies. *Robotics and Autonomous Systems*, 2012, **60**(3): 424–440
- 66 Ureche A L P, Umezawa K, Nakamura Y, Billard A. Task parameterization using continuous constraints extracted from human demonstrations. *IEEE Transactions on Robotics*, 2015, **31**(6): 1458–1471



- 67 Gams A, Nemec B, Ijspeert A J, Ude A. Coupling movement primitives: interaction with the environment and bimanual tasks. *IEEE Transactions on Robotics*, 2014, **30**(4): 816–830
- 68 Bruno D, Calinon S, Caldwell D G. Learning autonomous behaviours for the body of a flexible surgical robot. *Autonomous Robots*, 2017, **41**(2): 333–347
- 69 Sung J, Selman B, Saxena A. Learning sequences of controllers for complex manipulation tasks. In: Proceedings of the 30th International Conference on Machine Learning. Atlanta, Georgia, USA: JMLR, 2013.
- 70 Chernova S, Veloso M. Confidence-based policy learning from demonstration using Gaussian mixture models. In: Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems. Honolulu, Hawaii: ACM, 2007. Article No. 233
- 71 Edmonds M, Gao F, Xie X, Liu H X, Qi S Y, Zhu Y X, et al. Feeling the force: integrating force and pose for fluent discovery through imitation learning to open medicine bottles. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, BC, Canada: IEEE, 2017. 3530–3537
- 72 Inoue T, De Magistris G, Munawar A, Yokoya T, Tachibana R. Deep reinforcement learning for high precision assembly tasks. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, BC, Canada: IEEE, 2017. 819–825
- 73 Deisenroth M P, Rasmussen C E, Fox D. Learning to control a low-cost manipulator using data-efficient reinforcement learning. In: Proceedings of the 2011 Robotics: Science and Systems VII. Los Angeles, CA, USA: University of Southern California, 2011. 57–64
- 74 Deisenroth M P, Fox D, Rasmussen C E. Gaussian processes for data-efficient learning in robotics and control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(2): 408–423
- 75 Levine S, Wagener N, Abbeel P. Learning contact-rich manipulation skills with guided policy search. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 156–163
- 76 Han W Q, Levine S, Abbeel P. Learning compound multi-step controllers under unknown dynamics. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: IEEE, 2015. 6435–6442
- 77 Finn C, Tan X Y, Duan Y, Darrell T, Levine S, Abbeel P. Learning visual feature spaces for robotic manipulation with deep spatial autoencoders. arXiv: 1509.06113v1, 2015.
- 78 Lee J, Ryoo M S. Learning robot activities from first-person human videos using convolutional future regression. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, HI, USA: IEEE, 2017. 472–473
- 79 Gu S X, Holly E, Lillicrap T, Levine S. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation. Singapore, Singapore: IEEE, 2017. 3389–3396
- 80 Levine S, Finn C, Darrell T, Abbeel P. End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, 2016, **17**(1): 1334–1373
- 81 Sasaki K, Ogata T. End-to-end visuomotor learning of drawing sequences using recurrent neural networks. In: Proceedings of the 2018 International Joint Conference on Neural Networks. Rio de Janeiro, Brazil: IEEE, 2018. 1–2
- 82 Kase K, Suzuki K, Yang P C, Mori H, Ogata T. Put-in-box task generated from multiple discrete tasks by a humanoid robot using deep learning. In: Proceedings of the 2018 IEEE International Conference on Robotics and Automation. Brisbane, QLD, Australia: IEEE, 2018. 6447–6452
- 83 Wolpert D M, Diedrichsen J, Flanagan J R. Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 2011, **12**(12): 739–751
- 84 Ghadiri-zadeh A, Maki A, Kragic D, Björkman M. Deep predictive policy training using reinforcement learning. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, BC, Canada: IEEE, 2017. 2351–2358
- 85 Schou C, Andersen R S, Chrysostomou D, Bogh S, Madsen O. Skill-based instruction of collaborative robots in industrial settings. *Robotics and Computer-Integrated Manufacturing*, 2018, **53**: 72–80
- 86 Bekiroglu Y, Laaksonen J, Jorgensen J A, Kyrki V. Assessing grasp stability based on learning and haptic data. *IEEE Transactions on Robotics*, 2011, **27**(3): 616–629
- 87 Dang H, Allen P K. Learning grasp stability. In: Proceedings of the 2012 IEEE International Conference on Robotics and Automation. Saint Paul, MN, USA: IEEE, 2012. 2392–2397
- 88 Levine S, Pastor P, Krizhevsky A, Ibarz J, Quillen D. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research*, 2018, **37**(4–5): 421–436
- 89 Finn C, Goodfellow I, Levine S. Unsupervised learning for physical interaction through video prediction. In: Proceedings of the 30th Neural Information Processing Systems. Barcelona, Spain: MIT Press, 2016: 64–72
- 90 Finn C, Levine S. Deep visual foresight for planning robot motion. In: Proceedings of the 2017 IEEE International Conference on Robotics and Automation. Singapore, Singapore: IEEE, 2017. 2786–2793
- 91 Petrić T, Gams A, Colasanto L, Ijspeert A J, Ude A. Accelerated sensorimotor learning of compliant movement primitives. *IEEE Transactions on Robotics*, 2018, **34**(6): 1636–1642
- 92 Huang P C, Hsieh Y H, Mok A K. A skill-based programming system for robotic furniture assembly. In: Proceedings of the 16th IEEE International Conference on Industrial Informatics. Porto, Portugal: IEEE, 2018. 355–361
- 93 Qin F, Xu D, Zhang D, Li Y. Robotic skill learning for precision assembly with microscopic vision and force feedback. *IEEE/ASME Transactions on Mechatronics*, **24**(3): 1117–1128
- 94 Ni Zi-Qiang, Wang Tian-Miao, Liu Da. Vision guide based teaching programming for industrial robot. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, **42**(3): 562–568  
(倪自强, 王田苗, 刘达. 基于视觉引导的工业机器人示教编程系统. *北京航空航天大学学报*, 2016, **42**(3): 562–568)
- 95 Hu D Y, Gong Y Z, Hannaford B, Seibel E J. Semi-autonomous simulated brain tumor ablation with RavenII surgical robot using behavior tree. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 3868–3875
- 96 Ewerton M, Neumann G, Lioutikov R, Amor H B, Peters J, Maeda G, et al. Learning multiple collaborative tasks with a mixture of interaction primitives. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 1535–1542

- 97 Silvério J, Calinon S, Rozo L, Caldwell D G. Bimanual skill learning with pose and joint space constraints. In: Proceedings of the 2018 IEEE/RAS International Conference on Humanoid Robots. Beijing, China: IEEE, 2018. 153–159
- 98 Figueroa N, Ureche A L P, Billard A. Learning complex sequential tasks from demonstration: a pizza dough rolling case study. In: Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction. Christchurch, New Zealand: IEEE, 2016. 611–612
- 99 Calinon S, Sardellitti I, Caldwell D G. Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies. In: Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, China: IEEE, 2010. 249–254
- 100 Ureche A L P, Billard A. Analyzing human behavior and bootstrapping task constraints from kinesthetic demonstrations. In: Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts. Portland, Oregon, USA: ACM, 2015: 199–200
- 101 Muhlig M, Gienger M, Hellbach S, Steil J J, Goerick C. Task-level imitation learning using variance-based movement optimization. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan: IEEE, 2009. 1177–1184
- 102 Gupta A, Eppner C, Levine S, Abbeel P. Learning dexterous manipulation for a soft robotic hand from human demonstrations. In: Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Daejeon, South Korea: IEEE, 2016. 3786–3793
- 103 Peters J, Schaal S. Reinforcement learning of motor skills with policy gradients. *Neural Networks*, 2008, **21**(4): 682–697
- 104 Xu W J, Chen J, Lau H Y K, Ren H L. Automate surgical tasks for a flexible serpentine manipulator via learning actuation space trajectory from demonstration. In: Proceedings of the 2016 IEEE International Conference on Robotics and Automation. Stockholm, Sweden: IEEE, 2016. 4406–4413
- 105 Murali A, Sen S, Kehoe B, Garg A, McFarland S, Patil S, et al. Learning by observation for surgical subtasks: multilateral cutting of 3D viscoelastic and 2D orthotropic tissue phantoms. In: Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Seattle, WA, USA: IEEE, 2015. 1202–1209
- 106 Ureche L P, Billard A. Constraints extraction from asymmetrical bimanual tasks and their use in coordinated behavior. *Robotics and Autonomous Systems*, 2018, **103**: 222–235
- 107 Salehian S S M, Khoramshahi M, Billard A. A dynamical system approach for softly catching a flying object: theory and experiment. *IEEE Transactions on Robotics*, 2016, **32**(2): 462–471
- 108 Kalashnikov D, Irpan A, Pastor P, Ibarz J, Herzog A, Jang E, et al. Scalable deep reinforcement learning for vision-based robotic manipulation. In: Proceedings of the 2nd Conference on Robot Learning. Zurich, Switzerland: PMLR, 2018. 651–673
- 109 Deng J, Dong W, Socher R, Li L J, Li K, Li F F. Imagenet: a large-scale hierarchical image database. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009. 248–255
- 110 Du Z H, He L, Chen Y N, Xiao Y, Gao P, Wang T Z. Robot cloud: bridging the power of robotics and cloud computing. *Future Generation Computer Systems*, 2015, **21**(4): 301–312
- 111 Kehoe B, Patil S, Abbeel P, Goldberg K. A survey of research on cloud robotics and automation. *IEEE Transactions on Automation Science and Engineering*, 2015, **12**(2): 398–409
- 112 Hu G Q, Tay W P, Wen Y G. Cloud robotics: architecture, challenges and applications. *IEEE Network*, 2012, **26**(3): 21–28
- 113 Hunziker D, Gajamohan M, Waibel M, D'Andrea R. Rapyuta: the RoboEarth cloud engine. In: Proceedings of the 2013 IEEE International Conference on Robotics and Automation. Karlsruhe, Germany: IEEE, 2013. 438–444
- 114 Saxena A, Jain A, Sener O, Jami A, Misra D K, Koppula H S. Robobrain: large-scale knowledge engine for robots. arXiv: 1412.0691, 2014.
- 115 Wang Fei-Yue. Knowledge Robot and Industry 5.0. In: Proceedings of the 2015 China National Robotics Development Forum. Beijing, China: Chinese Association of Automation, 2015.  
(王飞跃. 知识机器人与工业 5.0. 2015 年国家机器人发展论坛. 北京: 中国自动化学会, 2015.)
- 116 Bai Tian-Xiang, Wang Shuai, Shen Zhen, Cao Dong-Pu, Zheng Nan-Ning, Wang Fei-Yue. Parallel robotics and parallel unmanned systems: framework, structure, process, platform and applications. *Acta Automatica Sinica*, 2017, **43**(2): 161–175  
(白天翔, 王帅, 沈震, 曹东璞, 郑南宁, 王飞跃. 平行机器人与平行无人系统: 框架、结构、过程、平台及其应用. 自动化学报, 2017, **43**(2): 161–175)
- 117 Wang Fei-Yue. Software-defined systems and knowledge automation: a parallel paradigm shift from Newton to Merton. *Acta Automatica Sinica*, 2015, **41**(1): 1–8  
(王飞跃. 软件定义的系统与知识自动化: 从牛顿到默顿的平行升华. 自动化学报, 2015, **41**(1): 1–8)



秦方博 中国科学院自动化研究所博士研究生. 2013 年获得北京交通大学电子信息工程学院学士学位. 主要研究方向为机器人视觉感知与控制, 精密装配.  
E-mail: qinfangbo2013@ia.ac.cn

(QIN Fang-Bo Ph.D. candidate at the Institute of Automation, Chinese Academy of Sciences. He received his

bachelor degree from the School of Electronic and Information Engineering, Beijing Jiaotong University in 2013. His research interest covers robot vision based perception and control, and precision assembly.)



徐德 中国科学院自动化研究所研究员. 于 1985 年和 1990 年获得山东工业大学学士和硕士学位, 2001 年获得浙江大学博士学位. 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. 本文通信作者.  
E-mail: de.xu@ia.ac.cn

(XU De Professor at the Institute of Automation, Chinese Academy of Sciences. He received his bachelor and master degrees from Shandong University of Technology in 1985 and 1990, respectively, and received his Ph.D. degree from Zhejiang University in 2001. His research interest covers robotics and automation such as visual measurement, visual control, intelligent control, visual positioning, microscopic vision, and microassembly. Corresponding author of this paper.)