

DOI: 10.19650/j.cnki.cjsi.J2413114

一种基于 DP-KMP 的机器人避障交互式学习方法*

肖 洒, 吕勇明, 吴海彬

(福州大学机械工程及自动化学院 福州 350108)

摘要:为了使机器人在执行复杂任务时能够及时避开障碍,提出了一种基于 DP-KMP 的机器人避障交互式学习方法。首先构建了该方法的整体框架,采用分割泛化策略,实现对示教轨迹的快速分段学习和对分段轨迹的避障规划;针对学习阶段,提出了基于 DP 算法的轨迹分割策略以提高分割效率,并使用高斯混合模型策略提取各子轨迹的参考数据库;针对轨迹规划阶段,使用 KMP 模型完成轨迹复现与泛化,并引入基于人机交互反馈的参考数据库更新策略,提升了人机交互避障的成功率;针对该更新策略可能失效导致避障轨迹规划失败的问题,提出了两个相应的适用条件用于检验分割生成的子轨迹。最后,通过仿真验证了所述适用条件的有效性;真实实验结果表明,使用所提出的方法分割两个实验的示教轨迹分别仅用时 0.084 和 0.107 s, KUKA 协作机器人在执行不同搬运任务的过程中通过与用户的多次交互成功避开了所有静止和突然变化的障碍。

关键词: 模仿学习;核化运动基元;道格拉斯-普克算法;轨迹分割;人-机器人交互;避障

中图分类号: TH89 TP242 **文献标识码:** A **国家标准学科分类代码:** 510.8050

Interactive learning approach for robot obstacle avoidance based on DP-KMP

Xiao Sa, Lyu Yongming, Wu Haibin

(School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China)

Abstract: To enable the robot avoid obstacles in real-time while performing complex tasks, an interactive learning approach for robot obstacle avoidance based on DP-KMP is proposed. First, the whole framework of this approach is constructed, which adopts the segmentation-generalization strategy to implement the learning of demonstrated trajectories with rapid segmentation and the planning of sub-trajectories for obstacle avoidance. During the learning phase, a trajectory segmentation strategy based on the DP algorithm is proposed to improve the efficiency of segmentation, while a Gaussian mixture model strategy is used to extract the reference database from each sub-trajectory. In the trajectory planning phase, the KMP model is used to implement the trajectory reproduction and generalization, while the reference database update strategy based on human-robot interaction feedback is introduced to enhance the success rate of obstacle avoidance through human-robot interaction. Aiming at the issue that this update strategy may fail to plan a successful obstacle avoidance trajectory, two available conditions are proposed for inspecting the sub-trajectories generated by segmentation. Finally, the effectiveness of mentioned available conditions is verified by simulation. Experimental results show that the proposed approach segments the demonstrated trajectories in just 0.084 and 0.107 s for two different experiments, respectively. Additionally, the KUKA cobot successfully avoids all static and suddenly changing obstacles through multiple interactions with user during the execution of the different lift-place tasks.

Keywords: imitation learning; kernelized movement primitives; Douglas-Peucker algorithm; trajectory segmentation; human-robot interaction; obstacle avoidance

0 引言

与传统的机器人编程技术相比,以强化学习^[1]、深度学习^[2]和模仿学习^[3]等为代表的机器人技能学习方法可以有效提高机器人对不同工作环境的适应能力。在这些方法中,模仿学习^[3-5]由于可以帮助机器人直接从人类的示教行为中学习相应的操作技能,具备较高的编程效率,因此得到了广泛的关注。目前已经有许多具备泛化能力的模仿学习模型被开发了出来,例如动态运动基元(dynamic movement primitives, DMP)^[6]、任务参数化高斯混合模型(task-parameterized Gaussian mixture model, TP-GMM)^[7]、概率运动基元(probabilistic movement primitives, ProMP)^[8]以及核化运动基元(kernelized movement primitives, KMP)^[9]等。然而,将机器人学到的轨迹应用到新任务中还需要考虑可能变化的约束条件,例如工作场景中出现的示教时不存在的障碍等。如何使机器人末端在运动中及时避开各种障碍一直是模仿学习领域研究的热点和难点。

针对避障问题,许多研究学者提出的解决方案是将基于避障的轨迹自适应策略集成到模仿学习框架中来应对。Huang等^[10]对TP-GMM的局部坐标系用强化学习进行优化,然后生成自适应的避障轨迹;Lu等^[11]结合DMP模型和障碍李雅普诺夫函数的控制概念,提出了一种包含各种分类约束的DMP通用框架以解决受约束的路径规划问题,可用于静止和动态环境下的避障任务;Zhai等^[12]提出了一种基于DMP模型和改进避障算法的运动规划方法,可利用机器人与障碍物的相对速度和距离应对动态避障问题;Frank等^[13]将中间通过点(via-point)和相互避障等各种类型的避障自适应技术集成到了一个通用的ProMP框架中,用于解决复杂的机器人避障问题;Davoodi等^[14]将通过设置机器人避开障碍的距离阈值并使用基于优化的ProMP模型来规划避障轨迹,但该方法不适用于动态避障。这些方法虽然可以通过预定义的自适应策略来规划避障轨迹,然而却难以应对无法预测的突发状况,例如突然出现并占据任务空间的未知障碍或突然移动的动态障碍等。

由于人类在非结构化环境中能够做出可靠反应,因此让机器人根据用户的指导调整所学到的轨迹,可以有效应对上述无法预测的状况,从而保障机器人的安全^[9]。Lu等^[15]提出了一种基于DMP和人体肌电信号反馈的增量式技能学习方法,可以通过肌电信号预测人类的意图并指导机器人避开未知障碍,但其避障轨迹无法在线实时调整;Zhang等^[16]提出了一种基于神经网络和改进的稳定动态系统法的学习框架,能够通过设置中间通过点约束或根据用户施加的外力使机器人躲开障碍物;

Fu等^[17]提出了一种基于高斯混合回归的非参数模仿学习算法,也可以实现与文献[16]方法类似的避障效果;然而,文献[16-17]的方法在使用中间通过点约束避障轨迹时需要预设该点信息,即用户无法在交互过程中实时指定期望中间通过点的输入与输出信息,不利于应对动态障碍。Huang等^[9]利用人机交互的反馈获得期望的中间通过点信息,然后使用KMP模型生成受该点约束的自适应避障轨迹,在实验中成功通过人工引导使机器人及时躲开了多个未知障碍;Sun等^[18]在实验中也直接通过人机交互获得的中间通过点约束基于第2类模糊模型的运动基元(type-2 fuzzy model-based movement primitives, T2FMP)模型生成的避障轨迹。然而,文献[9,18]的交互方法都只能在预设时刻使用,例如机器人即将碰到障碍物时。Xiao等^[19]在文献[9]的基础上提出了一种基于人机交互反馈的KMP参考数据库更新策略,使得用户可以在机器人执行当前任务的任何时刻出手实时调整其末端轨迹以避开各种静止和动态障碍物。然而,该策略在应对某些复杂轨迹时会出现失效的情况,需要对此类轨迹采取分割泛化策略进行处理。

分割泛化策略通常被用于处理复杂的多步骤运动轨迹,其思路是将整个示教轨迹分割为一系列对应不同步骤的子轨迹,然后通过每个步骤的子轨迹单独建模和泛化,使机器人可以应对多个步骤的目标泛化需求,也可以针对外界环境的变化在各步骤作出不同的泛化响应^[20-21]。该策略的核心步骤是轨迹分割。在这方面,近年来广泛使用的分割方法是将示教轨迹看作一个隐马尔可夫模型(hidden Markov model, HMM)过程^[3],采用非参数贝叶斯的聚类方法确定该HMM所含的不同潜在状态,即不同的基本动作,并以此作为分割轨迹的依据,例如分层狄利克雷过程隐马尔可夫模型(Hierarchical dirichlet process HMM, HDP-HMM)^[22]、贝塔过程自回归隐马尔可夫模型(Beta process autoregressive HMM, BP-AR-HMM)^[21]等。此类方法无需先验设置分段数量便可获得分割结果,然而在学习过程中所使用的基于采样的方法估算相关参数需要消耗大量时间。

针对上述基于人机交互反馈的KMP参考数据库更新策略和轨迹分割策略各自存在的问题,本文提出了一种基于道格拉斯-普克-核化运动基元(Douglas-Peucker-kernelized movement primitives, DP-KMP)的机器人避障交互式学习方法,可以让执行复杂任务的机器人在用户的指导下实时避开障碍。在KUKA协作机器人平台进行的实验结果及仿真对比验证了该方法的有效性。本文的主要贡献如下:

- 1) 提出了一种将分割泛化策略、基于人机交互的反馈与KMP模型相结合来规划机器人末端避障轨迹的学习方法,可以帮助机器人在执行复杂多步骤任务时根据

任意时刻人机交互反馈得到的期望中间通过点来约束相应的子轨迹,有效提升人机交互避障的成功率;

2) 提出了一种基于道格拉斯-普克 (Douglas-Peucker, DP) 算法的轨迹分割策略,有效提高了分割效率;

3) 提出了基于人机交互反馈的 KMP 参考数据库更新策略的适用条件,用于检验示教轨迹的分割方案,避免

避障轨迹规划失效。

1 机器人避障交互式学习方法概述

本文提出了基于 DP-KMP 的交互式学习方法,用于规划机器人末端自适应避障轨迹,其框架包括示教、学习和避障轨迹规划等阶段,具体流程如图 1 所示。

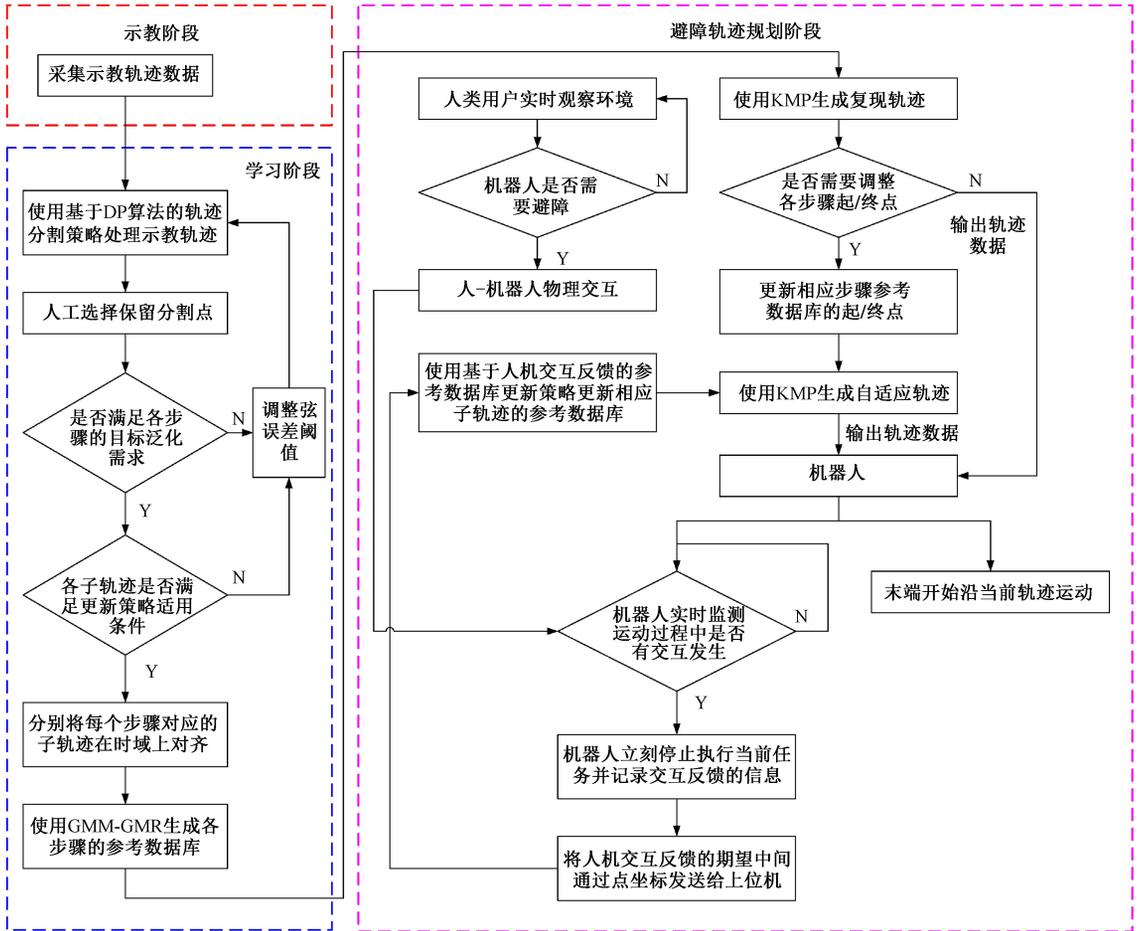


图 1 基于 DP-KMP 的机器人避障交互式学习框架流程

Fig. 1 Flow chart of interactive learning framework for robot obstacle avoidance based on DP-KMP

在示教阶段,由人类示教者(用户)演示执行特定任务的动作,并采集相应的运动轨迹数据供后续建模学习。在这个阶段需要对该任务进行多次完整的演示。

在学习阶段,首先使用基于 DP 算法的轨迹分割策略(详见第 2.1 节算法 1)对所有示教轨迹进行分割;接下来,分别将每个步骤的子轨迹在各自时域上对齐^[23],并利用高斯混合模型与回归 (Gaussian mixture model-Gaussian mixture regression, GMM-GMR) 提取各子轨迹的参考数据库。请注意,在确定分割方案时需满足以下约

束条件:1) 满足所有步骤各自的目标泛化需求;2) 所有子轨迹都满足所提出的更新策略适用条件(详见第 3.3 节式(17)和(18))。

在避障轨迹规划阶段,用 KMP 模型分别对各步骤的当前参考数据库进行编码。在这个阶段,除了直接复现示教轨迹,还可以根据具体的任务需求调整相应步骤的起点/终点,从而泛化初始轨迹(详见第 3.1 节);在机器人执行任务期间,可以通过任意时刻人机交互的反馈(添加中间通过点约束)实时单独调整各步骤轨迹以应对障

碍(详见第3.2节算法2)。在此期间,由上位机(计算机)计算出当前所需的轨迹,然后将相关数据发送给下位机(机器人);同时,上位机全程监听通讯端口,随时接收机器人反馈的人机交互信息。

2 基于 DP-GMM 的示教轨迹分段学习策略

针对传统的轨迹自动分割方法由于使用基于采样的参数估计策略导致计算效率较低的问题,本文提出了一种基于 DP 算法的轨迹分割策略以提升计算效率,并结合 GMM-GMR 算法对生成的子轨迹分别建模,从而获得相应的参考数据库供后续初始化 KMP 模型。

2.1 基于 DP 算法的轨迹分割策略

DP 算法^[24-25]通常用于简化曲线,其核心思路是基于给定的误差阈值 d_{th} 从初始曲线中提取一系列关键点,然后通过相邻关键点间的连线来拟合该曲线。具体而言,设曲线的位置序列 $\{\mu_n\}_{n=1}^N = \{(x_n, y_n, z_n)\}_{n=1}^N$, 其中 N 表示序列长度。首先以连接起点 μ_1 与终点 μ_N 之间的直线为该曲线的弦,然后计算出其余样本点 $\mu_i, \forall i \in \{2, 3, \dots, N-1\}$ 与该弦的距离 d_i :

$$d_i = \frac{\|A_i\|_2}{\|\mu_N - \mu_1\|_2} \quad (1)$$

式中: $A_i = \begin{vmatrix} i & j & k \\ x_1 - x_i & y_1 - y_i & z_1 - z_i \\ x_N - x_1 & y_N - y_1 & z_N - z_1 \end{vmatrix}$, 且 i, j, k 分别是

X, Y, Z 轴方向的单位向量; $\|\cdot\|_2$ 表示 2-范数; 接下来, 搜索其中的最大值 $d_{\max} = d_s$ 以及对应的样本点 μ_s ($2 \leq s \leq N-1$)。如果 $d_{\max} < d_{th}$, 则将 μ_1 和 μ_N 标记为关键点, 并删除其余样本点; 如果 $d_{\max} \geq d_{th}$, 则将 μ_s 作为关键点将当前曲线分成两段, 然后对这两段曲线分别进行递归处理。通过不断地提取关键点并删除其余样本点, 最终得到一条由关键点组成的简化曲线。

本文参考 DP 算法的原理提出了一种轨迹分割策略, 以所提取的关键点作为分割点(除了起点与终点), 将初始轨迹分割成一系列子轨迹。与 DP 算法不同的是, 该分割策略保留所有样本点, 最终输出的是分割点在原序列中的索引号集合, 其算法伪代码如下所示。

算法 1 基于 DP 算法的轨迹分割策略

输入: 初始轨迹样本点 $H = \{\mu_n\}_{n=1}^N$, 误差阈值 d_{th} ;

输出: 分割点在 H 中索引号集合 L ;

1) Function $L = \text{DPS}(H, d_{th})$

2) 初始化 $L = \text{NULL}$, 其中 L 用于存放分割点索引号;

3) 设 H 的始末样本点索引号分别为 a 和 b , 按式(1) 计算中各点到 $\overrightarrow{\mu_a \mu_b}$ 的距离 d_i ($a+1 \leq i \leq b-1$), 并检索其中最大

值 $d_{\max} = \max\{d_{a+1}, \dots, d_{b-1}\}$ 及其对应点 μ_s

if $d_{\max} < d_{th}$

$L = \{L, a, b\}$

else

$L = \{L, \text{DPS}(\{\mu_j\}_{j=a}^s, d_{th})\}$

$L = \{L, \text{DPS}(\{\mu_j\}_{j=s}^b, d_{th})\}$

end if

4) end Function

5) $L = \{L / \{1, N\}\}$, 其中符号‘/’表示排除操作

算法 1 进行轨迹分割的原理如图 2 所示。图中的初始轨迹样本点集合为 $\{\mu_i\}_{i=1}^N$, 其中 $N=8$ 。首先由式(1) 计算各点到初始弦的距离 d_i , 得 $d_{\max} = d_5$ 。假设给定的 $d_{th} < d_5$, 则 μ_5 被标记为分割点, 即当前轨迹被其分割成 2 段, 所对应的弦分别如图 2(a) 中的弦 1 和弦 2 所示。接下来, 设分别计算得点 μ_2, μ_3 和 μ_4 各自到弦 1 的距离以及点 μ_6 和 μ_7 各自到弦 2 的距离均小于 d_{th} , 因此初始轨迹有一个分割点 μ_5 。

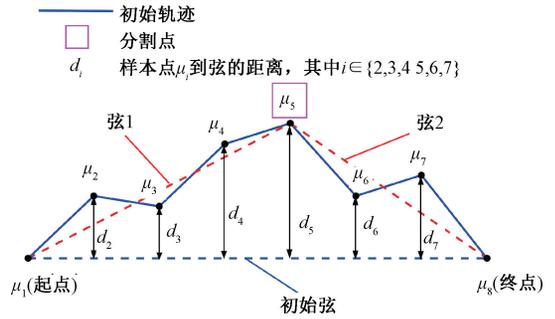


图 2 基于 DP 算法的轨迹分割策略原理

Fig. 2 Principle illustration of the trajectory segmentation strategy based on DP algorithm

2.2 基于 GMM-GMR 的参考数据库提取策略

GMM^[26-27]是一种基于统计理论的概率模型, 具有强大的信息表达能力, 常被用于对示教轨迹的编码。给定 M 条示教轨迹 $\{\{\xi_{n,m}\}_{n=1}^{N_m}\}_{m=1}^M = \{\{\xi_j\}_{j=1}^F\}$, 其中 N_m 表示第 m 条轨迹的样本点数量, $F = \sum_{m=1}^M N_m$; 样本点 $\xi_j = [\xi_j^1, \xi_j^0]$, 其中 ξ_j^1 和 ξ_j^0 分别表示输入和输出信息。使用 GMM 对样本建模, 其概率密度函数为 $p(\xi)$ 为:

$$p(\xi) = \sum_{k=1}^K \pi_k \mathcal{N}(\xi; \mu_k, \Sigma_k) \quad (2)$$

式中: $\pi_k, \mu_k = \begin{bmatrix} \mu_k^1 \\ \mu_k^0 \end{bmatrix}$ 以及 $\Sigma_k = \begin{bmatrix} \Sigma_k^1 & \Sigma_k^{10} \\ \Sigma_k^{01} & \Sigma_k^0 \end{bmatrix}$ 分别表示第 k

个高斯成分的先验概率、均值和协方差矩阵; K 表示高斯成分的个数, 可通过 BIC 算法 (Bayesian information criterion, 贝叶斯信息准则) 预估。为了确保提取的示教

行为特征信息更准确,在训练 GMM 之前可使用动态时间规整(dynamic time warping, DTW)算法^[23]对示教轨迹进行对齐预处理。

GMM 的参数 $\{\pi_k, \mu_k, \Sigma_k\}_{k=1}^K$ 可以使用期望最大值(expectation maximum, EM)算法进行迭代求解^[27]。对于以时间为输入,空间坐标为输出的示教轨迹,获得 GMM 参数后,给定时间输入序列 $\{t_n\}_{n=1}^N$ 便可利用 GMR^[28]得到对应的概率参考轨迹 $D = \{t_n, \hat{\mu}_n, \hat{\Sigma}_n\}_{n=1}^N$ 。此轨迹以期望均值 $\hat{\mu}_n$ 和协方差 $\hat{\Sigma}_n$ 的形式封装了示教行为的所有特征,也被称为参考数据库(the reference database^[9]),可用 KMP 模型对其进行进一步学习与泛化。

3 基于人机交互反馈的 KMP 避障轨迹规划

在使用 KMP 模型结合人机交互反馈的期望中间通过点规划避障轨迹时,需要用该点的信息更新当前参考数据库,然而在更新过程中存在难以确定该点的时间输入,无法满足用户在任意时刻与机器人交互的需求等问题,因此难以根据用户的指导实时生成期望的避障轨迹。为此本章首先介绍 KMP 模型学习并泛化以时间为输入的示教轨迹的原理,接下来引入基于人机交互反馈的参考数据库更新策略,提升人机交互避障的成功率;最后,提出该更新策略相应的适用条件,用于检验分割生成的子轨迹,避免避障轨迹规划失效的问题。

3.1 基于 KMP 的轨迹复现与泛化

KMP 模型^[9]的核心思想是利用 KL 散度(Kullback-Leibler divergence, KL-divergence)来优化参数轨迹的概率分布,使其与参考轨迹的概率分布相匹配。具体而言,首先构造时间输入 t 和空间坐标输出 s 的参数化模型,如下所示:

$$s(t) = \Phi^T(t)w \quad (3)$$

式中: $\Phi(t) = I_0 \otimes \varphi(t) \in \mathbf{R}^{B \times O}$; $\varphi(t) \in \mathbf{R}^B$ 表示 B 维基函数;KMP 假设 w 服从正态分布,即 $w \sim \mathcal{N}(\mu_w, \Sigma_w)$,其中 μ_w 和 Σ_w 是未知的。接下来为了估算 μ_w 和 Σ_w ,KMP 对由式(3)生成轨迹的概率分布和参考数据库之间的 KL 散度进行最小化,即:

$$J_{ini}(\mu_w, \Sigma_w) = \sum_{n=1}^N D_{KL}(\mathcal{P}_{para}(s|t_n) \parallel \mathcal{P}_{ref}(s|t_n)) \quad (4)$$

其中, $\mathcal{P}_{para}(s|t_n) = \mathcal{N}(s|\Phi^T(t_n)\mu_w, \Phi^T(t_n)\Sigma_w\Phi(t_n))$, $\mathcal{P}_{ref}(s|t_n) = \mathcal{N}(s|\hat{\mu}_n, \hat{\Sigma}_n)$ 。接下来对式(3)进行分解,利用向量和矩阵求导并使用核技巧便可获得任意时间输入 t^* 对应轨迹 $s(t^*)$ 的均值以及协方差:

$$E(s(t^*)) = k^*(K + \lambda_1\Sigma)^{-1}\mu \quad (5)$$

$$D(s(t^*)) = \frac{N}{\lambda_2} (k(t^*, t^*) - k^*(K + \lambda_2\Sigma)^{-1}k^{*T}) \quad (6)$$

其中,

$$k^* = [k(t^*, t_1) \quad k(t^*, t_2) \quad \cdots \quad k(t^*, t_N)] \quad (7)$$

$$K = \begin{bmatrix} k(t_1, t_1) & k(t_1, t_2) & \cdots & k(t_1, t_N) \\ k(t_2, t_1) & k(t_2, t_2) & \cdots & k(t_2, t_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(t_N, t_1) & k(t_N, t_2) & \cdots & k(t_N, t_N) \end{bmatrix} \quad (8)$$

$$\mu = [\hat{\mu}_1^T, \hat{\mu}_2^T, \cdots, \hat{\mu}_N^T]^T \quad (9)$$

$$\Sigma = \text{blockdiag}(\hat{\Sigma}_1, \hat{\Sigma}_2, \cdots, \hat{\Sigma}_N) \quad (10)$$

式中: $\lambda_1 > 0$ 以及 $\lambda_2 > 0$ 均为正则化系数; $k(\cdot, \cdot)$ 则被定义为:

$$k(t_i, t_j) = \begin{bmatrix} k_u(i, j)I_0 & k_{ud}(i, j)I_0 \\ k_{du}(i, j)I_0 & k_{dd}(i, j)I_0 \end{bmatrix} \quad (11)$$

其中,

$$k_u(i, j) = k(t_i, t_j)$$

$$k_{ud}(i, j) = \frac{k(t_i, t_j + \delta) - k(t_i, t_j)}{\delta}$$

$$k_{du}(i, j) = \frac{k(t_i + \delta, t_j) - k(t_i, t_j)}{\delta}$$

$$k_{dd}(i, j) =$$

$$\frac{k(t_i + \delta, t_j + \delta) - k(t_i + \delta, t_j) - k(t_i, t_j + \delta) + k(t_i, t_j)}{\delta^2} \quad (12)$$

式中: $k(t_i, t_j) = \exp(-l(t_i - t_j)^2)$ 是超参数 $l > 0$ 的高斯核函数,而 $\delta > 0$ 是一个很小的常数。

对于示教轨迹的复现,可以直接使用式(5)进行轨迹规划;对于轨迹的泛化,KMP 提供了通用的更新策略将任意期望点(例如起点/终点/中间通过点)更新到当前参考数据库中。具体而言,给定 M 个期望通过点集合 $\{\bar{t}_m, \bar{\mu}_m, \bar{\Sigma}_m\}_{m=1}^M$,其中 $\bar{t}_m, \bar{\mu}_m$ 和 $\bar{\Sigma}_m$ 分别为第 m 个期望中间通过点的时间输入、空间坐标输出以及自适应精度,KMP 按照如下公式将 $\{\bar{t}_m, \bar{\mu}_m, \bar{\Sigma}_m\}$ 更新进当前 D :

$$\begin{cases} D \leftarrow \{D / \{\bar{t}_r, \hat{\mu}_r, \hat{\Sigma}_r\}\} \cup \{\bar{t}_m, \bar{\mu}_m, \bar{\Sigma}_m\}, \\ \quad \text{if } d(\bar{t}_m, t_r) < \zeta \\ D \leftarrow D \cup \{\bar{t}_m, \bar{\mu}_m, \bar{\Sigma}_m\}, \quad \text{其他} \end{cases} \quad (13)$$

其中, $r = \text{argmin}_n d(\bar{t}_m, t_n)$, $n \in \{1, 2, \cdots, N\}$,而符号‘/’和‘ \cup ’分别表示排除和并集操作; $d(\cdot)$ 表示欧式距离,阈值 $\zeta > 0$ 。然后应用式(5),可以根据更新后的 D 生成受该期望通过点约束的泛化轨迹。

3.2 基于人机交互反馈的参考数据库更新策略

在文献[9,18]中的人机交互避障实验中,对于第 m 个期望中间通过点 $\{\bar{t}_m, \bar{\mu}_m, \bar{\Sigma}_m\}$,其中时间输入 \bar{t}_m 是由对应的交互发生时刻 t_m^* 修正而来,即:

$$\bar{t}_m = t_m^* + \delta_t \quad (14)$$

其中, $\delta_t > 0$ 是时间调节参数。然后结合由交互作用力换算的期望点坐标 $\bar{\mu}_m$ 与自适应精度 $\bar{\Sigma}_m$ 应用式(13)即可完成对当前参考数据库 $D = \{t_n, \hat{\mu}_n, \hat{\Sigma}_n\}_{n=1}^N$ 的更新。然而这种方法需要预设 δ_t 和交互发生时刻 t_m^* 。

文献[19]针对末端避障的需求在式(13)的基础上提出了一种基于人机交互反馈的参考数据库更新策略,其核心思路是:当机器人末端按照当前轨迹 D 移动,预计将在 t_c 时刻碰到障碍,其中 $t_1 \leq t_c \leq t_N$,因此只要将 t_c 赋予用户期望机器人避开碰撞的位置坐标 $\bar{\mu}_m$,再结合指定的自适应精度 $\bar{\Sigma}_m$ 组成一个新的中间通过点 $\{t_c, \bar{\mu}_m, \bar{\Sigma}_m\}$ 来约束泛化轨迹就能避开当前障碍。需要注意的是,这里 t_c 作为一个未知时刻可通过以下假设由 $\{t_n\}_{n=1}^N$ 中的样本点来近似替代:

假设1:预期碰撞点C的坐标 μ_c 可以用 D 的位置序列 $\{\hat{\mu}_n\}_{n=1}^N$ 中最接近前者的样本点 $\hat{\mu}_r$ 来近似,其中 $r = \operatorname{argmin}_n d(\mu_c, \hat{\mu}_n), n \in \{1, 2, \dots, N\}$ 。

由假设1可知,所求的期望通过点 $\{t_c, \bar{\mu}_m, \bar{\Sigma}_m\}$ 可用 $\{t_r, \bar{\mu}_m, \bar{\Sigma}_m\}$ 近似替代。为了得到 t_r ,首先根据参考数据库 $D = \{t_n, \hat{\mu}_n, \hat{\Sigma}_n\}_{n=1}^N$ 构建相应的虚拟参考数据库(the virtual reference database^[19]) $D' = \{t_n, \hat{\mu}'_n, \hat{\Sigma}_n\}_{n=1}^N$,其中位置序列 $\{\hat{\mu}'_n\}_{n=1}^N = \{(x'_n, y'_n, z'_n)\}_{n=1}^N$ 可以按照如下方式计算获得:

$$\begin{cases} x'_n = x'_1 + \frac{(x'_N - x'_1)(n-1)}{N-1}, \\ y'_n = y'_1 + \frac{(y'_N - y'_1)(n-1)}{N-1}, \\ z'_n = z'_1 + \frac{(z'_N - z'_1)(n-1)}{N-1}, \end{cases} \forall n \in \{1, 2, \dots, N\} \quad (15)$$

式中: $\{\hat{\mu}'_n\}_{n=1}^N$ 的起点 $\hat{\mu}'_1 = (x'_1, y'_1, z'_1) = (x_1, y_1, z_1)$,且其终点 $\hat{\mu}'_N = (x'_N, y'_N, z'_N) = (x_N, y_N, z_N)$ 。即 $\{\hat{\mu}'_n\}_{n=1}^N$ 与 $\{\hat{\mu}_n\}_{n=1}^N$ 有共同的起终点, $\{\hat{\mu}'_n\}_{n=1}^N$ 也被称为参考直线。

接下来按下列步骤选取期望中间通过点坐标 $\bar{\mu}_m$:

1) 通过观察机器人与障碍的状态来预测可能发生碰撞的位置C,如图3(a)所示;

2) 选择一个期望的中间通过点位置 $\bar{\mu}_m$ 以约束泛化轨迹从而避开障碍。请注意,该点与所预测的碰撞位置之间的连线应大致垂直于参考直线, $\bar{\mu}_m$ 可由机器人自动记录,如图3(b)、(c)所示。

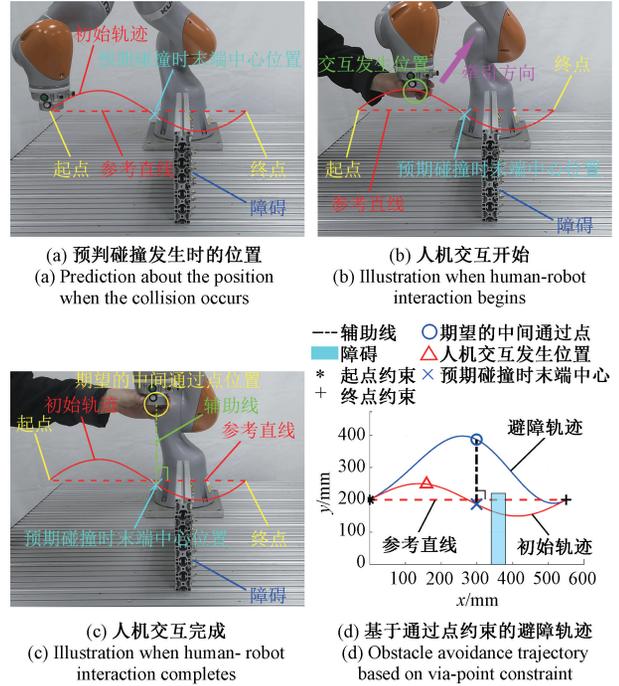


图3 基于人机交互反馈的KMP避障轨迹规划

Fig.3 Example illustration of KMP trajectory planning for obstacle avoidance based on the feedback from human-robot interaction

最后, $\bar{\mu}_m$ 对应的时间输入可由以下假设确定:

假设2:计算 $\bar{\mu}_m$ 与 $\{\hat{\mu}'_n\}_{n=1}^N$ 中各点的欧式距离获得其中最小值对应的点为 $\hat{\mu}'_r, r = \operatorname{argmin}_n d(\bar{\mu}_m, \hat{\mu}'_n)$,则该点与 $\bar{\mu}_m$ 、 $\hat{\mu}'_r$ 的输入时刻均为 t_r 。

综上所述,可按如下方式更新当前 D :

$$D \leftarrow \{D / \{t_r, \hat{\mu}'_r, \hat{\Sigma}_r\}\} \cup \{t_r, \bar{\mu}_m, \bar{\Sigma}_m\} \quad (16)$$

其中, $r = \operatorname{argmin}_n d(\bar{\mu}_m, \hat{\mu}'_n), n \in \{1, 2, \dots, N\}$ 。随后,将更新后的 D 用于训练KMP模型,即可生成在 t_r 时刻以 $\bar{\Sigma}_m$ 约束的精度经过 $\bar{\mu}_m$ 的泛化轨迹。图3(a)~(c)所示人机交互避障实例对应的KMP仿真避障轨迹如图3(d)所示;生成该轨迹所用的基于人机交互反馈的KMP避障轨迹规划方法如算法2所示。

算法2 基于人机交互反馈的KMP避障轨迹规划

输入:参考数据库 $D = \{t_n, \hat{\mu}_n, \hat{\Sigma}_n\}_{n=1}^N$,超参数 λ_l (见式(5)),超参数 l (见式(5)的 $k(t_i, t_j)$),基于人机交互反馈获得的期望中间通过点坐标 $\bar{\mu}_m$,自适应精度 $\bar{\Sigma}_m$;

输出:受未知时间输入的中间通过点 $\{\bar{\mu}_m, \bar{\Sigma}_m\}$ 约束的KMP自适应避障轨迹 $\{t_n, s(t_n)\}_{n=1}^N$;

1) 根据式(15)和 $\{\hat{\mu}_n\}_{n=1}^N$ 构建参考直线 $\{\hat{\mu}'_n\}_{n=1}^N$;

2) for $n = 1$ to N

计算 $\hat{\mu}'_n$ 与 $\bar{\mu}_m$ 的欧式距离 $d(\bar{\mu}_m, \hat{\mu}'_n)$;

end for

return $\hat{\mu}'_r, r = \operatorname{argmin}_n d(\bar{\mu}_m, \hat{\mu}'_n), n \in \{1, 2, \dots, N\}$

3) 根据式(16)将当前 D 中的样本点 $\{t_r, \hat{\mu}'_r, \hat{\Sigma}'_r\}$ 替换成 $\{t_r, \bar{\mu}_m, \bar{\Sigma}_m\}$;

4) 根据式(7)~(10)分别计算 k^* 、 $K\mu$ 和 Σ ;

5) 根据式(5)计算 KMP 自适应避障轨迹的位置期望 $\{s(t_n)\}_{n=1}^N$;

6) return $\{t_n, s(t_n)\}_{n=1}^N$

可以看出,算法 2 使用基于式(15)和(16)的更新策略仅需用户为机器人指定期望中间通过点坐标 $\bar{\mu}_m$ 和相应的自适应精度矩阵 $\bar{\Sigma}_m$ 便可完成对当前参考数据库 D 的更新,解决了使用基于式(13)和(14)的更新策略时难以预设时间参数的问题,很适合非专业的用户使用。此外,与基于式(13)和(14)的更新策略只能在预设时刻使用不同,基于式(15)和(16)的更新策略对人机交互的发生时刻没有限制,这使得它很适合本文研究的基于任意时刻人机交互反馈的机器人实时避障问题。

3.3 基于人机交互反馈的参考数据库更新策略适用条件

基于式(15)和(16)的更新策略在更新表征复杂轨迹的参考数据库时存在失效的可能,即无法通过搜索最小欧氏距离从参考直线 $\{\hat{\mu}'_n\}_{n=1}^N$ 中正确找到 $\bar{\mu}_m$ 对应样本点,其中一种情况是当 $\{\hat{\mu}'_n\}_{n=1}^N$ 中有样本点在参考直线上的垂足不在起点与终点之间时。为防止更新失效,当前参考数据库 D 应该满足如下适用条件:

1) 给定参考数据库 $D = \{t_n, \hat{\mu}'_n, \hat{\Sigma}'_n\}_{n=1}^N$, 对于任意点坐标 $\hat{\mu}'_i \in \{\hat{\mu}'_n\}_{n=2}^{N-1}, \forall i \in \{2, 3, \dots, N-1\}$, 须满足以下条件:

$$\|\hat{\mu}'_i - \hat{\mu}'_1\|_2^2 < \|\hat{\mu}'_i - \hat{\mu}'_N\|_2^2 + \|\hat{\mu}'_1 - \hat{\mu}'_N\|_2^2$$

$$\text{且 } \|\hat{\mu}'_i - \hat{\mu}'_N\|_2^2 < \|\hat{\mu}'_i - \hat{\mu}'_1\|_2^2 + \|\hat{\mu}'_1 - \hat{\mu}'_N\|_2^2 \quad (17)$$

其中, $\|\cdot\|_2^2$ 表示 2-范数的平方。

另一种失效的情况是当 $\{\hat{\mu}'_n\}_{n=1}^N$ 中存在多个样本点在参考直线上的垂足相同。对此,同样给出相应的适用条件,如 2) 所示:

2) 给定参考数据库 $D_r = \{t_n, \hat{\mu}'_n, \hat{\Sigma}'_n\}_{n=1}^N$, 对于任意点坐标 $\hat{\mu}'_i \in \{\hat{\mu}'_n\}_{n=2}^{N-1}, \forall i \in \{2, 3, \dots, N-1\}$, 须满足以下条件:

$$d(\hat{\mu}'_i, \hat{\mu}'_1) < d(\hat{\mu}'_{i+1}, \hat{\mu}'_1) \quad (18)$$

其中, $\hat{\mu}'_i = (x''_i, y''_i, z''_i)$ 是 $\hat{\mu}'_i = (x_i, y_i, z_i)$ 在参考直线上的真实垂足坐标,即:

$$\begin{cases} x''_i = x_1 + k_i(x_N - x_1) \\ y''_i = y_1 + k_i(y_N - y_1) \\ z''_i = z_1 + k_i(z_N - z_1) \end{cases} \quad (19)$$

式中: (x_1, y_1, z_1) 与 (x_N, y_N, z_N) 分别为参考直线起点坐标 $\hat{\mu}'_1$ 与终点坐标 $\hat{\mu}'_N$ (注: $\hat{\mu}'_1 = \hat{\mu}'_1$ 且 $\hat{\mu}'_N = \hat{\mu}'_N$), 且 $k_i =$

$$\frac{(x_1 - x_i)(x_N - x_1) + (y_1 - y_i)(y_N - y_1) + (z_1 - z_i)(z_N - z_1)}{(x_N - x_1)^2 + (y_N - y_1)^2 + (z_N - z_1)^2} \quad (20)$$

因此,在本文提出的机器人避障交互式学习框架中,为防止避障轨迹规划失效,只有当轨迹同时满足式(17)和(18)所述的适用条件时才可以根据式(15)和(16)用中间通过点更新对应的参考数据库。

4 实验分析与讨论

本节首先通过仿真实验分别验证两个更新策略适用条件(式(17)和(18))的有效性;接下来通过两个不同的搬运实验来检验所提出的机器人避障交互式学习方法;最后,对实验结果进行了分析与讨论。

4.1 仿真实验

1) 仿真避障实例 1

图 4(a) 显示了一条与固定障碍冲突的初始轨迹,设其参考数据库为 $D = \{t_n, \hat{\mu}'_n, \hat{\Sigma}'_n\}_{n=1}^N$ 。按照图 3(c) 展示的方式根据预测的碰撞点 C 以及参考直线 AB 选择期望中间通过点 M 的坐标 $\bar{\mu}_m$, 然而使用算法 2 生成的 KMP 泛化轨迹却失去了终点的约束(用“+”标记),如图 4(a) 所示。

由图 4(b) 所示几何关系可知,虽然向量 \overrightarrow{MC} 与 \overrightarrow{AB} 垂直,然而由于 $\angle MBA > 90^\circ$, 这使 M 点在 \overrightarrow{AB} 的投影点并不在 AB 之间,因此根据式(16)通过搜索最小欧式距离 $\min d(\bar{\mu}_m, \hat{\mu}'_n)$ 从 $\{\hat{\mu}'_n\}_{n=1}^N$ 中得到 M 的对应点是终点 B , 即原轨迹终点 $\{t_N, \hat{\mu}'_N, \hat{\Sigma}'_N\}$ 被更新成了 $\{t_N, \hat{\mu}'_m, \hat{\Sigma}'_m\}$, 最终导致泛化轨迹失效。

由上述分析可知,当 $\angle MAB > 90^\circ$ 或 $\angle MBA > 90^\circ$, 应用式(16)规划的避障轨迹将会失效。为了避免这种情况发生,应满足 $\angle MAB < 90^\circ$ 且 $\angle MBA < 90^\circ$; 由于 $\overrightarrow{MC} \perp \overrightarrow{AB}$, 可得 $\angle CAB < 90^\circ$ 且 $\angle CBA < 90^\circ$ 。因此,由锐角三角形的三边关系可得 $\triangle CAB$ 应满足以下必要条件:

$$|\overrightarrow{AC}|^2 < |\overrightarrow{BC}|^2 + |\overrightarrow{BA}|^2 \text{ 且 } |\overrightarrow{BC}|^2 < |\overrightarrow{AC}|^2 + |\overrightarrow{BA}|^2$$

即应满足式(17)表述的适用条件。

为了解决轨迹避障失效的问题,采用分割泛化的策略来应对。如图 4(c) 所示,标记为 I 和 II 的两段曲线分别代表分割后由 KMP 生成的子轨迹 1 和 2,经计算二者均满足式(17)和(18)。由于与障碍干涉的是子轨迹 2,因此根据式(15)和(16)更新后者的参考数据库,其中期望的中间通过点参照预期碰撞点以及子轨迹 2 的参考直线选取。如图 4(d) 所示,泛化子轨迹 2 成功避开了障碍。

激光跟踪仪(型号:AT960-MR)跟踪固定在示教者右手手套上的反射球,并以 100 Hz 的采样频率记录其轨迹,如图 6 所示;在轨迹规划阶段,用户指导 KUKA 机器人(型号:IIWA_14_R820)避障的做法是直接将其末端牵引到期望的通过点位置,而机器人在检测到超过阈值的外力时会立刻停止执行当前任务,随后自动记录该期望点坐标并将其发送给上位机;由 MATLAB 实时计算生成受其约束的避障轨迹再发送给机器人,后者按最新收到的轨迹继续执行当前任务。

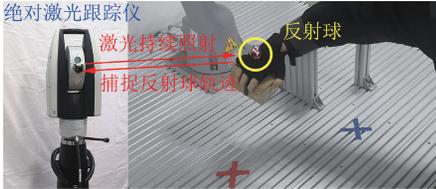


图 6 示教轨迹数据采集

Fig. 6 Demonstrated trajectory data acquisition

示教的工作场景如图 7(a) 所示。在这个场景中,用户从起点提起一个盒子,然后依次避开障碍 1 和障碍 2,最后将其放在终点。机器人需要学习人类的动作,并在泛化场景中执行搬运任务。请注意泛化场景调整了盒子的初始与终点位置,还分别增加了障碍 1 的宽度与障碍 2 的高度,如图 7(b) 所示。

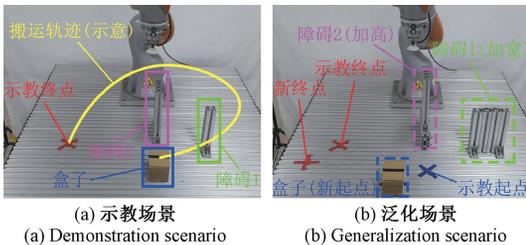


图 7 第 4.2 节示教场景与泛化场景的对比

Fig. 7 Comparison between the demonstration scenario and the generalization one in section 4.2

重复执行示教动作 5 次后,使用算法 1 对所有示教轨迹进行分割,其中误差阈值 $d_{th} = 60 \text{ mm}$ 。图 8(a) 显示了算法 1 生成的 4 处分割点位置;图 8(b) 显示了用户最终决定的分割方案,即保留了图 8(a) 中第 2 处分割点位置以应对障碍 1 的尺寸变化,因此示教动作被分为 2 个步骤。经检验所有子轨迹均满足式 (17) 和 (18)。接下来分别将每个步骤的子轨迹各自对齐,并使用 GMM-GMR 获得相应的参考轨迹(即复现轨迹),其中 $K=6$ 。如图 9 的点线所示,该轨迹由复现子轨迹 1 和 2 组成,分别对应避开障碍 1 的步骤 1 和避开障碍 2 的步骤 2。

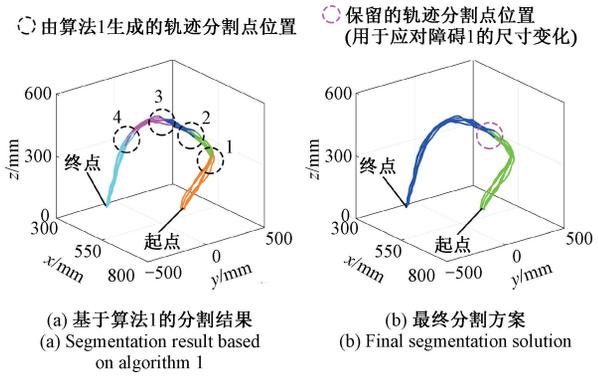


图 8 第 4.2 节轨迹分割方案

Fig. 8 Solution of trajectory segmentation in section 4.2

图 9 中的虚线表示由 KMP 生成的初始轨迹,其起点、终点以及两条子轨迹之间的分割点位置都做了调整,以分别适应泛化场景中盒子放置的起始、目标位置以及加宽后的障碍 1,相关的 KMP 模型参数设置为 $l = 0.3$, $\lambda_1 = 0.001$ 以及 $\Sigma_m = 1 \times 10^{-10} \mathbf{I}_3$ 。整个搬运任务的执行过程如图 10 所示。为了避免按照该轨迹移动的机器人与障碍 2 发生碰撞,用户手动牵引机器人末端到了新位置(图 10-F 中用圆圈标记),这个位置被机器人记录为期望中间通过点的坐标;当机器人根据受该点约束的避障轨迹提着盒子越过障碍 2 后,另一个用户突然将后者朝终点方向拖动,使其重新与当前盒子的轨迹相冲突,如图 10-G ~ H 所示。于是监督机器人的用户在盒子与障碍 2 碰撞前再次将末端牵引到新的期望通过点位置(图 10-J 中用圆圈标记)。机器人根据获得的最新避障轨迹提着盒子再次成功避开了障碍 2,随后将其放置在新终点处。

- i 复现子轨迹1(步骤1)
- ii 复现子轨迹2(步骤2)
- I 初始子轨迹1(步骤1)
- II 初始子轨迹2(步骤2)
- * 起点
- + 终点
- A 期望的中间通过点1
- B 期望的中间通过点2
- 泛化子轨迹2(由点1和2约束的避障轨迹)
- 分割点

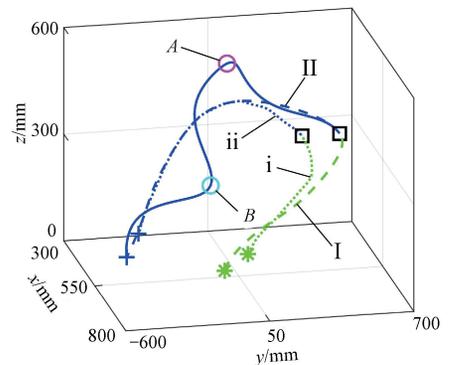


图 9 针对已知障碍(静止和突然移动)的自适应避障轨迹
Fig. 9 Adaptive trajectory for known obstacles (static and suddenly moving)

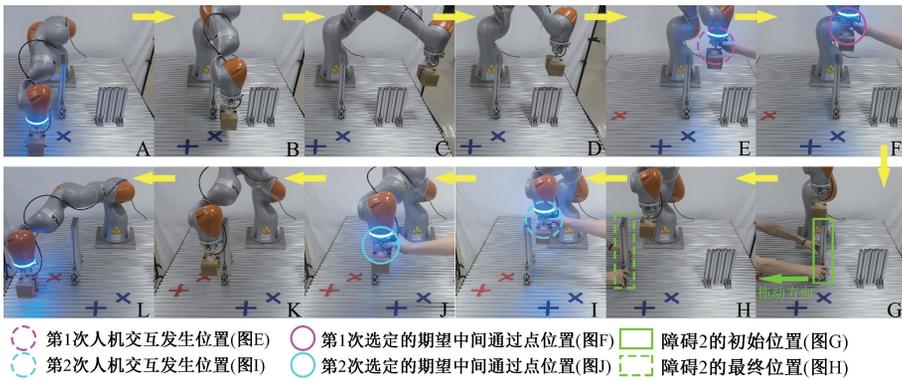


图 10 KUKA 协作机器人在用户的指导下避开不同的已知障碍的过程

Fig. 10 Process of KUKA cobot avoiding different known obstacles with the user's instruction

实验中基于人机交互反馈生成的自适应避障轨迹如图 9 中的实曲线所示。可以看出, KMP 模型成功地根据泛化任务的初始轨迹和两次交互反馈的期望中间通过点坐标对步骤 2 的子轨迹进行了调整。

4.3 应对未知障碍的真实实验

本节通过另一个搬运实验来进一步验证所提出方法的有效性,其教场景设置如图 11(a)所示。整个搬运过程分为 3 个步骤:1) 首先提起放置在起点的盒子 1, 将其移动到高台上并放置于“x”标记处;2) 接下来, 移动到盒子 2 上方并将其提起;3) 最后将盒子 2 放置于终点处。作为对比, 在泛化场景中盒子 1 的初始位置、高台位置、盒子 2 初始位置以及终点都被更改了。此外, 起点与高台之间新增了一个障碍 1。图 11(b) 显示了泛化场景的设置, 请注意其与图 11(a) 的区别。

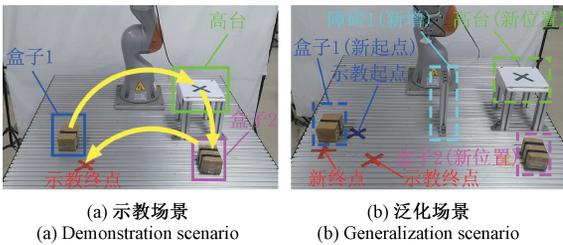


图 11 第 4.3 节示教场景与泛化场景的对比

Fig. 11 Comparison between the demonstration scenario and the generalization one in section 4.3

在示教阶段, 用户重复执行了搬运任务 5 次, 每次的动作顺序如图 11(a) 中箭头所示。设置误差阈值 $d_{th} = 60 \text{ mm}$, 算法 1 生成的 5 处分割点位置如图 12(a) 所示; 人工选择保留第 2、4 处分割点, 分别应对高台与盒子 2 位置的变化, 则整个示教动作被分为 3 个步骤, 如图 12(b) 所示, 其中所有子轨迹经检验均满足式 (17) 和 (18)。接下来, 由 GMM-GMR 从对齐后的示教轨迹中获得相应的参考轨迹 (即复现轨迹), 其中 $K=6$ 。如图 13

中的点线所示, 该轨迹由 3 段复现子轨迹组成, 分别对应步骤 1~3。

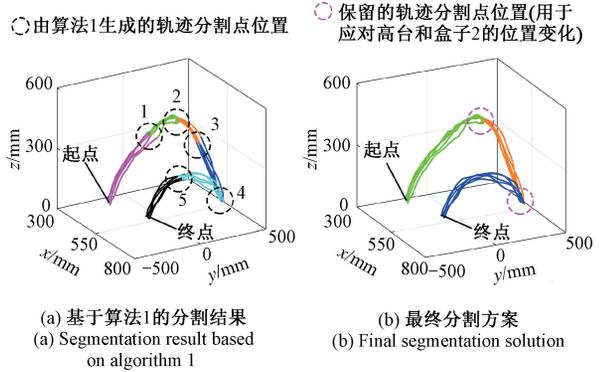


图 12 第 4.3 节轨迹分割方案

Fig. 12 Solution of trajectory segmentation in section 4.3

- i 复现子轨迹1(步骤1) * 起点
- ii 复现子轨迹2(步骤2) + 终点
- iii 复现子轨迹3(步骤3) □ 分割点
- I 初始子轨迹1(步骤1) A 期望的中间通过点1
- II 初始子轨迹2(步骤2) B 期望的中间通过点2
- III 初始子轨迹3(步骤3) C 期望的中间通过点3
- ① 泛化子轨迹1(第1次交互生成, 由中间通过点1约束)
- ② 泛化子轨迹3(第2次交互生成, 由中间通过点2约束)
- ③ 泛化子轨迹3(第3次交互生成, 由中间通过点3约束)

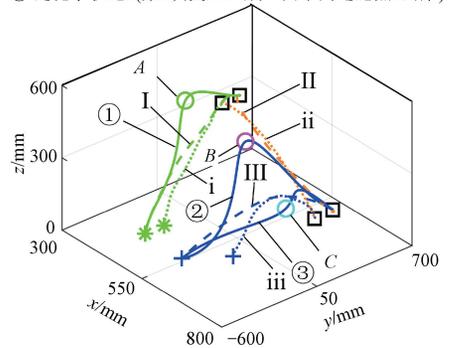


图 13 针对未知障碍 (静止和突然出现) 的自适应避障轨迹
Fig. 13 Adaptive trajectory for unknown obstacles (static and sudden)

图 13 中的虚线表示初始轨迹,即根据泛化场景的设置调整了每个步骤起点和终点的 KMP 泛化轨

迹,相关参数设置与第 4.2 节相同。整个搬运过程如图 14 所示。

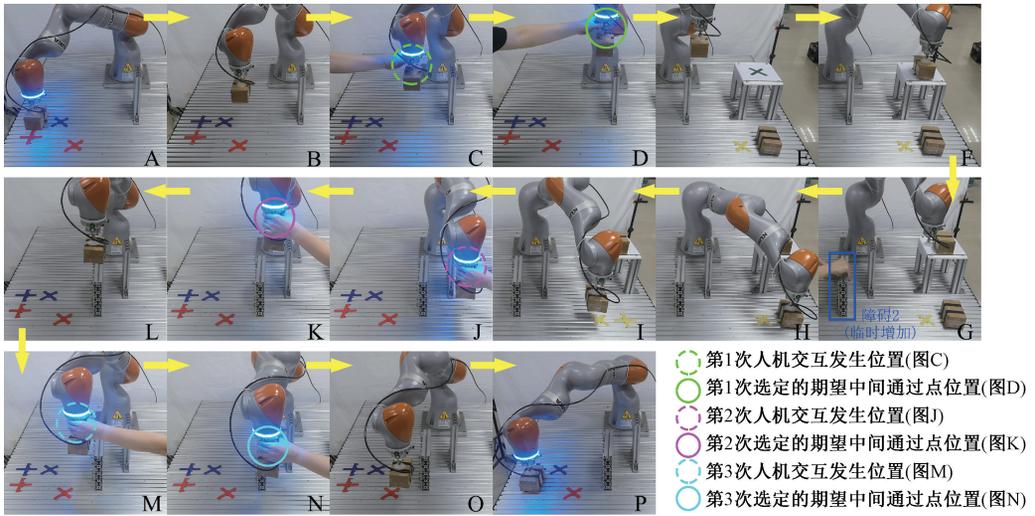


图 14 KUKA 协作机器人在用户的指导下避开不同的未知障碍的过程

Fig. 14 Process of KUKA cobot avoiding different unknown obstacles with the user's instruction

当机器人在泛化场景中提起盒子 1 后,沿着用户通过人机交互实时指定的中间通过点位置(图 14-D 中用圆圈标记)约束的避障轨迹越过了障碍 1,将其放置在了高台上的新终点位置;接下来,机器人朝放置在新位置的盒子 2 移动。然而此时另一用户突然在盒子 2 与新终点之间随机设置了一个未知尺寸的障碍 2,盒子 2 的移动路径将与其冲突,如图 14-G 所示。

为了应对突然出现的障碍 2,负责监督机器人的用户先后测试了两种不同的避障方案:1) 首先,机器人末端被手动牵引到一个期望的中间通过点位置(图 14-K 中用圆圈标记),并以此规划避障轨迹;2) 当机器人按照实时调整后的避障轨迹提着盒子 2 越过障碍物 2 的顶部后,机器人再次被手动牵引到障碍 2 的另一侧并重新记录了一个期望的中间通过点位置(图 14-N 中用圆圈标记),而之前设置的中间通过点被删除。机器人按照最新期望通过点约束的避障轨迹再次越过障碍 2,将盒子 2 放置在新的终点位置。

KMP 生成的自适应避障轨迹如图 13 的实曲线所示。可以看出,经过步骤 1 的交互,机器人根据用户指定的中间通过点坐标调整了子轨迹 1,越过了障碍 1;经过步骤 3 的两次交互,机器人先后从不同位置避开了障碍 2。

4.4 分析与讨论

通过第 4.1 节的仿真实验,不仅验证了所提出的基于式(17)和(18)的更新策略适用条件的有效性,还证明了用分割泛化策略处理不满足该条件的轨迹的可行性。此外,通过第 4.2~4.3 节的真实实验可以看出分割泛化

策略在轨迹规划中还起到了 2 个作用:1) 满足各步骤的目标泛化需求,如针对各盒子的起点与终点位置调整、第 4.2 节中步骤 1 的终点与步骤 2 的起点调整(应对障碍 1 尺寸变化)等;2) 针对实时避障需求的分段泛化响应,即人机交互避障是在不同的步骤内执行的,而不是泛化整条轨迹。更重要的是,这种分段调整的方式可以使各步骤的执行互不影响,所表现出的灵活性是对整体运动轨迹一次性建模的方法难以达到的。

算法 1 在第 4.2~4.3 节的真实实验中有效地执行了对示教轨迹的分割任务。作为对比,使用传统的 HDP-HMM 算法分别对 2 个实验的示教轨迹进行处理,设置吉布斯采样器迭代 1 500 次。所得的分割结果如图 15 所示。对比图 15(a)和图 8(a),可以看出 HDP-HMM 和算法 1 分别得到的分段数量均为 5;对比图 15(b)和

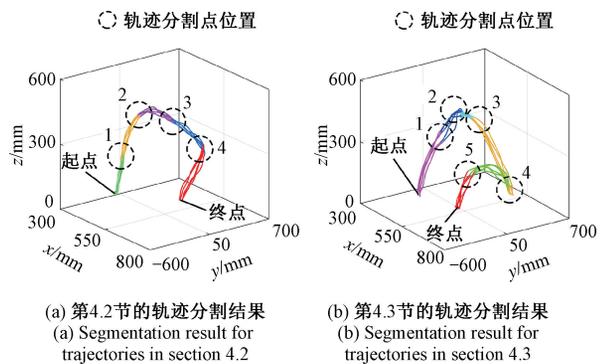


图 15 基于 HDP-HMM 分割示教轨迹的结果

Fig. 15 Segmentation results for the demonstrated trajectories based on HDP-HMM

12(a),可以看出二者分别得到的分段数量均为6。表1统计了两种算法各自处理两个实验的示教轨迹所消耗的时间,可以看出本文提出的算法1(基于DP算法的轨迹分割策略)在计算效率方面优于HDP-HMM算法。

表1 分割所用时间对比

Table 1 Comparison of time-consuming in segmentation

算法名称	第4.2节实验轨迹	第4.3节实验轨迹	迭代次数
	处理时间/s	处理时间/s	
HDP-HMM	59.07	72.8	100
	362.6	477.9	500
	789.3	1 119.6	1 000
	1 295.6	1 761.6	1 500
算法1	0.084	0.107	无

需要注意的是,最终保留的轨迹分割点除了要满足前述各步骤的目标泛化需求外,还需要各子轨迹全部满足式(17)和(18),否则需要调整误差阈值 d_{th} 再执行算法1重新分割示教轨迹,如图1学习阶段所示。不过,多次调整误差阈值 d_{th} 进行分割计算并不会显著增加用户的时间成本。图16展示了一个应用算法1的仿真实例,分别设置 d_{th} 为8、10、15和20 mm处理同一条轨迹(由2 561个样本点组成),相应的分割结果分别如图16所示。表2统计了算法1对图16所示轨迹进行分割的结果和相应求解时间。可以看出,在本例中分段数量 C 随着 d_{th} 取值的增加而减少,而不同取值的 d_{th} 对应的求解时间均在0.1 s以内。由于算法1的执行效率高,因此在实际操作中,用户可以按照具体需求通过调整 d_{th} 的取值来选择合适的分割方案。

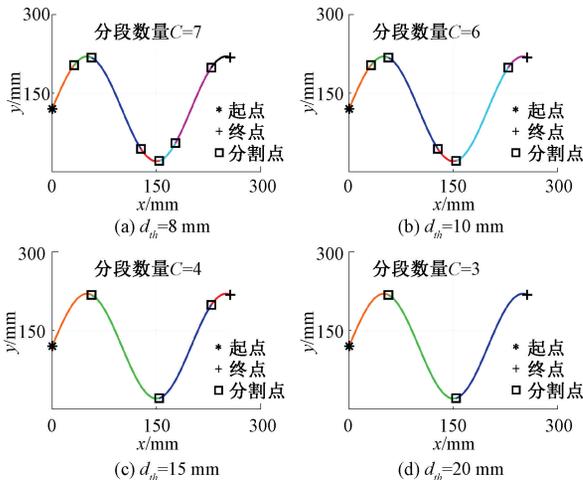


图16 不同 d_{th} 取值下基于DP算法的轨迹分割策略

Fig. 16 Example illustration of the trajectory segmentation strategy based on DP algorithm with different values of d_{th}

表2 算法1的分割结果统计

Table 2 Segmentation result statistics for algorithm 1

轨迹	d_{th}/mm	分段数量 C	求解时间/s
图16所示	8	7	0.069
	10	6	0.061
	15	4	0.057
	20	3	0.052

在所有步骤子轨迹均满足式(17)和(18)的情况下,使用本文的方法(算法2)在真实实验中通过人机交互实时规划避障轨迹成功避开了各种障碍,例如静止的已知障碍(第4.2节实验步骤1),突然移动的已知障碍(第4.2节实验步骤2),静止的未知障碍(第4.3节实验步骤1)以及突然出现的未知障碍(第4.3节实验步骤3)等。此外,用户还可以在当前任务中通过人机交互指导机器人针对某个障碍反复测试不同的避障轨迹,直到找到满意的解决方案,如第4.3节实验步骤3(图14-J~P)所示。作为对比,分别用文献[9,18]的方法结合第4.2节和第4.3节实验记录的交互发生时刻与中间通过点等数据,设置式(14)中的 $\delta_i=1\text{ s}$ (与文献[9]设置相同),复现各步骤的避障轨迹,先后共5次人机交互避障成功率如表3所示。可以看出,与其他人机交互避障方法相比,本文的方法有效提升了人机交互避障成功率。

表3 不同方法的避障成功率对比

Table 3 Comparison of success rate in obstacle avoidance of different approaches

方法	第4.2节实验		第4.3节实验			避障成功率/%
	步骤2	步骤2	步骤1	步骤3	步骤3	
	第1次交互	第2次交互	第1次交互	第2次交互	第3次交互	
本文	成功	成功	成功	成功	成功	100
文献[9]	失败	成功	成功	失败	失败	40
文献[18]	失败	成功	成功	失败	失败	40

5 结论

本文提出了一种基于DP-KMP的交互式学习方法,用于解决机器人在执行复杂任务时的避障问题。为了满足复杂多步骤任务的目标泛化及分段避障的需求,基于分割泛化的思路构建了该方法的整体框架。在学习阶段,提出了基于DP算法的轨迹分割策略以解决传统轨迹分割方法计算效率低的问题,并用GMM-GMR提取各

步骤相应的参考数据库;在轨迹规划阶段使用 KMP 模型规划目标泛化及实时避障轨迹,并引入基于人机交互反馈的参考数据库更新策略,解决了使用 KMP 结合人机交互反馈的期望中间通过点规划避障轨迹时难以确定该点的时间输入以及无法满足用户在任意时刻与机器人交互的需求等问题,进一步提升了人机交互避障的成功率;为了避免出现避障轨迹规划失效的情况,提出了该更新策略相应的适用条件,用于在学习阶段检验示教轨迹的分割方案,后者需满足该适用条件方可进行下一步的 GMM-GMR 学习。

通过真实机器人与仿真实验可以得到以下结论:

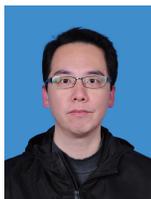
1) 本文方法采用的分割泛化策略使机器人不仅可以应对多个步骤的目标泛化需求,还可以在用户的指导下针对不同位置的障碍实时调整相应步骤的子轨迹,有效保证了任务的顺利完成;2) 所提出的基于 DP 算法的轨迹分割策略在分割两个实验的示教轨迹分别仅用时 0.084 和 0.107 s,在计算效率方面优于传统的 HDP-HMM 算法;3) 在满足所提出的基于人机交互反馈的参考数据库更新策略适用条件的情况下,使用本文的方法可以让运动中的机器人在用户的帮助下及时避开已知和未知的障碍(包括各种静止和突然变化的障碍),从而有效保证机器人的安全。

参考文献

- [1] 刘宜成, 杨迦凌, 梁斌, 等. 基于强化学习的多段连续体机器人轨迹规划[J]. 电子测量技术, 2024, 47(5): 61-69.
LIU Y CH, YANG J L, LIANG B, et al. Trajectory planning of multi-stage continuum robot based on reinforcement learning[J]. Electronic Measurement Technology, 2024, 47(5): 61-69.
- [2] 张仪, 冯伟, 王卫军, 等. 融合 LSTM 和 PPO 算法的移动机器人视觉导航[J]. 电子测量与仪器学报, 2022, 36(8): 132-140.
ZHANG Y, FENG W, WANG W J, et al. Visual navigation of mobile robots based on LSTM and PPO algorithms[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 132-140.
- [3] 黄艳龙, 徐德, 谭民. 机器人运动轨迹的模仿学习综述[J]. 自动化学报, 2022, 48(2): 315-334.
HUANG Y L, XU D, TAN M. On imitation learning of robot movement trajectories: A survey[J]. Acta Automatica Sinica, 2022, 48(2): 315-334.
- [4] 迟明善, 姚玉峰, 刘亚欣. 基于示教编程的共融机器人技能学习方法的研究进展[J]. 仪器仪表学报, 2020, 41(1): 71-83.
CHI M SH, YAO Y F, LIU Y X. Advances in coexisting-cooperative-cognitive robots skill learning approach based on learning from demonstration[J]. Chinese Journal of Scientific Instrument, 2020, 41(1): 71-83.
- [5] RAVICHANDAR H, POLYDOROS A S, CHERNOVA S, et al. Recent advances in robot learning from demonstration[J]. Annual Review of Control, Robotics, and Autonomous Systems, 2020, 3: 297-330.
- [6] IJSPEERT A J, NAKANISHI J, HOFFMANN H, et al. Dynamical movement primitives: Learning attractor models for motor behaviors[J]. Neural Computation, 2013, 25(2): 328-373.
- [7] CALINON S, BRUNO D, CALDWELL D G. A task-parameterized probabilistic model with minimal intervention control[C]. 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014. 3339-3344.
- [8] PARASCHOS A, DANIEL C, PETERS J, et al. Probabilistic movement primitives[C]. Advances in Neural Information Processing Systems (NIPS), 2013, 26(4): 2618-2626.
- [9] HUANG Y L, ROZO L, SILVÉRIO J, et al. Kernelized movement primitives[J]. International Journal of Robotics Research, 2019, 38(7): 833-852.
- [10] HUANG Y L, SILVÉRIO J, ROZO L, et al. Generalized task-parameterized skill learning[C]. 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, 5667-5674.
- [11] LU ZH Y, WANG N, YANG CH G. A constrained DMPs framework for robot skills learning and generalization from human demonstrations [J]. IEEE/ASME Transactions on Mechatronics, 2021, 26(6): 3265-3275.
- [12] ZHAI D H, XIA ZH Q, WU H C, et al. A motion planning method for robots based on DMPs and modified obstacle-avoiding algorithm [J]. IEEE Transactions on Automation Science and Engineering, 2023, 20(4): 2678-2688.
- [13] FRANK F, PARASCHOS A, VAN DER SMAGT P, et al. Constrained probabilistic movement primitives for robot trajectory adaptation[J]. IEEE Transactions on Robotics, 2022, 38(4): 2276-2294.
- [14] DAVOODI M, IQBAL A, CLOUD J M, et al. Rule-based safe probabilistic movement primitive control via control barrier functions[J]. IEEE Transactions on Automation Science and Engineering, 2023, 20(3): 1500-1514.
- [15] LU ZH Y, WANG N, LI M, et al. Incremental motor skill learning and generalization from human dynamic

- reactions based on dynamic movement primitives and fuzzy logic system[J]. *IEEE Transactions on Fuzzy Systems*, 2022, 30(6): 1506-1515.
- [16] ZHANG Y, CHENG L, CAO R, et al. A neural network based framework for variable impedance skills learning from demonstrations[J]. *Robotics and Autonomous Systems*, 2023, 160, 104312.
- [17] FU J Y, JIN ZH H, LIU AN D, et al. Non-parametric Gaussian process movement primitive with via-point constraint for effective and safe robot skill learning[J]. *Neurocomputing*, 2024, 589, 127711.
- [18] SUN D, LIAO Q F, LOUFI A. Type-2 fuzzy model-based movement primitives for imitation learning[J]. *IEEE Transactions on Robotics*, 2022, 38(4): 2462-2480.
- [19] XIAO S, CHEN X Y, LU Y K, et al. A KMP-based interactive learning approach for robot trajectory adaptation with obstacle avoidance[J]. *Industrial Robot*, 2024, 51(2): 326-339.
- [20] 曾超, 杨辰光, 李强, 等. 人-机器人技能传递研究进展[J]. *自动化学报*, 2019, 45(10): 1813-1828.
ZENG CH, YANG CH G, LI Q, et al. Research progress on human-robot skill transfer[J]. *Acta Automatica Sinica*, 2019, 45(10): 1813-1828.
- [21] YANG CH G, ZENG CH, CONG Y, et al. A learning framework of adaptive manipulative skills from human to robot[J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(2): 1153-1161.
- [22] NAGANO M, NAKAMURA T, NAGAI T, et al. Sequence pattern extraction by segmenting time series data using GP-HSMM with hierarchical dirichlet process[C]. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018: 4067-4074.
- [23] MUHLIG M, GIENGER M, HELLBACH S, et al. Task-level imitation learning using variance-based movement optimization[C]. *IEEE International Conference on Robotics and Automation*, 2009. 1635-1642.
- [24] DOUGLAS D H, PEUCKER T K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature[J]. *Canadian Cartographer*, 2006, 10(2): 112-122.
- [25] LI X CH, GAO X F, ZHANG W, et al. Smooth and collision-free trajectory generation in cluttered environments using cubic B-spline form[J]. *Mechanism and Machine Theory*, 2022, 169, 104606.
- [26] WU Y Q, ZHAO F, TAO T, et al. A framework for autonomous impedance regulation of robots based on imitation learning and optimal control[J]. *IEEE Robotics and Automation Letters*, 2021, 6(1): 127-134.
- [27] 刘成菊, 林立民, 刘明, 等. 一种基于 DTW-GMM 的机器人多机械臂多任务协同策略[J]. *自动化学报*, 2022, 48(9): 2187-2197.
LIU CH J, LIN L M, LIU M, et al. A multi-task collaborative strategy for multi-arm robot based on DTW-GMM[J]. *Acta Automatica Sinica*, 2022, 48(9): 2187-2197.
- [28] 吴晓敏, 贺苗, 刘瞰东, 等. 基于狄利克雷过程聚类的机器人演示学习研究[J]. *仪器仪表学报*, 2023, 44(1): 265-274.
WU X M, HE M, LIU T D, et al. Study of robot demonstration learning based on the Dirichlet process clustering[J]. *Chinese Journal of Scientific Instrument*, 2023, 44(1): 265-274.

作者简介



肖洒, 2008 年于中南大学获得学士学位, 2017 年于福州大学获得硕士学位, 现为福州大学博士研究生, 主要研究方向为机器人模仿学习以及人-机器人交互技术。

E-mail: xiaosa1014@126.com

Xiao Sa received his bachelor's degree from Central South University in 2008, and master's degree from Fuzhou University in 2017. He is currently a Ph. D. candidate at Fuzhou University. His main research interests include robot imitation learning and human-robot interaction technology.



吴海彬(通信作者), 2002 年于浙江大学获得博士学位, 现为福州大学教授, 博士生导师, 主要研究方向为机器人控制、机器人触觉传感器等。

E-mail: wuhb@fzu.edu.cn

Wu Haibin (Corresponding author) received his Ph. D. degree from Zhejiang University in 2002. He is currently a professor and a Ph. D. advisor at Fuzhou University. His main research interests include robot control, robotic tactile sensor and so on.