

DOI: 10.11992/tis.201807018

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20190520.1428.010.html>

基于动态系统的机器人模仿学习方法研究

于建均^{1,2}, 姚红柯^{1,2}, 左国玉^{1,2}, 阮晓钢^{1,2}, 安硕^{1,2}

(1. 北京工业大学 信息学部, 北京 100124; 2. 北京工业大学 计算智能与智能系统北京市重点实验室, 北京 100124)

摘要: 针对当前机器人模仿学习过程中, 运动模仿存在无法收敛到目标点以及泛化能力差的问题, 引入一种基于动态系统 (dynamical system, DS) 的模仿学习方法。该方法通过高斯混合模型 (gaussian mixture model, GMM) 将示教运动数据建模为一非线性动态系统; 将 DS 全局稳定的充分条件作为约束, 以保证 DS 所生成的所有轨迹收敛到目标点; 将动态系统模型的参数学习问题转化为求解一个约束优化问题, 从而得到模型参数。以 7bot 机械臂为实验对象, 进行仿真实验和机器人实验, 实验结果表明: 该方法学习的 DS 模型从不同起点生成的所有轨迹都收敛到目标点, 轨迹平滑, 泛化能力好。

关键词: 机器人; 模仿学习; 轨迹层面; 高斯混合模型; 动态系统; 参数学习; 7bot 机械臂; 泛化能力

中图分类号: TP 242.6 **文献标志码:** A **文章编号:** 1673-4785(2019)05-1026-09

中文引用格式: 于建均, 姚红柯, 左国玉, 等. 基于动态系统的机器人模仿学习方法研究 [J]. 智能系统学报, 2019, 14(5): 1026-1034.

英文引用格式: YU Jianjun, YAO Hongke, ZUO Guoyu, et al. Research on robot imitation learning method based on dynamical system[J]. CAAI transactions on intelligent systems, 2019, 14(5): 1026-1034.

Research on robot imitation learning method based on dynamical system

YU Jianjun^{1,2}, YAO Hongke^{1,2}, ZUO Guoyu^{1,2}, RUAN Xiaogang^{1,2}, AN Shuo^{1,2}

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China)

Abstract: In the current robot imitation learning process, the motion imitation cannot converge to the target point, and the generalization ability is poor. To solve this problem, an imitation learning method based on dynamical system (DS) is introduced. First, the demonstration motion data is modeled as a nonlinear DS by Gaussian mixture model (GMM). Second, the sufficient condition of DS global stability is used as a constraint to ensure that all the DS-generated trajectories converge to the target. Finally, the parameter learning problem of the DS model is transformed into seeking for a solution to a constrained optimization problem to obtain the model parameters. Simulation experiments and robot experiments were carried out using the 7bot manipulator. The experimental results show that all the trajectories generated by the DS model from different starting points converged to the target point, and the trajectory was smooth and the generalization performance was improved.

Keywords: robot; imitation learning; trajectory level; Gaussian mixture model; dynamical system; parameter learning; 7bot manipulator; generalization performance

目前, 在机器人技术的研究中, 如何使机器人具备类似于人类的智能行为已成为一大研究热点^[1]。

而从人类学习中发展而来的模仿学习, 作为机器人直接获取知识和技能的一种方式, 在机器人的智能性的提升方面越来越发挥出巨大的作用, 越来越多地受到学术界的关注和研究^[2-6]。一方面, 相对于传统的机器人编程控制, 模仿学习将使机器人编程更加容易, 提高了机器人的作业效率;

收稿日期: 2018-07-18. 网络出版日期: 2019-05-21.

基金项目: 国家自然科学基金项目 (61773027); 北京市自然科学基金项目 (4182008); 北京市自然科学基金项目/北京市教育委员会科技计划重点项目 (KZ201610005010).

通信作者: 左国玉. E-mail: zuoguoyu@bjut.edu.cn.

另一方面,模仿学习赋予机器人获取知识、学习知识的能力,使其更加具备智能性^[7]。

机器人模仿学习一般分为3个步骤完成:示教过程、学习过程、再现过程。当前,模仿学习的研究,主要集中在对学习过程的研究,已经出现了多种模仿学习的算法用来实现机器人的学习过程^[8-11]。其中基于轨迹匹配的机器人模仿学习算法主要有高斯混合模型(GMM)^[12]和高斯混合回归GMR(gaussian mixture regression)^[12-13]、局部加权回归LWR(locally weighted regression)^[14-15]、人工神经网络ANN(artificial neural networks)^[16-17]等,这些算法的特点是将示教运动轨迹数据建模为回归过程,以期获得控制策略指导机器人对示教运动进行再现。虽然基于轨迹匹配的回归算法具有较强的数据表征能力和抗噪能力,但是也存在泛化能力差、稳定性差等缺点^[18]。Ijspeert等^[19]提出的一种基于动态运动基元DMP(dynamic movement primitives)的机器人模仿学习方法。该方法将运动行为轨迹建模为一组微分方程,并通过线性动态系统保证了目标的全局稳定性,可以从一个简单的示教中学习较复杂的运动行为,已有广泛的使用^[19-22]。然而,DMP对时间具有依赖性,使其对时间扰动较为敏感,稳定性受到一定影响;而且泛化能力有限,其背后的物理意义也比较难于理解。

动态系统(DS)法^[13, 23]是近些年出现的一种对运动数据建模的机器学习方法。DS可以表征更丰富的行为,可以在状态空间的不同部分学习不同的运动。然而,由于DS存在稳定性的问题,常规的基于GMR、GPR(gaussian process regression)^[24]以及LWR的回归方法与DS结合也没有实现在目标点上的稳定收敛^[25]。其往往是找到关于系统的一个局部最优模型,此局部最优模型的约束条件无法保证系统收敛到稳定的目标点,仍然存在稳定性差和泛化能力不足的问题。

针对以上问题,本文通过GMM将示教运动建模为非线性DS,附加约束条件以确保DS的全局渐近收敛,将动态系统模型的学习问题转化为求解一个约束优化问题,通过优化问题的求解得到DS的参数,获得关于示教运动的非线性DS模型,并将其作为控制策略应用于机器人以实现示教运动的模仿。通过将动态系统方法引入模仿学习,并考虑其在目标点处全局渐近收敛的约束条件,相对于基于轨迹匹配的机器人模仿学习方法,其除了具有较强的数据编码能力和抗噪能力,而且可以保证生成的轨迹收敛到目标点,

避免了先前方法稳定性差的问题。同时,从未示教区域开始的再现运动轨迹也可以保证收敛到目标点,泛化能力将大大提高。以7bot机械臂为实验对象,进行仿真实验和机器人实验研究,实验结果表明:基于DS的模仿学习方法再现生成的模仿轨迹具有全局渐近收敛的特点,从不同起始位置开始的生成轨迹最终都可以到达目标点,学习得到的运动模型在稳定性和泛化能力两方面表现较好。

1 基于动态系统的示教运动建模

1.1 示教运动的表示及动态系统模型

在机器人模仿示教者示教运动的过程中,示教者通过直接或者间接的方式指导机器人手臂完成一次或多次的运动,机器人通过其内部的关节传感器或者位置传感器,或者是外部传感器,如:Kinect视觉传感,来捕获关于示教运动的相关信息。其中示教运动的轨迹信息可以表示为如下的集合形式

$$\{\zeta_{t,n}, \dot{\zeta}_{t,n}\}_{t=0,n=1}^{T_n,N} \quad (1)$$

式中:状态变量 ζ 表示机器人手臂末端执行器在笛卡尔坐标系下的位置;对于不同的示教任务, N 表示在某个模仿任务下的示教次数。而对于每一次的示教轨迹数据集可以表示为 $X_n = [\zeta_0, \zeta_1, \zeta_2, \dots, \zeta_{T_n}]$ 。一阶动态系统被用来建模所记录的示教轨迹数据,动态系统的一阶自治常微分方程表示为:

$$\dot{\zeta} = f(\zeta) + \varepsilon \quad (2)$$

式中: $f: \mathbf{R}^d \rightarrow \mathbf{R}^d$ 是非线性连续可微函数; ζ 表示机器人末端的位置; $\dot{\zeta}$ 对应于速度; ε 表示均值为0的高斯噪声项,代表了传感器测量的不准确性以及示教过程的不稳定所带来的示教数据的误差。且函数 f 可以由一组参数 θ 来描述。具体地,采用相应的方法,建模式(2)为具体表达形式,进而可以求解 θ 的值。

1.2 基于统计方法的高斯混合模型建模动态系统

为了得到动态系统具体的表达形式,考虑使用基于统计方法的GMM来建立对于DS的概率描述/估计。高斯混合模型是一种有限混合模型,其通过有限的高斯函数的混合来建立对示教运动数据的粗略的表示,同时消除数据的噪声。给定一组 N 次示教轨迹数据 $\{\zeta_{t,n}, \dot{\zeta}_{t,n}\}_{t=0,n=1}^{T_n,N}$, GMM由以下概率密度函数表示

$$p(\zeta_n) = \sum_{k=1}^K \omega_k \text{GP}(\zeta_n | \mu_k, \Sigma_k) \quad (3)$$

$$\text{GP}(\zeta_n|\mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^{2d}|\Sigma_k|}} \exp\left\{-\frac{1}{2}(\zeta_n - \mu_k)^T \Sigma_k^{-1}(\zeta_n - \mu_k)\right\} \quad (4)$$

式(3)、(4)中: $\text{GP}(\cdot)$ 为高斯过程概率密度函数, 或可简化表示为 $N(\cdot)$; K 为 GMM 中高斯分布的个数/混合模型的个数; ω_k 为第 k 个高斯分布的权重, 且 $\sum_{k=1}^K \omega_k = 1$; μ_k 为该密度函数的均值向量; Σ_k 为协方差矩阵; 相应的 GMM 的待确定参数可以表示为 $\theta_k = \{\omega_k, \mu_k, \Sigma_k\}, k = 1, 2, \dots, K$; 且全部参数表示为 $\theta = \{\theta_1, \theta_2, \dots, \theta_K\}$ 。

给定 $\zeta, \dot{\zeta}$ 的条件概率分布为

$$p(\dot{\zeta}|\zeta, k) = N(\dot{\zeta}|\tilde{\zeta}_k, \tilde{\Sigma}_{\dot{\zeta}k}) \quad (5)$$

其中:

$$\tilde{\zeta}_k = \mu_{\dot{\zeta}k} + \Sigma_{\dot{\zeta}k} \Sigma_{\zeta k}^{-1}(\zeta - \mu_{\zeta k}) \quad (6)$$

$$\tilde{\Sigma}_{\dot{\zeta}k} = \Sigma_{\dot{\zeta}k} - \Sigma_{\dot{\zeta}k} \Sigma_{\zeta k}^{-1} \Sigma_{\zeta k} \quad (7)$$

由此, $\dot{\zeta}$ 对应于 ζ 的条件概率可以定义为

$$p(\dot{\zeta}|\zeta) = \sum_{k=1}^K \beta_k N(\dot{\zeta}|\tilde{\zeta}_k, \tilde{\Sigma}_{\dot{\zeta}k}) \quad (8)$$

$$\beta_k = \frac{\omega_k p(\dot{\zeta}|k)}{\sum_{k=1}^K \omega_k p(\dot{\zeta}|k)} = \frac{\omega_k N(\dot{\zeta}|\mu_{\dot{\zeta}k}, \Sigma_{\dot{\zeta}k})}{\sum_{i=1}^K \omega_i N(\dot{\zeta}|\mu_{\dot{\zeta}i}, \Sigma_{\dot{\zeta}i})} \quad (9)$$

由条件期望可求得整个高斯混合模型中 $\dot{\zeta}$ 的总的期望为

$$\dot{\zeta} = \tilde{f}(\zeta) = \sum_{k=1}^K H_k(\zeta)(A_k \zeta + B_k) \quad (10)$$

其中:

$$\begin{cases} A_k = \Sigma_{\dot{\zeta}k} \Sigma_{\zeta k}^{-1} \\ H_k(\zeta) = \frac{\omega_k p(\dot{\zeta}|k)}{\sum_{i=1}^K \omega_i p(\dot{\zeta}|i)} \\ B_k = \mu_{\dot{\zeta}k} - A_k \mu_{\zeta k} \end{cases} \quad \forall k = 1, 2, \dots, K \quad (11)$$

式(10)、(11)得到的动态系统表示为一组线性动态系统的非线性加权的形式, 也即得到了关于示教运动的动态系统模型的具体表达形式。

2 动态系统全局稳定的充分条件

由以上得到的动态系统模型, 为了保证动态系统模型在示教的目标点具有全局收敛的特性以及提高其泛化能力, 也即将在不同的初始点开始最终到达同一目标点的模仿学习任务转化为具有全局收敛的学习任务, 则需要考虑其稳定性问题。对于由 $\dot{\zeta} = f(\zeta)$ 定义的动态系统, 根据 Lyapunov 稳定性理论: 其在点 ζ^* 处全局渐近稳定的充分条件是存在一个连续可微的 Lyapunov 函数 $V(\zeta): \mathbf{R}^d \rightarrow \mathbf{R}$ 满足:

$$\begin{cases} V(\zeta) > 0, \forall \zeta \in \mathbf{R}^d, \zeta \neq \zeta^* \\ \dot{V}(\zeta) < 0, \forall \zeta \in \mathbf{R}^d, \zeta \neq \zeta^* \\ V(\zeta^*) = 0, \dot{V}(\zeta^*) = 0 \end{cases} \quad (12)$$

为了得到式(10)中 DS 的稳定条件, 可以构造如下的 Lyapunov 函数:

$$V(\zeta) = \frac{1}{2}(\zeta - \zeta^*)^T(\zeta - \zeta^*), \forall \zeta \in \mathbf{R}^d \quad (13)$$

由式(12)、(13)可以求得保证 DS 全局稳定的充分条件

$$\begin{cases} (A_k \zeta^* + B_k) = 0 \\ A_k + (A_k)^T < 0 \end{cases} \quad \forall k = 1, 2, \dots, K \quad (14)$$

将式(14)作为 DS 的约束条件, 则 DS 满足全局收敛的特性。同时, 考虑高斯混合模型的性质和式(14)的条件, 其共同组成了具有全局渐进稳定的动态系统的完整的约束条件。此后, 需要求解动态系统模型的相应参数, 得到对于动态系统完整的描述。

3 动态系统参数学习

为了学习得到具有全局渐近稳定的 DS, 需要对其参数进行求解。由以上可知 DS 的未知参数为 $\theta = \{\omega_1, \omega_2, \dots, \omega_K; \mu_1, \mu_2, \dots, \mu_K; \Sigma_1, \Sigma_2, \dots, \Sigma_K\}$ 。本文通过将动态系统参数学习的过程转化为求解一个非线性优化问题, 在保证模型全局渐近稳定的约束条件下求解优化问题来计算 θ 的最优值。使用对数似然函数作为优化的目标函数

$$\arg \min_{\theta} J(\theta) = -\frac{1}{\mathcal{T}} \sum_{n=1}^N \sum_{t=0}^{T_n} \log p(\zeta_{t,n}|\theta) \quad (15)$$

对应的约束条件为

$$\begin{cases} (A_k \zeta^* + B_k) = 0 \\ A_k + (A_k)^T < 0 \\ \Sigma_k > 0 \\ 0 < \omega_k \leq 1 \\ \sum_{k=1}^K \omega_k = 1 \end{cases} \quad \forall k = 1, 2, \dots, K \quad (16)$$

式中 $\mathcal{T} = \sum_{n=1}^N T_n$ 是训练数据点总的数量, 也即示教数据的数量。

对于以上优化问题, 可以将其化为非线性规划问题, 借助于标准的优化约束技术可以很好地解决, 最终得到具有全局渐近稳定的动态系统的最优参数 θ , 实现对于动态系统的完整描述。

4 机器人模仿学习系统的设计

在以上内容的基础上, 设计了基于 DS 的机器人模仿学习系统。所设计的机器人模仿学习系

统工作流程如图1所示。在机器人模仿学习系统中,本文主要集中在对示教过程和学习过程的研究。

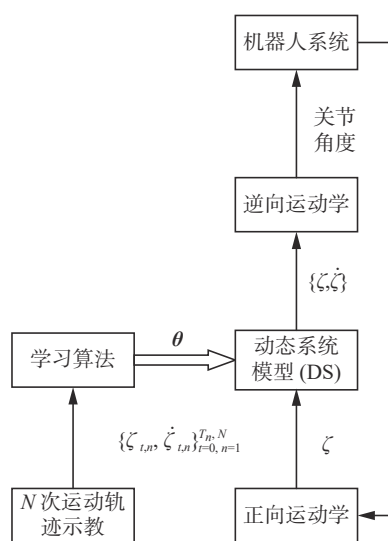


图1 模仿学习系统工作流程

Fig. 1 Workflow of imitation learning system

对应于模仿学习的3个基本的步骤,模仿学习系统工作流程,首先,对于某个固定的模仿任务(如:绘画或书写、拿放物品、篮球投篮等任务),通过手把手地抓取机器人手臂末端执行器进行多次目标点相对固定的轨迹示教,获得示教轨迹数据;然后,将获得的示教轨迹数据预处理后送入学习算法学习关于动态系统模型参数,得到动态系统模型。最后,将动态系统模型作为机器人的控制策略,设置机器人手臂初始的运动位置,结合机器人的控制系统完成对示教任务的模仿。

5 机器人模仿学习实验研究

为了验证基于动态系统的机器人模仿学习方法的有效性,在已设计的模仿学习系统的框架下进行仿真实验和机器人实验研究。

5.1 机器人平台和示教数据的获取

实验中的机器人平台是7bot机械臂,如图2所示。7bot机械臂是一款桌面应用级的机械臂,也是一款6轴的全金属智能机械臂。借助于Processing和Arduino Due实现开发和控制,使用USB线与计算机之间完成数据通信,能够实现在三维空间的多种运动,使用方便灵活。

实际中,对于示教数据的获取,是在机械臂的运动空间内通过抓取7bot机械臂末端进行N次运动轨迹示教。由机器臂自身携带的关节角度传感器记录运动轨迹信息。另外,也可以通过在仿真环境中拖动虚拟机械臂,进行示教并获取示教数据。

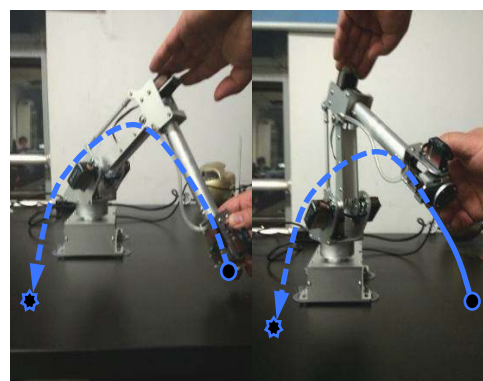
具体地,实际中的示教数据是由上位机程序处理示教轨迹信息,得到关于机械臂末端执行器在笛卡尔坐标系中的位置 $\mathbf{x}(\mathbf{x} \in \mathbf{R}^d, d=2/d=3)$,即为示教数据。对应于算法中的表述,经过计算,对于示教数据有 $\zeta = \mathbf{x}, \dot{\zeta} = \dot{\mathbf{x}}$ 。



图2 7bot机械臂

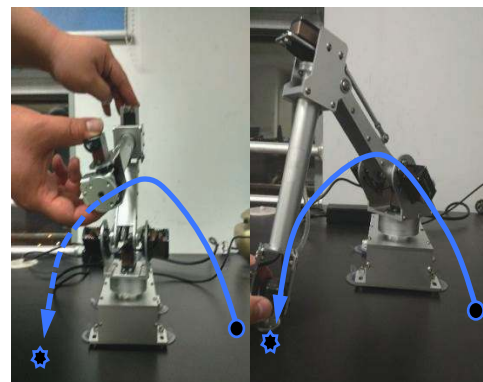
Fig. 2 7bot manipulator

而对于具体的运动轨迹的示教,模仿学习示教过程如图3所示。示教者通过抓取机械臂末端进行运动示教,此运动可以是一个拿放物品的运动。图中的几个关键子过程可以说明整个示教过程是如何完成的,黑色的圆点代表示教运动的起始点,黑色的星形点代表示教运动的终点或目标点,实线代表已完成的示教运动过程,虚线代表未完成的示教运动过程。



(a) 过程1

(b) 过程2



(c) 过程3

(d) 过程4

图3 模仿学习示教过程示意

Fig. 3 Imitation learning demonstration process diagram

5.2 实验步骤

本文在 MATLAB 环境中进行相关仿真实验, 为了简化实验操作的复杂性, 主要在 MATLAB 环境中进行二维平面的模仿学习仿真实验, 具体步骤如下:

1) 将实际采集的或在仿真环境中生成的多次示教运动轨迹信息在 MATLAB 中进行预处理并可视化, 完成模仿学习的示教过程。

2) 设置算法的相关初始值, 将示教运动数据由 GMM 模型编码为一动态系统模型。

3) 通过求解带有全局渐近稳定性约束条件的动态系统相关参数集合 $\theta_k = \{\omega_k, \mu_k, \Sigma_k\}$, 学习得到动态系统模型, 完成模仿学习的学习过程。

4) 由学习得到的动态系统模型生成新的模仿运动轨迹, 并分析生成的运动轨迹的相关结果。

5) 最后将学习得到的动态系统模型作为控制策略, 结合机器人控制系统, 完成对模仿学习过程中示教运动的模仿。

5.3 仿真实验

本部分主要针对简单示教运动和较复杂示教运动进行实验仿真。通过将示教运动轨迹与提出的方法生成的轨迹相比较来验证方法的有效性; 并结合仿真环境中的两连杆机械臂系统, 展示最终的学习效果。其中, 规定简单示教运动是具有一定弯曲度的类直线运动。而较复杂示教运动相比于简单示教运动其示教运动体现出更多的示教变化, 如二次及以上的曲线形状的运动、自相交运动或者它们的组合等。

模仿学习性能的量化指标可采用机械臂末端位置与目标点的距离 d 、示教与再现运动的路径积分长度之差 $\text{diff} = \left| \int_{\zeta} ds - \int_{\zeta} ds \right|$ 以及机器人完成模仿任务的时间与示教时间的差 $|\Delta t|$, 以此来衡量。

5.3.1 简单示教运动实验

图 4 给出了单次简单示教的仿真结果。示教运动代表的任务类型可以是绘画任务中的一笔或者是拿和放物品的运动等。图 4 中, (a) 为在 MATLAB 环境中简单示教运动轨迹, 起点是 $[0 \ 1]^T$, 目标点 $[1 \ 0.5]^T$; (b) 为使用文中的学习方法, 模型生成的使机器人到达目标点的平滑轨迹流图, 可见所有生成轨迹均收敛到目标点; (c)、(d) 中蓝色虚线为在 MATLAB 中模拟的一个两连杆机械臂系统, 将学习得到的运动模型作为控制策略, 从不同起点执行模仿的结果, 起点分别为 $[-0.25 \ 1]^T$ 、 $[0 \ 0.5]^T$ 。最终, 机械臂末端与目标点的距离分别为 $d = 0.0113 \text{ m}$ 、 $d = 0.0091 \text{ m}$ 。

在实际示教过程中, 机器人可能需要多次示教。图 5 分别为多次示教运动轨迹和经算法学习后生成的平滑的轨迹流图, 可见生成的轨迹均收敛到示教运动的目标点。

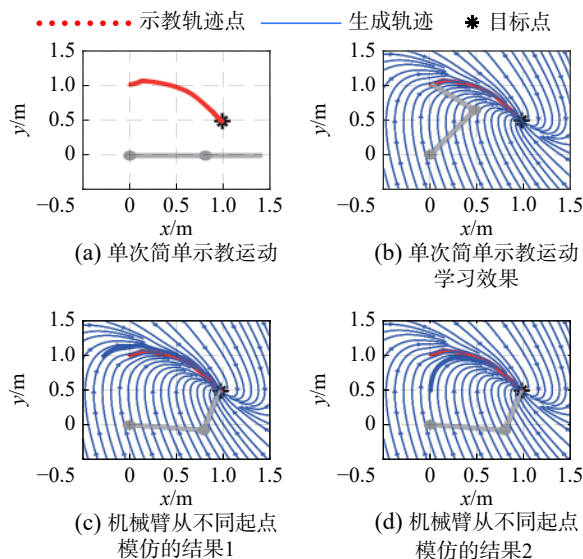


图 4 单次简单示教运动与仿真结果

Fig. 4 Single simple demonstration motion and simulation results

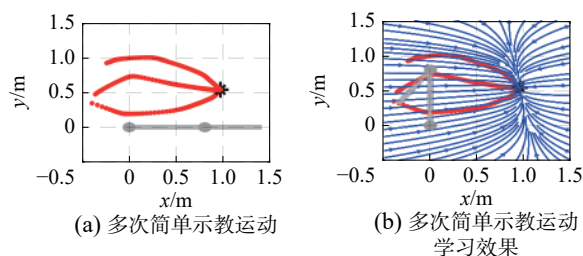


图 5 多次简单示教运动与仿真结果

Fig. 5 Multiple simple demonstration motion and simulation results

由以上可知, 学习得到的运动模型一方面具有很强的稳定性, 轨迹均收敛到同一个目标点。另一方面, 由未示教区域开始的再现运动也可以在具有一定模仿相似性的前提下到达目标点附近, 误差较小, 体现出模型具有较好的泛化能力。

5.3.2 较复杂示教运动实验

对于较复杂的示教运动, 相应的仿真结果如图 6、7 所示。其中图 6 是对机器人进行单次示教后的实验结果。单次示教的起点为 $[-0.25 \ 0.5]^T$, 目标点为 $[0.9 \ 0.4]^T$ 。通过算法学习后, 模型生成的轨迹均收敛到目标点 $[0.9 \ 0.4]^T$ 。类似地, 分别从不同的起点 $[-0.25 \ 0.75]^T$ 和 $[0 \ 0.5]^T$ 开始再现模仿, 实验结果如图 6(c)、(d) 所示。机械臂末端与目标点的距离分别为 $d = 0.0162 \text{ m}$ 、 $d = 0.0146 \text{ m}$, 误差较小。

同样地,在实际的机器人示教过程中,需要多次示教时,实验结果如图7所示。其中,经过算法学习后,学习得到的运动模型生成的轨迹收敛到同一个目标点。

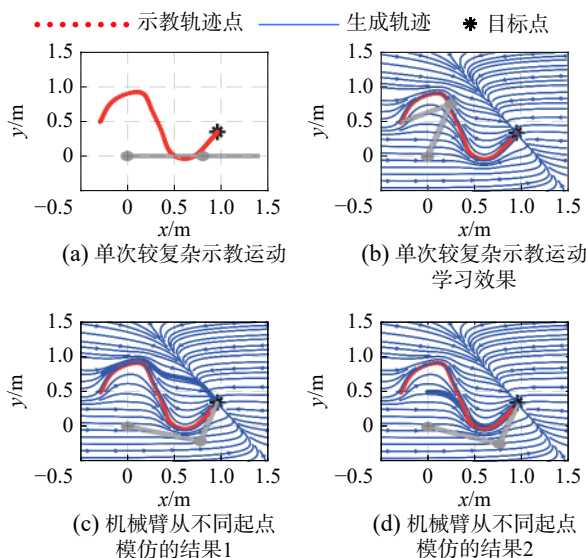


图6 单次较复杂示教运动与仿真结果

Fig. 6 Single more complex demonstration motion and simulation results

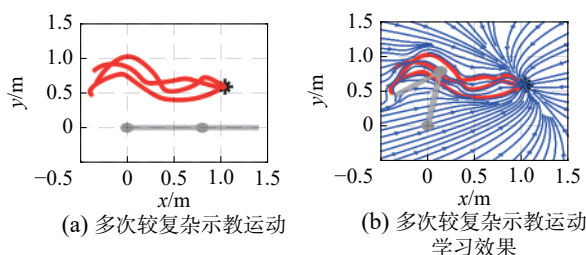


图7 多次较复杂示教运动与仿真结果

Fig. 7 Multiple more complex demonstration motion and simulation results

由此可知,学习得到的运动模型对于此类较复杂的示教运动也有不错的学习效果,模型的稳定性较好,未示教空间的泛化能力也较好。

5.4 机器人实验

此部分结合实际的7bot机器人系统来验证学习方法的有效性。

图8、9分别展示了二维平面中单次示教时,机器人真实示教运动轨迹和经算法学习后的学习效果。红色虚线为示教轨迹,蓝色实线为生成轨迹。由图可知,所有生成轨迹最终都收敛到目标点,体现了运动模型具有全局渐进稳定性的特点;但对于有交叉点的自相交运动,其在交叉点的学习效果并不明显,忽略了部分的学习,学习效果有待提高。

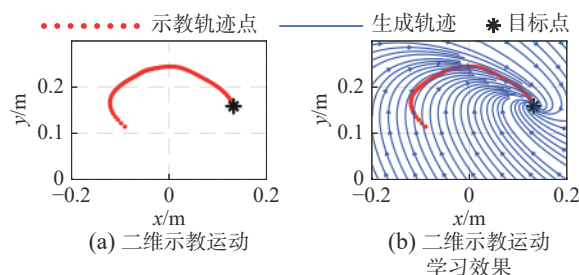


图8 二维真实示教运动与学习效果

Fig. 8 Two-dimensional real demonstration motion and learning results

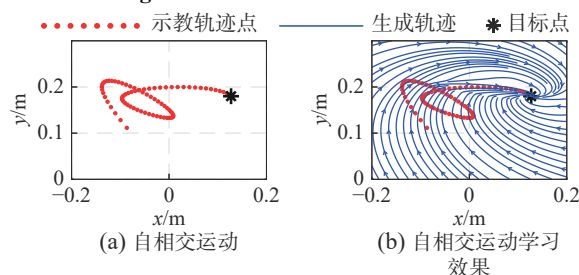


图9 自相交运动与学习效果

Fig. 9 Self-intersection motion and learning results

图10、11展示了通过7bot机械臂系统对示教运动进行模仿的结果。将学习得到的运动模型作为控制策略,设定机器人的初始位置,由机器人系统进行模仿实现。实验结果显示其成功实现了对运动的模仿,到达目标点附近。

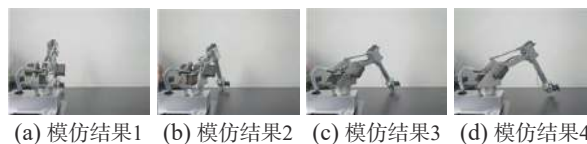


图10 机械臂模仿的结果

Fig. 10 Results of the imitation of a manipulator

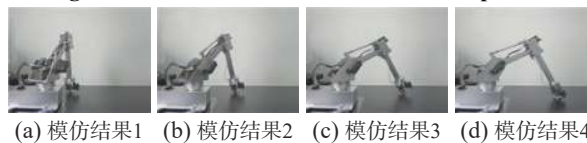


图11 机械臂对于自相交运动的模仿结果

Fig. 11 Results of the imitation of a manipulator for self-intersection motion

同样,图12中分别为在三维空间中的示教及学习效果。可以看出从起始点开始的三条生成轨迹均收敛到目标点,表明运动模型的稳定性很好。图13是7bot机械臂对示教运动进行模仿的结果,可见其也很好地完成对示教运动的模仿,到达目标点附近。

表1给出了机器人实验中机器人模仿性能的量化结果。可以看出运动模仿结束后,机械臂末端距离目标点较近,误差较小;示教时间和模仿时间相差不大;但是在自相交运动上轨迹长度之差比较大。

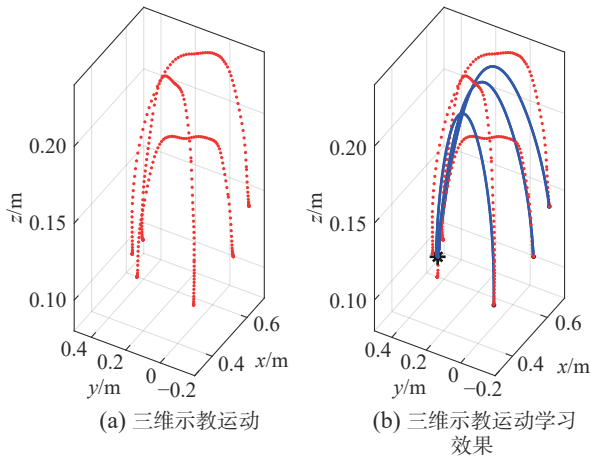


图12 三维真实示教运动与学习效果

Fig. 12 Three-dimensional real demonstration motion and learning results



(a) 模拟结果1 (b) 模拟结果2 (c) 模拟结果3 (d) 模拟结果4

图13 机械臂对于三维运动的模拟结果

Fig. 13 Results of the imitation of a manipulator for three-dimensional motion

表1 机器人模仿性能的量化结果

Table 1 Quantitative results of robot imitative performance

运动	$ \Delta t /s$	d/m	$diff/m$
运动1(图10)	3.79	0.027	0.101
运动2(图11)	6.57	0.021	0.455
运动3(图13)	5.82	0.039	0.375

综上实验结果,基于DS的模仿学习方法在简单示教运动、较复杂示教运动以及真实示教运动的学习上,通过附加稳定性约束条件的方法,再现的生成轨迹均收敛到目标点,一方面使运动模型的稳定性更强;另一方面,使运动模型具有较好的泛化能力。

6 结束语

本文针对当前机器人模仿学习过程中,运动模仿稳定性差,泛化能力不足,引入了一种基于动态系统的模仿学习方法来解决以上问题。首先,通过使用高斯混合模型将示教运动建模为一动态系统;然后,考虑稳定性问题,基于Lyapunov稳定性理论,给出动态系统全局渐近稳定的充分条件;最后,建立一个包含多约束的非线性优化问题,迭代求解以得到最优参数,进而得到动

态系统模型,并作为机器人的控制策略实现对示教运动的模仿。从实验中可以得到以下结论:

1) 文中提出的模仿学习方法在简单示教运动和较复杂的示教运动的仿真实验中,学习得到的运动模型从示教起始点生成的轨迹均能够到目标点,轨迹平滑,稳定性好。

2) 将示教运动建模为动态系统,通过附加稳定性约束条件,保证了学习得到的运动模型生成的轨迹都收敛到目标点。

3) 文中的方法可以实现从单一到多次示教运动的学习,对于未被示教的状态空间其模型也可以生成平滑的轨迹收敛到目标点,具有较强的泛化能力。

4) 存在的问题是对于复杂的自相交的示教运动,模型忽略了运动的关键信息,运动再现相似度将会稍差。同时,在其他机器人系统上的应用还有待进一步验证。对于非精密要求场合、目的为导向的机器人工作场合,此方法的通用性和稳定性具有较大优势。

在未来,可以结合具有高精度控制的机器人来实现实时调整再现轨迹提高模仿的相似性;同时,可以通过提高系统的阶次来学习复杂的自相交运动;另外,结合快速控制技术可以将其扩展到快速捕获运动目标的领域。

参考文献:

- [1] SCHAAL S. Is imitation learning the route to humanoid robots?[J]. Trends in cognitive sciences, 1999, 3(6): 233–242.
- [2] 杨俊友, 马乐, 白殿春, 等. 机器人模仿学习的非接触观测控制图模型[J]. 机器人, 2014, 36(3): 309–315.
YANG Junyou, MA Le, BAI Dianchun, et al. Cybernetic-graphic model for robot imitation learning based on non-contact observation[J]. Robot, 2014, 36(3): 309–315.
- [3] 曾华琳, 黄雨轩, 晁飞, 等. 书写机器人研究综述[J]. 智能系统学报, 2016, 11(1): 15–26.
ZENG Hualin, HUANG Yuxuan, CHAO Fei, et al. Survey of robotic calligraphy research[J]. CAAI transactions on intelligent systems, 2016, 11(1): 15–26.
- [4] UGUR E, NAGAI Y, SAHIN E, et al. Staged development of robot skills: behavior formation, affordance learning and imitation with motionese[J]. IEEE transactions on autonomous mental development, 2015, 7(2): 119–139.
- [5] BOBOC R G, TOMA M I, PANFIR A N, et al. Learning new skills by a humanoid robot through imitation[C]//Proceedings of the 2013 IEEE 14th International Symposium

- on Computational Intelligence and Informatics. Budapest, Hungary, 2013: 515–519.
- [6] 于建均, 门玉森, 阮晓钢, 等. 基于 Kinect 的 Nao 机器人动作模仿系统的研究与实现 [J]. 智能系统学报, 2016, 11(2): 180–187.
- YU Jianjun, MEN Yusen, RUAN Xiaogang, et al. CAAI transactions on intelligent systems[J]. CAAI transactions on intelligent systems, 2016, 11(2): 180–187.
- [7] BILLARD A, CALINON S, DILLMANN R, et al. Robot programming by demonstration[M]//SICILIANO B, KHATIB O. Springer Handbook of Robotics. Berlin, Heidelberg: Springer, 2008: 1371–1394.
- [8] ATKESON C G, SCHAAL S. Learning tasks from a single demonstration[C]//Proceedings of International Conference on Robotics and Automation. Albuquerque, NM, USA, 1997: 1706–1712.
- [9] MAAREF M, REZAZADEH A, SHAMAEI K, et al. A gaussian mixture framework for co-operative rehabilitation therapy in assistive impedance-based tasks[J]. IEEE journal of selected topics in signal processing, 2016, 10(5): 904–913.
- [10] KULIĆ D, TAKANO W, NAKAMURA Y. Incremental learning, clustering and hierarchy formation of whole body motion patterns using adaptive hidden markov chains[J]. The international journal of robotics research, 2008, 27(7): 761–784.
- [11] ANTONELO E A, SCHRAUWEN B. Supervised learning of internal models for autonomous goal-oriented robot navigation using reservoir computing[C]//Proceedings of 2010 IEEE International Conference on Robotics and Automation. Anchorage, AK, USA, 2010: 2959–2964.
- [12] CALINON S, BILLARD A. Incremental learning of gestures by imitation in a humanoid robot[C]//Proceedings of the ACM/IEEE international conference on Human-robot interaction. Arlington, VA, USA, 2007: 255–262.
- [13] HERSCH M, GUENTER F, CALINON S, et al. Dynamical system modulation for robot learning via kinesthetic demonstrations[J]. IEEE transactions on robotics, 2008, 24(6): 1463–1467.
- [14] SCHAAL S, ATKESON C, VIJAYAKUMAR S. Scalable locally weighted statistical techniques for real time robot learning[J]. Applied intelligence-special issue on scalable robotic applications of neural networks, 2002, 17(1): 49–60.
- [15] PETERNEL L, OZTOP E, BABIČ J. A shared control method for online human-in-the-loop robot learning based on locally weighted regression[C]//Proceedings of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Daejeon, South Korea, 2016: 3900–3906.
- [16] MCCORMICK J, VINCS K, NAHAVANDI S, et al. Teaching a digital performing agent: Artificial neural network and hidden markov model for recognising and performing dance movement[C]//Proceedings of the 2014 International Workshop on Movement and Computing. Paris, France, 2014: 70.
- [17] 于建均, 吴鹏申, 左国玉, 等. 基于 RNN 的机械臂任务模仿系统 [J]. 北京工业大学学报, 2018, 44(11): 1401–1408.
- YU Jianjun, WU Pengshen, ZUO Guoyu, et al. Robot arm task imitation system based on RNN[J]. Journal of Beijing University of Technology, 2018, 44(11): 1401–1408.
- [18] 于建均, 门玉森, 阮晓钢, 等. 在书写任务中的基于轨迹匹配的模仿学习 [J]. 北京工业大学学报, 2016, 42(8): 1144–1152.
- YU Jianjun, MEN Yusen, RUAN Xiaogang, et al. Imitation learning based on trajectory matching in the writing task[J]. Journal of Beijing University of Technology, 2016, 42(8): 1144–1152.
- [19] 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.
- [20] PARASCHOS A, RUECKERT E, PETERS J, et al. Model-free probabilistic movement primitives for physical interaction[C]//Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany, 2015: 2860–2866.
- [21] KOCH K H, CLEVER D, MOMBAUR K, et al. Learning movement primitives from optimal and dynamically feasible trajectories for humanoid walking[C]//Proceedings of 2015 IEEE-RAS 15th International Conference on Humanoid Robots. Seoul, South Korea, 2015: 866–873.
- [22] KOBER J, GIENGER M, STEIL J J. Learning movement primitives for force interaction tasks[C]//Proceedings of 2015 IEEE International Conference on Robotics and Automation. Seattle, USA, 2015: 3192–3199.
- [23] WIGGINS S. Introduction to applied nonlinear dynamical systems and chaos[M]. 2nd ed. New York: Springer Science & Business Media, 2003.
- [24] SEEGER M. Gaussian processes for machine learning[J]. International journal of neural systems, 2004, 14(2): 69–106.
- [25] KHANSARI-ZADEH S M, BILLARD A. BM: an iterat-

ive algorithm to learn stable non-linear dynamical systems with gaussian mixture models[C]//Proceedings of 2010 IEEE International Conference on Robotics and Automation. Anchorage, USA, 2010: 2381–2388.

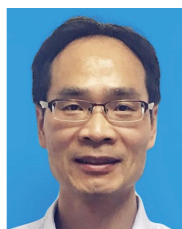
作者简介:



于建均, 女, 1965 年生, 副教授, 主要研究方向为智能机器人的仿生自主控制、智能计算与智能优化控制、复杂过程建模、优化与控制。主持或参与国家“863”计划项目、国家自然科学基金等省部级科研项目以及横向科研课题多项。获得国家发明专利、实用新型专利、国家软件著作权 10 余项。发表 SCI、EI、ISTP 收录论文 40 余篇。



姚红柯, 男, 1991 年生, 硕士研究生, 主要研究方向为机器学习、机器人行为模仿和控制。



左国玉, 男, 1971 年生, 副教授, 博士, 主要研究方向为机器人学习与控制。主持科研项目 10 余项, 取得国家发明专利 20 余项。发表学术论文 40 余篇。