

# 基于改进动态系统稳定估计器的机器人技能学习方法

金聪聪<sup>1</sup> 刘安东<sup>1</sup> LIU Steven<sup>2</sup> 张文安<sup>1</sup>

**摘要** 提出一种基于改进动态系统稳定估计器的机器人技能学习方法。现有的动态系统稳定估计器方法可以通过非线性优化来确保学习系统的全局稳定性，但是存在确定高斯混合分量个数困难以及稳定性和精度无法兼顾的问题。因此，根据贝叶斯非参数模型可以自动确定合适分量个数的特性，采用狄利克雷过程高斯混合模型对演示进行初始拟合。随后利用参数化二次李雅普诺夫函数重新推导新的稳定性约束，有效地解决了动态系统稳定估计器方法中稳定性和精度难以兼顾的问题。最后，在 LASA 数据库和 Franka-panda 机器人上的实验验证了新方法的有效性和优越性。

**关键词** 示教学习，动态系统，贝叶斯非参数模型，高斯混合模型，李雅普诺夫函数

**引用格式** 金聪聪，刘安东，LIU Steven，张文安. 基于改进动态系统稳定估计器的机器人技能学习方法. 自动化学报, 2022, 48(7): 1771-1781

**DOI** 10.16383/j.aas.c200341

## A Robot Skill Learning Method Based on Improved Stable Estimator of Dynamical Systems

JIN Cong-Cong<sup>1</sup> LIU An-Dong<sup>1</sup> LIU Steven<sup>2</sup> ZHANG Wen-An<sup>1</sup>

**Abstract** This paper presents a novel robot skill learning method based on improved stable estimator of dynamical systems (SEDS). The original SEDS method can ensure the global stability of the learning system through nonlinear optimization. However, it cannot automatically determine the optimal number of Gaussian components and is difficult to make a trade-off between reconcile the stability and accuracy. Therefore, note that the Bayesian nonparametric model can be used to determine the appropriate number of components, the Dirichlet process Gaussian mixture model is applied to perform the initial fitting of the demonstrations in this paper. Then, the stability constraints are reformulated by using the parameterized Lyapunov function. The problems of stability and accuracy in the SEDS method are solved effectively. Finally, experiments on a LASA dataset and a Franka-panda cooperative robot validate the effectiveness and superiority of the proposed method.

**Key words** Learning from demonstration, dynamical system, Bayesian nonparametric model, Gaussian mixture model, Lyapunov function

**Citation** Jin Cong-Cong, Liu An-Dong, Liu Steven, Zhang Wen-An. A robot skill learning method based on improved stable estimator of dynamical systems. *Acta Automatica Sinica*, 2022, 48(7): 1771-1781

近年来，机器人产业高速发展，整体规模持续增长，在制造业和服务业等众多领域都有广泛应用。随着工业产品迭代速度日益增长，个性化需要与日俱增，传统依靠手工编程完成特定任务的方法难以

适应新的需求。因此，迫切需要开发简单实用，且可以灵活适用于多种任务的机器人技能学习方法。

机器人示教学习 (Learning from demonstration, LfD) 的灵感最初来源于人类的模仿学习，近几年获得学术界和工业界的广泛关注<sup>[1-4]</sup>。机器人通过观察用户演示来学习新技能，同时将新技能泛化到不同场景下的相似任务中去，一般包含演示、学习、复现三个阶段。演示阶段需要解决的问题是如何向机器人进行技能演示，常用的方法有视觉示教、动觉示教、遥操作示教和双臂示教。学习阶段是对技能进行建模，抽象的技能符号由示教数据具体表示，然后利用示教数据训练模型参数。复现阶段的性能主要体现在稳定性、复现精度、泛化能力和抗干扰性能 4 个方面。

动态系统 (Dynamical system, DS) 是对机器

收稿日期 2020-05-22 录用日期 2020-09-30  
Manuscript received May 22, 2020; accepted September 30, 2020

浙江省自然科学基金重大项目 (LD21F030002)，国家自然科学基金 (61822311, 61973275) 资助

Supported by Zhejiang Provincial Natural Science Foundation of China (LD21F030002) and National Natural Science Foundation of China (61822311, 61973275)

本文责任编辑 陈德旺

Recommended by Associate Editor CHEN De-Wang

1. 浙江工业大学信息工程学院 杭州 310023 中国 2. 德国凯泽斯劳滕工业大学电气与计算机工程系 凯泽斯劳滕 67663 德国

1. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China 2. Department of Electrical and Computer Engineering, University of Kaiserslautern, Kaiserslautern 67663, Germany

人技能进行建模的常用方法<sup>[5]</sup>, 该方法将规划和执行集成到一起, 并将所有可能的解决方案嵌入到模型中以实现目标<sup>[6]</sup>. 在非线性 DS 基础上发展起来的动态运动原语只要进行一次演示就可以对运动进行建模<sup>[7]</sup>, 动态运动原语描述的运动模型由非线性 DS 和线性 DS 组成, 其中非线性部分保证了轨迹复现的相似度, 线性部分则确保了模型全局稳定性, 两者的切换依靠相位变量平稳进行. 尽管动态运动原语提供了一种有效而精确的方法来对复杂的动态进行编码, 但是单变量编码丢失了各自由度之间的相关信息, 而且该方法本质上仍依赖于时间, 在面对时间扰动时需要用启发式方法重置相位变量<sup>[8]</sup>.

为弥补动态运动原语的缺陷, 文献 [9] 提出了动态系统的稳定估计器 (Stable estimator of dynamical systems, SEDS). 它首先利用高斯混合模型 (Gaussian mixture models, GMM) 和高斯混合回归 (Gaussian mixture regression, GMR) 的概率学习方法对轨迹进行初步拟合. 概率学习方法是轨迹编码中的常用方法, 它可以保留演示的固有可变性<sup>[10]</sup>, 但是无法确保训练得到的动态系统具有全局稳定性. 因此 SEDS 在后续优化中加入了稳定性约束, 确保机器人在不受扰动的情況下能够到达目标点. 然而, 过于严格的稳定性约束可能会在学习过程中限制模型的精度. 针对 SEDS 方法中稳定性和精度难以平衡的问题, 文献 [11] 利用微分同胚变换改进了 SEDS, 称作  $\tau$ -SEDS. 该方法在保证系统稳定性的同时, 很大程度上克服了模型的精度问题. 但是增加了模型复杂度, 导致学习更加耗时, 仅限于离线学习. DS 方法的快速稳定学习<sup>[12]</sup> 同时考虑了估计 DS 的稳定性、准确性和学习速度三个因素. DS 方法快速稳定学习的快速学习能力在很大程度上方便了它的实际应用, 但在稳定性和准确性方面还不够优越. 文献 [13] 提出了一种基于流形浸没和淹没的学习方法来解决精度与稳定性的矛盾, 该方法保证了有效提取动力学特征和稳定形式的高精度, 而且能处理交叉运动的情形, 但模型复杂度较高.

SEDS 中另一个值得注意的问题是混合高斯分量个数的选取, 但是对于该问题的相关研究较少. 过多或者过少的分量个数选取都会导致模型无法有效提取演示的动力学特征, 因此该问题具有一定的研究价值. 通常用于确定有限混合模型的最佳分量的方法是贝叶斯信息准则<sup>[14]</sup>, 然而这种模型选择方法存在一些明显的缺陷<sup>[15]</sup>, 常常过高估计模型分量的个数, 导致过拟合. 贝叶斯非参数模型是一种定义在无限维参数空间上的贝叶斯模型, 其利用在适当数量的模型分量密度上产生后验分布来调整模型

大小, 因此可以根据数据自适应聚类个数, 其中狄利克雷过程混合模型是最常用的贝叶斯非参数模型之一<sup>[16-18]</sup>. Figureoa 等<sup>[19]</sup> 提出了一种物理一致的贝叶斯非参数混合模型, 该方法可以自动估计最佳的混合分量个数, 并且将相似性测度融入先验信息, 提高了复现和泛化的精度. 但是该方法使用吉布斯采样计算模型的后验概率, 计算复杂度较高.

鉴于 SEDS 存在的上述缺陷, 本文提出了改进的 SEDS (Improved SEDS, *i*-SEDS). 该方法有效地解决了 SEDS 中稳定性和精度无法兼顾的问题, 并且可以自动确定合适的分量个数. 仿真以及 Franka-panda 协作机器人的实验结果验证了本文方法的有效性和优越性. 本文的主要贡献有: 1) 使用狄利克雷过程高斯混合模型 (Dirichlet process GMM, DPGMM) 代替 GMM 拟合演示, 并利用变分推断 (Variational inference, VI) 训练模型, 该模型可以根据演示数据自动确定合适的混合分量个数. 仿真分析超参数对 DPGMM 的影响, 为超参数的选择提供了指导意义; 2) 采用参数化的李雅普诺夫函数修改了原 SEDS 中的稳定性约束条件, 从而提高了学习轨迹的精度, 解决了稳定性和精度难以兼顾的问题.

## 1 问题描述

将机器人离散运动公式化为由自治动态系统驱动的控制律. 考虑一个状态变量  $\xi \in \mathbb{R}^d$ ,  $\xi$  通常表示笛卡尔空间中末端执行器的坐标. 有时  $\xi$  也可用于定义机器人系统的广义坐标 (例如, 机械手的关节角度或轮式机器人的姿势)

$$\dot{\xi} = f(\xi), \quad f: \mathbb{R}^d \mapsto \mathbb{R}^d \quad (1)$$

式中,  $f(\xi)$  是一个非线性连续可微函数, 用于编码特定行为, 例如拾取搬运物品, 伸手拿杯子等. 从初始位置  $\xi_0$  开始, 机器人运动  $\xi_t, t \in [0, +\infty)$  由式 (1) 积分得到, 运动最后收敛至单一吸引子  $\xi^*$ ,  $f(\xi^*) = 0$ . 本文需要估计该吸引子上全局渐近稳定的动态系统  $f(\xi)$ .

通常, 将基于一组  $M$  条演示轨迹的数据  $\left\{ \xi_{t,m}, \dot{\xi}_{t,m} \right\}_{t=0, m=1}^{T^m, M}$  对  $f(\xi)$  进行估计,  $T^m$  表示第  $m$  条演示轨迹的序列长度, 估计得到的 DS 为  $\hat{f}(\xi)$ . 一般采用回归算法进行估计, 如高斯过程回归<sup>[20]</sup>、局部加权映射回归算法<sup>[21]</sup> 和混合高斯回归等. 但这些标准的回归技术在训练过程中没有考虑 DS 的渐近稳定性,  $\hat{f}(\xi)$  很可能是不稳定的或收敛至伪吸引子. 因此 SEDS 方法将不稳定 DS 的建模转换为非线性优化问题, 构造了全局稳定的运动估计, 具体

流程如图 1 所示.

SEDS 使用 GMM-GMR 回归方法从演示中估计初始 DS. 首先利用具有  $K$  个分量的 GMM 估计联合概率分布  $p(\xi, \dot{\xi}|\Theta) = \sum_{k=1}^K \pi_k \mathcal{N}(\xi, \dot{\xi}|\mu_k, \Sigma_k)$ , 其中  $\pi_k$ 、 $\mu_k$  和  $\Sigma_k$  分别是各高斯分量的权重、均值和协方差, 即  $\Theta_k = \{\pi_k, \mu_k, \Sigma_k\}$ ,  $\Theta = \{\Theta_1, \dots, \Theta_K\}$ , 然后 GMR 部分根据联合概率分布推导出后验概率  $p(\dot{\xi}|\xi, \Theta)$ , 对其取均值得到:

$$\dot{\xi} = \hat{f}_0(\xi) = \sum_{k=1}^K h_k(\xi)(A_k \xi + b_k) \quad (2)$$

式中,  $A_k = \Sigma_k^{\xi \dot{\xi}} (\Sigma_k^{\xi \xi})^{-1}$ ,  $b_k = \mu_k^{\dot{\xi}} - A_k \mu_k^{\xi}$ ,  $h_k(\xi) = \pi_k p(\xi|k) / \sum_{i=1}^K \pi_i p(\xi|i)$ .  $A_k \xi + b_k$  表示第  $k$  个线性 DS,  $h_k(\xi)$  是状态相关的非线性系数,  $\sum_{k=1}^K h_k(\xi) = 1$ , 且  $h_k(\xi) > 0$ .

上述方法有足够的灵活性来建模各种运动, 但是无法确保 DS 的渐近稳定性, 为此, SEDS 给出了在保证系统稳定性下的参数学习方法, 即通过最小化带非线性项约束的目标函数来估计参数  $\Theta$ .

$$\begin{cases} \min_{\Theta} J(\Theta) \\ \text{s.t.} \\ \begin{cases} A_k + (A_k)^T \prec 0, & \forall k \in \{1, \dots, K\} \\ b_k = -A_k \xi^* \end{cases} \end{cases} \quad (3)$$

式中,  $\prec 0$  表示矩阵的负定,  $\mathcal{T} = \sum_{m=1}^M T^m$  是全部训练数据的总数量. 目标函数  $J(\Theta)$  的选取有均方误差 (Mean square error, MSE) 和对数似然 (Log likelihood, LL) 两种.

$$J(\Theta) = \begin{cases} -\frac{1}{2\mathcal{T}} \sum_{m=1}^M \sum_{t=0}^{T^m} \|\dot{\xi}_{t,m} - \hat{f}(\xi_{t,m})\|^2, & \text{MSE} \\ -\frac{1}{\mathcal{T}} \sum_{m=1}^M \sum_{t=0}^{T^m} \ln p(\xi_{t,m}, \dot{\xi}_{t,m}|\Theta), & \text{LL} \end{cases} \quad (4)$$

式中, 对数  $\ln$  也可换成以其他常数为底的对数, 上式的约束是  $\hat{f}$  全局稳定收敛的充分条件, 这些条件由 Lyapunov 第二稳定性方法推导得到, 见定理 1.

**定理 1.** 如果存在径向无界且满足式 (5) 的 Lyapunov 函数  $V(\xi) : \mathbb{R}^d \mapsto \mathbb{R}$ , 则 DS 在  $\xi^*$  点处全局渐近稳定.

$$\begin{cases} V(\xi^*) = 0 \\ V(\xi) > 0, \quad \forall \xi \in \mathbb{R}^d \setminus \xi = \xi^* \\ \dot{V}(\xi^*) = 0 \\ \dot{V}(\xi) < 0, \quad \forall \xi \in \mathbb{R}^d \setminus \xi = \xi^* \end{cases} \quad (5)$$

SEDS 方法可以有效地从演示中估计出具有全局渐近稳定性的 DS, 是 LfD 领域中的常用方法, 但是该方法存在两个问题: 1) 用于估计初始 DS 的 GMM-GMR 方法需要人工指定混合分量个数  $K$ , 当状态  $\xi$  维度升高或演示轨迹变得复杂时, 通过经验确定超参数  $K$  将变得很困难; 2) 在 SEDS 中,  $V(\xi)$  是二次李雅普诺夫函数 (Quadratic Lyapunov function, QLF), 即  $V(\xi) = (\xi - \xi^*)^T (\xi - \xi^*)$ . 在几何学上, QLF 只允许  $L_2$  范数 (即  $\|\xi - \xi^*\|_2$ ) 距离是单调递减的轨迹<sup>[11]</sup>. 这导致了 SEDS 在包含高曲率或非单调 (暂时远离吸引子) 的高度非线性运动中

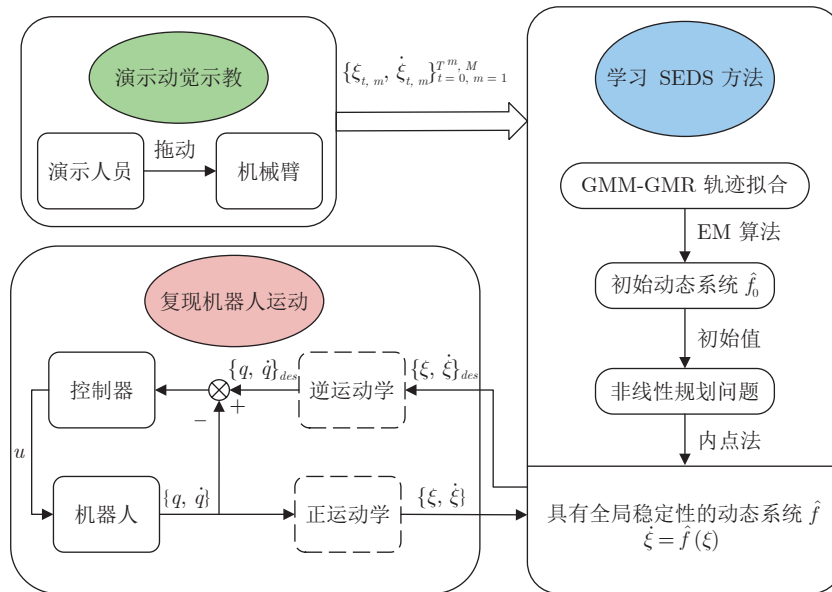


图 1 基于 SEDS 的机器人示教学习流程图

Fig.1 Flow chart of LfD based on SEDS

表现不佳, 不能同时保证稳定性和精度性.

针对问题 1), 本文利用贝叶斯非参数模型能根据数据自动确定最佳聚类个数的特性, 采用了 DP-GMM 拟合演示数据, 并推导 GMR 估计初始 DS; 针对问题 2), 本文选择参数化 QLF (Parametrized QLF, P-QLF) 来确保估计 DS 的稳定性, 即  $V(\xi) = (\xi - \xi^*)^T P(\xi - \xi^*)$ . 其中  $P$  的作用是将简单的 QLF 重塑为“椭圆”形式, 放松了约束条件, 从而能够精确复现不收缩 (高曲率且朝向目标具有非单调性) 的演示轨迹.

## 2 DPGMM-GMR 算法

DPGMM-GMR 算法由 DPGMM 的建模过程、利用变分推断进行模型求解和推导 GMR 估计初始 DS 三部分组成.

### 2.1 狄利克雷过程高斯混合模型

狄利克雷过程 (Dirichlet process, DP) 一般用  $DP(G_0, \alpha)$  表示, 其中  $G_0$  是连续的基分布,  $\alpha$  是集中参数, 用来控制从 DP 中抽取的分布的离散化程度<sup>[22]</sup>. 假设从 DP 中随机抽取一个样本分布  $G$ , 然后再从  $G$  中独立地抽取  $N$  个随机变量  $\{\theta_n^*\}_{n=1}^N$ :

$$\begin{cases} G | \{G_0, \alpha\} \sim DP(G_0, \alpha) \\ \theta_n^* | G \sim G, \quad k = 1, \dots, N \end{cases} \quad (6)$$

折棍过程 (Stick breaking process, SBP) 通常用于建模和推断 DP<sup>[23]</sup>. 给定两个无限随机变量  $v = \{v_k\}_{k=1}^\infty$  和  $\{\theta_k^*\}_{k=1}^\infty$ , 其中  $v_k$  服从 Beta 分布, 即  $v_k \sim \text{Beta}(1, \alpha)$ , 则 SBP 可以表示为:

$$G = \sum_{k=1}^\infty \pi_k(v) \delta_{\theta_k} \quad (7)$$

式中,  $\pi_k(v) = v_k \prod_{i=1}^{k-1} (1 - v_i)$ ,  $\sum_{k=1}^\infty \pi_k(v) = 1$ .

通过 SBP, 可以将独立于  $G_0$  绘制的原子  $\theta_k$  视为混合模型分量的分布参数, 该模型包含无限数量的分量, 并且每个分量的权重为  $\pi_k(v)$ , 从而建立了狄利克雷过程混合模型.

为了便于表述, 将全部演示数据表示为  $X = \{\xi_t, \dot{\xi}_t\}_{t=1}^T$ , 其中  $x_t = \{\xi_t, \dot{\xi}_t\}$ . 假设每个观测值  $x_t$  都对应一个参数  $\theta_t^*$ , 服从概率密度  $p(x_t | \theta_t^*)$ . 整个参数集  $\{\theta_t^*\}_{t=1}^T$  服从一个共同的 DP 先验, 由于分布  $G$  是离散的, 因此  $\{\theta_t^*\}_{t=1}^T$  可以划分到相同簇内, 以  $\pi_k(v)$  的概率共享相同的  $\theta_k$ . 然后引入指示变量  $Z = \{z_t\}_{t=1}^T$ , 也称作隐变量. 对于每个观测  $x_t$ , 都对应一个指示变量  $z_t$ , 用来指示数据  $x_t$  所属分量. 因此狄利克雷过程混合模型满足下式概率分布:

$$\begin{cases} v_k | \alpha \sim \text{Beta}(1, \alpha) \\ \theta_k | G_0 \sim G_0 \\ z_t | \pi(v) \sim \text{Cat}(\pi(v)) \\ x_t | z_{tk} = 1 \sim p(x_t | \theta_k) \end{cases} \quad (8)$$

式中,  $\text{Cat}(\cdot)$  表示分类分布.

在狄利克雷过程混合模型的基础上明确定义基分布  $G_0$  为 Normal-Wishart 分布, 即  $\theta_k \sim \mathcal{N}\mathcal{W}(\theta_0)$ , 就得到了 DPGMM, 这意味着每个高斯分量的均值和精度都服从一个独立的 Normal-Wishart 分布, 根据式 (8), DPGMM 的公式化表示如下:

$$\begin{cases} p(v | \alpha) = \prod_{k=1}^\infty \text{Beta}(v_k | 1, \alpha) \\ p(\mu, \Lambda) = \prod_{k=1}^\infty \mathcal{N}(\mu_k | m_0, (\beta_0 \Lambda_k)^{-1}) \times \mathcal{W}(\Lambda_k | W_0, \nu_0) \end{cases} \quad (9)$$

$$\begin{cases} p(Z | v) = \prod_{t=1}^T \prod_{k=1}^\infty \left( v_k \prod_{i=1}^{k-1} (1 - v_i) \right)^{z_{tk}} \\ p(X | Z, \mu, \Lambda) = \prod_{t=1}^T \prod_{k=1}^\infty \mathcal{N}(x_n | \mu_k, \Lambda_k^{-1})^{z_{tk}} \end{cases} \quad (10)$$

式中,  $\mu_k$  和  $\Lambda_k$  分别表示第  $k$  个高斯分量的均值和精度,  $m_0$  和  $\beta_0$  分别是  $\mu_k$  的高斯先验的均值和方差,  $W_0$  和  $\nu_0$  分别是  $\Lambda_k$  的 Wishart 先验的精度矩阵和自由度.

式 (9) 是给定的共轭先验分布, 式 (10) 是模型的似然函数. DPGMM 的概率结构如图 2 所示. 然而在无限维的设定下, 实际中模型不易处理. 因此本文采用一种常见的策略, 即基于 DP 的截断 SBP<sup>[24]</sup>. 预定义最大的混合分量数  $\hat{K}$ , 并且令  $q(v_{\hat{K}+1}) = 1$ . 即对于所有  $k > \hat{K}$ ,  $\pi_k(v)$  等于零.

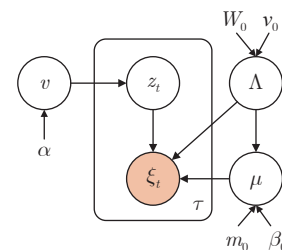


图 2 DPGMM 概率结构图  
Fig.2 DPGMM probability structure diagram

### 2.2 变分推断

第 2.1 节已为所有随机变量明确定义了共轭先验和模型的似然函数, 本节采用变分贝叶斯方法得出参数可靠的后验分布. 将隐变量和未知参数表示



为  $\Theta = \{v, Z, \mu, \Lambda\}$ , 超参数表示为  $\Xi = \{\alpha, m_0, \beta_0, W_0, \nu_0\}$ . 由于实际后验  $p(\Theta|X, \Xi)$  难以计算, 因此变分推断方法引入任意分布  $q(\Theta)$  以逼近  $p(\Theta|X, \Xi)$ . 在此假设下, 导出模型的对数似然值  $\ln p(X)$ , 即:

$$\ln p(X) = \mathcal{L}(q) + KL(q||p) \quad (11)$$

其中

$$\mathcal{L}(q) = \int q(\Theta) \ln \frac{p(X, \Theta|\Xi)}{q(\Theta)} d\Theta \quad (12)$$

$KL(q||p)$  表示变分后验  $q(\Theta)$  和实际后验  $p(\Theta|X, \Xi)$  之间的 KL(Kullback-Leibler) 散度, 刻画了两个分布之间的相似程度. 为降低求解难度, 一般采用平均场理论限制概率分布  $q(\Theta)$  的范围, 然后寻找这个类别中使得 KL 散度最小化的概率分布<sup>[25]</sup>. 因此需要将  $q(\Theta)$  分解为:

$$q(\Theta) = \prod_{k=1}^{\hat{K}} q(v_k) \prod_{k=1}^{\hat{K}} q(\mu_k, \Lambda_k) \prod_{t=1}^{\tau} q(z_t) \quad (13)$$

同时, 考虑 DPGMM 的共轭先验配置, 根据共轭的定义, 期望后验  $q(\Theta)$  与先验  $p(\Theta)$  具有相同的函数形式<sup>[26]</sup>, 可得:

$$\begin{cases} q(v_k) = \text{Beta}(v_k|\alpha) \\ q(\mu_k, \Lambda_k) = \mathcal{N}(\mu_k|m_k, (\beta_k \Lambda_k)^{-1}) \times \\ \quad \mathcal{W}(\Lambda_k|W_k, \nu_k) \end{cases} \quad (14)$$

指示变量  $z_t$  满足<sup>[27]</sup>:

$$q(z_{tk} = 1) \propto \exp \left\{ \sum_i^{k-1} \langle \ln(1 - v_i) \rangle + \langle \ln v_k \rangle + \frac{1}{2} \langle \ln |\Lambda_k| \rangle - \frac{D}{2} \ln 2\pi - \frac{1}{2} \langle (x_t - \mu_k) \Lambda_k (x_t - \mu_k)^T \rangle \right\} \quad (15)$$

式中,  $\langle \cdot \rangle$  表示数学期望.

采用变分贝叶斯期望最大化 (Expectation maximization, EM) 算法. 对该模型参数进行迭代求解. 该算法类似于 EM 算法, 在 E 步中, 计算指示变量的期望; 在 M 步中, 该期望值用于重新计算其他参数的变分分布.

期望通过变分推断得到 DPGMM 模型参数的后验分布后, 可以得到如下的模型联合概率密度:

$$p(\xi_t, \dot{\xi}_t) = \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \text{St}(\xi_t, \dot{\xi}_t | m_k, R_k, \nu_k + 1 - D) \quad (16)$$

式中,  $\text{St}(\cdot)$  是学生  $t$  分布, 各分布的权重  $\pi_k(v)$  按下式计算:

$$\langle \pi_k(v) \rangle = \langle v_k \rangle \prod_{i=1}^{k-1} (1 - \langle v_i \rangle) \quad (17)$$

式中,  $\langle v_k \rangle = 1/(1 + \alpha)$ .

### 2.3 GMR 估计初始 DS

与基于 GMM 的 GMR 一样, 先对式 (16) 的 DPGMM 中的均值和方差进行划分.

$$m_k = \begin{bmatrix} m_k^\xi \\ m_k^{\dot{\xi}} \end{bmatrix}, \quad R_k = \begin{bmatrix} R_k^\xi & R_k^{\xi\dot{\xi}} \\ R_k^{\dot{\xi}\xi} & R_k^{\dot{\xi}} \end{bmatrix} \quad (18)$$

给定输入位置  $\xi_t$  下输出速度  $\dot{\xi}_t$  的条件概率分布为:

$$p(\xi_t | \dot{\xi}_t, X) \sim \text{St} \left( \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \hat{m}_k, \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \hat{R}_k, \hat{\nu}_k \right) \quad (19)$$

其中

$$\hat{m}_k = m_k^\xi + R_k^{\dot{\xi}\xi} (R_k^\xi)^{-1} (\xi - m_k^\xi) \quad (20a)$$

$$\hat{R}_k = R_k^{\dot{\xi}} - R_k^{\dot{\xi}\xi} (R_k^\xi)^{-1} R_k^{\xi\dot{\xi}} \quad (20b)$$

$$\hat{\nu}_k = \nu_k + 1 - D \quad (20c)$$

然后将  $\dot{\xi}_t$  的条件概率期望作为预测值:

$$\hat{\xi} = \langle \dot{\xi} | \xi, X \rangle = \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \hat{m}_k \quad (21)$$

DPGMM 在拟合演示轨迹时, 自动确定最佳分量个数  $K$ , 其余不相关高斯分量权重趋向于 0, 因此可以丢弃  $\langle \pi_k(v) \rangle$  值过低的高斯分量来简化计算. 同时将式 (20a) 代入式 (21) 得到动态系统估计.

$$\begin{aligned} \hat{f}_0(\xi) &= \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \hat{m}_k \approx \\ &= \sum_{k=1}^{\hat{K}} \langle \pi_k(v) \rangle \hat{m}_k = \sum_{k=1}^{\hat{K}} h_k(\xi) (A_k \xi + b_k) \end{aligned} \quad (22)$$

式中,  $A_k = R_k^{\dot{\xi}\xi} (R_k^\xi)^{-1}$ .  $b_k = m_k^{\dot{\xi}} - A_k m_k^\xi$ ,  $h_k(\xi) = \langle \pi_k(v) \rangle$ .

### 3 基于 P-QLF 的稳定 DS 估计

第 2 节利用 DPGMM-GMR 从演示中估计初始动态系统  $\hat{f}_0$ , 同其他的回归算法一样, 训练得到的  $\hat{f}_0$  不具有全局稳定性. SEDS 方法针对该问题构造了具有稳定性约束的非线性规划问题, 如式 (3) 所示. 其不等式约束  $A_k + (A_k)^T \prec 0$  通过 QLF 推



数具有相同变化趋势, 这意味着 DPGMM 超参数具有可复用性, 一组超参数能够适用于相同场景下的不同演示数据集, 例如 LASA 数据集, 同一个机械臂操作空间下的不同任务演示等. DPGMM 超参数的不敏感性和可复用性使其在实际任务中的使用更加便捷.

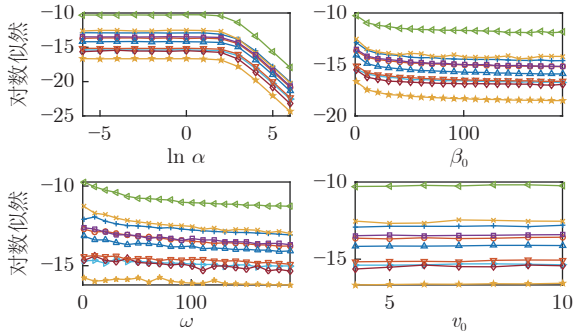


图 3 DPGMM 对数似然值关于超参数的变化趋势

Fig.3 The change trend of DPGMM log-likelihood value on hyperparameters

接着, 在整个 LASA 数据集上定量评估基于 VI 的 DPGMM-GMR 算法拟合初始 DS 的效果, 超参数均设置为  $\alpha = \beta_0 = \omega = 1$ ,  $\nu_0 = 4$ . 对比算法包括基于贝叶斯信息准则的 GMM-GMR (EM 算法)、狄利克雷分布 GMM-GMR (VI) 和 DPGMM-GMR, 其中贝叶斯信息准则高斯混合模型最大评估分量设定为 15 个, 其余方法的截断分量也设置为 15. 性能指标采用以下形式均方根误差 (Root mean square error, RMSE):

$$RMSE = \frac{1}{M} \sum_{t=1}^{\mathcal{T}} \left\| \hat{\xi}_t - f(\xi_t) \right\| \quad (27)$$

式中,  $M$  和  $\mathcal{T}$  分别表示一组演示中轨迹的条数和所有的轨迹点个数.

## 4.2 $i$ -SEDS 学习方法性能分析

本节  $i$ -SEDS 的性能分析同样在数据集 LASA 上进行, 如果算法能够使复现的轨迹紧密 (准确) 地跟随演示轨迹, 则表明该算法可以令人满意地解决稳定性和精度的难题.

由图 4 可以看出, 所有复现轨迹紧密跟随了演示轨迹, 仿真结果说明了  $i$ -SEDS 的有效性. 图 5 给出了 SEDS 和  $i$ -SEDS 在部分不收缩轨迹上的复现结果. 可以看出, SEDS 生成的轨迹完全偏离了演示, 这是由于遵循 QLF 稳定性结果导致的. 图 5 定性说明了  $i$ -SEDS 可以较好地解决 SEDS 稳定性和精度的难题.

为了定量分析该方法的有效性, 使用文献 [28] 提出的扫描误差区域 (Swept error area, SEA). SEA 方法计算了演示轨迹和复现轨迹之间的面积, 计算公式如下:

$$\mathcal{E} = \frac{1}{M} \sum_{m=1}^M \sum_{t=0}^{\mathcal{T}^m} \mathcal{A}(\hat{\xi}_{t,m}, \hat{\xi}_{t+1,m}, \xi_{t,m}, \xi_{t+1,m}) \quad (28)$$

式中,  $\hat{\xi}_{t,m}$  和  $\xi_{t,m}$  分别表示复现点和演示点. 如图 6 所示, 函数  $\mathcal{A}(\cdot)$  计算由 4 个点生成的封闭四边形的面积, 这反映了复现和演示的一致性.

考虑到原始 SEDS 中 2 种不同的目标函数 MSE 和 LL, 以及  $i$ -SEDS 在学习不收缩演示轨迹时具有明显优势. 因此利用 SEA 方法定量分析  $i$ -SEDS (MSE)、 $i$ -SEDS (LL)、SEDS (MSE)、SEDS (LL) 四种方法分别在收缩和不收缩演示轨迹上的性能. 从 LASA 数据集中取收缩轨迹和不收缩轨迹

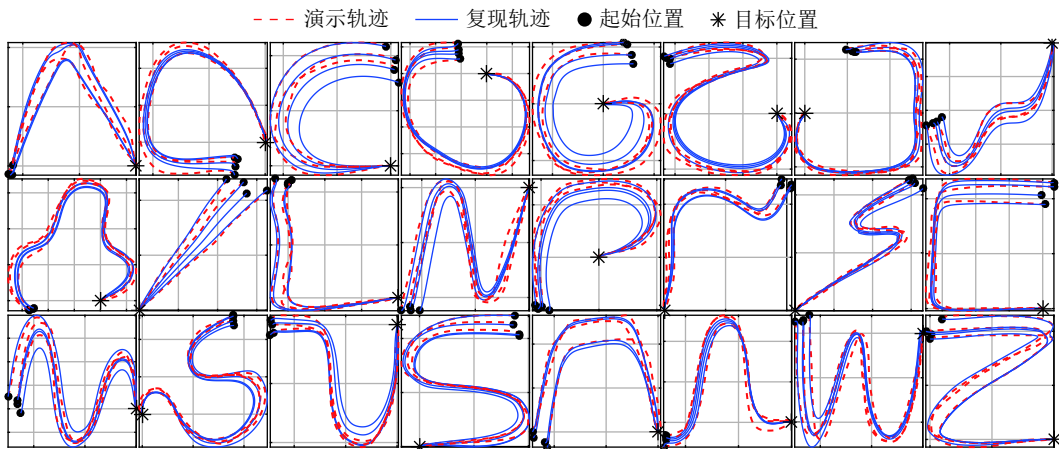


图 4  $i$ -SEDS 方法在 LASA 数据集上的复现效果

Fig.4 The reproductions of  $i$ -SEDS method on LASA dataset

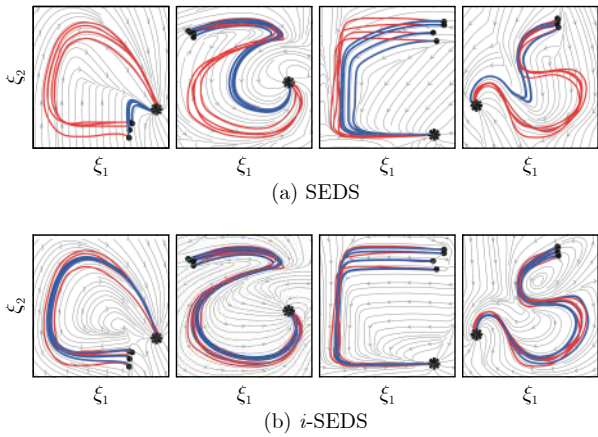


图 5 SEDS 和 *i*-SEDS 在不收缩轨迹上的复现结果  
 Fig.5 *i*-SEDS and SEDS reproductions on the non-contractive demonstrations

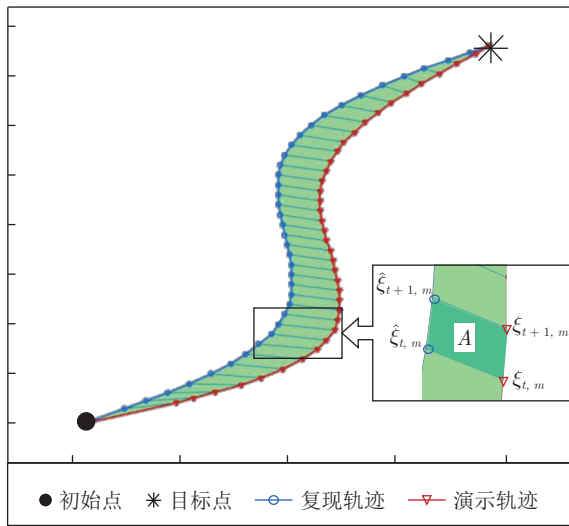


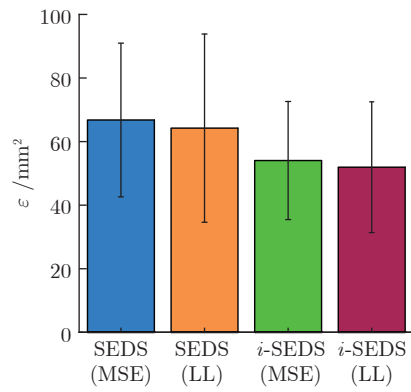
图 6 扫描误差区域 (SEA) 示意图  
 Fig.6 Schematic diagram of Swept error area (SEA)

各 10 组, 分别使用上述 4 种方法学习稳定 DS, 计算相应的 SEA, 结果如结果如图 7 所示. 从图 7(a) 可以看出, 四种方法的 SEA 在收缩轨迹上差别并不明显, *i*-SEDS 仅略优于 SEDS 方法. 但在不收缩轨迹组的图 7(b) 中, *i*-SEDS 方法的复现精度要明显优于 SEDS 方法. 基于 LL 的方法精度要略高于 MSE, 但 MSE 的计算速度更快, 因此, 在实际使用中可以根据需求进行选择.

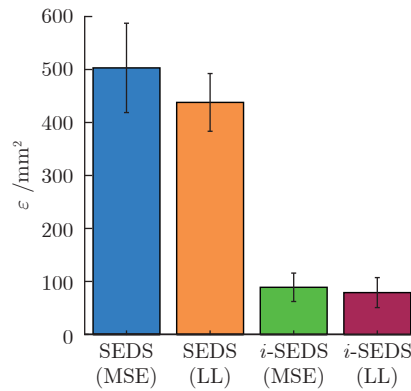
### 4.3 机器人实验

为了验证本文方法的有效性, 在 7 自由度的机械臂 Franka-panda 机器人上进行了实验验证.

机器人实验是在操作空间中学习物块搬运任务, 如图 8 所示. 这是一个点到点运动的任务, 机械



(a) 收缩轨迹上的 SEA  
 (a) SEA of the contractive demonstrations



(b) 不收缩轨迹上的 SEA  
 (b) SEA of the non-contractive demonstrations

图 7 4 种方法在收缩和不收缩轨迹上的 SEA 分析  
 Fig.7 Performance analysis of four methods on contraction and non-contraction demonstrations with SEA

臂需要学会如何从机械臂初始位置到达物块放置位置, 然后拾取物块将其放入目标容器中, 整个过程由拾取物块-放置物块 2 部分组成, 并利用 Kinect 摄像头对物体进行跟踪.

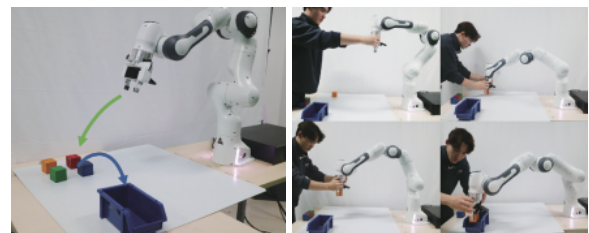


图 8 物品搬运任务示教流程图  
 Fig.8 The process of object transport task teaching

1) 示教: 首先设置 panda 机械臂为示教模式, 再由演示人员拖动机械臂完成拾取物块-放置物块 2 个步骤, 整个示教过程如图 8 所示. 然后不断调整物块起始位置重复上述示教过程共 8 次, 同时以 10 Hz 频率记录机械臂末端位置和速度, 获取 8 条演示轨



迹数据, 记作  $\{\xi_1, \xi_2, \xi_3, \dot{\xi}_1, \dot{\xi}_2, \dot{\xi}_3\}_{t=0, m=1}^{T^m, M=8}$ .

2) 学习: 在对演示轨迹建模前, 需要对数据进行一些预处理. a) 轨迹平滑: 使用移动均值滤波对采集到的演示数据进行平滑处理. b) 任务分割: 为了降低任务的学习成本, 将物品搬运任务分解成拾取物块-放置物块 2 个子任务. 根据 2 个子任务切换过程中机械臂末端速度为零这一特点将演示轨迹分成 2 段, 然后各自进行学习. c) 目标位置原点化: 为保证动态系统的参数不会随目标位置的变化而变化, 需要将参考坐标系的原点设在目标位置处. 2 个子任务的演示轨迹都转化为以原点为终点的轨迹. 预处理后的演示轨迹如图 9 所示.

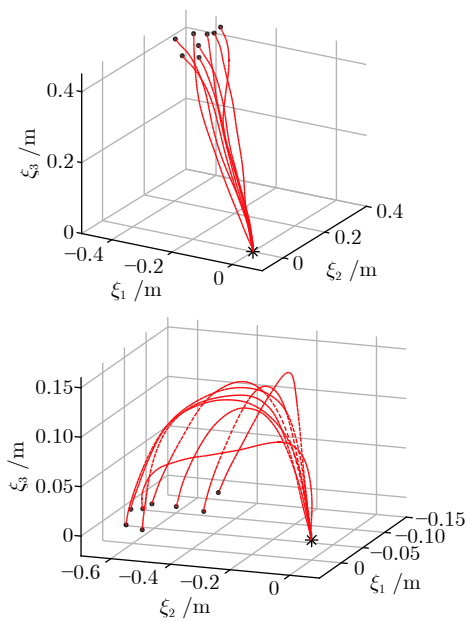


图 9 预处理后的物品搬运任务的演示轨迹  
Fig.9 Preprocessed demonstrations of object transport task

使用  $i$ -SEDS 方法对预处理后的演示数据进行建模, 用于 2 个子任务学习的超参数均设置为  $\alpha = 0.1$ ,  $\beta_0 = \omega = 1$ ,  $\nu_0 = 6$ .

3) 复现. 机械臂从初始位置开始, 通过模型不断迭代得到当前位置下的期望速度  $\{\dot{\xi}_1, \dot{\xi}_2, \dot{\xi}_3\}_{des}$ , 将其作为逆运动学的输入驱动机械臂按照期望轨迹到达目标位置. 记录每次生成的轨迹, 获得的结果如图 10 所示. 所有复现轨迹都与演示具有相似的位置和速度轮廓, 说明两者均遵循相似的动力学.

a) 目标位置的变化. 图 11 显示了模型对环境变化的鲁棒性, 其中虚线表示原始复现轨迹, 实线表示在  $t = 4$  s 时对目标位置进行转移后由动态系统重新规划的轨迹, 可以看到该模型可以快速适应

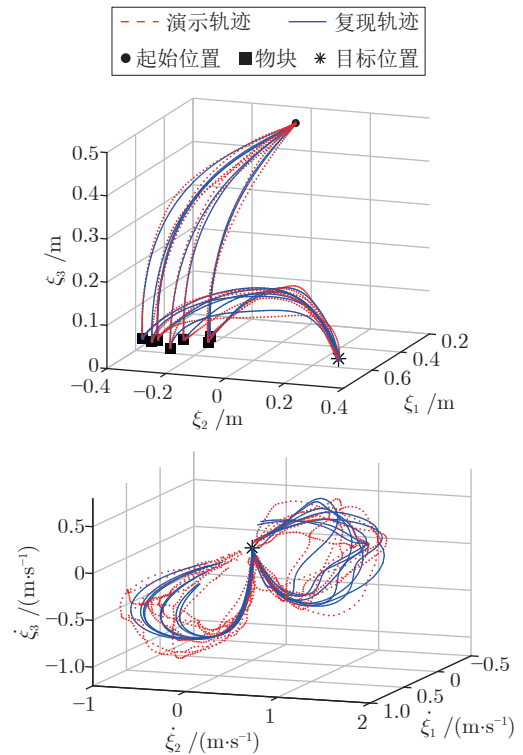


图 10 物品搬运任务复现位置和速度轨迹  
Fig.10 Position and velocity trajectories of reproduction in object transport task

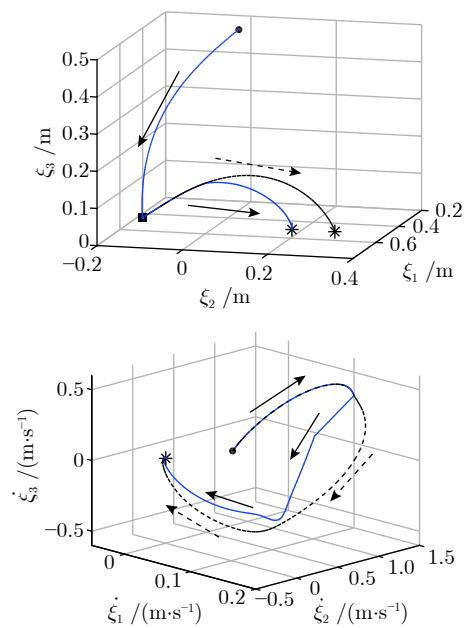


图 11 机械臂运行时改变目标位置后的轨迹变化  
Fig.11 Trajectory changes after moving the target container when the robot arm is running

目标位置的变化, 重新规划轨迹使机械臂到达目标.

b) 抗扰能力. 图 12 显示了模型对外界扰动

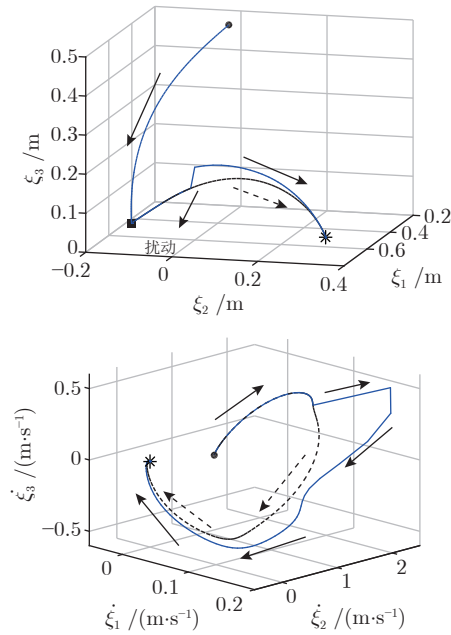


图 12 对机械臂施加外界扰动时的轨迹变化

Fig. 12 Trajectory changes when external disturbance is applied to the robot arm

的鲁棒性, 在  $t = 4$  s 时, 施加扰动使机械臂脱离原始轨迹, 其中虚线表示原始轨迹, 实线表示机械臂在脱离原始轨迹后模型重新规划的轨迹, 可以看到机械臂在受到外界扰动后仍能平稳到达目标位置, 因此该模型具备一定的抗扰动能力。

上述机器人实验展示了  $i$ -SEDS 方法在确保稳定性的前提下, 仍具有良好的复现精度以及泛化能力和抗扰动能力。

## 5 结束语

本文分析了 SEDS 方法中存在的缺陷, 并给出了合理的解决方案. 首先利用 DPGMM 对演示数据进行拟合, 并使用变分推断求解模型参数, 推导 GMR 拟合初始 DS. 解决了基于 EM 算法的 GMM 难以确定混合分量个数的难题. 然后利用 P-QLF 推导新的宽稳定性约束取代原有约束, 最后结合目标函数优化得到能精确复现的稳定 DS. 在 LASA 数据集中验证了 DPGMM 超参数的性质和其推导的 GMR 算法的性能, 以及  $i$ -SEDS 方法在不收缩轨迹上的优越性能. 最后通过 Franka-panda 协作机器人上的实验, 验证了本文方法在实际场景的有效性. 在未来工作中, 将降低本文方法模型的复杂度。

## References

1 Zeng Chao, Yang Chen-Guang, Li Qiang, Dai Shi-Lu. Research progress on human-robot skill transfer. *Acta Automatica Sinica*, 2019, **45**(10): 1813–1828

(曾超, 杨辰光, 李强, 戴诗陆. 人-机器人技能传递研究进展. *自动化学报*, 2019, **45**(10): 1813–1828)

2 Liu Nai-Jun, Lu Tao, Cai Ying-Hao, Wang Shuo. A review of robot manipulation skills learning methods. *Acta Automatica Sinica*, 2019, **45**(3): 458–470  
(刘乃军, 鲁涛, 蔡莹皓, 王硕. 机器人操作技能学习方法综述. *自动化学报*, 2019, **45**(3): 458–470)

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

4 Ravichandar H, Polydoros A S, Chernova S, Billard A. Recent advances in robot learning from demonstration. *Annual Review of Control, Robotics, and Autonomous Systems*, 2020, **3**(1): 297–330

5 Brock O, Khatib O. *Elastic Strips: A Framework for Integrated Planning and Execution*. London: Experimental Robotics VI, 2000. 329–338

6 Billard A, Hayes G. DRAMA, A connectionist architecture for control and learning in autonomous robots. *Adaptive Behavior*, 1999, **7**(1): 35–63

7 Ijspeert A J, Nakanishi J, Hoffmann H, Pastor P, Schaal Stefan. Dynamical movement primitives: Learning attractor models for motor behaviors. *Neural Computation*, 2012, **25**(2): 328–373

8 Gribovskaya E, Khansari-Zadeh S M, Billard A. Learning nonlinear multi-variate dynamics of motion in robotic manipulators. *International Journal of Robotics Research*, 2011, **30**(1): 80–117

9 Khansari-Zadeh S M, Billard A. Learning stable nonlinear dynamical systems with Gaussian mixture models. *IEEE Transactions on Robotics*, 2011, **27**(5): 943–957

10 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

11 Neumann K, Steil J J. Learning robot motions with stable dynamical systems under diffeomorphic transformations. *Robotics and Autonomous Systems*, 2015, **70**: 1–15

12 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

13 Jin S K, Wang Z Y, Ou Y S, Feng W. Learning accurate and stable dynamical system under manifold immersion and submerision. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, **30**(12): 3598–3610

14 Schwarz G. Estimating the dimension of a model. *Annals of Statistics*, 1978, **6**(2): 461–464

15 Chatzis S P, Korkinof D, Demiris Y. A nonparametric Bayesian approach toward robot learning by demonstration. *Robotics and Autonomous Systems*, 2012, **60**(6): 789–802

16 Blei D M, Jordan M I. Variational methods for the Dirichlet process. In: *Proceedings of the twenty-first international conference on Machine learning*. New York, USA: Association for Computing Machinery, 2004. 12

17 Müller P, Quintana F A. Nonparametric Bayesian data analysis. *Statistical Science*, 2004, **19**(1): 95–110

18 Blei D M., Kucukelbir A, McAuliffe J D. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 2017, **112**(518): 859–877

19 Figueroa N, Billard A. A physically-consistent Bayesian non-parametric mixture model for dynamical system learning. In: *Proceedings of the Machine Learning Research*. Zurich, Switzerland: CoRL, 2018. 927–946

20 Deisenroth M P, Rasmussen C E, Peters J. Gaussian process dy-

- namatic programming. *Neurocomputing*, 2009, **72**(7): 1508–1524
- 21 Vijayakumar S, D' Souza A, Schaal S. Incremental online learning in high dimensions. *Neural Computation*, 2005, **17**(12): 2602–2634
  - 22 Ferguson T S. A Bayesian analysis of some nonparametric problems. *Annals of Statistics*, 1973, **1**(2): 209–230
  - 23 Ahmed A, Xing E P. Dynamic non-parametric mixture models and the recurrent Chinese restaurant process: With applications to evolutionary clustering. In: Proceedings of the 2008 SIAM International Conference on Data Mining. Atlanta, USA: 2008. 219–230
  - 24 Blei D M, Jordan M I. Variational inference for Dirichlet process mixtures. *Bayesian Anal.*, 2006, **1**(1): 121–143
  - 25 Chatzis S P, Kosmopoulos D I, Varvarigou T A. Signal modeling and classification using a robust latent space model based on  $t$  distributions. *IEEE Transactions on Signal Processing*, 2008, **56**(3): 949–963
  - 26 Bishop C M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006. 461–522
  - 27 Zhu J L, Ge Z Q, Song Z H. Variational Bayesian Gaussian mixture regression for soft sensing key variables in non-Gaussian industrial processes. *IEEE Transactions on Control Systems Technology*, 2017, **25**(3): 1092–1099
  - 28 Khansari-Zadeh S M, Billard A. Learning control Lyapunov function to ensure stability of dynamical system-based robot reaching motions. *Robotics and Autonomous Systems*, 2014, **62**(6): 752–765

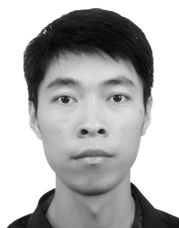


**金聪聪** 浙江工业大学信息工程学院硕士生. 主要研究方向为人体动作识别和机器人技能学习.

E-mail: jcc19960602@gmail.com

**(JIN Cong-Cong** Master student at the College of Information Engineering, Zhejiang University of Technology. His research interest covers human motion recognition and robot skills learning.)

His research interest covers human motion recognition and robot skills learning.)



**刘安东** 浙江工业大学信息工程学院副教授. 主要研究方向为模型预测控制, 网络化控制系统和移动机器人.

E-mail: lad@zjut.edu.cn

**(LIU An-Dong** Associate professor at the College of Information Engineering, Zhejiang University of Technology. His research interest covers model predictive control, networked control systems and mobile robots.)

His research interest covers model predictive control, networked control systems and mobile robots.)



**LIU Steven** 凯泽斯劳滕工业大学电气与计算机工程系教授. 主要研究方向为机电和电力系统控制, 机器人技术, 网络控制和基于模型的故障诊断.

E-mail: sliu@eit.uni-kl.de

**(LIU Steven** Professor in the Department of Electrical and Computer Engineering, University of Kaiserslautern. His research interest covers control of mechatronic and power systems, robotics, networked control and model-based fault diagnosis.)

His research interest covers control of mechatronic and power systems, robotics, networked control and model-based fault diagnosis.)



**张文安** 浙江工业大学信息工程学院教授. 主要研究方向为多源信息融合和机器人技能学习. 本文通信作者.

E-mail: wazhang@zjut.edu.cn

**(ZHANG Wen-An** Professor at the College of Information Engineering, Zhejiang University of Technology. His research interest covers information fusion estimation and robot skill learning. Corresponding author of this paper.)

His research interest covers information fusion estimation and robot skill learning. Corresponding author of this paper.)