

DOI: 10.19650/j.cnki.cjsi.J2210522

基于狄利克雷过程聚类的机器人演示学习研究

吴晓敏, 贺苗, 刘瞰东, 张馨月, 邵桂芳
(厦门大学航空航天学院 厦门 361005)

摘要:针对演示学习中高斯混合模型参数估计效率低,泛化能力不足的问题,提出一种基于狄利克雷过程聚类和高斯混合模型的复合动态运动基元算法。为实现高斯混合模型参数的实时估计,使用基于距离阈值的狄利克雷聚类算法进行演示轨迹点在线聚类,并引入Welford公式更新参数以提高参数估计效率。获得轨迹分布特征后,使用动态运动基元进行高斯混合回归轨迹的编码,以提高轨迹泛化能力。为了验证算法的有效性,引入了轨迹可达性和相似性指标评价算法的学习泛化能力,设计了基于手写体字母轨迹和机器人动觉示教的演示学习实验。实验结果表明,所提复合动态运动基元算法参数估计平均时间仅0.052 ms,具备快速轨迹复现和泛化能力。

关键词: 演示学习;狄利克雷过程;高斯混合模型;动态运动基元

中图分类号: TH241.2 **文献标识码:** A **国家标准学科分类代码:** 510.8050

Study of robot demonstration learning based on the Dirichlet process clustering

Wu Xiaomin, He Miao, Liu Tundong, Zhang Xinyue, Shao Guifang
(School of Aerospace Engineering, Xiamen University, Xiamen 361005, China)

Abstract: A composite dynamic movement primitives algorithm based on Dirichlet process clustering and Gaussian mixture model is proposed to address the problems of low efficiency of parameter estimation and insufficient generalization ability in demonstration learning. To achieve the real-time estimation of Gaussian mixture model parameters, the Dirichlet clustering algorithm based on the distance threshold is used to perform online clustering of demo trajectory points, and the Welford formula is introduced to update the parameters to improve the efficiency of parameter estimation. After obtaining the trajectory distribution characteristics, the Gaussian mixture regression trajectories are encoded by using the dynamic movement primitives to improve the trajectory generalization. To evaluate the effectiveness of the algorithm, trajectory reachability and similarity metrics are introduced to evaluate the learning generalization ability of the algorithm, and demonstration learning experiments based on handwritten letter trajectories and robot kinesthetic demonstrations are designed. Experimental results show that the average parameter estimation time of the proposed composite dynamic movement primitive algorithm is only 0.052 ms, which has the ability of fast trajectory reproduction and generalization.

Keywords: demonstration learning; Dirichlet process clustering; Gaussian mixture model; dynamic movement primitives

0 引言

工业机器人具有成本低、效率高、安全性好等优点,常被用于特定的和结构化的环境中执行任务。然而,不同场景下的机器人示教编程存在过程繁琐、示教器种类多和学习成本高等问题。因此,机器人技能快速学习与泛化受到了越来越多的关注。目前,强化学习、模仿学习和深度学习^[1-2]等机器学习技术已成功应用于机器人的

技能学习。在各种学习方法中,从演示中学习(learning from demonstration, LfD)已被证明是一种有效的方法,可以轻松地将从人类转移到机器人^[3-5]。

尽管演示学习已经得到了大量的应用,但对于实时性要求很高的工业应用中,技能的编码效率、轨迹复现精度和泛化能力仍然面临很大的挑战。因此,本文主要研究演示学习中的轨迹编码和泛化方法,以期提高从人类示教到机器人自主完成演示技能学习过程的效率和复现精度。

演示学习的核心问题是演示轨迹的编码与泛化,主要分为基于动态系统的方法和基于统计学习的方法^[3]。基于动态系统方法的典型代表为动态运动基元(dynamic movement primitives, DMP)^[6]。DMP 是一种基于二阶动力学系统的运动基元算法,由一个线性弹簧阻尼系统和一个可学习的非线性项组成。它可以保证泛化轨迹以原有轨迹特征收敛到给定目标位置。该算法通过基函数的线性组合组成的非线性项来拟合轨迹,其参数通常使用局部加权最小二乘法^[7]、递归最小二乘法^[8]或局部加权映射回归法^[9]进行求解。DMP 能够适应不确定性和扰动,以及空间和时间上的扩展,已被成功地应用于各种轨迹规划和避障任务。然而,DMP 是一种单样本学习算法,单个演示轨迹具有高度随机性和不确定性,因此很难直接用于演示轨迹学习。为了学习多演示轨迹的特征,Ginesti 等^[10]通过引入额外变量对多个相似轨迹进行建模,以学习多条演示轨迹的特征,但该方法对轨迹的要求高,需要保证多条演示轨迹具有相同的起点和终点。

基于统计的学习方法能充分利用多组演示轨迹进行建模^[7],如高斯混合模型与回归^[11](Gaussian mixture model-gaussian mixture regression, GMM-GMR)及其相关改进算法等^[12-14]。其中,Calinon 等^[11]提出的 GMM-GMR 作为一种可以处理多条演示轨迹的学习算法,其基本思想是将人类技能编码为聚类问题,使用期望最大化(expectation-maximization, EM)^[15]迭代学习算法以 GMM 的形式对数据集的联合概率进行建模,并使用高斯分布的协方差来描述不同阶段示教轨迹的不确定性。文献[16]提出了一种增量式的 GMM-GMR 学习算法,渐进式地教授类人机器人任务技能。同时 GMM-GMR 也被广泛应用于其它机器人演示学习任务中,如协同搬运^[17]、倒饮料^[18]和烹饪^[19]等。但是,GMM 的参数估计需要预先指定高斯分布的数量和初始解,通常使用 K-means 对数据进行预处理,将聚类中心作为初始高斯均值,然后使用 EM 算法进行迭代求解。由于 EM 算法每次迭代需使用所有演示轨迹点,导致了迭代时间会随着轨迹长度增加而增加。

综上所述,尽管上述算法均已成功应用于不同的演示学习任务,但算法的学习过程繁琐且低效,需要根据不同任务预先设定专家知识,难以简单地应用于实际场景。因此,本文提出一种基于狄利克雷过程聚类(Dirichlet process-means, DP-means)和 GMM 的复合动态运动基元算法。首先,为解决 GMM 参数估计耗时问题,通过设置狄利克雷聚类算法的距离阈值,在采集演示轨迹的过程中,使用 Welford 公式增量式更新 GMM 参数,当轨迹点和聚类中心的距离大于距离阈值时产生新的高斯分布,演示完成的同时即可获得 GMM。其次,引入 DMP 对 GMR 生成的轨迹进行学习泛化,保证了泛化轨迹的特征

不变性和目标可达性。最后,开展了基于手写体字母轨迹和机器人动觉示教的演示学习实验,通过仿真和实际实验验证了所提复合动态运动基元方法的快速轨迹复现和泛化能力。

1 高斯混合模型参数辨识与回归

1.1 高斯混合模型

GMM^[20]作为一种经典的基于统计理论的时间序列和运动表示方法,被广泛运用于机器人演示轨迹编码中。GMM 是 K 个高斯分布的线性组合,可以用来表征机器人运动轨迹的变化特征。

假设示范轨迹数据集包含 M 条轨迹,可以表示为 $\mathbf{X} = \{\mathbf{X}^1, \dots, \mathbf{X}^M\}$ 。其中,第 m 条轨迹 $\mathbf{X}^m \in \mathbf{X}$, $m \in \{1, \dots, M\}$ 包含 T_m 个时刻的机器人末端位置序列,即 $\mathbf{X}^m = [\mathbf{X}^m(1), \dots, \mathbf{X}^m(T_m)]$ 。对于一个有限的 GMM,通常假设数据服从如下概率分布:

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

式中: \mathbf{x} 代表状态量; K 表示 GMM 中子分布的个数; π_k 表示混合系数(即权重)且满足 $0 \leq \pi_k \leq 1$ 和 $\sum_{k=1}^K \pi_k = 1$; $N(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ 代表高斯分布, $\boldsymbol{\mu}_k$ 和 $\boldsymbol{\Sigma}_k$ 表示第 k 个分模型的均值和协方差。

GMM 中,待估计参数为 $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$ 。由于事先不知道每个观测数据点是属于哪个子分布,无法像单个高斯模型那样使用最大似然法来求导得到参数,因此通常采用 EM 算法^[21]来进行参数的估计。EM 算法是一种广泛适用于计算不完整数据的极大似然估计算法^[22],能够迭代更新 GMM 的参数。然而,EM 算法每次迭代需使用所有数据点,计算效率低,并且求解性能高度依赖初始解的选择。为了避免陷入局部最优解,通常需要指定一个较好的初始解。对于一组观测数据 $\{\mathbf{x}_i\}_{i=1}^N$,其中 $N = \sum_{m=1}^M T_m$,一般使用 K-means 聚类算法进行混合系数 π_k 、均值 $\boldsymbol{\mu}_k$ 和协方差矩阵 $\boldsymbol{\Sigma}_k$ 的初始化,然后使用 EM 算法进行迭代求解。EM 算法首先计算联合分布的条件概率期望,通常称为 E 步,计算公式为:

$$\gamma_{z_{ik}} = \frac{\pi_k N(\mathbf{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{k=1}^K \pi_k N(\mathbf{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} \quad (2)$$

E 步计算完成后,根据 $\gamma_{z_{ik}}$ 重新估计分布参数,称为 M 步,计算公式为:

$$\boldsymbol{\mu}_k^{\text{new}} = \frac{1}{T_k} \sum_{i=1}^N \gamma_{z_{ik}} \mathbf{x}_i \quad (3)$$

$$\boldsymbol{\Sigma}_k^{\text{new}} = \frac{1}{T_k} \sum_{i=1}^N \gamma_{z_{ik}} (\mathbf{x}_i - \boldsymbol{\mu}_k^{\text{new}})(\mathbf{x}_i - \boldsymbol{\mu}_k^{\text{new}})^T \quad (4)$$

$$\pi_k^{\text{new}} = \frac{T_k}{N} \quad (5)$$

式中: $T_k = \sum_{i=1}^N \gamma_{z_{ik}}$, $\gamma_{z_{ik}}$ 表示 \mathbf{x}_i 属于第 k 个高斯分布的概率。获得新参数后,利用新参数继续执行式(2)~(5)的 E 步和 M 步,如此重复进行 GMM 的参数估计,直到算法收敛,获得近似最优解。由于 EM 算法每次迭代都会使用全部轨迹点,迭代时间会随着轨迹点数量的增加而增加,因此,本文将针对 GMM 参数估计效率低的问题进行改进。

1.2 基于 DP-means 的在线 GMM 参数估计算法

1) DP-means 在线聚类算法

对于 GMM 的参数估计问题,除了上述的 EM 算法外,还可以采用非参数贝叶斯统计中的重要方法——狄利克雷过程^[23](Dirichlet process, DP)。DP 在参数估计的过程中,可以很好地解决 EM 算法重复使用所有轨迹点迭代导致计算效率低的问题。文献[24]给出的 DP-means 聚类算法伪代码如下。

算法1 DP-means 聚类算法

输入:数据集 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, 聚类惩罚参数 λ

输出:点集 I_1, I_2, \dots, I_K 和聚类个数 K

1. 初始化参数 $K = 1, I_1 = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, 聚类中心 $\boldsymbol{\mu}_1$

2. 初始化聚类指标 $z_i = 1, i = 1, \dots, N$

3. 循环开始,直到收敛:

对于每个数据点 \mathbf{x}_i , 计算数据点到聚类中心的欧式距离

$$d_{ic} = \|\mathbf{x}_i - \boldsymbol{\mu}_c\|, c = 1, \dots, K$$

如果 $\min_c d_{ic} > \lambda$,

$$\text{令 } K = K + 1, z_i = K, \boldsymbol{\mu}_K = \mathbf{x}_i$$

否则,令 $z_i = \operatorname{argmin}_c d_{ic}$

根据 $z_1, \dots, z_K: I_j = \{\mathbf{x}_i | z_i = j\}$, 生成点集 I_1, I_2, \dots, I_K

对于每个点集 I_j , 计算 $\boldsymbol{\mu}_j = \frac{1}{|I_j|} \sum_{\mathbf{x} \in I_j} \mathbf{x}$

从算法1可以看出,DP-means 仅需要设置一个惩罚参数 λ 即可实时输出点集 I_1, I_2, \dots, I_K 和聚类个数 K , 解决了 K-means 算法需要预先指定聚类个数的问题,在应用过程中实现了在线聚类。与依赖于数据初始聚类的 K-means 不同,DP-means 算法依赖于数据点的处理顺序,很适合处理时间序列数据,这使得它很适合本文研究的机器人在线演示学习问题。

2) GMM 在线参数估计算法

由于 DP-means 聚类算法依赖于距离度量,能够对演示学习的数据进行在线聚类,非常适合于时间序列数据的在线聚类,而 GMM 的求解,本质上也是对数据进行软聚

类。因此,可以将 DP-means 计算得到的每类点集归于某个高斯子分布,然后采用 Welford 公式进行增量式计算均值和协方差,从而更新 GMM 的参数。具体计算公式如下:

$$\boldsymbol{\mu}_t = \frac{t-1}{t} \boldsymbol{\mu}_{t-1} + \frac{1}{t} \mathbf{x}_t \quad (6)$$

$$\boldsymbol{\Sigma}_t = \frac{t-2}{t-1} \boldsymbol{\Sigma}_{t-1} + \frac{1}{t-1} (\mathbf{x}_t - \boldsymbol{\mu}_t)(\mathbf{x}_t - \boldsymbol{\mu}_{t-1})^T \quad (7)$$

式中: $\mathbf{x}_t, \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t$ 分别表示 t 时刻的数据点、均值和协方差。利用此方法可在线求得 GMM 参数的近似解。

在演示学习的应用场景中,基于 DP-means 的 GMM 在线参数估计能够在演示完成后直接生成对应 GMM 的参数,实现在线演示学习,算法伪代码如下。

算法2 基于 DP-means 的在线 GMM 参数估计算法

输入:每个时间步的数据点 \mathbf{x}_t , 长度阈值 λ

输出:GMM 参数 $K, \pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k (k = 1, \dots, K)$

1. 初始化演示次数 $D = 0$, 数据点总数 $N = 0$, GMM 参数 $K = 1, \boldsymbol{\mu}_1 = \mathbf{x}_0, \boldsymbol{\Sigma}_1 = 0, \pi_1 = 1$, 高斯分布点集长度 $n_1 = 0$

2. 循环开始,直到完成所有演示:

$$D = D + 1$$

循环开始,直到完成一次演示:

计算当前 t 时刻数据点总个数 $N = N + 1$

计算 $d_{ik} = \|\mathbf{x}_t - \boldsymbol{\mu}_k\|, k = 1, \dots, K$

如果 $\min_k d_{ik} > \lambda$ 且 $D = 1$:

$$K = K + 1, \boldsymbol{\mu}_K = \mathbf{x}_t, \boldsymbol{\Sigma}_K = 0, \pi_K = \frac{1}{N}, n_K = 1$$

否则:

$$k = \operatorname{argmin}_k d_{ik}$$

$$n_k = n_k + 1$$

$$\boldsymbol{\mu}_k^{\text{new}} = \frac{n_k - n_k^{\text{old}}}{n_k} \boldsymbol{\mu}_k^{\text{old}} + \frac{n}{n_k} \mathbf{x}_t$$

$$\boldsymbol{\Sigma}_k^{\text{new}} = \frac{n_k - 2}{n_k - 1} \boldsymbol{\Sigma}_k^{\text{old}} + \frac{1}{n_k - 1} (\mathbf{x}_t - \boldsymbol{\mu}_k^{\text{new}})(\mathbf{x}_t - \boldsymbol{\mu}_k^{\text{old}})^T$$

$$\text{令 } \pi_k = \frac{n_k}{N}, k = 1, \dots, K$$

基于 DP-means 的参数估计方法运行过程如图1所示,仅靠长度阈值 λ 就可以输出聚类个数 K 和每个类别的点集,再结合 Welford 公式即得到 GMM 的参数估计。使得能在获取演示轨迹的同时进行在线参数估计,演示学习结束的同时就能够自动生成相应的 GMM,达到了在线演示学习的效果。同时,算法也可运行在离线模式下,即在演示结束后再进行参数估计。与传统 EM 算法相比,基于 DP-means 的 GMM 参数估计不重复使用演示数据,计算效率高,实际操作和交互十分友好,更加贴合实际应用需求。

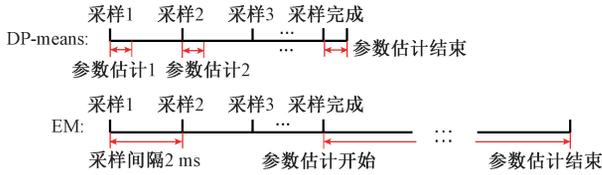


图1 DP-means 和 EM 算法参数估计过程

Fig.1 DP means and EM algorithm parameter estimation process

1.3 高斯混合回归

GMM 的参数估计完成后,可以通过 GMR 进行回归处理,来获得期望的运动轨迹。对于每个高斯分布,将均值和协方差进行分块处理:

$$\mu_k = \{u_k^I, u_k^O\} \tag{8}$$

$$\Sigma_k = \begin{Bmatrix} \Sigma_k^{II} & \Sigma_k^{IO} \\ \Sigma_k^{OI} & \Sigma_k^{OO} \end{Bmatrix} \tag{9}$$

对于每个输入数据的条件概率分布为:

$$p(x_i^O | x_i^I) = \sum_{k=1}^K \sigma_k(x_i^I) N(\hat{\mu}_k^O(x_i^I), \hat{\Sigma}_k^O) \tag{10}$$

其中:

$$\sigma_k(x_i^I) = \frac{\pi_k N(x_i^I | u_k^I, \Sigma_k^I)}{\sum_{i=1}^K \pi_i N(x_i^I | u_i^I, \Sigma_i^I)} \tag{11}$$

根据高斯分布的线性变换特性,由式(10)可知,

$p(x_i^O | x_i^I)$ 的期望 $\hat{\mu}_k^O$ 和协方差 $\hat{\Sigma}_k^O$ 为:

$$\hat{\mu}_k^O = u_k^O + \Sigma_k^{IO} \Sigma_k^{II}^{-1} (x_i^I - u_k^I) \tag{12}$$

$$\hat{\Sigma}_k^O = \Sigma_k^{OO} - \Sigma_k^{OI} \Sigma_k^{II}^{-1} \Sigma_k^{IO} \tag{13}$$

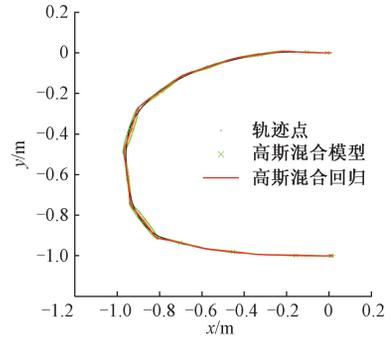
根据输入数据,计算得到期望输出的公式如下:

$$u_k(u_k^O | u_k^I) = \sum_{k=1}^K \sigma_k(x_i^I) \hat{\mu}_k^O \tag{14}$$

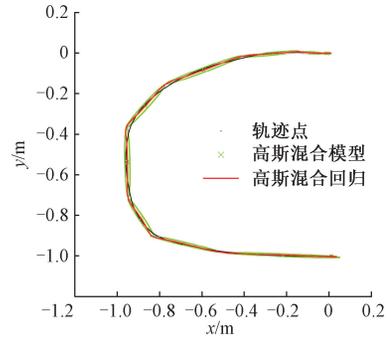
因此,假设已知机器人演示轨迹的一系列数据点 $x = [x_i^O | x_i^I]$,首先对数据点的概率分布 $p(x)$ 建立 GMM,再利用 GMR 计算条件概率 $p(x_i^O | x_i^I)$ 的期望 $E[p(x_i^O | x_i^I)]$ 和协方差 $cov[p(x_i^O | x_i^I)]$,最后将期望值作为泛化的重构数据点,并在该协方差的约束下进行机器人轨迹的生成。本文利用 GMR 来对改进参数估计方法后的得到的 GMM 进行轨迹生成。

1.4 参数估计效率对比

为了验证所提方法的计算效率,使用 3.1 节的工控机进行基于 DP-means 和基于 EM 算法的公共数据集^[25] 手写体字母 C 的 GMM 参数估计实验。由于数据集的演示轨迹只有 200 个数据点,这里使用样条插值将数据点分别扩充至 2 500 点、5 000 点、10 000 点、15 000 点和 20 000 点,即模拟在 500 Hz 采样频率下,采集 5、10、20、30 和 40 s 的演示轨迹。实验中,基于 DP-means 的 GMM 参数估计算法将在离线和在线两种模式下运行。实验结果如图 2 和表 1 所示。



(a) DP-means 辨识结果
(a) DP-means identification result



(b) EM 辨识结果
(b) EM identification result

图2 辨识结果对比

Fig.2 Comparison of identification results

表1 GMM 参数估计时间对比

Table 1 GMM parameter estimation		ms			
算法类别	2 500 点	5 000 点	10 000 点	15 000 点	20 000 点
离线 DP	153	252	515	685	989
在线 DP	0.056	0.059	0.055	0.056	0.051
EM	4 861	9 470	23 940	30 701	44 975

从图 2 和表 1 可以看出,两种估计方法都能得到较好的近似解。但 EM 算法在演示轨迹点增加的情况下,计算效率明显下降。而基于 DP-means 的离线辨识不重复使用数据点进行计算,仍然保持一定的计算效率。由于单个数据点计算时间平均只有 0.056 ms,远小于 2 ms 采样周期。因此,基于 DP-means 的在线辨识能在每个采样周期内完成参数计算,演示结束后即可马上完成参数估计,对于实际应用交互十分友好。

2 基于 DP-means 和 GMM 的复合 DMP 算法

2.1 动态运动基元算法

DMP 具有出色的轨迹复现和泛化能力,如式(15)所示。通过对 DMP 的非线性强制项进行建模,能够将轨迹泛化到新的目标位置,同时保持所学轨迹的形状。DMP

可用于建模周期性和离散运动轨迹。这项工作重点在关注离散运动轨迹。

$$\tau \ddot{\mathbf{x}} = \alpha(\beta(\mathbf{x}_g - \mathbf{x}) - \dot{\mathbf{x}}) + f(t) \quad (15)$$

式中: τ 是时间常数, α 和 β 是正的常数; $f(t)$ 是强迫项。一般来说, 为了获得多功能的动力学, 强迫项 $f(t)$ 被选择为:

$$f(t) = \frac{\sum_{i=1}^{N_1} \varphi_i(t) w_i}{\sum_{i=1}^{N_1} \varphi_i(t)} \quad (16)$$

式中: φ_i 是固定的基函数; N_1 是基函数的个数; w_i 是可调节的权重。为了减少 f 项对时间的明确依赖, 可以用一个相位变量来表示:

$$f(w_i, s) = \frac{\sum_{i=1}^{N_1} \varphi_i(s) w_i s}{\sum_{i=1}^{N_1} \varphi_i(s)} \quad (17)$$

相位变量 s 来自一个一阶系统, 满足 $\dot{s} = -\gamma s$ 。这使得系统在 s 从 1 衰减到 0 的过程中收敛到目标点。通常情况下, 一次演示就足以估算出 f 。此外, 为了获得 w_i , 可以用给定的 $f(s)$ 解决回归问题。得到 w_i 的估计值后, 输入起始位置和新的目标位置, 并重置相位变量, 使用式 (15) 可以计算得到运动到新目标位置的轨迹。

2.2 复合动态运动基元

DMP 是单样本学习方法, 在学习多条演示轨迹时, 先通过 GMM-GMR 获得回归轨迹, 再使用 DMP 进行回归轨迹的学习, 完成后即可进行不同目标位置的轨迹泛化。复合动态运动基元兼具了 GMM 多轨迹编码的能力和 DMP 出色的轨迹复现和泛化能力, 算法过程如算法 3 所示。

算法 3 复合动态运动基元算法

输入: 每个时间步的数据点 \mathbf{x}_i , 长度阈值 λ , DMP 参数 α, β 基函数个数 N

输出: DMP 泛化轨迹

1. 初始化参数 $\lambda, N, \alpha, \beta$
2. 运行基于 DP-means 的在线 GMM 参数估计算法(算法 2)
3. 根据式 (6) ~ (12) 计算 GMR 轨迹
4. 使用局部加权回归计算 DMP 强迫项参数
5. 给定新目标点, 计算 DMP 泛化轨迹

算法运行过程如图 3 所示, 参数需根据机器人类型和作业半径进行选取。本文设置距离阈值 $\lambda = 0.06$ 、DMP 基函数个数 $N = 100$, α, β 使用文献[6]推荐参数, 分别为 6.25、4。

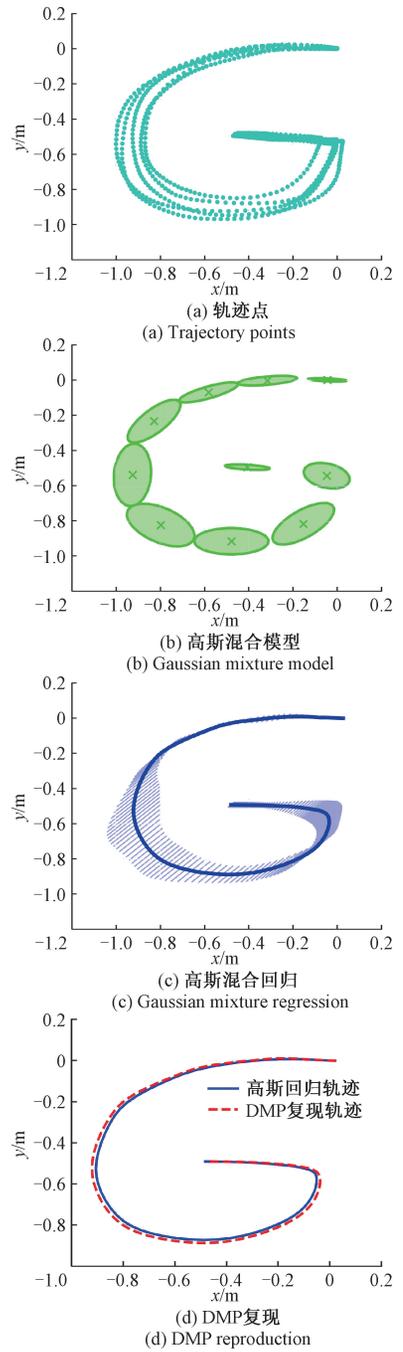


图 3 复合 DMP 学习过程

Fig. 3 Composite DMP learning process

3 实验分析与讨论

分别通过手写体轨迹复现实验和机器人搬运演示实验来验证本文所提方法的有效性。首先验证基于 DP-means 的 GMM 的参数估计性能和轨迹复现效率, 并将所提方法与基于 K-means 的 EM 算法进行对比。然后

进行了机器人物体搬运任务的动作学习实验,验证所提出的复合 DMP 方法的有效性。

3.1 实验平台

本文实验平台如图 4 所示,由机械本体、伺服驱动器、力矩传感器和上位工业控制计算机等组成。机器人关节由三相交流伺服电机和谐波减速器驱动。上位控制采用自主研发的总线式机器人控制系统,使用 EtherCAT 总线和伺服驱动器进行通信,控制频率 500 Hz,具有基于位置控制的直接示教模式和基于力矩控制的动觉示教模式,可在控制软件中自由增加算法。

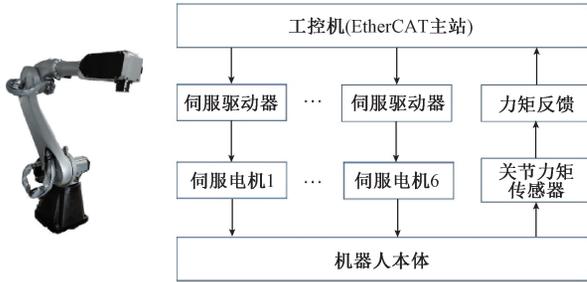


图 4 机器人实验平台

Fig. 4 Robot experimental platform

机器人控制系统搭载在嵌入式工控板上,CPU 为 Intel J1900,内存 8 G。实验采集 5 条机器人末端轨迹进行建模与泛化,采样周期为 500 Hz,所有算法均内置于机器人控制系统中。由于抓取物体和放置物体的行为具有两个目标点,将分别对两段轨迹进行建模与泛化,在完成参数估计后,改变目标位置进行多次泛化实验。实验参数设置如表 2 所示。

表 2 参数设置

Table 2 Parameter settings

参数	参数描述	参数值
N	DMP 基函数个数	100
α	DMP 系数	6.25
β	DMP 系数	4
K	GMM 个数(基于 EM)	10
λ	距离阈值	0.06

3.2 定量评价指标

为了量化学习效果,定义了可达性和相似性评价指标。可达性使用由末端状态下目标和实际位置之间的绝对位置误差 e_x (absolute position error, APE) 来表示。使用均方根误差 (root mean square error, RMSE) 和皮尔逊相关系数 ρ_x (Pearson's correlation coefficient) 来衡量演示轨迹和泛化轨迹之间的相似性。各个评价指标的计算公

式如式 (18) ~ (20) 所示。

$$e_x = |\hat{\mathbf{x}}_g - \hat{\mathbf{x}}_{end}| \quad (18)$$

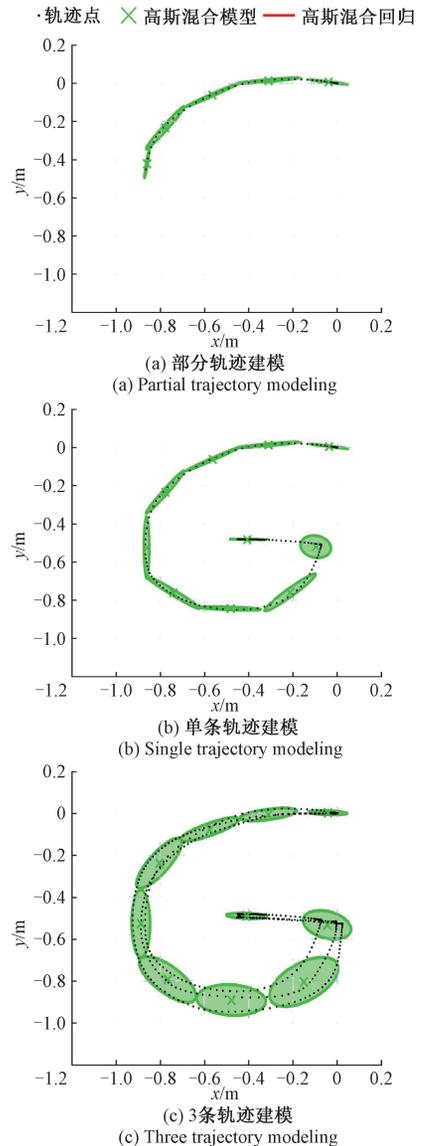
$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{\mathbf{x}} - \mathbf{x})^2}{N}} \quad (19)$$

$$\rho_x = \frac{\text{cov}(\boldsymbol{\eta} \{ \mathbf{x} \}, \{ \hat{\mathbf{x}} \})}{\sigma_{\boldsymbol{\eta} \{ \mathbf{x} \}} \sigma_{\{ \hat{\mathbf{x}} \}}} \quad (20)$$

式中: $\hat{\mathbf{x}}_g$ 为目标期望位置, $\hat{\mathbf{x}}_{end}$ 为目标实际位置, $\hat{\mathbf{x}}$ 为泛化轨迹数据点, \mathbf{x} 为演示轨迹数据点, N 为演示轨迹点个数。

3.3 轨迹复现性能对比

为了验证基于 DP-means 的 GMM 的参数估计性能和轨迹复现效率,本节使用公共数据集^[25]中字母 G 的 5 条轨迹进行实验。实验过程将模拟实际演示过程,以 2 ms 为周期进行轨迹点采样和在线 GMM 参数估计,并将其参数估计过程和轨迹复现结果与基于 K-means 优化的 EM 算法进行比较。两种算法的对比实验结果如图 5 和 6 所示。



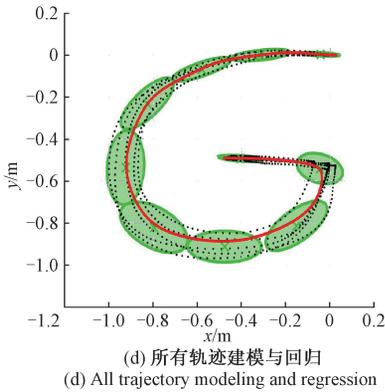


图 5 基于 DP-means 的在线 GMM 参数估计过程
Fig. 5 Online GMM parameter estimation process based on DP-means

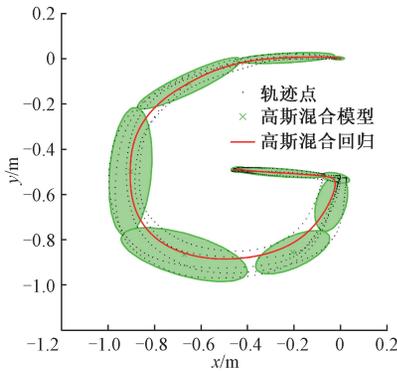


图 6 基于 EM 算法的 GMM 参数估计
Fig. 6 GMM parameter estimation based on the EM algorithm

可以看出,两种方法都能够很好的提取数据有效特征,并实现轨迹复现。基于 DP-means 的在线 GMM 参数估计算法能在轨迹采样过程中增量式计算当前 GMM 参数,演示结束时即可完成参数估计。然而,基于 K-means 的 EM 算法需要在所有轨迹采样完成后才能进行迭代求解。

为了量化说明两种方法生成的回归轨迹和演示轨迹的相似度,表 3 和 4 分别为采用两种方法的生成轨迹的皮尔逊系数和均方误差。

表 3 回归轨迹皮尔逊系数

Table 3 Regression trajectory Pearson's coefficient

算法	轨迹 1	轨迹 2	轨迹 3	轨迹 4	轨迹 5	均值
DP	0.924 0	0.988 7	0.955 8	0.996 7	0.987 3	0.970 5
EM	0.925 2	0.988 2	0.954 0	0.996 8	0.988 0	0.970 4

从表 3 和 4 可以看到,两种算法都能很好地复现演示轨迹。其中,基于 DP-means 方法的皮尔逊系数和均方误差的平均值分别为 0.970 5 和 0.069 1 mm,基于 EM 方

表 4 回归轨迹均方误差

Table 4 RMSE of the regression trajectory mm

算法	轨迹 1	轨迹 2	轨迹 3	轨迹 4	轨迹 5	均值
DP	0.109 0	0.046 2	0.094 6	0.036 9	0.059 0	0.069 1
EM	0.109 2	0.047 0	0.095 7	0.035 6	0.057 2	0.068 9

法的皮尔逊系数和均方误差的平均值分别为 0.970 4 和 0.068 9 mm,两种算法都具有很强的轨迹编码学习能力。

在计算效率上,采用 DP-means 方法的计算时间仅需 0.051 ms,远小于 2 ms 采样周期,能在不影响采样的前提下完成在线参数估计。而基于 K-means 的 EM 算法参数估计的平均计算时间为 4 626 ms。可以看出,基于 DP-means 的参数估计算法计算效率高,更加适合需要快速生成泛化轨迹的场景。

3.4 物体搬运实验

通过对水杯搬运任务的学习来测试本文方法的任务复现能力和计算效率。实验使用图 4 实验平台。采用动觉示教模式拖动机器人末端运动完成预定的相关操作,并记录运动轨迹。演示过程如图 7 所示,采样频率为 500 Hz。在实验过程中,算法参数与前面的保持一致,如表 2 所示。

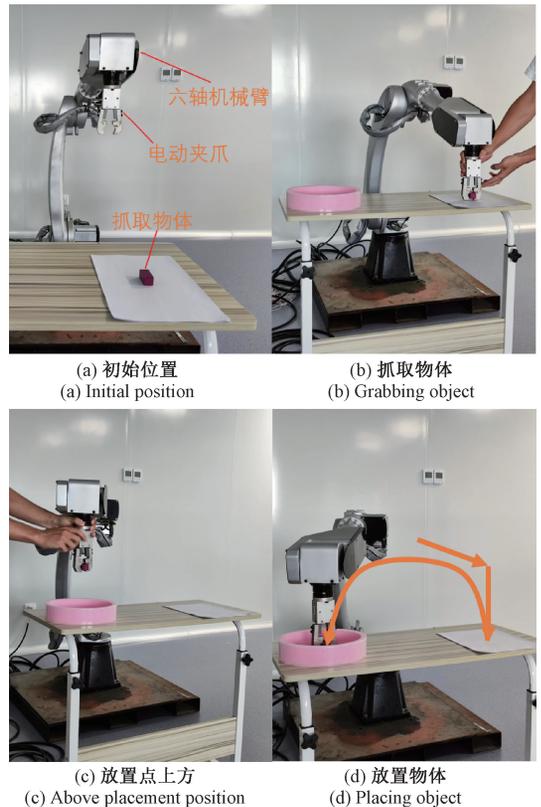


图 7 机器人动觉示教

Fig. 7 Robot kinesthetic demonstrations

为了验证基于 DP-means 的复合 DMP 算法在实际问题中的计算效率,实验将分别使用在线和离线两种运行模式进行参数辨识,并与 EM 算法进行对比。完成参数估计后,再使用 DMP 进行不同目标点的轨迹泛化实验。

图 8 和表 5 所示为复合 DMP 算法对物体搬运任务轨迹编码泛化的整体效果。可以看出,3 种算法在 x 、 y 和 z 轴上泛化位置的平均绝对误差和皮尔逊相关系数差别不大,都具有很好的多轨迹编码和泛化能力。

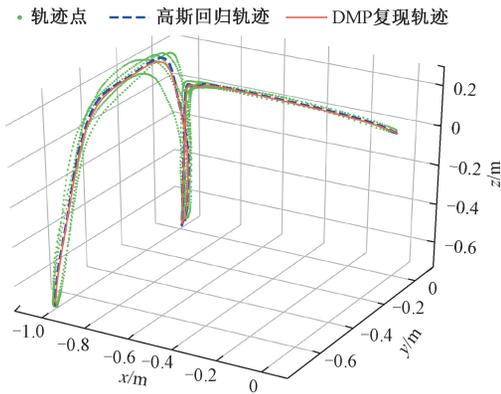


图 8 复合 DMP 轨迹建模

Fig. 8 Composite DMP trajectory modelling

表 5 泛化结果量化指标对比

Table 5 Comparison of generalization indicators

算法	绝对位置误差			皮尔逊系数		
	x/m	y/m	z/m	x	y	z
在线 DP+DMP	0.003 3	0.002 5	0.001 3	0.993	0.998	0.995
离线 DP+DMP	0.003 3	0.002 5	0.001 3	0.993	0.998	0.995
EM+DMP	0.004 1	0.002 2	0.001 7	0.991	0.995	0.996

图 9 所示为轨迹泛化的结果。从图中可以看出,当目标点发生变化时,DMP 能保持轨迹原有特征完成演示任务的泛化,具有很好的轨迹复现能力。

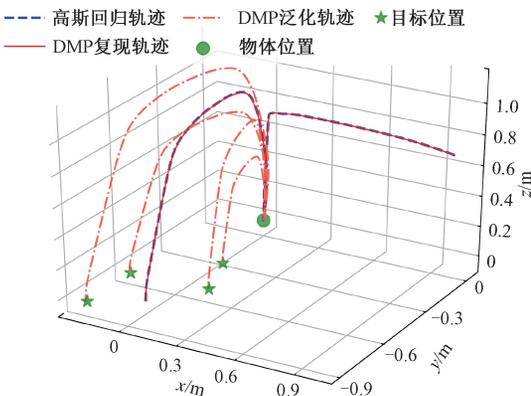


图 9 复合 DMP 轨迹泛化

Fig. 9 Composite DMP trajectory generalization

3 种算法重复运行 10 次的平均时间和稳定性如图 10 所示。其中,在线 DP-means、离线 DP-means 和 EM 算法的平均运行时间分别为 0.052、321 和 31 459 ms。可以看出,基于 DP-means 的参数估计算法无论在线运行或者离线运行,都有很高的计算效率。而 EM 算法需要在每个 M 步骤重新估计 GMM 的协方差,并且每次迭代需要使用所有的数据点,因此随着轨迹复杂程度增加,算法收敛时间将会线性增加。

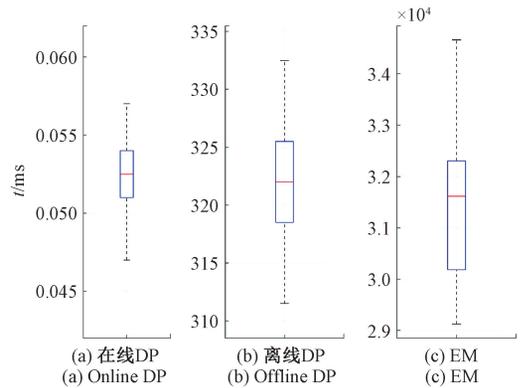


图 10 算法运行时间对比

Fig. 10 Algorithm runtime comparison

4 结 论

针对传统 EM 算法参数估计效率低的问题,本文提出一种基于 DP-means 的 GMM 在线参数估计算法,在保证轨迹编码精度的同时,克服了 EM 算法参数估计时间随着轨迹长度增加而增加的问题。为了提高轨迹泛化能力,进而提出了复合 DMP 算法,在获得高斯混合回归轨迹后,使用 DMP 进行回归轨迹学习与泛化。最后进行了手写体字母轨迹复现和实际机器人搬运任务的实验。结果表明,复合 DMP 算法计算效率高,并且兼具 DMP 的轨迹泛化能力和 GMM 的多轨迹编码能力。

参考文献

[1] 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.

[2] KROEMER O, NIEKUM S, KONIDARIS G. A review of robot learning for manipulation: Challenges, representations, and algorithms [J]. The Journal of Machine Learning Research, 2021, 22(1): 1395-1476.

- [3] 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(1):297-330.
- [4] 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.
- [5] 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.
- [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] CLEVELAND W S. Robust locally weighted regression and smoothing scatterplots[J]. *Journal of the American Statistical Association*, 1979, 74(368): 829-836.
- [8] SCHAAL S, ATKESON C G. Constructive incremental learning from only local information [J]. *Neural Computation*, 1998, 10(8): 2047-2084.
- [9] VIJAYAKUMAR S, D' SOUZA A, SCHAAL S. Incremental online learning in high dimensions [J]. *Neural Computation*, 2005, 17(12): 2602-2634.
- [10] GINESI M, SANSONETTO N, FIORINI P. Dmp ++: Overcoming some drawbacks of dynamic movement primitives[C]. *CoRR*, 2019.
- [11] CALINON S, GUENTER F, BILLARD A. On learning, representing, and generalizing a task in a humanoid robot[J]. *IEEE Transactions on Systems Man & Cybernetics Part B*, 2007, 37(2):286-298.
- [12] ARDUENGO M, COLOME A, BORRAS J, et al. Task-adaptive robot learning from demonstration with Gaussian process models under replication [C]. *International Conference on Robotics and Automation (ICRA)*, IEEE, 2021.
- [13] IODICE F, WU Y Q, KIM W, et al. Learning cooperative dynamic manipulation skills from human demonstration videos [J]. *Mechatronics*, 2022, DOI: 10.1016/j.mechatronics.2022.102807.
- [14] 盛敏, 刘双庆, 王婕, 等. 基于 GMM-HMM 模型的下肢假肢运动意图识别[J]. *仪器仪表学报*, 2019, 40(5):169-178.
SHENG M, LIU SH Q, WANG J, et al. Motion intent recognition of intelligent lower limb prosthesis based on GMM-HMM [J]. *Chinese Journal of Scientific Instrument*, 2019, 40(5): 169-178.
- [15] MOON T K. The expectation-maximization algorithm[J]. *IEEE Signal Processing Magazine*, 1996, 13(6): 47-60.
- [16] 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*, 2007:255-262.
- [17] EVRARD P, GRIBOVSKAYA E, CALINON S, et al. Teaching physical collaborative tasks: Object-lifting case study with a humanoid [C]. *2009 9th IEEE-RAS International Conference on Humanoid Robots*, IEEE, 2009: 399-404.
- [18] MÜHLIG M, GIENGER M, STEIL J J. Interactive imitation learning of object movement skills [J]. *Autonomous Robots*, 2012, 32(2): 97-114.
- [19] LEE S H, SUH I H, CALINON S, et al. Learning basis skills by autonomous segmentation of humanoid motion trajectories [C]. *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*, IEEE, 2012: 112-119.
- [20] DEB P. Finite mixture models[J]. *Partha Deb*, 2000, 39(4):521-541.
- [21] REDNER R A, WALKER H F. Mixture density, maximum likelihood and the EM algorithm [J]. *SIAM Review*, 1984, 26(2):195-239.
- [22] OHAGAN A. Curve fitting and optimal design for prediction[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1978, 40(1):1-42.
- [23] FERGUSON T S. A Bayesian analysis of some

nonparametric problems [J]. The Annals of Statistics, 1973, 1(2): 209-230.

- [24] KULIS B, JORDAN M I. Revisiting K-means: New algorithms via Bayesian nonparametrics [C]. CoRR, 2011.
- [25] PIGNAT E, CALINON S. Learning adaptive dressing assistance from human demonstration [J]. Robotics and Autonomous Systems, 2017, 93: 61-75.

作者简介



吴晓敏, 2014 年于东北大学获得学士学位, 2018 年于厦门大学获得硕士学位, 现为厦门大学博士研究生, 主要研究方向为工业机器人控制。

E-mail: wuxiaomin@stu.xmu.edu.cn

Wu Xiaomin received his B. Sc. degree from Northeast University in 2014, and M. Sc. degree from Xiamen University in 2018. He is currently a Ph. D. candidate at Xiamen University. His main research interest is industrial robot control.



刘瞰东(通信作者), 2003 年于中国科学技术大学获得博士学位, 现为厦门大学教授, 博士生导师, 主要研究方向为工业机器人控制等。

E-mail: ltd@xmu.edu.cn

Liu Tundong (Corresponding author) received his Ph. D. degree from University of Science and Technology of China in 2003. He is currently a professor and a Ph. D. advisor at Xiamen University. His main research interest is industrial robot control and so on.