



文章

机器人任务约束优化与适应
概率运动基元

丁冠文,

西哲藏

张学和 *

李长乐*、朱彦河和赵杰

哈尔滨工业大学机器人技术与系统国家重点实验室, 哈尔滨 150001, 黑龙江; 17b908042@stu.hit.edu.cn (GD); zangxizhe@hit.edu.cn (XZ); yhzhu@hit.edu.cn (YZ); jzhao@hit.edu.cn (JZ)

*通讯地址: zhangxuehe@hit.edu.cn (XZ); lichangle@hit.edu.cn (CL)

摘要: 让机器人能够向人类学习技能并适应不同的任务场景, 将使机器人在制造业中的使用能够提高效率。运动基元 (MP) 是编码技能的重要工具。本文研究了如何从少量人类演示中学习 MP 并适应不同的任务约束, 包括航点、关节限制、虚拟墙和障碍物。概率运动基元 (ProMP) 使用分布对运动进行建模, 从而为机器人提供了执行任务的额外自由。我们为机器人提供了三种移动模式, 每种模式只需要一个人类演示。我们提出了一种改进的经点泛化方法, 以使用编码的 ProMP 概括平滑轨迹。此外, 我们提出了一种有效的任务约束优化方法, 该方法将所有任务约束分析地纳入概率框架。我们在每个时间步将 ProMP 分离为高斯函数, 并最小化 Kullback-Leibler (KL) 散度, 并执行梯度上升下降算法以获得优化的 ProMP。鉴于优化的 ProMP, 我们概述了一种统一的机器人运动适应方法, 可从单个障碍物扩展到多个障碍物。我们通过一系列运动适应实验, 使用 7-DOF Xarm 机器人验证了我们的方法。

关键词: 人机技能转移; 从示范中学习; 概率运动原语; 任务约束优化; 运动适应



引文: Ding, G.; Zang, X.; Zhang, X.; Li, C.; Zhu, Y.; Zhao, J. 机器人任务约束优化与概率自适应

运动原语。仿生学

2024 年 9 月 738 日, <https://doi.org/10.3390/biomimetics9120738>

学术编辑: 余俊杰

收到日期: 2024 年 10 月 29 日

修订日期: 2024 年 11 月 26 日

接受日期: 2024 年 12 月 2 日

发布日期: 2024 年 12 月 3 日



版权: © 2024 作者。

被许可方: 瑞士巴塞尔 MDPI。

本文是一篇开放获取文章

根据条款分发

知识共享条款

署名 (CC BY) 许可证 ([https://](https://creativecommons.org/licenses/by/4.0/)

[creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)

4.0/).

1. 简介

随着现代技术的进步, 人工智能 (AI) 在改变各个领域 (包括智慧城市和工业制造) 方面发挥着关键作用。在智慧城市中, 人工智能与物联网 (IoT) 和区块链相结合, 实现了数据驱动的决策、实时监控和安全的数据共享。这些功能使智慧城市能够优化资源配置、增强公共安全、改善交通管理并为居民提供个性化服务[1,2]。通过分析从传感器、物联网设备和其他来源收集的大量数据, 人工智能促进了智能自动化, 帮助城市系统有效应对复杂挑战[3,4]。

同样, 在工业制造领域, 机器人技术也在经历一场由适应性和自动化需求推动的平行变革[5]。传统上, 机器人在制造业中的任务是在精心构建的环境中以重复的轨迹执行的。然而, 当环境发生变化时, 机器人往往需要进行大量的重新编程才能适应。为了解决这一限制, 人工智能驱动的方法 (例如从人类示范中学习) 使机器人能够适应不同的任务场景, 从而使其更加灵活和有弹性[6]。在此背景下, 运动基元 (MP) 应运而生, 成为机器人技术中一种成熟的方法。MP 通过时间和空间调制实现运动的泛化, 使机器人能够对不同的基元进行排序和组合, 以高效地完成复杂任务[7]。

核心挑战在于如何从有限的人类演示中学习和调整 MP。当前的方法可分为两类：(i)概率方法和 (ii)动态系统。概率方法通常涉及隐马尔可夫模型 - 高斯混合回归 (HMM-GMR) [8]和高斯混合模型 - 高斯混合回归 (GMM-GMR) [9]等方法,它们对多个演示的联合概率分布进行建模并通过回归得出所需的轨迹。虽然这些方法捕捉到了演示的统计变化,但它们很难适应新的任务场景。为了解决这个问题,任务参数化高斯混合模型 (TP-GMM) [10]将任务参数合并到 GMM 中,使其能够适应新条件。但是,当学习过程中不包含任务参数时,TP-GMM 会失败。概率运动基元 (ProMP) [11]通过具有概率条件的分层贝叶斯模型来表示运动,以实现点泛化,从而捕获来自不同演示的相关性。

然而,ProMP 需要大量的演示才能获得可靠的分布,并且一旦途经点超出原始范围,外推能力就会受到限制。

另一方面,动态系统方法,例如动态运动原语 (DMP) [12,13],使用受外部强迫项调制的非线性微分方程来描述运动。DMP 非常适合从单个演示中概括起点和终点,但无法有效处理中间的中途点。为了增强其能力,一些研究将 DMP 与强化学习 (RL) 相结合,以学习不同控制变量之间的耦合[14]。

最近,已经提出了结合 DMP 和 ProMP 的混合方法。

概率动态运动基元 (ProDMP) [15]将 DMP 的数值积分转换为基函数,以表示具有 ProMP 的轨迹。但是,学习 ProMP 所需的 DMP 轨迹数量会影响性能。通过点运动基元 (VMP) [16]表示具有基本和形状轨迹的运动,适用于单个或多个演示。定义了一个阈值来区分 DMP外推和 ProMP 内插。但是,VMP 可能会导致在具有非平滑轨迹的过渡点处产生高加速度。

在将 MP 适应不同的任务场景时,通过点泛化是不够的。任务约束表现为位置和速度限制、障碍物和虚拟墙等不平等,这些约束被施加在 MP 上。目标是在满足任务约束的同时生成与原始轨迹非常接近的轨迹。MP 适应方法得到了广泛的研究;例如,DMP 通过在加速度域中施加排斥力来扩展避障能力[17]。通过将轨迹转换为外生状态来确保位置限制[18]。然而,当多种力相互作用时,DMP很容易陷入局部最小值。此外,DMP 适应特定的轨迹,失去了以不同方式执行任务的额外自由。

相比之下,ProMP 使用协方差矩阵对变化进行编码,并且可以进一步利用这些变化来满足任务约束。参考文献[19]组合了多个 ProMP 来实现避障。然而,通过添加基元来预见所有情况是不切实际的。另一种方法是通过优化从原始 ProMP 中排除特定区域;例如,参考文献[20]将优化表述为策略搜索问题,其中定义了 Kullback-Leibler (KL) 散度和奖励函数来推导策略,而[21]中的工作通过最小化与原始 ProMP 的马哈拉诺比斯距离并结合障碍物距离计算来实现优化。然而,上述方法仅针对机器人末端执行器优化 ProMP,而没有考虑其他机器人链接的可能碰撞。此外,它们仅处理障碍物约束,而没有考虑其他任务约束,例如虚拟墙和关节限制。参考文献[22]将多个任务约束纳入一个集成概率框架;他们直接用梯度上升下降算法来优化ProMP,但是ProMP在每个时间步都会影响高斯分布,满足任务约束的时间步才会被优化,导致优化时间过长。

在本文中,我们提出了一种改进的点泛化方法,允许机器人仅从一次演示中获取技能。ProMP 通过泛化来编码路线图与中途点相结合,而不是提供更多示范。此外,我们提出了一种改进的任务约束 ProMP 自适应方法,通过分离每个采用高斯时间步长,并且仅优化违反任务约束的特征时间步长。

总而言之,我们的贡献如下:

(1)我们提出了一种改进的经点泛化方法,允许机器人仅通过一次人类示范即可获得技能,且轨迹平滑,可以推广到学习 ProMP;

(2)我们提出了一种有效的任务约束 ProMP 优化方法,该方法与将所有任务约束分析地纳入概率框架中;

(3)鉴于优化的 ProMP,我们提出了统一的机器人运动适应从单一障碍延伸至多种障碍的方法。

本文的结构如下:第2部分推导了理论
本文涉及的公式和方法,第2.1节介绍通过点轨迹泛化和 ProMP 编码,第2.2节制定任务约束的
优化问题,第2.3节介绍优化程序;第3.2节
概述了完整的机器人运动适应过程。我们验证并
通过第 4 节中的实验比较我们的方法。最后,我们总结了这项工作,并
概述未来的研究方向。为清楚起见,本文使用的所有定义和符号
论文总结在表1中。

表 1.本文中的定义和符号。

方面	描述
电视	轨迹的时间步长
n	移动平均滤波器 (MAF) 的窗口宽度
否	示威人数
德	机器人的 DoF 数
钾	高斯基函数的数量

特征空间	描述
$p = [x, y, z], o = qw, qx, qy, qz$	人手质心的位置p和方向o集
$y(p), y(o)$	结合 3D 和方向轨迹
$h(p), h(o)$	基本 3D 和方向轨迹
$f(p), f(o) p1,$	3D 和方向形状调制
pend	起点和终点 3D 点
r	表示投影点比例的 3D 比例因子
TVIA、P VIA	途经点的时间步长和 3D 位置
K	三维形状调制系数
否	采用逆运动学 (IK) 算法的机器人所有关节组
{gt} 我=1 分钟 最大幅度	第 k 个关节自由度的最小和最大关节位置
问, t, q k,t	预期航点位置
Xt	超平面法线和偏差向量
nk,t, bk,t	机器人驱赶器之间的最小距离
dmin	排斥器安全半径
安全	

分布	描述
$\Psi(t) w_i$	高斯基函数
en	第 i 个高斯核的权重向量
	零均值高斯噪声
$\{\mu w, \Sigma w\} \Sigma n$	权重向量的均值和协方差
	高斯噪声协方差
$p0(w) = N w; \mu$	ProMP 的初始高斯分布
页 (w) = N (w; $\mu_{\bar{w}}, \sum_{\bar{w}}^0 \sum_{\bar{w}}^0 * w$)	优化 ProMP 的高斯分布

表 1.续

分布	描述
$\theta_0 = \mu_{\sigma_0, t}^0$	每个时间步的初始平均值和标准差
$\theta = (\mu_{\sigma_t}, \sigma_t)$	每个时间步的优化平均值和标准差
$c_{k,t} = C(\mu_t, \sigma_t)$	与 (μ_t, σ_t) 相关的第 k 个任务约束函数
$F_{c_{k,t}}(\mu_t, \sigma_t)$	第 k 个任务约束的累积分布函数 (CDF)
霍克	单侧或双侧不等约束
σ_t 霍克	任务约束概率
分布	描述
q_k	表示任务约束概率的置信水平 α_k, t $d_{k,t}$ 表示任务约束概率的距离阈值 $S = \{c_{k,t}(q_1, q_2, \dots, q_D),$
$\in \{\mu_k - 2\sigma_k, \mu_k + 2\sigma_k\}$	非线性任务约束的估计集 非线性任务约束的估计均值和方差 $E_{c_{k,t}}, V_{c_{k,t}} \Gamma(\alpha, \beta)$
	形状为 α 且速率为 β 的伽玛分布 $DKL[\theta \theta_0]$
	两个单变量高斯分布之间的 KL 散度 $\lambda_{k,t}$ 拉格朗日乘数 $L_{\theta, \lambda_{k,t}}$ 拉格朗日函数 收敛速度参数 η_k

2. 理论表述和方法2.1. 生成轨迹和编码分布 2.1.1. 经点轨迹泛化

通过点轨迹泛化,机器人能够为给定的技能快速生成一组不同的可行动作,同时保留人类演示的关键特征。

此过程包括泛化位置和方向,通过教授机器人新技能,使机器人能够高效地生成特定于任务的轨迹,同时保持任务的多样性。如图1 所示,我们提供了三种模式,即在障碍物上方、左侧和右侧移动。每种模式只需要一个演示,它是通过使用 Kinect FORTH 系统跟踪人手运动获得的[23]。我们选择人手质心作为跟踪点,并在T时间步内记录位置和方向集 $x, y, z, q_x, q_y, q_z, q_w$ 。为了确保轨迹平滑,我们应用窗口宽度为n 的移动平均滤波器 (MAF) ,并将轨迹变换到机器人基准框架中。为了学习和泛化人类演示特征,我们将轨迹表示为基本轨迹 h 和形状调制 f 的组合。

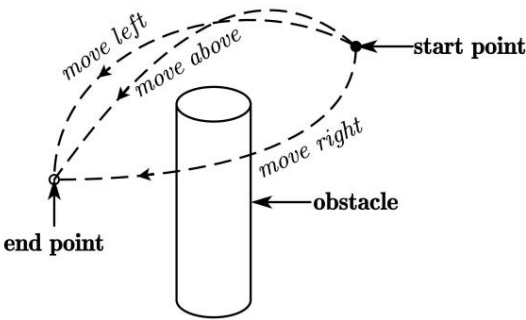


图 1.机器人基于人类示范绕过障碍物的三种技能。

(1)立场概括

对于三维位置,我们用 $y(p) = h(p) + f(p)$ 表示位置集 p 。基本轨迹 $h(p)$ 是一条直接连接起点 p_1 和终点 p_{end} 的线,形状调制定义为 $f(p) = p - p_1 + r$,其中 r 是 p_1 p_{end} ,比例因子表示直线 $h(p)$ 上投影点的比例。如果不需要中间经由点,我们直接根据 r 在 a 和终点 p 之间生成新的投影点,学习到的形状调制为新的起点 p

新的 = p1 新的 + 角色扮演 1 页 结尾

$$k = p_{via} - p_{\text{新的}} = p_{\text{新的}} + \text{角色扮演的新的} - \text{新的结束} \quad |f(p)|^{-1} \text{。随后,新的轨迹集将通用}$$

$$\text{化为} p_{\text{新的}} = p_{\text{新的}} + r p_{\text{新的结束}} + k f(p) \text{。}$$

考虑到方向维度相互影响,我们将轨迹定义为 $y(o) = h(o) \times f(o)$,其中 $o = q_w, q_x, q_y, q_z$ 表示四元数集, \times 表示四元数乘法。球

一旦在泛化过程中需要新的起始和终止四元数集, 我们就会计算新的基本轨迹 $h(o)$, 并将学习到的形状调制 $f(o)$ 叠加起来, 以生成新的方向轨迹 $v(o)$ new。请注意, 无需泛化方向经点, 因为我们可以使用经点泛化 3D 位置, 同时保留原始方向。

概率运动基元 (ProMP) [24]是在关节或笛卡尔空间中编码方差的概率方法。为了以分析方式整合所有任务约束,我们在关节空间中对 ProMP 进行编码。给定第2.1.1节中概括的位置和方向集,我们通过雅可比逆运动学 (IK) 算法获得所有关节集 $\{q_t\}$,其中 $q_t \in \mathbb{R}^D$, $t \in [0, T]$, N 是演示数, D 是 DoF 数。

电视
 $w_i + \varepsilon n_i$
$$p(\tau; \theta) = p(\tau|w)p(w; \theta)dw$$

$$= \prod_{t=1}^T N(q(t)|\Psi(t), \Sigma(t)) \prod_{t=1}^T \mathcal{N}(\Psi(t) | \mu_w, \Sigma_w, \Psi(t) + \Sigma_n) \quad (1)$$

其中 μ_{wv} 和 Σ_{wv} 表示权重向量的均值和协方差,它们由最大似然估计得出, Σ_n 是高斯噪声协方差。为了使用 ProMP 建模关节分布,我们首先将人手轨迹转换为机器人基准坐标系,并得出相应的关节值,然后,这些关节值表示为 K 个高斯基函数及其相关权重的组合。

在执行任务的过程中,机器人在每个时间步从关节分布中采样轨迹,确保适应性,同时保留所表现出的可变性。

给定初始 ProMP $p_0(w) = N(w; \mu, \Sigma)$ 主要目标是获得优化的 ProMP,使其尽可能接近原始 ProMP,同时满足任务约束。实现这种权衡需要一种方法来量化分布之间的相似性,同时有效地编码任务约束。在这种情况下,KL 散度特别适合,因为它对高斯分布有闭式解,确保了计算效率。此外,累积分布函数 (CDF) 提供了一个

考虑到 ProMP 的随机性,我们引入了概率框架,将可变性和不确定性纳入约束中。因此,我们使用 KL 散度来表示相似性,并使用累积分布函数 (CDF) 来制定任务约束。

由于 ProMPs 将每个时间步建模为高斯分布,我们通过将初始 ProMPs 分别转换为在任务约束下优化的平均值和标准差 μ), 将每个时间步与高斯分布分开。
令 $ck,t = C(\mu_t, \sigma_t)$ 表示与 (μ_t, σ_t) 相关的第 k 个任务约束函数,令 $Fck,t(\mu_t, \sigma_t)$ 表示 ck,t 的 CDF。我们可以将优化问题重写为拉格朗日函数:

每个时间步 t 的新平均值和标准差 (μ σ_t ,

$$L(\mu_t, \sigma_t, \lambda_{k,t}) = DKL(N(q_t; \mu_t, \sigma_t) || N(q_t; \mu_t^0, \sigma_t^0)) + \sum_{k,t} \lambda_{k,t} (a_{k,t} - Fck,t(\mu_t, \sigma_t)) \tag{2}$$

其中 $a_{k,t}$ 是置信水平,表示任务约束概率 $P(\mu_t, \sigma_t | H(ck,t))$ 是最小值
 H 表示单边或双边不等式约束,优化的 $(\mu_{a_{k,t}}, \sigma_{a_{k,t}})$ 通过梯度上升下降法得出: $\mu_{a_{k,t}}, \sigma_{a_{k,t}}$
随后,给定 $(\mu_{timized}, \sigma_{timized})$ 作为 $\mu_{t,t=最小\mu_{w}}, \sigma_{t,t=最小\sigma_{w}}$, 我们实施另一个优化步骤来获得 $op-\Psi T = \min \sum w \Psi t - \sigma$

tion, 我们通过定义与 (μ_t, σ_t) 相关的 $\Psi T_{\mu w} - \mu \sum t \bar{w}$ 在接下来的一秒里 Σw

函数 $ck,t = C(\mu_t, \sigma_t)$ 、表达不等式约束函数 H 和近似 CDF $Fck,t(\mu_t, \sigma_t)$ 来制定每个任务约束。

2.2.2. 关节和笛卡尔空间中的约束定义 (1) 关节范围限制在本文中,我们重点保留从 ProMP 中学到的原始关节范围。

我们建立约束函数 $ck,t = qt(\mu_t, \sigma_t)$, 并将双边约束表示为, 其中 k 表示第 k 个 $P(\mu_t, \sigma_t)$ 关节的自由度, $a_{k,t}$ 是 $a_{k,t}$, $q_{k,t}$ 分钟 $q_{k,t} < qt(\mu_t, \sigma_t) < q_{k,t}$ 最大限度 $q_{k,t}$ 分钟 $q_{k,t}$ 最大限度 $q_{k,t}$
置信水平, q 分别为最小和最大关节位置。我们用高斯函数近似 $q_{k,t}$ 的关节位置,并计算 CDF 中的差异以获得 $Fck,t(\mu_t, \sigma_t)$ 。

(2) 路点我们期望机器人在时间步 t 到达的路点在笛卡尔空间中用 FK 函数 T 定义。根据预期点和实际点之间的距离平方,我们将路点约束函数表示为 $ck,t = |T(q_t(\mu_t, \sigma_t)) - X_t|^2$, $ex-$
将单侧约束设为 $P(\mu_t, \sigma_t | |T(q_t(\mu_t, \sigma_t)) - X_t| \leq dk,t - a_{k,t}, k,t)$ 为距离阈值, $a_{k,t}$ 为置信水平。由于的非线性特性,我们估计分布如下: ck,t , d^2

$$S = \{ck,t(q_1, q_2, ..., q_D), q_k \in [\mu_k - 2\sigma_k, \mu_k + 2\sigma_k]\}$$
$$E(ck,t) \approx \text{均值}[S]$$
$$V(ck,t) \approx \text{Var}[S]$$

(3)

其中 D 是自由度,我们利用 2 倍标准差构成集合 S , 并使用矩匹配计算均值和方差。随机变量 ck,t 分布是广义的 χ^2 , 我们用形状为 α 和速率为 β 的更简单的 Gamma 分布来近似它,以分析表示 CDF。然后使用 $Fck,t = F(dk,t; \Gamma(\alpha, \beta))$ 计算 CDF 以表示航点约束概率。

(3) 超平面超平面约束

束,称为虚拟墙,在笛卡尔空间中定义,以将机器人限制在特定平面上。我们将超平面约束函数公式化为 $T(q_t(\mu_t, \sigma_t) - b_{k,t})$,其中 $n_{k,t}$ 和 $b_{k,t}$ 是超平面法向量和偏差向量,

$$T_{k,t} = n_{k,t}^T$$

具有单边不等式约束,表示为 $P_{\mu_t, \sigma_t}(T(q_t(\mu_t, \sigma_t) - b_{k,t}) \leq 0)$, $k, t \in \{1, \dots, N\}$, $c_{k,t}$ 分布是笛卡尔三维位置 $T(q_t(\mu_t, \sigma_t))$ 的线性变换;然而, T 在 (μ_t, σ_t) 中是非线性的。因此,我们使用矩匹配法(等式(3))来获得估计的均值 $E[T]$ 和方差 $V[T]$,因此整体分布变为 $c_{k,t} \sim N(\mu_{k,t}, \Sigma_{k,t})$, N 表示概率。

$$\mu_{k,t} = b_{k,t} + T_{k,t}^T \mu_{k,t}, \Sigma_{k,t} = T_{k,t}^T \Sigma_{k,t} T_{k,t}, \text{使用高斯分布计算 CDF,}$$

(4) 排斥器排斥器约束

加在机器人末端执行器或其他连杆上。我们定义危险距离时仅考虑机器人容易发生碰撞的危险情况。我们计算机器人连杆与障碍物之间的最短距离,找出最小距离 d_{min} 。如果 d_{min} 大于障碍物安全半径 R_{safe} ,则危险距离为0;否则,危险距离为 $d_{danger} = R_{safe} - d_{min}$ 。我们建立约束函数并表达单边不等式,以表明 $d_{min} \leq R_{safe}$ 的概率为 P_{μ_t, σ_t} , $d_{danger} \leq 0$, $c_{k,t}$ 分布在 (μ_t, σ_t) 中是非线性的,因此我们利用公式(3)估计并计算高斯分布的CDF来表示概率。

2.3. 优化过程基于2.2.2节中定义

的拉格朗日函数和约束函数,我们具体化了获得优化参数的过程。为方便起见,我们使用符号 $\theta = \mu$ 和 $\theta = (\mu, \sigma)$ 表示初始参数和优化参数。我们将拉格朗日函数重写为公式(2):

$$L(\theta, \lambda, t) = DKL[\theta || \theta_0] + \sum_{k,t} \lambda_{k,t} \alpha_{k,t} - F_{c_{k,t}}(\theta) \quad (4)$$

$DKL[\theta || \theta_0]$ 表示两个单变量高斯分布之间的 KL 散度,优化参数为 $\theta = L(\theta, \lambda, t)$ 。

$$\lambda_{k,t} = \begin{cases} \lambda_{k,t}^{\max} & \text{if } \lambda_{k,t} > \lambda_{k,t}^{\max} \\ \lambda_{k,t} & \text{if } \lambda_{k,t} \leq \lambda_{k,t}^{\max} \\ 0 & \text{if } \lambda_{k,t} < 0 \end{cases}$$

用双循环梯度上升下降算法,因为它非常适合解决涉及最大化和最小化目标的优化问题。

在这种情况下,外循环优化拉格朗日乘数 $\lambda_{k,t}$,以强制执行任务约束,而内循环最小化参数 θ 以实现与所需 ProMP 的最佳拟合。通过解耦对偶变量和原始变量的更新,这种方法确保了高效收敛,使算法能够有效地处理任务约束和分布相似性之间的固有耦合。双循环结构还提供了一种系统的方法来平衡满足约束和最小化 KL 散度之间的权衡。具体来说,对于外循环,我们使用梯度上升步骤优化 $\lambda_{k,t}$ 。我们使用指数乘法法(EMM) [26]通过指数因子求解不等式约束,将每次迭代(s+1)后的优化后的 $\lambda_{k,t}$ 表示为 $\lambda_{k,t} = \exp(\eta_k \partial \lambda_{k,t})$,其中 η_k 决定收敛速度, k, t 和 $\lambda_{k,t}$ 的梯度通过 $\partial \lambda_{k,t} = \alpha_{k,t} - F_{c_{k,t}}(\theta)$ 获得。对于内循环,我们使用梯度下降法优化 θ 。我们采用 LBFGS [27]方法进行低内存成本的非线性优化。我们执行几个 LBFGS 步骤,使用 Tensorflow 自动微分框架[28]计算梯度 ∂L 。

$$= \lambda_{k,t}^{\max} \quad (4)$$

每次迭代后都会检查收敛条件,以加快该过程。给定优化的 (μ, σ) 。对于参数 (μ, σ) ,我们执行几个 LBFGS 步骤来获得 (μ, σ) 。

选择,我们首先选择参数 θ_0 、 λ_0 、 k_t 、 η_k 的初始值,然后通过内环和外环优化迭代更新它们,直到收敛。

为了整合本节提出的方法,算法1 总结了整体优化过程。

算法1 ProMPs的任务约束优化算法。

输入: μ^0, Σ^0

输出: (μ^*, Σ^*)

初始化: μ^0, Σ^0

计算 μ^0 推导 k_t, H 和 σ_t

$F_{k,t}(\mu_t, \sigma_t)$ (参见第2.2.2节)以构建方程 (2)。在方程 (4)中定义。

$\theta \leftarrow \theta_0, \lambda \leftarrow \lambda_0, k_t \leftarrow k_t$

可学习参数1: $\mu_t, \Sigma_t, k_t, \lambda$

优化 1: 重复 θ (s+1) \leftarrow LBFGS $L, \partial L, \theta(s)$

直到内循环条件 (r+1) (r) λ, k_t 直到收敛 可学习, λ, k_t (r)

参数 2: $(\mu_t, \Sigma_t) \leftarrow \lambda, k_t, \exp(\eta_k k_t - F_{k,t}(\theta))$

直到收敛 μ_t, Σ_t

优化2: 重复 (s+1) $\mu_t, \Sigma_t \leftarrow \Sigma_t^{(s+1)}$ (s) \leftarrow LBFGS μ_t, Σ_t

直到收敛

3. 使用 ProMP 调整运动3.1. 单一障碍物下的移动调整给定一个障

碍物,我们提供三种模式来绕过障碍物。我们将轨迹推广到相同的起点和终点,没

有途经点 (参见第 2.1.1节) ,并将它们规范化为相同的时间步长 T ,作为新的模板轨迹。为了生成不会发生碰撞的可行轨迹集,我们在适当范围内以相等的间隔建立特征途经点。我们计算机器人末端执行器和障碍物之间的距离,并在必要时微调轨迹以确保安全。为了方便距离计算和轨迹推广,我们将机器人末端执行器建模为球体。如果机器人末端执行器最小 z 高于障碍物的上表面,则轨迹是安全的,如果最小 z 低于障碍物的上表面,则当轴向距离大于总安全半径时,轨迹是安全的。当轨迹容易发生碰撞时,我们以不同模式确定的方式调整特征途经点以概括安全轨迹。

对于每种模式,我们概括了 N 条安全的机器人末端执行器轨迹。我们使用雅可比矩阵 J_K 获得所有可行的关节集 $\{q\}$ 我=1 并编码 ProMPs $p_0(w) = N w; \mu^0, \Sigma^0$ 。为了加快优化速度,我们首先找到实际上违反任务约束的特征时间步长,并使用算法1 优化分布。优化的 ProMP 被视为安全联合分布。

在执行任务时,机器人首先确定起点和终点,然后选择模式 $k \in \{\text{上方,左侧,右侧}\}$,并使用高斯 $p_{via} = N(\mu_{xyz}, \sigma_{xyz})$ 采样路径点。平均值 μ_{xyz} 和标准偏差 σ_{xyz} 是根据初始路径点范围计算得出的。我们利用采样的路径点对轨迹进行概括和微调,以确保安全。优化的 μ 和 σ (μ

$q_i - 2\sigma_{t,t} + 2\sigma_{t,t}$),初始联合值表示为 q_i t 表示关节极限范围为 $q_i^0 = \Psi^T w_{q_i}$ 。关节有限安全的 \in

范围和初始值被传递给雅可比 IK 算法,以测试解决方案是否计算成功。所有特征时间步的成功次数决定机器人是否可以成功执行任务:即,如果成功次数大于阈值,机器人直接运行广义轨迹;否则,机器人继续采样直到成功。为了获得阈值,我们从初始的路径点范围进行等距采样,记录成功结果和相应的成功次数,然后确定从成功到失败的阈值。

3.2. 有各种障碍物的运动适应基于我们试图主动避开

环境中的每个障碍物的假设,我们将M个障碍物的适应性定义为障碍物对的顺序组合。因此,我们使用通过 ProMP 学习的M个段来分割轨迹,并通过算法 1 分别进行优化。对于每个段,我们首先根据两个障碍物的相对位置确定起点和终点。当轴向距离能够容纳末端执行器时,我们取两个障碍物之间的中点;如果不能,我们取上方的中点。机器人首先从不同模式中概括安全轨迹集并编码 ProMP,然后在任务约束下优化 ProMP,最后从每个段的不同模式中采样并将所有段连接在一起以执行任务。统一的运动适应方法在算法 2 中提出。

算法2机器人任务约束运动适应算法。

```
输入: 障碍物信息{obsm} 模式模板轨迹x, y, z, qx, qy, qz, qw, m=1, p1, pend, k=1

输出: 任务执行轨迹x, y, z, qx, qy, qz, qw

1. 确定局部起点和终点 (参见第3.2 节) 并生成N 条安全轨迹。 获取可行的关节集{q} 查找违反任务约束的特征时间步长{tfeature}。
2. 优化 ProMP 分布(μ, σ) 获取优化的关节限制范围和初始关节 q0, σ0 并编码 ProMP μ, σ。
3. 值q (μ, σ) 确定成功阈值 (参见第3.1 节)。重复 采样模式 k {上方、左侧、右侧}。 采样途经点pvia = N μxyz, σxyz。 概括和微调轨迹。 通过 w, w)使用算法1。
4. 雅可比 IK测试{tfeature}的成功次数 安全的 ∈
   -2σ t,t + 2σ q, q = ΨT w。
5. 直至成功。end for
6. 任务执行
7. 连接各段轨迹,执行任务。
```

4.实验验证

在本节中,我们将在工业场景中评估所提出的方法,在该场景中,机器人尝试在满足任务约束的同时抓住紧急停止装置。我们将场景从单个障碍物扩展到多个障碍物,并与其他方法进行了比较实验,以验证所提出的方法。在从人类演示中学习不同的技能后,机器人需要在不同的任务约束下自主执行抓住紧急停止装置的任务。

4.1. 单一障碍物的实验

在本研究中,我们选择了 Ufactory Xarm 7-DoF 机械臂作为实验平台。Kinect V2 摄像头策略性地放置在侧面,以跟踪人类手部运动,捕捉连续的质心姿势。这种设置确保即使在动态条件下也能准确记录人类演示。此外,Zed2i立体摄像头安装在头顶,可以精确获取物体姿势,这对于理解工作空间内的空间环境和交互要求至关重要。这种双摄像头配置提供了互补的视角,增强了系统的整体稳健性。

给定工作空间中的单个障碍物,我们提供三种模式来绕过障碍物。每种模式都涉及对人体手部质心姿势的连续跟踪,由 Kinect FORTH 系统实现,确保轨迹获取的流畅性和实时性。为了改进从人类演示中获得的原始轨迹,我们应用 $n = 5$ 的移动平均滤波器 (MAF)来消除噪音并确保时间步之间的平滑过渡。随后,这些轨迹被标准化为 30 个时间步的统一长度,从而能够在所有模式下一致执行。然后对轨迹进行

转化为机器人基础框架,并进一步转化为定义的桌子框架
在桌面表面。为了保证机器人的可执行轨迹,我们在夹持器中心的最小点处偏移了轨迹。

在任务执行过程中,起点是固定的,以确保可重复性,而终点则根据 YOLO [29] 检测到的紧急停止姿势动态确定。
这种实时物体检测的集成增强了系统的安全性和适应性,使其在不可预测的场景中也能保持稳健的性能。轨迹获取过程和相应的模板轨迹如图2a、b 所示。

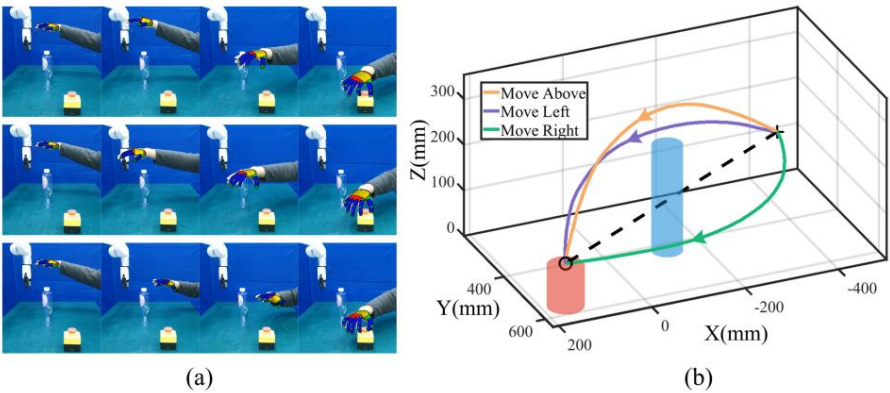


图 2. 轨迹获取过程。(a)跟踪人手运动的快照。第一行、中间行和最后一行表示向上、向左和向右移动的模式。(b)相同起点 (符号 “+”)和终点 (符号 “o”)处的三种模式的模板轨迹。基本轨迹用黑色虚线表示。

如图 2a 所示,Kinect FORTH 系统展示了实时跟踪人手运动的能力,提供了稳定而准确的手部质心估计。彩色指尖可以快速匹配人手的演示并实现准确的姿势估计。这种稳定性确保机器人仅通过一次演示就能有效地掌握人类的技能。经过预处理后,我们将模板轨迹概括为具有相同起点和终点的三种模式,如图2b 所示。模板轨迹概括没有中间的中途点。

机器人通过动态调整不同的起点和终点来适应实际任务条件。广义轨迹保持平滑度并保留人类演示的关键特征,从而使机器人动作与演示的技能非常相似。

生成每种模式的模板轨迹后,我们继续概括无碰撞笛卡尔轨迹集,以用于夹持器中心的最小点。我们首先确定时间步长和模板轨迹上每种模式的最大偏差对应的点。该点被选为经点。随后,我们确定经点的 XYZ 范围,并通过等间距离散化来调整范围。

对于上述模式,X 和 Y 在障碍物周围 ± 36.23 mm (安全障碍物半径)内,Z 在周围 ± 100 mm;对于左、右模式,我们将 X 固定,Y 和 Z 在周围 ± 100 mm。我们将范围与 Y 和 Z 的限值并集应用于任务执行。我们以相等的间距遍历 X、Y 和 Z 范围以获得特征经点集并概括初始轨迹集。我们根据机器人与障碍物之间的距离进一步微调轨迹。微调过程包括找到最危险的点,即最靠近障碍物的点,并调整轨迹以确保该点不会进入危险区域。对于上述模式,这通常涉及将 Z 坐标设置为安全值。在左、右模式下,Y 坐标在避开障碍物方面起着更关键的作用,我们调整 Y 坐标以保持安全距离。然后将这些微调点视为新的特征经由点,以细化和概括图 3a 所示的每种模式的无碰撞轨迹集。

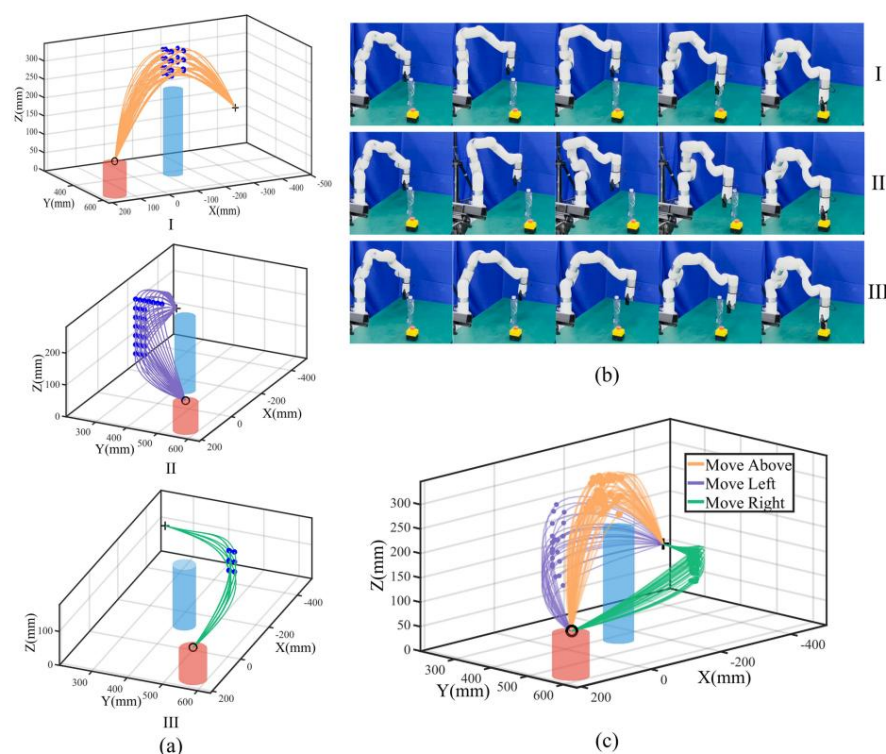


图 3.单个障碍物的实验。(a)广义无碰撞轨迹集。I、II 和 III 分别对应于具有蓝色特征途经点的上方、左侧和右侧模式。(b)任务执行快照。第一行、中间行和最后一行分别显示上方、左侧和右侧模式。(c)任务执行轨迹,其中橙色、紫色和绿色线(点)表示轨迹(特征途经点)。

我们进一步通过雅可比 IK 计算相应的无碰撞关节集并对 ProMP 进行编码。我们设置了 35 个基数来学习原始的 7-DoF ProMP,并制定任务约束以查找违反任务约束的特征时间步长。任务约束包括几个关键因素:(1) 关节范围限制确保每个关节在整个任务过程中保持在其指定范围内,这对于防止运动过程中的机械损坏或错误

至关重要。

(2)路点约束用于确保起点和终点。这些路点在任务空间中定义,对于确保机器人准确到达所需位置至关重要。正向运动学(FK)函数 T 根据关节配置计算末端执行器位置,我们设置精度要求 $dk, t = 1 \text{ mm}$,以确保高精度完成任务。

(3)超平面约束定义了机器人运动的空间限制。例如,最大 Z 平面被约束到由 $n_z, t = (0, 0, -1)$ 、 $b_z, t = (0, 0, 575)$ 定义的超平面,确保机器人的末端执行器在 Z 方向上保持在桌子下方。在实验中,我们将 Y 和 Z 的超平面约束定义为 $100 < Y < 700, 10 < Z < 575$ 。

(4)排斥器约束考虑了机器人各连杆之间可能发生的碰撞。例如,我们为机器人连杆6至连杆1定义排斥器约束,确保机器人在运动过程中避免与障碍物发生碰撞。本任务中的障碍物是一个水瓶,其安全障碍半径为 $R_{\text{safe}} = 36.23 \text{ mm}$ 。

对于所有任务约束,我们设置一个置信度参数 $\alpha_k, t = 0.95$ 来解释2倍标准差的不确定性,并使用矩匹配方法来估计非线性任务约束,例如航点、超平面和排斥器。此方法允许根据概率模型估计任务约束,如第2.2.2节所述。

我们用 λ_0 执行算法1 $\lambda = 100, \eta_k = 1$ 以获得优化的ProMP,并且 k, t 执行算法2执行任务。在每次实验中,我们从三种不同的模式中随机选择,使用高斯分布对路径点进行采样,以确保任务执行的多样性。我们重复实验100次以评估三种模式的性能;我们并行运行运动优化和自适应算法,总运行时间为1.79664秒,成功率为100%。结果如图3b,c所示。

结果表明,我们可以仅通过一次人工演示并借助经点轨迹泛化有效地生成可行的无碰撞轨迹集。

此外,通过使用ProMP对关节位置进行编码,离散的途经点可以生成连续的分布,代表执行任务的各种技能以及执行相同技能的灵活性。通过应用不同的任务约束,可以得出针对特定任务的新分布。这些分布可直接用于机器人任务执行,使机器人能够保留人类演示的特征,同时运用多种技能有效地完成任务。在本实验中,我们发现没有违反任务约束的特征时间步长。原因是航点约束和超平面约束由轨迹泛化确保,满足约束的概率始终为1。对于排斥器约束,水瓶高度不高,因此如果链接7是安全的,机器人链接6到链接1就不会发生碰撞。

同时,对于上面和左边的模式,如果机器人连杆7是安全的,考虑到机器人的配置,机器人连杆6到连杆1不会发生碰撞。因此,为了进一步评估所提出的方法,我们通过提高障碍物的高度、限制机器人绕着障碍物移动以及将场景扩展到多个障碍物来增加任务难度。

4.2. 有各种障碍的实验

为了测试所提出的任务约束运动自适应算法在涉及任务约束违规和多个障碍物的关键场景中的性能,我们设置了一个危险场景,其中存在三个高障碍物,机器人只能在障碍物周围移动。我们设置了适当的局部起点和终点以及特征经过点,以便链接7的部分无碰撞轨迹集显然会在其他链接处发生碰撞。

为了探索和验证机器人在这些条件下适应运动的能力,整个轨迹被分成三段。对于每一段,我们固定特征点的 X 和 Y 坐标,并减小 Z 坐标,以确保其他机器人连杆会与障碍物发生碰撞,如ROS Rviz模拟所示(图4a)。我们定义了与单障碍场景中相同的任务约束

第4.1节分为三个部分,不同之处在于某些时间步会违反任务约束,并可能与机器人链接 4、5 和 6 发生碰撞。

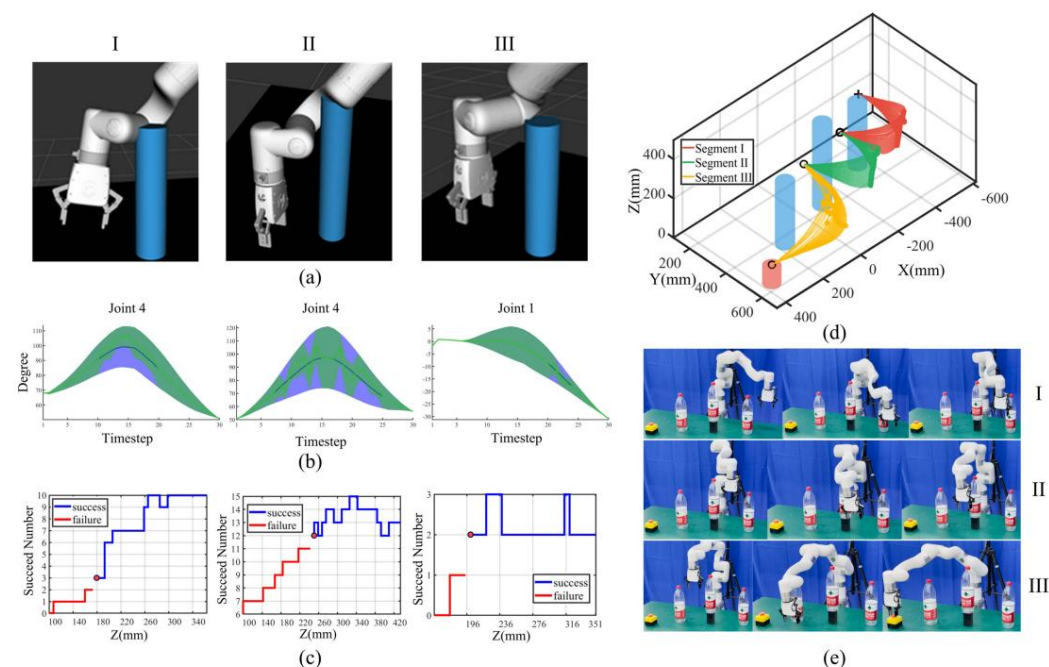


图 4.危险场景中的实验。I、II 和 III 表示三个轨迹段。(a) ROS rviz 中显示的危险情况,蓝色圆柱体表示障碍物。(b)任务约束下的典型原始(蓝色)和优化(绿色)ProMP。联合平均值(2 倍标准差)用粗线(阴影区域)表示。(c)阈值确定结果。蓝线和红线是导致成功和失败的 Z 位置,红点表示阈值。(d)三个段的任务执行轨迹,红线、绿线和黄线(点)表示广义轨迹(特征通过点)。(e)机器人任务执行快照。第一行、中间行和最后一行表示 I、II 和 III 轨迹段。

对于这三个段,在对模板轨迹进行经点轨迹泛化之后,我们获得了机器人连杆 7 的可行无碰撞轨迹集,并编码了初始 7-DoF 关节 ProMP 分布。然后,将此分布转换为每个时间步的高斯均值和标准差。利用矩匹配法,我们确定机器人连杆 4、5 和 6 可能与障碍物发生碰撞的特征时间步。我们估计这些时间步的任务约束概率,并将任务约束(包括关节约束、超平面约束、航路点约束和排斥器约束)纳入统一概率框架。利用算法 1 中的双循环 LBFGS 和 EMM 优化,我们获得了满足任务约束的优化 ProMP,如图 4b 所示。优化的关节分布表示机器人执行任务的可行关节范围,可作为雅可比逆运动学求解器的参考。

然而,考虑到 ProMP 的概率性质,优化过程中可能会排除一些满足任务约束的分布。为了改进这一点,我们进一步为可行关节范围设置了一个阈值。具体来说,我们允许机器人从经点轨迹泛化生成初始无碰撞轨迹,从而获得每个时间步对应的关节逆运动学解。然后,我们确定这些时间步的值中有多少个落在优化的关节范围内。设置一个阈值,如果有效时间步的数量超过此阈值,机器人将继续执行轨迹;否则,机器人停止。

对于阈值确定,机器人自动选择与初始泛化过程相同的路径点范围,并记录有效时间步数

关节范围和相应的避障结果。然后建立一个阈值,如果有效时间步数超过阈值,则认为机器人执行任务成功,如果低于阈值,则认为机器人执行任务失败。

阈值判定结果如图4c所示,机器人任务执行结果如图4d,e所示。

违反排斥器约束的特征时间步长分别为 10、15 和 3 个,在三个轨迹段中阈值分别为 3、12 和 2。在段 I、II 和 III 上并行运行算法 2,总时间为 315.07 秒。三个段中从 100 个特征经点执行的轨迹分别为 71、54 和 58 条,所有执行的轨迹均满足任务约束,成功率为 100%。

结果表明,机器人可以利用稀疏和离散特征经点对初始 ProMP 进行编码,并通过将所有任务约束以分析方式集成到概率框架中来优化 ProMP。由于该方法的概率性质,在优化过程中可能会排除一些满足任务约束的分布。因此,通过确定阈值进一步细化可行关节范围至关重要。此过程是自动化的,并利用轨迹泛化期间使用的相同离散和稀疏经点范围。通过为优化的关节范围添加额外的阈值,机器人可以仅使用一组离散的经点来得出在连续范围内执行任务的可行关节轨迹。具体而言,机器人生成广义轨迹,计算相应的关节值,并确定给定阈值时有多少时间步长的关节值落在优化范围内。然后将这些结果直接应用于任务执行,使机器人能够有效区分成功和失败的情况。总之,通过从连续高斯分布中采样,将阈值和优化的 ProMP 直接用于任务执行。因此,我们为机器人提供了以通用方式执行任务的能力;也就是说,机器人不需要在每次执行任务时确定任务约束是否满足。

另外,我们设置了不同的相对障碍物位置,并将障碍物高度设置为接近无限大,选择了两种情况下三个不同颜色的喷漆罐,夹持器位于两个相邻障碍物之间和上方,并将水瓶高度设置为1000 mm。重复100次实验,记录总时间分别为3.25874 s、2.45945 s和0.87882 s,结果如图5所示。对于不同的相对障碍物位置,直接用于任务执行的轨迹从三种模式下采样,不存在违反任务约束的时间步。对于接近无限高的障碍物,机器人链6、5和4总会在某些时间步发生碰撞,满足任务约束的概率始终为0。因此我们排除这些模式,只保留剩下的模式,所有实验的成功率均为100%。

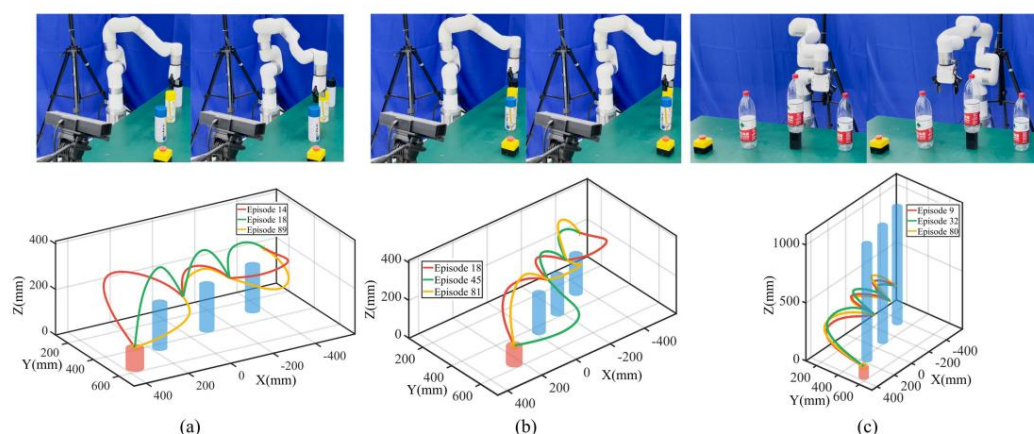


图 5.具有不同相对位置和几乎无限高的障碍物的实验。夹持器 (a)位于两个障碍物之间, (b)位于两个障碍物上方。 (c)具有几乎无限高的障碍物的情况。第一个

行表示局部的起点和终点,第二行显示用红色、黄色和绿色表示的典型轨迹。

结果表明,机器人能够自主选择面对不同相对障碍位置的局部泛化点,排除那些不能满足任务约束的模式,从而实现灵活的运动适应。机器人多样化的任务执行技能在此过程中起着至关重要的作用。当单一技能不足以完成任务时,机器人可以利用替代技能进行优化和适应。此外,如果所有可用技能都无法满足任务约束,则引入人工演示来提供额外指导,促进迭代演示和优化以应对这些挑战。

4.3. 对比实验 4.3.1. 经点轨

迹泛化性能

为了评估途经点轨迹泛化性能,我们比较了两个过程:没有途经点的人类演示模板轨迹生成(见图2)和具有特征途经点的初始无碰撞轨迹集泛化(见图3)。我们将我们的方法与没有途经点的DMP [12]和具有途经点的VMP [16]进行了比较。对于模板轨迹生成,考虑到没有途经点,使用VMP泛化的轨迹与我们的方法相同;因此我们选择DMP进行比较。对于初始无碰撞集泛化,考虑到DMP无法处理中间途经点,我们选择VMP进行比较。

对于模板轨迹生成,我们选择了第4.1节中的单障碍任务场景,并同时利用我们的方法和DMP生成模板轨迹,可视化轨迹如图6a所示。结果表明,即使任务的起点和终点发生变化,我们的方法也能保留人类的演示特征。然而,DMP会丢失演示信息;附加到上面和左边模式的轨迹将被推广到右边模式。原因是DMP对起点和终点的相对位置很敏感,导致轨迹与演示有显著偏差。

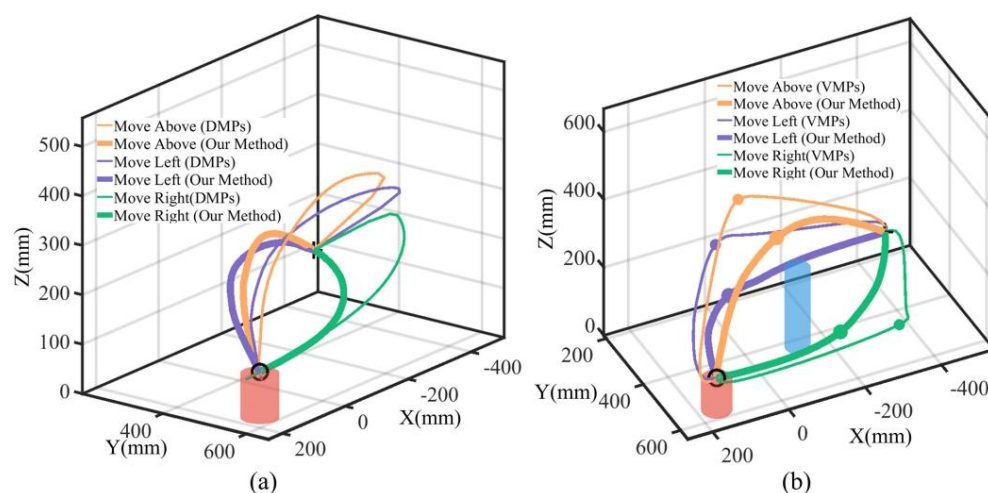


图 6.途经点轨迹泛化比较。(a) DMP 与我们的模板轨迹生成方法之间的比较。同一模式下的轨迹对用不同宽度的线表示。(b) VMP 与我们的无碰撞轨迹泛化方法之间的比较。同一模式下的轨迹和途经点用不同的宽度表示。

此外,我们设计了五种任务条件的实验,通过改变起点和终点,使它们的相对距离逐渐增大。对于每种任务条件,我们在 20 种不同的配置中改变了起点和终点。为了定量评估

生成的轨迹和演示之间的相似性,我们采用了余弦
通过计算每个轨迹段之间的相似度来衡量

$$M_{\text{相似度}} = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{(x_{t+1} - x_t) \cdot x_{t+1} - x_t}{\|x_{t+1} - x_t\| \|x_{t+1} - x_t\|}$$

(5)

其中 $M_{\text{similarity}}$ 表示计算的轨迹相似度, x_t 和 x_{t+1} 表示广义和演示轨迹向量, T 是总时间步长。该指标近似计算两个向量之间的余弦相似度,其中值越接近 1 表示相似度较高,表2给出了轨迹相似度度量结果。

表 2. DMP 和我们的方法的轨迹相似性测量结果（相对距离 r 表示实验期间观察到的归一化相对距离）。对于每个任务条件，我们在 20 种不同的配置中改变了起点和终点。

轨迹相似性	任务 1 ($r = 0-0.2$)	任务 2 ($r = 0.2-0.4$)	任务 3 ($r = 0.4-0.6$)	任务 4 ($r = 0.6-0.8$)	任务 5 ($r = 0.8-1$)
决策管理平台	0.9563 ± 0.1064	0.8973 ± 0.1564	0.7654 ± 0.1318	0.6424 ± 0.1389	0.5575 ± 0.2289
我们的方法	0.9845 ± 0.0184	0.9732 ± 0.0264	0.9667 ± 0.0187	0.9674 ± 0.0204	0.9574 ± 0.0584

从表2 中我们观察到,当开始和结束之间的位置偏移演示中的点很少,DMP 的轨迹相似性指标和我们的方法可以比较。在这种情况下,DMP 保留了人类的大部分特征演示。然而,随着位置偏移的增加,轨迹相似性度量我们的方法的效率始终保持较高水平,而 DMP 的效率则迅速下降。这证明我们的方法受任务初始和终止时间变化的影响较小条件,适应不同的任务条件。

对于带有中途点的初始无碰撞轨迹概括,给定模板轨迹,我们评估了不同任务场景下的表现,包括单个障碍物（见图3）以及分为I、II、III三个段的各各种障碍物（见图4）。我们选择了具有代表性的轨迹进行可视化,结果如图6b 所示。我们观察到 VMP 在途经点产生不自然的过渡,产生不连续的轨迹,并在途经点造成严重的中断,从而偏离人类演示的特征。定量测量轨迹平滑度,我们计算加速度的导数如下:

$$M_{\text{平滑度}} = \frac{1}{T-3} \sum_{t=1}^{T-3} \|x_{t+3} - x_t\|$$

(6)

其中 $M_{\text{smoothness}}$ 表示计算的轨迹平滑度,并且...是加速器 x_t 轨迹平滑度越接近 0,表示轨迹越平滑。我们评估了四种任务场景中的平滑度,结果总结如表3所示。结果表明,我们的方法在不同场景中始终能实现类似的平滑度。相比之下,在单障碍物情况下,VMP 的平滑度相对较低情况下,但在多障碍物场景中显着增加。与我们的方法相比,VMP 表现出的过渡不太平稳。

表 3. VMP 和我们的方法的轨迹平滑度结果。给定模板轨迹,我们评估了不同任务场景中的表现,包括单个障碍（见图3）以及I、II、III三个段中的各种障碍物（见图4）。

轨迹平滑度 单一障碍物	各种障碍		
	第一部分	第二部分	第三部分
病毒载体	0.1043 ± 0.0275	0.1843 ± 0.0395	0.2864 ± 0.0286
我们的方法	0.0752 ± 0.0146	0.0864 ± 0.0123	0.0796 ± 0.0452
			0.0852 ± 0.0204

对比实验表明,我们的方法可以从单个演示中生成适应不同任务条件的模板轨迹。它有效地保留了人类演示的关键特征,同时利用点泛化来生成平滑且无碰撞的初始轨迹集。这些轨迹实现了不同任务场景之间的平滑过渡,为后续优化过程提供了高质量的初始数据集。

4.3.2. 任务约束运动适应性能

本文针对机器人任务约束运动自适应问题,提出了算法1和算法2。算法1从初始ProMP分布开始,将其转换为每个时间步的高斯分布,并提前确定违反任务约束的时间步进行优化。然后对每个时间步的优化高斯分布进行再次优化,以获得优化的ProMP分布。算法2引入了额外的阈值作为ProMP优化关节范围的参考,使机器人能够直接执行任务,而无需重复进行任务约束验证。

该方法还可以扩展到多障碍物场景,确保在更复杂的环境中高效执行任务。为了进一步评估和比较不同任务场景中的稳健性,我们在不同的任务场景中对类似方法进行了全面比较,包括单个障碍物(见图3)、多个危险障碍物(见图4)、无限高的障碍物(见图5c)以及多个障碍物的不同相对位置(见图5a、b)。

(1)单障碍场景验证

重点关注算法1, [22]中提出了一种类似的方法,其中通过动作教学展示多种技能并将其编码为 ProMP 分布。

然而,与我们的方法不同,他们的方法通过预定义所有任务约束并使用无迹变换估计分布的均值和协方差来直接优化 ProMP 参数。然后,机器人直接使用优化后的 ProMP 分布来采样关节轨迹以执行任务。相比之下,我们的方法概括离散的路径点以生成连续的 ProMP 分布,并在优化 ProMP 分布后采用基于阈值的机制来执行任务。

为了确保公平比较,我们对两种方法使用了相同的初始 ProMP 分布。然后,我们使用我们的方法和[22]中概述的方法优化了 ProMP 分布,并将得到的优化分布应用于任务执行。在这次比较中,我们优先考虑优化性能,而不是解决多障碍环境中的运动适应问题。因此,我们首先考虑单障碍情况,因为多障碍任务可以分解为连续的单障碍阶段。具体来说,我们设计了四种任务场景:一个障碍物(见图3)和三个连续的段(I、II 和 III),其中有一个高危险障碍物(见图4)。为了评估算法在不同任务场景中的稳健性,对于每种场景,我们广泛改变了障碍物的位置以及起点和终点,总共产生了 30 个测试用例。

使用优化参数数量、优化时间、任务成功率以及初始分布与优化分布之间的损失等指标来评估机器人的任务执行情况。这里,初始分布与优化分布之间的损失是指优化后所有时间步的关节范围与原始关节范围的总体比例。我们记录指标并在表4中总结结果。

表4表明,尽管与[22]中的方法相比,我们的方法总体上优化了更多参数,但优化时间却显著缩短。原因是我们只优化了那些违反任务约束的时间步。在单一障碍的场景中,在大多数时间步中,高斯分布已经满足任务约束,因此我们的方法可以直接使用优化后的分布来执行任务。

相比之下,参考文献[22]在每个时间步上优化所有 ProMP 参数,这可能会改变最初满足任务约束的分布。这种过度优化可能导致

任务失败,因为机械臂可能会消除能够完成任务的动作。
虽然机器人仍可以执行一些任务,但其在执行过程中所展示的运动范围训练减少了。

表 4.与单一障碍情景的比较。我们设计了四个任务情景:一个涉及单个障碍物 (见图3)和三个连续的段 (I、II和III),其中I、II和III具有三个高危险障碍 (见图4)。每个场景涉及30种任务条件的变化。

任务场景	方法	优化参数	优化时间	成功率	配送损失
一个障碍	[22]中的算法	665	20.4586 秒± 4.2367 秒	96.67%	85.46% ± 3.24%
	我们的方法	679 ± 14	1.7904 秒± 0.1986 秒	100%	98.74% ± 0.16%
第一部分	[22]中的算法	665	1204.32 秒± 48.52 秒	63.33%	64.72% ± 5.79%
	我们的方法	749 ± 56	346.08 秒± 86.34 秒	93.33%	85.48% ± 3.56%
第二部分	[22]中的算法	665	1365.38 秒± 49.37 秒	73.33%	62.62% ± 4.28%
	我们的方法	805 ± 70	479.28 秒± 65.28 秒	90.00%	79.49% ± 4.68%
第三部分	[22]中的算法	665	864.37秒± 32.58秒	80.00%	73.52% ± 5.62%
	我们的方法	707 ± 42	273.58秒± 57.42秒	96.67%	89.37% ± 3.56%

为了防止概率优化排除某些分布,我们的方法为优化的联合分布引入了一个额外的阈值,并用于作为雅可比逆运动学求解器的参考。当逆运动学求解器的数量运动学解超过此阈值,机器人成功执行轨迹。这种基于阈值的机制确保机器人保留来自优化过程,允许更灵活的技能选择。因此,我们的方法通过更好地利用已证明的技能实现了减少分布损失。

在 I、II 和 III 段中,对于三个高障碍物,参考文献[22]往往会失去显著优化过程中的信息,导致任务失败,尤其是在移动障碍物时在右侧。我们的方法在任务执行中仍然成功,即使在高风险的情况下条件,随着场景难度的增加,产生更多可行的动作。请注意,我们的方法在阈值与边界相交的边缘情况下可能会失败连续空间。换句话说,一旦定义了阈值,接近该阈值的轨迹最初可能会执行,机器人执行后,就会发生故障。增加中途点数量或牺牲某些分布可能会克服这个缺点,提高任务成功率。

(2)多障碍场景验证

我们还在多障碍场景中评估了我们的方法,并将其与现成的方法进行了比较
机器人运动自适应算法,如基于采样的 RRT-Connect [30]和基于概率的 STOMP [31]。重点关注多个危险障碍物 (见图4),无限高的障碍物 (见图5c),以及多个障碍物的不同相对位置 (见图5a、b),我们改变了 30 种配置中障碍物的相对位置,每个场景。

为了便于可视化,图 7 显示了一条代表性轨迹,其中相同的起点和终点用于比较我们的方法与 STOMP 和 RRT-Connect。生成的轨迹根据预定义类别,结果如图7 所示。机器人从三个类别中随机选择模式,同时记录 RRT-Connect 和 STOMP 的轨迹并分为三种模式。如果这些轨迹不属于这三种模式中的任何一种模式,我们将它们分开记录,并将它们定义为模糊模式。我们发现 RRT-Connect 和 STOMP 轨迹混乱且不稳定,有时机器人的轨迹蜿蜒曲折 (见图7b),导致相邻的时间步长。

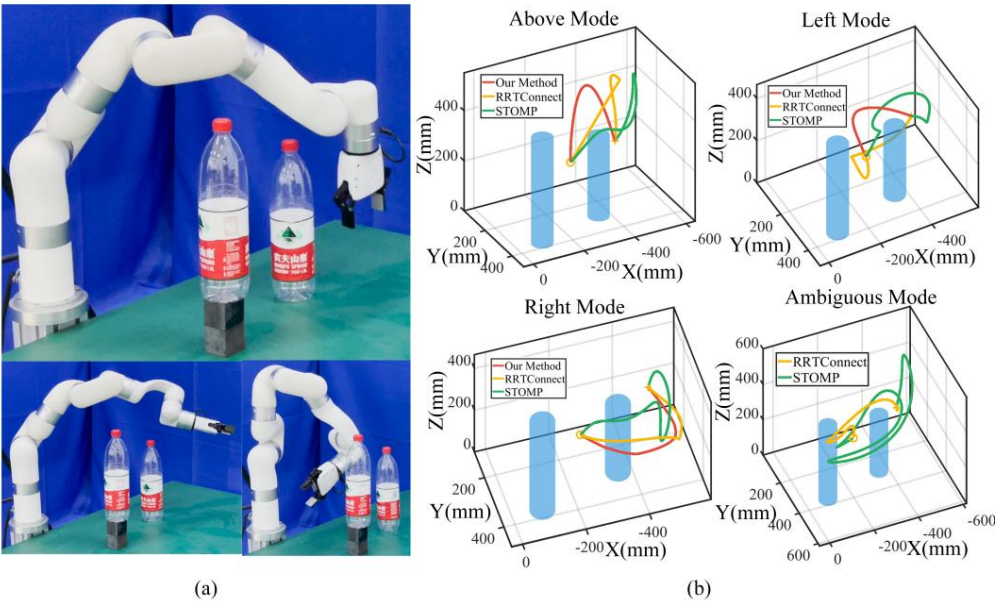


图 7.运动适应性比较。(a)任务执行姿势的快照。第一个第一行是我们的方法,第二行是STOMP和RRTConnect。(b)不同的模式轨迹。算法 2、RRTConnect 和 STOMP 生成的轨迹以红色表示,黄色和绿色。不属于任何模式的轨迹被视为模糊模式。

为了定量评估性能,我们选择了几个指标来测试上述任务场景中,障碍物的相对位置随机变化。我们将两个相邻时间步之间的位姿误差归一化和 ΔP 定义为
$$\Delta P = \sum_{i=1}^{T-1} q_{i+1} - q_i \text{ 且调制路径 } L \text{ 为 } L = \sum_{i=1}^{T-1} |q_{i+1} - q_i|$$
 相应的能源效率定义为 $\eta = \frac{\text{最小路径长度}}{L_0}$ 能量效率越接近1,越高 机器人执行任务的效率。我们还评估了不同的任务场景,以任务成功率来衡量。此外,由于我们的方法涉及离线优化,然后进行在线执行,我们计算了离线优化时间和在线执行时间。结果如表5所示。

表 5.多障碍物场景比较。对于多个危险障碍物（见图4），无限高的障碍物（见图5c），以及多个障碍物的不同相对位置（见图5a、b)中,我们在每个场景中的 30 种配置中改变了障碍物的相对位置。

任务场景	方法	姿势错误	路径长度	能源效率	成功率	离线时间	在线时间
危险	踩踏	0.81 ± 0.19	10.37 米 ± 1.52 米	63.25% ± 5.76%	RRT-Connect 0.87	96.67%	0.063 秒 ± 0.023 秒
		± 0.13	11.36 米 ± 2.34 米	59.25% ± 6.38%	我们的方法 0.15 ± 0.05	100%	0.054 秒 ± 0.068 秒
		84.62% ± 2.34%				96.67%	304.2 秒 ± 42.4 秒
无限高	STOMP	0.75 ± 0.14	6.72 米 ± 0.96 米	68.64% ± 3.28%	7.48 米 ± 2.31 米	83.33%	0.123 秒 ± 0.032 秒
	RRT-Connect	0.82 ± 0.25	6.72 米 ± 0.96 米	68.64% ± 3.28%	7.48 米 ± 2.31 米	80.00%	0.086 秒 ± 0.029 秒
	我们的方法	0.22 ± 0.12	6.72 米 ± 0.96 米	68.64% ± 3.28%	7.48 米 ± 2.31 米	93.33%	0.017 秒 ± 0.006 秒
变化	踩踏	0.76 ± 0.13	9.52 米 ± 1.46 米	56.72% ± 4.38%	RRT-Connect 0.68 ± 0.25	100%	0.076 秒 ± 0.024 秒
		± 0.25	10.28 米 ± 2.74 米	53.47% ± 2.42%	我们的方法 0.09 ± 0.02	100%	0.046 秒 ± 0.013 秒
		± 1.59%				100%	0.012 秒 ± 0.003 秒

结果表明,我们的方法可以生成更短的调制路径,但需要相当多的离线规划时间。这可以归因于我们的方法充分考虑了任务约束,从而代表了多种可行的技能完成任务。优化后,我们直接执行更短的轨迹在线执行时间。STOMP 和 RRT-Connect 计算并采样单个轨迹无需考虑多任务执行技能即可执行。它们表现出显著的

随机性,有时为了完成任务而绕了很长的弯路,而直接路线就足够了。虽然施加额外的约束可以减少这种随机性,但机器人的行为往往缺乏一致性和连续性。相比之下,我们的方法保持了轨迹的连续性,姿势变化较小,从而促进了更平稳的过渡并降低了能耗。

同时,我们的方法在涉及无限高障碍物的场景中表现出更高的成功率,原因是 STOMP 和 RRT-Connect 通常由于障碍物的高度无限而难以找到成功的轨迹。在某些情况下,它们会在任务中途失败或在接近高障碍物时陷入局部最小值。相比之下,我们的方法能够排除不可行的技能并选择替代技能来完成任务,从而实现更灵活和适应性更强的动作。

然而,在通过点泛化无法生成有效的无碰撞轨迹的关键情况下,我们的方法很容易失败。

5. 结论5.1. 意义本文

提出了一种基于

ProMP 的机器人任务约束优化和适应方法。该方法允许机器人在各种任务约束下调整其运动并扩展到多个障碍物,重点关注不同技能的任务执行灵活性。我们提出了一种通过点轨迹泛化方法,该方法仅从每个技能类别的一个人类演示中学习。该方法有效地适应了不同的任务条件,同时保留了人类演示的关键特征。

离散的经由点生成用ProMP编码的连续分布,代表执行任务的各种技能以及执行相同技能的灵活性。

给定初始 ProMP,我们将所有任务约束纳入统一的概率框架,将 ProMP 解耦为每个时间步的高斯分布,并仅优化违反任务约束的时间步,从而显著缩短优化时间。为了防止概率优化排除某些分布,我们的方法为优化的关节分布引入了一个额外的阈值,作为雅可比逆运动学求解器的参考。当逆运动学解的数量超过此阈值时,机器人成功执行轨迹。

这种基于阈值的机制确保机器人从优化过程中保留有用的信息,从而实现更灵活的技能选择。

同时,我们提出了一种针对多障碍物的运动适应方法,使机器人能够根据不同的相对障碍物位置自主选择局部泛化点,并排除无法满足任务约束的模式。当单一技能不足以完成任务时,机器人可以利用替代技能进行优化和适应。此外,如果所有可用技能都无法满足任务约束,则引入人工演示以提供额外指导,促进迭代演示和优化以应对这些挑战。

5.2. 限制

然而,我们的方法有一定的局限性,在特定场景下可能会失败。它依赖于初始无碰撞末端执行器轨迹集的质量。在有多障碍物的密闭空间中,概括轨迹以生成大量无碰撞轨迹是一项挑战。因此,后续优化缺乏高质量的数据集,使机器人难以优化能够用有限的数据集完成任务的分布。

此外,我们使用离散的路径点来定义阈值。如果所选的路径点太稀疏,则很难在连续空间中表示分布。在极端阈值位置,优化结果可能表明机械臂可以执行任务,但实际任务执行通常会失败。此外,即使机械臂生成大量初始无碰撞轨迹,尽管进行了优化,它仍可能无法满足所有任务约束。遇到障碍物时尤其明显

无限高度。虽然末端执行器可以避免碰撞,但其他关节的分布可能会违反任务约束,导致无法选择适当的技能来完成任务。

此外,我们的方法专门针对简单的拾取和放置任务。
演示数量有限可能会影响在需要多样化场景的复杂任务中的通用性。因此,对于涉及对象操作的更复杂任务,需要更细粒度的途径点泛化方法来生成高质量的数据集以供后续优化。在有多个障碍物的情况下,我们的方法根据相对位置确定局部起点和终点。然而,更复杂的障碍物分布需要更启发式的运动模式切换策略来确保任务成功。同时,我们的方法无法处理多个实时动态障碍物。

5.3. 未来工作

为了解决上述限制,未来的工作应侧重于更复杂的物体操作任务,探索从人类演示中生成高质量初始可行运动集的方法。这包括改进路径点泛化方法,并结合运动分割技术将运动划分为多个轨迹段,每个轨迹段都将被泛化和组合以生成高质量数据集。此外,还需要进一步探索阈值确定过程,确保离散路径点与连续空间的分布更好地对齐。此外,未来的研究应该探索受生物启发的运动模式切换技术,例如从甲虫的自然行为中衍生的无梯度控制方法,用于同时进行静态和动态避障,如[32]中所述。将这种方法与我们的任务约束相结合可以增强其对实时应用的适用性。

作者贡献:概念化,GD,XZ (Xizhe Zang) 和 XZ (Xuehe Zhang);方法论,GD;软件,XZ (Xizhe Zang) 和 XZ (Xuehe Zhang);验证,GD;形式分析, GD;调查,GD;资源,XZ (Xizhe Zang)、XZ (Xuehe Zhang)、CL 和 JZ;数据管理,GD;写作初稿,GD;写作 审查和编辑,GD;可视化,GD和 CL;监督,XZ (Xizhe Zang)、XZ (Xuehe Zhang)、CL 和 YZ;项目管理, YZ 和 JZ;资金获取,YZ 和 JZ 所有作者均已阅读并同意稿件的出版版本。

资金资助:本研究得到先进技术船舶研究项目 (CBZ02N23-05)、国家自然科学基金优秀青年科学基金项目 (52025054) 和国家自然科学基金重大研究计划 (92048301)的资助。

机构审查委员会声明:不适用。

数据可用性声明:研究中提出的原始贡献包含在文章中,进一步的询问可以直接联系相应的作者。

利益冲突:作者声明没有利益冲突。

参考

1. Alrashdi, I.;Alqazzaz, A. 协同人工智能、物联网和区块链在智慧城市中诊断大流行性疾病:挑战与机会。 可持续发展。 机器智能杂志2024,7,1-28。 [CrossRef]

2. Hussein, HI;Mohammed, AO;Hassan, MM;Mstafa, RJ 基于轻量级深度 CNN 的模型,用于早期检测通过胸部 X 光片识别 COVID-19 患者。Expert Syst. Appl. 2023, 223, 119900。 [CrossRef] [PubMed]

3. Hong, Q.;Lin, L.;Li, Z.;Li, Q.;Yao, J.;Wu, Q.;Liu, K.;Tian, J. 一种用于 cxr 图像分类的混合特征融合距离变换深度森林框架。IEEE 神经网络学习系统学报, 2023,13,14633–14644。 [CrossRef] [PubMed]

4. Hashem, IAT;Usmani, RSA;Almutairi, MS;Ibrahim, AO;Zakari, A.;Alotaibi, F.;Alhashmi, SM;Chiroma, H.可持续智慧城市的城市计算 :最新进展、分类和开放研究挑战。《可持续性》 2023,15,3916。
[交叉引用]

5.福尔加多,FJ;卡尔德隆,D.;冈萨雷斯,I.; Calderón,AJ 从自动化和监管角度回顾工业 4.0 系统 :定义、架构和最新趋势。电子2024,13,782。 [CrossRef]

6. Barekatain, A.;Habibi, H.;Voos, H. 制造业机器人操作人员从演示中学习的实用路线图。Robotics 2024,13,100。 [CrossRef]

7. Jaquier, N.; Zhou, Y.; Starke, J.; Asfour, T. 学习通过可微分优化对机器人技能进行排序和混合。IEEE 机器人。自动。莱特。 2022, 7, 8431–8438。 [交叉参考]

8. Wang, B.;Gong, X.;Lyu, P.;Liang, S. 基于迭代学习的联网协作运动规划与决策以及匝道处的自动驾驶汽车协调。IEEE 智能交通系统汇刊2024,25,8105–8120。 [CrossRef]

9.肖萌; 张晓玲; 张婷; 陈珊; 邹艳; 吴伟. 基于GMM/GMR算法的机器人力控制研究融合不同的补偿策略。前沿。神经机器人。2024,18,1290853。 [CrossRef]

10. von Hartz, JO;Welschhold, T.;Valada, A.;Boedecker, J. 模仿的艺术 :从一些演示。IEEE Robot.&Autom.&Lett.2024,9,11369–11376。 [CrossRef]

11. Fu, J.;Ye, X. 基于 ProMPs 的机器人可变阻抗技能习得框架。2024 年国际智能机器人与系统研讨会 (ISIRS) 论文集,中国长沙,2024 年 6 月 14-16 日;第 108-112 页。

12.艾斯皮特,AJ;中西,J.;霍夫曼,H.;牧师,P.; Schaal, S. 动态运动基元 :学习吸引子模型运动行为。神经计算。2013,25,328–373。 [CrossRef] [PubMed]

13. Ding, G.;Liu, Y.;Zang, X.;Zhang, X.;Liu, G.;Zhao, J. 一种从人类演示中学习机器人装配任务的任务学习策略。传感器2020,20,5505。 [CrossRef] [PubMed]

14.袁勇; 余志良; 华玲; 程颖; 李建军; 桑鑫. 层次化动态运动基元在机器人运动平稳性方面的应用基于深度强化学习的机器人。Appl. Intell. 2023,53,1417–1434。 [CrossRef]

15. Li, G.;Jin, Z.;Volpp, M.;Otto, F.;Lioutikov, R.;Neumann, G. ProDMP:动态和概率的统一视角运动原语。IEEE Robot.&Autom.&Lett. 2023,8,2325–2332。 [CrossRef]

16. Zhou, Y.;Gao, J.;Asfour, T. 学习具有内插和外推功能的点移动原语。2019年 IEEE/RSJ 国际智能机器人与系统会议 (IROS) 论文集,中国澳门,2019 年 11 月 3 日至 8 日;第 4301-4308 页。

17.胡颖;崔梅;段建军;刘伟;黄丹;Knoll A;陈刚. 模仿学习的模型预测优化演示。机器人自动系统。2023,163,104381。 [CrossRef]

18. Duan, A.;Camoriano, R.;Ferigo, D.;Calandriello, D.;Rosasco, L.;Pucci, D. 受约束的 DMP 用于人形机器人的可行技能学习。2018 年 IEEE-RAS 第 18 届人形机器人 (Humanoids)国际会议论文集,中国北京,2018 年 11 月 6 日至 9 日;第 1-6 页。

19. Paraschos, A.;Lioutikov, R.;Peters, J.;Neumann, G. 移动原语的概率优先级排序。IEEE Robot. Autom. Lett. 2017,2,2294–2301。 [交叉引用]

20. Koert, D.;Maeda, G.;Lioutikov, R.;Neumann, G.;Peters, J. 基于演示的可推广机器人运动轨迹优化。2016 年 IEEE-RAS 第 16 届人形机器人 (Humanoids)国际会议论文集,墨西哥坎昆,2016 年 11 月 15-17 日;第 515-522 页。

21.Koert ,D.;帕贾里宁,J.;肖特施奈德,A.;特里克,S.;罗斯科普夫,C.; Peters, J. 学习意图感知在线适应运动原语。IEEE Robot. Autom. Lett. 2019,4,3719–3726。 [CrossRef]

22. Frank, F.;Paraschos, A.;van der Smagt, P.;Cseke, B. 用于机器人轨迹自适应的受限概率运动原语。IEEE 机器人学报。2021,38,2276–2294。 [CrossRef]

23. Oikonomidis, I.;Kyriazis, N.;Argyros, AA 使用 Kinect 实现基于模型的高效 3D 手部关节跟踪。论文集英国机器视觉会议,英国邓迪,2011 年 8 月 29 日至 9 月 2 日;第 1 卷,第 3 页。

24. Paraschos, A.;Daniel, C.;Peters, J.;Neumann, G. 在机器人技术中使用概率运动原语。Auton. Robot. 2018, 42,529–551。 [交叉引用]

25. McDonald,GC Ridge 回归。Wiley Interdiscipline。Rev. Comput. Stat.2009,1,93–100。 [CrossRef]

26. Bertsekas,DP 非线性规划。J. Oper. Res. Soc. 1997,48,334。 [CrossRef]

27. Berahas, AS;Taká v c, M. 一种用于机器学习的稳健多批次 L-BFGS 方法。优化方法软件。2020,35,191–219。
[交叉引用]

28.阿巴迪,M.;阿加瓦尔,A.;巴勒姆,P.;布雷夫多,E.;陈,Z.;雪特罗,C.;科拉多,GS;戴维斯,A.;迪安,J.;德文,M.;等人。张量流 :异构分布式系统上的大规模机器学习。arXiv 2016, arXiv:1603.04467。

29. Adarsh, P.;Rathi, P.;Kumar, M. YOLO v3-Tiny :使用单阶段改进模型进行对象检测和识别。2020年第 6 届国际先进计算和通信系统会议 (ICACCS) 论文集,印度哥印拜陀, 2020 年 3 月 6-7 日;第 687-694 页。

30. Klemm, S.;Oberländer, J.;Hermann, A.;Roennau, A.;Schamm, T.;Zollner, JM;Dillmann, R. Rrt-connect :更快、渐近最优的运动规划。2015 年 IEEE 机器人与仿生学国际会议 (ROBIO) 论文集,中国珠海,2015 年 12 月 6-9 日;第 1670-1677 页。

31. Kalakrishnan, M.; Chitta, S.; Theodorou, E.; Pastor, P.; Schaal, S. STOMP: 运动规划的随机轨迹优化。2011 年 IEEE 国际机器人与自动化会议论文集, 中国上海, 2011 年 5 月 9-13 日; 第 4569-4574 页。

32. Khan, A.T.; Li, S.; Cao, X. 智能家居中利用甲虫触角搜索进行人类引导的合作机器人代理。中国科学: 信息与通信技术, 第 51 卷, 第 2899-2915 页。Sci. 2022, 65, 122204. [[CrossRef](#)]

免责声明/出版商说明: 所有出版物中包含的声明、观点和数据仅代表个人作者和贡献者的观点, 不代表 MDPI 和/或编辑的观点。MDPI 和/或编辑对因内容中提及的任何想法、方法、说明或产品而导致的任何人身伤害或财产损失不承担任何责任。