

一种小样本支持向量机控制器在两足机器人步态控制的研究

王丽杨^{1,2}, 刘 治¹, 赵之光¹, 章 云¹

(1. 广东工业大学 自动化学院, 广东 广州 510006; 2. 顺德职业技术学院 电子工程系, 广东 顺德 528300)

摘要: 神经网络等传统的机器学习方法是基于样本数目无穷大的经验风险最小化原则, 这对非确定环境下有限样本的步态学习控制非常不利. 针对两足机器人面临的非确定环境适应性难题, 提出了一种基于支持向量机(SVM)的两足机器人步态控制方法, 解决了小样本条件下的步态学习控制问题. 提出了一种基于混合核的步态回归方法, 仿真研究表明了这种方法比全局核和局部核分别单独用于步态学习时有优越性. SVM以踝关节及髋关节的轨迹作为输入, 相应的满足ZMP判据的上体轨迹作为输出, 利用有限的理想步态样本对机器人上体轨迹与腿部轨迹之间的动态运动关系进行学习, 然后将训练好的SVM置入机器人控制系统, 从而增强了步态控制的鲁棒性, 有利于实现两足机器人在非结构环境下的稳定步行. 仿真结果表明了所提方法的优越性.

关键词: 两足机器人; 步态; 小样本; 学习控制; 支持向量机

中图分类号: TP273 **文献标识码:** A

Support-vector-machines learning controller based on small sample sizes for biped robots

WANG Li-yang^{1,2}, LIU Zhi¹, ZHAO Zhi-guang¹, ZHANG Yun¹

(1. Faculty of Automation, Guangdong University of Technology, Guangzhou Guangdong 510006, China;

2. Department of Electronic Engineering, Shunde Polytechnic, Shunde Guangdong 528300, China)

Abstract: Conventional machine learning methods such as neural network(NN) use empirical risk minimization(ERM) based on infinite samples, which is disadvantageous to the gait learning control based on small sample sizes for biped robots walking in unstructured, uncertain and dynamic environments. To deal with the stable walking control problem in the dynamic environments for biped robots, we put forward a method of gait control based on support-vector-machines(SVM), which provides a solution for the learning control issue based on small sample sizes. A support vector machine regression(SVR) method for gait control with mixed kernel functions is proposed, and the proposed method shows superior performance when compared with SVR with radial basis function(RBF) kernels or polynomial kernels, respectively. Using ankle trajectory and hip trajectory as inputs, and the corresponding trunk trajectory which guarantees the ZMP criterion as outputs, the SVM is trained based on small sample sizes to learn the dynamic kinematic relationships between the legs and the trunk of the biped robot. Then the trained SVM is incorporated into the control system of the robots. Robustness of the gait control is enhanced, which is advantageous to realizing stable biped walking in unstructured environments. Simulation results demonstrate the superiority of the proposed methods.

Key words: biped robots; gait; small sample sizes; learning control; support-vector-machines(SVM)

1 引言(Introduction)

非结构化步行环境会使机器人腿部轨迹偏离规划值, 导致机器人因上体位置与下肢姿态不协调而摔倒. 如何利用有限样本进行步态学习, 使机器人具备在非确定环境中步行的自适应能力, 是两足机器人步态控制领域面临的难题.

目前已有多种机器学习方法在机器人学习控制方面取得进展^[1~5], 包括神经网络^[6~9]、强化学习^[10~12]、进化学习^[13,14]、支持向量机(SVM)^[15,16]等. 其中, 神经网络等传统的机器学习方法采用

基于样本数目无穷大的经验风险最小化原则(ERM), 这对非确定环境下有限样本的步态学习控制非常不利. Vapnik等人^[17]提出的SVM理论广泛用于解决分类、回归和概率密度估计等问题. 与传统基于经验风险最小化原则的方法不同, SVM方法基于结构风险最小化原则, SVM最小化泛化误差上界而ERM最小化训练误差上界, 因此SVM在解决学习精度和推广性之间的矛盾方面有明显的优势. 尤其值得关注的是, 传统学习方法需要获得大量具有泛化能力的样本, 而SVM能有效解决小样本问题, 具有很强的泛

化能力. 这对于解决机器人在不确定环境下稳态步行的小样本问题非常合适.

本文针对两足机器人面临的非确定环境适应性难题, 提出了一种基于SVM的两足机器人步态控制方法, 解决了小样本条件下的步态学习控制问题. 提出了一种基于混合核的步态回归方法, 解决了全局核和局部核分别单独用于步态学习时逼近能力不足或泛化能力较差的问题.

2 基于小样本数据的两足机器人SVM学习控制(SVM learning control based on small sample sizes for biped robots)

小样本条件下, 为实现两足机器人在非结构化环境下的稳态步行, 提出了一种基于SVM的步态控制方法. 在步态规划的基础上, 利用理想步态样本

集对SVM进行离线训练, 采用一种混合核函数保SVM学习机具备小样本前提下的最佳学习能力.

如图1所示, 基于SVM的步态控制机制可以描述为一个双闭环的反馈控制系统. 其中, 内环关节PID控制器跟踪给定的关节参考轨迹, 目前已有比较成熟的实现方法. 外环步态控制方面, 传统的步态控制器能够实现结构化环境下的稳态步行, 然而如何利用有限样本进行步态学习, 使机器人在非结构化环境下保持步态稳定, 是两足机器人步态控制领域面临的难题. 本文提出的基于SVM的步态学习控制方法, 构造一种由全局核和局部核加权构成的混合核函数, 实现了小样本情况下控制器逼近能力和泛化能力的平衡优化, 从而增强了步态控制的鲁棒性, 解决了小样本条件下的步态学习控制问题.

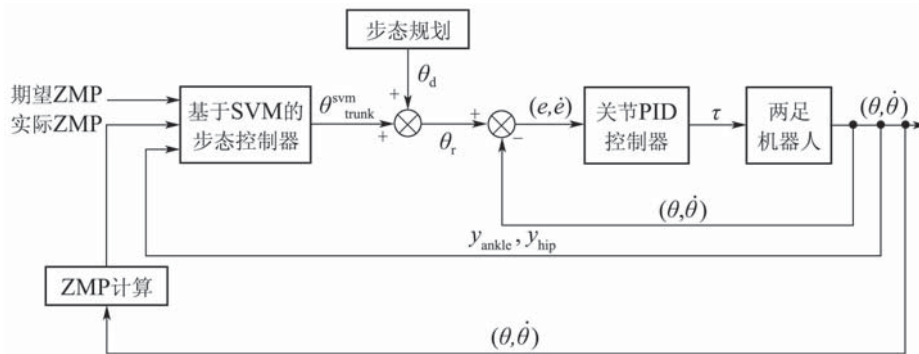


图1 基于SVM的步态控制机制示意图

Fig. 1 Schematic of SVM-based gait control mechanism

2.1 基于SVM的步态控制机制(Gait control mechanism based on SVM)

基于SVM的步态控制机制也即图1所示的双闭环反馈控制机制. 内环PID控制器跟踪机器人关节的参考轨迹; 外环基于SVM的步态控制器则根据机器人踝、髋实时位置在线调整机器人的上体位置, 通过满足两足步行机构特有的ZMP稳定判据来实现机器人系统的稳定步行运动.

2.1.1 内环关节PID控制(Joint control in the internal-loop based on PID)

内环的关节PID控制器实施关节位置伺服控制. 内环控制器的输入为各关节跟踪误差及误差变化, 输出为关节控制力矩. 内环控制器为实现关节位置跟踪误差为零的控制目标, 采用如下的工作方式:

如图1所示, 离线规划的关节期望轨迹为 θ_d , 此期望轨迹经外环控制器在线修正后为 θ_r . 位置传感器采样并经A/D转换得到第 i 时刻关节位置 $\theta(i)$, 而后与第 i 时刻关节参考位置 $\theta_r(i)$ 相比, 求得当前位置误差 $e(i) = \theta_r(i) - \theta(i)$, 以及 $i - 1$ 时刻位置误

差 $e(i - 1)$, 并得出误差变化 $\dot{e}(i)$. 将关节位置误差及误差变化 (e, \dot{e}) 输入给关节PID控制器, 经过控制算法, 输出控制力矩 τ , 控制电机转轴驱动减速器和关节运动.

2.1.2 外环基于SVM的步态控制(Gait control in the outer-loop based on SVM)

外环基于SVM的步态控制器实施关节位置协调控制. 外环控制器的输入为踝关节位置和髋关节位置, 输出为在线修正的期望上体轨迹. 外环控制器为实现ZMP轨迹跟踪误差为零的控制目标, 采用如下的工作方式:

位置传感器和速度传感器采样并经A/D转换得到第 i 时刻关节状态 $(\theta(i), \dot{\theta}(i))$, 经ZMP计算后得到当前实际ZMP位置 $y_{zmp}(i)$, 将它与第 i 时刻期望ZMP位置 $y_{zmp}^d(i)$ 相比较. 当 $y_{zmp}^d(i) - y_{zmp}(i) \neq 0$ 时, 考虑到外部环境及机器人状态等不可抗因素, 控制器并不直接调整偏离规划的机器人踝关节及髋关节位置, 而是通过调整上体轨迹使实际ZMP向期望ZMP移动. 通过基于SVM的步态控制算法, 外环控制器输出期望上体轨迹 θ_{trunk}^{svm} .

与步态规划给出的期望上体轨迹 θ_{trunk}^d 不同, 外环控制器输出的 θ_{trunk}^{svm} 为在线修正的期望上体轨迹. 离线规划的期望上体轨迹 θ_{trunk}^d 与在线修正的期望上体轨迹 θ_{trunk}^{svm} 二者合并构成外环控制器的参考轨迹 θ_{trunk}^r , 也即 $\theta_{trunk}^r = \theta_{trunk}^d + \theta_{trunk}^{svm}$.

总的来说, 作为离线步态规划的补充, 基于SVM的在线调整控制方法充分发挥了上体姿态对于两足机器人步行稳定的平衡作用. SVM以踝关节及髋关节的轨迹作为输入, 相应的满足ZMP判据的上体轨迹作为输出, 利用有限的理想步态样本对机器人上体轨迹与腿部轨迹之间的动态运动关系进行学习, 其本质在于控制上体轨迹, 使两足机器人的实际ZMP向期望ZMP移动, 从而实现两足机器人在非结构环境下的稳定步行.

2.2 步态样本分析(Analysis of the gait samples)

地面特性、障碍物尺寸、外力等环境因素, 以及负重、能量等机器人状态的变化, 常常导致机器人原有的步态和控制方法不再适用, 这也决定了非确定环境下机器人步态学习控制的小样本特性. 这种情况下, 机器人必须充分利用原有的理想步态样本, 通过合适的学习机制将有限样本中的客观规律应用到新的环境和状态中去, 才能够确保两足机器人在非确定环境下的稳态步行.

满足ZMP稳定条件的上体轨迹与腿部轨迹之间存在某种可以学习的非线性运动动态^[4,6,18]. 为了获取满足这种动态关系的步态样本, 可控制机器人处于ZMP稳定的步行周期内, 按照时间序列对两足机器人的支撑腿髋关节、摆动腿踝关节及上体轨迹实施采样. θ_{trunk}^{svm} 与 y_{ankle} , y_{hip} 等变量之间存在某种动态关系:

$$\theta_{trunk}^{svm} = f(y_{ankle}, y_{hip}). \quad (1)$$

符合ZMP稳定性判据的理想步态样本集可表示为

$$\{(y_{ankle}^{(1)}, y_{hip}^{(1)}, \theta_{trunk}^{svm(1)}), \dots, (y_{ankle}^{(N)}, y_{hip}^{(N)}, \theta_{trunk}^{svm(N)})\}, \quad (2)$$

其中: SVM学习机的输入变量是 y_{ankle} 和 y_{hip} , 输出变量是 θ_{trunk}^{svm} , N 为样本数量. y_{ankle} 和 y_{hip} 分别表示机器人摆动腿踝关节和支撑腿髋关节轨迹的当前值, θ_{trunk}^{svm} 为能够保持机器人ZMP稳定的上体轨迹.

需要说明的是, 仿真获取符合ZMP判据的样本数据并不难, 然而对于非结构化环境中行走的实际机器人, 已获取的样本总有可能不适用于新的环境条件和机器人状态. 相比于环境的不确定而言, 仿真取得的理想步态样本集是典型的小样本集, 非确定环境下的机器人步态学习控制问题是典型的小样本学习问题.

2.3 基于步态控制的SVM模型构建(SVM modeling based on gait control)

在有限的理想步态样本条件下, 所采用的学习方法不仅要逼近已有的步态数据, 还要具备在恶劣地形下推广所学步态的能力. 在SVM学习机中, 步态的逼近和自适应性能很大程度上取决于核函数及其参数选择, 这就是基于步态控制的SVM模型构建问题.

2.3.1 基于步态控制的核函数构造(Kernel selection based on gait control)

核函数选择是SVM应用的关键^[17], 已有的核函数的构造理论包括黎曼几何分析和Boosting法等. 实际应用中, 往往结合问题的先验知识或领域知识来选择核函数. 针对两足机器人的步态控制问题, 在有限步态样本条件下, SVM不仅要较好地逼近已有的步态数据, 而且要具备不确定环境中的步态预测能力. 而核函数主要分为两大类: 全局核函数和局部核函数. 全局核函数具有全局特性, 泛化性能强、逼近能力较弱, 其典型代表为多项式核; 而局部核函数具有局部性, 逼近能力强、泛化性能较弱, 其典型代表为RBF核. 混合核函数结合二者的优点, 其支持向量机的回归性能优于由单核函数构成的支持向量机, 能够实现逼近能力和泛化能力的平衡优化, 因此本文选择的混合核函数表达式如下:

$$k = ak_{poly} + (1 - a)k_{rbf}, \quad a \in (0, 1), \quad (3)$$

$$k_{poly}(x_i, x) = [\langle x_i, x \rangle + 1]^d, \quad d = 1, 2, \dots, \quad (4)$$

$$k_{rbf}(x_i, x) = \exp\left(-\frac{(x_i - x)^2}{2\sigma^2}\right), \quad (5)$$

其中 k_{poly} 和 k_{rbf} 分别表示多项式核函数和RBF核函数. 容易证明混合核也是一个Mercer核. 调节混合系数 a , 相当于调节两种核函数在混合核函数中的权重. 文献[19]通过仿真指出, 上述混合核中多项式核的阶次取为2, RBF核的宽度取为0.15, 混合系数 a 取为0.95时可同时获得较好的逼近能力和泛化能力.

2.3.2 基于步态控制的SVM参数设计(Design of SVM parameters based on gait control)

目前, SVM参数确定中应用较多的方法有重采样技术和贝叶斯证据框架等. 对于传统的多项式核函数, 多项式的阶次 d 越大, 回归函数的逼近性能越好; d 越小, 回归函数的泛化性能越好, 然而笔者找不到一个能同时满足这两种要求的 d 值. 对于传统RBF核函数, 较小的宽度 σ 可以保证较好的逼近能力, 宽度 σ 过大时其逼近能力明显下降, 且无论 σ 如何取值, 都不能获得稳定的泛化能力.

对于本文采用的混合核函数 $K = aK_{\text{poly}} + (1-a)K_{\text{rbf}}$, 学习机的性能主要取决于混合系数 a , 而多项式的阶次 d 和 RBF 函数的宽度 σ 可以选取某一范围内固定的参数值^[19], 使得学习机同时表现出良好的逼近能力和泛化能力. 因而, 应用混合核的 SVM 学习机, 参数优化并不会变得复杂, 而逼近能力和泛化能力却同时得到了保证, 这对两足机器人的步态学习与控制是十分有利的.

2.4 步态回归(Gait regression)

2.4.1 步态的回归函数(Function of gait regression)

地面环境和机器人状态的变化会直接导致机器人腿部轨迹偏离规划值. 为避免机器人因上体位置与下肢姿态不协调而摔倒, 可在相应的步态回归函数表达式基础上, 依腿部轨迹的实际位置调整上体轨迹, 使机器人满足 ZMP 稳定判据并维持稳定步行. 采用仿真手段获取符合 ZMP 稳定性判据的理想步态样本集

$\{(y_{\text{ankle}}^{(1)}, y_{\text{hip}}^{(1)}, \theta_{\text{trunk}}^{\text{svm}(1)}), \dots, (y_{\text{ankle}}^{(N)}, y_{\text{hip}}^{(N)}, \theta_{\text{trunk}}^{\text{svm}(N)})\}$, 其中: SVM 学习机的输入变量是 y_{ankle} 和 y_{hip} , 输出变量是 $\theta_{\text{trunk}}^{\text{svm}}$, N 为样本数量. y_{ankle} 和 y_{hip} 分别表示机器人摆动腿踝关节和支撑腿髋关节轨迹的当前值, $\theta_{\text{trunk}}^{\text{svm}}$ 为能够保持机器人 ZMP 稳定的上体轨迹.

SVM 步态回归的目标函数是

$$\begin{aligned} \min_{w, b, \xi, \xi^*} & \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^*, \quad (6) \\ \text{s.t.} & \begin{cases} w^T \phi[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T] + b - \theta_{\text{trunk}}^{\text{svm}(i)} \leq \varepsilon + \xi_i, \\ \theta_{\text{trunk}}^{\text{svm}(i)} - w^T \phi[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T] - b \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, N, \end{cases} \quad (7) \end{aligned}$$

其中: 第 i 对样本 $\{x_i, y_i\}$ 的输入和输出分别是 $x_i = (y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T$, $y_i = \theta_{\text{trunk}}^{\text{svm}(i)}$. 方程(6)的求解可通过对偶形式的 Lagrange 多项式, 将问题转化为

$$\begin{aligned} \min_{\alpha, \alpha^*} & \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \\ & \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N \theta_{\text{trunk}}^{\text{svm}(i)} (\alpha_i - \alpha_i^*), \quad (8) \\ \text{s.t.} & \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, N. \quad (9) \end{aligned}$$

其中:

$$Q_{ij} = k[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T (y_{\text{ankle}}^{(j)}, y_{\text{hip}}^{(j)})^T] \equiv \phi(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T \cdot \phi[(y_{\text{ankle}}^{(j)}, y_{\text{hip}}^{(j)})^T]$$

为满足 Mercer 条件的核函数, 本文取 $K = aK_{\text{poly}} + (1-a)K_{\text{rbf}}$, $a \in (0, 1)$. α_i 和 α_i^* 为 Lagrange 乘子. 二

次规划式(8)的解中, 由于 SVM 的稀疏性, 只有少量步态样本的系数 (α_i^*, α_i) 不为零, 系数 (α_i^*, α_i) 不为零所对应的步态样本 $(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)}, \theta_{\text{trunk}}^{\text{svm}(i)})$ 即为支持向量. 设支持向量个数为 g , 则机器人的步态回归函数为

$$f(x) = \sum_{i=1}^g (\alpha_i^* - \alpha_i) k[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T (y_{\text{ankle}}, y_{\text{hip}})^T] + b, \quad (10)$$

$$\begin{aligned} b = & \frac{1}{2} \{ \min \{ \theta_{\text{trunk}}^{\text{svm}(i)} - \\ & \sum_{i=1}^g (\alpha_i^* - \alpha_i) k[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T (y_{\text{ankle}}, y_{\text{hip}})^T] \} \cdot \\ & \max \{ \theta_{\text{trunk}}^{\text{svm}(i)} - \sum_{i=1}^g (\alpha_i^* - \alpha_i) \cdot \\ & k[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T (y_{\text{ankle}}, y_{\text{hip}})^T] \} \}. \quad (11) \end{aligned}$$

以上通过 SVM 回归的方法获得了步态回归函数表达式(10). 其中: $y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)}, \theta_{\text{trunk}}^{\text{svm}(i)}$ 为已知的步态样本, α_i^* 和 α_i 为需要优化学习的 Lagrange 乘子. 优化学习的算法将在下一节具体描述.

2.4.2 步态的学习算法(Algorithm of gait learning)

算法的目标是求解式(8)中的 Lagrange 乘子, 将其代入步态回归式(10), 从而得到稳定条件下机器人上体轨迹与腿部轨迹之间的非线性运动动态. 这里采用离线的步态学习方式. 由 $\beta_i = \alpha_i^* - \alpha_i$ 且 $\alpha_i^* \alpha_i = 0$, 二次规划式(8)等价于

$$\begin{aligned} \max & \sum_{i=1}^n \theta_{\text{trunk}}^{\text{svm}(i)} \beta_i - \varepsilon \sum_{i=1}^n |\beta_i| - \\ & \frac{1}{2} \sum_{i,j=1}^n \beta_i \beta_j k[(y_{\text{ankle}}^{(i)}, y_{\text{hip}}^{(i)})^T (y_{\text{ankle}}^{(j)}, y_{\text{hip}}^{(j)})^T], \quad (12) \\ \text{s.t.} & \sum_{i=1}^n \beta_i = 0, -C \leq \beta_i \leq C, i = 1, 2, \dots, n. \quad (13) \end{aligned}$$

采用序贯最小优化算法(SMO)求解方程(12). 将问题分解成一系列最小子集的二次优化, 每次迭代只优化其中两个乘子 β_1 和 β_2 . 当所有的乘子 β_i 均满足 KKT 条件时, 便得到了二次规划式(8)的解, 将其代入式(10), 就得到了有限步态样本下机器人上体轨迹与腿部轨迹之间的非线性模型.

3 步行稳定性分析(Stability analysis for the biped walking)

3.1 基于 ZMP 判据的稳定域(Stable region according to the ZMP criterion)

两足机器人步行稳定性通常用 ZMP 点在双脚

支撑多边形中的位置来衡量. 如果ZMP点在支撑多边形内, 则步行稳定; 如果ZMP点在支撑多边形外, 则步行不稳定. 不失一般性的, 本文将主要讨论单脚支撑期的稳定性问题.

在单脚支撑期, 机器人前进方向的稳定域可以描述为

$$y_{zmp} \in (y_{heel}, y_{toe}), \quad (14)$$

$$y_{zmp} = \frac{\sum_{i=1}^n m_i(\ddot{z}_i - g_z)y_i - \sum_{i=1}^n m_i(\ddot{y}_i - g_y)z_i}{\sum_{i=1}^n m_i(\ddot{z}_i - g_z) - \sum_k F_{z,k}} - \frac{\sum_{i=1}^n J_{x,i}\ddot{\phi}_{x,i}}{\sum_{i=1}^n m_i(\ddot{z}_i - g_z) - \sum_k F_{z,k}} + \frac{\sum_k (z_k F_{y,k} - y_k F_{z,k})}{\sum_{i=1}^n m_i(\ddot{z}_i - g_z) - \sum_k F_{z,k}}, \quad (15)$$

其中: y_{zmp} 是ZMP点在轴方向的坐标值, y_{heel} 和 y_{toe} 分别是支撑脚后跟及脚尖坐标值, n 是两足机器人模型的连杆数量, $J_{x,i}$ 和 $\phi_{x,i}$ 分别表示连杆 i 在 x 轴方向的惯性力矩和倾角. F_x, F_y, F_z 分别为 x, y, z 方向上的惯性力, m_i 为各连杆的质量, $\ddot{x}_i, \ddot{y}_i, \ddot{z}_i$ 分别为各连杆质心在 x, y, z 方向上的加速度. 详细的分析则须考虑控制系统对期望ZMP的跟踪误差, 以及稳定裕度等两项指标.

3.2 基于ZMP的稳定性指标设计(Design of ZMP-based stability indexes)

基于ZMP稳定判据的两足机器人稳定性分析, 包括动态步行的稳定性分析及稳定裕度分析等两个方面. 在一个步行周期内研究如下两个指标:

对期望ZMP的跟踪误差

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [y_{zmp}(i) - y_{zmp.ref}]^2}; \quad (16)$$

ZMP稳定裕度(stability margin)的均值

$$\overline{SM} = \frac{1}{n} \sum_{i=1}^n \begin{cases} [y_{toe} - y_{zmp}(i)], \\ \text{如果 } y_{zmp}(i) > \frac{y_{heel} + y_{toe}}{2}; \\ [y_{zmp}(i) - y_{heel}], \\ \text{如果 } y_{zmp}(i) \leq \frac{y_{heel} + y_{toe}}{2}, \end{cases} \quad (17)$$

式中: $y_{zmp}(i)$ 是 y 轴方向上ZMP坐标的第 i 个样本, n 是一个步行周期内所取的样本数. $y_{zmp.ref}$ 是ZMP的期望位置, y_{heel} 和 y_{toe} 分别是支撑脚后跟及脚尖的位置. 这里, 跟踪误差RMSE取值越小则步行越稳定, 稳定裕度均值SM越大则步行越稳定.

4 仿真研究(Simulations research)

4.1 仿真准备(Preparation for the simulations)

为了获取满足这种动态关系的步态样本, 依据人类的步行模式规划了两种典型步态供SVM进行学习. 规划的两种典型步态中, 步态1的单脚支撑期占整个步行周期的100%; 步态2的单脚支撑期占整个步行周期的80%. 令机器人的步行周期为 T_c , 取 $T_c = 1$ s. 上体始终保持与地面垂直.

对两种样本步态分别应用3组不同的步态参数. 其中: 第1组参数为步长0.5 m, 步高0.1 m; 第2组参数为步长0.6 m, 步高0.1 m; 第3组参数为步长0.7 m, 步高0.1 m. SVM的参数取值为: 惩罚系数 $C = 1000$, 不敏感损失参数 $\varepsilon = 0.001$. 多项式核的阶次 $d = \{1, 2, 3, 4, 10, 15, 20, 25\}$, RBF核的宽度 $\sigma = \{0.01, 0.05, 0.1, 0.15, 0.2, 0.25\}$, 混合系数 $a = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99\}$. 在步态回归环节, 取采样周期 $T = 0.025$ s, 理想步态样本集为 $\{(y_{ankle}^{(1)}, y_{hip}^{(1)}, \theta_{trunk}^{svm(1)}), \dots, (y_{ankle}^{(41)}, y_{hip}^{(41)}, \theta_{trunk}^{svm(41)})\}$. 所有的仿真实验均选取前11组数据作为测试样本, 后30组数据作为训练样本. 建立了3层神经网络模型用以学习步态, 神经网络的训练和测试均方误差(MSE)目标均为0, 隐层传递函数为Sigmoid, 输出层传递函数为线性函数. 在步态控制环节, 为模拟环境影响或机器人状态变化导致的腿部轨迹突变, 不失一般性, 选取第16个采样点的髋关节位置偏离规划值, 具体数据见表1.

表1 第16个采样点髋关节的位置偏离表
Table 1 Hip deviation of the biped robot at the 16th sampling point

步长/m	规划髋坐标/m	突变髋坐标/m	变化量与步长的比值/%
0.5	0.213764	0.236005	4.45
0.6	0.181256	0.204477	3.87
0.7	0.158059	0.171805	3.29

4.2 基于SVM步态控制的有效性分析(Effectiveness analysis of the proposed SVM-based controller)

为证明本文所提方法的有效性, 实施了两个方面的性能比较. 对步态回归, 将基于混合核、基于多项式核、基于RBF核的3种SVM的回归误差相比较, 又将SVM与神经网络两者的回归误差相比较. 对步态控制, 将采用基于SVM步态控制方法与不采用此方法的两种控制效果相比较.

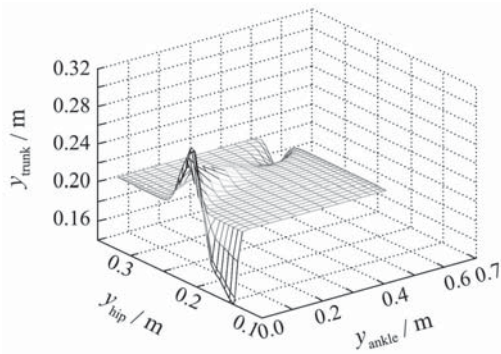
以步态2, 步长0.5 m的情况为例, 采用不同核函

数得到的步态回归函数见图2, 回归误差见表2. 其中, 图2(a)、图2(b)和图2(c)分别为基于RBF核、多项式核和混合核的SVM步态回归函数. 用以生成图2的RBF核宽度、多项式阶次以及混合核函数的混合系数均采用测试MSE性能最佳的参数. 由图2可以看出, RBF核和多项式核分别具有显著的局部特性和全局特性, 而混合核结合了两者的优点, 对于训练过和没有训练过的腿部轨迹都有较好的上体轨迹响应. 不同步态参数条件下的仿真图形相似, 限于篇幅未一一列出.

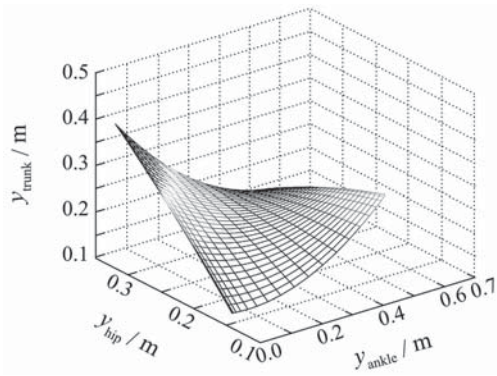
表2 步态回归函数MSE(步态2, 步长0.5 m)

Table 2 MSE of the regression(gait 2 with step length 0.5 m)

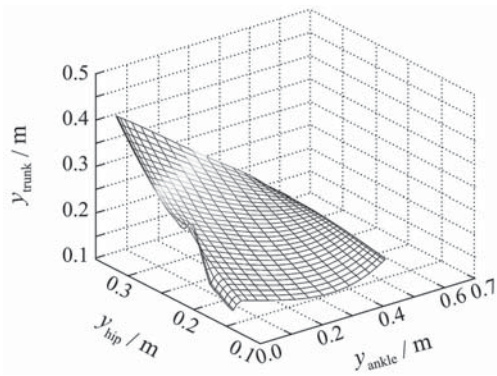
核函数	训练MSE(10^{-6})	测试MSE(10^{-3})
多项式核, $q = 2$	20.702	1.8088
多项式核, $q = 25$	3.9259	14.543
RBF核, $\sigma = 0.1$	3.0827	1.8055
RBF核, $\sigma = 0.01$	0.91712	20.334
混合核, $q = 20,$ $\sigma = 0.05, \alpha = 0.95$	—	—
混合核, $q = 20,$ $\sigma = 0.05, \alpha = 0.3$	2.3659	3.4709
神经网络	1.95179	5.6215



(a) RBF核



(b) 多项式核



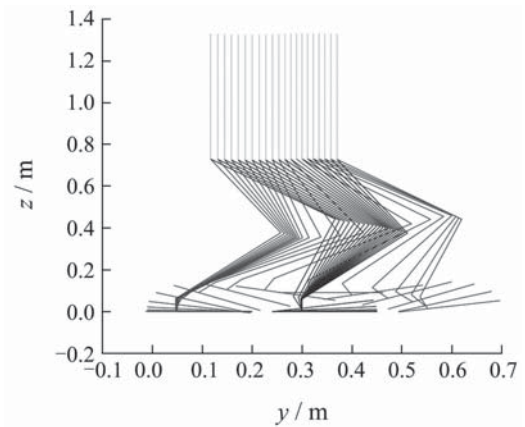
(c) 混合核

图2 采用不同核函数所得的步态回归函数(步态2, 步长0.5 m)

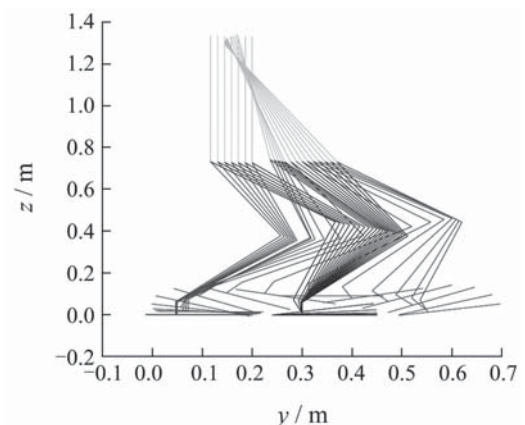
Fig. 2 Regression of biped gait using different kernel functions (gait 2 with step length 0.5 m)

基于混合核SVM的步态控制效果见图3(c). 与之相比较的, 图3(a)为规划步态; 图3(b)中, 髋关节轨迹在第16个采样点偏离规划值, 仅采用图1所示的内环关节控制而未采用外环SVM控制方法, 机器人步态逐渐失控; 图3(c)中, 髋关节轨迹仍然在第16个采样点发生同样大小的偏离,

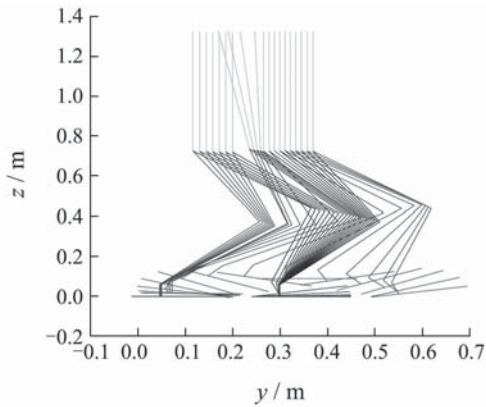
采用外环SVM控制方法调整上体轨迹, 机器人步态逐渐稳定.



(a) 规划步态



(b) 髋关节偏离, 不采用基于SVM的步态控制方法



(c) 髌关节偏离, 采用基于SVM的步态控制方法

图 3 步态效果图(步态2, 步长0.5 m)

Fig. 3 Walking simulation(gait 2 with step length 0.5 m)

4.3 基于ZMP的稳定性指标分析(Analysis of ZMP-based stability indexes)

在一个步行周期内进行周期为 T 的采样, 取 $T = 0.025\text{ s}$, 采用Runge-Kutta方法对式(15)求解, 并根据式(16)和式(17)求出控制系统对期望ZMP的跟踪误差、以及稳定裕度等两项指标. 在采用基于SVM步态控制方法与不采用此方法的两种情况下, 基于ZMP稳定性的两个指标值见表3. 表中: 内环为关节PID控制, 外环为基于SVM的步态控制.

由表3中数据可见, 采用基于SVM的步态控制方法可以显著降低ZMP的跟踪误差, 并且提高步行运动的平均稳定裕度.

表 3 基于ZMP稳定性的指标比较

Table 3 The comparison of ZMP-based performance indexes

	步态	步长/m	RMSE	\overline{SM}
内环+外环	1	0.5	0.0611	0.0840
	1	0.6	0.0538	0.0911
	1	0.7	0.0566	0.0899
	2	0.5	0.0531	0.0951
	2	0.6	0.0550	0.0902
	2	0.7	0.0674	0.0739
内环	1	0.5	0.4190	0.0050
	1	0.6	0.4005	0.0046
	1	0.7	0.4012	0.0104
	2	0.5	0.3992	0.0048
	2	0.6	0.4011	0.0045
	2	0.7	0.4308	0.0061

5 结论(Conclusion)

SVM理论很好地解决了机器学习方法研究中遇到的一些瓶颈难题, 如神经网络结构的确定、过

学习与欠学习、维数灾难、局部极小点等问题, 已成为继神经网络研究之后机器学习领域新的研究热点. 针对两足机器人面临的非确定环境适应性难题, 本文提出了一种基于SVM的两足机器人步态控制方法, 为小样本条件下的学习控制问题提供了新思路. 提出了一种基于混合核的步态回归方法, 仿真研究表明了这种方法比较于全局核和局部核分别单独用于步态学习的优越性. 后继研究还可以考虑使用其他更适合两足机器人步态学习的核函数, 以提高两足机器人步行的自适应性.

参考文献(References):

- [1] RATLIFF N D, SILVER D, BAGNELL J A. Learning to search: functional gradient techniques for imitation learning[J]. *Autonomous Robots*, 2009, 27(1): 25 – 53.
- [2] BYUNGCHAN K, JOOYOUNG P, SHINSUK P, et al. Impedance learning for robotic contact tasks using natural actor-critic algorithm[J]. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 2010,40(2): 433 – 443.
- [3] NGUYEN-TUONG D, SEEGER M, PETERS J. Model learning with local gaussian process regression[J]. *Advanced Robotics*, 2009, 23(15): 2015 – 2034.
- [4] KIM D W, SEO S J, SILVA C, et al. Use of support vector regression in stable trajectory generation for walking humanoid robots[J]. *Etri Journal*, 2009, 31(5): 565 – 575.
- [5] 段萍, 张建畅, 丁承君, 等. 基于模糊遗传算法的移动机器人墙跟踪控制策略[J]. *控制理论与应用*, 2006, 23(3): 416 – 420. (DUAN Ping, ZHANG Jianchang, DING Chengjun, et al. The fuzzy genetic algorithm for the mobile robot's wall tracking control[J]. *Control Theory & Applications*, 2006, 23(3): 416 – 420.)
- [6] FERREIRA J P, CRISOSTOMO M M, COIMBRA A P. SVR versus neural-fuzzy network controllers for the sagittal balance of a biped robot[J]. *IEEE Transactions on Neural Networks*, 2009, 20(12): 1885 – 1897.
- [7] LIU Z, LI C W. Fuzzy neural network quadratic stabilization output feedback control for biped robots via approach[J]. *IEEE Transactions on Systems Man and Cybernetics, Part B*, 2003, 33(1): 67 – 84.
- [8] MANOONPONG P, WORGOTTER F. Efference copies in neural control of dynamic biped walking[J]. *Robotics and Autonomous Systems*, 2009, 57(11): 1140 – 1153.
- [9] WIKLENDT L, CHALUP S K, SERON M M. Simulated 3d biped walking with an evolution-strategy tuned spiking neural network[J]. *Neural Network World*, 2009, 19(2): 235 – 246.
- [10] VARSHAVSKAYA P, KAEHLING L P, RUS D. Automated design of adaptive controllers for modular robots using reinforcement learning[J]. *International Journal of Robotics Research*, 2008, 27(3/4): 505 – 526.
- [11] NAKAMURA Y, MORI T, SATO M A, et al. Reinforcement learning for a biped robot based on a CPG-actor-critic method[J]. *Neural Networks*, 2007, 20(6): 723 – 735.
- [12] HITOMI K, SHIBATA T, NAKAMURA Y, et al. Reinforcement learning for quasi-passive dynamic walking of an unstable biped robot[J]. *Robotics and Autonomous Systems*, 2006, 54(12): 982 – 988.
- [13] DUAN Y, CUI B X, XU X H. Evolutionary reinforcement learning and its application in robot path tracking[J]. *Control and Decision*, 2009, 24(4): 532 – 536, 541.