

基于 PSO 算法的 RoboCup2D 机器人研究

包胜刚, 董春晨, 刘 钊

(武汉科技大学 计算机科学与技术学院, 武汉 430081)

摘要: 对机器人体系结构、动作学习及行为的组织方式进行了研究, 以演化计算为基本方法, 以 RoboCup2D 为平台, 设计了基于 PSO 算法的足球机器人的体系结构, 解决感知、动作、和规划问题; 在训练环境下, 形成感知规则, 优化感知相关参数, 得到对信息高效快速的感知方法, 并根据指定的粒度、功能、参数, 对 RoboCup2D 机器人的原子动作进行了组合优化, 得到一组带参数和执行效果描述的粒子动作; 最后在赛场环境和任务驱动下, 搜索粒子动作并进行组织规划, 得到完成特定任务的机器人行为; RoboCup2D 仿真实验表明, 演化计算方法不仅能利用原子动作进行组合优化, 得到适应于不同条件的粒子动作, 而且能通过其在搜索粒子动作, 动态组成机器人行为; 基于演化计算的足球机器人能更好地完成跑位、截球、带球、传球等任务, 具有更强的适应性。

关键词: 智能体; 机器人体系结构; 规划; 粒子群优化算法

Research on RoboCup2D Robot Planning Based on PSO

Bao Shengang, Dong Chunchen, Liu Zhao

(College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430081, China)

Abstract: Research on the robot system structure, the organization of action learning and behavior way, evolutionary computation as the basic method, RoboCup2D as platform, designs the architecture of soccer robot based on PSO algorithm, solving the problem of perception, action, and planning. By offline training, agents format perception rules and relevant parameters, to optimize perception method for the information, and according to the granularity, functions, and parameters manually specified, PSO builds a set of combo actions, which described by atomic actions, parameters and execution results. According to game environment and a few task rules, PSO searches for task, behavior, and combo actions, as a whole, to accomplish the game tasks. The simulation experiments on RoboCup2D platform show that, agent based on PSO is a robust and flexible robot control method: given evaluation methods and implementation frames, it is able to learn rapidly in real environment, and displays planning behavior without the use of classical planning techniques.

Keywords: agent; robot architecture; planning; PSO

0 引言

机器人行为的组织在智能机器人中占有重要地位, 而它与动作设计具有密切的关系。根据机器人的感知、规划、行为的方式和关系, 机器人体系结构分为: 分层、包容、分层-反应混合 3 大类结构。早在 1984 年 Nilsson 就在分层范式里提出“三层塔式结构”并在 Shakey 上实现, 将机器人动作分级为引导、导航、任务; 之后类似方法反复被使用^[2-3]; 2001 年 Nilsson 提出三层塔式体系结构, 并将 Teleo-Reactive (TR) 规划^[4]引入, 该结构中机器人的组成包括: 感知塔 (Perception Tower)、建模塔 (Model Tower) 和行为塔 (Action Tower), 每个塔都包含层次, 主要特点包括: 1) 将 TeleoReactive 规划应用于行为塔。TeleoReactive 规划融合了层次控制结构的自顶向下方法和基于主体的自底向上方法, 使用了 TR 树的数据结构, 在 TeleoReactive 循环中, 从树根任务条件开始, 依次评估每个任务条件, 直到找到第一个成立的任务节点, 然后就立即执行与之关联的动作或者序列。2) 将感知规则应用于感知塔。3) 将真知保留系统 (Truth Maintenance System, TMS) 应用于建模塔中, 负责获取可靠的知识。

收稿日期:2016-03-31; 修回日期:2016-05-10。

基金项目:国家自然科学基金资助项目(51174151)。

作者简介:包胜刚(1991-),男,湖北黄冈人,硕士研究生,主要从事人工智能和机器学习方向的研究。

刘 钊(1969-),男,教授,主要从事智能计算、人工智能和计算机视觉方向的研究。

Nilsson 的三层塔式体系结构如图 1。

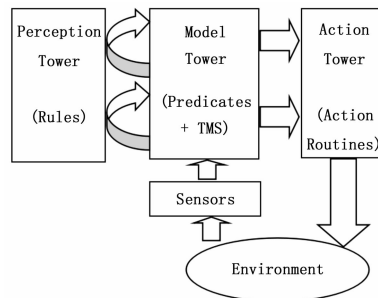


图 1 Nilsson 的三层塔式体系结构

该结构被成功应用于许多自动化领域, 包括机器人规划和航空控制 (Benson, 1996). 结构中的规则和动作设计是其关键点和难点, 包括 Nilsson 在内的许多研究者考虑过自动化这部分工作。2003 年 Kochenderfer 将遗传算法应用于上述体系结构, 利用演化计算的方法, 自动生成 Teleo-Reactive 程序中的规则和行为, 并可以用来求解积木世界中的不同问题规模的动作规划问题。遗传算法的使用提高了系统的自适应性, 这是获取动作的相关知识的新途径。

积木世界里的堆放问题状态空间不大, 对系统的智能性和适应性要求并不高, 相比而言, 机器人踢足球, 以及机器人足球比赛对智能机器人的挑战性更高, 更适用于研究和验证基于演化计算方法的通用性和灵活性。

在 RoboCup 挑战计划 [Noda et al. 1997] 中就指出, RoboCup 重点研究: ①在多智能体合作及对抗环境中的机器

学习, ②多主体体系结构, ③实时的多智能体规划和规划执行, 以及④对手建模。RoboCup2D 仿真平台利用网络互联的计算机来模拟人类足球比赛, 是一个动态、实时、不确定的环境, 比赛双方各自利用 11 个独立的程序来控制本方的 11 个具有指定质量和体力的球员(足球机器人)进行踢球比赛, 最终目的是按照指定规则(和人类足球比赛基本一致)踢球并获取比赛胜利。为了适度简化问题, 平台提供给队员一些最基本的动作, 如踢球、加速、转身等, 程序必需合理地组合这些动作, 形成更“智能”的行为。比赛上下半场各 5 分钟, 每个模拟周期只有 100 ms, 要求程序的实时性非常高。其中的足球机器人设计问题具有下面的特点: 1) 任务复杂: 足球机器人的任务包括截球、阻挡、带球、过人、射门、传球等, 不仅种类多, 而且彼此之间有密切联系, 要求机器人能够有效选择任务并且实现任务之间平滑的过度。2) 动作空间大: 足球机器人的动作力度和方向在实数空间取值, 搜索空间巨大。3) 执行效率高: 竞技对抗中的足球机器人, 要求其动作尽可能简洁、能耗少、行动敏捷。4) 多机器人合作: 踢球任务是典型的合作性任务, 要求多机器人相互配合完成。

本文将演化计算方法里的 PSO 算法应用于 RoboCup2D 机器人动作知识获取, 以及比赛过程中任务选择和动作规划问题, 将粒子群优化算法进行了适当的改进, 并将该算法用于足球机器人的动作选择。设计了一个足球机器人的动作集合; 然后, 根据赛场上的实际情况为机器人分配角色与任务; 利用粒子群优化算法为足球机器人进行动作选择。实验表明, 应用新算法的仿真足球机器人动作灵活、准确、效果好。

1 基于演化计算的足球机器人体系结构

在目前 RoboCup2D 足球机器人研究中, 智能机器人的构建方式一般采用任务分层并分别决策或规划的体系结构, 其示图如图 2 所示。

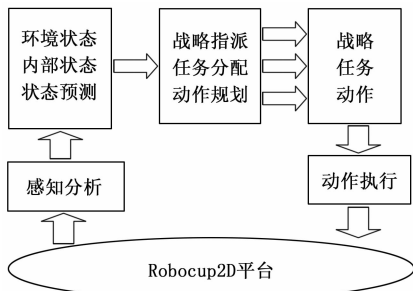


图 2 一种典型的足球机器人的体系结构

其中, 战略层是决策系统模块中的最高层, 用来从全队利益出发实现高层决策, 确定球员智能体的行为目标及总的行为方案; 任务层要制订出为了完成这一特定的目标, 所应该采取的行为策略, 包括抢占位置、抢球, 拦截, 带球, 传球, 射门, 扑球, 盯人等等; 动作规划是决策系统层最低的, 用以实现智能体的个人技能, 即将这些具体的行为决策细化为比赛平台可以接受的执行指令。

类似于 Nilsson 的三层次体系结构, 在这种体系结构中, 基本动作的描述、动作规划是机器人的核心模块。这种结构的主要优点在于:

- 1) 模块性强。各模块分工清晰, 可理解性强, 便于工程化实现。
- 2) 实时性好。计算明确, 响应迅速。

3) 可靠性高。各模块测试方便, 组织的系统可靠性高。

在足球机器人以及其他智能体的设计中我们发现, 采用这种体系结构来建立复杂环境下多级任务驱动的智能体系统, 具有下面缺点:

1) 动作设计困难。由于环境和任务的复杂性, 通用的、有效的动作难以设计, 动作的相关知识, 包括启动条件、执行周期、效果描述等等, 获取困难。动作集合的完备性难以保证。目前许多研究采用加强学习、人工神经网络等方法来建立机器人基本动作, 但是由于评价函数的设计困难, 动作使用过程中的整体效果并不好。

2) 适应性差。特定任务下, 机器人组织动作的方式单一, 对人工经验依赖性强, 可扩展性差。

3) 整体性差。实际上, 各级任务其实具有密切的联系, 忽略这些联系可能影响智能机器人行为的一致性和平滑性。

根据 RoboCup 足球机器人的环境和要求, 本文提出了机器人行为的 4 层(原子动作层、粒子动作层、行为层、任务层)组织方法, 设计了基于演化计算的足球机器人的体系结构。主要特点: 利用在感知分析、动作设计、任务规划部分引入演化计算方法。基于演化计算的足球机器人的体系结构如图 3 所示。

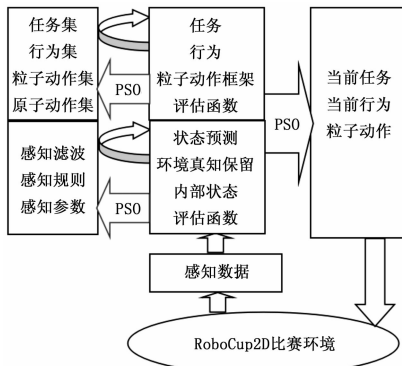


图 3 基于演化计算的足球机器人的体系结构

由于环境和任务的复杂性, 通用的、有效的动作难以设计, 动作的相关知识, 包括启动条件、执行周期、效果描述等等, 获取困难。动作集合的完备性难以保证。目前许多研究采用加强学习、人工神经网络等方法来建立机器人基本动作, 但是由于评价函数的设计困难, 动作使用过程中的整体效果并不好。本文设计了基于演化计算的足球机器人学习结构。该结构使用任务框架、行为框架和动作框架表达各种动作, 利用特定环境里的感知数据, 在一些规则的诱导下, 形成和丰富行为知识。基于演化计算的足球机器人学习结构如图 4 所示。

2 基于 PSO 算法的足球机器人动作设计

2.1 原子动作集合

在 RoboCup2D 环境中, 足球机器人一般拥有以下基本动作: dash、turn、kick、tackle、catch、move、turn_neck、change_view、say、pointto、attentionto 等, 这些动作添加 1 个或者 2 个参数后就直接可以作为动作输出并产生效应。其中, catch 仅仅应用于守门员, 而 move 仅仅应用于赛前移动队员, 但是 turn_neck、change_view、say、pointto、attentionto 等仅仅用于感知, 这些动作规划相对容易、可控性强, 本文将它们排除在原子动作集合之外, 并分别给予考虑和分配。本文设计的原子动作集合仅仅包含 4 个元素, 它们在与 Robo-

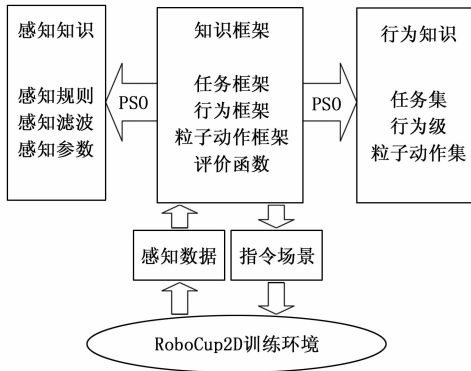


图 4 基于演化计算的足球机器人学习结构

Cup2D 服务器交互过程中, 每个周期里只容许使用一次, 在机器人动作规划过程中频繁地被搜索和执行, 对足球机器人的整体性能产生非常关键的影响。原子动作集合见表 1。

表 1 原子动作集合

Action	Action Parameters	
	1st	2nd
Dash	Power range: (-100,100)	null
Kick	Power range: (0,100)	Direction range: (-100,100)
Tackle	Power range: (-100,100)	null
Turn	Angle range: (-180,180)	null

2.2 粒子动作框架

本文按照足球机器人动作的特点和功能, 利用机器人原子动作, 设计了 7 类机器人粒子动作, 然后利用演化计算方法, 生成各种粒子动作, 组成粒子动作集合, 供比赛过程选择和调用。

1) 大力踢球 (Hard_Kick): 目的是向某确定方向射门或者大力传球; 在持有球的时候使用; 该动作由 turn、kick 和 dash 组成, 持续时间为 1~20 个周期; 优化目标为踢球速度尽可能快且射出球角度与期望射出方向的偏差小。

2) 带球移动 (Dribble_Ball): 目的是向某确定方向带球; 在持有球且无截球威胁时候使用; 该动作由 turn、kick 和 dash 组成, 持续时间为 1~20 个周期。优化目标为运球速度尽可能快。

3) 带球过人 (Dribble_Pass): 目的是绕过对手并向某确定方向带球; 在持有球且有截球威胁时候使用。该动作由 turn、kick 和 dash 组成, 持续时间为 1~20 个周期。优化目标为突破对手防线且丢球率尽可能低。

4) 位置部署 (Location_Arrive): 目的是到达某确定位置; 在没持有球、持有球的机会小时使用。该动作由 turn、dash 组成, 持续时间为 1~20 个周期。优化目标为到达位置的速度大, 而能量消耗速度小。

5) 位置抢占 (Location_Rushing): 目的是先于对手迅速到达某确定位置; 在没持有球、持有球的机会大时候使用。该动作由 turn 和 dash 组成, 持续时间为 1~20 个周期。优化目标为到达位置的速度。

6) 阻挡运行 (Block_Robot): 目的是阻挡对手到达某确定位置; 在没持球、球在附近、持球的机会小时使用。该动

作由 turn 和 dash 组成, 持续时间为 1~20 个周期。优化目标为阻挡对方机器人的成功率。

7) 阻挡铲球 (Block_Tackle): 目的是破坏对手的控球; 在没持有球、球在附近、铲球的机会比较大时候使用。该动作由 turn、kick、tackle 和 dash 组成, 持续时间为 1~20 个周期。优化目标为截球的成功率。

在利用演化计算方法优化粒子动作问题中, 解的结构设计, 也就是粒子动作的表达方法, 非常关键。本文利用一个框架来描述每种粒子动作, 主要包括事先状态、原子动作序列规划、动作目的要求、动作执行评价等方面的信息。用于描述每种粒子动作的框架设计如下。

框架代码: 1-7

事先的状态:

- 球的速度: (dx, dy)
- 对手相对于球的位置: (x, y)
- “我”相对于球的位置: (x, y)
- “我”的体力: stamina

事后期望的状态:

- 球的速度: (dx, dy)
- “我”相对于球的位置: (x, y)
- “我”相对于球的位置: (x, y)
- “我”的体力: stamina

动作执行规划:

- 预计执行周期: t_now
- 原子动作序列: act[t_now]
- 原子动作参数: cl[t_now], c2[t_now]

动作执行结果:

- 效果评价: value

动作修正:

- 下次执行周期: t_next
- 下次原子动作序列: act[t_next]
- 下次原子动作参数: cl[t_next], c2[t_next]

2.3 基于 PSO 算法的粒子动作生成

粒子动作的设计目的是通过训练得到通用的技能, 包括: 大力踢球 (Hard_Kick)、带球移动 (Dribble_Ball)、带球过人 (Dribble_Pass)、位置部署 (Location_Arrive)、位置抢占 (Location_Rushing)、阻挡运行 (Block_Robot)、阻挡铲球 (Block_Tackle) 及它们的启动条件、时间耗费、精力耗费、效果描述等。

基于演化计算的粒子动作设计的基本过程是: 先根据人工经验设计表达上述各类动作的框架, 对于每类动作, 初始化 n 个框架作为演化计算的种群, 然后在指定训练环境中应用这些框架所表达的动作, 利用演化计算对框架进行遗传和变异, 再根据期望效果和动作产生的实际效果进行动作效果评价、筛选, 形成新的种群。在这种迭代中选择出某种动作的最优表达和序列。基于演化计算的粒子动作设计如图 5 所示。

3 仿真实验及分析

为了验证本文所提出的基于 PSO 算法的 RoboCup2D 机器人及其动作学习结构的可行性和有效性, 进行了 RoboCup2D 环境下的仿真实验, 实验设计如下:

1) 指定任务下的动作搜索: 给定了特定任务, RoboCup2D 机器人搜索和选择动作序列, 验证动作搜索的效率和结果。为完成上述目的, 本文设计了专门用于机器人动作学习的智能体程序, 利用历年 RoboCup2D 国际比赛中表现优秀的

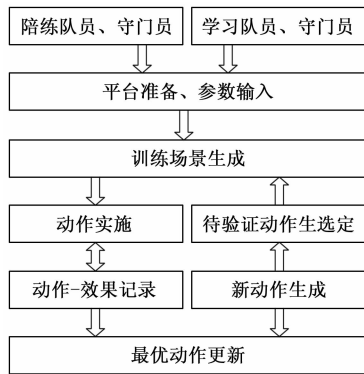


图 5 基于演化计算的粒子动作设计

智能体程序，组成 2 对 2 的对战平台（双方各 1 名守门员和 1 名普通队员），在实际对战中捕捉学习机会并学习 7 类粒子动作，形成踢球过程中通用的动作技能知识系统。

2) 特定环境下的任务规划：给定了特定环境，规定了双方机器人的位置和状态，RoboCup2D 机器人搜索和选择任务，验证任务搜索的效率和结果。

3) 实际比赛：与其他方法设计的机器人进行比赛，评价整体的效率和结果，验证比赛过程中的在线学习效果。

实验结果证明了基于 PSO 算法的 RoboCup2D 机器人的可行性。PSO 算法在指定任务下的动作搜索的效果很好，特定环境下的任务规划能力一般，而且机器人在比赛过程中具有一定的在线学习能力。

4 结束语

本文将演化计算方法里的 PSO 算法应用于 RoboCup2D 机

（上接第 226 页）

2) 锚节点数目比例变化的定位性能比较：

假设总节点的个数不变，伴随着锚节点个数的比率不断变化，本文算法，文献 [5] 算法和文献 [6] 算法的定位误差变化曲线如图 7 所示。从图中发现，3 种算法在锚数目比较少的时候相差不是很大，这说明 3 种算法都能适合锚节点数目少的时候，但伴随着锚节点的数目增多，3 种算法的误差逐渐拉大，文献 [5] 和文献 [6] 的算法图像波动比较大，而本文算法的图像平缓，这说明本文算法收到锚节点数量的影响不大。

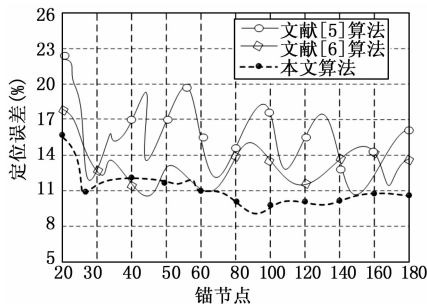


图 7 不同锚节点个数的 3 种算法定位性能

5 结束语

针对无线传感定位中的质心定位算法存在的不足，首先采用了序列加权的概念对未知节点的定位方法进行了改进，其次，引入 LSSVR 概念对改进后的节点定位的精度进一步改

进。仿真实验说明本文算法与基本的 LSSVR 算法相比误差明显减低，通过与其他算法在锚节点，未知节点比例方面的定位误差相比有了显著的降低。

参考文献：

[1] Nilsson N. Teleo-Reactive Programs for AgentControl [J]. Journal of Artificial Intelligence Research, 1994.

[2] Nils J. Nilsson, Teleo-Reactive Programs and the Triple-Tower Architecture [J]. Electronic Transactions on Artificial Intelligence, 2001, 5, Section B, 99-110.

[3] Rajan K, Py F, McGann C. Adaptive control of AUVs using on-board planning and execution [N]. Sea Technology Magazine, April, 2010.

[4] Russell S E, Carr D, Dragone M, et al. From bogtrotting to herding: a UCD perspective [J]. Annals of Mathematics and Artificial Intelligence 61: 349-368, 2011.

[5] Sánchez P, Alonso D, Morales J M, et al. From Teleo-Reactive specifications to architectural components: A model-driven approach [J]. Journal of Systems and Software, 85, 2012.

[6] Smith G, Sanders J W, Winter K. Designing adaptive systems using teleo-reactive agents [J]. Transactions on Computational Collective Intelligence. Springer-Verlag, 2014 (to appear).

[7] Soto F, Sánchez P, Mateo A, et al. An Educational Tool for Implementing Reactive Systems Following a Goal-Driven Approach [J]. Computer Applications for Engineering Education, DOI: 10.1002/cae. 21568, 2012.

进。仿真实验说明本文算法与基本的 LSSVR 算法相比误差明显减低，通过与其他算法在锚节点，未知节点比例方面的定位误差相比有了显著的降低。

参考文献：

[1] 何艳丽. 无线传感器网络质心定位算法研究 [J]. 计算机仿真, 2011, 28 (5): 163-166.

[2] 杨新宇, 孔庆茹, 戴湘军. 一种改进的加权质心定位算法 [J]. 西安交通大学学报, 2010, 44 (8): 1-3.

[3] 赵栋栋, 赵菊敏, 李灯熬. 基于质心迭代的 DV-Hop 定位改进算法 [J]. 计算机测量与控制, 2013, 21 (10): 2764-2766.

[4] 吕振, 赵鹏飞. 一种改进的无线传感器网络加权质心定位算法 [J]. 计算机测量与控制, 2013, 21 (4): 1102-1104.

[5] 刘玉军, 蔡猛, 高立恒. 基于 RSSI 测距的传感器节点质心定位修正算法 [J]. 计算机测量与控制, 2014, 22 (9): 2860-2862.

[6] 施伟, 高军. 无线传感器网络中基于 RSSI 的改进加权质心定位算法 [J]. 计算机应用与软件, 2015, 32 (12): 68-70.

[7] 王缓缓, 邱建文. 基于区域分割的二次质心定位算法 [J]. 2015, 32 (4): 279-281.

[8] 王振朝, 张琦, 张峰. 基于 RSSI 测距的改进加权质心定位算法 [J]. 电测与仪表, 2014, 51 (21): 63-66.

[9] 李兆斌, 魏占祯, 徐凤麟, 等. 无线传感器网络增强的质心定位算法及性能分析 [J]. 传感技术学报, 2009, 22 (4): 1247-1250.

[10] Bulusu B, Heidemann J, Estrin D. Denstiny adaptive algorithms for beacon placement in wireless sensor networks [A]. In: Proceedings of IEEE ICDES [C]. Mesa, AZ, USA, IEEE Press, 2001, 489-498.