

文章编号:1006-2467(2006)07-1069-05

基于短期预测的追捕方法

罗 真, 曹其新

(上海交通大学 机器人研究所, 上海 200030)

摘 要: 针对具有安全区的追捕逃避问题, 抛开参与人是否理性的假设, 提出了一种基于短期预测的追捕方法. 由参与人短期内状态空间有限和伪蒙特卡罗方法原理阐述了进行短期预测的可行性. 追捕方法的流程为: 追捕方先根据物理条件限制预测双方的短期状态; 根据反映对抗态势的特征计算出对手处于某状态时自身某状态的收益, 并由此进一步获得针对同时刻对手各个状态的自身某状态的综合收益; 估算出自身各个预测状态的综合收益后, 选取某个状态作为短期目标, 执行能到达该状态的行为. 结合足球机器人一对一对抗的实践进行了仿真, 结果表明该方法有效. 介绍了在自主式足球机器人上应用基于短期预测的追捕方法的初步实验, 证实该方法具有一定的实际应用价值.

关键词: 追捕-逃避; 短期预测; 收益函数; 伪蒙特卡罗

中图分类号: TP 242.6 **文献标识码:** A

A Pursuing Method Based on Short-Term Prediction

LUO Zhen, CAO Qi-xin

(Research Inst. of Robotics, Shanghai Jiaotong Univ., Shanghai 200030, China)

Abstract: For the pursuit-evasion games with safety zone, this paper put away the assumption of whether game players are rational or not, and proposed a short-term prediction based pursuing method. With the fact that players' possible states are limited in short-term due to physical limits and the principle of quasi-Monte Carlo, the feasibility of short-term prediction was analyzed. The decision-making procedure is: According to physical limits, at first the pursuer predicts its own and evader's possible short-term future states; then based on features reflecting game situation, the pursuer evaluates the single cost of certain state VS certain evader's state; after that, the pursuer computes the aggregate cost of certain state VS all evader's states at the same time; after all aggregate costs of all states have been gained, the pursuer selects a proper state as its short-term goal and executes the action leading to the goal. The simulation result illustrates that the short-term prediction based pursuing method is effective. It also presented an experiment with middle size soccer robots. The experiment reveals that the method can be applied in real robots effectively.

Key words: pursuit-evasion; short-term prediction; cost function; quasi-Monte Carlo

收稿日期: 2005-06-25

基金项目: 国家自然科学基金资助项目 (60443007)

作者简介: 罗 真 (1974-), 男, 浙江金华人, 博士生, 主要从事对抗性环境中多智能机器人协作研究. 曹其新 (联系人), 男, 教授, 博士生导师, 电话 (Tel.): 021-62932750; E-mail: qxcao@sjtu.edu.cn.

追捕逃避问题是比较热门的研究议题,如多机器人围捕问题^[1]、视线发现问题^[2]等.本文研究具有安全区的追捕逃避问题:逃避者存在一个目标区或者说安全区,追捕者的任务是拦截和阻止逃避者进入该安全区.该类追捕逃避问题的典型实例有导弹拦截问题^[3-5].在巡航导弹拦截问题研究中,往往认为逃避者具有追捕者事先不知道的航迹指令序列.此外,随着军事技术的发展,未来会有越来越多的无人战机走向战场,具有比导弹更高智能的无人战机的拦截围捕将是未来的一大挑战.本文以具有某些与未来无人战机相似特征的自主式足球机器人作为对象,探讨一种基于短期预测的追捕智能体的方法.

在追捕逃避问题中,很多论文假设对抗双方都是理性的^[1,6],参与人“知道对手知道自己所知的”.但在实践中,对抗双方受制于自身的感知能力、智能水平,并非总是“理性”的:参与人会规避一些已知的明显对自己不利的情势,趋向对自己“有利”的情势,但并不总能肯定该“有利”是否是真正的“有利”.由此,本文避开对抗双方是否理性的假设,而从参与人的物理条件限制角度出发,对参与人进行短期未来状态预测,然后根据反映对抗态势的特征因子评估这些预测出的短期未来状态,最后根据评估结果决定自身行为.

1 问题描述

本文所研究的具有安全区的追捕逃避问题满足如下假设:

(1) 已知活动场地 F 中存在一区域 G , 逃避方 E 负责将所携带的物件 B 送入 G ; 追捕方 P 的任务是防止 E 将 B 送入 G ; 若 E 携带 B 时被 P 拦截或 E 将 B 带到活动场地 F 之外, P 拦截成功; 游戏无时间限制.

(2) P 、 E 具有动力, 分别具有运动参数: 最大运行线速率 v_P 和 v_E 、最大角速度 $\omega_{P,\max}$ 和 $\omega_{E,\max}$ 、最大线加速度 $a_{P,\max}$ 和 $a_{E,\max}$; 且这些参数事先互知.

(3) P 、 E 具有一定感知力, 能准确感知到自身位置坐标 (x, y) 、朝向 φ_P 和 φ_E 、自身速度 v_P 和 v_E , 以及对方的当前位置 (x, y) , 但对方的运动方向和速度不能直接准确获得, 需要通过计算对方的位置变化来估计; 此外, P 能准确感知到 E 抛射 B 这一事件.

(4) B 不具备动力、感知力和智能, 只能在惯性和阻力作用下做减速运动, 线加速度为 a_B ; 当其被 E 抛射时, 获得最大初速度 $v_{B,\max}$, 朝向 φ_B 等于 E 抛射 B 时的朝向 φ_E .

(5) P 、 E 都采用数字控制系统, 行为决策周期分别为 T_P 和 T_E .

RoboCup 中型组的实践表明: 因规则禁止安装粘球机构, 机器人运球时, 改变线速率易导致丢球, 匀速率运行有助于机器人保持控球. 故本文假定 E 携带 B 时线速率保持不变, E 的智能主要体现在能根据情势决定以哪种角速度转弯; 为了比较追捕效果, 仿真中 P 在预测自身未来时也假定自身线速率不变, 行为差异通过不同的角速度体现. P 、 E 的行为决策主要体现在下一时段以多大角速度等运行; 此外, E 还决定是否抛射物件 B .

为叙述方便起见, 根据上述假设, 可将 E 和 B 统一为一个新的混合智能体“ E ”: 在 E 携带 B 运行时, “ E ”的行为特点按 E 进行, 当 E 抛射 B 以后, “ E ”的行为特点按 B 进行. 下文的 E 表示该新的混合智能体.

2 短期预测策略的可行性

智能体行为决策的灵活性会使其经长期运行后的可能状态空间变得很大, 预测由此失去实际意义. 但对智能体进行短期预测是可行的: 受制于物理条件限制, 在某一个较短时期内, 智能体可能出现的区域、朝向等都有限; 正是这种“有限”, 使得针对智能体的短期预测具有某种有效性.

定理 假设智能体可以恒定线速率 v 和不大于最大角速度的任意角速度 ω 运行. 如图 1 所示, 以 O 点为运行原点运行 T 时间, 若智能体沿直线运行, 其轨迹为线段 OC , 若以最大角速度运行, 其可能轨迹为弧线 OA 和弧线 OB , 则该智能体运行 T 时间后: 必然出现在阴影部分所示的圆扇区内, 其内环半径为线段 OA 、 OB 的长度, 外环半径为线段 OC 的长度; 以初始方向的角度为 0° , 智能体的运行方向在区间 $[-\angle AOB, \angle AOB]$ 的范围内.

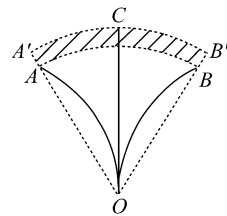


图1 短期运行后机器人的区域范围

Fig. 1 Position range of robot after short-term running

限于篇幅, 定理证明省略. 该定理意味着短期内智能体的分布区域和状态空间有限.

根据蒙特卡罗方法的基本原理, 如果随机地选

取参与人的可能的行为序列,执行这些行为序列后产生的参与人状态的统计学特性反映出参与人经过短期运行之后的状态空间特性.进一步,如果按照均匀分布原理来选择参与人的可能的决策序列,根据“伪蒙特卡罗方法”的原理,也能较好地近似反映出参与人经过短期运行之后的状态空间特性.这意味着参与人虽然不知道对手真正的行为序列,但可通过假设一个均匀分布的行为序列对对手进行预测来有效地近似反映出对手的未来状态空间.

上述理由表明参与人对自身和对手做短期预测是可行的.

3 基于短期预测追捕应对

在每个系统指令周期, P 都根据当前对抗形势对双方进行新的短期预测,并根据这种预测决定当前采用何种应对行为.基于短期预测的追捕决策过程包括预测、评估和决策 3 个阶段:预测是预测短期可能状态;评估是对预测出的状态进行评估;决策是决定自身当前行为.

3.1 预 测

现实中,许多系统在控制上分多个控制档位,即行为空间是离散的.在本文 RoboCup 中型组自主机器人开发实践的一种智能系统实现中,运行参数也被离散化为一个个档位,如 1.0、2 rad/s 等角速度档位,机器人根据情势决定用哪一个档位运行.

基于这种考虑, P 对自身的预测基于自身离散的控制档位:计算出在当前状况下,分别施加各档位的控制序列,经过时间 T 后的位置和状态.

P 对 E 的预测时,则是基于伪蒙特卡罗方法的原理,先在 E 的极限运动参数范围内均匀地选取用于预测的控制档位.预测前 P 要先检测 E 是处于射门状态还是处于带球状态.如果 E 处于射门状态,按匀减速运动进行预测;如果 E 处于带球状态,预测过程为:计算出 E 在当前状况下,分别执行各档位的控制序列,经过时间 T 后的位置和状态.

图 2 示意了 P、E 都只有 n 个控制档位情况下的短期预测结果,分别是个 n 叉树.图中:圆方框表的存储某个预测状态的节点; $1 \sim K_0$ 表示是某步预测.其中“步”表示连续执行某控制档位的时间单元.

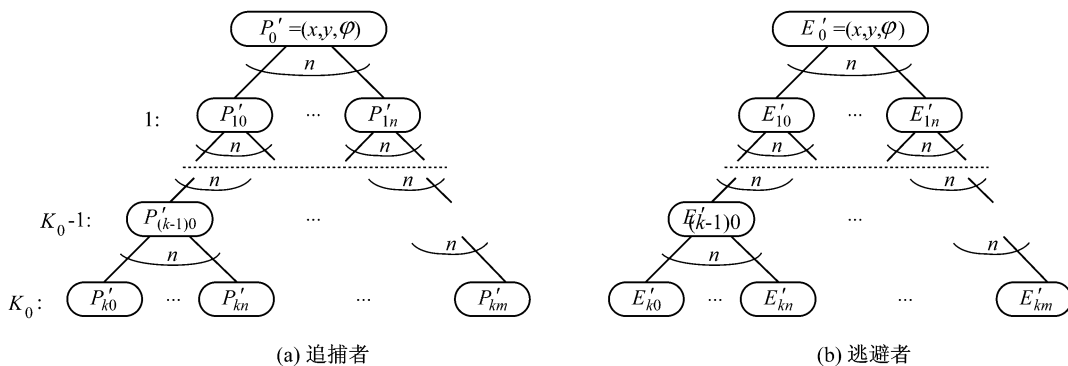


图 2 预测结果状态树图

Fig. 2 The predicted states tree

3.2 评估和决策

追捕逃避对抗中,一方的获利意味着另一方的失利,是零和的,故 P 只需针对预测出的状态估算自身收益. P 估算收益的过程分两步:第 1 步先估算某步某个 P 状态针对该步某个 E 状态的个别收益;第 2 步算出针对该步所有 E 状态的该 P 状态的综合收益.

在第 1 步个别收益估算中,针对某个 E 状态 j 的某个 P 状态 i 的收益 f_{ij} 估算过程为:先分解出反映对抗态势的各特征因子值,根据事先给出的与这些特征因子关联的收益评估函数分别计算出各特征因子项的收益,然后综合各个特征因子的收益获得

针对 E 状态 j 的状态 i 的个别收益.在足球机器人对抗中可抽取出下列特征因子: E 到球门的距离; E 的朝向和其所处位置的可行射门角度范围之间的偏角; P 离拦截成功点的距离; P 离 E 可能的射门线段的距离; P 与 E 相互接近或远离的速度等等.单个特征因子项的收益函数是通过经验知识和优化调整过程获得的.如, E 到球门的距离越小,对 P 威胁越大,且该距离越大,距离的增加对威胁的降低的影响越不明显,由此特征因子“E 到球门的距离”的收益函数可用指数函数形式:

$$f_d = e^{-k/d} \tag{1}$$

式中: k 为待定的参数; d 为 E 到球门的距离.

假设 n 项特征因子的收益函数分别为 $f_x, f_y, \dots, f_z$, 则个别收益函数既可用线性加权方式构造, 也可用相乘方法, 它们本质上是相同的. 本文仿真中采用的是相乘汇总方法:

$$f_{ij} = f_x \cdot f_y \cdot \dots \cdot f_z \tag{2}$$

获得针对同步的各个 E 预测状态的某 P 预测状态 i 的个别收益后, 可进行收益估算的第 2 步: 统计汇总出该 P 预测状态 i 的综合收益 f_i . f_i 可通过一个多元组来表示, 如四元组 $\bar{a}, n, n_v, n_f$. 其中: $\bar{a}$ 表示针对该步所有 E 预测状态的平均收益; n 表示该步有多少个 E 预测状态; n_v 表示 P 获胜的 E 预测状态数; n_f 表示 P 失败的 E 预测状态数. 由

$$f_i \bar{a} = \sum_j \frac{f_{ij} n}{f_i n} \tag{3}$$

确定出 P 各个预测状态的综合收益, 也就获得了一棵完整的决策树. 目前, 有很多成熟的决策树遍历计算技术可用, P 很容易根据这一决策树确定出当前的合理行为.

4 仿 真

本文以一对一的 RoboCup 中型组机器人足球比赛为对象进行了仿真, 比较了在同样的物理参数、起始条件和逃避策略情况下, 两种追捕策略的拦截效果: 一是沿最短路径直接追球的追捕策略; 二是基于短期预测的追捕策略.

在该仿真实验中, E 的基本策略是: 如果满足射门条件, 则射门; 否则, 尽量往门前带球, 如果正前方一定范围内出现 P, 则带球转弯以避开 P. 该仿真实验中, E 不做预测.

仿真实验中, 场地为 10 m × 5 m, 球门宽 2 m. P 的物理参数为: $v_P = 1.5 \text{ m}$, $\omega_{P,\max} = 2 \text{ rad/s}$. E 的物理参数为: $v_E = 1.5 \text{ m}$, $\omega_{E,\max} = 1 \text{ rad/s}$. 球踢出去的初速度为 5 m/s, 球自由滚动的加速度为 2 m/s². P 的起始位置坐标为 (7.5 m, 2.5 m), 起始方向为 3.142 rad (以朝向门一侧的水平方向为 0), E 的起始位置坐标为 (5 m, 2.5 m), 起始角为 $(-\pi/2, \pi/2)$ 范围内的任意值. E 和 P 的系统指令周期都是 0.1 s.

P 的胜利条件是: 出现在 E 正前方 0.5 m × 0.6 m 的范围内; E 将球带到界外.

在仿真实验, 收益评估中使用了 4 个特征因子, 相应的收益评估函数分别为:

$$f_{d_1} = \begin{cases} e^{-500/d_1} & d_1 \geq 500 \\ d_1 e^{-1/500} & d_1 < 500 \end{cases} \tag{4}$$

$$f_{\alpha} = (2 - e^{-2\alpha\pi})/2 \tag{5}$$

$$f_{d_2} = e^{-d_2 \times 10^{-4}} \tag{6}$$

$$f_{d_3} = e^{-d_3 \times 10^{-4}} \tag{7}$$

$$f_{ij} = f_{d_1} \cdot f_{\alpha} \cdot f_{d_2} \cdot f_{d_3} \tag{8}$$

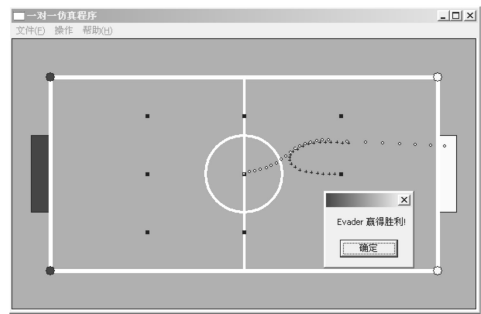
式中: 距离单位为 mm; 角度的单位为 rad; d_1 为 E 到球门的距离; α 为 E 朝向与 E 所处位置的可行射门角度之间的偏角; d_2 为 P 离拦截成功点的距离; d_3 为 P 离 E 可能的射门线段的距离.

仿真实验中, P 只是分别针对各第 1 步预测状态, 计算出与其后续状态的平均综合收益, 选择能达到具有最好平均综合收益的第 1 步预测状态的行为执行. 实验结果为:

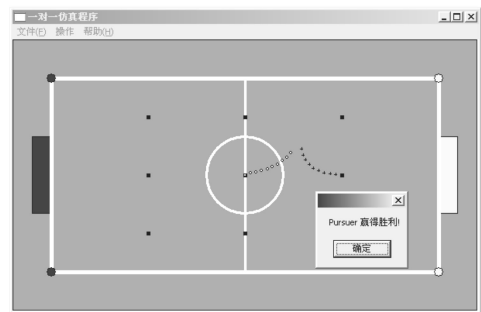
(1) 在 P 采用直接追球策略时, 只有当 E 的起始角落在 $[-0.205, -0.164]$ $[-0.160, -0.149]$ $[-0.039, 0.038]$ $[0.148, 0.158]$ $[0.163, 0.203]$ 内时, P 才能拦截成功;

(2) P 采用基于短期预测策略时, E 以 $(-\pi/2, \pi/2)$ 范围内任意角起动, P 都能取得胜利. 其中, 预测步数为 3, 每步预测含 3 个指令周期, 预测时用的行为数为 6.

图 3 为 E 起始角度为 0.35 rad 下的两种策略的仿真结果. 图 3(a) 中, E 末端的轨迹是射门后球的轨迹.



(a) P 直接追球



(b) P 做短期预测
o - E 轨迹, + - E 轨迹

图 3 E 起始角度为 0.35 rad 时的两种仿真结果
Fig. 3 The two simulation results with 0.35 rad starting angle of E

5 实 验

将基于短期预测的追捕方法应用到实际的中型自主足球机器人中的主要要求是:机器人的自定位准确、可靠,同时能准确估计出对手的位置朝向以及知道对方运球时的线速率。目前,在“交龙”机器人上对该方法进行了初步应用实践:通过融合全视觉和里程计信息,获得了比较稳定的自定位,在 $8\text{ m} \times 5\text{ m}$ 的实验场地中获得的自定位精度可达 0.1 m 级;通过对最近几个时刻球位置进行曲线拟合,近似估计出对手的速度和方向,并以此作为预测的基础。

由于多种因素的影响,并不是在任何时刻都能获得足够多的可靠信息用于估计对手的速度和方向,所以在初步应用实践中,采用混合方法:在能够近似估计出对手速度和朝向的情况下应用基于短期预测的追捕方法,在无法估计出对手速度的情况下使用其他追捕方法。

在 $8\text{ m} \times 5\text{ m}$ 的实验场地中,做了一个对方运球进攻时的拦截比较实验:一是机器人直接朝球运动进行拦截;二是采用混合拦截方法:机器人在能估计出对手的当前速度和朝向的情况下,采用基于短期预测的追捕方法进行拦截,否则采用直接朝球运动的拦截方式。

实验方法:在 $8\text{ m} \times 5\text{ m}$ 的场地中(见图 4),拦截方 P 的初始位置为 $(2.5\text{ m}, 2\text{ m})$, 朝向为 $\pi/2$; 进攻方 E 初始位置为 $(2.5\text{ m}, 4.5\text{ m})$, 朝向在 $[-\pi/2, 0]$ 范围内随机选取(由于在实际实验中,影响拦截最终结果的带有随机性的因素很多,所以采用在此角度范围内随机摆放机器人 E 初始朝向);分别应用两种方法进行拦截,各进行 50 次实验,记录拦截成功的次数。

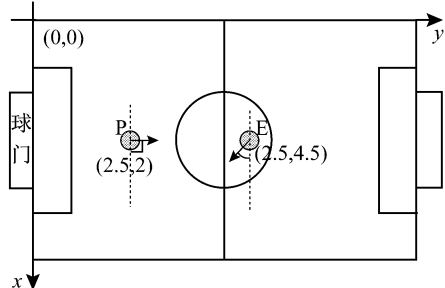


图 4 实验初始位姿示意图

Fig. 4 The robots' initial poses

实验中, $v_{P,\max} = 1\text{ m/s}$, $v_E = 1\text{ m/s}$, $\omega_{P,\max} = 1.5\text{ rad/s}$, $\omega_{E,\max} = 1\text{ rad/s}$ 。在应用基于短期预测的追捕方法时, P 的线速率均匀地分为 5 档, 角速度也均匀分为

5 档, 即 P 共有 25 种行为; E 的角速度被均匀分为 11 档。P 进行的是两步预测, 每步预测时间为 800 ms 。

基于预测的追捕方法中, 针对某个 E 状态 j 的某状态 i 下 P 的个别收益 f_{ij} 计算方法如下:

$$f_{ij} = \begin{cases} 1 & \text{球出界} \\ 0 & \text{球进门} \\ 0.6e^{-d_1/8} + 0.4e^{-d_2/4} & \text{其他} \end{cases} \quad (9)$$

拦截实验结果为: 采用直接追球的拦截方式, 共拦截成功 13 次, 拦截成功率为 26%; 采用混合拦截方法, 共拦截成功 21 次, 拦截成功率为 42%。

实验结果表明, 结合了基于预测追捕方法的拦截成功率与直接追球方法相比有所提高, 这说明基于预测的追捕方法有一定的实际价值, 但效果提高还不是很明显, 可能原因有: 对对手的运动估计不够精准; 机器人自定位存在一定误差; 收益函数的构建不尽合理等。

6 结 语

抛开参与人是否理性的假设, 本文假定追捕者难以知道逃避者到底有多聪明、下一步到底会做什么。以此为基础, 针对具有安全区的追捕逃避问题探讨了一种基于短期预测的追捕方法: 根据参与人所受物理条件限制对参与人进行短期状态预测, 对预测出的结果进行评估并以评估结果为依据做出行为决策。从参与人因受物理条件限制, 短期内所能出现的位置和状态空间有限的角度和伪蒙特卡罗方法的原理, 分析了进行短期预测的可行性。

本文在总结自主式足球机器人开发实践的基础上, 构造了仿真软件, 在相同物理参数、相同逃避者算法策略的情况下, 追捕者分别采用直接追球策略和基于短期预测的追捕策略的仿真。比较两种策略的仿真结果, 证实基于短期预测的追捕策略能大大提高追捕效能。

本文还介绍了将基于预测的方法应用于实际机器人的初步实践过程中的一个实验, 该实验表明基于预测的追捕方法有一定的实际价值。下一步的工作是准备进一步提高机器人环境建模的性能, 同时找出更多的反映对抗形势的特征因子, 构建更合理的收益评估函数, 以进一步通过实践来探究基于短期预测的追捕方法。

(下转第 1078 页)

由图9中最后一幅照片可见,机器人最终并没有准确到达指定位置.实际上更多的实验揭示出最后机器人大多停止在以目标点为中心的一个区域中.造成这种结果的主要原因是由于履带与地面复杂的接触、打滑的存在,实际系统与简化理论模型之间的差异等等.

4 结 语

本文提出一种实用的低成本自主移动机器人导航设计方案,以研制的模型样机为例阐述了其具体实现技术.仿真实验及模型样机的实际运行效果均验证了这种系统构成的可行性.由于这种方案本质上是模块化的,因此便于对已有移动机器人系统进行改造以实现自主导航功能.

本文研究内容是为了进一步提高遥操作排爆机器人 PBJ-1^[9] 的移动操作性与可靠性,而使其能够具有一定局部自主导航功能的前期研究工作.成熟的结果将应用于 PBJ-1.

参考文献:

- [1] Lian S H. Fuzzy logic control of an obstacle avoidance robot[C]// *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*. New Orleans, USA: IEEE, 1996: 26 - 30.
- [2] Tan K C, Tan K K, Lee T H, *et al.* Autonomous robot navigation based on fuzzy sensor fusion and reinforcement learning [C]// *Proceedings of the 2002 IEEE International Symposium on Intelligent Control*. Vancouver, Canada: IEEE, 2002: 182 - 187.
- [3] Maaref H, Barret C. Sensor-based fuzzy navigation of an autonomous mobile robot in an indoor environment[J]. *Control Engineering Practice*, 2000, 8(7): 757 - 768.
- [4] Grabowski B. The CMU millibots [EB/OL]. http://www.contrib.andrew.cmu.edu/~rjg/millibots/millibot_project.html.
- [5] 宁柯军, 杨汝清. 基于多智能体的机械臂嵌入式系统控制[J]. *上海交通大学学报*, 2005, 39(12): 2016 - 2019.
NING Ke-jun, YANG Ru-qing. The manipulator control by multi-agent based embedded system[J]. *Journal of Shanghai Jiaotong University*, 2005, 39(12): 2016 - 2019.
- [6] Doitsidis L, Valavanis K P, Tsourveloudis N C. Fuzzy logic based autonomous skid steering vehicle navigation[C]// *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*. Washington DC, USA: IEEE, 2002: 2171 - 2177.
- [7] Yang S X, Li H, Meng M Q, *et al.* An embedded fuzzy controller for a behavior-based mobile robot with guaranteed performance[J]. *IEEE Transactions on Fuzzy Systems*, 2004, 12(4): 436 - 446.
- [8] Borenstein J, Feng L. Measurement and correction of systematic odometry errors in mobile robots [J]. *IEEE Transactions on Robotics and Automation*, 1996, 12(6): 869 - 880.
- [9] 宁柯军. 面向机器人的基于多智能体嵌入式控制系统设计方法及其应用研究[D]. 上海: 上海交通大学机械与动力工程学院, 2006.
- [10] Montaner M B, Ramirez-Serrano A R. Fuzzy knowledge-based controller design for autonomous robot navigation [J]. *Expert Systems with Applications*, 1998, 14(1 - 2): 179 - 186.

(上接第 1073 页)

参考文献:

- [1] 周浦城, 洪炳容. 基于对策论的群集人追捕-逃跑问题研究[J]. *哈尔滨工业大学学报*, 2003, 35(9): 1056 - 1059.
ZHOU Pu-cheng, HONG Bing-rong. Group robot pursuit-evasion problem based on game theory [J]. *Journal of Harbin Institute of Technology*, 2003, 35(9): 1056 - 1059.
- [2] Lavelle S M, Hinrichsen J E. Visibility-based pursuit-evasion: The class of curved environments[J]. *IEEE Transactions on Robotics and Automation*, 2001, 17(2): 196 - 202.
- [3] Shima T. Capture zones in a pursuit-evasion game [C]// *Proceedings of the 42nd IEEE Conference on Decision and Control*. Maui, Hawaii USA: IEEE, 2003: 5450 - 5455.
- [4] Shinar J, Lipman Y, Zarkh M. Mixed strategies in missile versus missile interception scenarios [C]// *Proceedings of the American Control Conference*. Seattle, Washington: IEEE, 1996: 4116 - 4120.
- [5] Turetsky V, Shinar J. Missile guidance laws based on pursuit-evasion game formulations [J]. *Automatica*, 2003, 39: 607 - 618.
- [6] Hespanha J P, Prandini M, Sastry S. Probabilistic pursuit-evasion games: A one-step Nash approach [C]// *Proceedings of the 39th IEEE Conference on Decision and Control*. Sydney, Australia: IEEE, 2000: 2272 - 2277.