

文章编号: 1006-4710(2015)03-0310-06

基于学习自动机的移动机器人导航行为协调控制

郝大鹏^{1,2}, 傅卫平¹, 王雯¹

(1. 西安理工大学 机械与精密仪器工程学院, 陕西 西安 710048; 2. 西安航空学院 理学院, 陕西 西安 710077)

摘要: 针对移动机器人导航控制中的行为协调问题, 提出了一种学习速率可控的学习自动机。该方法将机器人与障碍物之间的接触时间变化作为奖惩信号, 通过主动控制机器人线速度来调节学习决策时间, 根据环境动态特性调整行为动机, 并通过学习决策时间和行为动机控制共同控制学习速率, 保证机器人在学习决策时间内完成导航行为的协调执行。仿真实验提出的学习自动机应用于移动机器人导航控制是可行的, 与动力学分岔控制方法对比, 在未知动态环境中进行导航行为协调控制, 提出方法的安全性更高。

关键词: 移动机器人; 行为动力学方法; 行为协调; 学习自动机

中图分类号: TP24 **文献标志码:** A

Behavior coordination control based on mobile robot navigation of learning automata

HAO Dapeng^{1,2}, FU Weiping¹, WANG Wen¹

(1. Faculty of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an 710048, China; 2. School of Science, Xi'an Aeronautical University, Xi'an 710077, China)

Abstract: With an aim at behavior coordination problem in the mobile robot navigation control, this paper suggests a kind of learning automata with the controllable rate. The learning automata uses the contact time variation between the robot and obstacles as the reward and penalty signals to regulate the learning and decision time through the active control over robot linear velocity and to adjust the behavior intension in terms of environ dynamic behaviors and to control over the shared management learning rate through the learning decision time and the behavior intension control, whereby ensuring that the robot can complete the coordination implementation of navigation behaviors within the learning decision time. The simulation results indicate that the learning automata suggested in this paper in applying to the navigation control by mobile robot is feasible, and in comparison with the dynamic bifurcation control method, its safety in unknown dynamic environ to carry out navigation behavior coordination control is much higher.

Key words: mobile robot; behavior dynamics method; behavior coordination; learning automata

行为动力学导航方法是一种基于行为的移动机器人导航方法, 具有很高的自治性和鲁棒性^[1-2], 该方法已取得了一些研究成果, 例如, 郝大鹏^[3]提出跟随吸引子控制律解决线速度受限问题; Monteiro^[4]利用该方法实现了多机器人编队问题; 雷艳敏^[5]、郝大鹏^[6-7]等分别针对该方法在动态未知环境中导航

安全问题提出了解决方案。

经典行为动力学方法^[1]通过线性叠加方式完成整体行为输出。线性叠加存在行为相互抵消的问题, 针对该问题, Large^[8]和 Ijspeert^[9]分别提出了动力学分岔控制方法调节行为竞争输出, 但这些方法都没有融入学习策略, 不能很好地适应动态未知

收稿日期: 2014-11-28

基金项目: 国家自然科学基金资助项目(10872160, 51275407, 51475365); 陕西省自然科学基金基础研究计划重点资助项目(2011JZ012)。

作者简介: 郝大鹏, 男, 博士生, 讲师, 研究方向为机器人行为动力学导航控制。E-mail: haodpwork@163.com。

通讯作者: 傅卫平, 男, 博士, 教授, 主要研究方向为智能机器人控制、现代物流系统工程与技术、智能车辆控制理论与技术。E-mail: weipingf@xaut.edu.cn。

环境。

本文针对行为动力学导航方法,提出一种带有行为动机机制的学习自动机,可以动态调节学习自动机的学习速率,完成避障行为之间的协调输出,避免避障行为之间相互抵消。仿真结果表明本文方法在动态未知环境下进行导航控制比 Large^[8] 提出的动力学分岔控制方法更加安全。

1 行为动力学导航行为

行为动力学导航方法通过航向角和线速度控制移动机器人导航,根据文献[1,6-7],给出航向角和线速度控制的描述方法。

航向角控制包括趋向目标行为和避障行为,趋向目标行为表示为:

$$f_0 = -\lambda_0 \sin(\theta - \theta_0) \quad (1)$$

式中, θ 为机器人的航向角, θ_0 为目标和机器人连线与 x 轴的夹角, λ_0 为航向角控制强度。

避障行为表示为:

$$f_i = \lambda_i (\theta - \psi_i) e^{-\frac{(\theta - \psi_i)^2}{2\sigma_i^2}} \quad (2)$$

$$\sigma_i = \arctan \left[\tan \left(\frac{\Delta\psi_i}{2} \right) + \frac{R_{\text{robot}}}{R_{\text{robot}} + d_i} \right] \quad (3)$$

其中 $i \in \mathbf{N}^+$ (正自然数), ψ_i 是障碍物 i 和机器人连线与 x 轴的夹角, $\Delta\psi_i$ 是障碍物 i 遮蔽机器人视野产生的角度, R_{robot} 为机器人的半径, d_i 为机器人与障碍物 i 的距离。

以上变量的几何表示如图1所示。

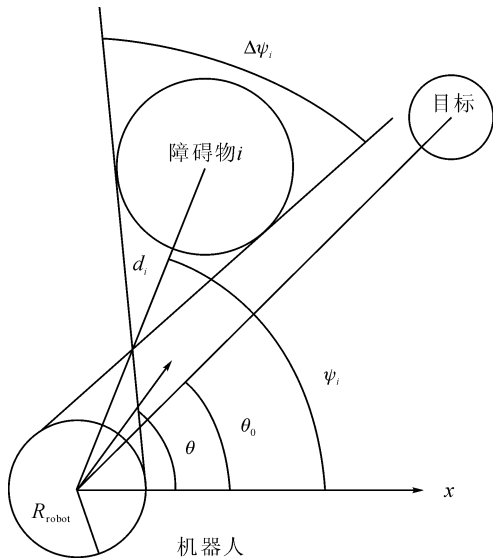


图1 航向角控制方程中变量的几何表示

Fig.1 The geometrical representation of head angle control equation

线速度控制表示为:

当未感知到障碍物时,

$$dv/dt = -\gamma_0 (v - V_0) \quad (4)$$

当感知到障碍物时,

$$dv/dt = -\gamma_i (v - \min(V_i)) \quad (5)$$

式中, v 为机器人线速度, V_0 与 V_i 分别为目标和障碍物产生的速度吸引子, γ_0 和 γ_i 为权值。速度吸引子计算的方法是 $V_0 = C_0 d_0 \dot{\theta}_{\text{max}}$ 和 $V_i = C_i d_i \dot{\theta}_{\text{max}}$,其中 $\dot{\theta}_{\text{max}}$ 为机器人航向角最大转角速度, d_0 与 d_i 分别为机器人与目标和障碍物 i 的距离, C_0 与 C_i 为权值。

2 带有行为动机机制的学习自动机

学习自动机^[10](Learning automata, LA)是一种重要的基于学习策略的行为选择方法。该方法具有计算复杂度低、通用性强等特点,已被广泛地应用于无线认知网络控制^[11]、能源控制^[12]、股票投资^[13]等领域。但经典学习自动机存在行为收敛速度慢等问题,本文提出一种学习速率可控的学习自动机,可以根据环境变化动态调节学习自动机的学习速率,改善了学习自动机实时控制能力。

2.1 学习自动机框架

行为动力学方法设计的导航行为包括趋向目标行为和避障行为,学习自动机可以作为导航行为协调选择控制器。

初始时刻,学习自动机选择趋向目标行为控制机器人奔向目标;当机器人感知环境中障碍物时,学习自动机根据机器人与环境的交互信息选择恰当避障行为;当危险解除时,学习自动机将选择趋向目标行为控制机器人继续奔向目标。

选择恰当的避障行为是设计学习自动机的重点。在学习自动机选择避障行为时应该注意:第一,时间控制,即学习自动机必须在机器人陷入不可避免碰撞状态前完成避障行为选择;第二,潜在危险控制,即学习自动机在选择瞬时最佳行为更新执行概率的同时,应注意潜在危险。

根据以上分析,本文提出的学习自动机相对于经典学习自动机增加了决策时间控制和行为动机控制(潜在危险控制),如图2所示。学习策略在学习自动机与环境交互过程中,首先计算行为选择所需时间限制,根据环境的动态变化更新行为动机,确定学习自动机当前学习速率,之后更新行为执行的概率。

2.2 学习自动机建模

学习自动机一般可以定义为六元组 $\langle \mathbf{I}, \mathbf{A}, \mathbf{S}, \mathbf{O}, \mathbf{R}, \xi \rangle$,其中 \mathbf{I} 为环境感知, \mathbf{A} 为行为集, \mathbf{S} 为行为状态集, \mathbf{O} 为行为执行函数, \mathbf{R} 为奖惩信息, ξ 为状态过渡函数。

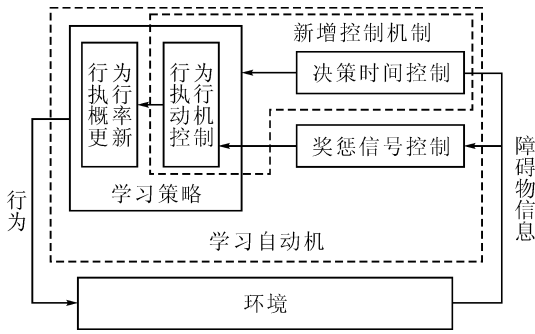


图2 学习自动机框架

Fig. 2 The framework of learning automata

环境感知 \mathbf{I} 存储着环境反馈信息,是行为选择的依据。针对导航问题,本文将机器人与障碍物的接触时间^[7]作为导航环境信息,用于描述障碍物危险程度。观察图1,第 t 次采样时刻机器人与障碍物 i 的接触时间 $T_i(t)$ 为:

$$T_i(t) = \frac{d_i \sin(\psi_i)}{v_i \sin(\theta) + v_i \sin(\psi_i)} \quad (6)$$

式中 v_i 为障碍物 i 的瞬时速度。

利用式(6)计算时,存在 $T_i(t) = \infty$ 的情况,表示障碍物 i 对于机器人不存在危险。

环境感知 \mathbf{I} 表示为:

$$\mathbf{I} = \begin{bmatrix} T_1(1) & \cdots & T_n(1) \\ \vdots & \ddots & \vdots \\ T_1(m) & \cdots & T_n(m) \end{bmatrix} \quad (7)$$

向量 $\mathbf{I}[t, \cdot]$, $t \in \{1, \dots, m\}$ 为环境感知 \mathbf{I} 的行向量,存储着第 t 次采样时所有障碍物的接触时间, $T_i(t)$ 为 $\mathbf{I}[t, \cdot]$ 的第 i 个分量, $i \in \{1, \dots, n\}$, 是第 t 次采样时刻机器人与障碍物 i 的接触时间。 $\mathbf{I}[\cdot, i]$ 为环境感知 \mathbf{I} 的列向量,存储着障碍物 i 的接触时间历史。

存在第 t 次采样时,首次感知到障碍物 i 的情况,此时环境感知 \mathbf{I} 增加新列,并规定: $[T_i(1), \dots, T_i(t-1)] = \infty$, 表示第 1 到第 $t-1$ 次采样时障碍物 i 不存在危险。

为了避免环境感知矩阵 \mathbf{I} 的列无限增长,规定当某个障碍物持续一定时间(该时间为常值)对机器人不存在危险时, $\mathbf{I}[\cdot, i]$ 将从环境感知 \mathbf{I} 中剔除。

行为集 $\mathbf{A}[a_0, a_1, a_2, \dots, a_n]$ 为机器人导航行为,包括始终存在的趋向目标行为 a_0 及避障行为 $a_1 \sim a_n$, 行为由上节提出行为动力学导航方法实现。

行为状态集 $\mathbf{S}[P_0, P_1, \dots, P_n]$ 存储着行为的执行状态,下标与行为集中行为的下标一一对应, P_0 为趋向目标行为执行状态, $P_1 \sim P_n$ 为避障行为的执行状态,本文利用概率描述行为状态,且所有行为执

行概率和为 1, 即 $\sum_{j=0}^n P_j = 1$ 。

当行为执行概率为 1 时,该行为将被执行,可以定义行为执行函数 \mathbf{O} 为:

$$\mathbf{O}(p_j = 1) = \text{执行 } j \text{ 行为}, j \in \{0, 1, \dots, n\} \quad (8)$$

学习自动机根据环境给予的瞬时奖惩信息完成行为状态过渡,本文采用 P 型环境,即奖惩为离散值 0 或 1,其中 1 为奖励,0 为惩罚。

奖惩信息向量 $\mathbf{R}[0, \dots, n]$ 为行为的瞬时奖惩信息,其中 $\mathbf{R}(0)$ 为趋向目标行为的瞬时奖惩信息, $\mathbf{R}(i)$, $i \in \{1, \dots, n\}$ 为避障行为的瞬时奖惩信息。

趋向目标行为与避障行为的特点不同,奖惩方式也不同。对于趋向目标行为而言:

$$\begin{cases} \mathbf{R}_0(t) = 1 & \text{当 } T_i(t-1) \neq \infty, T_i(t) = \infty \text{ 时} \\ \mathbf{R}_0(t) = 0 & \text{其他情况时} \end{cases} \quad (9)$$

式(9)表示在第 t 次采样时刻,如果任意障碍物 i 对机器人危险解除时,趋向目标行为获得奖励。

对于避障行为而言:

$$\begin{cases} \mathbf{R}_i(t) = 1 & \text{当 } V(t) \neq \infty, i = D(t) \text{ 时} \\ \mathbf{R}_i(t) = 0 & \text{当 } V(t) \neq \infty, i \neq D(t) \text{ 时} \end{cases} \quad (10)$$

$$D(t) = \operatorname{argmin}(I(t, \cdot)) \quad (11)$$

$$V(t) = \min(I(t, \cdot)) \quad (12)$$

式中 $i \in \{1, \dots, n\}$, $D(t)$ 为第 t 次采样时刻接触时间最小的障碍物编号; $V(t)$ 是最小的接触时间。式(10)表示在第 t 次采样时刻,最危险的障碍物获得奖励。

行为状态 $\mathbf{S}[P_0, P_1, \dots, P_n]$ 根据奖惩信息向量 $\mathbf{R}[0, \dots, n]$ 的变化而改变。任意采样时刻,首先更新趋向目标行为的行为状态,之后再更新避障行为的行为状态。提出的状态过渡函数 ξ 为:

当 $\mathbf{R}_0(t) = 1$ 时,

$$\begin{cases} P_0(t) = P_0(t-1) + \sum P_k(t-1) \\ P_k(t) = 0 \end{cases} \quad (13)$$

式中 $k \in \{1, \dots, n\}$, 且 $T_k(t-1) \neq \infty, T_k(t) = \infty$ 。

当 $\mathbf{R}_0(t) = 0$ 时,

$$P_0(t) = P_0(t-1) \quad (14)$$

式(13)和(14)描述了趋向目标行为状态的更新方式,即在障碍物危险解除时,该行为的执行概率立即分配给趋向目标行为,保证趋向目标行为状态可以获得更新。

对于避障行为状态的更新,采取传统学习自动机的追击学习(pursuit learning)策略^[14]实现,该算法可以保证行为选择是 ϵ 最优(ϵ -optimality)的。

当 $R_i(t)=1$ 时,

$$P_i(t) = P_i(t-1) + \kappa[1 - P_i(t-1)] \quad (15)$$

当 $R_i(t)=0$ 时,

$$P_i(t) = P_i(t-1) - \kappa P_i(t-1) \quad (16)$$

其中 $i \in \{1, \dots, n\}$, 且 $i \neq k$, $P_i(t-1) \neq 0$, κ 是学习速率。

2.3 决策时间及行为动机控制

状态过渡函数 ξ 的核心部分是避障行为状态更新。学习自动机必须在机器人与障碍物发生碰撞前完成行为选择。本文提出决策时间和行为动机控制动态调节学习自动机的学习速率 κ , 保证学习自动机的实时性。

决策时间控制实时计算学习自动机的决策时间范围。首先,环境中存在接触时间较大障碍物,此障碍物对机器人的危险较小,根据环境需要设定一时间 $T_{\text{allowable}}$ 作为允许接触时间,当障碍物的接触时间大于 $T_{\text{allowable}}$ 时,认为该障碍物不存在危险。

其次,学习自动机必须在机器人陷入不可避免碰撞状态^[14-15]之前完成避障行为的选择。

设机器人当前速度为 v , 以允许的最大加速度值减速至 0 时,所需时间为 T_{break} , 障碍物 i 的接触时间将变为 T_{max}^i , 那么必须在 T_{decision}^i 时间内完成避障行为 a_i 选择, T_{decision}^i 可以表示为:

$$T_{\text{decision}}^i = T_{\text{max}}^i - T_{\text{break}} \quad (17)$$

假设机器人避障时始终运动,机器人最小运动速度为 v_{small} , T_{max}^i 可以近似表示为:

$$T_{\text{max}}^i \approx \frac{d_i \sin(\psi_i)}{\frac{(v - v_{\text{small}})}{2} \sin(\theta) + v_i \sin(\psi_i)} \quad (18)$$

计算所有存在碰撞危险障碍物的 T_{decision}^i , 将最小的 T_{decision}^i 作为学习自动机的学习决策时间 T_{decision} 。在任意第 t 采样时刻时,决策时间范围为 $(T_{\text{decision}}(t), T_{\text{allowable}})$ 。

行为动机控制目标是评价环境中障碍物危险的动态变化。

观察环境感知历史 $I[\cdot, i]$, 存在接触时间加速变化的情况,可以利用行为动机描述危险的动态变化。

当第 t 次采样时刻, $T_i(t) \neq \infty$, 且接触时间历史 $I[\cdot, i]$ 中第 t 次采样之前的 k 次接触时间都不为无穷大时, $k \geq 1$, 即:

$$T_i(t) = T_i(t-1) = \dots = T_k(t-k) = \infty$$

表示连续 $k+1$ 次传感器采样,障碍物 i 都存在危险,则行为动机 M 可以表示为:

$$M_i(t) = \sum_{j=t-k}^t \left(\frac{\Delta T_i(j) - \Delta T_i(j-1)}{T_{\text{decision}}(j)} \right) \quad (19)$$

式(15)表示如果第 t 次采样之前障碍物存在连续危险时,行为动机累积接触时间的变化率。

学习速率 κ 可以通过提出的决策时间控制和行为动机控制共同调节,即:

$$\kappa_i(t) = 1 - (1 - N_i(t))^{\tan^{-1} \frac{M_i(t)\pi}{2}} \quad (20)$$

式中 $N_i(t) = \frac{T_i(j) - T_{\text{decision}}(t)}{T_{\text{allowable}} - T_{\text{decision}}(t)}$ 为第 t 次采样时刻接触时间在决策时间范围上的归一化。

当障碍物 i 获得奖励,则根据式(20)计算 $\kappa_i(t)$, 作为此刻式(15)、(16)行为状态更新的瞬时学习速率。图 3 是学习速率 $\kappa_i(t)$ 随决策时间和行为动机变化的示意图。

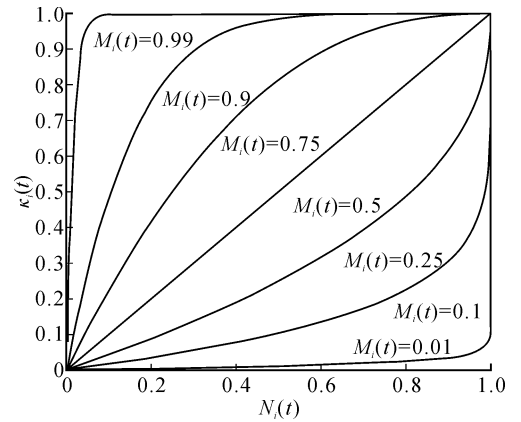


图 3 学习速率变化示意图

Fig. 3 The sketch of learning rate changing

根据以上分析,给出本文提出的学习自动机主体算法。

ALGORITHM

初始化:

$$t=0$$

$\mathbf{A}=[a_0]$, 初始时仅有趋向目标行为

$\mathbf{S}=[p_0=1]$, 初始时趋向目标行为状态为 1

Loop

利用式(8)执行行为

传感器感知环境

IF 感知障碍物

对感知的 n 个障碍物编号

利用式(6)计算 T_i

利用式(7)更新环境感知 \mathbf{I}

$$\mathbf{A}=[a_0]+[a_1, \dots, a_n]$$

$$\mathbf{S}=[p_0=1]+[\underbrace{0, \dots, 0}_n]$$

End IF

利用式(9)和式(10)计算环境奖惩信息 R

利用式(13)和(14)更新行为状态

利用式(17)计算行为决策时间

利用式(19)计算行为动机 M

利用式(20)计算瞬时学习速率 κ

利用式(15)、(16)更新行为状态

$t = t + 1$

End Loop, 直到机器人到达目标

END ALGORITHM

3 仿 真

仿真环境为 $50\text{ m} \times 50\text{ m}$ 的正方形平面。环境中包括4个障碍物,圆形静态障碍物1初始位置为 $(20\text{ m}, 20\text{ m})$, 半径为 2 m ; 圆形静态障碍物2初始位置为 $(30\text{ m}, 28\text{ m})$, 半径为 2 m 。圆形动态障碍物1位于 $(28\text{ m}, 35\text{ m})$, 半径为 2 m , 线速度为 2 m/s , 运动方向为 -90° ; 圆形动态障碍物2位于 $(35\text{ m}, 40\text{ m})$, 半径为 2 m , 线速度为 1 m/s , 运动方向为 120° 。圆形目标位于 $(40\text{ m}, 35\text{ m})$, 半径为 2 m 。圆形机器人初始位置为 $(10\text{ m}, 5\text{ m})$, 半径为 2 m , 初始航向角为 120° , 初始线速度为 3 m/s 。行为动力学导航模型参数包括 $\lambda_0 = 0.5, \lambda_i = 1, \gamma_0 = 0.4, \gamma_i = 1.2$ 。本文提出方法仿真机器人运动轨迹如图4所示, 参考文献[8]仿真机器人运动轨迹图5所示, 图6给出了本文提出方法与参考文献方法行为协调的时间序列图。

参考文献[8]在利用行为动力学方法导航时并未增加线速度控制, 本文在仿真时为其增加了与本文相同的行为动力学速度控制。

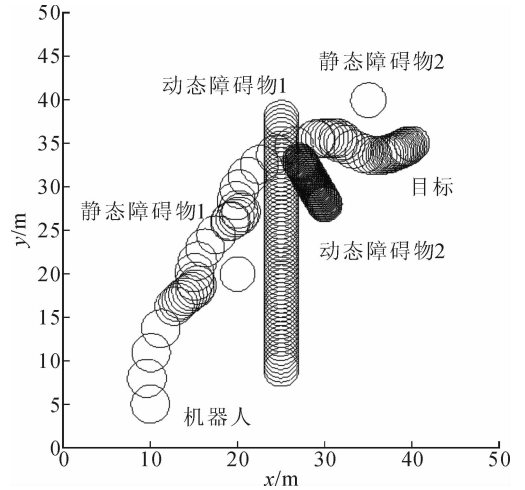


图4 本文提出方法在动态环境的仿真结果
Fig. 4 The dynamic environment simulation results of the method proposed in this paper

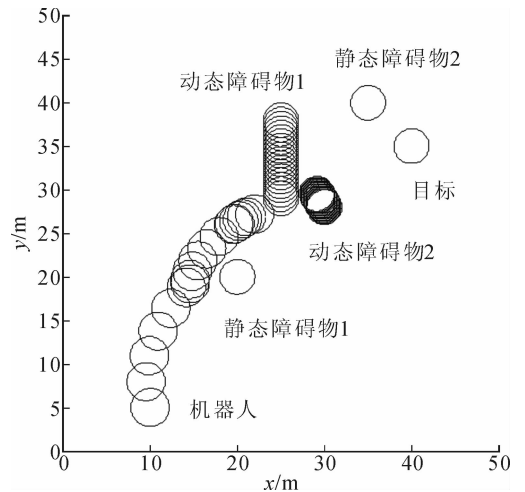


图5 参考文献[8]中动态环境的仿真结果
Fig. 5 The dynamic environment simulation results of the method proposed in reference [8]

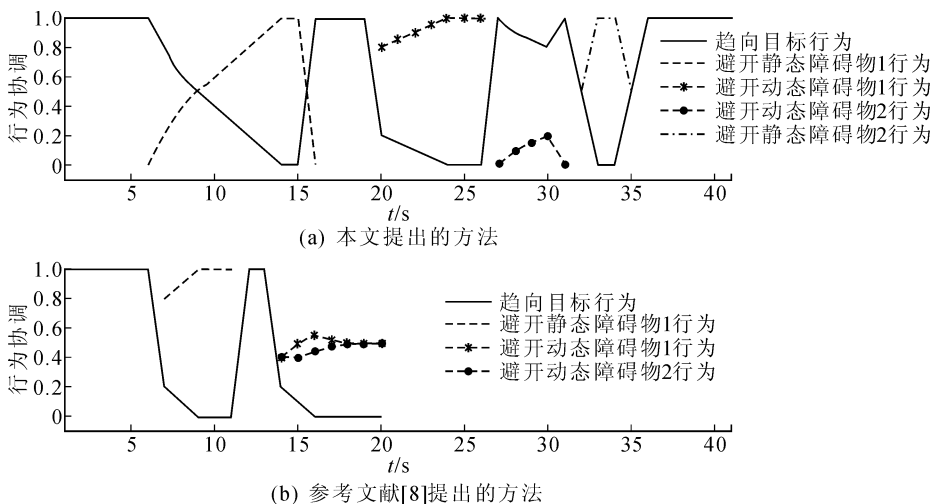


图6 行为协调的时间序列图仿真对比

Fig. 6 Simulation comparison of the time sequence diagram for behavior coordination

由公式(17)可知,本文提出的学习自动机为了增大决策时间,采取控制机器人减速策略,当接近障碍物时缓慢调整位姿,即选择障碍物进行避障时采用“轨迹最优”的优劣评价标准。如图4所示在避开静态障碍物1时,在接近 T_{decision}^1 时选择了避开障碍物1,比图5所示的参考文献[8]提出的方法机器人运动轨迹更优。

观察图4,本文提出的学习自动机在避开动态障碍物1和2时,由于动态障碍物1和2的线速度不同,对机器人的威胁不同,对动态障碍物1的行为动机快速增大,使得学习自动机的学习速率也快速增大,促使机器人在距离动态障碍物1较远的位置便开始执行对该障碍物的避障行为,而对动态障碍物2的学习速率增长得很慢,该行为的执行概率始终未达到1,所以并未执行避障行为,而执行的是趋向目标行为,机器人保持3 m/s的线速度越过动态障碍物2。观察图5,参考文献[8]提出的方法在动态障碍物1和2相遇时,避障行为竞争输出最大,始终未选择合适障碍物避障,最终与障碍物1发生碰撞。

4 结 论

本文提出的学习自动机与经典学习自动机^[10]相比,增加了决策时间及行为动机控制,并利用它们调节学习自动机的学习速率,通过仿真实证提出的方法可以应用于实时性较强的未知环境移动机器人动态导航控制。但在未知环境中障碍物的运动通常是任意的,故而在本文提出的学习自动机基础上增加障碍物运动模式识别是进一步研究的方向。

参考文献:

- [1] Bicho E, Mallet P, Schöner G. Target representation on an autonomous vehicle with low-level sensors [J]. *The International Journal of Robotics Research*, 2000, 19(5):424-447.
- [2] Warren W H. The dynamics of perception and action [J]. *Psychological Review*, 2006, 113(2):358-389.
- [3] 郝大鹏,傅卫平,杨世强. 移动机器人行为动力学导航方法研究 [J]. *机械科学与技术*, 2013, 32(10): 1488-1491.
Hao Dapeng, Fu Weiping, Yang Shiqiang. Study on the navigation method of behavior dynamics in mobile robot [J]. *Mechanical Science and Technology for Aerospace Engineering*, 2013, 32(10): 1488-1491.
- [4] Monteiro S, Bicho E. Attractor dynamics approach to

formation control: theory and application [J]. *Autonomous Robots*, 2010, 29(3):331-355.

- [5] 雷艳敏,朱齐丹,冯志彬. 基于速度障碍和行为动力学的动态路径规划 [J]. *华中科技大学学报:自然科学版*, 2011, 39(4): 15-19.
Lei Yanmin, Zhu Qidan, Feng Zhibin. Dynamic path planning using velocity obstacles and behavior dynamics. [J]. *Journal of Huazhong University of Science and Technology(Natural Science Edition)*, 2011, 39(4): 15-19.
- [6] 郝大鹏,傅卫平,王雯. 基于行为动力学的移动机器人安全导航方法 [J]. *系统工程与电子技术*, 2014, 36(1): 136-142.
Hao Dapeng, Fu Weiping, Wang Wen. Mobile robot safe navigation base on behavior dynamics [J]. *Systems Engineering and Electronics*, 2014, 36(1): 136-142.
- [7] 郝大鹏,傅卫平,王雯. 基于接触时间的移动机器人安全导航 [J]. *计算机应用*, 2014, 34(4): 1209-1212.
Hao Dapeng, Fu Weiping, Wang Wen. Mobile robot safety navigation based on time to contact [J]. *Journal of Computer Applications*, 2014, 34(4): 1209-1212.
- [8] Large E W, Christensen H I, Bajcsy R. Scaling the dynamic approach to path planning and control: competition among behavioral constraints [J]. *The International Journal of Robotics Research*, 1999, 18(1): 37-58.
- [9] 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.
- [10] Thathachar M. Learning systems: stochastic automata models [J]. *Defence Science Journal*, 1985, 35(3): 361-366.
- [11] Golestanian M, Iranmanesh S, Ghazizadeh R, et al. A learning automata based spectrum prediction technique for cognitive radio networks [J]. *Science and Education*, 2014, 2(3): 93-97.
- [12] Zhu J, Jiang P, Gu W, et al. Finite action-set learning automata for economic dispatch considering electric vehicles and renewable energy sources [J]. *Energies*, 2014, 7(7): 4629-4647.
- [13] Mozafari M, Alizadeh R. A cellular learning automata model of investment behavior in the stock market [J]. *Neurocomputing*, 2013, (122): 470-479.
- [14] Lefèvre S, Vasquez D, Laugier C. A survey on motion prediction and risk assessment for intelligent vehicles [J]. *Robomech Journal*, 2014, 1(1): 1-14.
- [15] Fraichard T, Kuffner J J. Guaranteeing motion safety for robots [J]. *Autonomous Robots*, 2012, 32(3): 173-175.

(责任编辑 王卫勋)