doi:10.6041/j.issn.1000-1298.2019.11.003

基于启发式动态规划的履带机器人路径跟随控制方法

宋　彦[1,2]　张羊阳[1,2]　姚　琦[1,2]　袁　胜[3]　廖　娟[1,2]　刘　路[1,2]

(1.安徽农业大学工学院,合肥 230036;2.安徽省智能农机装备工程实验室,合肥 230036;
3.合肥中科智驰科技有限公司,合肥 230088)

摘要:针对移动机器人传统路径跟随控制方法需要人工调校参数、缺乏自主优化能力的问题,提出了一种基于启发式动态规划(Heuristic dynamic programming, HDP)的路径跟随控制方法。首先,设计履带式机器人路径跟随控制系统结构,建立了误差状态方程;其次,提出了一种基于 HDP 算法的路径跟随控制方法,综合误差性能指标和跟随稳定性指标设计了回报函数,采用多层前馈神经网络逼近评价器和执行器,并推导了网络参数的在线优化规则;最后,通过数值仿真和系统试验验证了 HDP 方法的路径跟随性能。试验结果证明,基于 HDP 算法的控制器跟随直线的平均误差绝对值为 0.04 m、均方根误差为 0.06 m;跟随钝角转向曲线的平均误差绝对值为 0.01 m、均方根误差为 0.06 m;跟随锐角转向曲线的平均误差绝对值为 0.03 m、均方根误差为 0.09 m。该方法不需要对控制参数进行反复调试就能够获得较好的控制效果,提高了移动机器人路径跟随控制方法的环境适应性和自主优化能力。

关键词:履带式机器人;路径跟随;启发式动态规划;前馈式神经网络

中图分类号:S24;TP242.6　　**文献标识码**:A　　**文章编号**:1000-1298(2019)11-0024-10

Path Following Control Method of Tracked Mobile Robot Based on Heuristic Dynamic Programming

SONG Yan[1,2]　ZHANG Yangyang[1,2]　YAO Qi[1,2]　YUAN Sheng[3]　LIAO Juan[1,2]　LIU Lu[1,2]

(1. *School of Engineering*, *Anhui Agricultural University*, *Hefei* 230036, *China*
2. *Anhui Province Engineering Laboratory of Intelligent Agricultural Machinery and Equipment*, *Hefei* 230036, *China*
3. *Hefei CAS-Juche Technology Co.*, *Ltd.*, *Hefei* 230088, *China*)

Abstract: Aiming at the problem that the traditional path following control method of mobile robot requires manual tuning parameters and lacks of self-optimization ability, a path following control method based on heuristic dynamic programming (HDP) was proposed. Firstly, the error state equation of the path following system was established based on the structure of tracked robot system. Secondly, a path following controller based on HDP algorithm was proposed. The return function was designed by synthesizing the error performance index and the following stability index. A multi-layer feedforward neural network was used to establish the actuator and evaluation modules, and the online optimization rules of network parameters were derived. Thirdly, the path following performance of HDP method was verified by numerical simulation and system test. The experimental results showed that the controller based on the HDP algorithm of the mean absolute error was 0.04 m and the root mean square error was 0.06 m when following a straight line, 0.01 m and 0.06 m when following a obtuse angle steering curve, and 0.03 m and 0.09 m when following an acute angle steering curve, and the method based on HDP can obtain high precision control effect without repeated debugging control parameters, and it improved the environmental adaptability and autonomous optimization ability of the path following control method for mobile robots.

Key words: tracked robot; path following; heuristic dynamic programming; feedforward neural network

收稿日期:2019-04-30　修回日期:2019-06-18

基金项目:国家自然科学基金项目(61503363)

作者简介:宋彦(1983—),男,副教授,博士,主要从事智能农业技术与装备研究,E-mail: songyan@ahau.edu.cn

0　引言

随着移动机器人在农业领域的广泛应用，复杂的地形和多变的环境对移动机器人的运动控制能力提出了更高的要求，如何提高移动机器人在未知、复杂环境的路径跟随性能成为研究的热点。

路径跟随的控制机理多采用“预瞄—跟随”原理[1-2]，传统的控制方法有纯追踪[3]和 PID 控制[4-5]，但是上述方法均忽略了系统的数学模型，难以取得高精度的控制效果。一些学者提出了基于车辆动力学模型[6]或运动学模型[7]的控制方法，例如：基于期望横摆角速度[8]以及横向速度[9]的状态反馈控制方法，基于模型预测控制[10-11]的方法等。此类方法在模型较为精确的条件下，提高了控制精度，但是需要调节和优化的控制参数较多，尤其是当机器人的工作环境发生变化时，需要重新选择控制参数，优化控制性能。

近年来，如何提高移动机器人的环境适应性和自主行为能力成为学界的研究热点，出现了基于智能控制的方法[12-13]，如基于模糊控制[14-15]和神经网络方法[16]等，如文献[17]采用模糊控制理论调整前视距离，文献[18]提出了一种自适应模糊变结构控制方法。但上述方法仍然需要将专家经验抽象为模糊控制规则，并调校模糊控制参数。文献[19]提出了一种根据车辆速度和道路弯度动态调整预瞄距离的改进纯追踪方法。也有学者采用了一些智能优化算法，如文献[20]采用了遗传算法自动优化控制参数，文献[21]提出了一种基于增量 Q 学习的控制参数优化方法。

自适应动态规划[22]（Adaptive dynamic programming，ADP）理论采用了强化学习机制，通过计算环境对不同策略的评价性反馈信号来优化控制策略，使智能体具备了自主优化能力[23]。ADP 理论包含了不同的实现方法，如启发式动态规划（Heuristic dynamic programming，HDP）、对偶启发式规划（Dual heuristic programming，DHP）等。ADP 理论的实现框架有执行器、评价器两个模块，评价器基于性能指标函数评价当前控制动作的性能优劣，执行器根据评价结果优化控制动作，并通过“执行—评价”迭代运算寻找最优控制策略。近年来，这一方法被用于过程控制、机器人控制等领域。文献[24]提出了一种基于回声状态网络的 HDP 控制方法，用于污水处理过程控制。在移动机器人的应用领域，文献[25]提出了一种基于 HDP 的路径规划方法。文献[26]较早地将 ADP 理论应用于移动机器人的路径跟踪控制问题，但是其建立的回报函数仅考虑了速度误差，且未对该方法进行试验验证。文献[27]采用一种基于 DHP 的方法用于机器人的路径跟随控制，该方法能够估计性能指标函数的导数，但是仍然缺乏相关的试验验证。

为提高移动机器人路径跟随控制方法对环境的适应性和自主优化能力，本文提出一种基于 HDP 算法的路径跟随控制方法，构建综合跟随精度指标和稳定性指标的回报函数，采用神经网络模型逼近评价器和执行器，通过开展对不同线形路径的跟随试验，验证该方法对未知路径线形的适应性。

1　控制系统结构与数学模型

1.1　控制系统结构

本文研究对象为履带式机器人，其路径跟随控制系统结构如图 1 所示，其中参考路径为一系列离散的 GPS 坐标，机器人通过车载 GPS 传感器得到实际位姿。坐标转换模块的功能是通过当前实际位姿和预瞄点信息得到位姿误差信息。路径跟随控制器根据位姿误差由控制算法计算产生控制动作。控制量转换模块将控制动作转换为两侧电机的转速，驱动机器人稳定跟随参考路径。

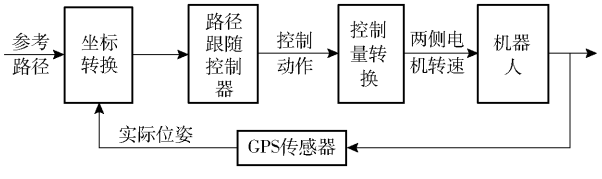


图 1　控制系统结构

Fig. 1　Control system structure

1.2　系统数学模型

本文建立的履带机器人路径跟随示意图如图 2 所示，首先定义大地坐标系为 XOY，其中 X 轴为正东方向，Y 轴为正北方向。定义机器人局部坐标系为 xoy，其中 x 轴方向为机器人前进方向，将 y 轴方向定义为机器人的前进方向逆时针旋转 90°。其中 L 为预瞄距离，(X_c, Y_c) 为移动机器人在全局坐标系

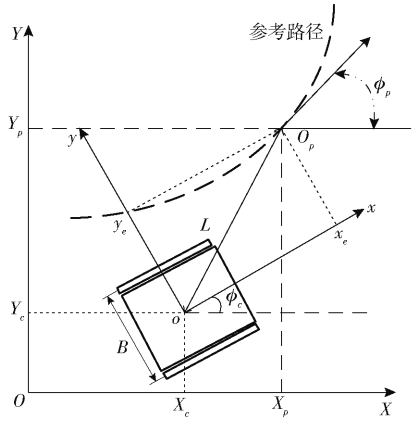


图 2　履带机器人路径跟随示意图

Fig. 2　Tracked robot path following diagram

中的位置坐标，ϕ_c 为车辆纵轴线与大地坐标系横坐标的夹角，(X_c, Y_c, ϕ_c) 定义为机器人的当前位姿。(X_p, Y_p) 为预瞄点在全局坐标系中的位置坐标，ϕ_p 为预瞄点处的切线与大地坐标系横坐标的夹角，(X_p, Y_p, ϕ_p) 定义为预瞄点处的期望位姿。

在不考虑履带滑动的情况下，定义履带机器人的驱动轮半径为 r_d，履带中心距为 B，当前机器人前进速度为 v_c，角速度为 ω_c，两侧电机的转速为 n_L、n_R，v_c 和 ω_c 与两侧电机转速的转换关系为

$$\begin{bmatrix} v_c \\ \omega_c \end{bmatrix} = \begin{bmatrix} \pi r_d & \pi r_d \\ \dfrac{-2\pi r_d}{B} & \dfrac{2\pi r_d}{B} \end{bmatrix} \begin{bmatrix} n_L \\ n_R \end{bmatrix} \tag{1}$$

则履带式机器人的运动学模型可表示为

$$\begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{\phi}_c \end{bmatrix} = \begin{bmatrix} \pi r_d \cos\phi_c & \pi r_d \cos\phi_c \\ \pi r_d \sin\phi_c & \pi r_d \sin\phi_c \\ \dfrac{-2\pi r_d}{B} & \dfrac{2\pi r_d}{B} \end{bmatrix} \begin{bmatrix} n_L \\ n_R \end{bmatrix} \tag{2}$$

移动机器人相对预瞄点的误差状态方程为

$$\begin{bmatrix} x_e \\ y_e \\ \phi_e \end{bmatrix} = \begin{bmatrix} \cos\phi_c & \sin\phi_c & 0 \\ -\sin\phi_c & \cos\phi_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p - X_c \\ Y_p - Y_c \\ \phi_p - \phi_c \end{bmatrix} \tag{3}$$

式中 x_e——纵向偏差

y_e——横向偏差

ϕ_e——航向角偏差

根据图 2 的几何关系，由式(2)、(3)可得移动机器人相对预瞄点 O_p 的横向偏差和航向角偏差的变化率为

$$\begin{cases} \dot{y}_e = -v_y - L\omega + v_x\phi_e \\ \dot{\phi}_e = v_x\rho - \omega \end{cases} \tag{4}$$

式中 ω——横摆角速度

v_x——移动机器人纵向速度

v_y——移动机器人横向速度

ρ——预瞄点处的道路曲率

2 基于 HDP 算法的路径跟随控制策略设计

首先采用马尔可夫决策过程(Markov decision process, MDP)模型描述移动机器人路径跟随控制问题，然后采用 HDP 算法求解该控制问题的近似最优控制策略，并给出算法的实现框架，最后推导路径跟随控制器参数的在线学习规则。

2.1 路径跟随控制问题的 MDP 模型

ADP 算法[28]的理论框架是以马尔可夫决策过程(MDP)为基础的。一个 MDP 模型由四元组 $\{\boldsymbol{S}, \boldsymbol{U}, r(\boldsymbol{s}, u, \boldsymbol{s}'), P(\boldsymbol{s}, u, \boldsymbol{s}')\}$ 表示，其中 $\boldsymbol{S}$ 为状态集合，$\boldsymbol{U}$ 为动作集合，$r(\boldsymbol{s}, u, \boldsymbol{s}')$ 为机器人采用动作 u 使其从状态 $\boldsymbol{s}$ 转移到 $\boldsymbol{s}'$ 时获得的回报，$P(\boldsymbol{s}, u, \boldsymbol{s}')$ 为机器人采用动作 u 使其从状态 $\boldsymbol{s}$ 转移到 $\boldsymbol{s}'$ 时的概率。

针对路径跟随问题，选取状态 $\boldsymbol{s} = (y_e, \phi_e, \omega_c, \rho)^{\mathrm{T}}$，动作 $u = \omega$。由式(4)离散化后可以得到状态转移方程为

$$\boldsymbol{s}(k+1) = \begin{bmatrix} y_e(k) \\ \phi_e(k) \\ \omega(k) \\ \rho(k) \end{bmatrix} + T_s \begin{bmatrix} -L\omega(k) + v_x\phi_e(k) - v_y \\ v_x\rho - \omega(k) \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

式中 T_s——系统的控制周期

基于 MDP 模型，将路径跟随最优控制策略描述为：寻找合适的控制动作 u 使得机器人从状态 $\boldsymbol{s}(k)$ 转移到 $\boldsymbol{s}(k+1)$，且使得性能指标函数最小。性能指标函数公式为

$$J = E\left(\sum_{t=0}^{\infty} \gamma^t r_t \right) \quad (0 < \gamma < 1) \tag{6}$$

式中 $E(\,)$——数学期望

r_t——t 时刻的回报

γ——折扣因子

路径跟随过程中，不但期望减小横向偏差，也期望尽量减少控制动作波动，因此选取的回报函数为

$$r(\boldsymbol{s}(k), u(\boldsymbol{s}(k))) = \boldsymbol{s}(k)^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{s}(k) + (u(\boldsymbol{s}(k)) - \omega_p)^{\mathrm{T}} R (u(\boldsymbol{s}(k)) - \omega_p) \tag{7}$$

其中 $\omega_p = v_x\rho$

式中 $\boldsymbol{Q}$——半正定对角矩阵，表征对横向偏差和航向角偏差指标的权重

R——正常数，表征对期望横摆角速度补偿的权重

2.2 路径跟随控制器设计

本文定义的性能指标函数式(6)也可描述为

$$J(\boldsymbol{s}(k)) = r(\boldsymbol{s}(k), u(\boldsymbol{s}(k))) + \gamma J(\boldsymbol{s}(k+1)) \tag{8}$$

根据 Bellman 优化原理，最优的控制策略为

$$u^*(\boldsymbol{s}(k)) = \underset{u}{\operatorname{argmin}} [r(\boldsymbol{s}(k), u(\boldsymbol{s}(k))) + \gamma J(\boldsymbol{s}(k+1))] \tag{9}$$

定义 $u^*(\boldsymbol{s}(k))$ 为使性能指标函数极小化的最优控制动作，式(8)可以写作

$$J^*(\boldsymbol{s}(k)) = r(\boldsymbol{s}(k), u^*(\boldsymbol{s}(k))) + \gamma J^*(\boldsymbol{s}(k+1)) \tag{10}$$

式中 $J^*(\boldsymbol{s}(k))$——最优性能指标函数

$J^*(\boldsymbol{s}(k+1))$——机器人在下一状态 $\boldsymbol{s}(k+1)$ 时的最优性能指标函数

令 $\partial J^*(\boldsymbol{s}(k))/\partial u^*(\boldsymbol{s}(k)) = 0$，则最优控制动

作可表示为

$$u^*(\boldsymbol{s}(k)) = \frac{-1}{2R}\gamma E^{\pi^*}\left[\left(\frac{\partial \boldsymbol{s}(k+1)}{\partial u^*(\boldsymbol{s}(k))}\right)^{\mathrm{T}}\frac{\partial J^*(\boldsymbol{s}(k+1))}{\partial \boldsymbol{s}(k+1)}\right] \tag{11}$$

为实现上述最优控制，构建如图 3 所示基于 HDP 算法的控制框架。算法实现通常包括 3 部分[21]：评价器、模型以及执行器，评价器的作用为根据状态变量估计性能指标函数 $J(\boldsymbol{s}(k))$ 的值；执行器的作用为根据状态变量计算控制动作 $u(\boldsymbol{s}(k))$；模型对移动机器人下一状态 $\boldsymbol{s}(k+1)$ 进行预测。评价器 1 和评价器 2 的结构及参数均相同，但是分别用于估计当前状态的性能指标函数值以及预测状态的性能指标函数值。为了按照时间正向求解控制动作，HDP 方法往往采用函数近似的方法估计性能指标函数值[21]，本文采用多层前馈神经网络逼近评价器，同样也采用这一模型逼近最优动作。算法执行过程中，采用梯度下降算法依次对动作网络和评价网络里的神经网络权值进行反馈调节，最小化性能指标函数 $J(\boldsymbol{s}(k))$，进而获得使 $J(\boldsymbol{s}(k))$ 最优的控制策略 $u(\boldsymbol{s}(k))$。

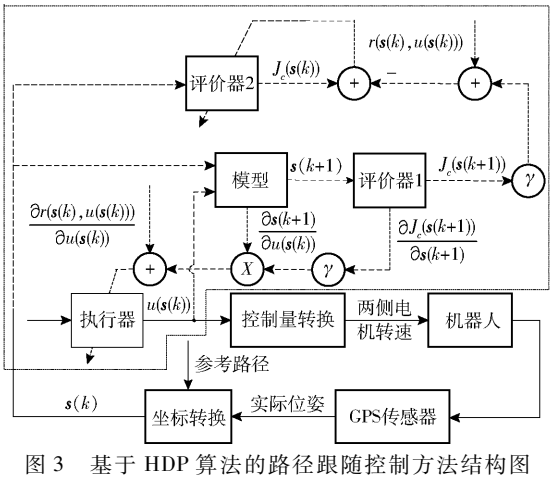


图 3　基于 HDP 算法的路径跟随控制方法结构图

Fig. 3　Structural diagram of path following control method based on HDP algorithm

2.3　控制器参数的在线学习

控制器参数在线学习过程实质上是对神经网络权值的在线调整过程，其目标是寻求一组最优权值使 $J(\boldsymbol{s}(k))$ 最小化，进而获取近似最优控制动作。HDP 算法中评价器网络的优化目标是对最优性能指标函数值的近似，执行器网络优化的目标是找到最优控制动作 $u^*(\boldsymbol{s}(k))$ 使得目标函数 $J(\boldsymbol{s}(k))$ 最小。

本文采用 3 层前馈神经网络实现评价器和执行器，激活函数为 Sigmoid 函数，从隐含层到输出层的函数为线性函数，此时执行器和评价器的表达式分别为

$$u(\boldsymbol{s}(k)) = \boldsymbol{\Phi}_a(\boldsymbol{W}_a\boldsymbol{\sigma}_a(\boldsymbol{V}_a\boldsymbol{s}(k)+\boldsymbol{b}_{a1})+\boldsymbol{b}_{a2}) \tag{12}$$

$$J_c(\boldsymbol{s}(k)) = \boldsymbol{\Phi}_c(\boldsymbol{W}_c\boldsymbol{\sigma}_c(\boldsymbol{V}_c\boldsymbol{s}(k)+\boldsymbol{b}_{c1})+\boldsymbol{b}_{c2}) \tag{13}$$

式中　$J_c(\boldsymbol{s}(k))$——评价器对状态 $\boldsymbol{s}(k)$ 估计的性能指标函数值

$\boldsymbol{V}_a$、$\boldsymbol{V}_c$——神经网络从输入层到隐含层的权值

$\boldsymbol{b}_{a1}$、$\boldsymbol{b}_{c1}$——神经网络从输入层到隐含层的偏置

$\boldsymbol{W}_a$、$\boldsymbol{W}_c$——神经网络从隐含层到输出层的权值

$\boldsymbol{b}_{a2}$、$\boldsymbol{b}_{c2}$——神经网络从隐含层到输出层的偏置

$\boldsymbol{\sigma}_a$、$\boldsymbol{\sigma}_c$——Sigmoid 函数

$\boldsymbol{\Phi}_a$、$\boldsymbol{\Phi}_c$——线性函数

2.3.1　执行器网络参数优化

执行器的学习目标是最小化误差函数

$$E_a(\boldsymbol{s}(k)) = \frac{1}{2}(u(\boldsymbol{s}(k)) - u^*(\boldsymbol{s}(k)))^2 \tag{14}$$

定义 $\zeta_d(k) = u(\boldsymbol{s}(k)) - u^*(\boldsymbol{s}(k))$，则执行器网络的权值和偏置更新规则为

$$\boldsymbol{\Psi}_a(k+1) = \boldsymbol{\Psi}_a(k) - \alpha_a(k)\zeta_d(k)\frac{\partial u(\boldsymbol{s}(k))}{\partial \boldsymbol{\Psi}_a(k)} \tag{15}$$

其中　$\boldsymbol{\Psi}_a(k) = (\boldsymbol{W}_a, \boldsymbol{b}_{a2}, \boldsymbol{V}_a, \boldsymbol{b}_{a1})$

式中　$\alpha_a(k)$——执行器网络的学习率

$\boldsymbol{\Psi}_a(k)$——执行器网络的权值和偏置

由式(12)可得

$$\frac{\partial u(\boldsymbol{s}(k))}{\partial \boldsymbol{W}_a(k)} = \boldsymbol{\sigma}_a(\boldsymbol{V}_a(k)\boldsymbol{s}(k)+\boldsymbol{b}_{a1}(k)) \tag{16}$$

$$\frac{\partial u(\boldsymbol{s}(k))}{\partial \boldsymbol{b}_{a2}(k)} = 1 \tag{17}$$

$$\frac{\partial u(\boldsymbol{s}(k))}{\partial \boldsymbol{V}_a(k)} = \boldsymbol{W}_a(k)\boldsymbol{\sigma}_a'(\boldsymbol{V}_a(k)\boldsymbol{s}(k)+\boldsymbol{b}_{a1}(k))\boldsymbol{s}(k) \tag{18}$$

$$\frac{\partial u(\boldsymbol{s}(k))}{\partial \boldsymbol{b}_{a1}(k)} = \boldsymbol{W}_a(k)\boldsymbol{\sigma}_a'(\boldsymbol{V}_a(k)\boldsymbol{s}(k)+\boldsymbol{b}_{a1}(k)) \tag{19}$$

其中　$\boldsymbol{\sigma}_a'(x) = \mathrm{e}^{-x}/(1+\mathrm{e}^{-x})^2$

式中　$\boldsymbol{\sigma}_a'(x)$——Sigmoid 函数导数

2.3.2　评价器网络参数优化

定义误差

$$\delta_d(\boldsymbol{s}(k)) = J_c(\boldsymbol{s}(k)) - r(\boldsymbol{s}(k),u(\boldsymbol{s}(k))) - \gamma J_c(\boldsymbol{s}(k+1)) \tag{20}$$

式中　$J_c(\boldsymbol{s}(k+1))$——评价器对状态 $\boldsymbol{s}(k+1)$ 估计的性能指标函数值

因此评价器网络的学习目标是最小化误差函数

$$E_c(\boldsymbol{s}(k)) = \frac{1}{2}\delta_d^2(\boldsymbol{s}(k)) \tag{21}$$

则评价器网络的权值和偏置更新规则如下

$$\boldsymbol{\Psi}_c(k+1)=\boldsymbol{\Psi}_c(k)-\alpha_c(k)\delta_d(k)\frac{\partial J_c(\boldsymbol{s}(k))}{\partial \boldsymbol{\Psi}_c(k)} \tag{22}$$

其中 $\boldsymbol{\Psi}_c(k)=(\boldsymbol{W}_c,\boldsymbol{b}_{c2},\boldsymbol{V}_c,\boldsymbol{b}_{c1})$

式中 $\alpha_c(k)$——评价器网络的学习率

$\boldsymbol{\Psi}_c(k)$——评价器网络的权值和偏置

由式(13)可得

$$\frac{\partial J_c(\boldsymbol{s}(k))}{\partial \boldsymbol{W}_c(k)}=\sigma_c(\boldsymbol{V}_c(k)\boldsymbol{s}(k)+\boldsymbol{b}_{c1}(k)) \tag{23}$$

$$\frac{\partial J_c(\boldsymbol{s}(k))}{\partial \boldsymbol{b}_{c2}(k)}=1 \tag{24}$$

$$\frac{\partial J_c(\boldsymbol{s}(k))}{\partial \boldsymbol{V}_c(k)}=\boldsymbol{W}_c(k)\sigma_c'(\boldsymbol{V}_c(k)\boldsymbol{s}(k)+\boldsymbol{b}_{c1}(k))\boldsymbol{s}(k) \tag{25}$$

$$\frac{\partial J_c(\boldsymbol{s}(k))}{\partial \boldsymbol{b}_{c1}(k)}=\boldsymbol{W}_c(k)\sigma_c'(\boldsymbol{V}_c(k)\boldsymbol{s}(k)+\boldsymbol{b}_{c1}(k)) \tag{26}$$

其中 $\sigma_c'(x)=\mathrm{e}^{-x}/(1+\mathrm{e}^{-x})^2$

式中 $\sigma_c'(x)$——Sigmoid 函数导数

3 仿真分析

为了验证本文提出算法的有效性，在 Matlab/Simulink 环境下，通过数值仿真验证该方法的性能。定义跟随轨迹点到参考路径的垂线距离为横向偏差，采用横向偏差平均值的绝对值和横向偏差的均方根值评价跟随性能。

仿真参数设置为：机器人履带中心距 0.485 m，驱动轮半径 0.2 m，控制周期 $T_s=0.2$ s。为了对比控制效果，同时设计了基于 LQR 和基于传统纯追踪算法的控制器，并且在不同速度下对 3 种控制器进行仿真。参考路径采用了组合线（直线和圆）和螺旋线。

为了观察误差收敛过程，将机器人的起点偏离轨迹起点一段距离，在跟随组合线时，组合线起点为(0 m，−20 m)，机器人的起点为(2 m，−20 m)。当速度 $v=0.3$ m/s 时，LQR 算法和纯追踪算法的预瞄距离为 2.5 m，HDP 算法的预瞄距离为 3 m。选取 LQR 算法的反馈矩阵 $\boldsymbol{K}=[-0.15\quad -0.02]^{\mathrm{T}}$，HDP 算法的参数为 $\boldsymbol{Q}=\mathrm{diag}\{4,1,0,0\}$，$R=1$，$\gamma=0.9$，学习步长为 0.2。基于纯追踪算法、HDP 算法以及 LQR 算法的路径跟随轨迹如图 4 所示，对应的横向偏差如图 5 所示。为准确评价其跟随误差，剔除初始误差的收敛过程，分析机器人进入稳态跟随时的横向偏差，结果如表 1 所示。

在螺旋线的参考路径中，螺旋线可表示为

$$\begin{cases}x=53-1.5t\cos t\\ y=-29-1.5t\sin t\end{cases}\quad(0<t<20) \tag{27}$$

螺旋线的起点为(53 m，−29 m)，也将机器人的起点偏离轨迹起点，设定为(53 m，−27 m)。基于纯追踪算法、HDP 算法以及 LQR 算法的路径跟随轨迹如图 6 所示，对应的横向偏差如图 7 所示。表 2 为剔

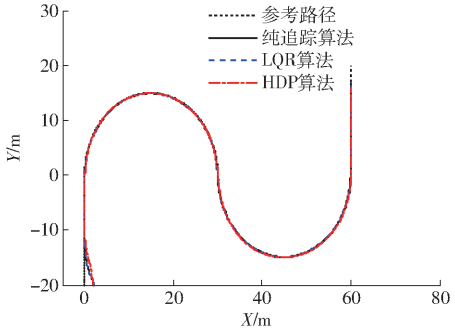


图 4 速度为 0.3 m/s 时不同控制器跟随组合线的轨迹

Fig. 4 Trajectory of robot using different controllers on combination line at speed of 0.3 m/s

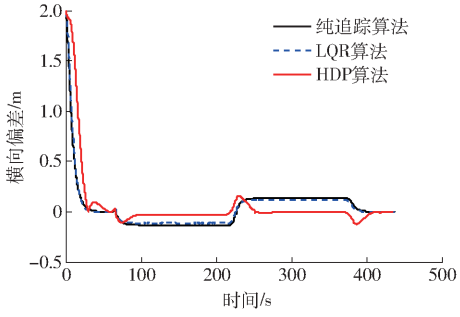


图 5 速度为 0.3 m/s 时不同控制器跟随组合线的横向偏差

Fig. 5 Lateral error of different controllers on combination line at speed of 0.3 m/s

表 1 速度为 0.3 m/s 时不同控制器跟随组合线的误差

Tab. 1 Tracking error of different controllers on combination line at speed of 0.3 m/s m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0	0	0.01
均方根误差	0.11	0.09	0.06

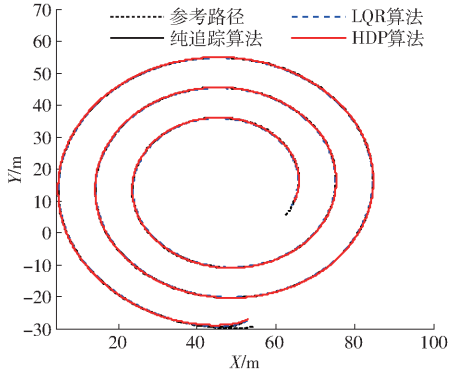


图 6 速度为 0.3 m/s 时不同控制器跟随螺旋线的轨迹

Fig. 6 Trajectory of robot using different controllers on helical line at speed of 0.3 m/s

除初始误差的收敛过程，进入稳态跟随状态后的误差统计结果。

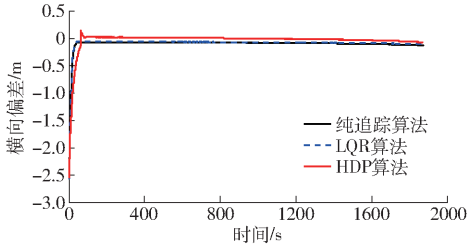


图 7　速度为 0.3 m/s 时不同控制器跟随螺旋线的横向偏差

Fig. 7　Lateral error of different controllers on helical line at speed of 0.3 m/s

表 2　速度为 0.3 m/s 时不同控制器跟随螺旋线的误差

Tab. 2　Tracking error of different controllers on helical line at speed of 0.3 m/s　m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0.02	0.03	0.01
均方根误差	0.09	0.07	0.02

为了验证移动机器人在不同速度下的跟随效果，改变机器人移动速度 $v = 1$ m/s 进行仿真。LQR 算法、纯追踪算法以及 HDP 算法的预瞄距离均为 3 m。选取 LQR 算法的反馈矩阵 $\boldsymbol{K} = [-0.3 \quad -0.005]^{\mathrm{T}}$，HDP 算法参数不变，纯追踪算法的预瞄距离不变，机器人起点不变。基于纯追踪算法、HDP 算法以及 LQR 算法的路径跟随轨迹如图 8 所示，对应的横向偏差如图 9 所示。表 3 为剔除初始误差的收敛过程，进入稳态跟随状态的误差统计结果。

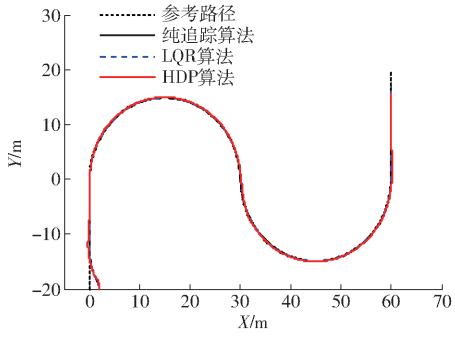


图 8　速度为 1 m/s 时不同控制器跟随组合线的轨迹

Fig. 8　Trajectory of robot using different controllers on combination line at speed of 1 m/s

在跟随螺旋线时，机器人起点不变。基于纯追踪算法、HDP 算法以及 LQR 算法的路径跟随轨迹如图 10 所示，对应的横向偏差如图 11 所示。表 4 为剔除初始误差的收敛过程，进入稳态跟随状态的误差统计结果。

从仿真结果可以得出，3 种控制算法在 2 种跟随速度条件下，均能够实现对两种复杂线形的跟随控制，设置的初始误差均得到快速收敛。在跟随组合线时，3 种控制方法的稳态平均误差绝对值均在 0 附近，基于 HDP 控制方法的均方根误差比纯追踪算法和 LQR 算法要小。组合线曲率在直线与半圆线交接处、两条半圆曲线交接处均发生了变化，虽然基于 HDP 算法在此时的瞬态误差较大，但随即收敛到 0 附近，而纯追踪算法和 LQR 算法，在跟随半圆线

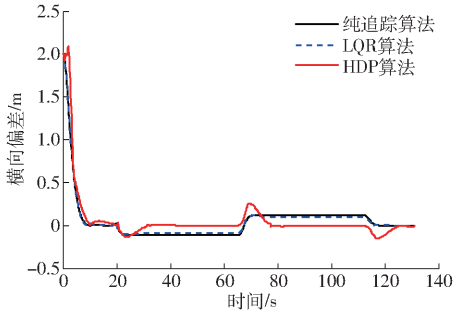


图 9　速度为 1 m/s 时不同控制器跟随组合线的横向偏差

Fig. 9　Lateral error of different controllers on combination line at speed of 1 m/s

表 3　速度为 1 m/s 时不同控制器跟随组合线的误差

Tab. 3　Tracking error of different controllers on combination line at speed of 1 m/s　m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0.01	0	0
均方根误差	0.12	0.10	0.06

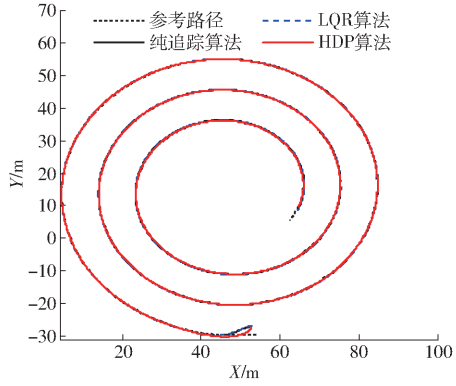


图 10　速度为 1 m/s 时不同控制器跟随螺旋线的轨迹

Fig. 10　Trajectory of robot using different controllers on helical line at speed of 1 m/s

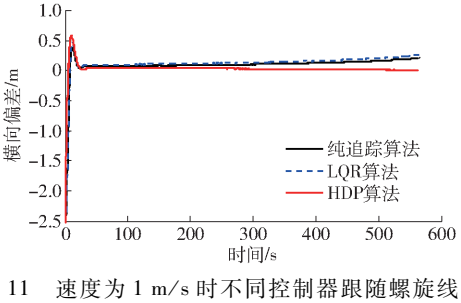


图 11　速度为 1 m/s 时不同控制器跟随螺旋线的横向偏差

Fig. 11　Lateral error of different controllers on helical line at speed of 1 m/s

表 4 速度为 1 m/s 时不同控制器跟随螺旋线的误差

Tab.4 Tracking error of different controllers on helical line at speed of 1 m/s m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0.11	0.14	0.03
均方根误差	0.04	0.04	0.04

时,误差基本不变。在跟随螺旋线时,可以看出随着螺旋线曲率的增大,3 种控制算法的横向偏差均有增大的趋势,但是基于 HDP 算法的平均误差绝对值依然比纯追踪算法和 LQR 算法小,基于 HDP 算法的均方根误差小于或等于基于纯追踪算法和 LQR 算法的均方根误差。上述现象说明,相对于其他两种方法,HDP 算法可以通过参数的在线学习,自主优化路径跟随性能,减小控制误差。

4 试验结果分析

4.1 路径跟随试验结果

本文使用的履带机器人试验平台如图 12 所示,其长度为 0.8 m,宽度为 0.57 m,质量约为 60 kg,履带中心距 B 为 0.485 m,驱动轮半径 r_d 为 0.2 m。采用 48 V、20 A·h 的锂电池供电,驱动电机功率为 400 W。传感器为北斗星通 BDM680,工作在差分定位模式,移动站与基准站之间的距离小于 5 km,位置类型状态为 Narrow_Int,按照产品手册,在该工作模式下定位误差为 0.01 m(RMS),通过双天线定位提供移动平台的航向角信息。计算平台采用工控机和基于 STM32 单片机的控制卡组成,工控机计算控制动作;单片机控制卡通过数学模型将控制动作分解为两侧履带转速,并给电机驱动器发送指令。



图 12 履带机器人

Fig. 12 Tracked robot

试验路径包括了直线、钝角转向曲线以及锐角转向曲线,试验地点为安徽农业大学农萃园。机器人运动速度 $v=0.3$ m/s,系统控制周期 $T_s=0.5$ s。为验证 HDP 算法的路径跟随性能,同时采用 LQR 算法和纯追踪算法作为比较。为了获得对不同线形的理想跟随效果,对 LQR 算法和纯追踪算法的控制参数作了针对性调试。为了验证 HDP 算法对不同线形的自适应能力,在跟随不同的参考路径时,不改变 HDP 算法的参数配置,其参数配置为 $\boldsymbol{Q}=\mathrm{diag}\{2,1,0,0\}$,$R=5$,$\gamma=0.9$,控制器的神经网络隐含层的神经元均为 5 个,学习步长均为 0.2。采用基于 LQR 算法获得的路径跟随数据训练评价器和执行器网络,获得初始网络权值。

在跟随直线时,3 种方法的预瞄距离均为 3 m,其中 LQR 算法的反馈矩阵 $\boldsymbol{K}=[-0.11\quad -0.015]^{\mathrm{T}}$。图 13 为 3 种方法跟随直线轨迹。图 14 为不同控制器跟随直线的横向偏差,其误差统计量如表 5 所示。

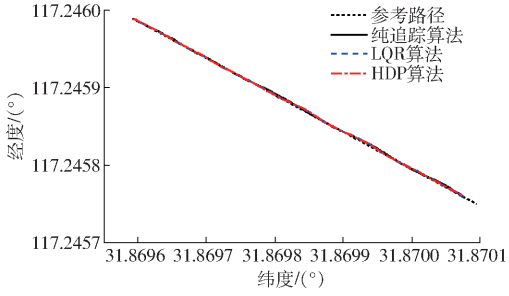


图 13 基于不同控制器跟随直线的轨迹

Fig. 13 Trajectory of robot using different controllers on straight line

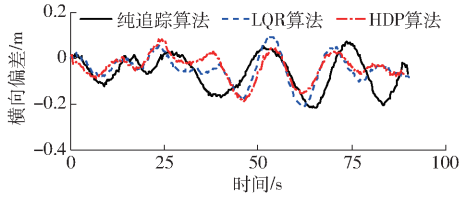


图 14 基于不同控制器跟随直线的横向偏差

Fig. 14 Lateral error of different controllers on straight line

表 5 不同控制器跟随直线试验的误差

Tab.5 Tracking error of different controllers on straight line m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0.06	0.05	0.04
均方根误差	0.08	0.06	0.06
最大误差绝对值	0.22	0.21	0.18

在跟随钝角转向曲线时,纯追踪算法的预瞄距离为 2.5 m,LQR 算法为 3 m,其中 LQR 算法的反馈矩阵 $\boldsymbol{K}=[-0.05\quad -0.055]^{\mathrm{T}}$。图 15 为 3 种方法跟随钝角转向曲线的轨迹,图 16 为不同控制器跟随钝角转向曲线的横向偏差,其统计量如表 6 所示。

在跟随锐角转向曲线时,预瞄距离均为 3 m,其中 LQR 算法的反馈矩阵 $\boldsymbol{K}=[-0.25\quad -0.075]^{\mathrm{T}}$。图 17 为 3 种方法跟随锐角转向曲线的轨迹,图 18 为不同控制器跟随锐角转向曲线的横向偏差,其统计量如表 7 所示。

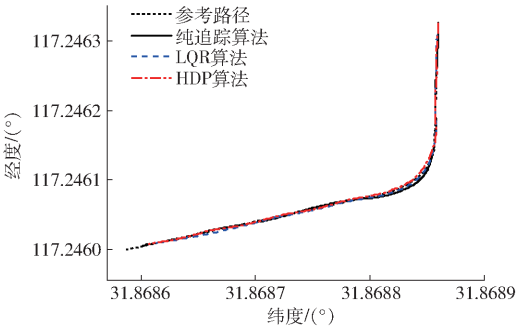


图 15　基于不同控制器跟随钝角转向曲线的轨迹

Fig. 15　Trajectory of robot using different controllers on obtuse angle steering curve

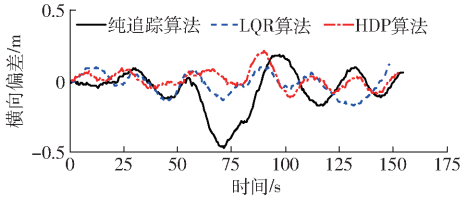


图 16　基于不同控制器跟随钝角转向曲线的横向偏差

Fig. 16　Lateral error of different controllers on obtuse angle steering curve

表 6　不同控制器跟随钝角转向曲线试验的误差

Tab. 6　Tracking error of different controllers on obtuse angle steering curve　m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0. 06	0. 02	0. 01
均方根误差	0. 15	0. 08	0. 06
最大误差绝对值	0. 47	0. 18	0. 21

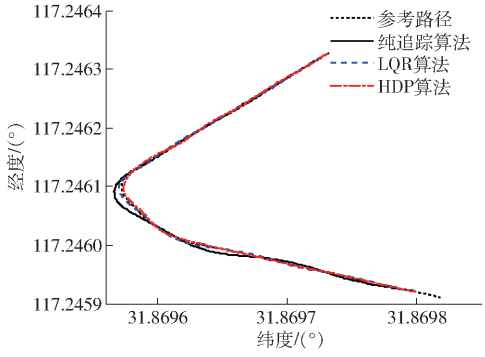


图 17　基于不同控制器跟随锐角转向曲线的轨迹

Fig. 17　Trajectory of robot using different controllers on acute angle steering curve

4.2　结果分析

总结上述试验结果可以发现，3 种方法均能够顺利跟随直线、钝角转向曲线以及锐角转向曲线。对比 LQR 和纯追踪算法，在跟随直线时，HDP 算法除均方根误差等于 LQR 算法的同类指标外，其他统计指标均为最小；在跟随钝角转向曲线时，HDP 算法除最大误差绝对值略大于 LQR 算法同类指标 0. 03 m，其余两项误差指标均最小；在跟随锐角转向曲线时，HDP 算法除平均误差绝对值略大于 LQR 算法同类指标 0. 01 m，其余两项误差指标均最小。为了完成 3 种曲线的路径跟随任务并取得较好的控制效果，试验过程中对纯追踪算法和 LQR 算法都进行了反复参数优化。而 HDP 算法仅通过历史跟随数据对评价器和执行器网络进行了离线训练，变换跟随路径时控制参数未做人工调试，仅通过 HDP 算法的参数优化机制对控制网络进行在线优化，即可取得较高精度的跟随结果，说明该方法具有较好的自主优化能力。

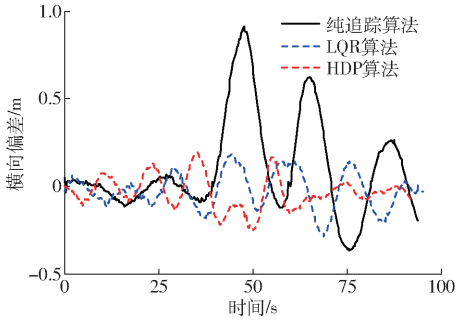


图 18　基于不同控制器跟随锐角转向曲线的横向偏差

Fig. 18　Lateral error of different controllers on acute angle steering curve

表 7　不同控制器跟随锐角转向曲线试验的误差

Tab. 7　Tracking error of different controllers on acute angle steering curve　m

评价参数	纯追踪算法	LQR 算法	HDP 算法
平均误差绝对值	0. 10	0. 02	0. 03
均方根误差	0. 27	0. 11	0. 09
最大误差绝对值	0. 90	0. 28	0. 24

5　结论

（1）针对移动机器人传统路径跟随控制算法环境适应能力弱、需要人工优化控制参数的问题，以履带式移动机器人为研究对象，提出了一种基于启发式动态规划的路径跟随控制方法，并试验验证了路径跟随性能，结果表明，跟随直线的均方根误差为 0. 06 m，跟随钝角转向曲线的均方根误差为 0. 06 m，跟随锐角转向曲线的均方根误差为 0. 09 m。该方法可以实现对多种复杂线形的路径跟随，且具有较高的控制精度。

（2）基于启发式动态规划算法框架，构建了综合跟随精度指标和稳定性指标的评价函数，通过对当前控制动作的评价性反馈信号优化控制器参数，仿真和试验结果表明，在跟随路径线形变化时，该方法不需要对控制参数进行调试，仅通过 HDP 算法的参数优化机制对网络权值进行在线优化就能够获得高精度的控制效果，具有较好的环境适应性和自主优化能力。

参 考 文 献

[1] RAJAMANI R. Vehicle dynamics and control[M]. Second Edition. Berlin: Springer Science & Business Media, 2011.

[2] 魏爽,李世超,张漫,等. 基于 GNSS 的农机自动导航路径搜索及转向控制[J]. 农业工程学报,2017,33(增刊 1):70 - 77.
WEI Shuang, LI Shichao, ZHANG Man, et al. Automatic navigation path search and turning control of agricultural machinery based on GNSS[J]. Transactions of the CSAE,2017,33(Supp. 1):70 - 77. (in Chinese)

[3] 孙怀江,杨静宇. 关于纯追踪模型的讨论[J]. 计算机学报,2002,25(12):1445 - 1449.
SUN Huaijiang, YANG Jingyu. On the pure pursuit model[J]. Chinese J. Computers, 2002, 25(12): 1445 - 1449. (in Chinese)

[4] 丁幼春,詹鹏,周雅文,等. 北斗定位田间信息采集平台运动控制器设计与试验[J]. 农业工程学报,2017,33(12):178 - 185.
DING Youchun, ZHAN Peng, ZHOU Yawen, et al. Design and experiment of motion controller for information collection platform in field with Beidou positioning[J]. Transactions of the CSAE,2017,33(12):178 - 185. (in Chinese)

[5] MARINO R, SCALZI S, NETTO M. Nested PID steering control for lane keeping in autonomous vehicles[J]. Control Engineering Practice, 2011,19(12): 1459 - 1467.

[6] 武星,沈伟良,楼佩煌,等. 视觉导引 AGV 鲁棒特征识别与精确路径跟踪研究[J/OL]. 农业机械学报,2016,47(7):48 - 56.
WU Xing, SHEN Weiliang, LOU Peihuang, et al. Robust feature recognition and precise path tracking for vision-guided AGV [J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2016,47(7):48 - 56. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20160708&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.07.008. (in Chinese)

[7] KAPANIA N R, GERDES J C. Design of a feedback-feedforward steering controller for accurate path tracking and stability at the limits of handling[J]. Vehicle System Dynamics, 2015,53(12): 1687 - 1704.

[8] 王家恩,陈无畏,王檀彬,等. 基于期望横摆角速度的视觉导航智能车辆横向控制[J]. 机械工程学报,2012,48(4): 108 - 115.
WANG Jiaen, CHEN Wuwei, WANG Tanbin, et al. Vision guided intelligent vehicle lateral control based on desired yaw rate [J]. Journal of Mechanical Engineering,2012, 48(4): 108 - 115. (in Chinese)

[9] WANG Rongrong, HU Chuan, YAN Fengjun, et al. Composite nonlinear feedback control for path following of four-wheel independently actuated autonomous ground vehicles[J]. IEEE Transactions on Intelligent Transportation Systems,2016,17(7): 2063 - 2074.

[10] YANG Lu, YUE Ming, MA Teng. Path following predictive control for autonomous vehicles subject to uncertain tire-ground adhesion and varied road curvature[J]. International Journal of Control, Automation and Systems,2019,17(1): 193 - 202.

[11] REN Yue, ZHENG Ling, KHAJEPOUR A. Integrated model predictive and torque vectoring control for path tracking of 4-wheel-driven autonomous vehicles[J]. IET Intelligent Transport Systems,2019,13(1):98 - 107.

[12] CLEMENTE E, MEZA-SANCHEZ M, BUGARIN E, et al. Adaptive behaviors in autonomous navigation with collision avoidance and bounded velocity of an omnidirectional mobile robot[J]. Journal of Intelligent & Robotic Systems,2018,92(2): 359 - 380.

[13] XIAO Hanzhen, LI Zhijun, YANG Chenguang, et al. Robust stabilization of a wheeled mobile robot using model predictive control based on neurodynamics optimization[J]. IEEE Transactions on Industrial Electronics,2017,64(1):505 - 516.

[14] 熊中刚,叶振环,贺娟,等. 基于免疫模糊 PID 的小型农业机械路径智能跟踪控制[J]. 机器人,2015,37(2):212 - 223.
XIONG Zhonggang, YE Zhenhuan, HE Juan, et al. Small agricultural machinery path intelligent tracking control based on fuzzy immune PID[J]. Robot,2015,37(2):212 - 223. (in Chinese)

[15] HWANG C L, YANG C C, HUNG J Y. Path tracking of an autonomous ground vehicle with different payloads by hierarchical improved fuzzy dynamic sliding-mode control[J]. IEEE Transactions on Fuzzy Systems,2018,26(2): 899 - 914.

[16] LU Xingyang, ZHANG Xiangyin, ZHANG Guoliang, et al. Neural network adaptive sliding mode control for omnidirectional vehicle with uncertainties[J]. ISA Transactions, 2019,86: 201 - 214.

[17] 唐小涛,陶建峰,李志腾,等. 自动导航插秧机路径跟踪系统稳定性模糊控制优化方法[J/OL]. 农业机械学报,2018,49(1):29 - 34.
TANG Xiaotao, TAO Jianfeng, LI Zhiteng, et al. Fuzzy control optimization method for stability of path tracking system of automatic transplanter[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(1):29 - 34. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20180103&flag=1. DOI:10.6041/j.issn.1000-1298.2018.01.003. (in Chinese)

[18] BEGNINI M, BERTOL D W, MARTINS N A. A robust adaptive fuzzy variable structure tracking control for the wheeled mobile robot: simulation and experimental results[J]. Control Engineering Practice,2017,64: 27 - 43.

[19] 李革,王宇,郭刘粉,等. 插秧机导航路径跟踪改进纯追踪算法[J/OL]. 农业机械学报,2018,49(5):21 - 26.
LI Ge, WANG Yu, GUO Liufen, et al. Improved pure pursuit algorithm for rice transplanter path tracking [J/OL].

Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 21 - 26. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20180502&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.05.002. (in Chinese)

[20] 孟宇,汪钰,顾青,等.基于预见位姿信息的铰接式车辆 LQR - GA 路径跟踪控制[J/OL].农业机械学报,2018,49(6):375 - 384.
MENG Yu, WANG Yu, GU Qing, et al. LQR - GA path tracking control of articulated vehicle based on predictive information [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6): 375 - 384. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20180645&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2018.06.045. (in Chinese)

[21] CARLUCHO L, PAULA M D, VILLAR S A, et al. Incremental Q-learning strategy for adaptive PID control of mobile robots [J]. Expert Systems with Applications, 2017, 80(1): 183 - 199.

[22] 张化光, 张欣, 罗艳红, 等. 自适应动态规划综述[J]. 自动化学报, 2013, 39(4): 303 - 311.
ZHANG Huaguang, ZHANG Xin, LUO Yanhong, et al. An overview of research on adaptive dynamic programming[J]. Acta Automatica Sinica, 2013, 39(4): 303 - 311. (in Chinese)

[23] 徐昕,沈栋,高岩青,等.基于马氏决策过程模型的动态系统学习控制:研究前沿与展望[J].自动化学报,2012,38(5):673 - 687.
XU Xin, SHEN Dong, GAO Yanqing, et al. Learning control of dynamical systems based on makov decision processes: research frontiers and outlooks[J]. Acta Automatica Sinica, 2012, 38(5): 673 - 687. (in Chinese)

[24] 薄迎春,乔俊飞.启发式动态规划在污水处理过程控制中的应用[J].控制理论与应用,2013,30(7):828 - 833.
BO Yingchun, QIAO Junfei. Application of heuristic dynamic programming to wastewater treatment process control[J]. Control Theory & Applications, 2013, 30(7): 828 - 833. (in Chinese)

[25] DABOONI S A, WUNSCH D. Heuristic dynamic programming for mobile robot path planning based on dyna approach[C]// International Joint Conference on Neural Networks (IJCNN), 2016: 3723 - 3730.

[26] LIN Weisong, YANG PingChieh. Adaptive critic motion control design of autonomous wheeled mobile robot by dual heuristic programming[J]. Automatica, 2008, 44(11): 2716 - 2723.

[27] HUANG Zhenhua, LIAN Chuanqiang, XU Xin, et al. Lateral control for autonomous land vehicles via dual heuristic programming[J]. International Journal of Robotics and Automation, 2016, 31(6): 539 - 547.

[28] OSINENKO P, GÖHRT T, DEVADZE G, et al. Stacked adaptive dynamic programming with unknown system model[C]// IFAC-Papers OnLine, 2017, 50(1): 4150 - 4155.