

基于双启发动态规划的预分解窑控制器设计

杨宝生¹, 马修水²

(1. 宿州学院 智能信息处理实验室, 安徽 宿州 234000; 2. 浙江大学 宁波理工学院, 浙江 宁波 315100)

(bsyang@ymail.com)

摘要:针对水泥预分解窑熟料煅烧过程多变量、多扰动、非线性,难以建立精确的数学模型,实际生产中对工人的生产经验依赖性较强等问题,提出采用误差反向传播(BP)神经网络建立烧成系统数学模型,设计预分解窑双启发动态规划(DHP)控制器。DHP评价网络输出代价函数 J 关于状态量的偏导数,获得最优或者次优的控制信号,然后由动作网络输出控制信号,使系统得到期望的控制轨迹。仿真结果表明控制器响应时间较快,各参量超调量均不大,有助于实际系统的稳定运行。

关键词:双启发动态规划;神经网络;煅烧工段;控制器

中图分类号: TP273.21 **文献标志码:** A

Design of cement calcination process controller based on dual heuristic programming algorithm

YANG Bao-sheng¹, MA Xiu-shui²

(1. Laboratory of Intelligent Information Processing, Suzhou University, Suzhou Anhui 234000, China;

2. Ningbo Institute of Technology, Zhejiang University, Ningbo Zhejiang 315100, China)

Abstract: For the multiple variables, disturbances, nonlinearity and other properties of the cement clinker kiln process, it is very difficult to establish an accurate model of the cement kiln system. It is strongly dependent on the experience of workers in the actual production. The error Back-Propagation (BP) neural network was used to establish the firing system model, and a controller was designed for the kiln based on Dual Heuristic Programming (DHP). DHP critic network output the partial derivative of cost function J with the state to obtain the optimal or sub-optimal control signal. The action network output control actions to control the system to achieve desired trajectory. The simulation results show that the controller has faster response time and less overshoot. These features contribute to the stable operation of the real system.

Key words: Dual Heuristic Programming (DHP); neural network; clinker calcination process; controller

0 引言

水泥熟料煅烧过程是一个关于传质、传热和物理化学反应的复杂的多变量、多扰动对象的过程,涉及到机械、电器、控制、压力、温度、流量和成分等多种变量。各种参数的影响都是不相同的,这些参数之间还会产生相互影响;并且,这些参数都处于不断变化之中,是否能够随时掌握它们的具体数据并且随时传送到指定位置显得非常重要^[1-2]。人工神经网络被较早引入到水泥生产过程控制或预测,均有良好表现。Lin等人首先尝试将自适应动态规划(Adaptive Dynamic Programming, ADP)方法应用到分解炉温度控制,达到了其预期的稳定控制目标^[3];刘唐波应用启发式动态规划(Heuristic Dynamic Programming, HDP)和执行依赖启发式动态规划(Action-dependent HDP, ADHDP)方法来控制回转窑烧成带温度,仿真控制结果较为理想^[4]。分析发现,这些研究均是对烧成系统个别设备进行的局部研究,对于水泥熟料最终煅烧质量还有待分析。熟料生产稳定性直接影响水泥质量的好坏。本文采用启发式动态规划控制算法,针对预分解窑系统最大的非线性环节煅烧工段做整体优化控制,探索提高我国预分解窑生产效率的方法。

1 水泥预分解窑系统 BP 神经网络建模

水泥预分解窑系统主要由预热器、分解炉、回转窑和篦冷机四部分组成。图1是经过主元分析的预分解窑系统的简化原理图。以带有五级旋风预热器的预分解窑为例,整个工艺过程分成两个方向:一个是从上而下的物料流向;另一个是从下而上的烟气流向。灰色箭头表示物料流向,虚线表示气体流向,黑色表示煤粉喷入燃烧流向。实际生产中,这几个部分的位置关系是上下或倾斜放置,以使物料经过所有环节后被煅烧成水泥熟料,从篦冷机卸出。

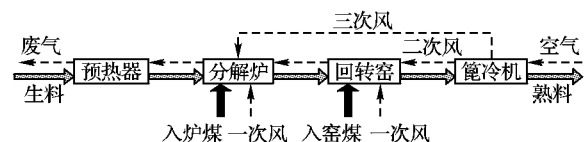


图1 预分解窑系统简图

对现场采集到的数据进行分析,然后对预分解窑系统进行简化,获知影响系统的主要因素就是风、煤、料。因此,控制量的选取就集中在对风、煤、料的控制上。分解炉出口是系统的一个交汇点,它是系统多个状态综合指标的反映。因此,选取分解炉出口温度作为其中一个状态量。排出废气中的氧含

收稿日期:2011-03-07;修回日期:2011-05-07。 基金项目:安徽省自然科学基金资助项目(10040606Q64);安徽省教育厅高校省级科学研究基金资助项目(KJ2010B224);宿州学院智能信息处理实验室开放课题资助项目(2011YKF11)。

作者简介:杨宝生(1981-),男,安徽宿州人,讲师,硕士,主要研究方向:智能优化控制; 马修水(1963-),男,安徽庐江人,教授,博士生导师,博士,主要研究方向:智能检测。

量能够表示燃烧程度。氧含量不足说明燃烧不充分;氧含量过大表示通风量过大,空气流量大会加大热损失。适度的氧含量既能保证燃烧充分又避免过多热损耗。因此,本文选取一级旋风筒出口废气氧含量作为本文的另一个状态量。将二者稳定控制在一个合理的范围之内是所要实现的目标。鉴于前述预分解窑系统的复杂性,难以用数学模型来实现,本文采用人工神经网络(Artificial Neural Network, ANN)对预分解窑系统进行建模。基于BP神经网络的模型能够较好地描述该非线性系统的主要性能,并且具有一定泛化能力^[5]。选择控制量 $u_i(t)$, $i = 1, 2, \dots, 5$,表示系统操作量,即生料量、分解炉喂煤量、回转窑喂煤量、回转窑转速,旋风筒C1出口负压。表1是本文从实际生产线采集到的用来训练模型的数据(采样时间间隔为2 min),数据单位t/d表示吨/天,t/h表示吨/小时,r/min表示转/分钟,kPa表示千帕。由于采集的各数据单位不一致,并且在数值上差距较大,为避免神经网络隐层权值调整过程中大数据对小数据的淹没现象,因而须对数据进行 $[-1, 1]$ 归一化处理。同时,归一化也是为了加快训练网络的收敛性。相反地,神经控制器输出的控制信号需要进行反归一化处理,使其恢复本来的量纲与单位。

表1 5000 t/d 预分解窑生产线采样数据(部分)

时间步	$u_1/(t \cdot h^{-1})$	$u_2/(t \cdot h^{-1})$	$u_3/(t \cdot h^{-1})$	$u_4/(r \cdot min^{-1})$	u_5/kPa
t	444.4	19.0	15.1	3.7	-5.3
$t+1$	452.1	17.9	15.1	3.8	-5.3
$t+2$	453.4	19.7	15.1	3.7	-4.9
$t+3$	450.3	19.7	15.1	3.8	-4.9
$t+4$	426.2	19.7	15.6	3.7	-5.0
$t+5$	457.4	19.8	15.6	3.8	-5.0
$t+6$	452.9	19.4	15.0	3.7	-4.7
$t+7$	384.1	19.8	15.0	3.7	-5.2
$t+8$	426.3	19.7	14.6	3.7	-5.4

2 预分解窑 DHP 控制器设计

动态规划(Dynamic Programming, DP)是解决多阶段决策过程最优化问题的一种常用方法,它把比较复杂的问题划分成若干个阶段,通过逐段求解,最终求得全局最优解。它的核心是贝尔曼最优性原理,既可以用来求解约束条件下的函数极值问题,也可用于求解约束条件下的泛函极值问题,尤其对解决线性时间离散系统二次型性能指标最优控制问题特别有效。经典的动态规划算法尽管是解决最优化问题的一个有力工具,但在实施过程中会遇到“维数灾”问题^[6],即代价函数的计算量随状态变量成指数增长,所以一般只限于解决小规模问题,从而限制了其应用范围。为了避免“维数灾”问题,一般采用近似的方法来计算代价函数,自适应动态规划就是在这样的背景下产生的。

DHP算法是高级形态的自适应动态规划方法,它用神经网络来满足修正的Bellman方程^[7],主要包括评价网络(Critic)、动作网络(Action)和模型网络(Model)三个模块。图2是本文所设计的预分解窑DHP控制系统,其中Action作为控制器输出控制信号,Model为预分解窑系统模型,Critic用于评价控制效果,指导控制信号修正调整。它的评价网络输出代价函数 J 关于状态量的偏导数,所以具有更高的精确度,与之相应的是其计算量相应要复杂得多。它的目的是当最小化cost-to-go函数 $J(t)$ 时,发现最优或者次优的控制信号

$u(t)$,以使离散时间形式的非线性系统 $x(k+1) = f[x(k), u(k), k]$ 到期望的控制轨迹^[8]。

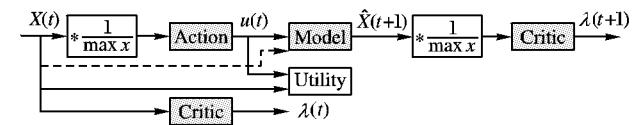


图2 DHP方法结构

图2中 $X(t)$ (n 维)是对象在 t 时刻的状态。控制信号 $u(t)$ (m 维)通过将 $X(t)$ 输入Action产生。控制信号 $u(t)$ 接着被应用到对象且对象演变到状态 $X(t+1)$,见式(1)。Critic的作用是帮助设计一个控制器(Action),Critic估计 $J(t)$ (总的效用函数)关于 $X(t)$ 的梯度; λ 表示这样一个梯度的缩写。 $*1/\max x$ 模块将 $X(t)$ 按比例缩放到 n 维状态空间 \mathbb{R}^n 的 $[-1, +1]^n$ 区间内,即对输入参数进行归一化处理。

$$X(t+1) = f_x(X(t), u(t)) \quad (1)$$

对于一个无限时域问题,通过构想一个主要效用函数 $U(t)$ 的方法,效用函数定义为 $U(t) = f_u(X(t), x(t))$,体现出对于控制的系统中一个或多个可测变量目标性。评价函数是用来估计总的代价函数的值,在状态空间任何容易取得的点。根据这样的假设,评价是用来精确估计通过动作函数参数值所指定策略总的代价,评价函数的梯度可以用来调整策略参数,以便达成一个局部最优的参数策略。该过程使用人工神经网络实现控制和评价函数,以及使用模糊系统作为控制器^[9]。总的效用函数,即代价函数 $J(t)$,式(2)给出在时间步 k 时的 $J(t)$:

$$J(t) = \sum_{k=0}^{\infty} \gamma^k U(t+k) \quad (2)$$

其中: γ 是折扣因子, $0 < \gamma < 1$; $U(k)$ 是效用函数,其允许设计者来衡量当前与将来效用的相关性。对式(2)进一步推导可以得到一个有用的特性,即式(3),该式表明了 t 与其下一时刻代价函数之间的关系:

$$J(t) = U(t) + \gamma J(t+1) \quad (3)$$

该方法利用两个不同的训练循环:一个为了控制策略;一个为了评价估计。调整被控系统控制策略是用来最优化总的效用函数 $J(t)$ 。因为其控制动作输出 $u(t)$,控制器训练是基于梯度的学习算法,要求估计导数 $\frac{\partial J(t)}{\partial u_i(t)}$ 。训练评价函数是基于其估计值的一致性,在整个时间域上使用式(3)来判断。对于DHP方法,评价网络估计 $J(t)$ 关于系统状态的导数,也就是 $\lambda_i(t) = \frac{\partial J(t)}{\partial X_i(t)}$,通过对式(3)两边微分有:

$$\frac{\partial}{\partial X_i(t)} J(t) = \frac{\partial}{\partial X_i(t)} (U(t) + \gamma J(t+1)) \quad (4)$$

得到用于评价训练的式(5):

$$\lambda_i(t) = \frac{\partial U(t)}{\partial X_i(t)} + \frac{\partial u_j(t)}{\partial X_i(t)} \frac{\partial U(t)}{\partial u_j(t)} + \gamma \lambda_k(t+1) \cdot \left[\frac{\partial X_k(t+1)}{\partial X_i(t)} + \frac{\partial u_i(t)}{\partial X_i(t)} \frac{\partial X_k(t+1)}{\partial u_m(t)} \right] \quad (5)$$

计算方程的右边需要一个系统动态模型,其包含来自对象控制器系统的一对Jacobian矩阵,也就是 $\frac{\partial X_k(t+1)}{\partial X_i(t)}$ 和 $\frac{\partial X_k(t)}{\partial u_i(t)}$ 。

控制策略使用链式规则,系统模型被更新到转换评价输出,得到 $\frac{\partial J(t)}{\partial u_i(t)}$ 的估计值:

$$\frac{\partial J(t)}{\partial u_i(t)} = \frac{\partial U(t)}{\partial u_i(t)} + \gamma \sum \lambda_i(t+1) \frac{\partial X_i(t+1)}{\partial u_i(t)} \quad (6)$$

整个过程可描述为一个同步最优化问题:基于 $J(t)$ 的估计值的梯度的评价函数估计值的最优化,从评价获得的基于梯度的控制策略参数最优化。通过使用不同的策略来使得这两个最优化得到收敛。一个可行的方法是轮流进行,即最优化评价估计值和控制策略的最优化交替,也可以同时做两个过程的最优化。

动作网络是代表产生控制信号。本系统中使用的动作模块是一个三层前馈神经网络加上一个固定的比例缩放模块。神经网络有5个输入,一个带有10个 symmetrical sigmoidal 神经元的单隐层和2个 symmetrical sigmoidal 输出神经元,分别为 $u_i(t)$, $i=1,2,\dots,5$,表示系统控制量,即生料量、分解炉喂煤量、回转窑喂煤量、回转窑转速和旋风筒 C1 出口负压; $x_j(t)$, $j=1,2$,分别表示分解炉出口温度和旋风筒 C1 出口废气氧含量。模型网络是一个预先训练或者在线训练的神经网络,用来近似非线性系统的动态特性。本系统模型模块使用一个三层前馈神经网络,7个输入,采用系统 t 时刻的控制量与被控量。本系统中评价模块使用 2-10-5 结构。

3 DHP 控制器训练

3.1 效用函数定义与各参数范围设定

在训练过程中,为了获得更快的收敛速度,效用函数定义采用二次型形式(7):

$$U = \sum \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} \quad (7)$$

其中: \mathbf{Q} 、 \mathbf{R} 为分别为 n 维和 m 维单位矩阵, \mathbf{x} 为所要达到状态变量, \mathbf{u} 为控制器输出的操作变量。控制器训练过程中预先设定状态量变化范围,见式(8),系统运行过程中通过此范围将认为控制失效。

$$(|x_1(t) - 890| < 40) \cup (|x_2(t) - 3| < 2) \quad (8)$$

3.2 动作和评价网络训练策略

DHP 控制器的动作和评价网络训练策略采用单阶段训练过程,所谓单阶段也就是动作和评价网络同时调整。与之对应的双阶段,是动作和评价网络的调整分别进行^[10]。单阶段,动作网络和评价网络的并发训练步骤如下:

- 步骤1 按比例缩放 $X(t)$, 将它输入到动作网络, 获得 $u(t)$;
- 步骤2 将 $u(t)$ 输入到被控对象, 获得下一时刻状态 $X(t+1)$;
- 步骤3 按比例缩放 $X(t+1)$, 将它输入到评价网络, 获得 $\lambda(t+1)$;
- 步骤4 计算评价网络的期望输出 $\lambda^d(t)$;
- 步骤5 计算动作网络的权值变化;
- 步骤6 执行动作网络的权值变化;
- 步骤7 按比例缩放 $X(t)$, 将它输入到评价网络;
- 步骤8 计算/执行评价网络权值变化;
- 步骤9 增加 t 并回到步骤1。

训练好的动作网络即可以作为实际生产中预分解窑控制器,用来输出控制信号。

4 DHP 控制器输出结果

通过仿真实验,该控制器输出的控制信号如图3、4所示,在该控制器的控制下,系统状态指标始终处在3.1节所要求

的范围内。图3是分解炉喂煤量、回转窑喂煤量调整变化曲线,图4是生料量、回转窑转速、旋风筒 C1 出口负压调整变化曲线,每一个时间步为采样数据的时间间隔。各控制量初期波动较大,而后伴随被控量的稳定缓慢变化。从控制结果可以看出系统响应时间较快,除废气氧含量外其余各参量超调量均不大,这有助于实际系统的稳定运行,可见 DHP 控制器表现出较好的控制特性。

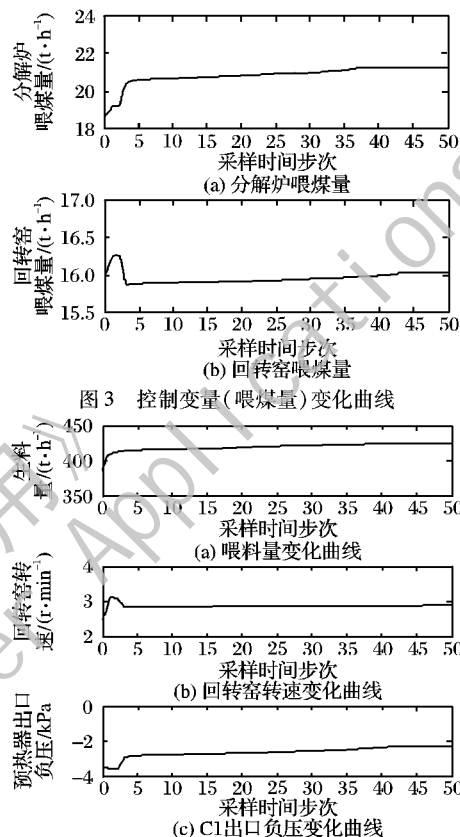
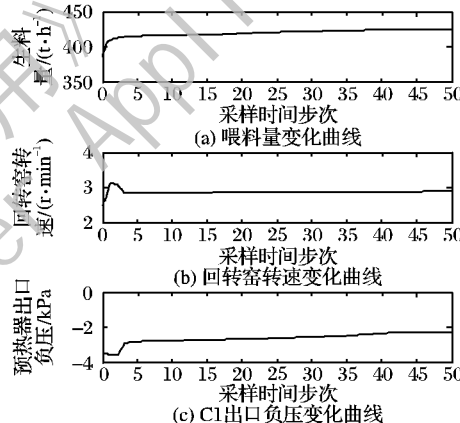


图3 控制变量(喂煤量)变化曲线



5 结语

本文将 DHP 算法应用到预分解窑系统多变量控制。预训练的模型网络被用来近似迭代被控对象。在训练过程中,可以施加统一的随机噪声到被控系统,来改进 DHP 控制器的表现和鲁棒性。一旦评价和动作网络权值已经收敛,训练停止,把动作网络作为被控系统的 DHP 神经网络控制器。带有预训练的模型网络的 DHP 控制器能够有效减轻被控系统扰动和显示出较好的控制特性。进一步地,DHP 控制器对于带有无模型的、不确定的非线性系统的鲁棒特性也有很好表现。分析可以看出,用 DHP 方法处理预分解窑系统多变量控制是一个有效的途径,显示出实时最优控制的潜力。

参考文献:

- [1] 赵应武, 过伦祥, 张先成, 等. 预分解窑水泥生产技术与操作 [M]. 北京: 中国建材工业出版社, 2004: 71-92.
- [2] 韩文. 水泥生产过程仿真培训平台开发及分解炉温度控制系统研究 [D]. 济南: 济南大学, 2007.
- [3] LIN XIAOFENG, ZHANG ZHIGANG, LIU DERONG. Temperature control in precalcinator with dual heuristic dynamic programming [C]// IJCNN 2007: Proceedings of the International Joint Conference on Neural Networks. Piscataway, NJ: IEEE Press, 2007: 344-349.

由图 7~9 可见,在 GPS/SINS 紧耦合系统框架下,新息检测法和新息外推法在 3 颗星时仍能检测卫星故障,这是利用 GPS/SINS 紧耦合系统进行完好性监测的优势所在。根据图 8、9,新息检测法检验统计量在 556 s 时达到 34.961,新息外推法检验统计量在 529 s 达到 28.543,因此新息外推法在 3 颗星时也比新息检测法更快地检测到故障。

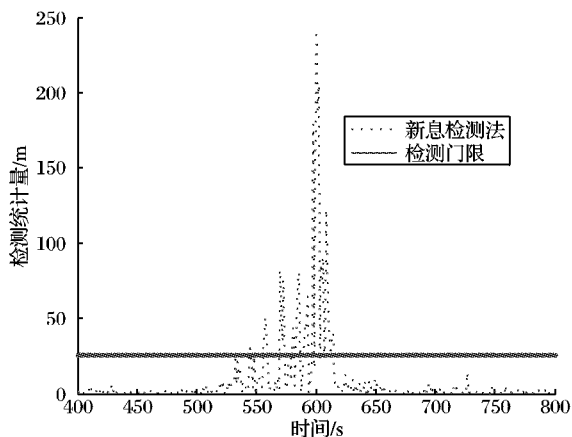


图 8 2 m/s 故障时新息检测法检测结果

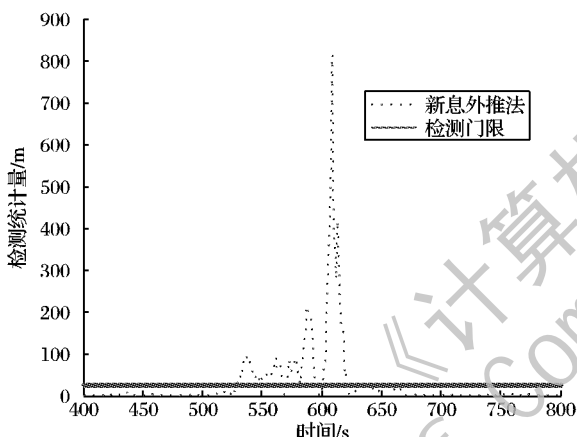


图 9 2 m/s 故障时新息外推法检测结果

通过上述仿真分析,新息外推法的特点总结如下:如果递推过程系统未发生故障,则新息外推法检验统计量为递推周期内所有新息及其方差阵的加权平均求得,当系统存在野值时,新息检测法的检验统计量将超出检测门限而报警,而新息外推法将不反映故障,因而改进的新息检测法有屏蔽野值的功能;如果递推过程中系统发生故障或递推起始时刻即发生

故障,则新息外推法将有放大故障的功能,其检验统计量将比新息检测法的检验统计量大数倍,因而更容易检测到故障,新息外推法的这种特点尤其适合对慢变故障的检测。

4 结语

在 GPS/SINS 紧耦合系统框架下进行故障检测,即使可见卫星数少于 4 颗时完好性算法仍可用。基于 GPS/SINS 紧耦合系统的新息外推法,不仅可以缩短检测故障(尤其是慢变故障)的时间,而且在一定程度上具有辨识野值和故障的功能。

参考文献:

- [1] 秦永元,张洪钺,汪叔华.卡尔曼滤波与组合导航原理[M].西安:西北工业大学出版社,1998.
- [2] 刘建业,曾庆化,赵伟.导航系统理论与应用[M].西安:西北工业大学出版社,2010.
- [3] BABU R, WANG JINLING, RAO G. Analysis of ultra-tight GPS/INS integrated system for navigation performance [C]// ICSCN'08: Proceedings of the 2008 IEEE International Conference on Signal Processing, Communications and Networking. Piscataway, NJ: IEEE Press, 2008: 234-237.
- [4] McDONALD J, KENDRICK J. Benefits of tightly coupled GPS/IRS for RNP operations in terrain challenged airports [C]// Proceedings of the 2008 IEEE/ION Position Location and Navigation Symposium. Washington, DC: IEEE Computer Society, 2008: 294-303.
- [5] 孙学磊,赵忠,夏家和.基于滤波残差的组合导航系统故障检测研究[J].计测技术,2009,29(2):5-7.
- [6] 白宇,黄智刚,李锐.残差检测法和状态多步外推法的 GI-RAIM 性能比较[J].航空电子技术,2008,39(4):1-4.
- [7] BHATTI U I. Improved integrity algorithms for integrated GPS/INS systems in the presence of slowly growing errors [D]. London, UK: Civil and Environmental Engineering Imperial College, 2007.
- [8] LUPASH L O. Horizontal / vertical exclusion level determination scheme for RAIM fault detection and exclusion implementation: USA, US6847893 [P]. 2005-11-25.
- [9] RTCA DO-229D. Minimum operational performance standards for global positioning system/wide area augmentation system airborne equipment [S], 2001.
- [10] 刘承香,孙枫,陈小刚,等.野值存在情况下组合导航系统的容错技术研究[J].中国惯性技术学报,2002,10(6):12-17.

(上接第 2288 页)

- [4] 刘唐波.基于自适应评价方法的水泥回转窑神经控制器[D].南宁:广西大学电气工程学院,2008.
- [5] 郝晓弘,段晓燕,李恒杰.基于 BP 神经网络的迭代学习初始控制策略研究[J].计算机应用,2009,29(4):1025-1027.
- [6] LIAO XIAO-FENG, CHEN GUAN-RONG, SANCHEZ E N. Delay-dependent exponential stability analysis of delayed neural networks: An LMI approach [J]. Neural Networks, 2002, 15(7): 855-866.
- [7] 李国勇,张翠平,郭红戈,等.最优控制理论及参数最优化[M].北京:国防工业出版社,2006:203-234.
- [8] BAI XUERUI, YI JIANQIANG, ZHAO DONGBIN. Approximate

- dynamic programming for ship course control [C]// ISNN'07: Proceedings of the 4th International Symposium on Neural Networks: Advances in Neural Networks, LNCS 4491. Berlin: Springer-Verlag, 2007: 349-357.
- [9] SI J, WANG Y T. On-line learning control by association and reinforcement [J]. IEEE Transactions on Neural Networks, 2001, 12(2): 264-276.
- [10] YANG BAOSHENG, CAO DEGUANG. Action-dependent adaptive critic design based neurocontroller for cement precalciner kiln [J]. International Journal of Computer Network and Information Security, 2009, 1(1): 62-68.