

# 基于自适应动态规划的矿渣微粉生产过程跟踪控制

王康<sup>1</sup> 李晓理<sup>2</sup> 贾超<sup>1</sup> 宋桂芝<sup>3</sup>

**摘要** 矿渣微粉是一种新型绿色环保型建材,可以大大提高水泥混凝土的力学性能.本文以矿渣微粉生产过程为研究对象,针对该过程难以通过机理建模进行辨识和控制的特点,利用数据驱动的思想,建立矿渣微粉生产过程的递归神经网络模型.在此基础上,利用自适应动态规划,设计具有控制约束的跟踪控制器,并将其应用到矿渣微粉生产过程中.仿真分析表明,建立的数据驱动模型能够有效地辨识矿渣微粉生产过程,同时,本文提出的控制方法能够实现输入受限的微粉比表面积及磨内压差的最优跟踪控制.

**关键词** 矿渣微粉, 数据驱动, 自适应动态规划, 最优跟踪控制, 输入有界

**引用格式** 王康, 李晓理, 贾超, 宋桂芝. 基于自适应动态规划的矿渣微粉生产过程跟踪控制. 自动化学报, 2016, 42(10): 1542–1551

**DOI** 10.16383/j.aas.2016.c150808

## Optimal Tracking Control for Slag Grinding Process Based on Adaptive Dynamic Programming

WANG Kang<sup>1</sup> LI Xiao-Li<sup>2</sup> JIA Chao<sup>1</sup> SONG Gui-Zhi<sup>3</sup>

**Abstract** Super fine slag powder is a new kind of green environmental-friendly construction material, which can greatly improve the mechanical properties of cement concrete. However, the slag powder grinding process is hard to identify by a mechanism model. In this paper, a data-driven based recurrent neural network model is constructed utilizing the information measured from slag grinding system. Based on this model, an adaptive dynamic programming algorithm is proposed to realize the optimal tracking control with constrained control input. Further, this algorithm is applied to the slag grinding process. Simulation examples show that the data-based model can effectively identify the grinding process, and the control method can realize the optimal tracking control of specific surface area and mill differential pressure with control constraints.

**Key words** Slag grinding process, data driven, adaptive dynamic programming, optimal tracking control, input constrained

**Citation** Wang Kang, Li Xiao-Li, Jia Chao, Song Gui-Zhi. Optimal tracking control for slag grinding process based on adaptive dynamic programming. *Acta Automatica Sinica*, 2016, 42(10): 1542–1551

矿渣粉磨生产是一个复杂多变量、强耦合的非

线性过程,实质是将废弃矿渣经过粉磨,及时选出物料中达到一定粒径要求的微粉.高炉矿渣废弃物经过粉磨,比表面积在 400 m<sup>2</sup>/kg 以上的细粉称为矿渣微粉.矿渣微粉可以部分代替水泥,以矿物参合料的形式参与混凝土配置,大大提高水泥混凝土的力学性能<sup>[1–2]</sup>.近年来,建筑业对高质量的矿渣微粉需求十分紧迫,然而在实际生产过程中,高效率地生产超细矿渣微粉还面临许多控制上的难题.例如,微粉的比表面积往往会受到风速、电机转速、压力、温度等多方面因素的影响,矿渣粉磨系统的复杂性造成了对其进行有效控制十分困难.实际生产过程中,往往是凭借工人经验、长期监控设备状态和及时调整控制参数实现生产控制.然而由于工人操作水平参差不齐,往往造成产品质量波动较大.如何建立微粉生产过程的有效模型,并在此基础上实现生产过程的高效自动控制,是一个极为重要的研究方向.

矿渣粉磨是一个多变量、强耦合、非线性的生产

收稿日期 2015-11-30 录用日期 2016-03-02  
Manuscript received November 30, 2015; accepted March 2, 2016

国家自然科学基金(61473034, 61673053), 高等学校博士学科点专项科研基金(20130006110008), 北京工业大学内涵发展—引进人才科研启动经费, 北京科技新星计划跨学科合作项目资助

Supported by National Natural Science Foundation of China (61473034, 61673053), Specialized Research Fund for the Doctoral Program of Higher Education (20130006110008), Beijing University of Technology Start-up Funding of Content Development and the Introduced Talent Research, and Beijing Nova Programme Interdisciplinary Cooperation Project

本文责任编辑 魏庆来  
Recommended by Associate Editor WEI Qing-Lai

1. 北京科技大学自动化学院 北京 100083 2. 北京工业大学电子信息与控制工程学院 北京 100124 3. 济南鲁新新型建材股份有限公司 济南 250109

1. School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083 2. College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124 3. Jinan Luxin Materials Company Limited, Jinan 250109

过程, 并且立磨内部存在复杂的物理化学变化, 难以建立整个微粉生产过程的准确机理模型. 陈远<sup>[3]</sup>对立磨中微粉颗粒进行了流体力学分析, 并讨论了产品细度与各单一变量之间的关系, 对揭示立磨内部现象和粉磨机理起到了重要作用, 但由于产品细度受到多变量的影响, 难以建立生产过程的完整机理模型. 另一方面, 立磨粉磨过程中, 具有丰富的离线和在线数据, 例如主机电流、磨机压差、选粉机转速、入磨温度、比表面积等, 这些数据仅被工人记录并作为检修参考数据, 并没有得到很好的利用和挖掘.

数据驱动控制不依靠生产过程的系统模型信息, 仅利用在线和离线数据便能实现输出变量和过程可测变量之间的建模与控制<sup>[4-6]</sup>. 对于水泥生产过程中的立磨粉磨系统, 数据驱动控制正在被广泛研究, 试图对复杂的粉磨系统进行准确辨识, 以实现高效控制. 针对料层厚度与喂料量、磨机出口温度与热风阀开度之间的关系, 颜文俊等<sup>[7]</sup>采用最小二乘法建立了立磨控制回路模型, 并采用预测控制方法实现了优化控制. Cai 等<sup>[8]</sup>对水泥粉磨系统采用最小二乘支持向量机方法, 建立了料层厚度的软测量模型, 实现了料层厚度的间接测量. Lin 等<sup>[9]</sup>采用小波神经网络建立了立磨粉磨模型, 并通过蚁群算法实现参数最优设置. Umucu 等<sup>[10]</sup>利用多层感知器神经网络和径向基函数神经网络, 建立了水泥粒度模型, 并获得了较高的预测精度.

作为一种数据驱动控制方法, 自适应动态规划 (Adaptive dynamic programming, ADP) 利用离线和在线数据, 采用神经网络逼近动态规划中的性能指标函数和控制策略, 进而满足最优性原理<sup>[11-12]</sup>. 自适应动态规划解决了传统动态规划中的“维数灾难”难题, 为求解非线性系统的最优控制提供了一种切实可行的控制方案. 2002 年, Murray 等<sup>[13]</sup>首先提出了针对连续系统的迭代 ADP 算法. 迭代 ADP 通过策略迭代和值迭代来更新策略方程和值函数<sup>[14-15]</sup>. 然而迭代 ADP 算法往往迭代次数不确定, 计算时间较长且计算量大, 一旦系统发生变化, 需要重新进行离线计算. 近年来, 在线 ADP 算法<sup>[16-17]</sup>被广泛提出, 这些在线算法不需要进行离线迭代计算, 能够通过在线自适应的方式获得最优控制问题的解. 实际生产过程中, 各执行机构变量往往会受到范围约束, 存在饱和和非线性. 本文将在线自适应 ADP 应用到矿渣微粉生产过程中, 在输入受约束的前提下, 设计输入受限的在线 ADP 跟踪控制器, 实现生产过程的最优跟踪控制.

针对矿渣微粉生产难以机理建模的特点, 本文深入研究矿渣微粉生产过程机理, 分析并给出了系统的非线性动态过程结构. 采用递归神经网络, 建立基

于数据的生产过程的辨识模型. 根据辨识模型, 采用在线自适应动态规划的方法, 实现了矿渣微粉生产过程输入有界的最优跟踪控制. 本文结构如下: 第 1 节分析矿渣微粉生产流程并给出系统动态方程描述; 第 2 节建立微粉生产过程的数据驱动模型, 并在此基础上设计基于 ADP 的输入有界跟踪控制器; 第 3 节将提出的建模和控制方法应用到矿渣微粉生产过程中, 验证方法的有效性; 第 4 节对全文进行总结.

## 1 矿渣微粉生产过程

### 1.1 矿渣微粉生产流程

矿渣粉磨生产是将废弃矿渣经过粉磨, 及时选出物料中达到一定粒径要求的微粉. 比表面积为  $400 \text{ m}^2/\text{kg}$  以上的矿渣微粉可以部分代替水泥, 以矿物参合料的形式配置混凝土, 从而大大提高水泥混凝土的力学性能.

矿渣微粉生产系统主要由喂料皮带称、立磨机、物料传送带、收尘器、风机等设备构成. 矿渣原料经过预处理, 由称量皮带进行称量并由传送皮带传送至立磨, 原料由于磨盘旋转的离心力作用, 在粉磨过程中移向磨盘边缘, 在磨盘和磨辊之间的挤压下进行粉磨. 自下而上的热风将物料吹起烘干, 同时将物料提送至磨机上部进行筛选, 符合粒度要求的物料被抽出立磨进行收集, 没有被选出的粗颗粒被重新粉磨, 其现场生产实时监控画面及工艺流程如图 1 和图 2 所示.

### 1.2 矿渣微粉生产系统分析

矿渣微粉生产系统的控制目标是在保证矿渣微粉质量的同时, 最大限度地降低立磨振动, 以保证生产过程平稳运行.

#### 1.2.1 磨机振动

影响立磨稳定运行的因素很多, 立磨振动是最常见也是对立磨运行影响最大的一种异常工况. 根据微粉粉磨生产的控制要求, 立磨振动值仅允许在给定的范围内变动, 超出该范围就可能引起立磨停机故障. 影响立磨振动的因素主要有磨内压差、磨内温度及选粉机转速等.

1) 磨内压差. 磨内压差是指在粉磨过程中, 磨腔选粉机下部与热风入口处静压之差. 压差的变化直接反映了立磨内循环物料量的变化. 正常情况下, 磨内压差是稳定的, 标志着入磨物料量和出磨物料量达到了动态平衡. 压差低于稳定值表明物料入磨量小于出磨量, 料层厚度降低, 直至立磨因空磨而振动停机; 压差高于稳定值表明物料入磨量大于出磨量, 严重时导致料层不稳定或吐渣严重, 造成饱磨而振动停机. 一般情况下, 磨内压差可以通过调整喂料量来进行调节. 实际生产过程中, 磨内压差一般维持

矿渣粉磨工艺流程图

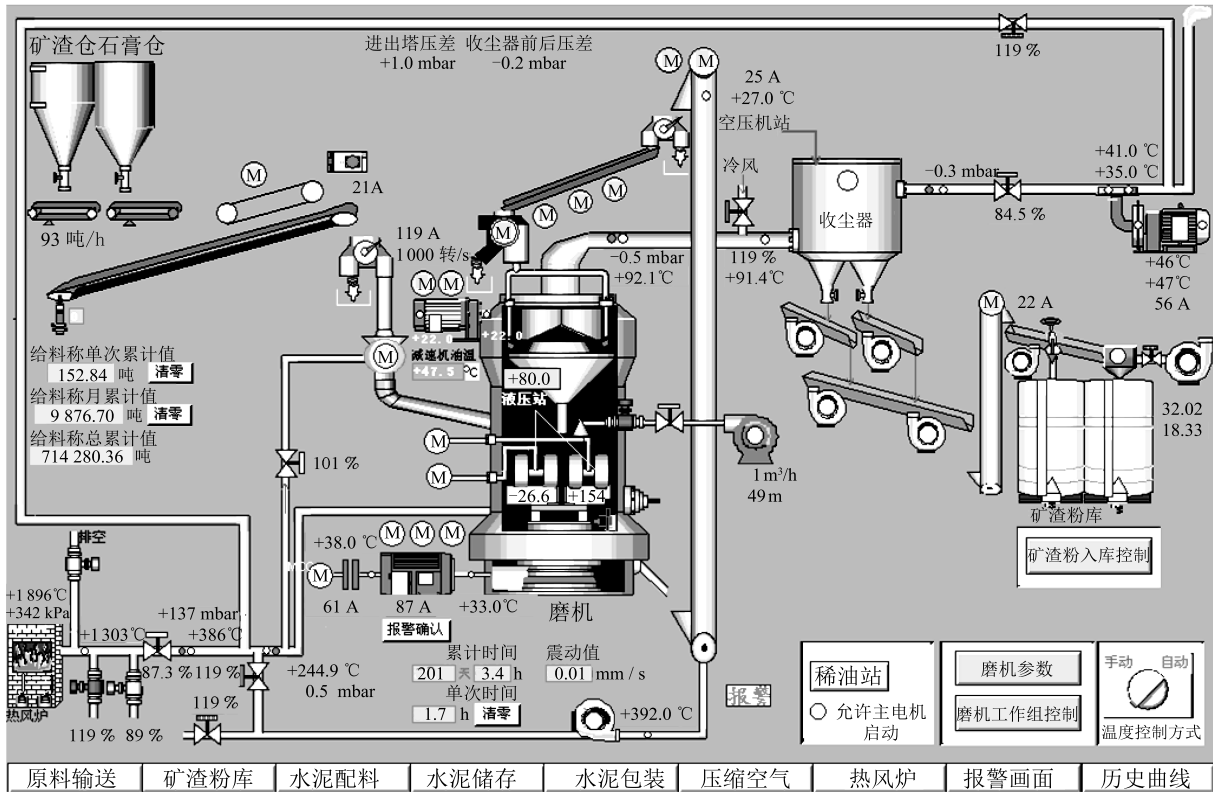


图1 矿渣微粉生产监控画面

Fig. 1 Monitor screen of slag grinding process

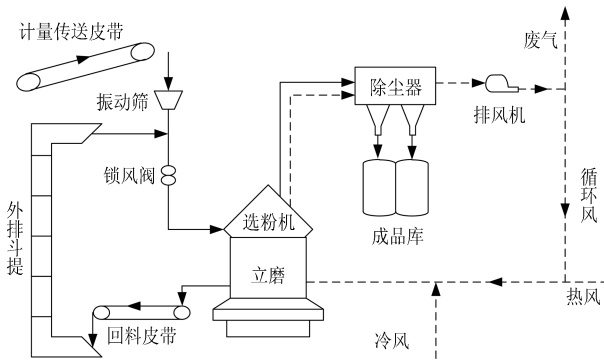


图2 矿渣微粉生产流程图

Fig. 2 Flow chart of slag grinding process

在 20~35 mbar, 并以磨内压差作为立磨振动大小及立磨稳定运行的指标。

2) 磨内温度. 磨内温度过高, 会造成磨内上部物料由于静电作用而大规模聚集, 掉入磨盘底部进行重新粉磨, 从而造成立磨振动值增大. 入磨热风温度是影响磨内温度的主要因素。

3) 选粉机转速. 选粉机转速过快, 大量微粉物料因不能满足细度要求而落入磨内重新研磨, 导致磨内物料增多, 磨内压差增大, 引起立磨振动. 选粉机转速一般要求小于 1 250 r/min.

### 1.2.2 微粉比表面积

微粉比表面积是衡量矿渣微粉产品质量 (即产品细度) 的重要指标, 当高炉矿渣细磨成比表面积为 400 m<sup>2</sup>/kg 以上时, 可以以 20%~70% 的比例替代水泥, 以矿物掺合料的形式参与混凝土配置, 大大提高了水泥混凝土力学性能. 超细矿渣微粉是指比表面积达到 500 m<sup>2</sup>/kg 以上更为细化的矿渣微粉, 其活性较普通的矿渣微粉更强, 具有水化热低、耐腐蚀、与钢筋粘结力强、后期强度高、防微缩等特点. 影响微粉比表面积的因素主要有:

1) 入磨物料. 入磨物料的硬度、含水量等直接影响影响矿渣微粉的产量, 同时也是影响产品细度的重要参数。

2) 入磨热风. 入磨热风不仅对立磨内物料进行烘干, 而且将粉磨后物料提升至选粉区域进行选粉, 热分的风量、风速以及出入磨风温等对选粉机的产量和产品细度都有着重要影响. 在风速一定的前提下, 风量过大或过小都会影响产品细度. 磨内风量增大, 则产品细度变粗, 微粉产量增大, 同时造成磨内回料量增多, 料层稳定性变小; 磨内风量过小, 导致产品细度变细, 产量降低, 同时由于磨内颗粒湿度增加, 造成回料量增多. 磨内风量一般通过控制入磨循

环风阀开度来进行调节。

3) 选粉机转速. 选粉机转速是直接影响微粉生产产量和产品细度的重要参数. 在选粉机分级力场中, 微粉颗粒主要受到两个力的作用: 选粉机转速决定的离心力  $F_u$  和风速决定的径向方向上气流对颗粒的阻力  $F_r$ . 如图 3 所示, 当  $F_u > F_r$  时, 颗粒向边壁沉降, 成为粗粉进一步粉磨; 当  $F_u < F_r$  时, 颗粒在气流的携带作用下穿过分级面形成细粉. 选粉机转速越快, 离心力越大, 产品细度越细. 在实际生产过程中, 出磨风量及风速变化不大, 控制选粉机转速是控制产品细度的重要手段. 立磨选粉机转子采用了变频调速系统, 能够实现平滑调速, 从而更好地控制产品细度.

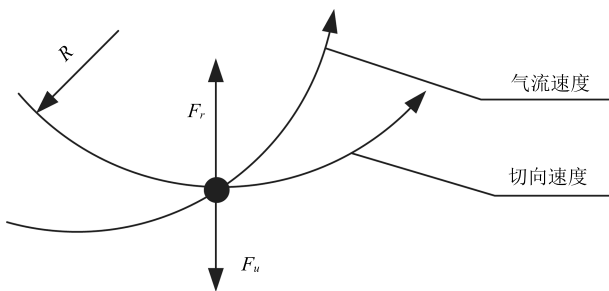


图 3 微粉颗粒受力图

Fig. 3 Stress analysis of slag powder

通过以上分析可知, 矿渣微粉的产量和质量, 以及立磨的正常运行, 受到磨内压差、料层厚度、喂料量、入磨循环风阀开度、选粉机转速等诸多参数的影响. 这些参数之间具有很强的耦合性, 互相影响、互相制约.

### 1.3 矿渣粉磨系统动态方程描述

通过对矿渣粉磨系统进行机理分析, 磨机控制过程主要与以下参数有关:

- 1) 主要被控量, 如比表面积  $s$ 、磨内压差  $p_d$  等;
- 2) 主要控制量, 如喂料量  $m$ 、选粉机转速  $v$ 、入磨热风温度  $T$  及入磨循环风阀开度  $p$ . 同时还受到矿渣湿度、热风风量、入口负压等参数  $\theta$  的影响.

系统动态过程可描述为

$$\dot{x} = f(x, \theta, u) \quad (1)$$

其中,  $u = [u_1, u_2, u_3, u_4]^T$ ,  $x = [x_1, x_2]^T$ ,  $u_1 = m$ ,  $u_2 = v$ ,  $u_3 = T$ ,  $u_4 = p$ ,  $x_1 = s$ ,  $x_2 = p_d$ ,  $\theta$  为常数.

在控制器的设计过程中, 既要保证系统的稳定性、收敛性和对设定值的有效跟踪, 又要保证控制器设计的可实施性. 在实际生产过程中, 各执行器受到自身物理条件的约束, 各输入变量不能任意选取, 只能在给定范围内取值; 从生产工艺来看, 立磨要稳定运行, 其热风温度、喂料量、选粉机转速等运行参数

也必须保持在一定范围内波动. 因此, 往往获得的最优控制并不是理论上的最优值, 而是满足约束条件 (2) 的次优控制.

$$\lim_{t \rightarrow \infty} |x - x_d| = 0$$

$$\text{s. t. } \begin{cases} \underline{m} \leq u_1 \leq \bar{m} \\ \underline{v} \leq u_2 \leq \bar{v} \\ \underline{T} \leq u_3 \leq \bar{T} \\ \underline{p} \leq u_4 \leq \bar{p} \end{cases} \quad (2)$$

其中,  $x_d$  为被控量的期望值,  $\underline{m}, \bar{m}, \underline{v}, \bar{v}, \underline{T}, \bar{T}, \underline{p}, \bar{p}$  为各控制量的变化范围.

## 2 基于 ADP 的输入有界跟踪控制器设计

从被控对象的数学模型描述 (1) 来看, 由于矿渣微粉生产系统具有可控变量多、各变量之间耦合性强、系统非线性等特点, 难以通过机理分析的方法建立全系统的精确数学模型. 本文利用生产过程中的大量数据, 建立有效的数据驱动模型. 并在此基础上, 考虑控制约束条件 (2), 设计在线 ADP 控制策略, 在保证系统稳定的前提下, 提高产品质量, 实现微粉生产过程的最优跟踪控制.

### 2.1 数据驱动建模

基于输入输出数据, 采用递归神经网络 (Recursive neural network, RNN)<sup>[16]</sup>, 系统 (1) 可以写成仿射非线性的形式

$$\dot{x}(t) = A^{*T}x(t) + B^{*T}f(x(t)) + C^{*T}u(t) + D^{*T} + \varepsilon(t) \quad (3)$$

其中, 系统状态  $x \in \mathbf{R}^n$ , 控制输入  $u \in \mathbf{R}^m$ ,  $A^*, B^*, C^*, D^*$  为未知的理想神经网络权值矩阵, 模型重建误差  $\varepsilon(t)$  是有界的. 激活函数  $f(\cdot)$  为单调递增函数, 且对于任意的  $x, y \in \mathbf{R}$ ,  $x \geq y$ , 存在  $k > 0$  满足

$$0 \leq f(x) - f(y) \leq k(x - y) \quad (4)$$

式中,  $\|f(x)\| \leq b_f \|x\|$ , 其中  $b_f$  为正常数. 本文令  $f(x) = \tanh(x)$ .

根据式 (3), 可以构建数据驱动模型

$$\hat{\dot{x}}(t) = \hat{A}^T \hat{x}(t) + \hat{B}^T f(\hat{x}(t)) + \hat{C}^T u(t) + \hat{D}^T - v(t) \quad (5)$$

其中,  $\hat{x}(t)$  为系统状态向量估计值,  $\hat{A}, \hat{B}, \hat{C}, \hat{D}$  分别为期望权值  $A^*, B^*, C^*, D^*$  的估计,  $v(t)$  满足

$$v(t) = S e_m(t) + \frac{\hat{\lambda}(t) e_m(t)}{e_m^T(t) e_m(t) + \eta} \quad (6)$$

其中,  $e_m(t) = x(t) - \hat{x}(t)$  为模型状态误差,  $S \in \mathbf{R}^{n \times n}$  为设计矩阵,  $\hat{\lambda}(t) \in \mathbf{R}$  为调节参数, 常数  $\eta > 1$ .

结合式 (3) 和式 (5), 可推导出模型误差动态方程

$$\begin{aligned} \dot{e}_m(t) = & A^{*T} e_m(t) + \tilde{A}^T(t) \hat{x}(t) + \\ & B^{*T} \tilde{f}(e_m(t)) + \tilde{B}^T f(\hat{x}(t)) + \\ & \tilde{C}^T(t) u(t) + \tilde{D}^T(t) + \\ & \varepsilon(t) + S e_m(t) - \frac{\tilde{\lambda}(t) e_m(t)}{e_m^T(t) e_m(t) + \eta} + \\ & \frac{\lambda^*(t) e_m(t)}{e_m^T(t) e_m(t) + \eta} \end{aligned} \quad (7)$$

其中,  $\tilde{A}(t) = A^*(t) - \hat{A}(t)$ ,  $\tilde{B}(t) = B^*(t) - \hat{B}(t)$ ,  $\tilde{C}(t) = C^*(t) - \hat{C}(t)$ ,  $\tilde{D}(t) = D^*(t) - \hat{D}(t)$ ,  $\tilde{f}(e_m(t)) = f(x(t)) - f(\hat{x}(t))$ ,  $\tilde{\lambda}(t) = \lambda^* - \hat{\lambda}(t)$ .

引理 1<sup>[16]</sup>. 若数据驱动模型 (5) 的网络权值矩阵和调节参数按照如下学习律进行更新:

$$\begin{aligned} \dot{\hat{A}}(t) = & \Lambda_1 \hat{x}(t) e_m^T(t) \\ \dot{\hat{B}}(t) = & \Lambda_2 f(\hat{x}(t)) e_m^T(t) \\ \dot{\hat{C}}(t) = & \Lambda_3 u(t) e_m^T(t) \\ \dot{\hat{D}}(t) = & \Lambda_4 e_m^T(t) \\ \dot{\hat{\lambda}}(t) = & -\Lambda_5 \frac{e_m^T(t) e_m(t)}{e_m^T(t) e_m(t) + \eta} \end{aligned} \quad (8)$$

其中,  $\Lambda_i, i = 1, 2, \dots, 5$  为相应维数的正定矩阵, 则模型辨识误差渐近收敛, 即  $\lim_{t \rightarrow \infty} e_m(t) = 0$ . 且当  $t \rightarrow \infty$  时,  $\hat{A}(t)$ ,  $\hat{B}(t)$ ,  $\hat{C}(t)$ ,  $\hat{D}(t)$  分别趋近于常值矩阵  $A, B, C, D$ .

因此, 采用大量离线数据, 经过足够长时间的模型辨识, 非线性系统 (1) 可表示为

$$\dot{x}(t) = A^T x(t) + B^T f(x(t)) + C^T u(t) + D^T \quad (9)$$

## 2.2 输入受限的 ADP 跟踪控制器设计

针对递归神经网络模型 (9), 本小节采用特殊指标函数解决输入有界问题, 采用评价网逼近指标函数, 采用执行网逼近最优控制, 从而设计出满足控制约束的 ADP 跟踪控制器.

### 2.2.1 问题描述

令  $u_i$  和  $\bar{u}_i$  为控制元素  $u_i$  的下界和上界, 根据式 (9), 矿渣微粉生产系统模型可表示为

$$\dot{x}(t) = A^T x(t) + B^T f(x(t)) + C^T u(t) + D^T, \quad u \in \Omega \quad (10)$$

其中,  $\Omega_u = \{u | u \in \mathbf{R}^m, \underline{u}_i \leq u_i \leq \bar{u}_i, i = 1, 2, \dots, m\}$  为满足约束条件的控制输入.

假设期望状态轨迹为  $x_d(t)$ , 由式 (10) 可得:

$$\dot{x}_d(t) = A^T x_d(t) + B^T f(x_d(t)) + C^T u_d(t) + D^T \quad (11)$$

其中,  $u_d(t)$  为期望输入量, 可通过下式直接计算获得:

$$u_d(t) = C^{-T} (\dot{x}_d(t) - A^T x_d(t) - B^T f(x_d(t)) - D^T) \quad (12)$$

其中,  $C^{-T} = C(C^T C)^{-1}$  为  $C$  的伪逆.

定义状态误差  $e(t) = x(t) - x_d(t)$ . 根据式 (10) 和式 (11), 误差系统可以写成

$$\dot{e}(t) = A^T e(t) + B^T f_e(t) + C^T u_e(t) \quad (13)$$

其中,  $f_e(t) = f(x(t)) - f(x_d(t))$ ,  $u_e(t) = u(t) - u_d(t)$ . 下文中  $e(t)$ ,  $u(t)$ ,  $u_e(t)$ ,  $u_d(t)$  将分别简写为  $e, u, u_e, u_d$ .

根据式 (10), 令  $u_e$  满足约束条件

$$|u_e^i| \leq \chi_i = \min(\bar{u}_i - \max(u_d^i), \underline{u}_i - \min(u_d^i)) \quad (14)$$

其中,  $u_e^i$  和  $u_d^i$  分别为向量  $u_e$  和  $u_d$  的第  $i$  个元素, 记  $\bar{U} = \text{diag}\{\chi_1, \chi_2, \dots, \chi_m\}$ .

对于带有约束的最优跟踪控制问题, 其目标是寻找满足约束条件的最优控制律  $u_e(t)$ , 使得系统 (13) 渐近稳定, 使状态误差  $e$  渐近收敛到零, 并且能够使如下性能指标函数最小:

$$V(e(t)) = \int_t^\infty r(e(\tau), u_e(\tau)) d\tau \quad (15)$$

其中, 效用函数  $r(e, u_e) = Q(e) + W(u_e)$ . 通常,  $Q(e) = e^T Q e$ ,  $W(u_e) = u_e^T R u_e$ ,  $Q$  和  $R$  为对称正定矩阵. 由于要求控制输入有界, 采用

$$W(u_e) = 2 \int_0^{u_e} \varphi^{-T}(\bar{U}^{-1} s) \bar{U} R ds \quad (16)$$

其中,  $R$  为对角正定矩阵,  $s \in \mathbf{R}^m$ ,  $\varphi(\cdot) \in \mathbf{R}^m$  为单调递增的有界奇函数, 满足  $|\varphi(\cdot)| \leq 1$ , 且其一阶导数的界为常数,  $\varphi^{-1}(\cdot)$  表示  $\varphi(\cdot)$  的反函数,  $\varphi^{-T} = (\varphi^{-1})^T$ . 由于  $\varphi^{-1}(\cdot)$  为单调奇函数且  $R$  是正定对角阵, 故  $W(u_e)$  也是正定的. 不失一般性, 本文令  $\varphi(\cdot) = \tanh(\cdot)$ ,  $R$  为  $m$  维单位矩阵.

定义如下 Hamilton 函数:

$$\begin{aligned} H(e, \nabla V_e, u_e) = & (\nabla V_e)^T \times \\ & (A^T e + B^T f_e + C^T u_e) + Q(e) + W(u_e) \end{aligned} \quad (17)$$

其中,  $\nabla V_e = \frac{\partial V(e)}{\partial e} \in \mathbf{R}^n$ .

**定义 1**<sup>[18]</sup>. 若  $u_e(e)$  在  $\Omega$  上连续,  $u_e(0) = 0$ ,  $u_e(e)$  能够使系统 (13) 在  $\Omega$  上稳定, 并且对于任意的  $e \in \Omega$ ,  $V(u_e(0))$  都是有界的, 则称控制量  $u_e(e) : \mathbf{R}^n \rightarrow \mathbf{R}^m$  对于  $\Omega$  上的函数 (15) 是容许的, 记作  $u_e \in \psi(\Omega)$ .

定义最优指标函数

$$V^*(e) = \min_{u_e \in \psi(\Omega)} \int_t^\infty r(e(\tau), u_e(\tau)) d\tau \quad (18)$$

满足如下 HJB 方程:

$$\min_{u_e \in \psi(\Omega)} H(e, \nabla V_e^*, u_e) = 0 \quad (19)$$

进一步, 通过求解  $\frac{\partial H(e, \nabla V_e^*, u_e)}{\partial u_e} = 0$ , 可求得满足约束条件的如下最优控制律:

$$u_e^* = -\bar{U} \tanh(\Phi(e)) \quad (20)$$

其中,  $\Phi(e) = \frac{1}{2}(\bar{U}R)^{-1}C\nabla V_e^*$ .

理论上,  $\nabla V_e^*$  可以通过求解 HJB 方程 (19) 获得, 然而, 式 (19) 实际上是关于  $\nabla V_e^*$  的非线性偏微分方程, 无法直接通过解析方法求解. 为了获得最优控制量, 本文采用在线 ADP 求解.

### 2.2.2 基于 ADP 的跟踪控制器设计

为了实现基于 ADP 的输入有界最优跟踪控制, 分别设计评价网络和执行网络.

1) 评价网络. 根据前馈神经网络的全局逼近特性, 指标函数可以通过如下网络精确表示:

$$V(e) = W_1^T \phi_1(e) + \varepsilon_1(e) \quad (21)$$

其中,  $W_1 \in \mathbf{R}^{N_1}$  为理想评价网络权值,  $N_1$  为神经元个数,  $\phi_1(e) = [\phi_{11}(e), \phi_{12}(e), \dots, \phi_{1N_1}(e)]^T \in \mathbf{R}^{N_1}$  为激活函数向量,  $\varepsilon_1(e)$  为有界评价网逼近误差, 假设其对于  $e$  的一阶偏导也是有界的.

指标函数  $V(e)$  对于  $e$  求导, 可得:

$$\nabla V_e = \nabla \phi_1^T W_1 + \nabla \varepsilon_1 \quad (22)$$

其中,  $\nabla \phi_1 = \frac{\partial \phi_1(e)}{\partial e}$ ,  $\nabla \varepsilon_1 = \frac{\partial \varepsilon_1(e)}{\partial e}$ .

定义  $\hat{W}_1$  为权值  $W_1$  的估计向量, 则指标函数  $V(e)$  的估计可以表示为

$$\hat{V}(e) = \hat{W}_1^T \phi_1(e) \quad (23)$$

根据式 (17) 和式 (23), 可以推导出如下逼近的 Hamilton 函数:

$$\begin{aligned} H(e, u_e, \hat{W}_1) &= \hat{W}_1^T \nabla \phi_1 \times \\ &(A^T e + B^T f_e + C^T u_e) + e^T Q e + W(u_e) = e_1 \end{aligned} \quad (24)$$

评价网络的目标是极小化如下目标函数:

$$E_1(\hat{W}_1) = \frac{1}{2} e_1^T e_1 \quad (25)$$

根据梯度下降法, 评价网的权值更新法则可以表示为

$$\dot{\hat{W}}_1 = -\alpha_c h_1 (h_1^T \hat{W}_1 + e^T Q e + W(u_e)) \quad (26)$$

其中,  $\alpha_c > 0$  为评价网的学习率,  $h_0 = \nabla \phi_1(A^T e + B^T f_e + C^T u_e)$ ,  $h_1 = h_0 / (h_0^T h_0 + 1)$ .

2) 执行网络. 为了求得严格受约束的反馈控制策略, 通过执行网络逼近函数  $\Phi(e)$ :

$$\Phi(e) = W_2^T \phi_2(e) + \varepsilon_2 \quad (27)$$

其中,  $W_2 \in \mathbf{R}^{N_2 \times m}$  为执行网的未知理想权值矩阵,  $\phi_2(e) \in \mathbf{R}^{N_2}$  为执行网激活函数,  $N_2$  为神经元个数,  $\varepsilon_2$  为执行网逼近误差.

定义  $\hat{W}_2$  为  $W_2$  的估计值, 则函数  $\Phi(e)$  的估计值  $\hat{\Phi}(e)$  可以表示为

$$\hat{\Phi}(e) = \hat{W}_2^T \phi_2(e) \quad (28)$$

根据式 (20) 和式 (23), 执行网络的误差函数为

$$e_2 = \hat{W}_2^T \phi_2(e) - \frac{1}{2}(\bar{U}R)^{-1}C\nabla \phi_1^T \hat{W}_1 \quad (29)$$

执行网络的目标是使如下的目标函数极小化:

$$E_2(\hat{W}_2) = \frac{1}{2} e_2^T e_2 \quad (30)$$

根据梯度下降法, 执行网络的权值更新律可以表示为

$$\dot{\hat{W}}_2 = -\alpha_a \frac{\phi_2(e)}{\phi_2^T(e)\phi_2(e) + 1} e_2^T \quad (31)$$

其中,  $\alpha_a$  为执行网的学习率.

执行网输出可表示为

$$\hat{u}_e = -\bar{U} \tanh(\hat{\Phi}(e)) \quad (32)$$

对于带有控制约束的系统 (10), 采用上述方法设计的 ADP 最优控制器, 有如下定理成立.

**定理 1.** 对于如式 (13) 描述的带有控制约束的非线性系统, 如果给定合适的学习律  $\alpha_c$ ,  $\alpha_a$ , 采用式 (23) 和式 (28) 描述的评价网和执行网, 且其权值分别按照式 (26) 和式 (31) 在线调整, 则当神经元个数足够多时, 系统状态误差  $e$  及评价网和执行网的估计误差均一致最终有界.

根据文献 [19] 中的收敛性证明过程, 易证对于误差系统 (13), 定理 1 成立, 具体过程从略.

### 3 仿真分析

本文以济南鲁新新型建材股份有限公司 3 号矿渣微粉生产线采集的数据为研究对象, 验证上述基于数据驱动的建模与最优跟踪控制的有效性.

#### 3.1 实验数据及预处理

实验针对 3 号矿渣微粉生产线运行数据, 每隔 2 min 进行采样, 共获得 325 组实验数据, 如表 1 所示. 由于测量误差或人为因素等原因, 直接从现场获得的数据中不可避免地存在显著误差. 当采样点的值与均值的偏差大于三倍标准差时, 删除该点数据以消除显著误差. 同时, 对最终得到的 250 组数据进行归一化处理.

由前述分析可知, 矿渣微粉生产过程是一个多变量、强耦合的非线性系统. 在实际生产过程的各项数据中, 比表面积作为微粉细度的指标, 是衡量产品质量的重要指标. 磨内压差对于正常生产过程的稳定运行具有重要意义. 本文将微粉的比表面积和磨内压差作为控制量, 将喂料量、选粉机转速、入磨热风温度及入磨循环风阀开度作为控制变量, 通过 RNN 进行模型辨识.

为了保证产品质量的同时, 维持生产过程的稳定, 各控制变量必须达到某一指定范围, 同时根据各执行机构自身约束和现场工程师经验, 各控制变量都有其容许的变化范围, 如表 2 所示.

#### 3.2 模型辨识效果

采用如式 (5) 的递归神经网络对输入输出数据进行辨识. 令递归神经网络初始值为  $\hat{x}(0) = [0; 0]$ , 各辨识参数分别为  $S = -30 I_2$ ,  $\eta = 1.5$ ,  $\Lambda_1 = [1 \ 0.1; 0.1 \ 1]$ ,  $\Lambda_2 = [1 \ 0.2; 0.2 \ 1]$ ,  $\Lambda_3 = [1 \ 1 \ 1 \ 0.1; 1 \ 1 \ 0.1 \ 1; 1 \ 0.1 \ 1 \ 1; 0.1 \ 1 \ 1 \ 1]$ ,  $\Lambda_4 = 0.2$ ,  $\Lambda_5 = 0.1$ . 辨识效果和辨识误差如图 4 和图 5 所示, 为了方便显示模型辨识效果, 图 4 仅显示前 40 组数据模型辨

识曲线. 从图 5 可以看出, 由于初始值选择问题, 模型误差在初始阶段较大, 但经过一段时间后, 模型误差趋近于 0, 得到的基于数据的模型能够有效地辨识出真实系统.

表 2 各控制变量容许变化范围

Table 2 Tolerance range of different variables

	水渣进料 (10 <sup>3</sup> kg/Hr)	电机转速 (r/min)	进口风温 (°C)	入磨循环风阀开度 (%)
最大值	160	1300	300	80
最小值	0	0	150	0

得到的辨识模型 (10) 中各稳定参数为

$$A = \begin{bmatrix} 0.3135 & -0.1570 \\ 0.3766 & 0.1374 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.4500 & 0.0675 \\ -0.4391 & -0.0244 \end{bmatrix}$$

$$C = \begin{bmatrix} -0.3522 & -0.1475 & 0.0581 & -0.2100 \\ -0.3834 & 0.0302 & 0.1384 & -0.6277 \end{bmatrix}^T$$

$$D = \begin{bmatrix} 0.0042 & 0.4030 \end{bmatrix}$$

#### 3.3 ADP 跟踪控制效果

根据现场工程师经验, 微粉比表面积维持在 440 m<sup>2</sup>/kg 左右时, 产品质量满足要求, 且产量最优; 磨内压差维持在 27 mbar 左右时, 保证出入磨物料达到动态平衡, 生产能够稳定运行. 基于以上经验数据, 通过归一化, 得到微粉生产过程的跟踪目标, 使输出状态跟踪到如下期望轨迹:

$$\dot{x}_{d1} = 0.7512 (1 - e^{-0.5t})$$

$$\dot{x}_{d2} = 0.7726 (1 - e^{-t})$$

表 1 济钢鲁新建材 3 号矿渣微粉生产线生产运行数据

Table 1 Production data of Luxin mill line 3

编号	水渣进料 (10 <sup>3</sup> kg/Hr)	电机转速 (r/min)	进口风温 (°C)	入磨循环风阀开度 (%)	比表面积 (cm <sup>2</sup> /g)	磨内压差 (mbar)
1	85.60	1250	230	65.13	438.5	27.60
2	84.81	1160	229	69.50	426.3	28.13
3	84.77	1240	235	66.17	430.7	26.97
⋮	⋮	⋮	⋮	⋮	⋮	⋮
323	99.63	1049	242	60.59	438.5	24.65
324	100.42	1050	243	60.53	426.3	24.94
325	101.20	1051	248	60.62	433.9	25.00

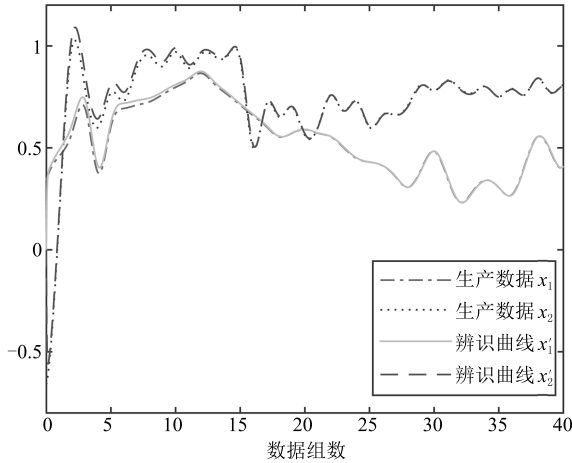


图 4 模型辨识曲线

Fig. 4 Curve of model identification

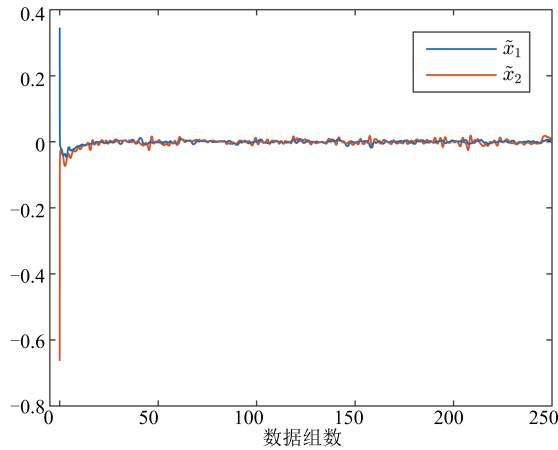


图 5 模型辨识误差曲线

Fig. 5 Curve of model identification error

同时各控制变量需满足表 2 所示的控制约束. 根据式 (12), 可以推导出期望状态  $u_d$ , 其变化范围为

$$\begin{aligned} -0.77949 &\leq u_{d1} \leq 0.94702 \\ -0.50552 &\leq u_{d2} \leq 0.61426 \\ -0.08557 &\leq u_{d3} \leq 0.05804 \\ -0.09014 &\leq u_{d4} \leq 0.38433 \end{aligned}$$

为了使最优控制量满足约束条件, 令  $|u_e| \leq [1.86 \ 0.59 \ 1.24 \ 1.27]^T$ .

令评价网激活函数  $\phi_1 = [e_1 e_2 \ e_1^2 \ e_2^2]^T$ , 执行网激活函数  $\phi_2 = [e_1 \ e_2]^T$ , 神经网络学习率分别为  $\alpha_c = 0.8$ ,  $\alpha_a = 0.5$ . 令评价网权值初始值为  $\hat{W}_1 = [1 \ 1 \ 1]^T$ , 执行网初始权值要反映初始容许控制, 令其初始值为  $\hat{W}_2 = [-2.263 \ -1.3483 \ -1.2997 \ 1.9595; -0.5561 \ 1.3 \ 2.3772 \ -0.6598]$ . 为了说明控制效果,

采用具有相同参数配置的无控制约束的最优控制算法<sup>[16]</sup> 作为对比.

采用带有控制约束的最优跟踪控制策略, 评价网和执行网权值曲线如图 6 和图 7 所示, 各权值均收敛到稳定值. 图 8 显示了期望控制曲线  $u_d$  和采用本文方法得到的各控制曲线  $u$ . 采用无控制量约束最优控制方法得到的各控制量曲线如图 9 所示. 图 10 显示了两种控制方法得到的状态跟踪曲线  $x^c$ ,  $x^u$  和期望状态曲线  $x^d$ .

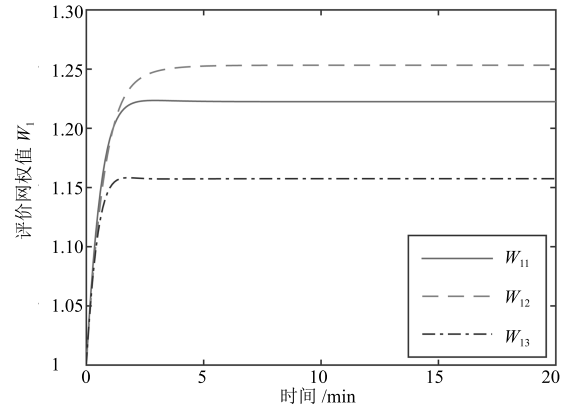


图 6 评价网权值曲线

Fig. 6 Critic network weights

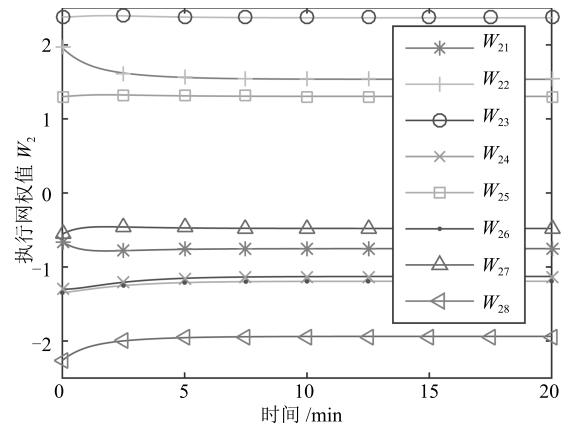


图 7 执行网权值曲线

Fig. 7 Actor network weights

由图 8 及对比表 2 所示的容许范围可知, 采用本文控制方法能够使控制误差  $u_e$  满足约束条件, 从而保证控制量  $u$  始终保持在约束范围内. 而无约束的最优控制由于  $u_e$  最大值过大, 造成控制量  $u$  超出约束范围, 如图 9 所示. 电机转速  $u_2$  在初始阶段为负值, 明显与实际意义不符; 进口风温  $u_3$  最大值远远超过了实际生产能够提供的最大风温. 因此, 图 9 所示的无约束控制结果仅能在理论上跟踪上期望轨迹, 实际生产过程中很难被采用. 同时, 如图 10 所示, 在控制过程中, 带有控制约束的输出状态  $x^c$  同样



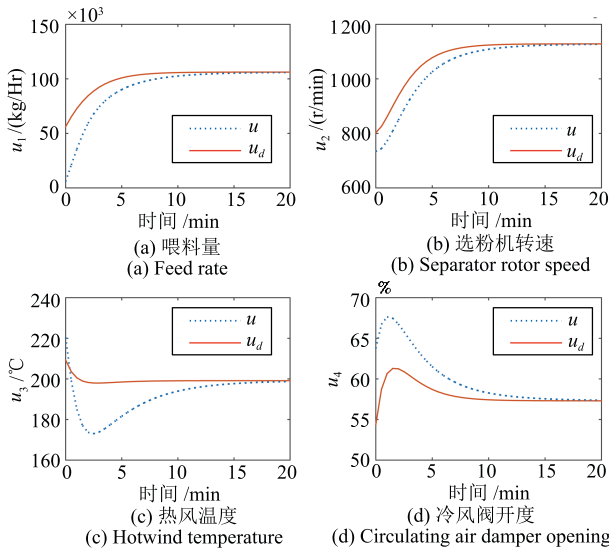


图 8 受约束控制曲线  
Fig. 8 Constrained control signal

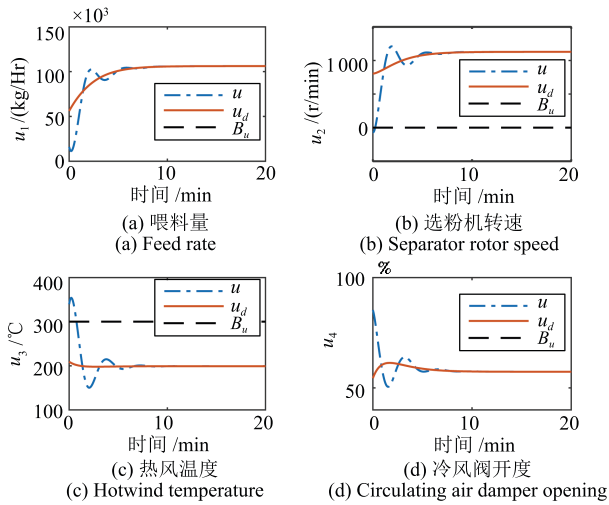


图 9 无约束控制曲线  
Fig. 9 Control signal without constraints

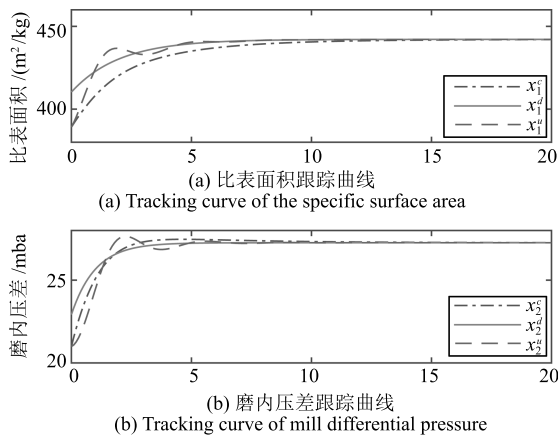


图 10 状态输出曲线  
Fig. 10 Output state signal

能够使比表面积跟踪上期望值  $440 \text{ m}^2/\text{kg}$ , 磨内压差稳定在  $27 \text{ mbar}$ . 同时, 由于控制量受约束, 跟踪曲线  $x^c$  超调量小于无约束最优控制曲线  $x^u$ , 但跟踪上期望轨迹需要更多的调整时间.

### 4 结论

针对矿渣微粉生产这一多变量、强耦合、控制有界的复杂非线性系统, 本文提出了一种基于数据的最优跟踪控制方法. 采用递归神经网络建立了生产过程的数据驱动模型, 并在该模型的基础上设计了基于近似动态规划的最优控制器, 在状态输出跟踪上期望轨迹的同时, 保证了控制量始终在约束范围内变化. 微粉生产过程的实验表明, 在各执行器满足约束的前提下, 该方法能够保证微粉质量和磨内压差稳定, 从而保证生产过程平稳运行. 在接下来的研究中, 将引入微粉产量这一重要指标, 针对产量-质量这两个互相矛盾的目标, 采用多目标优化算法求得最优解, 作为微粉生产过程的跟踪值, 进而采用本文方法实现微粉生产过程产量-质量的最优跟踪控制.

### References

- 1 Işıklıdağ, Topçu İ B. The effect of ground granulated blast-furnace slag on properties of Horasan mortar. *Construction and Building Materials*, 2013, **40**: 448–454
- 2 Zhang Y J, Zhang X. Grey correlation analysis between strength of slag cement and particle fractions of slag powder. *Cement and Concrete Composites*, 2007, **29**(6): 498–504
- 3 Chen Yuan. Study on Separator of Large-scale Vertical Mill [Master dissertation], Chongqing University, China, 2008. (陈远. 大型立磨选粉机研究 [硕士学位论文], 重庆大学, 中国, 2008.)
- 4 Xu J X, Hou Z S. Notes on data-driven system approaches. *Acta Automatica Sinica*, 2009, **35**(6): 668–675
- 5 Hou Zhong-Sheng, Xu Jian-Xin. On data-driven control theory: the state of the art and perspective. *Acta Automatica Sinica*, 2009, **35**(6): 650–667 (侯忠生, 许建新. 数据驱动控制理论及方法的回顾和展望. *自动化学报*, 2009, **35**(6): 650–667)
- 6 Dai Wei, Chai Tian-You. Data-driven optimal operational control of complex grinding processes. *Acta Automatica Sinica*, 2014, **40**(9): 2005–2014 (代伟, 柴天佑. 数据驱动的复杂磨矿过程运行优化控制方法. *自动化学报*, 2014, **40**(9): 2005–2014)
- 7 Yan Wen-Jun, Qin Wei. Modeling and control optimization in cement vertical roller mill process. *Control Engineering of China*, 2012, **19**(6): 929–943 (颜文俊, 秦伟. 水泥立磨流程的建模和控制优化. *控制工程*, 2012, **19**(6): 929–943)
- 8 Cai X Y, Meng Q J, Luan W L. Soft sensor of vertical mill material layer based on LS-SVM. In: Proceedings of the

2013 International Conference on Measurement, Information, and Control (ICMIC). Harbin, China: IEEE, 2013. 22–25

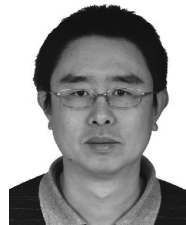
- 9 Lin X F, Qian Z. Modeling of vertical mill raw meal grinding process and optimal setting of operating parameters based on wavelet neural network. In: Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN). Beijing, China: IEEE, 2014. 3015–3020
- 10 Umucu Y, Çağlar M F, Gündüz L, Bozkurt V, Deniz V. Modeling of grinding process by artificial neural network for calcite mineral. In: Proceedings of the 2011 International Symposium on Innovations in Intelligent Systems and Applications (INISTA). Istanbul, Turkey: IEEE, 2011. 344–348
- 11 Zhang Hua-Guang, Zhang Xin, Luo Yan-Hong, Yang Jun. An overview of research on adaptive dynamic programming. *Acta Automatica Sinica*, 2013, **39**(4): 303–311  
(张化光, 张欣, 罗艳红, 杨珺. 自适应动态规划综述. 自动化学报, 2013, **39**(4): 303–311)
- 12 Xu X, Zuo L, Huang Z H. Reinforcement learning algorithms with function approximation: recent advances and applications. *Information Sciences*, 2014, **261**: 1–31
- 13 Murray J J, Cox C J, Lendaris G G, Saeks R. Adaptive dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 2002, **32**(2): 140–153
- 14 Wei Q L, Liu D R, Yang X. Infinite horizon self-learning optimal control of nonaffine discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2015, **26**(4): 866–879
- 15 Liu D R, Wei Q L. Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(3): 621–634
- 16 Zhang H, Cui L, Zhang X, Luo Y H. Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Transactions on Neural Networks*, 2011, **22**(12): 2226–2236
- 17 Modares H, Lewis F L, Naghibi-Sistani M B. Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems. *Automatica*, 2014, **50**(1): 193–202
- 18 Wei Q, Liu D. A novel iterative  $\theta$ -adaptive dynamic programming for discrete-time nonlinear systems. *IEEE Transactions on Automation Science and Engineering*, 2014, **11**(4): 1176–1190
- 19 Qin C B, Zhang H G, Luo Y H. Adaptive optimal control for nonlinear discrete-time systems. In: Proceedings of the 2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning. Singapore: IEEE, 2013. 13–18



**王康** 北京科技大学自动化学院博士研究生. 2012 年获得北京科技大学自动化系学士学位. 主要研究方向为最优控制, 自适应控制.

E-mail: wangkangustb@gmail.com

(**WANG Kang** Ph.D. candidate at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. He received his bachelor degree from University of Science and Technology Beijing in 2012. His research interest covers optimal control and adaptive control.)



**李晓理** 北京工业大学电子信息与控制工程学院教授. 1997 年获得大连理工大学控制理论与工程硕士学位, 2000 年获得东北大学博士学位. 主要研究方向为多模型自适应控制, 神经网络控制. 本文通信作者.

E-mail: lixiaolibjut@bjut.edu.cn

(**LI Xiao-Li** Professor at the College of Electronic Information and Control Engineering, Beijing University of Technology. He received his master degree in control theory and control engineering from Dalian University of Technology in 1997, and Ph.D. degree from North-eastern University in 2000, respectively. His research interest covers multiple model adaptive control and neural network control. Corresponding author of this paper.)



**贾超** 北京科技大学自动化学院博士研究生. 2011 年获得青岛理工大学学士学位. 主要研究方向为多模型控制, 模糊控制和神经网络控制.

E-mail: jiachaocharles@outlook.com

(**JIA Chao** Ph.D. candidate at the School of Automation and Electrical Engineering, University of Science and Technology Beijing. He received his bachelor degree from Qingdao Technological University in 2011. His research interest covers multiple model control, fuzzy control, and neural network control.)



**宋桂芝** 济南鲁新新型建材股份有限公司电气工程师. 2007 年获得山东大学电气工程及其自动化硕士学位. 主要研究方向为大型立磨系统的自动控制.

E-mail: luxinsonggz@163.com

(**SONG Gui-Zhi** Electrical engineer at Jinan Luxin Materials Company Limited. She received her bachelor degree in electric engineering and automation from Shandong University in 2007. Her research interest covers automatic control of large scale vertical mill.)