

基于多信号源的神经模糊 Hammerstein-Wiener 模型研究

贾立¹ 杨爱华¹ 邱铭森²

摘要 面对复杂工业过程控制的需求,设计一种结合数据信息的特殊模型结构,在保证控制系统有效性的前提下通过模型的结构来简化控制器的求解是亟待解决的问题.为此,本文提出一种基于多信号源的神经模糊 Hammerstein-Wiener 模型,突破传统的迭代分离方法,通过组合式多信号实现 Hammerstein-Wiener 模型中神经模糊非线性环节和线性环节的分离,同时设计了神经模糊模型参数的非迭代优化算法,将研究结果推广到分段非线性系统,改善了模型的适用范围.该算法保证了模型的预测精度,具有逼近较强非线性过程的能力.在此基础上设计了基于神经模糊 Hammerstein-Wiener 模型的控制策略,利用模型的特殊结构将非线性系统的控制问题简化为线性系统的控制问题,采用简单的 PID 控制器便能达到较好的控制效果.仿真结果验证了上述方法的有效性.

关键词 Hammerstein-Wiener 模型,神经模糊系统,非线性系统,信号分离

引用格式 贾立, 杨爱华, 邱铭森. 基于多信号源的神经模糊 Hammerstein-Wiener 模型研究. 自动化学报, 2013, 39(5): 690-696

DOI 10.3724/SP.J.1004.2013.00690

Research on Multi-signal Based Neuro-fuzzy Hammerstein-Wiener Model

JIA Li¹ YANG Ai-Hua¹ CHIU Min-Sen²

Abstract In order to solve the control problem of complex systems, it is important to design a special structure model with data information to simplify the question of designing control system. Thus, a multi-signal based neuro-fuzzy Hammerstein-Wiener model is proposed, which breaks through the traditional iterative separation method. The separation of the neuro-fuzzy nonlinear and linear parts of the Hammerstein-Wiener model is realized by one kind of multi-signals. And a noniterative neuro-fuzzy optimization algorithm is designed to expand the research results to piecewise nonlinear system, which can be applied to much more nonlinear systems. This algorithm guarantees the precision of the model. Moreover, it has the ability of approximating strong nonlinearity. Furthermore, a neuro-fuzzy Hammerstein-Wiener model based control system is designed to simplify the control problem of the nonlinear system into the problem of linear system by using the special structure of the model. As a result, the traditional PID controller can get a better control result. Simulated results show the effectiveness of the method.

Key words Hammerstein-Wiener model, neuro-fuzzy systems, nonlinear systems, signal separation

Citation Jia Li, Yang Ai-Hua, Chiu Min-Sen. Research on multi-signal based neuro-fuzzy Hammerstein-Wiener model. *Acta Automatica Sinica*, 2013, 39(5): 690-696

在基于数据的非线性动态系统建模研究领域,块结构非线性动态模型是一个典型的研究热点.该模型用静态(无记忆)非线性环节与动态线性环节通过内部级联的方式来模型化,具有易辨识、计算量少、能较好地反映过程特征的特点^[1-6],可以描述一大类工业非线性过程,如蒸馏塔、热交换器、pH 中和、聚烯烃树脂牌号切换、连续搅拌反应釜(Continuous stirred tank reactor, CSTR)、精对苯二甲酸(Pure terephthalic acid, PTA)生产中的浆料配置系统、冶金反应过程、生物系统等^[7].块结构模型具有传统非线性模型不可比拟的优越性,可以和先进控制算法结合,在保证控制系统有效性的前提下简化控制器的求解,把非线性控制问题简化为线性模型预测控制问题,避免了传统非线性控制方法计算量大、收敛性和稳定性不能得到保证等诸多问题;从而可以直接利用线性控制系统中的成熟理论,达到降低能耗、提高生产装置经济效益的目的,并且便于现场操作人员理解.

目前,这方面的研究主要分为迭代法、过参数法、分段常数输入信号法、二进制和多步混合输入信号法,如 Bai^[8]针对基于非线性函数基的 Hammerstein-Wiener 模型,提出了两阶段辨识法,首先采用递推最小二乘法辨识出模型的回归系数,再采用奇异值分解法辨识出线性和非线性部分的参数,该方法仅适用于离散系统中线性部分是最小相位动态的形式;Wang 等^[9]提出了增广随机梯度辨识方法,并给出了平均方法和奇异值分解两种参数分离方法,在这类重新参数化 Hammerstein-Wiener 模型中,获得参数估计包含了原系统参数的乘积项;Zhu^[10]采用松弛迭代法,通过最小化辨识误差对 Hammerstein-Wiener 模型进行参数估计;Crama 等^[11]首先通过傅里叶变换,将问题由时域空间转到频域空间,再采用迭代法对 Hammerstein-Wiener 模型进行辨识;Park 等^[12]研究了采用激励信号辨识 Hammerstein-Wiener 模型的方法;满红等^[13]提出了一种基于神经网络的预测控制策略,采用分段最小二乘支持向量机辨识 Hammerstein-Wiener 模型的方法;李妍等^[14-15]针对输入和输出受约束的 Hammerstein-Wiener 模型,提出了一种基于 T-S 模糊模型的非线性预测控制算法和一种基于多面体终端域的预测控制综合算法.总体上,90%的文献都采用迭代分离方法分离静态(无记忆)非线性环节与动态线性环节,需要在一定的假设条件下才能实施分离,因此在实际应用中具有一定的局限性.文献[12]通过组合式多信号源实现了 Hammerstein-Wiener 模型中非线性环节和线性环节的分离,但该方法采用多项式拟合静态非线性环节,仅适用于连续非线性函数的情况;并且多项式描述的非线性只能代表一类轻微的非线性过程,对于具有中等或严重非线性的过程,其精度和适应性将会有所下降^[16].

本文在上述研究工作的基础上,提出一种基于多信号源

Higher Education of China (20093108120013), and Shanghai University "11th Five-Year Plan" 211 Construction Project

本文客座编委 王宏

Recommended by Guest Editor WANG Hong

1. 上海大学机电工程与自动化学院 上海市电站自动化技术重点实验室 上海 200072, 中国 2. 新加坡国立大学工程学院 新加坡 119260, 新加坡

1. Shanghai Key Laboratory of Power Station Automation Technology, Department of Automation, College of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200072, China
2. Faculty of Engineering, National University of Singapore, Singapore 119260, Singapore

收稿日期 2012-05-11 录用日期 2012-10-12
Manuscript received May 11, 2012; accepted October 12, 2012
国家自然科学基金(61004019, 61074032), 教育部博士点基金(20093108120013), 上海大学“十一五”211 建设项目资助
Supported by National Natural Science Foundation of China (61004019, 61074032), Research Fund for the Doctoral Program of

的神经模糊 Hammerstein-Wiener 模型. 通过组合式多信号实现 Hammerstein-Wiener 模型中的非线性环节和线性环节的分离, 同时设计了神经模糊 Hammerstein-Wiener 模型的参数非迭代优化算法, 将研究结果拓广到分段非线性系统, 为神经模糊系统的参数求解提供了新思路. 该算法保证了模型的预测精度, 具有逼近较强非线性过程的能力. 在此基础上设计了基于神经模糊 Hammerstein-Wiener 模型的控制系

1 神经模糊 Hammerstein-Wiener 模型

考虑单输入单输出 Hammerstein-Wiener 系统, 模型结构如图 1 所示, 分成输入静态 (无记忆) 非线性环节、动态线性环节、输出静态 (无记忆) 非线性环节三部分, 通过内部级联的方式来模型化, 表示如下:

$$\begin{aligned}
 v(k) &= f(u(k)) \\
 z(k+1) &= \sum_{i=1}^{n_a} a_i z(k+1-i) + \sum_{j=1}^{n_b} b_j v(k+1-j) \\
 y(k) &= g(z(k))
 \end{aligned} \tag{1}$$

其中, $f(\cdot)$ 和 $g(\cdot)$ 分别表示输入静态非线性环节和输出静态非线性环节, $u(k)$ 和 $y(k)$ 分别表示模型在 k 时刻的输入和输出, $v(k)$ 表示输入静态非线性环节的输出, $z(k)$ 表示动态线性环节的输出, n_a 和 n_b 表示与模型阶数相关的参数.

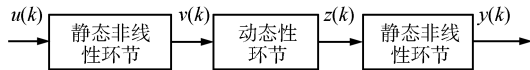


图 1 Hammerstein-Wiener 模型结构图
Fig. 1 The structure of Hammerstein-Wiener model

对于任意给定的 ε , 建立 Hammerstein-Wiener 模型就是要寻求满足如下条件的参数:

$$\begin{aligned}
 E(\hat{f}, \hat{\theta}, \hat{g}) &= \frac{1}{2N_P} \sum_{k=1}^{N_P} (\hat{y}(k) - y(k))^2 \leq \varepsilon \\
 \text{s.t. } \hat{v}(k) &= \hat{f}(u(k)) \\
 \hat{z}(k+1) &= \sum_{i=1}^{n_a} \hat{a}_i \hat{z}(k+1-i) + \sum_{j=1}^{n_b} \hat{b}_j \hat{v}(k+1-j) \\
 \hat{y}(k) &= \hat{g}(\hat{z}(k))
 \end{aligned} \tag{2}$$

其中, $\hat{\theta} = [\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{n_a}, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_{n_b}]$. 为简化后续讨论, 将输出静态非线性环节记为 $\hat{z}(k) = \hat{g}^{-1}(\hat{y}(k))$. 本文采用神经模糊网络拟合静态非线性环节, 模型由五层前向网络构成, 第 1 层为输入层, 由 1 个神经元构成, 神经元起着将输入信号 $x(k)$ 直接传给下一层的作用; 第 2 层为模糊化层, 由 N 个神经元组成 (N 表示模糊规则数), 每个神经元表示一个隶属度函数; 第 3 层为模糊条件层, 由 N 个神经元组成, 每个神经元表示一条模糊规则; 第 4 层为模糊决策层, 由两个神经元组成; 第 5 层为输出层, 由 1 个神经元组成. 模型的输出为

$$\begin{aligned}
 \text{output} &= \frac{\sum_{j=1}^N w_j \mu_j(x(k))}{\sum_{j=1}^N \mu_j(x(k))} = \\
 &= \frac{\sum_{j=1}^N w_j \exp\left(-\frac{(x(k)-c_j)^2}{b_j^2}\right)}{\sum_{j=1}^N \exp\left(-\frac{(x(k)-c_j)^2}{b_j^2}\right)} = \\
 &= \sum_{j=1}^N \phi_j(x(k)) w_j
 \end{aligned} \tag{3}$$

其中, $\phi_j(x(k)) = \frac{\exp\left(-\frac{(x(k)-c_j)^2}{b_j^2}\right)}{\sum_{j=1}^N \exp\left(-\frac{(x(k)-c_j)^2}{b_j^2}\right)}$, $\mu_j(x(k))$ 表示高斯型隶属度函数, c_j 表示高斯型隶属度函数的中心, b_j 表示高斯型隶属度函数的宽度, w_j 为后件参数.

因此, 基于神经模糊网络的输入静态非线性环节、输出静态非线性环节可分别表示为

$$\hat{v}(k) = \hat{f}(u(k)) = \sum_{l=1}^{N^{\text{input}}} \phi_l^{\text{input}}(u(k)) w_l^{\text{input}} \tag{4}$$

$$\hat{z}(k) = \hat{g}^{-1}(\hat{y}(k)) = \sum_{j=1}^{N^{\text{output}}} \phi_j^{\text{output}}(\hat{y}(k)) w_j^{\text{output}} \tag{5}$$

其中, 上标 “input” 表示输入非线性环节, “output” 表示输出非线性环节.

由于块结构模型由静态非线性环节和动态线性环节组合而成, 在实际工程应用中, 中间变量 (图 1 中 $v(k)$ 和 $z(k)$) 的信息不可测量, 因此这不同于传统的神经模糊非线性建模方式, 不是简单采用输入输出数据建立被控过程的数据库、知识库与规则库的问题. 本文将通过多源激励信号的组合设计间接获取中间不可测量, 有效分离各串联模块, 以降低模型的复杂性.

2 神经模糊 Hammerstein-Wiener 模型的分离

采用图 2 所示的多信号源^[12] 作为激励信号, 该信号源由 $u_1^{\text{binary}}(k)$ 、 $u_2^{\text{binary}}(k)$ 和 $u_3^{\text{random}}(k)$ 三种信号构成. 其中, $u_1^{\text{binary}}(k)$ 表示幅值为 0 或 λ 的二进制信号, $u_2^{\text{binary}}(k)$ 表示幅值为 0 或 $\alpha \cdot \lambda$ ($0 < \alpha < 1$) 的二进制信号, $u_3^{\text{random}}(k)$ 表示随机信号. 与此相对应的输入非线性环节的输出分别为 $v_1(k)$ 、 $v_2(k)$ 和 $v_3(k)$, 线性环节的输出分别为 $z_1(k)$ 、 $z_2(k)$ 和 $z_3(k)$, 输出非线性环节的输出分别为 $y_1(k)$ 、 $y_2(k)$ 和 $y_3(k)$, 如图 2 所示. 由于信号源 $u_1^{\text{binary}}(k)$ 和 $u_2^{\text{binary}}(k)$ 的幅值成倍数关系, 即 $u_2^{\text{binary}}(k) = \alpha \cdot u_1^{\text{binary}}(k)$, 因此可知静态非线性环节的输出 $v_1(k)$ 和 $v_2(k)$ 也是二进制信号, 并且幅值为 β 倍数关系: $v_2(k) = \beta \cdot v_1(k)$. 再根据线性系统的叠加定理, 可知动态线性环节的 β 倍数输入信号 $v(k)$ 必将导致 β 倍数输出信号 $z(k)$, 即 $z_2(k) = \beta \cdot z_1(k)$.

根据上述分析, 本文将通过多源激励信号的组合设计间接获取中间不可测量. 具体来讲, 可利用两组不同幅值的二进制信号源 $u_1^{\text{binary}}(k)$ 和 $u_2^{\text{binary}}(k)$ 及其相应的输出 $y_1(k)$ 和 $y_2(k)$ 求解输出非线性环节的参数. 在此基础上, 将输入非线性环节和动态线性环节看成 Hammerstein 系统, 利用一组二进制信号源 $u_1^{\text{binary}}(k)$ 及其相应的输出 $y_1(k)$ 求解动态

线性环节的参数。最后, 利用随机信号源 $u_3^{\text{random}}(k)$ 及其相应的输出 $y_3(k)$ 求解输入非线性环节的参数。本文的关键是如何在上述特殊信号的关系中推导出输入静态非线性神经模糊网络和输出静态非线性神经模糊网络的参数。

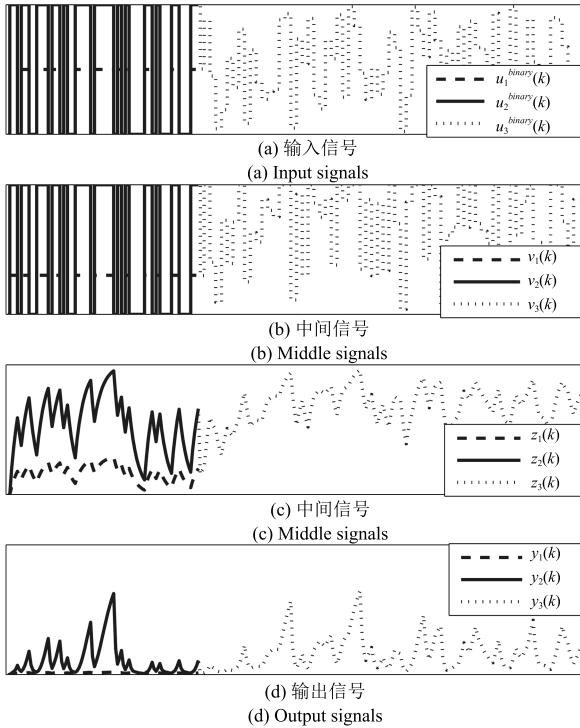


图 2 特殊输入信号源
Fig. 2 Special input signals

2.1 基于神经模糊网络的输出非线性模型

在两组不同幅值的二进制信号源 $u_1^{\text{binary}}(k)$ 、 $u_2^{\text{binary}}(k)$ 作用下, 输出静态非线性神经模糊模型可分别表示为

$$\hat{z}_1(k) = \hat{g}^{-1}(\hat{y}_1(k)) = \sum_{j=1}^{N^{\text{output}}} \phi_j^{\text{output}}(\hat{y}_1(k)) w_j^{\text{output}} \quad (6)$$

$$\hat{z}_2(k) = \hat{g}^{-1}(\hat{y}_2(k)) = \sum_{j=1}^{N^{\text{output}}} \phi_j^{\text{output}}(\hat{y}_2(k)) w_j^{\text{output}} \quad (7)$$

根据上述分析, 动态线性环节的输出, 即输出静态非线性神经模糊模型的输入有如下关系:

$$\hat{z}_2(k) = \hat{\beta} \cdot \hat{z}_1(k) \quad (8)$$

因此, 由式 (6) ~ (8) 可以得到:

$$\sum_{j=1}^{N^{\text{output}}} \phi_j^{\text{output}}(\hat{y}_2(k)) w_j^{\text{output}} = \hat{\beta} \cdot \sum_{j=1}^{N^{\text{output}}} \phi_j^{\text{output}}(\hat{y}_1(k)) w_j^{\text{output}} \quad (9)$$

将式 (9) 展开可得到:

$$\begin{aligned} \phi_1^{\text{output}}(\hat{y}_2(k)) - \hat{\beta} \cdot \phi_1^{\text{output}}(\hat{y}_1(k)) = & \\ \frac{1}{w_1^{\text{output}}} \cdot (\hat{\beta} \cdot \phi_2^{\text{output}}(\hat{y}_1(k)) w_2^{\text{output}} + \dots + & \\ \hat{\beta} \cdot \phi_{N^{\text{output}}}^{\text{output}}(\hat{y}_1(k)) w_{N^{\text{output}}}^{\text{output}}) - & \\ \frac{1}{w_1^{\text{output}}} \cdot (\phi_2^{\text{output}}(\hat{y}_2(k)) w_2^{\text{output}} + \dots + & \\ \phi_{N^{\text{output}}}^{\text{output}}(\hat{y}_2(k)) w_{N^{\text{output}}}^{\text{output}}) & \end{aligned} \quad (10)$$

记 $\bar{w}_i^{\text{output}} = \frac{w_i^{\text{output}}}{w_1^{\text{output}}}$ ($i = 2, 3, \dots, N^{\text{output}}$), 式 (10) 又可表示为

$$\begin{aligned} \phi_1^{\text{output}}(\hat{y}_2(k)) - \hat{\beta} \cdot \phi_1^{\text{output}}(\hat{y}_1(k)) = & \\ (\hat{\beta} \cdot \phi_2^{\text{output}}(\hat{y}_1(k)) - \phi_2^{\text{output}}(\hat{y}_2(k))) \bar{w}_2^{\text{output}} + \dots + & \\ (\hat{\beta} \cdot \phi_{N^{\text{output}}}^{\text{output}}(\hat{y}_1(k)) - \phi_{N^{\text{output}}}^{\text{output}}(\hat{y}_2(k))) \bar{w}_{N^{\text{output}}}^{\text{output}} & \end{aligned} \quad (11)$$

首先, 采用距离-角度相似性判据^[17]

$$S_L = \max_{1 \leq l \leq N} \left(\sqrt{e^{-\|\hat{y}(k) - c_j\|_2^2}} \right)$$

建立模糊规则前件参数 c_j , b_j , N . 该算法将非线性模型在空间中分成几个不同的区域, 从而减少了模糊规则数, 增强了区域分布的合理性。

在此基础上, 将输出静态非线性神经模糊模型参数求解转化为如下优化问题:

$$E_1(\hat{g}^{-1}(\cdot), \hat{\beta}, \Theta) = \min_{\hat{\beta}} \left(\frac{1}{2N_P} \sum_{j=1}^{N_P} (\hat{z}_2(k) - \hat{\beta} \hat{z}_1(k))^2 \right) \quad (12)$$

$$\text{s.t. } \Theta = (\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} \bar{\mathbf{X}}^T \bar{\mathbf{Y}} \quad (13)$$

其中

$$\begin{aligned} \Theta &= [\bar{w}_2^{\text{output}}, \dots, \bar{w}_{N^{\text{output}}}^{\text{output}}]^T \\ \bar{\mathbf{X}} &= [\bar{\mathbf{x}}_1 \ \bar{\mathbf{x}}_2 \ \dots \ \bar{\mathbf{x}}_{N^{\text{output}}-1}]_{N_P \times (N^{\text{output}}-1)} \\ \bar{\mathbf{Y}} &= [\phi_1^{\text{output}}(y_2(1)) - \hat{\beta} \cdot \phi_1^{\text{output}}(y_1(1)), \dots, \\ & \phi_1^{\text{output}}(y_2(N_P)) - \hat{\beta} \cdot \phi_1^{\text{output}}(y_1(N_P))]_{N_P \times 1}^T \\ \bar{\mathbf{x}}_1 &= [\hat{\beta} \cdot \phi_2^{\text{output}}(y_1(1)) - \phi_2^{\text{output}}(y_2(1)), \dots, \\ & \hat{\beta} \cdot \phi_2^{\text{output}}(y_1(N_P)) - \phi_2^{\text{output}}(y_2(N_P))]_{N_P \times 1}^T \\ \bar{\mathbf{x}}_{N^{\text{output}}-1} &= [\hat{\beta} \cdot \phi_{N^{\text{output}}}^{\text{output}}(y_1(1)) - \phi_{N^{\text{output}}}^{\text{output}}(y_2(1)), \dots, \\ & \hat{\beta} \cdot \phi_{N^{\text{output}}}^{\text{output}}(y_1(N_P)) - \phi_{N^{\text{output}}}^{\text{output}}(y_2(N_P))]_{N_P \times 1}^T \end{aligned} \quad (14)$$

对式 (12) 求偏导数可得:

$$\hat{\beta} = \frac{\sum_{k=1}^{N_P} \hat{z}_2(k) \hat{z}_1(k)}{\sum_{k=1}^{N_P} \hat{z}_1(k) \hat{z}_1(k)} \quad (15)$$

将式 (15) 代入式 (12) 中可求出输出静态非线性神经模糊模型的参数。

2.2 线性环节模型

仿真实验表明, 对于输入静态非线性环节, 如果输入 $u(k)$ 为二进制信号, 静态非线性环节的特性将不会被激发, 则输出 $v(k)$ 也是二进制信号, 即 $v_1(k) = \gamma u_1^{\text{binary}}(k)$, 其中 γ 为常数, 表示输入静态非线性环节在二进制信号作用下输入输出信号的幅值关系. 因此, 可以将二进制信号 $u_1^{\text{binary}}(k)$ 看成 $v_1(k)$, 即 $v_1(k) = u_1^{\text{binary}}(k)$. 需要指出的是, 幅值的差异 γ 可以通过静态非线性环节的增益进行补偿. 所以, 本节将二进制信号源 $u_1^{\text{binary}}(k)$ 及其相应的输出 $y_1(k)$ 作为输入输出信号, 再由第 2.1 节的输出非线性神经模糊模型求出中间信号 $z_1(k)$, 最后结合最小二乘法方法求解动态线性环节的参数.

$$\hat{\theta} = [\mathbf{X}_L^T \mathbf{X}_L]^{-1} \mathbf{X}_L^T \mathbf{Z}_1 \quad (16)$$

其中

$$\begin{aligned} \hat{\theta} &= [\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{n_a}, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_{n_b}] \\ \mathbf{X}_L &= [\mathbf{x}_L(1), \mathbf{x}_L(2), \dots, \mathbf{x}_L(N_P)]_{N_P \times (n_a + n_b)}^T \\ \mathbf{Z}_1 &= [\hat{z}_1(1), \hat{z}_1(2), \dots, \hat{z}_1(N_P)]_{N_P \times 1}^T \\ \mathbf{x}_L(k) &= [\hat{z}_1(k-1), \dots, \hat{z}_1(k-n_a), \\ &u_1^{\text{binary}}(k-1), \dots, u_1^{\text{binary}}(k-n_b)]_{(n_a + n_b) \times 1}^T \end{aligned} \quad (17)$$

2.3 基于神经模糊网络的输入非线性模型

在获得线性环节和输出非线性环节后, 利用随机信号源 $u_3^{\text{random}}(k)$ 及其相应的输出 $y_3(k)$ 求解输入非线性环节的参数. 模型需满足如下条件:

$$\min E_2(\hat{f}(\cdot)) = \frac{1}{2N_P} \sum_{k=1}^{N_P} (\hat{z}_3(k) - z_3(k))^2 \quad (18)$$

下面讨论模型参数的求解算法. 首先, 对式 (18) 求一次偏导数得到

$$\frac{\partial E_2}{\partial \mathbf{W}^{\text{input}}} = \frac{1}{N_P} \sum_{k=1}^{N_P} (\hat{z}_3(k) - z_3(k)) \frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} \quad (19)$$

其中

$$\frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} = \left[\frac{\partial \hat{z}_3(k)}{\partial w_1^{\text{input}}}, \frac{\partial \hat{z}_3(k)}{\partial w_2^{\text{input}}}, \dots, \frac{\partial \hat{z}_3(k)}{\partial w_{N^{\text{input}}}^{\text{input}}} \right]^T \quad (20)$$

根据式 (1) 和式 (4) 可以得到:

$$\begin{aligned} \frac{\partial \hat{z}_3(k)}{\partial w_l^{\text{input}}} &= \frac{\partial \hat{z}_3(k)}{\partial \hat{v}_3(k-1)} \cdot \frac{\partial \hat{v}_3(k-1)}{\partial w_l^{\text{input}}} + \\ &\frac{\partial \hat{z}_3(k)}{\partial \hat{v}_3(k-2)} \cdot \frac{\partial \hat{v}_3(k-2)}{\partial w_l^{\text{input}}} + \dots + \\ &\frac{\partial \hat{z}_3(k)}{\partial \hat{v}_3(k-n_b)} \cdot \frac{\partial \hat{v}_3(k-n_b)}{\partial w_l^{\text{input}}} = \\ &\hat{b}_1 \phi_l^{\text{input}} \left(u_3^{\text{random}}(k-1) \right) + \\ &\hat{b}_2 \phi_l^{\text{input}} \left(u_3^{\text{random}}(k-2) \right) + \dots + \\ &\hat{b}_{n_b} \phi_l^{\text{input}} \left(u_3^{\text{random}}(k-n_b) \right) \end{aligned} \quad (21)$$

其中, $l = 1, 2, \dots, N^{\text{input}}$.

接着对式 (18) 求二次偏导数得到:

$$\begin{aligned} \frac{\partial^2 E_2}{\partial (\mathbf{W}^{\text{input}})^2} &= \frac{1}{N_P} \sum_{k=1}^{N_P} (\hat{z}_3(k) - z_3(k)) \frac{\partial^2 \hat{z}_3(k)}{\partial (\mathbf{W}^{\text{input}})^2} + \\ &\frac{1}{N_P} \sum_{k=1}^{N_P} \left[\frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} \right] \left[\frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} \right]^T \end{aligned} \quad (22)$$

根据式 (4) 可以得到:

$$\frac{\partial^2 \hat{v}_3(k-1)}{\partial (\mathbf{W}^{\text{input}})^2} = 0 \quad (23)$$

因此

$$\frac{\partial^2 \hat{z}_3(k)}{\partial (\mathbf{W}^{\text{input}})^2} = \frac{\partial^2 \hat{z}_3(k)}{\partial^2 \hat{v}_3(k-1)} \frac{\partial^2 \hat{v}_3(k-1)}{\partial (\mathbf{W}^{\text{input}})^2} = 0 \quad (24)$$

基于上述分析, 可以得到如下结论:

$$\frac{\partial^2 E_2}{\partial (\mathbf{W}^{\text{input}})^2} = \frac{1}{N_P} \sum_{k=1}^{N_P} \left[\frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} \right] \left[\frac{\partial \hat{z}_3(k)}{\partial \mathbf{W}^{\text{input}}} \right]^T \quad (25)$$

$$\frac{\partial^j E_2}{\partial (\mathbf{W}^{\text{input}})^j} = 0, \quad j \geq 3 \quad (26)$$

Taylor 级数展开, 由于 Taylor 公式的高阶项为零, 因此得到:

$$\begin{aligned} \frac{\partial E_2}{\partial \mathbf{W}^{\text{input}}} &= \frac{\partial E_2}{\partial \mathbf{W}^{\text{input}}} \Big|_{\mathbf{W}^{\text{input}}=0} + \\ &\frac{\partial^2 E_2}{\partial (\mathbf{W}^{\text{input}})^2} \Big|_{\mathbf{W}^{\text{input}}=0} \times \mathbf{W}^{\text{input}} \end{aligned} \quad (27)$$

令 $\frac{\partial E_2}{\partial \mathbf{W}^{\text{input}}} = 0$, 再根据式 (27), 可以得到模型参数

$$\begin{aligned} \mathbf{W}^{\text{input}} &= - \left[\frac{\partial^2 E_2}{\partial (\mathbf{W}^{\text{input}})^2} \Big|_{\mathbf{W}^{\text{input}}=0} \right]^{-1} \times \\ &\frac{\partial E_2}{\partial \mathbf{W}^{\text{input}}} \Big|_{\mathbf{W}^{\text{input}}=0} \end{aligned} \quad (28)$$

基于多信号源的神经模糊 Hammerstein-Wiener 模型分离学习算法简单描述为:

步骤 1. 采用特殊信号源 $u_1^{\text{binary}}(k)$ 和 $u_2^{\text{binary}}(k)$ 及其对应的输出 $y_1(k)$ 和 $y_2(k)$, 按式 (12)~(15) 求解输出非线性环节模型参数 $\bar{w}_i^{\text{output}} (i = 2, 3, \dots, N^{\text{output}})$;

步骤 2. 采用特殊信号源 $u_1^{\text{binary}}(k)$ 及其对应的输出 $y_1(k)$, 由步骤 1 求解得到的输出非线性环节的模型参数 $\bar{w}_i^{\text{output}} (i = 2, 3, \dots, N^{\text{output}})$ 和 $\hat{\beta}$ 求出中间信号 $z_1(k)$, 再由式 (16) 和式 (17), 求得线性环节的参数 $\hat{\theta} = [\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{n_a}, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_{n_b}]$;

步骤 3. 在步骤 1 和步骤 2 的基础上, 采用信号源 $u_3^{\text{random}}(k)$ 及其对应的输出 $y_3(k)$, 按式 (28) 求解输入非线性环节的模型参数 $w_i^{\text{input}} (i = 2, 3, \dots, N^{\text{input}})$.

注 1. 本文提出的神经模糊 Hammerstein-Wiener 模型适用于具有中等或严重非线性以及分段非线性过程.

注 2. 假设输入输出信号, $u_3^{\text{random}}(k)$ 和 $y_3(k)$ 满足持续激励条件 (PE 条件). 根据逼近定理, 总是存在一个有限向量 $\mathbf{W}^{\text{input}*}$, 使得:

$$v_3(k) = \sum_{l=1}^{N^{input}} \phi_l^{input} \left(u_3^{random}(k) \right) w_l^{input*} + d(k)$$

$$z_3(k+1) = \sum_{i=1}^{n_a} a_i z_3(k+1-i) + \sum_{j=1}^{n_b} b_j v_3(k+1-j)$$

$$y_3(k) = g(z_3(k)), \quad \forall k \in \Omega_k$$

满足

$$\mathbf{W}^{input*} = \arg \min_{\mathbf{W} \in \Omega_w} \left\{ \sup_{t \in \Omega_{k_t}} \left| \sum_{l=1}^{N^{input}} \phi_l^{input} \times \left(u_3^{random}(k) \right) w_l^{input*} - v_3(k) \right| \right\}$$

事实上, $d(k)$ 是由神经模糊网络的结构所决定的, 如果模糊规则数足够大, 那么 $d(k)$ 的值可以任意小. 因此, 当 $d(k) \equiv 0$ 的时候, 总可以找到一个无限逼近于 \mathbf{W}^{input*} 的 \mathbf{W}^{input} .

传统的神经模糊模型采用迭代方式求解模型参数, 需要分析参数的收敛性问题; 而本文提出的神经模糊 Hammerstein-Wiener 模型参数算法 (28), 由于采用 Taylor 级数展开非迭代解析法直接求解, 因此避免了模型参数的收敛性问题.

3 基于神经模糊 Hammerstein-Wiener 模型的控制系统

在得到基于神经模糊 Hammerstein-Wiener 模型的基础上, 可利用模型特殊的结构将非线性系统的控制问题简化为线性系统的控制问题. 通过 $\hat{f}^{-1}(\cdot)$ 和 $\hat{g}^{-1}(\cdot)$ 的作用将原非线性系统近似为线性系统, 如图 3 所示.

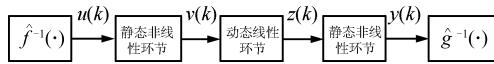


图 3 线性化模型示意图

Fig. 3 The diagrammatic sketch of linearized model

根据上述分析, 本文设计的控制系统如图 4 所示. 通过模型的特殊结构简化了控制系统的设计, 采用传统的 PID 控制器便能达到较好的控制效果.

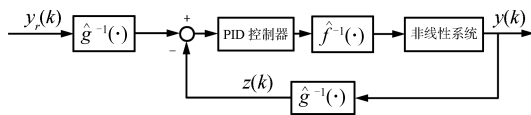


图 4 基于 Hammerstein-Wiener 模型的控制系统

Fig. 4 Control system based on Hammerstein-Wiener model

4 仿真研究

4.1 分段非线性 Hammerstein-Wiener 系统

考虑一类复杂的 Hammerstein-Wiener 非线性过程:

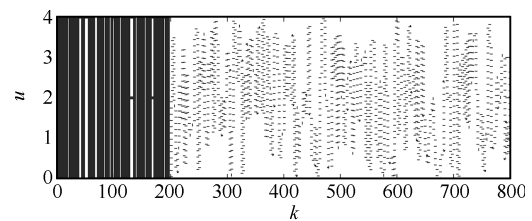
$$v(k) = \begin{cases} 2 - \cos(3u(k)) - \exp(-u(k)), & u(k) \leq 3.15 \\ 3, & u(k) > 3.15 \end{cases} \quad (29)$$

$$z(k) = 0.8z(k-1) + 0.4v(k-1) \quad (30)$$

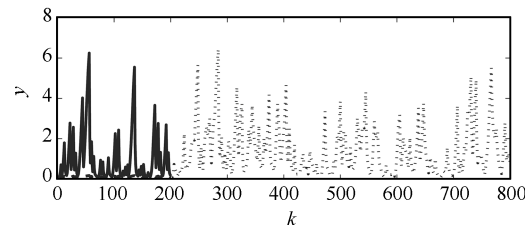
$$y(k) = \begin{cases} 0.1z(k), & z(k) \leq 2.5 \\ 0.25 \exp(z(k) - 2.5), & z(k) > 2.5 \end{cases} \quad (31)$$

分别按本文提出的方法和多项式描述法^[12] 建立 Hammerstein-Wiener 模型.

首先, 产生如图 5 所示的特殊信号源, 其中 $u_1(k)$ 是 200 组幅值为 0 或 4 二进制信号源; $u_2(k)$ 是 200 组幅值为 0 或 2 二进制信号源; $u_3(k)$ 是 600 组幅值为 0 和 4 之间的随机信号源. 采用特殊信号源 $u_1(k)$ 和 $u_2(k)$ 及其对应的输出 $y_1(k)$ 和 $y_2(k)$, 按式 (12)~(15) 求解输出非线性环节的模型参数 w_i^{output} ($i = 2, 3, \dots, 6$) 和 $\hat{\beta} = 0.3091$, 实际值为 $\beta = 0.3015$; 采用特殊信号源 $u_1(k)$ 和其对应的输出 $y_1(k)$, 求出中间信号 $z_1(k)$, 再按由式 (16) 和式 (17) 求得线性环节的参数 $\hat{\theta} = [0.8022, 0.2968]$; 最后采用信号源 $u_3^{random}(k)$ 和其对应的输出 $y_3(k)$, 按式 (28) 求解输入非线性环节的模型参数 w_i^{input} ($i = 2, 3, \dots, 8$).



(a) 特殊输入信号源
(a) Special input signals



(b) 输出信号
(b) Output signals

图 5 特殊信号源及输出信号

Fig. 5 Special input signals and output signals

由于 α 、 β 、 γ 的存在, 各串联模块的增益与真实值之间会存在倍数关系, 因此将仿真中的 Hammerstein-Wiener 非线性过程改写为

$$v(k) = \begin{cases} 1.3207 \times (2 - \cos(3u(k)) - \exp(-u(k))), & u(k) \leq 3.15 \\ 1.3207 \times 3, & u(k) > 3.15 \end{cases} \quad (32)$$

$$z(k) = 0.8z(k-1) + \frac{0.4}{1.3477}v(k-1) \quad (33)$$

$$y(k) = \begin{cases} 0.1\bar{z}(k), & \bar{z}(k) \leq 2.5 \\ 0.25 \exp(\bar{z}(k) - 2.5), & \bar{z}(k) > 2.5 \end{cases} \quad (34)$$

这里, $\bar{z}(k) = \frac{1.3477}{1.3207}z(k)$.

本文提出的方法和多项式描述法得到的 Hammerstein-Wiener 模型比较如图 6 和图 7 所示. 可以看到, 本文提出的方法在讨论域 $u(k) \in [0, 4]$ 上均能拟合原函数, 而多项式描

述的非线性只能拟合 $u(k) \in [0, 3.15]$ 范围上的原函数, 对于分段非线性其精度和适应性会下降. 因此, 本文提出的方法比多项式描述法能更好地逼近分段非线性过程.

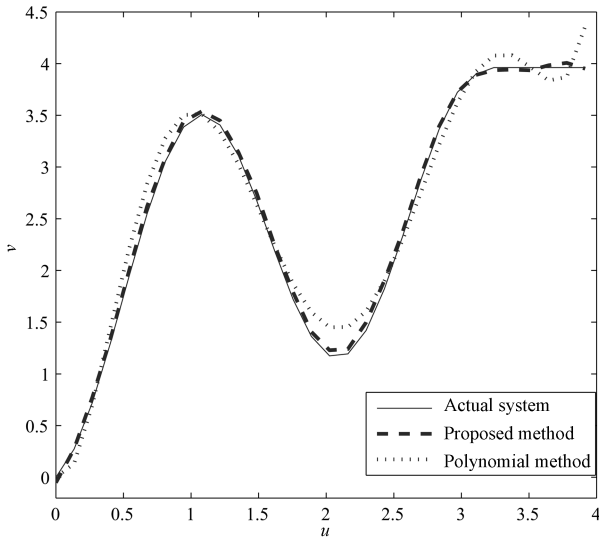


图 6 输入非线性环节
Fig. 6 Input nonlinear link

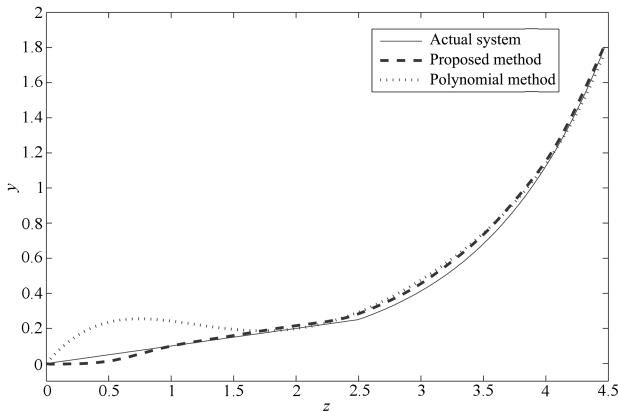


图 7 输出非线性环节
Fig. 7 Output nonlinear link

4.2 CSTR 连续反应搅拌釜过程

为进一步验证 Hammerstein-Wiener 模型的有效性, 考虑选用一个恒温的 CSTR 中的 van de Vusse 反应作为仿真研究对象^[17-18], 该过程具有动态反应: $A \rightarrow B \rightarrow C, A \rightarrow D$, 其动力学方程为

$$\frac{dC_A}{dt} = -k_1 C_A - k_3 C_A^2 + \frac{F}{V}(C_{Af} - C_A) \quad (35)$$

$$\frac{dC_B}{dt} = k_1 C_A - k_2 C_B - \frac{F}{V} C_B \quad (36)$$

该过程反应的参数和操作条件是: $k_1 = 50$ (1/h), $k_2 = 100$ (1/h), $k_3 = 10$ L/(mol · h), $C_{Af} = 10$ mol/L, $V = 1$ L, $F = 34.3$ L/h, $C_{A0} = 3.0$ mol/L, $C_{B0} = 1.12$ mol/L. 其中, F_0, C_{A0} 和 C_{B0} 分别为其一个稳态工作点对应的稳态值.

仿真中, 首先归一化数据: $u = (F - 34.3)/34.3, y = (C_B - 1.12)/1.12$, 其中 u 为输入, y 为输出. 采用复合式特殊信号源: $u_1(k)$ 是 200 组幅值为 0 或 1 二进制信号源; $u_2(k)$ 是

100 组幅值为 0 或 0.5 二进制信号源; $u_3(k)$ 是 600 组幅值为 -1 和 1 之间的随机信号源. 同仿真研究 (1), 可得到 $\hat{w}_i^{\text{output}}$ ($i = 2, 3, \dots, 6$), $\hat{\beta} = 0.6300, w_i^{\text{input}}$ ($i = 2, 3, \dots, 7$).

在获得的 Hammerstein-Wiener 模型基础上设计控制系统, 如图 3 所示, 线性控制器采用传统的 PI 控制器:

$$u(k) = u(k-1) + K_c \left[e(k) - e(k-1) + \frac{e(k)}{\tau_I} \right] \quad (37)$$

其中, $e(k) = y_r(k) - y(k)$, $y_r(k)$ 为控制系统的设定值.

设反应从原稳态工作点 $C_{B0} = 1.12$ mol/L 分别跃变到新的工作点 $C_B = 1.232$ mol/L 和 $C_B = 0.56$ mol/L, 仿真中 PI 控制器参数为 $K_c = 2.2, \tau_I = 9$. 为了进一步说明基于 Hammerstein-Wiener 模型的 PI 控制系统的优越性, 将其和文献 [17] 中的线性 PI 控制器比较, 比较结果如图 8 和图 9 所示. 可以看出, 对于该 CSTR 连续反应搅拌釜过程线性 PI 控制器不能取得较好的控制效果, 无论怎么调节 PI 的参数, 都不能平衡好设定点 -0.5 和 0.1 的控制效果. 具体来讲, 在设定点 -0.5 控制系统表现出震荡现象, 而在设定点 0.1 控制系统表现出迟缓现象. 如果为了避免反应从原稳态工作点 $C_{B0} = 1.12$ mol/L 跃变到新的工作点 $C_B = 0.56$ mol/L 时

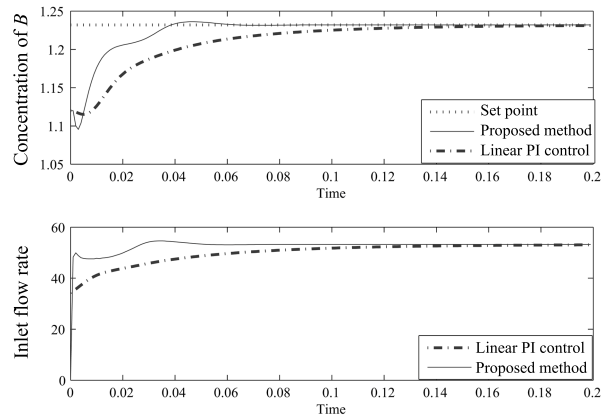


图 8 工作点在 1.132 mol/L 时的情况
Fig. 8 The work point in 1.132 mol/L

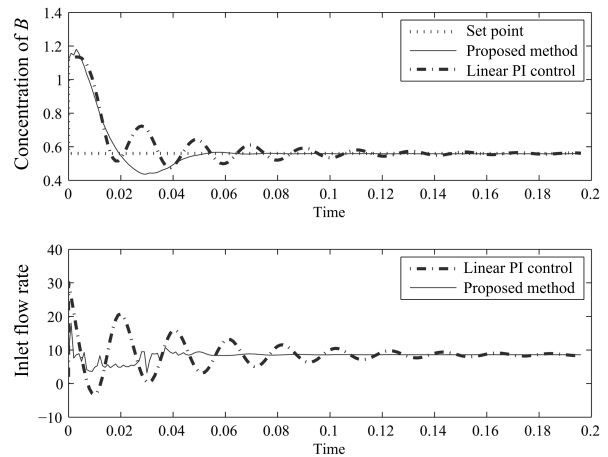


图 9 工作点在 0.56 mol/L 时的情况
Fig. 9 The work point in 0.56 mol/L

的震荡现象而重新设置一组保守的 PI 控制器参数, 必将导

致反应从原稳态工作点 $C_{B0} = 1.12 \text{ mol/L}$ 跃变到新的工作点 $C_B = 1.232 \text{ mol/L}$ 时的响应更加缓慢。即, 如果要加速后者的响应速度必将以前者的不稳定为代价。相比之下, 本文提出的基于 Hammerstein-Wiener 模型的 PI 控制系统能够平衡好两个工作点的响应, 具有较好的控制效果。

5 结论

本文提出了一种基于多信号源的 Hammerstein-Wiener 模型辨识方法, 该方法利用特殊的信号源把模型的非线性环节和线性环节分离开来, 大大简化了辨识过程, 然后用神经模糊模型拟合输入非线性和输出非线性环节, 该方法适用范围广、计算简单、辨识精度高、对具有分段非线性特性的系统具有较好的拟合精度。仿真结果验证了上述辨识方法对于分段非线性 Hammerstein-Wiener 系统的有效性。更为重要的是, 本文提出的方法把原非线性系统的控制问题分解为线性模块的动态优化问题和非线性模块的静态映射问题, 因此可以有效结合线性模型预测控制的成熟理论解决非线性过程的控制问题, 避免了传统非线性控制方法计算量大、收敛性和闭环稳定性不能得到保证等诸多问题。

References

- Liu Y, Bai E W. Iterative identification of Hammerstein systems. *Automatica*, 2007, **43**(2): 346–354
- Gui Wei-Hua, Song Hai-Ying, Yang Chun-Hua. Hammerstein-Wiener model identified by least-squares-support-vector machine and its application. *Control Theory and Applications*, 2008, **25**(3): 393–397
(桂卫华, 宋海鹰, 阳春华. Hammerstein-Wiener 模型最小二乘向量机辨识及其应用. 控制理论与应用, 2008, **25**(3): 393–397)
- Marzban H R, Tabrizidooz H R, Razzaghi M. A composite collocation method for the nonlinear mixed Volterra-Fredholm-Hammerstein integral equations. *Communications in Nonlinear Science and Numerical Simulation*, 2011, **16**(3): 1186–1194
- Chen H T, Hwang S H, Chang C T. Iterative identification of continuous-time Hammerstein and Wiener systems using a two-stage estimation algorithm. *Industrial and Engineering Chemistry Research*, 2009, **48**(3): 1495–1510
- Wang Feng, Xing Ke-Yi, Xu Xiao-Ping. Study on method for identification of Hammerstein model. *Journal of System Simulation*, 2011, **23**(6): 1090–1092, 1136
(王峰, 邢科义, 徐小平. 辨识 Hammerstein 模型方法研究. 系统仿真学报, 2011, **23**(6): 1090–1092, 1136)
- Ding Feng, Liu Jing-Fan, Xiao Yong-Song. Parameter estimation for a class of nonlinear systems. *Control of Engineering China*, 2011, **18**(3): 373–376, 409
(丁峰, 刘景璠, 肖永松. 一类非线性系统的参数估计. 控制工程, 2011, **18**(3): 373–376, 409)
- 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)
- Bai E W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 1998, **34**(3): 333–338
- Wang D Q, Ding F. Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems. *Computers and Mathematics with Applications*, 2008, **56**(12): 3157–3164
- Zhu Y C. Estimation of an N-L-N Hammerstein-Wiener model. *Automatica*, 2002, **38**(9): 1607–1614
- Crama P, Schoukens J. Hammerstein-Wiener system estimator initialization. *Automatica*, 2004, **40**(9): 1543–1550
- Park H C, Sung S W, Lee J. Modeling of Hammerstein-Wiener processes with special input test signals. *Industrial and Engineering Chemistry Research*, 2006, **45**(3): 1029–1038
- Man Hong, Shao Cheng. Neural network predictive control of continuous stirred-tank reactor based on Hammerstein-Wiener model. *CIESC Journal*, 2011, **62**(8): 2275–2280
(满红, 邵诚. 基于 Hammerstein-Wiener 模型的连续搅拌反应釜神经网络预测控制. 化工学报, 2011, **62**(8): 2275–2280)
- Li Yan, Mao Zhi-Zhong, Wang Yan, Yuan Ping. Fuzzy predictive control of Hammerstein-Wiener nonlinear systems. *Journal of Northeastern University (Natural Science)*, 2011, **32**(3): 322–326
(李妍, 毛志忠, 王琰, 袁平. Hammerstein-Wiener 非线性系统的模糊预测控制. 东北大学学报(自然科学版), 2011, **32**(3): 322–326)
- Li Yan, Mao Zhi-Zhong, Wang Yan, Yuan Ping, Jia Ming-Xing. Predictive control of Hammerstein-Wiener nonlinearity based on polytopic terminal region. *Acta Automatica Sinica*, 2011, **37**(5): 629–638
(李妍, 毛志忠, 王琰, 袁平, 贾明兴. 基于多面体终端域的 Hammerstein-Wiener 非线性预测控制. 自动化学报, 2011, **37**(5): 629–638)
- Wang Wei, Chai Tian-You, Zhao Li-Jie. Dynamic partial least squares modeling with recurrent neural networks of stable learning. *Control Theory and Applications*, 2012, **29**(3): 337–341
(王魏, 柴天佑, 赵立杰. 带有稳定学习的递归神经网络动态偏最小二乘建模. 控制理论与应用, 2012, **29**(3): 337–341)
- Jia L, Chiu M S, Ge S S. A noniterative neuro-fuzzy based identification method for Hammerstein processes. *Journal of Process Control*, 2005, **15**(7): 749–761
- Hahn J, Edgar T F. A Gramian based approach to nonlinearity quantification and model classification. *Industrial and Engineering Chemistry Research*, 2001, **40**(24): 5724–5731

贾立 上海大学机电工程与自动化学院教授。主要研究方向为复杂非线性系统的建模、优化与控制。本文通信作者。

E-mail: jiali@staff.shu.edu.cn

(JIA Li Professor at the School of Mechatronic Engineering and Automation, Shanghai University. Her research interest covers modeling and optimization controlling of complex nonlinear system. Corresponding author of this paper.)

杨爱华 上海大学机电工程与自动化学院硕士研究生。主要研究方向为 Hammerstein-Wiener 模型。E-mail: yangaihua@shu.edu.cn

(YANG Ai-Hua Master student at the School of Mechatronic Engineering and Automation, Shanghai University. His main research interest is Hammerstein-Wiener model.)

邱铭森 新加坡国立大学工程学院副教授。主要研究方向为鲁棒控制, 分散控制, 间歇过程优化控制。E-mail: checms@nus.edu.sg

(CHIU Min-Sen Associate professor at the Faculty of Engineering, National University of Singapore, Singapore. His research interest covers robust control, decentralized control, and optimization controlling for batch process.)