

文章编号: 1001-0920(2006)02-0139-04

## 一类非线性多变量系统的多模型自适应解耦控制

富月, 柴天佑, 岳恒

(东北大学 自动化研究中心, 沈阳 110004)

**摘要:** 针对一类多变量离散时间非线性动态系统, 分别设计线性鲁棒自适应解耦控制律和神经网络非线性自适应解耦控制律。根据指定的性能指标, 通过它们之间的切换对系统进行控制。理论分析和仿真结果表明, 该控制策略不但可以保证闭环系统 BIBO 稳定, 而且能够改善系统的性能。

**关键词:** 多模型; 自适应; 解耦; 神经网络; 切换

**中图分类号:** TP273

**文献标识码:** A

## Multiple Models Adaptive Decoupling Control for a Class of Nonlinear Multivariable Systems

FU Yue, CHAI Tian-you, YUE Heng

(Research Center of Automation, Northeastern University, Shenyang 110004, China Correspondent: FU Yue, Email: fu\_yue1004@163.com)

**Abstract:** A linear robust adaptive decoupling control law and a neural network nonlinear adaptive decoupling control law are designed respectively for a class of multivariable discrete time nonlinear dynamic systems. Via switching between these control laws by a suitably defined switching law, the system is controlled. Theory analysis and simulation results show that by using the control scheme, the BIBO stability of the closed-loop system and the improved performance can be achieved simultaneously.

**Key words:** Multiple models; Adaptive; Decoupling; Neural network; Switching

### 1 引言

在复杂的工业过程中, 大多数工业对象都是非线性、多变量、强耦合的动态系统。文献[1~3]针对这种系统将其在原点附近 Taylor 展开, 分别采用神经网络前馈补偿与反馈控制相结合以及多模型的方法, 实现了对系统的有效控制, 但缺乏系统的稳定性分析。文献[4, 5]针对单变量非线性系统, 采用多模型自适应控制方法, 通过模型切换不仅保证了系统输入输出信号有界, 而且改善了系统的跟踪性能。但该方法除了限于单变量系统外, 还要求系统在原点附近零动态渐近稳定。

本文针对一类非线性多变量离散时间动态系统, 将多变量前馈自适应解耦控制<sup>[6, 7]</sup>与多模型方

法<sup>[4]</sup>相结合, 提出一种非线性多变量自适应解耦控制算法, 并证明了该算法能够保证闭环系统 BIBO 稳定。

### 2 被控系统描述

设  $n$ - 输入  $n$ - 输出离散时间非线性系统如下:

$$y(t+1) = f[y(t), \dots, y(t-n_a+1), u(t), \dots, u(t-n_b)] \quad (1)$$

其中:  $u(t), y(t) \in R^n$  分别为输入和输出向量;  $f[\bullet] = [f_1[\bullet], \dots, f_n[\bullet]]^T$  是  $n$  维连续可微的非线性向量函数;  $n_a$  和  $n_b$  为系统的阶次; 原点为平衡点。

考虑到研究系统在平衡点附近的线性化模型的意义, 将系统(1)在原点附近转化为如下等价模型<sup>[8, 9]</sup>:

收稿日期: 2004-07-19; 修回日期: 2005-02-25

基金项目: 国家 973 计划项目(2002CB312201); 国家 863 计划项目(2004AA412030)。

作者简介: 富月(1978—), 女, 辽宁开原人, 博士生, 从事智能解耦、多模型智能控制的研究; 柴天佑(1947—), 男, 兰州人, 中国工程院院士, 教授, 博士生导师, 从事自适应控制、智能控制等研究。

$$\begin{aligned}
 A(z^{-1})y(t+1) = \\
 B(z^{-1})u(t) + v[y(t), \dots \\
 y(t-n_a+1), u(t), \dots, u(t-n_b)] \quad (2)
 \end{aligned}$$

其中:  $A(z^{-1})$  和  $B(z^{-1})$  是  $z^{-1}$  的  $n \times n$  矩阵多项式, 阶次分别为  $n_a$  和  $n_b$ ;  $v[\bullet]$  为高阶非线性项  
系统假设如下:

- 1) 构成  $A(z^{-1})$  和  $B(z^{-1})$  的参数矩阵在紧集  $\Sigma$  中变化, 并且  $A(z^{-1})$  为首一对角矩阵多项式;
- 2) 系统在原点附近开环稳定(即系统输出序列的增长不超过输入序列);
- 3) 非线性部分  $v[\bullet]$  全局有界, 即  $v[\bullet] \leq \Delta$ , 其中  $\Delta > 0$  且已知<sup>[4,5]</sup>.

### 3 线性鲁棒自适应解耦控制及稳定性分析

为进行解耦设计, 将式(2) 改写为

$$\begin{aligned}
 A(z^{-1})y(t+1) = \\
 \bar{B}(z^{-1})u(t) + \bar{\bar{B}}(z^{-1})u(t) + v[\bullet] \quad (3)
 \end{aligned}$$

其中:  $\bar{B}(z^{-1})$  为对角矩阵多项式,  $\bar{\bar{B}}(z^{-1})$  为对角元为零的矩阵多项式, 满足  $B(z^{-1}) = \bar{B}(z^{-1}) + \bar{\bar{B}}(z^{-1})$ .

由于  $v[\bullet]$  有界, 式(3) 可近似为一个带有界干扰的线性模型 控制器设计模型如下:

$$\begin{aligned}
 A(z^{-1})y(t+1) = \\
 \bar{B}(z^{-1})u(t) + \bar{\bar{B}}(z^{-1})u(t). \quad (4)
 \end{aligned}$$

易知在模型(4) 下, 使性能指标

$$\begin{aligned}
 J_L = P(z^{-1})y(t+1) - R w(t) + \\
 Q(z^{-1})u(t) + S u(t)^2 \quad (5)
 \end{aligned}$$

最小的最优控制律为

$$\begin{aligned}
 [F\bar{B}(z^{-1}) + Q(z^{-1})]u(t) + \\
 [F\bar{\bar{B}}(z^{-1}) + S]u(t) = \\
 R w(t) - G(z^{-1})y(t). \quad (6)
 \end{aligned}$$

其中:  $w(t) \in R^n$  是已知有界的参考输入向量,  $P(z^{-1})$  和  $Q(z^{-1})$  是  $n \times n$  对角加权矩阵多项式,  $R$  是  $n \times n$  对角加权阵,  $S$  是  $n \times n$  对角元为零的加权阵,  $F$  是  $n \times n$  常数阵,  $G(z^{-1})$  是  $n_a - 1$  阶的  $n \times n$  矩阵多项式, 由下式唯一确定:

$$P(z^{-1}) = F A(z^{-1}) + z^{-1} G(z^{-1}). \quad (7)$$

当系统参数未知时进行间接自适应控制, 系统参数辨识方程为

$$y_1(t) = \Theta^T X(t-1). \quad (8)$$

其中

$$\begin{aligned}
 \Theta_1 = [\theta_1, \theta_2, \dots, \theta_n] = \\
 [-A_1, \dots, -A_{n_a}, B_0, \dots, B_{n_b}]^T, \\
 X(t-1) = \\
 [y^T(t-1), \dots, y^T(t-n_a), \\
 u^T(t-1), \dots, u^T(t-n_b-1)]^T.
 \end{aligned}$$

对  $\theta_i (i = 1, 2, \dots, n)$  采用如下辨识算法<sup>[4]</sup>:

$$\begin{aligned}
 \hat{\theta}_i(t) = \\
 \hat{\theta}_i(t-1) + \frac{a_{1i}(t)X(t-1)e_{1i}(t)}{1 + X(t-1)^T X(t-1)}, \quad (9)
 \end{aligned}$$

$$a_{1i}(t) = \begin{cases} 1, & |e_{1i}(t)| > 2\Delta; \\ 0, & \text{else} \end{cases} \quad (10)$$

其中辨识误差

$$\begin{aligned}
 e_{1i}(t) = y_i(t) - \hat{y}_{1i}(t) = \\
 y_i(t) - X(t-1)^T \hat{\theta}_i(t-1) = \\
 X(t-1)^T (\theta_i - \hat{\theta}_i(t-1)) + v_i[X(t-1)] \quad (11)
 \end{aligned}$$

设  $t$  时刻的系统参数阵分别为  $A_{1i}(z^{-1})$  和  $B_{1i}(z^{-1})$ , 则线性自适应解耦控制律为

$$\begin{aligned}
 [F_1 \bar{B}_{1i}(z^{-1}) + Q(z^{-1})]u(t) + \\
 [F \bar{B}_{1i}(z^{-1}) + S]u(t) = \\
 R w(t) - G_{1i}(z^{-1})y(t). \quad (12)
 \end{aligned}$$

其中  $F_{1i}, G_{1i}(z^{-1}), Q(z^{-1}), S$  为控制器参数阵

**定理 1** 系统满足假设条件 1) ~ 3), 则通过离线选择  $P(z^{-1})$  和  $Q(z^{-1})$ , 使

$$\det\{P(z^{-1})B(z^{-1}) + A(z^{-1})[Q(z^{-1}) + S(z^{-1})]\} \neq 0, |z| > 1, \quad (13)$$

且当自适应算法(9) ~ (12) 用于系统(2) 时, 系统 BBO 稳定

**证明** 采用反证法及文献[4] 中定义的  $O[\bullet]$  (Large order) 和  $o[\bullet]$  (Small order).

假设存在  $k (0 < k < n)$  使  $u_k(t)$  无界 对于任意的  $i (0 < i < n)$ , 定义  $\theta_i(t) = \hat{\theta}_i(t) - \theta_i$ , 采用与文献[4] 类似的方法, 可得

$$\begin{aligned}
 |\theta_i(t)|^2 - |\theta_i(t-1)|^2 = \\
 \frac{a_{1i}(t)(e_{1i}(t))^2 - 4\Delta^2}{2(1 + X(t-1)^T X(t-1))}.
 \end{aligned}$$

由此可知,  $\{|\theta_i(t)|^2\}$  是一个单调非增序列,  $\hat{\theta}_i(t)$  有界, 并且

$$\lim_{t \rightarrow \infty} \frac{a_{1i}(t)(e_{1i}(t))^2 - 4\Delta^2}{2(1 + X(t-1)^T X(t-1))} = 0$$

因此若  $u_k(t)$  无界, 则  $e_{1i}(t) = o[X(t-1)]$ , 故  $\max_{1 \leq i \leq n} |e_{1i}(t)| = o[X(t-1)]$ . 由  $X(t-1)$  的表达式以及假设 2) 可知  $X(t-1) = O[u_k(t)]$ .

由式(11) 和(12) 得

$$\begin{aligned}
 F_{1i}(z^{-1})e_{1i}(t+1) = \\
 P(z^{-1})y(t+1) - R(z^{-1})w(t) + \\
 Q(z^{-1})u(t) + S(z^{-1})u(t).
 \end{aligned}$$

将其与系统(2) 联立, 由式(13) 及  $w(t)$  和  $v[\bullet]$  有界, 可知  $u_k(t) = O[\max_{1 \leq i \leq n} |e_{1i}(t)|]$ . 因此  $X(t-1) = O[\max_{1 \leq i \leq n} |e_{1i}(t)|]$  成立, 这与  $\max_{1 \leq i \leq n} |e_{1i}(t)| = o[X(t-1)]$  矛盾, 即与  $u_k(t)$  矛盾, 故系统 BBO 稳定

#### 4 神经网络自适应解耦控制

线性解耦控制律无法消除有界非线性项  $v[\bullet]$  对系统的干扰, 为提高系统的性能, 采用模型(3) 设计神经网络自适应解耦控制律

在模型(3) 下使性能指标

$$J_N = P(z^{-1})y(t+1) - Rw(t) + Q(z^{-1})u(t) + Su(t) + Kv[\bullet]^2 \quad (14)$$

最小的最优控制律<sup>[10]</sup> 为

$$[F\bar{B}(z^{-1}) + Q(z^{-1})]u(t) + [F\bar{B}(z^{-1}) + S]u(t) = Rw(t) - G(z^{-1})y(t) - [F + K]v[\bullet] \quad (15)$$

其中:  $w(t), P(z^{-1}), Q(z^{-1}), R, S$  如前所述,  $K$  为  $n \times n$  对角加权阵

将式(15) 作用于系统(3), 可得

$$[P(z^{-1})\bar{B}(z^{-1}) + Q(z^{-1})A(z^{-1})]y(t+1) = \bar{B}(z^{-1})Rw(t) + [Q(z^{-1})\bar{B}(z^{-1}) - \bar{B}(z^{-1})S]u(t) + [Q(z^{-1}) - \bar{B}(z^{-1})K]v[\bullet] \quad (16)$$

同时得到闭环系统方程

$$\{P(z^{-1})B(z^{-1}) + A(z^{-1})[Q(z^{-1}) + S]\}u(t) = A(z^{-1})Rw(t) - [A(z^{-1})K + P(z^{-1})]v[\bullet] \quad (17)$$

为使闭环系统稳定并达到稳态解耦和消除稳态误差,  $P(z^{-1}), Q(z^{-1}), R, S, K$  除了满足式(13) 外, 还应满足以下几个式子:

$$[P(1)\bar{B}(1) + Q(1)A(1)] = \bar{B}(1)R, \quad (18)$$

$$Q(1)\bar{B}(1) = \bar{B}(1)S, \quad (19)$$

$$Q(1) = \bar{B}(1)K. \quad (20)$$

当系统参数未知时进行间接自适应控制, 系统参数辨识方程为

$$y_2(t) = \Theta_2^T X(t-1) + v[X(t-1)], \quad (21)$$

其中

$$\Theta_2 = \Theta_1 = [\theta_{21}, \theta_{22}, \dots, \theta_{2n}]$$

对于  $\theta_i (i = 1, 2, \dots, n)$  除了要求  $\hat{\theta}_i$  属于紧集  $\Sigma$ , 并未对其算法进行任何限制  $v[X(t)]$  采用BP神经网络加以估计, 即

$$v[x(t)] = NN[W, \bar{X}(t)] \quad (22)$$

其中:  $NN[\bullet]$  表示网络结构,  $W$  为权矩阵,  $\bar{X}(t)$  为网络的输入量, 并且假设  $v[\bullet]$  有界

设  $A(z^{-1})$  和  $B(z^{-1})$  在  $t$  时刻的估计分别为  $A_{2t}(z^{-1})$  和  $B_{2t}(z^{-1})$ , 则神经网络自适应解耦控制律为

$$[F_2\bar{B}_{2t}(z^{-1}) + Q(z^{-1})]u(t) + [F_2\bar{B}_{2t}(z^{-1}) + S]u(t) = Rw(t) - G_{2t}(z^{-1})y(t) -$$

$$[F_{2t} + K]v[X(t)] \quad (23)$$

其中  $F_{2t}, G_{2t}(z^{-1}), P(z^{-1}), Q(z^{-1}), R, S, K$  为控制器参数阵

辨识误差

$$e_{2i}(t) = y_i(t) - \hat{y}_{2i}(t) = y_i(t) - (X(t-1)^T \hat{\Theta}_{2i}(t-1) + \hat{v}_i[X(t-1)]) \quad (24)$$

#### 5 多模型自适应解耦控制及稳定性分析

类似于文献[4], 定义切换函数准则

$$J_j(t) = \frac{c \prod_{i=1}^n (e_{ji}^2(t) - 4\Delta^2)}{2(1 + X(t-1)^T X(t-1))} + \prod_{i=1}^n (1 - a_{ji}(t)) e_{ji}^2(t), \quad (25)$$

$$a_{ji}(t) = \begin{cases} 1, & |e_{ji}(t)| > 2\Delta; \\ 0, & \text{else} \end{cases} \quad (26)$$

其中:  $e_{ji} = y_i(t) - y_{ji}(t), i = 1, 2, \dots, n, j = 1, 2; N$  是正整数,  $c$  是大于等于 0 的常数

比较  $J_1(t)$  和  $J_2(t)$ , 求出最小的  $J^*(t)$ , 选择与  $J^*(t)$  对应的自适应解耦控制律作为系统的控制输入, 即  $u(t) = u^*(t)$ .

**定理 2** 系统满足假设条件 1) ~ 3), 则通过离线选择  $P(z^{-1})$  和  $Q(z^{-1})$ , 使其满足式(13), 且当自适应算法(9) ~ (12), (22) ~ (24), 以及切换准则(25) 和(26) 用于系统(2) 时, 系统BBO 稳定

**证明** 采用反证法 假设存在  $k (0 < k < n)$  使  $u_k(t)$  无界 每一时刻系统的辨识误差

$$e_i(t) = e_{1i}(t) \text{ or } e_i(t) = e_{2i}(t), \quad i = 1, 2, \dots, n.$$

由式(26) 可知, 切换函数(25) 的第 2 项是有界的 根据定理 1 的证明过程可知  $J_1(t)$  是有界的, 并且  $e_{1i}(t) = o[X(t-1)]$  对于  $J_2(t)$  有两种情况:

1) 当  $J_2(t)$  有界时, 有  $e_{2i}(t) = o[X(t-1)]$ , 因此系统辨识误差  $e_i(t) = o[X(t-1)]$ ;

2) 当  $J_2(t)$  无界时, 由于  $J_1(t)$  有界, 存在时刻  $t_0$ , 使得当  $t > t_0$  时有  $J_1(t) < J_2(t)$ , 因此当  $t > t_0 + 1$  时, 系统辨识误差  $e_i(t) = o[X(t-1)]$

无论是 1) 还是 2), 都有  $e_i(t) = o[X(t-1)]$  其余部分的证明类似于定理 1, 这里不再赘述

#### 6 仿真实验

考虑如下两输入两输出非线性系统:

$$\begin{aligned} y_1(t+1) &= -0.2y_1(t) + \sin(u_1(t)) + 1.5\sin(u_2(t)) - \\ &0.5u_1(t-1) - 0.8u_2(t-1) + \frac{u_2(t-1)}{1+u_2(t-1)^2}, \\ y_2(t+1) &= \end{aligned}$$

$$Q \ 6y(t_2) + Q \ 2\sin(u_1(t)) + u_2(t) + \\ Q \ 3u_1(t-1) + 1.5u_2(t-1) + \frac{u_1(t-1)}{1+u_1(t-1)^2}$$

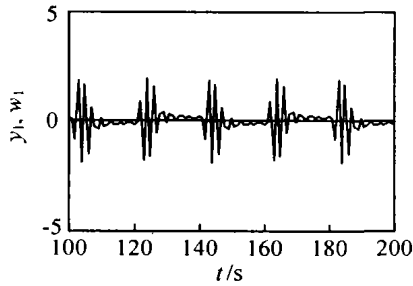
易知该系统在原点附近开环稳定

参考输入

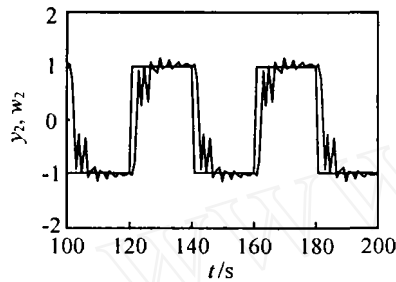
$$w_1 = 0, w_2 = \text{sign}(\sin(\pi t/20)).$$

控制器加权阵

$$Q(z^{-1}) = \begin{bmatrix} 2z^{-1} & 0 \\ 0 & 3.8 \end{bmatrix}, P(z^{-1}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

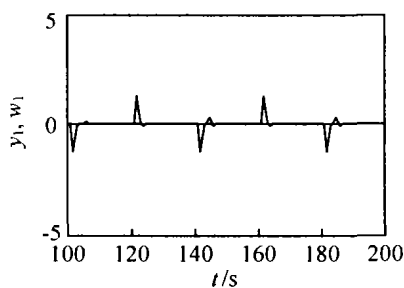


(a)  $y_1, w_1$

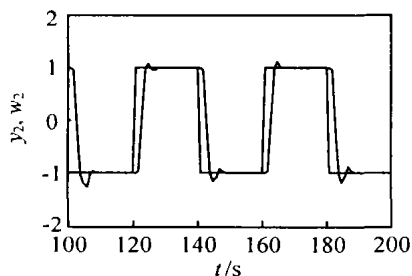


(b)  $y_2, w_2$

图1 线性自适应解耦控制时系统的输出



(a)  $y_1, w_1$



(b)  $y_2, w_2$

图2 多模型自适应解耦控制时系统的输出

图1为仅用线性鲁棒自适应解耦控制律时系统的输出响应。从图中可以看出,解耦效果不够显著,并且跟踪性能较差。这是由于线性解耦控制器缺乏补偿非线性的能力,同时由于非线性部分的存在,系统无法进入稳态。因此在控制非线性系统时,线性控制器具有一定的局限性。

图2为采用本文提出的多模型自适应解耦控制器时系统的输出响应。单隐层BP神经网络含有32个隐元,在线训练权值。从图中可以看出,系统的解耦效果有了一定的改善,跟踪性能也得到了明显的提高。

## 7 结论

本文将多变量前馈自适应解耦控制与多模型自适应控制算法相结合,提出一种非线性多变量系统的多模型自适应解耦控制方法,取消了系统零动态渐近稳定的要求,证明了闭环系统的BBO稳定性。从仿真结果可以看出,不但解耦效果显著,而且与线性鲁棒自适应解耦控制相比,多模型自适应解耦控制具有良好的控制效果。

## 参考文献(References)

- [1] Yue H, Chai T Y. Adaptive Decoupling Control of Multivariable Nonlinear Nonminimum Phase Systems Using Neural Networks[A]. Proc of the American Control Conf [C]. Philadelphia, 1998: 513-514.
- [2] 岳恒,柴天佑.一类多变量非线性系统的神经网络自适应解耦控制[J].信息与控制,1999,28(7):607-612. (Yue H, Chai T Y. Neural Network Adaptive Decoupling Control of Multivariable Nonlinear Systems[J]. Information and Control, 1999, 28(7): 607-612.)
- [3] 王昕,李少远,岳恒.一类非线性系统的多模型神经网络解耦控制器[J].控制与决策,2004,19(4):424-428. (Wang X, Li S Y, Yue H. Multiple Models Neural Network Decoupling Controller for a Nonlinear Systems[J]. Control and Decision, 2004, 19(4): 424-428.)
- [4] Chen L J, Narendra K S. Nonlinear Adaptive Control Using Neural Networks and Multiple Models[J]. Automatica, 2001, 37: 1245-1255.
- [5] Chen L J, Narendra K S. Intelligent Control Using Multiple Neural Networks[J]. Int J of Adaptive Control and Signal Process, 2003, 17: 417-430.
- [6] Chai T Y. A Self-tuning Decoupling Controller for a Class of Multivariable Systems and Global Convergence Analysis[J]. IEEE Trans on Automatic Control, 1988, 33(8): 767-771.
- [7] 柴天佑.多变量间接自适应解耦控制算法[J].自动化学报,1991,17(2):182-190.

(下转第147页)

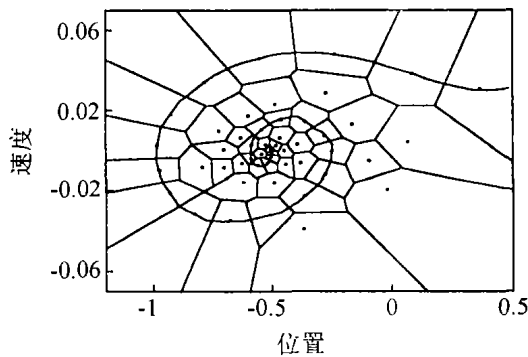


图 5 基于  $K$ -均值聚类的强化学习方法的学习结果

## 5 结 论

本文使用  $K$ -均值聚类算法实现连续状态空间的自适应离散化. 与替代合适迹 Sarsa 学习算法相结合, 得到基于  $K$ -均值聚类的强化学习方法. 该方法包含状态空间学习和策略学习两个过程. 前者决定了后者用于决策的 cost-to-go 函数估计的准确程度; 后者则影响前者训练数据的分布. 因此在选择参数时, 要通过尝试选择最佳参数. 将其应用于解决连续状态的 MountainCar 问题, 仿真实验表明, 该方法能实现状态空间的自适应离散化, 并学习到最优策略. 与基于 CMAC 网络的强化学习方法<sup>[12]</sup>进行比较, 学习效果相当, 但对存储空间需求更少, 计算时间更短, 实现更方便.

## 参考文献(References)

- [1] Sutton R S, Barto A G. *Reinforcement Learning: An Introduction* [M]. Cambridge: MIT Press, 1998
- [2] 高阳, 陈世福, 陆鑫. 强化学习研究综述[J]. *自动化学报*, 2004, 30(1): 86-100  
(Gao Y, Chen S F, Lu X. Research on Reinforcement Learning Technology: A Review [J]. *Acta Automatica Sinica*, 2004, 30(1): 86-100)
- [3] Moore AW, Atkeson C G. The Partitioned Algorithm for Variable Resolution Reinforcement Learning in Multidimensional State-spaces [J]. *Machine Learning*, 1995, 21(3): 199-233
- [4] Uther W T, Veloso M M. Tree Based Discretization

for Continuous State Space Reinforcement Learning [A]. *AAAI'98[C]*. Madison, 1998: 769-774

- [5] Smith A J. Applications of the Self-organising Map to Reinforcement Learning [J]. *Neural Networks*, 2002, 15(8-9): 1107-1124
- [6] 晏雄伟, 邓志东, 孙增圻. 竞争式 Takagi-Sugeno 模糊再励学习[J]. *自动化学报*, 2002, 28(6): 873-880  
(Yan X W, Deng Z D, Sun Z Q. Competitive Takagi-Sugeno Fuzzy Reinforcement Learning [J]. *Acta Automatica Sinica*, 2002, 28(6): 873-880)
- [7] 路兆梅, 匡文生. 自组织增强学习模糊神经网络控制器的设计[J]. *东南大学学报*, 1999, 29(4): 109-112  
(Lu Z M, Kuang W S. Design of Fuzzy Neural Network Controller Used Reinforcement Learning [J]. *J of Southeast University*, 1999, 29(4): 109-112)
- [8] 马勇, 杨煜普, 许晓鸣, 等. 一类再励学习控制器设计及其在倒车模型中的应用[J]. *上海交通大学学报*, 2000, 34(12): 1661-1663  
(Ma Y, Yang Y P, Xu X M, et al. Design of the Controller Based on Reinforcement Learning and Its Application on Truck Backer-upper [J]. *J of Shanghai Jiaotong University*, 2000, 34(12): 1661-1663)
- [9] Lee I S K, Lau H Y K. Adaptive State Space Partitioning for Reinforcement Learning [J]. *Engineering Applications of Artificial Intelligence*, 2004, 17(6): 577-588
- [10] 李春贵, 吴沧浦, 刘永信. 一种基于状态聚类的 SARSA ( $\lambda$ ) 强化学习算法[J]. *计算机工程*, 2003, 29(5): 37-98  
(Li C G, Wu C P, Liu Y X. SARSA ( $\lambda$ ) Algorithm of Reinforcement Learning Based on States Clustering [J]. *Computer Engineering*, 2003, 29(5): 37-98)
- [11] Haykin S. *Neural Networks: A Comprehensive Foundation* [M]. Beijing: Tsinghua University Press, 2001.
- [12] Singh S P, Sutton R S. Reinforcement Learning with Replacing Eligibility Traces [J]. *Machine Learning*, 1996, 22(2): 123-158

(上接第 142 页)

- (Chai T Y. Multivariable Indirect Adaptive Decoupling Control Algorithm [J]. *Acta Automatica Sinica*, 1981, 17(2): 182-190
- [8] Cabrera J B D, Narendra K S. Issues in the Application of Neural Networks for Tracking Based on Inverse Control [J]. *IEEE Transactions on Automatic Control*, 1999, 44(11): 2007-2027.
- [9] Chen L, Narendra K S. Identification and Control of a

Nonlinear Dynamical System Based on Its Linearization: Part II [A]. *Proc of the American Control Conf [C]*. Florida, 2002: 382-387.

- [10] 柴天佑. *多变量自适应解耦控制及应用* [M]. 北京: 科学出版社, 2001.  
(Chai T Y. *Multivariable Adaptive Decoupling Control and Its Applications* [M]. Beijing: Science Press, 2001.)