

# 神经网络异步自学习控制系统<sup>1)</sup>

邓志东 孙增圻

(清华大学计算机系智能技术与系统国家重点实验室 北京 100084)

刘建伟

(哈尔滨工业大学控制工程系 哈尔滨 150006)

## 摘 要

基于异步自学习控制方法,利用前馈网络对学习动态特性建模,从而将两者结合起来,既避免了前者对重复性的苛求,又避免了神经网络控制方法通常存在的分析与实时控制的困难.文中证明了整个系统的稳定性,并以机械手为例进行了仿真.

**关键词:** 神经网络与智能控制,异步自学习控制, Lyapunov 稳定性理论, 学习算法, 机器人控制.

## 1 引言

文 2) 基于迭代与重复自学习控制提出了一种统一的异步自学习控制理论 (ALCB). 但不管是在时域与频域讨论的迭代与重复自学习控制系统, 还是统一处理的 ALCS, 上述各种方法都必须假定被控系统具有重复性. 随机 ALCS<sup>[1]</sup> 部分地取消了这一限制, 但它仅研究重复性被反馈回路之外的量测噪声所破坏的情形, 且学习收敛速度也有待进一步提高.

对于多层前馈神经网络 (MLPs), 由于它具有学习任意非线性映射关系的能力, 以及其固有的自适应特点, 已得到各个领域的普遍重视. 将极具潜力的神经网络 (NN) 模型运用于具有非线性不确定性动态系统的直接自适应控制, 最近已取得了激动人心的结果<sup>[2,3]</sup>.

但 NN 控制方法通常存在稳定性证明与在线学习的困难, 考虑到上述 ALCS 对重复性的过度依赖, 本文尝试将两者结合在一起, 利用前者“学习”的鲁棒性与后者“异步”的优点, 以图从另一途径克服这一“瓶颈”.

1) 国防科技预研基金资助项目.  
本文于 1993 年 6 月 7 日收到.

## 2 典型的异步自学习控制系统<sup>1)</sup>

不失一般性,考虑如下非线性连续时间闭环系统

$$\dot{x}_k(t) = \Psi(x_k(t), u_k(t), t), \quad (1a)$$

$$y_k(t) = C(x_k(t), t). \quad (1b)$$

其中  $x_k(t) \in R^n$  为  $t$  时刻第  $k$  步学习时的状态,  $y_k(t) \in R^m$  为输出,  $u_k(t) \in R^r$  取 ALCS 律为

$$u_{k+1}(t) = u_k(t) + \Phi(e_k(t), t). \quad (2)$$

这里  $\Phi(\cdot, \cdot)$  为 ALCS 算子,且输出误差定义为

$$e_k(t) = y_d(t) - y_k(t). \quad (3)$$

其中  $y_d(t)$  为给定的期望输出,  $k = 0, 1, \dots$  为学习迭代次数.

ALCS 方法的基本思想,是第  $k+1$  次学习时的输入  $u_{k+1}(t)$  将基于第  $k$  次学习时的经验  $\Phi(e_k(t), t)$  和输入  $u_k(t)$  获得,并且随着其中“有效”经验的不断积累而使  $e_k(t) \rightarrow 0$  或  $y_k(t) \rightarrow y_d(t)$ ,  $k \rightarrow \infty$ . 从而可望使实际输出经过“学习”而逐渐逼近其期望输出.

典型的 ALCS 方法包括早期的 PID 型学习控制<sup>2)</sup>, 以及近期发展的最优学习控制<sup>[4]</sup>、随机学习控制<sup>[5]</sup>和自适应学习控制<sup>[6]</sup>等等,它们本质上区别于学习算子  $\Phi(\cdot, \cdot)$  的具体形式,后者的选择需保证相应的学习收敛性.

## 3 神经网络异步自学习控制系统

基于上述讨论,本文尝试利用 MLPs 对任意非线性映射的逼近能力,通过**定义二次型 Lyapunov 函数及相应的代价函数**,以实现对学习算子  $\hat{\Phi}(\cdot, \cdot)$  的逼近,从而构成所谓的神经网络异步自学习控制系统.

若  $P \in R^{m \times m}$  为正定对称加权阵,设 Lyapunov 函数为

$$v(e_k(t)) = e_k^T(t) P e_k(t), \quad (4)$$

相应的代价函数定义为

$$J = (v(e_{k+1}(t)) - v(\hat{e}_{k+1}(t)))^2 / v(\hat{e}_{k+1}(t)). \quad (5)$$

式中  $\hat{e}_{k+1}(t)$  为由学习动态特性的 NN 模型给出的输出误差,此 NN 模型由 MLPs 实现,可首先利用典型的 ALCS 方法,如 PID 型学习控制进行离线学习. 而  $e_{k+1}(t)$  则由如下收敛模型给出:

$$e_{k+1}(t) = A_c e_k(t). \quad (6)$$

其中  $A_c \in R^{m \times m}$  为 Hurwitz 矩阵,它可有预先配置的极点或期望的收敛性能.

易知,这时

$$\partial J / \partial u_k(t) = (\partial J / \partial \hat{y}_{k+1}(t)) \cdot (\partial \hat{y}_{k+1}(t) / \partial u_k(t)), \quad (7)$$

上式右端的第一项可由式(3)–(5)解析求出,即

1) 邓志东. 自学习控制理论与应用. 哈尔滨工业大学博士学位论文,1991.



$$\frac{\partial J}{\partial \hat{y}_{k+1}(t)} = \frac{\partial J}{\partial \hat{e}_{k+1}(t)} \frac{\partial \hat{e}_{k+1}(t)}{\partial \hat{y}_{k+1}(t)} = 2 \left[ \left( \frac{v(e_{k+1}(t))}{v(\hat{e}_{k+1}(t))} \right)^2 - 1 \right] \hat{e}_k^T(t) P. \quad (8)$$

而第二项则可由学习动态特性的 NN 模型反向传播给出, 这与 BP 算法思路类似, 只是对连接权  $w_{ij}(t)$  的修正改成了对输入  $u_k(t)$  的学习修正. 此时

$$\frac{\partial \hat{y}_{k+1}(t)}{\partial u_k(t)} = \sum_{j=1}^{m_H} \frac{\partial \hat{y}_{k+1}(t)}{\partial o_j^H(t)} \frac{\partial o_j^H(t)}{\partial net_j^H(t)} \frac{\partial net_j^H(t)}{\partial u_k(t)} = \sum_{j=1}^{m_H} \delta_j o_j^H (1 - o_j^H) w_{ij}^H, \quad (9)$$

其中  $\delta_j = \hat{y}_{k+1}(1 - \hat{y}_{k+1})w_{ji}^0$  为反向传播到隐层的广义误差. 相应的神经网络异步自学习控制律为

$$u_{k+1}(t) = u_k(t) - \eta \partial J / \partial u_k(t) + \alpha (u_k(t) - u_{k-1}(t)), \quad (10)$$

其中  $\eta > 0, 0 \leq \alpha < 1$  与 BP 算法类似, 分别为学习率和动量项参数, 它们可进一步采用文[6,7]的快速学习算法. 图 1 给出了神经网络异步自学习控制系统的方框图.

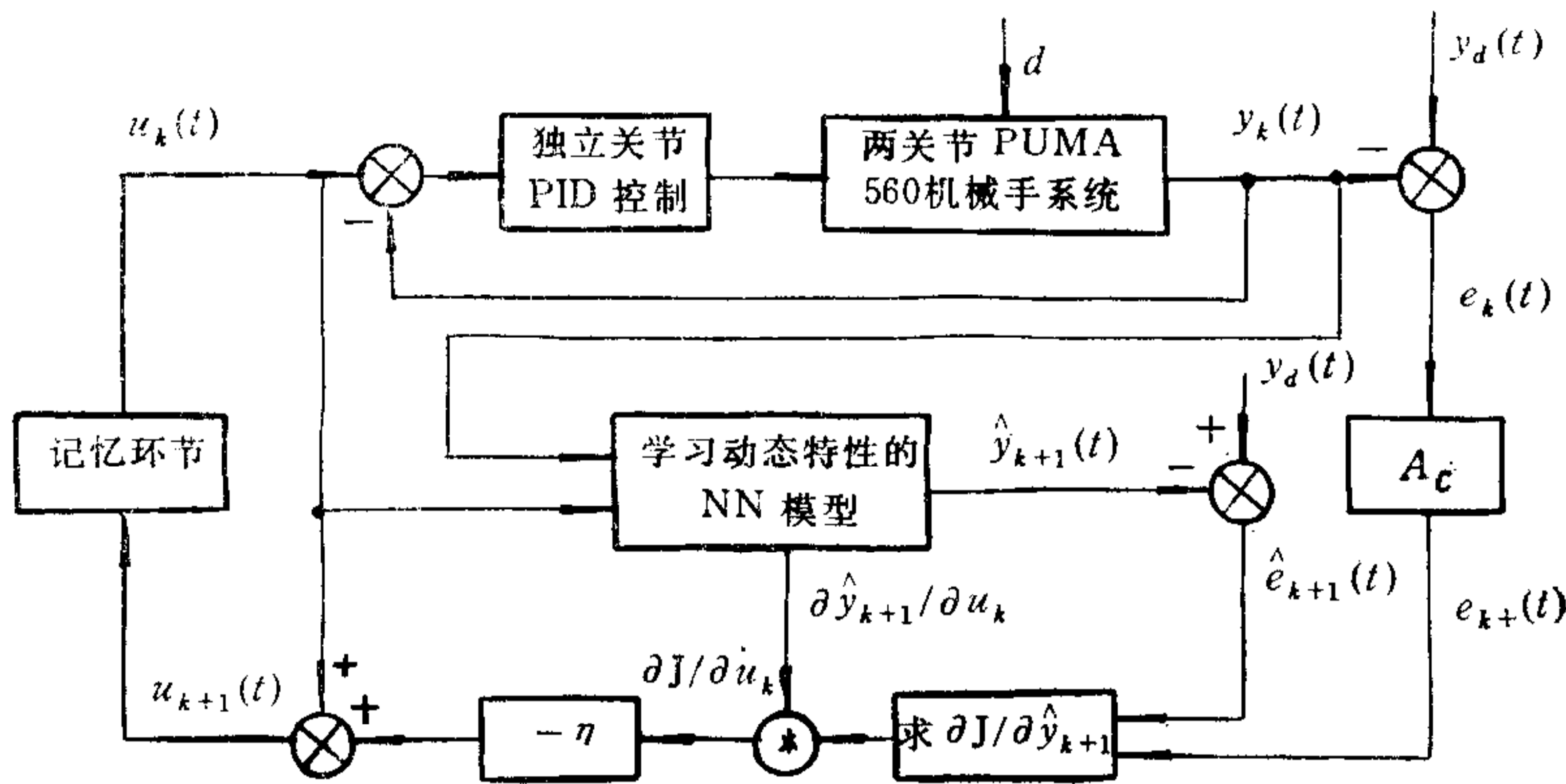


图 1 神经网络异步自学习控制系统方框图

现证明上述控制系统的稳定性. 考虑到式 (4) 的 Lyapunov 函数的正定性, 根据 Lyapunov 稳定性理论可知, 这时只需证明其时间导数负定即可.

事实上, 由式(4),(6)可知,

$$\begin{aligned} \Delta v(e_k(t)) &= v(e_{k+1}(t)) - v(e_k(t)) = e_{k+1}^T(t) P e_{k+1}(t) - e_k^T(t) P e_k(t) \\ &= e_k^T(t) (A_c^T P A_c - P) e_k(t) = -e_k^T(t) Q e_k(t). \end{aligned} \quad (11)$$

因要求  $\Delta v(e_k(t))$  负定, 故  $Q$  必须正定, 即 Lyapunov 方程  $A_c^T P A_c - P = -Q$ , 在给定正定对称阵  $Q$  时, 需存在唯一正定解  $P$ , 而这可由  $A_c$  为 Hurwitz 矩阵予以保证, 从而即可证得此神经网络异步自学习控制系统为学习渐近收敛(稳定).

### 4 两关节机械手的仿真结果

以 PUMA560 机械手关节 2 (肩关节) 和关节 3 (肘关节) 的轨迹跟踪控制为例. 图 2 给出了当随机干扰  $\sigma_d = 0.1$  时本文方法的仿真结果 (即非重复情形). 这里取仿真时间  $T_{TOL} = 2$  秒, 采样周期  $T = 0.001$  秒; 独立关节 PID 控制的各参数整定为  $K_{p1} = 76.6, K_{d1} = 40.8, K_{i1} = 40.0$  (关节 2),  $K_{p2} = 57.7, K_{d2} = 15.7, K_{i2} = 40.0$  (关节 3); PID

型 ALCS 各学习因子选择为  $KL_{p1} = 0.2$ ,  $KL_{d1} = KL_{i1} = 0$  (关节 2),  $KL_{p2} = 0.2$ ,  $KL_{d2} = KL_{i2} = 0$  (关节 3);  $A_c = 0.707$ ; MLPs 的学习率  $\eta = 0.87$ , 动量项参数  $\alpha = 0.7$ ; 初始条件  $\theta_{10} = \theta_{20} = 0$ ; 跟踪精度要求为  $E_{RMS,1} = E_{RMS,2} = 0.05$ ; 最大控制力矩限制为  $\tau_{max,1} = 520\text{H} \cdot \text{m}$  (关节 2),  $\tau_{max,2} = 260\text{H} \cdot \text{m}$  (关节 3).

单纯利用 PID 型 ALCS 和独立关节 PID 控制时的学习动态特性 (假定此时无干扰), 首先由 MLPs 进行离线学习.

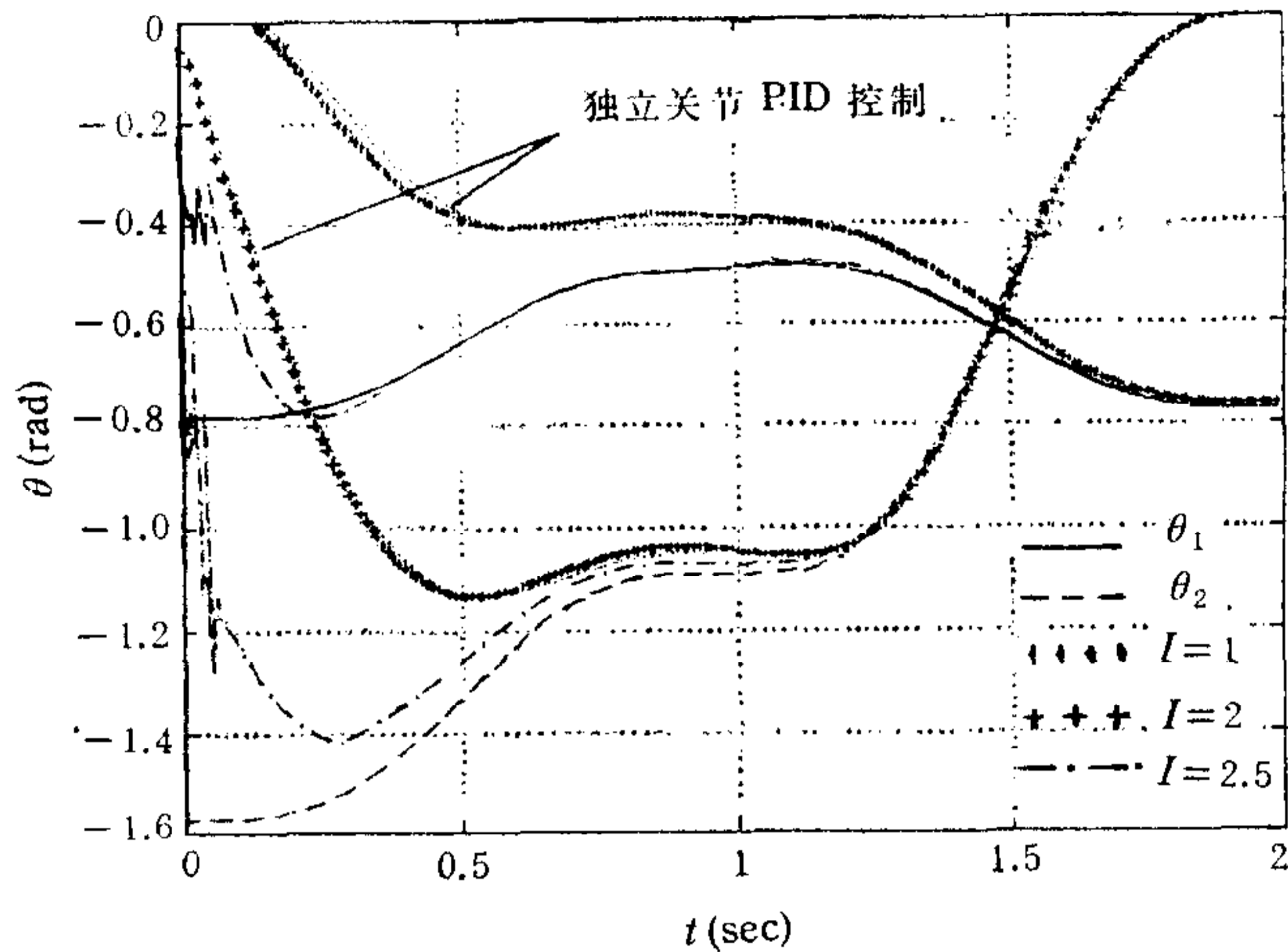


图 2 仿真结果

从上图可以明显看出, 由于各关节之间存在较强的耦合, 单纯利用独立关节 PID 控制不能取得较好的跟踪效果 ( $e_{RMS} \approx 0.5$ ), 但在采用本文的方法之后, 经过  $l = 25$  步的学习, 就可使相应的精度达到  $e_{RMS} \approx 0.05$ , 即大约提高了 1 个数量级之多.

## 5 结论

本文将 ALCS 方法与 MLPs 结合起来, 既可应用于非重复性的情形, 又可避免了 NN 控制方法通常存在的分析与实时控制的困难. 事实上, 由于采用异步自学习, 即离线学习的方式, 完成第  $k$  次典型任务时, 只需将各采样时刻的  $u_k(t)$ ,  $e_k(t)$  在线存储, 第  $k+1$  次的控制量  $u_{k+1}(t)$  可在此次典型任务完成后, 再离线计算; 如此下去, 最终可将满足跟踪精度要求之第  $l = L$  步的控制量  $u_L(t)$  作为实际控制之用, 而这仅通过查表即可实现, 因此又具有可实现性. 根据 Lyapunov 稳定性理论, 进一步给出了整个控制系统的稳定性证明, 并以 PUMA 560 机械手两关节的轨迹跟踪控制为例进行了仿真, 得到了满意的结果.

## 参 考 文 献

- [1] Deng Z D, Sun Z Q. On stochastic asynchronous learning control method, TENCON'93, (1993)
- [2] 邓志东, 孙增圻, 钱宗华. 多层前馈神经网络收敛性的比较研究, 第一届全球华人智能控制与智能自动化大会论文集, 科学出版社, 1993, 500—505.



- [3] Miller T. W, et al. (Eds.), *Neural networks for control*. MIT Press, Cambridge, MA, (1990).
- [4] 胡恒章, 邓志东. 一种最优异步自学习控制方法. *哈尔滨工业大学学报*, 1992, 24: 48-56.
- [5] Oh, S., et al., An Iterative Learning Control Method with Application for the Robot Manipulator. *IEEE J. Robotics and Automation*, 1988, (4): 508-514.
- [6] 邓志东, 孙增圻. 利用线性再励的自适应变步长快速 BP 算法, 模式识别与人工智能, 1993(已录用).
- [7] 邓志东, 孙增圻. MLPs 的高阶序贯非线性 Kalman 滤波学习算法, 控制理论与应用, 1993(已录用).

## NEURAL NETWORK BASED ASYNCHRONOUS LEARNING CONTROL SYSTEMS

DENG ZHIDONG SUN ZENGQI

(Dept. of Computer Sci. & Tech., Tsinghua University Beijing 100084 China)

LIU JIANWEI

(Dept. of Control Engineering, Harbin Institute of Technology Harbin 150006 China)

### ABSTRACT

In this paper the neural network based asynchronous learning control system, on the basis of the asynchronous learning control method given by ref. [1], is proposed. The gradient-type learning control algorithm is derived, the strict proof on stability convergency is provided by Lyapunov stability theory, and the simulation study of two links of PUMA 560 robot systems is given. The results show that the RMS track accuracy of this proposed method is improved greatly compared to classical PID control.

**Key words:** Neural network and intelligent control, asynchronous learning control, Lyapunov stability theory, learning algorithm, robot control.