

基于信度分配的并行集成CMAC及其在建模中的应用

潘 晔¹, 顾幸生¹, 卢胜利²

(1. 华东理工大学 自动化研究所, 上海 200237; 2. 天津工程师范学院 自动化工程系, 天津 300222)

摘要: Albus CMAC(cerebella model articulation controller) 神经网络是一种模拟人类小脑学习结构的小脑模型关节控制器, 它具有很强的记忆与输出泛化能力, 但对于在线学习来说, Albus CMAC仍难满足快速性的要求. 本文在常规CMAC神经网络的基础上, 针对其在学习精度与存储容量之间的矛盾, 引入信度分配概念, 提出了一种基于信度分配的并行集成CMAC. 它将大规模网络切割为多个子网络分别训练后再组合, 大大地提高了计算效率. 通过对复杂非线性函数建模的仿真研究表明, 该方案提高了系统建模的泛化能力和算法的收敛速度. 文章最后讨论了学习常数和泛化参数对该神经网络在线学习效果的影响.

关键词: 小脑模型关节控制器(CMAC); 信度分配; 并行; 神经网络集成; 建模

中图分类号: TP273 **文献标识码:** A

Credit-assignment-based parallel ensemble CMAC and its applications in modeling

PAN Ye¹, GU Xing-sheng¹, LU Sheng-li²

(1. Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China;
2. Department of Automation Engineering, Tianjin University of Technology and Education, Tianjin 300222, China)

Abstract: Albus CMAC (cerebella-model-articulation-controller) is a neural network that simulates the structure of the human cerebella and performs the articulation controller. Although it has a large memory capability and is capable of output generalization, Albus CMAC is still hard to meet the requirements of rapidity for online learning. To solve the conflict between the accuracy and memory capability of Albus CMAC, we introduce the concept of credit assignment and propose the parallel ensemble CMAC based on credit assignment. A large-scale network is separated into several sub-networks; these sub-networks are trained synchronously, and then are combined. It greatly improves the computational efficiency. In simulating the model of the complex nonlinear function, results show that the proposed scheme improves the generalization capability of the system model and raises the convergence rate of the improved arithmetic. Finally, how the learning parameter and the generalized parameter influence the effect of online learning of this neural network is discussed.

Key words: cerebella model articulation(CMAC); credit assignment; parallel; neural network ensemble; modeling

1 引言(Introduction)

小脑模型关节控制器CMAC(cerebella model articulation controller)是由Albus在小脑时空模型Eccles的基础上于1975年提出的. Albus CMAC神经网络是一种可以表达复杂非线性函数的表格查询型自适应神经网络, 它属于局部逼近神经网络, 具有固有的信息分类存储的能力和局部泛化能力, 学习速度快, 网络结构简单, 易于实现^[1]. 由于具有上述的优越性能, CMAC神经网络比一般神经网络具有更好的非线性逼近能力, 因此CMAC神经网络在非线性函数逼近、动态系统建模、神经网络软测量等领域得到了广泛应用. Cheng Wen等人提

出了一种自组织CMAC学习方法, 它是一种无监督的学习方式, 他们将其应用于心电图分类器上, 取得了比较好的结果^[2]; Huailin Zhao和Masanori Sugisaka在CMAC的原理基础上设计了一对机器人关节控制器, 其中一个将CMAC与PID控制相结合, 研究表明, 基于CMAC-PID的控制器具有更好的稳定性^[3]; Kuo-Hsiang Cheng提出了一种基于模糊神经系统的CMAC模型, 并将其用于复杂系统建模, 取得了较好的效果^[4]. 但是, 常规的CMAC仍需多个周期才能达到一定的精度^[5], 也就是说, 常规CMAC虽然收敛速度快, 但对于在线学习来说, 仍难满足其快速性的要求. 为提高CMAC的学习速度, Shun-Feng

Su和Ted Tao等人在分析常规CMAC权值调整规则的基础上,考虑到已学习知识的可信度,提出了一种基于信度分配的CA-CMAC算法(credit assigned cerebella model articulation controller)^[6].

神经网络集成(neural network ensemble)最初是由L. K. Hansen和P. Salamon等人于1990年提出的^[7].其主要思想是为要解决的对象问题训练多个神经网络,然后将这些神经网络的输出进行结合,形成最终结论.通过将多个神经网络进行组合,可以在更大范围内选择样本特征,多种信息相互补偿,从而减少选择错误预测的机会,因此它被视为一种非常有效的神经计算方法.

本文将上述思想引入标准Albus CMAC神经网络中,提出了一种并行集成CMAC神经网络模型,它将大规模网络切割为多个子网络分别训练后再组合,可以有效地降低计算成本,并得到更好的泛化能力.同时,网络训练在常规CMAC权值调整规则的基础上,考虑已学习知识的可信度,采用基于信度分配的CA-CMAC训练规则,以提高CMAC神经网络对于复杂非线性函数的逼近精度和学习速度.

2 Albus CMAC神经网络(Albus CMAC neural network)

Albus CMAC神经网络是一种模拟人类小脑学习结构的小脑模型关节控制器,它具有很强的记忆与输出泛化能力.常规CMAC神经网络的结构示意图如图1所示^[8].

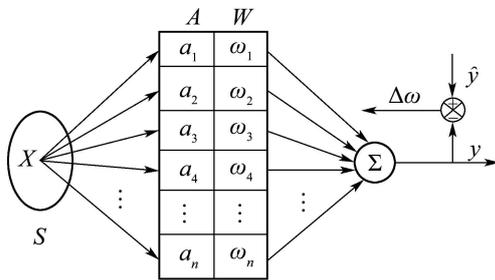


图1 CMAC神经网络结构示意图

Fig. 1 Structure of CMAC neural network

它包括4个基本组成部分:输入空间 S 、虚拟联想存储空间 A 、物理存储空间 W 、CMAC神经网络的输出 y .输入量经过量化后得到离散的输入矢量,每个输入矢量激活 A 中的 C 个单元(C 被称为泛化参数,其大小决定了网络的泛化能力).

CMAC神经网络的学习过程是通过一系列映射实现的:输入空间到虚拟联想存储单元的映射($S - A$);虚拟联想存储单元到物理存储单元的映射($A - W$);物理存储单元到输出空间的映射($W - y$).其中被激活的 C 个 a_i 的值为1,其余为0.

被激活的 a_i 对应的 w_i 之和就是网络的输出 y ,即

$$y = \sum_{i=1}^n a_i w_i. \quad (1)$$

Albus CMAC神经网络的训练过程是通过误差纠正算法实现的.其基本思想是根据期望输出值与网络实际输出值之间的误差最小原则来修正网络的权向量^[8].设 \hat{y} 为期望输出值, y 为网络的实际输出值, $e = \hat{y} - y$ 为输出误差,则权值修正公式可表示为

$$w_i(n+1) = w_i(n) + \frac{\eta(\hat{y} - y)}{C}. \quad (2)$$

其中: η 为学习常数, C 为泛化参数, $w_i(n+1)$ 是经过第 n 次迭代后存储在第 i 个存储单元中的权值.当 $\|e\| \leq \varepsilon$ (ε 为误差要求)时,停止训练;反之,当 $\|e\| > \varepsilon$ 时,继续训练,直到满足要求的误差精度为止.这里权值的校正方法是将误差平均分配给所有被激活的存储单元.

3 基于信度分配的CMAC神经网络(Credit-assignment-based CMAC, CA-CMAC)

常规CMAC训练算法直接忽略了历史学习信息对激活存储单元权值调整的影响,把误差的校正值平均分配给了所有激活的存储单元.但是事实上,由于每个存储单元并非具有相同的学习历史,所以这些存储单元也不应有相同的可信度^[8].为了提高学习速度,校正误差必须根据存储单元的可信度进行分配.文献^[6]提出权值要由该存储单元目前更新的次数来决定,存储单元学习更新次数越多,其存储的数值越可靠.因此存储单元的学习次数被看成其可信度,可信度越高,其权值修正越小,由此式(2)改写为

$$w_i(n+1) = w_i(n) + \eta \frac{f_i^{-1}(n)}{\sum_{i \in I} f_i^{-1}(n)} (\hat{y} - y). \quad (3)$$

其中: $f_i(n)$ 是第 i 个存储单元在第 n 次迭代时的激活次数, $\{I | a_i = 1, i = 1, 2, \dots, n\}$ 为被激活的存储单元集.

为了验证CA-CMAC与CMAC在函数逼近方面的有效性和差异性,选择如式(4)所示的2维非线性函数进行仿真研究,自变量从输入空间 $[-3, 3] \times [-3, 3]$ 中以间隔0.2取均匀分布的格栅点.设置CMAC神经网络的泛化参数 C 为6,两种网络训练算法的学习常数 η 为0.05,采用MATLAB 7.0实现软件编程,分别进行网络训练. CA-CMAC神经网络和CMAC神经网络迭代50次后的均方误差下降曲线如图2所示:

$$z = 0.5(x^2 + y^2). \quad (4)$$

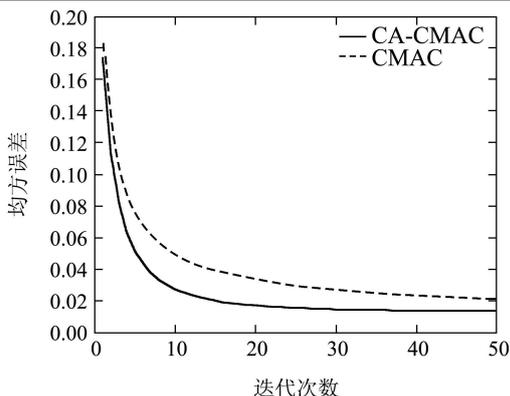


图 2 两种算法逼近2维函数的均方误差下降曲线
Fig. 2 Mean-square error curve of two algorithms approaching 2-dimension function

从以上仿真研究可以看出, 基于信度分配的CA-CMAC算法较常规CMAC算法的误差下降速度更快, 精度更高. 但是, 这两种算法在对高维复杂函数进行逼近时, 能够达到的精度并不一定能满足辨识要求. 因此, 保持CA-CMAC算法的快速收敛性并有效地提高其学习精度是有待解决的问题.

4 并行集成CMAC神经网络(Parallel ensemble CMAC)

为了提高函数逼近精度, 可以通过提高量化分辨率的方法, 但会导致内存容量以指数递增, 从而降低网络的运行速度. 对于一个大规模的问题, 单一CMAC神经网络需要一个很长的训练过程. 为了缩短网络学习时间, 采用并行的方法将大规模网络切割为多个子网络分别训练后再组合, 可以有效地降低计算成本, 并得到更好的泛化能力.

并行集成CMAC神经网络采用结构并行的方式, 使用多个低维输入CMAC按某种规律进行组合, 以构造一个与高维输入CMAC等价的网络模型. 每个CMAC神经网络分别完成自己的计算任务, 当多个网络的训练结束时, 并非选择其中的某一个神经网络, 而是通过组合这几个神经网络的结果实现网络的线性或非线性的集成, 以提高模型的精度和鲁棒性.

同典型的神经网络集成构造方法一样, 并行集成CMAC神经网络需要解决两个问题, 即: 个体生成方法和结论集成策略.

1) 个体生成方法.

组成集成神经网络的各子网差异越大, 集成的效果越好. 为此, 可以通过使用不同的训练样本集、不同的网络结构、不同的网络类型、不同的初始条件及不同的训练算法等方式来提高网络间的差异. 对于一个大规模的多输入模型, 本文采用逐级分解的

方法, 将一个高维输入CMAC模型分解成多个低维初级CMAC输入模型, 并将每个低维模型的输出作为次级模型的输入.

2) 结论集成策略.

按照集成学习的思想, 个体网络集合产生后, 还需要寻找一种方法将各个体的输出结果进行组合来生成较优的最终输出. 常用的结论集成方法有投票法、加权平均法、非线性组合方法等. 总之, 根据集成神经网络使用目标的不同, 则相应的结论集成方法也不尽相同. 本文不使用线性合成方法, 而是仍采用一个CMAC神经网络来对多个个体预测结果进行合成.

5 仿真研究(Simulation and research)

为了验证该并行集成CMAC神经网络的有效性, 选择如式(5)所示的3输入非线性函数进行仿真实验. 各子网采用基于信度分配的训练算法(CA-CMAC)进行权值修正.

$$P = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2. \quad (5)$$

将此非线性函数通过式(6)进行分解:

$$\begin{cases} R = x^{0.5} + y^{-1}, \\ V = 1 + z^{-1.5}, \\ P = (R + V)^2. \end{cases} \quad (6)$$

利用两个CMAC神经网络分别对 R, V 进行并行网络训练, 然后, 各个子网的输出结果再通过一个CMAC神经网络进行非线性集成, 以得到并行网络的最终结果, 实现 $x, y, z, \rightarrow P$ 的学习, 结构如图3所示.

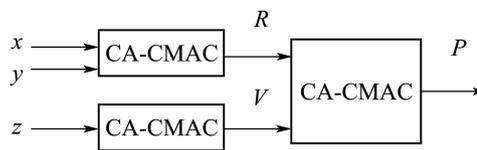


图 3 3输入并行CA-CMAC结构图
Fig. 3 Three-input parallel CA-CMAC neural network

从输入空间 $[1, 6] \times [1, 6] \times [1, 6]$ 中以间隔0.5均匀采样, 输入数据根据归一化方法进行预处理. 设置各子网的泛化参数 C 为5, 选择学习常数 η 为0.02. 采用MATLAB 7.0实现软件编程, 进行网络训练. 并行集成CA-CMAC神经网络(parallel ensemble CA-CMAC neural network, 简称为P-CA-CMAC)、CA-CMAC神经网络和CMAC神经网络迭代50次后的均方误差下降曲线及均方误差列表分别如图4和表1所示.

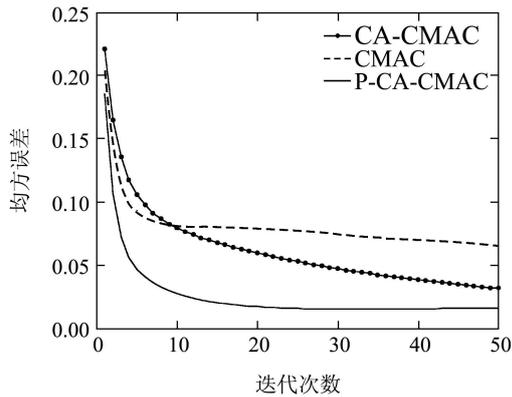


图4 3种算法逼近3维函数的均方误差下降曲线
Fig. 4 Mean-square error curve of three algorithms approaching 3-dimension function

表1 3种算法逼近3维函数的均方误差
Table 1 Mean-square error of three algorithms approaching 3-dimension function

迭代次数	均方误差		
	P-CA-CMAC	CA-CMAC	CMAC
1	0.18531	0.22108	0.20318
3	0.072357	0.13538	0.11137
5	0.046532	0.10569	0.089971
7	0.035854	0.091133	0.083529
10	0.027211	0.079166	0.081091
20	0.016841	0.059393	0.078718
30	0.015033	0.047034	0.074233
50	0.01579	0.03157	0.065508

结果可知,3种算法经过一定次数的迭代后,都能够达到一定的学习精度,然后加以保持.由图表可以清楚地看到,CA-CMAC训练算法较常规CMAC算法具有明显的收敛速度快的优势.而P-CA-CMAC训练算法相对CA-CMAC训练算法具有明显更快的收敛速度,在经过相对较少的迭代次数后便达到了更高的学习精度.这充分显示了P-CA-CMAC算法在函数逼近中的可行性和有效性.可见,对于高维复杂函数,P-CA-CMAC算法能够更加有效地完成网络学习.

为了研究学习常数 η 和泛化参数 C 对P-CA-CMAC的在线学习效果的影响,以上面式(5)所示的非线性函数为例,采用3输入并行CA-CMAC结构,从输入空间 $[1, 6] \times [1, 6] \times [1, 6]$ 中以间隔0.5均匀采样,采用MATLAB7.0实现软件编程,进行网络训练.

设置泛化参数 C 为5,则对于不同的学习常数 η ,P-CA-CMAC神经网络迭代50次后的均方误差下降曲线如图5所示.

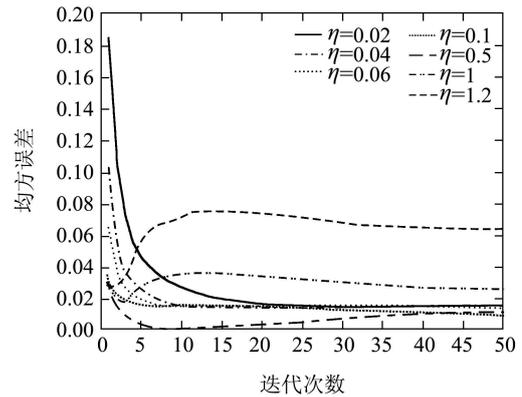


图5 不同 η 的P-CA-CMAC逼近3维函数的均方误差下降曲线
Fig. 5 Mean-square error curve of P-CA-CMAC with different η approaching 3-dimension function

设置学习常数 η 为0.02,则对于不同的泛化参数 C ,P-CA-CMAC神经网络迭代50次后的均方误差下降曲线如图6所示.

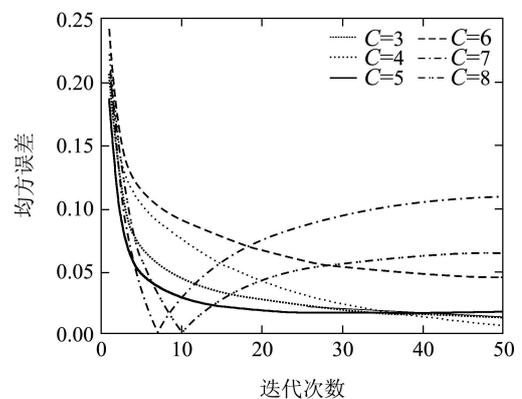


图6 不同 C 的P-CA-CMAC逼近3维函数的均方误差下降曲线
Fig. 6 Mean-square error curve of P-CA-CMAC with different C approaching 3-dimension function

从图5可以看出,在网络学习的初始阶段(前10个周期),不同的学习常数 η 使均方误差下降的速度和精度有很大差别:在 $\eta = 0.02$ 时,P-CA-CMAC网络的收敛速度较慢;随着 η 的增加,收敛速度变快,在 $\eta = 0.5$ 时收敛速度达到最快; η 值继续增大,网络学习精度变差.这是因为在网络的最初阶段,输出误差 $e = \hat{y} - y$ 较大,随着学习常数 η 的增大,权值调整的量也随之增大;当权值调整量超过一定范围后,会导致误差振荡,从而影响网络的学习精度.在网络学习的后续阶段(10个周期后),对于不同的 η 值,网络的均方误差都将趋于一个稳定值.从图上可以看出 η 在区间 $[0.02, 0.5]$ 时,网络的学习效果较好.

从图6可以看出,网络的泛化参数 C 不同,网络的收敛速度和学习精度也不相同.这是因为在保持网

络结构参数不变的情况下, 当泛化参数 C 逐渐增大时, 在网络学习中权值更新的次数也会不同, 这种更新次数的不同使得那些更新次数多的权值对网络输出的影响大于更新次数少的权值, 从而产生泛化误差, 使学习精度下降. 从图上可以看出, 当泛化参数 C 取3~6时, 网络的学习速度和精度比较好, 其中当 $C = 5$ 时, 收敛速度最快, 当 $C = 4$ 时, 学习精度最高.

对于逼近高维复杂函数, P-CA-CMAC算法最初几个周期的学习速度和最终的学习精度极为重要. 因此, 选取适当的学习常数 η 和泛化参数 C 对提高P-CA-CMAC神经网络的性能具有一定的意义.

6 结论(Conclusions)

Albus CMAC神经网络是一种模拟人类小脑学习结构的小脑模型关节控制器, 它具有很强的记忆与输出泛化能力. 而基于信度分配的CA-CMAC算法较常规CMAC算法的误差下降速度更快, 精度更高. 本文在标准Albus CMAC神经网络的基础上, 针对其在学习精度与储存容量之间的矛盾, 提出了基于信度分配的并行集成CMAC神经网络. 仿真结果证明, 基于信度分配的并行集成CMAC神经网络在保持CA-CMAC神经网络快速性的基础上, 有效地提高了网络的学习精度. 该方法在一定意义上解决了Albus CMAC在学习精度与存贮容量之间的矛盾, 使其在系统建模、软测量等领域具有实际应用价值.

参考文献(References):

[1] ALBUS J S. A theory of cerebellar function[J]. *Mathematics Biosciences*, 1971, 10(1): 25 – 61.

- [2] WEN C, LIN T C, CHANG K C, et al. Classification of ECG complexes using self-organizing CMAC[J]. *Measurement*, 2009, 42(3): 399 – 407.
- [3] ZHAO H L, SUGISAKA M. Simulation study of CMAC control for the robot joint actuated by McKibben muscles[J]. *Applied Mathematics and Computation*, 2008, 203(1): 457 – 462.
- [4] CHENG K H. CMAC-based neuro-fuzzy approach for complex system modeling[J]. *Neurocomputing*, 2009, 72(7): 1763 – 1774.
- [5] IIGUNI Y. Hierarchical image coding via cerebellar model arithmetic computers[J]. *IEEE Transactions on Image Processing*, 1996, 5(6): 1393 – 1401.
- [6] SU S F, TAO T, HUNG T H. Credit assigned CMAC and its application to online learning robust controllers[J]. *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, 2003, 33(2): 202 – 213.
- [7] HANSEN L K, SALAMON P. Neural network ensembles[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993 – 1001.
- [8] 鲍琼, 顾幸生. 基于信度分配的串行集成CMAC及其仿真[J]. 华东理工大学学报, 2008, 34(5): 751 – 754.
(BAO Qiong, GU Xing-sheng. Series ensemble CMAC based on credit assignment[J]. *Journal of East China University of Science and Technology*, 2008, 34(5): 751 – 754.)

作者简介:

潘晔 (1985—), 女, 主要研究方向为智能优化算法、软测量建模, E-mail: selina_panye@163.com;

顾幸生 (1960—), 男, 博士, 教授, 博士生导师, 在国内外学术刊物上发表学术论文200多篇, 目前研究方向为控制理论与应用、生产计划与调度、工业工程建模、控制与优化、故障检测与诊断等, E-mail: xsgu@ecust.edu.cn;

卢胜利 (1956—), 男, 博士, 教授, 主要研究领域为智能信息处理技术和嵌入式测控系统等, E-mail: LSL5612@163.com.