

## 协同神经网络中参数的优化\*

王海龙 戚飞虎 任庆生

(上海交通大学计算机系, 上海, 200030)

**摘要** 提出了一种基于遗传算法的协同神经网络参数的优化算法, 该算法利用遗传算法的全局最优搜索能力, 在协同神经网络的参数空间搜索最优解. 对从“车牌识别系统”中采集得到的数字样本进行的测试表明: 优化算法能有效提高协同神经网络的识别性能, 使识别率达到了较为实用的水平(98.4%). 另外, 还对协同神经网络中各个参数在识别过程中的作用进行了讨论.

**关键词** 协同神经网络, 神经网络优化, 遗传算法, 目标识别.

## PARAMETERS OPTIMIZATION OF SYNERGETIC NEURAL NETWORK\*

WANG Hai-Long QI Fei-Hu REN Qing-Sheng

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

**Abstract** An optimization algorithm based on genetic algorithm for parameters of synergetic neural network was proposed, which searched the optimal solutions in the parameters space of synergetic neural network using the globally optimal searching ability of genetic algorithm. The test upon the numeral samples from 'Car License Recognition System' shows that the optimization algorithm is able to improve the recognition performance of synergetic neural network effectively, and make the recognition rate reach the level of real application (98.4%). Additionally, the functions of parameters of synergetic neural network in object recognition were discussed.

**Key words** synergetic neural network, optimization of neural network, genetic algorithm, target recognition.

### 引言

协同神经网络自 H. Haken 教授提出<sup>[1]</sup>以来, 人们已进行了较为广泛的研究. 胡<sup>[2]</sup>等对协同方法的基本动力学方程中注意参数不平衡情况下的动力学行为进行了研究, 并得到了注意参数与神经网络解的不动点之间的关系. Wang, Lever 和 Pu<sup>[3]</sup>研究了注意参数的特性并提出了一种简单的基于奖惩学习机制的注意参数训练方案, 使协同神经网络的识别性能有所提高. 但实际上, 协同神经网络中的参数不仅仅只有注意参数  $\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ , 还包括定义了各原型模式吸引域之间边界的参数  $B$  和限制项参数  $C$ , 所以真正提高协同神经网络性能的方法是提出一种对全部 3 个网络参数  $(\bar{\lambda}, B, C)$  进行网络学习的训练方案.

近年来出现了利用进化计算来优化协同神经网络的方法<sup>[4]</sup>, 并日益成为智能计算领域中的研究热

点, 而且已经在某些领域中得到成功的应用<sup>[5~8]</sup>. 神经网络的优化包括网络结构的优化和网络参数或权重的优化, 网络优化是将网络的学习过程看作是在结构空间(或参数空间)中搜索最优结构(或参数集合)的过程. 由于进化计算是一类全局随机搜索算法, 能够在复杂的、多峰的、不可微的大矢量空间中快速有效地找到全局最优解.

鉴于协同神经网络结构相对固定, 本文提出了一种基于遗传算法(GA)来优化协同神经网络中网络参数  $(\bar{\lambda}, B, C)$  的优化方案, 该算法能有效地在参数空间搜索全局最优解, 使协同神经网络的识别性能达到最佳. 这一点在对从“车牌识别系统”中得到的数字样本进行的实验中得到证实.

### 1 协同神经网络

假设有  $M$  个原型模式  $v_i$  和试验模式  $q$ , 它们分

\* 国家自然科学基金(编号 69772002)资助项目  
稿件收到日期 2000-03-08, 修改稿收到日期 2000-06-07

\* The project supported by the National Natural Science Foundation of China (No. 69772002)

Received 2000-03-08, revised 2000-06-07

别被归一化为具有零均值和单位长度的  $N$  维列向量。为了满足原型模式间线性无关条件,要求  $M < N$ 。Haken 教授指出协同识别方法可描述为一个有势的动力学过程,其势函数表达式为

$$V = -\frac{1}{2} \sum_{i=1}^M \lambda_i (v_i \cdot q)^2 - \frac{1}{4} \sum_{i=1}^M \sum_{j=1, j \neq i}^M B_{ij} (v_i \cdot q) \cdot (v_j \cdot q)^2 - \frac{1}{4} C (q^+ \cdot q)^2 \quad (1)$$

其中  $\lambda_i$  为注意参数,  $B_{ij}$  和  $C$  为常数。势能函数  $V$  由 3 项组成: 第一项定义在势能面上对应原型模式的极小值, 每个极小值的深度由注意参数  $\lambda_i$  决定; 第二项定义了各原型模式之间的竞争并决定各原型模式吸引域之间的边界, 这些靠参数  $B_{ij}$  来定义; 第三项是限制项, 由常数  $C$  决定。

为了降低系统维数, 引入序参量  $\xi_i$ , 序参量定义为试验模式  $q$  在伴随模式上的投影, 即

$$\xi_i = v_i \cdot q, \quad k = 1, \dots, M \quad (2)$$

根据势动力学可以得到协同系统的经过简化的模式识别动力学方程

$$\dot{\xi}_i = \lambda_i \xi_i - \sum_{j=1}^M B_{ij} \xi_j \xi_i - C \sum_{i=1}^M \xi_i \xi_i \quad (3)$$

目标识别的动力学方程式(3)可以表示为 3 层网络, 最后网络的输出被投影到第三层, 根据

$$q(t) = \sum_{i=1}^M \xi_i v_i \quad (4)$$

实现网络的联想功能, 最终完成协同神经网络的识别过程。

## 2 协同神经网络参数的优化

### 2.1 $V$ 的稳定不动点

当动力学方程式(3)中的 3 个参数取值为

$$B_{ij} = C = \lambda > 0 \quad (5)$$

时, Haken 证明了动力学方程式(3)具有如下性质: (1) 各原型目标模式向量是  $V$  的稳定不动点; (2) 没有其它的目标模式向量  $V$  的稳定不动点; (3) 最初最大的序参量将在网络竞争中最终获胜, 而其他

序参量将消亡。

其中性质(1)和(2)决定了网络最终将收敛到某个原型目标模式, 而不会出现传统神经网络中的“伪状态”, 性质(3)意味着根据序参量的初值可以预测网络最终的输出, 式(3)失去了通过进一步学习使网络正确识别试验目标模式的能力。

当动力学方程式(3)中的 3 个参数取值为

$$B_{ij} = B > 0, C > 0, \lambda_i > 0 \quad (6)$$

时, 胡<sup>[2]</sup>证明式(3)有如下定理。

定理 1

$$\text{假设 } (B - C)\lambda_i - C\lambda_j > 0, \quad 1 \leq i \neq j \leq M \quad (7)$$

则式(3)描述的有势动力学系统的静态点可以分为 3 类: (1) 零点  $\xi = 0$  为不稳定不动点; (2) 有  $2M$  个稳定不动点  $\xi = (0, \dots, \pm(\lambda_k/C)^{1/2}, \dots, 0)^T, 1 \leq k \leq M$ , 此时的势函数值为:  $V_k = -\lambda_k/4C$ ; (3) 有  $3^M - 2M - 1$  个鞍点。

定理 1 表明  $V$  的稳定不动点对应于  $M$  个原型模式, 并按正负对称分布; 不平衡注意参数条件下  $V$  的稳定不动点并非均匀分布, 对应的吸引域浓度各不相同, 且由注意参数  $\lambda$  和  $C$  决定。稳定不动点的吸引域大小由引域深度和参数  $B$  决定, 所以不同  $(\lambda, B, C)$  将导致稳定不动点的吸引域深度和吸引域大小不同, 影响到各原型目标模式之间的竞争态势, 从而达到最佳的识别效果。

### 2.2 GA 参数优化方法

遗传算法(GA)是 J. Holland 在 70 年代受生物进化论思想启发提出的一种基于自然选择和群体遗传机理搜索最优解的概率算法, 它模拟了自然选择和自然遗传过程中发生的繁殖、杂交和变异现象, 通过复制(duplication)、杂交(crossover)和变异(mutation)3 种基本遗传算子实现种群进化<sup>[9]</sup>。本文提出的 GA 参数优化方法目的在于寻找协同神经网络参数  $(\lambda, B, C)$  的最优解, 文中对参数  $(\lambda, B, C)$  采用二进制编码, 并排成一行, 取值范围为  $[0, 6]$ , 二进制编码的精度为  $1 \times 10^{-5}$ , 所以  $(\lambda, B, C)$  中的每个参数

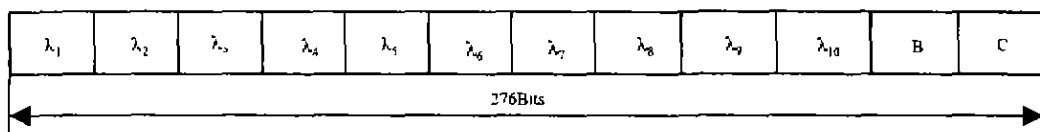


图 1 协同神经网络的 GA 参数优化算法中染色体表示的结构图

Fig. 1 Structure representation of chromosome of parameter optimization algorithm of synergistic neural network based on GA

用 23Bits 来表示, 每个染色体占  $(M+2) \times 23\text{Bits}$ , 文中  $M=10$ , 其结构如图 1 所示. 为了能够利用 GA 求解参数空间  $(\bar{\lambda}, B, C)$  中的全局最优参数, 还必须建立适当的适应度函数. 由于参数最优是针对给定的样本训练集, 在我们这里是从“车牌识别系统”中得到的 421 幅数字图像. 所以本文把利用协同神经网络对 421 幅数字图像进行成识别率作为该染色体的适应度值. 其中协同神经网络对数字图像进行识别的参数设置为: 迭代步长  $\Delta t=0.1$ , 迭代步数  $n=80$ . 另外, GA 算法的参数和算子的设置如下: 每一代种群规模为 20, 最大繁衍代数 (Maximum generation) 设置为 200 代. 本文中交叉概率  $p_c$  和变异概率  $p_m$  都分别取 0.8 和 0.05. 另外, 本文采用的遗传算子分别是转盘选择算子、简单交叉算子和二进变异算子. 停止准则为 EPSILON-DELTA 停止准则. 具体的遗传算法流程如下: (1) 随机产生初始解群  $P \{x_1, x_2, \dots, x_N\}$  作为第一代种群, 并给出  $P$  中的所有个体的适应度值; (2) 通过选择算子挑选出合适的新个体, 产生新的解群  $P'$ ; (3) 对解群  $P'$  以概率  $P_c$  进行交叉算子操作, 得到解群  $P''$ ; (4) 对解群  $P''$  以概率  $p_m$  进行变异算子操作, 得到解群  $P'''$ ; (5) 计算解群  $P'''$  中所有个体的适应度, 若计算代数已到或解群中产生满意解, 算法结束, 否则, 转到 (2).

### 3 实验结果和讨论

本文将 421 幅从“车牌识别系统”中得到的数字二值图像作为协同神经网络的 GA 参数优化算法的训练样本集, 所有实验都在 CELERON 300A 上的 Matlab 5.2 环境下完成.

首先计算出协同神经网络中的原型模式及其对应的伴随模式, 然后利用协同神经网络的 GA 参数优化算法来训练参数  $(\bar{\lambda}, B, C)$ , 图 2 就是 GA 参数优化算法在训练参数  $(\bar{\lambda}, B, C)$  时的训练性能曲线, 其中  $N$  表示迭代步数,  $F$  表示适应值. 表一是利用 GA 参数优化协同神经网络时得到的最佳参数  $(\bar{\lambda}, B, C)$ . 从表中容易得出参数  $(\bar{\lambda}, B, C)$  满足定理中的

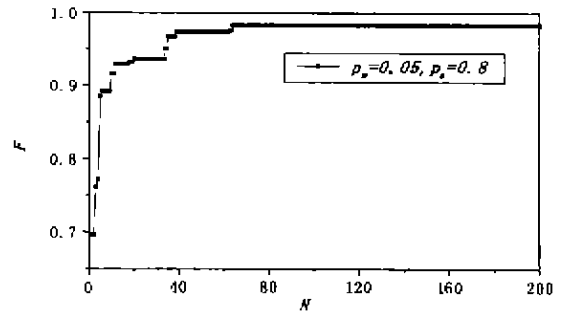


图 2 协同神经网络的 GA 参数优化算法在训练参数  $(\bar{\lambda}, B, C)$  时的训练性能曲线.

Fig. 2 Performance of training parameters  $(\bar{\lambda}, B, C)$  using parameter optimization algorithm of synergetic neural network based on GA

条件, 所以定理成立, 即协同动力学系统只存在  $M$  个以原型模式为稳定不动点的静态点, 协同神经网络可以用于目标识别. 从图 2 中的关系曲线可以看出: 参数优化算法具有很强的全局最优搜索能力, 并确实能够找到最优的参数  $(\bar{\lambda}, B, C)$  使协同神经网络的识别性能得到较大提高, 我们对“车牌自动识别系统”数字样本的测试得到识别率为 98.4%, 达到了较为实用的水平. 为了研究参数  $(\bar{\lambda}, B, C)$  中各个参数的作用, 文中设计了 4 个试验, 4 个试验的参数设置和试验结果分别是: (1) 令参数  $B=C=1.0$ , 而参数  $\bar{\lambda}$  保持不变, 协同神经网络对数字样本集的测试的识别率仅为 92.4%; (2) 令参数  $\bar{\lambda}$  中的  $\lambda_k=1.0, k=1, 2, \dots, M$ . 而参数  $B$  和  $C$  保持不变, 则对数字样本的测试结果与参数  $\lambda_k=B=C=1.0$  的设置的结果相同, 为 95.6%; (3) 令参数  $B=1.0$ , 而参数  $\bar{\lambda}$  和  $C$  保持不变, 协同神经网络对数字样本集的测试的识别率仅为 93.1%; (4) 令参数  $C=1.0$ , 而参数  $\bar{\lambda}$  和  $B$  保持不变, 协同神经网络对数字样本集的测试的识别率为 98.1%.

由以上 4 个试验可以看出: 在参数  $B$  和  $C$  取某个值时得到的最优注意参数  $\lambda'_k, (k=1, 2, \dots, M)$ . 在参数  $B$  或  $C$  取另外某个值时, 最优注意参数  $\lambda'_k, (k=1, 2, \dots, M)$  将不再是最优的了, 所以对注意参数

表 1 利用 GA 参数优化算法优化协同神经网络时得到的最佳参数  $(\bar{\lambda}, B, C)$

Table 1 Optimal parameters  $(\bar{\lambda}, B, C)$  obtained by optimization of synergetic neural network using parameter optimization algorithm based on GA

$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$\lambda_8$	$\lambda_9$	$\lambda_{10}$	$B$	$C$
3.03903	2.28236	4.07689	3.26969	2.68916	3.04207	2.7872	2.73728	3.03336	3.38050	4.51414	0.85393

寻优算法只解决了在  $B$  和  $C$  取某个值时如何获得具有较好识别性能的注意参数的问题,然而参数  $(\bar{\lambda}, B, C)$  的这种设置不一定是协同神经网络识别性能最佳的参数配置,而文中的参数优化算法却很好的解决了此问题.上文的试验证明了我们的分析是正确的,即协同动力学系统的稳定不动点的吸引域的深度和大小是由参数  $(\bar{\lambda}, B, C)$  共同决定的.当然,如果能够证明式(1)中势函数对应的协同动力学系统具有类似于定理的结论,即  $B_m (1 \leq k, 1 \leq M)$  取矩阵形式,而不再是单一的常数项,这意味着不仅能够对协同动力学系统的每个稳定不动点的吸引域深度进行控制,而且对吸引域的大小也能进行逐个控制,那么就能够对协同神经网络进行最一般的优化,也是我们即将进行的工作.

#### 4 结论

本文提出了一种协同神经网络的 GA 参数优化算法,该算法利用 GA 的全局搜索能力,在协同神经网络的参数空间搜索最优参数  $(\bar{\lambda}, B, C)$ .对“车牌识别系统”采集到的数字样本进行的试验证明了新算法的有效性,协同神经网络的识别性能也达到了较为实用的水平.另外,文中还讨论了参数  $\bar{\lambda}, B$  和  $C$  在协同动力学系统中各自的作用,并在实验上进行了验证.

#### REFERENCES

- [1] Haken H. *Synergetic Computers and Cognition-Top-down Approach to Neural Nets*. Berlin: Springer-Verlag, 1991
- [2] HU Dong-Liang, QI Fei-Hu. Study on unbalanced attention parameters in synergetic approach on pattern recognition. *Journal of Electronics* (胡栋梁, 戚飞虎. 协同模式识别中不平衡注意参数的研究, *电子学报*), 1999, **27**(5): 15—17
- [3] Wang F Y, Lever P J A, Pu B. A robotic vision system for object identification and manipulation using synergetic pattern recognition. *Robot. Comput. Integrated Manufacturing*, 1993, **10**(6): 445—459
- [4] Yao X. A review of evolutionary artificial neural networks. *International Journal of Intelligent Systems*, 1993, **8**(4): 539—567
- [5] Porto V W, Fogel D B. Alternative neural network training methods. *IEEE Expert*, 1995, **10**(3): 16—22
- [6] Yao X, Liu Y. A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks*, 1997, **8**(3): 694—713
- [7] ZHU Hong, ZHAO Yi-Gong. Fast image correlative matching based on genetic algorithm. *Journal of Infrared Millim. Waves* (朱红, 赵亦工. 基于遗传算法的快速图像相关匹配. *红外与毫米波学报*), 1999, **18**(2): 145—150
- [8] HUANG Hai-Long, QI Fei-Hu. A genetic algorithm approach to accurate calibration of camera. *Journal of Infrared Millim. Waves* (黄海龙, 戚飞虎. 一种精确标定摄像机的遗传算法方案. *红外与毫米波学报*), 2000, **19**(1): 1—6
- [9] LIU Yong, KANG Yu-Shan, CHENG Ping. *Nonnumerical Parallel Algorithms-Genetic Algorithm*. Book 2. Beijing: Science Press (刘勇, 康毓山, 陈屏. 非数值并行算法—遗传算法. 第2册. 北京: 科学出版社), 1997: 5—23
- [1] Haken H. *Synergetic Computers and Cognition-Top-*