

基于信息叠加的学习算法*

王海龙 戚飞虎

(上海交通大学计算机科学与工程系, 上海, 200030)

TP301.6

摘要 提出了一种基于信息叠加的迭代学习算法. 该算法利用协同神经网络中的原型模式具有信息的可加性, 将学习中误识率最高的模式作为反馈量来修正原型模式. 利用实际采集得到的样本对新算法进行的测试表明: 新算法具有最优搜索能力强, 训练时间短的特点. 另外, 将新算法与基于遗传算法的原型模式选取算法在网络训练性能上进行了比较.

关键词 原型模式, 序参量, 协同神经网络, 遗传算法.

信息叠加. 学习算法

LEARNING ALGORITHM BASED ON THE SUPERPOSITION OF INFORMATION*

WANG Hai-Long QI Fei-Hu

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

Abstract A new iterative learning algorithm based on the superposition of information was proposed. Because the prototype patterns of synergetic neural network (SNN) has the ability of superposition of information, the new algorithm can modify the prototype patterns using the pattern, of which the recognition rate is the lowest during training as the feedback. The test upon the samples from real environment shows that the new algorithm has the characteristic of strong ability of optimal searching and the shorteness of training time. Additionally, the comparison of training performance between the new algorithm and selection algorithm of prototype patterns based on genetic algorithm (SAPPGA) was made.

Key words prototype pattern, order parameter, synergetic neural network (SNN), genetic algorithm.

引言

协同神经网络自 Haken 教授^[1]提出以来, 人们提出了多种协同学习算法. Wagner 和 Bobe^[2]提出了利用简单的数学平均在协同神经网络的原型模式中引入多样本信息的 SCAP 算法. 王海龙等^[3,4]分别提出了运用 C-均值聚类算法和遗传算法来获得具有多样本信息的原型模式的学习算法. 这 3 种学习算法都是通过某种算法得到具有多样本信息的原型模式, 其中王海龙在文献[4]中提出的基于遗传算法的原型模式选取算法具有较好的效果, 而且在无穷迭代条件下, 由遗传算法的性质可保证该算法可获得最优原型模式. 但由于用计算机仿真时的迭代总归是有限次的, 所以实际应用中该算法并不能保证获得最优解. 实验上也发现标准遗传算法在搜索最优原型模式上表现一般, 训练时间又长, 所以训练效率很低. 改进的办法有两种: 一

是改进遗传算法, 使其具有更强的寻优能力; 二是提出更为高效的算法, 使其具有强的搜索能力, 并且缩短训练时间. 本文采用了第二种改进. 上面 3 种学习算法都有一个共同的特点, 就是都利用了原型模式具有信息的可加性这个特点. 本文正是利用这个特点提出了一种极其高效的学习算法——基于信息叠加的学习算法, 该算法将学习中误识率最高的模式作为反馈量来修正原型模式. 新算法在较短的时间内进行迭代计算, 最后得到的原型模式的识别率较之前 3 种算法都有较大幅度的提高, 也是目前协同学习算法中训练效果最好、效率最高、实用性最强的算法.

1 协同基本原理

假设原型模式数为 M , 原型模式向量的维数为 N , 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

* 国家自然科学基金(编号 69772002)资助项目
稿件收到日期 1999-12-06, 修改稿收到日期 2000-03-06

* The project supported by the National Natural Science Foundation of China (No. 69772002)
Received 1999-12-06, revised 2000-03-06

力学过程^[1],其势函数表达式为

$$V = -\frac{1}{2} \sum_{k=1}^M \lambda_k (v_k^+ q)^2 + \frac{1}{4} \sum_{i \neq j} \sum_{k=1}^M B_{ik} (v_i^+ q)^2 + \frac{1}{4} C (q^+ q)^2. \quad (1)$$

势能函数 V 由三项组成,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 λ_k 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_{ik} 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 ξ_k ,序参量定义为试验模式 q 在伴随模式上的投影,即

$$\xi_k = v_k^+ q, \quad k = 1, \dots, n \quad (2)$$

根据势动力这可以得到协同系统的经过简化的模式识别动力学方程

$$\dot{\xi}_k = \lambda_k \xi_k - \sum_{i \neq k} B_{ik} \xi_i \xi_k - C \left(\sum_{i=1}^n \xi_i^2 \right) \xi_k. \quad (3)$$

2 基于信息叠加的学习算法

协同神经网络中的原型模式具有信息叠加能力,所谓具有信息叠加能力就是假如有 M 个原型模式 v_i , $i=1, 2, \dots, M$,现在有一个测试样本 q 属于第 i 个模式,但送到协同神经网络中被识别为第 j 个模式.如果将第 i 个原型模式 v_i 修改为 $v_i = (v_i + q)/2$,只要测试样本 q 变形不是非常大,这时将 q 送到协同神经网络中就可能被正确识别为模式 i .本文利用这个特点,提出了一种基于信息叠加的迭代学习算法.该算法先将所有训练样本送进协同神经网络中识别,把误识率最高的模式作为反馈量来修正原型模式,然后用修正后的原型模式将训练样本集识别一遍,再把误识率最高的模式作为反馈量来修正原型模式,就这样通过“识

别-修正”的循环迭代过程来训练原型模式,直到算法结束条件满足或迭代步数到为止.假设训练样本集 $\{q_k\}$,其中 i 代表属于哪个模式, k 表示属于模式 i 的第 k 个训练样本, n_i 表示模式 i 的训练样本数,采用下式来修正原型模式 $v_i(n)$:

$$v_i(n) = \frac{\left[v_i(n-1) \cdot r + \sum_{k=1}^{n_i(n-1)} \frac{q_k^i(n-1)}{e_i} \right]}{(r+1)}. \quad (4)$$

其中 $v_i(n)$ 为第 n 次迭代得到的原型模式 v_i , $e_i(n)$ 为用 $v_i(n)$ 识别所有训练样本时误识率最高的模式的误识训练样本数.显然可以看出,式(4)中存在一个未知参数 r ,它决定了一次循环中误识率最高的模式作为反馈量来修正原型模式的力度,所以称之为力度参数,其取值为 $r \geq 1$ 的正整数.力度参数 r 对新算法的训练性能有较大的影响, r 过大,会导致新算法振荡,搜索性能差,但如果 r 过小,会使每次循环新算法对原型模式的修正力度很小,从而导致收敛速度缓慢,文中通过实验上的比较发现 $r=6$ 比较合适,但这仅对我们特定的训练样本有效,如果改变训练样本,力度参数 r 需重新确定.基于信息叠加的学习算法的具体算法描述如下:

- (1) 计算原型模式 $v_i(0) = \sum_{k=1}^{n_i} q_k^i / n_i, i = 1, 2, \dots, M$;
- (2) 将所有训练样本 q_k^i 都送进协同神经网络识别,如果识别率达到要求或迭代步数满则算法结束,输出结果原型模式;
- (3) 否则根据式(4)修改原型模式,然后转(2)继续.

3 实验结果与讨论

本文将 400 多幅通过实际采集得到的数字二值图作为训练样本集,协同神经网络识别的参数设置为: $\lambda_k = B_{kl} = C = 1, 0, k, l = 1, \dots, M$. 所有实验都在 PII400 的 Matlab5.2 环境下完成.

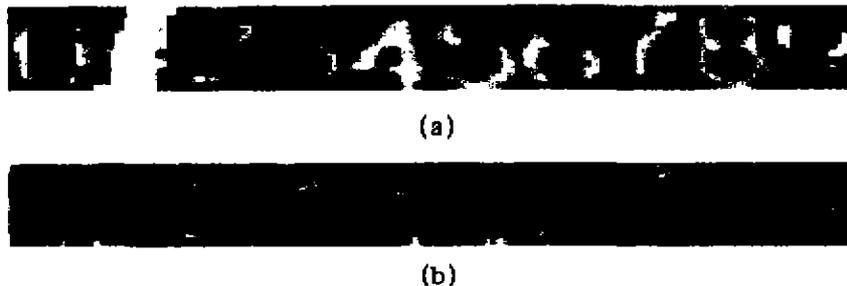


图1 根据基于信息叠加的学习算法计算得到的
(a)原型模式及(b)伴随模式

Fig.1 Calculated prototype patterns and adjoint patterns through learning algorithm based on the superposition of information

(a) prototype patterns, (b) adjoint patterns

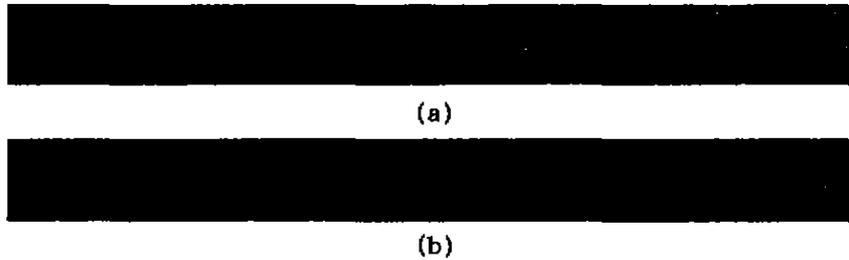


图 2 根据基于遗传算法的原型模式选取算法计算得到的(a)原型模式及(b)伴随模式

Fig. 2 Calculated prototype patterns and adjoint patterns through selection algorithm of prototype patterns based on genetic algorithm
(a) prototype patterns, (b) adjoint patterns

图 1 和 2 分别是利用基于遗传算法的原型模式选取算法^[4]和新算法在给定训练样本集条件下学习获得的原型模式及其相应的伴随向量。图 3 是利用新算法进行网络学习时的训练性能曲线,从图中曲线容易看出,新算法对实际应用中的数字样本的训练性能达到了 97.8% 的识别率,对测试样本的识别率达到了 97.3%。仔细分析训练时未能正确识别的样本可以发现样本具有缺损和平移特点,我们知道协同神经网络的抗缺损能力很强,但其对平移很敏感。造成这种情况的主要原因是图像分割算法有问题,所以提高协同神经网络的识别性能的最简单办法就是改进分割算法,使送进协同网络的样本质量提高。另外可考虑在协同神经网络中加入不变性识别能力,从而屏蔽掉平移对识别造成的影响。图 4 是利用基于遗传算法的原型模式选取算法进行网络学习时的训练性能曲线。对比图 3 和图 4 中的训练性能曲线可以看出:首先,新算法和

基于遗传算法的原型模式选取算法在训练原型模式时都迭代了 200 代,但新算法的训练性能达到了 97.8% 的识别率,而基于遗传算法的原型模式选取算法的训练性能仅达到 92.4%;其次,新算法迭代 200 代,仅需要不到 10min 时间,而基于遗传算法的原型模式选取算法同样迭代 200 步就需要 50h,由上面的对比容易发现新算法具有比遗传算法更强的全局搜索能力,所以其最优解的搜索效率比遗传算法高得多,而且其搜索速度快得多。基于信息叠加的学习算法是目前协同神经网络的原型选取算法和学习算法中最为高效和实用的算法。遗传算法是一种较为普适的搜索算法,在无穷迭代条件下可得到全局最优解,但是由于在计算机模拟时迭代次数肯定是有限次,所以遗传算法并不一定能够得到全局最优解,其效率甚至远不如为特定问题开发的搜索算法。这主要是由于遗传算法没有能够利用特定问题所具有特定信息和规律来优化搜索路

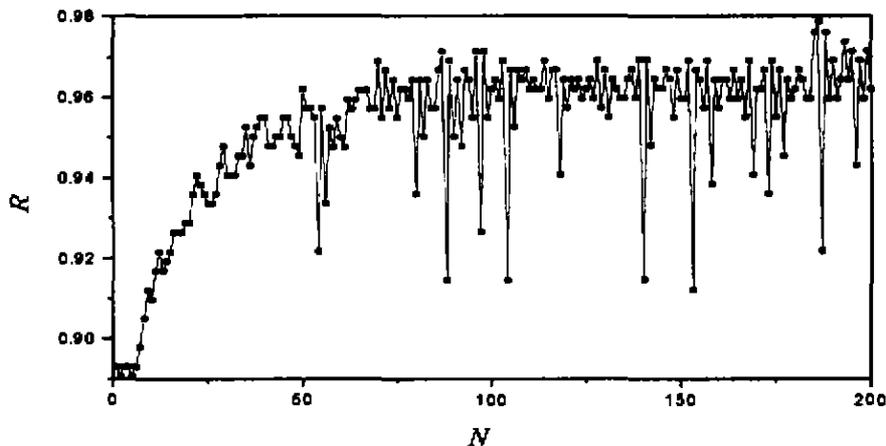


图 3 基于信息叠加学习算法的训练性能(横坐标 N 代表迭代步数,纵坐标 R 代表识别率)

Fig. 3 Training performance of learning algorithm based on superposition of information (N means iteration number, R means recognition rate)

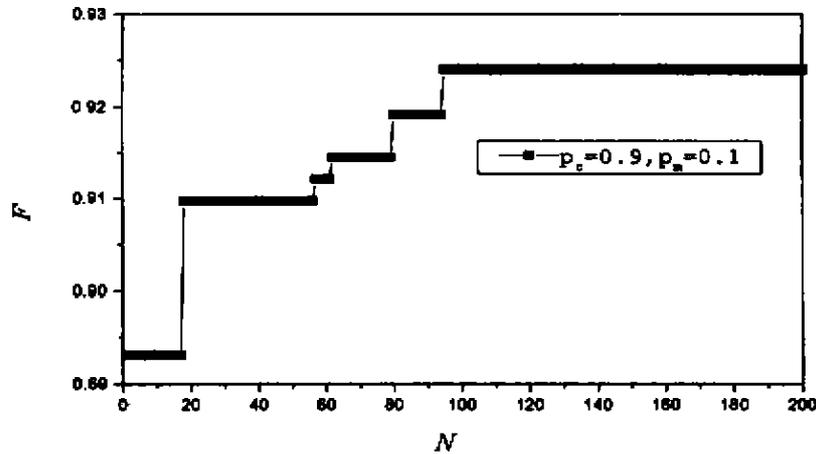


图4 基于遗传算法的原型模式选取算法的训练性能(横坐标 N 代表迭代步数, 纵坐标 F 代表适应值)

Fig. 4 Training performance of selection algorithm of prototype patterns based on genetic algorithm (N means iteration number, F means fitness)

径. 解决的方法之一是遗传算法的参数动态化或开发搜索能力更强的遗传算子, 但遗传算法的搜索仍旧是随机的. 方法之二是设法将特定问题的有用信息和规律作为遗传算法的几条规则而加以利用, 这将大大强化遗传算法的搜索能力, 当然这将遗传算法不再是纯粹的随机搜索算法. 如果我们将具有明显平移的样本剔除, 新算法能够很快搜索到识别率达到 100% 的原型模式, 这说明加入不变性识别能力后新算法将具有比较好的性能. 最后, 在利用新算法时有一个参数 r 需设定, 应用中发现参数 r 是比较重要的参数, 直接影响训练效果. 本文中根据经验设定为 $r=6$.

4 结论

本文提出了一种新的基于信息叠加的迭代学习算法, 该算法利用协同神经网络中的原型模式具有信息的可加性, 将学习中误识率最高的模式作为反馈量来修正原型模式. 利用实际采集得到的样本对算法进行的测试表明: 新算法具有最优搜索能力强, 训练时间短的特点, 具有很强的实用性. 将新算法与基于遗传算法的原型模式选取算法进行了性能比较, 结果表明新

算法具有比遗传算法更强的全局搜索能力, 训练效率高得多. 新算法中有一个待定参数 r , 本文是根据经验确定的, 没有确定怎样的参数选取可使训练效果最佳, 为了进一步提高算法的学习效果, 应该考虑用一种算法来确定参数 r .

REFERENCES

- [1] Haken H. *Synergetic Computers and Cognition-A Top-down Approach to Neural Nets*. Berlin: Springer-Verlag, 1991: 21-40
- [2] Wagner T, Boebel F G. Testing synergetic algorithms with industrial classification problems, *Neural Networks*, 1994, 7(8): 1313-1321
- [3] WANG Hai-Long, QI Fei-Hu. A learning algorithm of synergetic neural network based on cluster algorithm, *Journal of Shanghai Jiaotong University* (王海龙, 戚飞虎. 基于聚类法的协同神经网络学习算法, 上海交通大学学报), 1998, 32(10): 39-41
- [4] WANG Hai-Long, QI Fei-Hu, ZHAN Jin-Feng. Selection algorithm of prototype pattern based on genetic algorithm, *Computer Engineering* (王海龙, 戚飞虎, 詹劲峰. 基于遗传算法的原型模式选取算法, 计算机工程), 2000, 26(10) (to be published)