

Min-Max 模糊神经网络的应用研究

周 越

(上海交通大学 图像处理与模式识别研究所 上海 200030)

相敬林

(西北工业大学 航海工程学院 西安 710072)

杨 杰

(上海交通大学 图像处理与模式识别研究所 上海 200030)

摘 要 本文从结构和算法上研究了 Min-Max 模糊神经网络, 找出了其固有的局限性, 相应地提出了一系列的改进措施, 得到了好的仿真结果, 并运用了它作为分类器对实际信号进行了分类, 得到了满意的结果. 由于网络参数对网络的性能影响很大, 本文应用遗传算法对网络算法中的参数进行了优化处理, 与优化前相比, 网络性能有明显地提高.

关键词 模糊隶属度函数, 神经网络, 分类器, 遗传算法

中图法分类号 TP183

1 引 言

本文研究的 Min-Max 模糊网络(简称 MMFN), 是一种模糊决策器, 其模糊隶属度函数是根据样本的空间分布决定的, 每一类样本在空间形成一系列的超盒, 当不同类的超盒发生重叠时, 网络可自动调整超盒的大小, 消除重叠区. 虽然 MMFN 的自学习和自拓展功能不及人工神经网络(ANN)强大, 但它的优势在于: 算法较为简单, 不存在迭代运算; 运算量不会因为输入向量的维数的增大而剧增; 网络结构固定(即隐层节点确定); 另外, 具有继承性, 训练集增加新的训练样本时无需重新学习, 只学习新的样本即可.

本文通过研究 MMFN 的结构和算法, 解决了原有算法中存在的不足之处, 使得 MMFN 形成空间超盒更加合理. 文中将通过仿真实验说明改进算法(IMMFN), 并且将其作为分类器验证分类效果, 得到了满意的结果. MMFN 和 IMMFN 的性能关键取决于网络的两个参数, 因此在一定代价下应用遗传算法对它们进行优化, 能使分类效果明显提高.

2 Min-Max 神经网络

2.1 Min-Max 模糊神经网络结构和基本算法

本文将此模糊神经网络作为分类器. 它主要分为以下几个部分^[1](如图 1 所示): 第一层为输入层; 第二层为超盒的形成和调整层; 第三层为模糊决策层(输出层). 它的工作机理及算法^[2]如下所述.

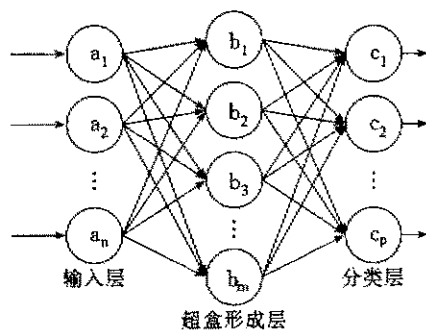


图 1 模糊神经网络结构

2.1.1 超盒形成与扩展

若训练集 $T = \{X_k, d_k\}$, $d_k \in \{1, 2, \dots, m\}$
 $X_k = \{x_{k1}, x_{k2}, \dots, x_{kn}\} \in I^n$;
 其中, X_k 是特征向量, d_k 是每一个 X_k 所对应的类别, 共有 m 个类. 这一步是该分类器的关键, 不同类的特征向量空间将由各自一系列的超盒构成. 而超盒是由输入向量的每一个分量的最小值 v_{ji} 和最大值 w_{ji} ($i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$) 构成. 在建立超盒的过程中, 形成新的超盒的条件是

$$n\theta \geq \sum_{i=1}^n (\max(w_{ji}, x_{ki}) - \min(v_{ji}, x_{ki})) \quad (1)$$

若不满足条件, 原来的超盒将自动按下面方式扩展而不重新生成超盒:

超盒的最小点修正为

$$v_{ji}^{\text{new}} = \min(v_{ji}^{\text{old}}, x_{ki}) \quad \forall i = 1, 2, \dots, n$$

超盒的最大点修正为

$$w_{ji}^{\text{new}} = \max(w_{ji}^{\text{old}}, x_{ki}) \quad \forall i = 1, 2, \dots, n$$

2.1.2 检测重叠区

所谓重叠区是指超盒在扩展时, 不同类别的超盒的重叠部分. 给定一个尺度量 $0 < \theta < 1$, 对于不同类的超盒由以下四种情况来判断:

- (1) $v_{ji} < v_{ki} < w_{ji} < w_{ki}$,
 $\theta^{\text{new}} = \min(w_{ji} - v_{ki}, \theta^{\text{old}})$;
- (2) $v_{ki} < v_{ji} < w_{ki} < w_{ji}$,
 $\theta^{\text{new}} = \min(w_{ki} - v_{ji}, \theta^{\text{old}})$;
- (3) $v_{ji} < v_{ki} < w_{ki} < w_{ji}$,
 $\theta^{\text{new}} = \min(\min(w_{ki} - v_{ji}, w_{ji} - v_{ki}), \theta^{\text{old}})$;
- (4) $v_{ki} < v_{ji} < w_{ji} < w_{ki}$,
 $\theta^{\text{new}} = \min(\min(w_{ji} - v_{ki}, w_{ki} - v_{ji}), \theta^{\text{old}})$.

若 $\theta^{\text{old}} - \theta^{\text{new}} > 0$, 则 $\theta^{\text{old}} = \theta^{\text{new}}$, 记 $s = i$; 同时说明有重叠区.

2.1.3 消除重叠区

- (1) $w_{ji}^{\text{new}} = v_{ks}^{\text{new}} = \frac{w_{ji}^{\text{old}} + v_{ks}^{\text{old}}}{2}$;
- (2) $w_{ki}^{\text{new}} = v_{js}^{\text{new}} = \frac{w_{ki}^{\text{old}} + v_{js}^{\text{old}}}{2}$;
- (3a) 且 $(w_{ks} - v_{js}) < (w_{js} - v_{ki})$,
 $v_{js}^{\text{new}} = w_{ks}^{\text{old}}$;
- (3b) 且 $(w_{ks} - v_{js}) > (w_{js} - v_{ki})$,
 $v_{js}^{\text{new}} = w_{ks}^{\text{old}}$;
- (4a) 且 $(w_{ks} - v_{js}) > (w_{js} - v_{ki})$,
 $v_{ks}^{\text{new}} = w_{js}^{\text{old}}$;
- (4b) 且 $(w_{ks} - v_{js}) < (w_{js} - v_{ki})$,

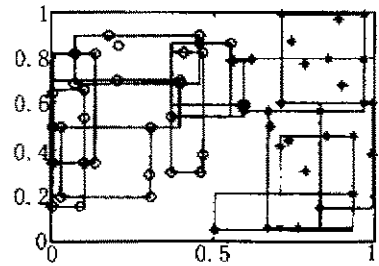
$$v_{ks}^{\text{new}} = w_{js}^{\text{old}};$$

这样就形成了不同类的一系列超盒 B_{jk} .

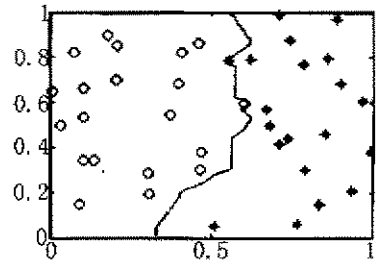
以上过程是学习过程, 它形成了向量 V 和 W , 由向量 V 和 W 就构成了模糊隶属函数:

$$b_j(A_k) = \frac{1}{2n} \sum_{i=1}^n [\max(0, 1 - \max(0, \gamma \cdot \min(1, a_{ki} - w_{ji}))) + \max(0, 1 - \max(0, \gamma \cdot \min(1, v_{ji} - a_{ki})))] \quad (2)$$

其中 γ 为自定义常数, 它控制隶属函数的上下沿的斜率, γ 这里可以取为 $1 \sim 6$. 图 2 给出了由此算法得到的超盒分布与两类样本的边界线.



(a) 两类样本形成的超盒



(b) 样本分布与形成的分类线

图 2 仿真实验结果(图中坐标是归一化无量纲的)

训练完毕之后, 在判决过程中, 将被测试点 x_p 分别代入不同类的超盒 b_j 中求出各自的隶属度, 判该点为隶属度值最大的那一类, 即

若有

$$u_{jk} = \begin{cases} 1 & b_j \in c_k \\ 0 & b_j \notin c_k \end{cases}$$

则可得

$$c_k(x_p) = \max_{j=1}^m b_j(x_p) \cdot u_{jk};$$

$$\text{那么 } C = \max_{k=1}^m c_k(x_p).$$

其中 K 表示总类别数, m 表示对于某一类超盒的个数.

2.2 MMFN 分析研究与改进方法 IMMFN

2.2.1 超盒的初始化

本文将训练集样本随机排序. 考虑到超盒形成

的合理性,在形成超盒时,产生第一个超盒时加入以下两个约束条件:

第一,由距离其它类各点最远的点来初始化 V 和 W .

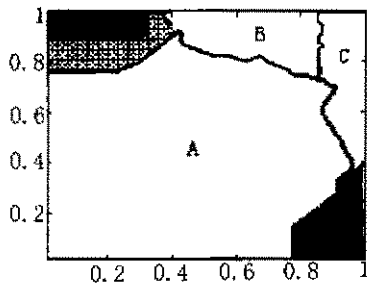
第二,第一个超盒由第一个中选取的点和离其最近的点形成;是否产生超盒由(1)式决定,若不能产生超盒,与同集中次最近的点产生超盒.这一方法减少了不同类别超盒的重叠.

2.2.2 错判区域的消除

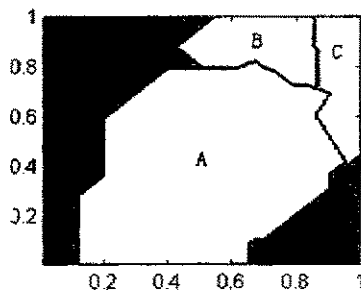
对(2)式的研究可知,当 $x - w > 1/\gamma$ 或 $v - x > 1/\gamma$ 时,隶属度在该方向的值恒为 1,即 x 只要落入此区域内,不管与 w 或 v 的绝对距离有多远,隶属度值都没有区别,这将导致不可判决区变为错判区,如图 3(a) 所示.图中的浅色阴影部分被错判为 B 类,深色阴影部分错判为 C 类.为了解决这一问题,本文将(2)式修正为下式:

$$b_i(A_k) = \left[1 - \frac{1}{\alpha_i} \sum_{j=1}^n \max(1, (1 - \gamma(v_{ji} - x_i))) \right] \times \left[1 - \frac{1}{\beta_j} \sum_{i=1}^n \max(1, (1 - \gamma(x_i - w_{ji}))) \right] \quad (3)$$

其中, α 和 β 分别是 x 向量中 $x_i > v$ 和 $x_i < w$ 的个数.如此修改的目的是为了使 w 和 x 或 v 和 x 之间的距离大小直接与隶属度函数发生关系,从而解决将不可判区域变为错判区的问题.图 3(b) 给出了修改后的判决区域图.



(a) 阴影部分表示误判区



(b) 阴影部分表示不可判区

图3 算法改进前与修改后的对比

2.2.3 “飞点”的处理方案

对训练样本集中“飞点”(即该点远离同类样本集的其它点)的解决方法为,第一步,检测出“飞点”,由于该点与同类中其它的点建立超盒时,超盒会因此而产生很大误判区和死区(即某类中的两个点在各个维上包含其它类中的两个点,这时超盒的重叠将无法去除);第二步,将“飞点”单独建立一个超盒.

2.2.4 θ 的选取

由式(1)可知 θ 的选取影响超盒的形成.实验研究发现, θ 选取得过小,会使判决区域很小, θ 选取得过大会使误判区域增大.由于不同类别的样本在空间中散布是不一样的,那么对不同类样本用同一种 θ 来形成超盒显然是不合适的,这样可能造成某类可判决区过小或过大.因此本文在训练样本的预处理阶段对每一类样本做统计处理求得样本的空间散布量 ζ (即单位空间体积 θ 中包含的样本点个数 N),制定 θ 与 ζ 的关系 Γ , $\theta = \Gamma(\zeta)$,使得对于不同的类利用不同的 θ .

2.2.5 代价函数的形成

由式(1)和式(3)可以看出 θ 和 γ 决定了最终的模糊隶属度函数. θ 直接影响超盒的大小,从前面的公式知 θ 还影响超盒的重叠和重叠的消除,对于给定一个 θ ,就有 N 个超盒与之对应,在超盒中各点的模糊隶属度值为 1.在盒外的延拓是由 γ 决定的,它将影响非盒内样本点的隶属度,也决定该样本点的类别. γ 过大会导致错判区域增大(即错判点数 M 增大),而 γ 过小使得可判区域过小,即空间的自延拓能力低.由此可知, θ 和 γ 是具有一定的矛盾性.据此建立一个代价函数 $\Psi(\theta, \gamma)$:

$$\Psi(\theta, \gamma) = (1 - \lambda) \cdot M + \lambda \cdot N \quad (4)$$

改变 θ 和 γ 寻找 $\Psi(\theta, \gamma)$ 的全局最小值,也就是 θ 和 γ 的最优值, λ ($0 < \lambda < 1$) 为一常数.

2.2.6 需要说明的两点

当 K 类的隶属度值相等时,本文将该点定义为不可分点.

当 K 类的隶属度值很小时,即 $c_1 \approx c_2 \approx \dots \approx c_K$ 时,本文将该点定义为不可分点,这是由于训练集样本分布决定的.

2.2.7 训练集样本排序

IMMFN 方法对训练集样本的排列顺序极为敏感,形成超盒顺序不同也会导致判决域不同从而使得分效果不同.为此本文对一组实验数据做了 100 次和 1000 次的分类研究.样本集一定,样本顺序随机排列;在此条件下得到了在不同超盒形成顺序下该网络得到的正确识别率的置信度.从表 1 可以看

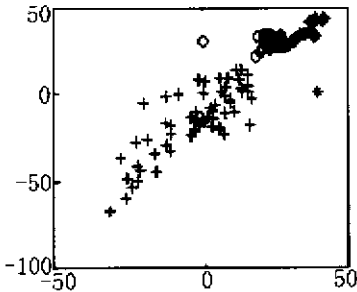
出错误概率基本是一致的.可以这样认为,当单次由该 IMMFN 识别后得到的正确识别率为 85% 时,其置信度为 95%.

表 1 识别率及相应置信度

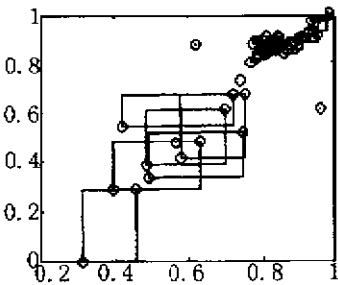
错判率	100 次	1 000 次
5%	27%	27%
10%	75%	74%
15%	95%	94%
20%	96%	96%

2.3 分类实验研究

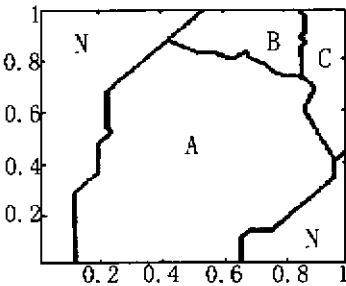
将 2 维三类 180 个样本进行归一化处理,将每一类中随机抽取 20 个样本作为训练集,其余 40 个作为检验集.将训练集中的样本也随机排序,送入 IMMFN 进行训练.



(a) 降二维特征向量分布



(b) 训练样本产生的超盒



(c) 类划分区域

图 4 降二维特征向量分布与超盒、类划分区域

4(a) 中可以看出 A 类散布较大, B 和 C 散布较小, N 为不可分区域, 在形成超盒时应用变 θ 的方法得到较好的分类效果. 根据实验统计结果可知, 在三类识别中, 对于训练集的检验正确识别率均为 100%. 对于检验样本集, 平均正确识别率达到 95%, 如表 2 所示, 与以多层感知机作为分类器的分类效果是相似的.

表 2 正确识别率比较

训练样本数				对训练样本的 正确识别率(%)				检验样本数				对检验样本的 正确识别率(%)			
I	II	III		I	II	III	R_{cor}	I	II	III		I	II	III	R_{cor}
20	20	20		100	100	100	100	40	40	40		98	95	92	95

3 基于遗传算法的网络优化

从上一部分的讨论可以明显看出 θ 与 γ 直接影响到隶属度函数, 也就是说决定了类边界, 因此对两者的优化是很有必要的. 遗传算法是一种智能式搜索、渐进式优化寻找全局最优解, 并且其搜索算法是一种并行式算法, 因此搜索速度快且不易落入局部极大点. 本文采用此方法对在 $[0, 1] \times [0, 1]$ 上分布均匀的两类样本进行分类试验, 将 θ 与 γ 进行优化 (由于是均匀分布, 故考虑两类样本公用同一个 θ). 具体步骤如下:

首先, 对变量进行编码. 我们可以设定两者的大致范围, 即 $0 < \theta < 1$ 和 $0.1 < \gamma < 10$. 而实际上为了编码方便, 本文对 θ' 和 γ' 进行编码, 而 $\theta' = A_1 \cdot \theta$, $\gamma' = A_2 \cdot \gamma$. 考虑到编码分辨率的要求, 本文用八位二进制数来对其编码. $[00000000, 11111111]$ 与 $[\theta'_{min}, \theta'_{max}]$ 和 $[\gamma'_{min}, \gamma'_{max}]$.

根据模式定理, 遗传过程中存活的都是长度短、阶次低、适应度高的模式, 因此在多参量编码时应采用交叉编码的方法:

假设, 对 θ 的单独编码为: $a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8$, 对 γ 的单独编码为: $b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8$, 那么对 θ 与 γ 的联合编码为: $a_1 b_1 a_2 b_2 \cdots a_8 b_8$.

第二、建立适应度函数. 根据代价函数, 本文将其转化为一个最大值问题. 若每一类训练集的样本个数为 M_1 , 共有 M_c 类, 则最多建立 $M_c \cdot M_1$ 个超盒; 若每一类检验集的样本个数为 M_2 , 则错判的个数不会超过 $M_c \cdot M_2$. 这样存在一常数 $M_c \cdot (M_1 + M_2)$, 使得当 θ 与 γ 取到最优值时, Γ 取到最小值 Γ^* , 其中,

为了较直观地分析, 给出二维时的情况, 从图

$$F = M_c \cdot (M_1 + M_2) - (1 - \lambda) \cdot M - \lambda \cdot N \quad (5)$$

第三、确定种群的规模. 遗传算法中一个重要参数是每一代种群的个体数. 很明显, 个体数目越多, 搜索范围越广, 容易找到全局最优解. 然而个体数目太多, 每次迭代时间也长. 至今对这一参数的选取从理论上还未有严格和科学的方法. 本文依照 Min-Max 神经网络的具体情况, 随机选取 30 个个体组成种群.

第四、复制. 每一代种群中, 被复制的个体数目由复制概率 P_r 决定, P_r 一般为 0.1 ~ 0.2, 将适应度大的 10% ~ 20% 的个体复制到下一代, 相应地有 10% ~ 20% 适应度小的个体被淘汰, 以保持种群的大小.

第五、交换. 交换概率 P_e 一般为 0.5 ~ 0.8, 也就是说, 约 50% ~ 80% 的个体要执行交换. 然后, 采用轮盘选择的方法, 按适应度的大小选择被交换的个体, 依次两两进行交换. 交换点是随机选取的. 本文采用两点交换法, 将两点之间的位全部交换.

第六、突变. 本文选取的突变概率为 0.01%, 突变的个体的选择和该个体发生突变的位都是随机确定的, 突变位按照 50% 的概率随机变化. 实验中发现, 突变难得作用比交换的作用小得多.

第七、终止条件. 本文采用最大迭代次数和判断适应度的变化两个条件共同决定终止操作.

表 3 分类性能比较

参 数	经验值	优化值
θ	0.2	0.1435
γ	4.0	2.4673
超盒个数	(28, 30)	(26, 25)
错判概率	20%	10%

通过应用遗传算法对 θ 与 γ 的优化, 使得在超盒数最少、错判数最小的意义下通过 100 次迭代收敛后确定了一个最优的隶属度函数即产生了最优的空间边界曲面. 由于每一次迭代时, 训练样本的顺序是随机的, 为了使结果具有可比性, 记录下与最优参数 (θ^*, γ^*) 对应的训练样本顺序. 在训练样本顺序相同的情况下, 最优参数对于多类变 θ 的情况, 可对每

一类的 θ 进行编码, 同时最优参数 (θ^*, γ^*) 和经验参数 (θ_0, γ_0) 形成的隶属度函数在分类时的性能大不一样, 明显前者优于后者, 如表 3.

4 分析与讨论

本文采用的分类器的优越性在于:

(1) 该神经网络的学习过程没有迭代运算, 因此运算量小, 大大节省了时间, 从而可以作到在线实时处理.

(2) 对于已完成的学习网络进行增补学习时, 无需重新学习所有样本, 只需学习新添的样本, 并且对以前的学习结果不会破坏, 具有继承性.

但是这种方法也存在着不足, 在研究过程中发现, 它要求学习样本要具有较为广泛的代表性, 这样才能使得超盒能覆盖整个特征子空间; 其次, 对于类中嵌套问题以及 θ 和 γ 的选择仍需要做进一步的研究; 由于训练样本的排序是随机的, 这也给空间边界曲面带来了不确定性, 从而分类结果具有一定的随机性.

5 结 论

本文详细研究了 Min-Max 模糊神经网络的结构和算法, 提出并详尽地论述了一系列的改进措施, 通过仿真实验和对有限实船辐射噪声的分类实验, 表明改进措施明显提高了分类的正确识别率; 研究发现, 由于这种方法算法相对于神经网络简单, 因此具有实际的应用前景. 另外, 通过应用遗传算法对网络参数优化使得网络性能得到提高.

参 考 文 献

- [1] Simpson P K. Fuzzy Min-Max Networks - Part 1: Classification. IEEE Trans on Neural Networks, 1992, 3(5): 776-786
- [2] Simpson P K. Fuzzy Min-Max Networks - Part 2: Clustering. IEEE Trans on Fuzzy Systems, 1993, 1(1): 32-45

STUDY ON THE APPLICATION OF MIN-MAX FUZZY NEURAL NETWORK

Zhou Yue

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030)

Xiang Jingling

(College of Marine Engineering, Northwestern Polytechnical University, Xi'an 710072)

Yang Jie

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030)

ABSTRACT

In this paper, the structure and algorithm of Min-Max fuzzy neural network (MMFN) are studied in detail. A series of steps for overcoming some intrinsic limitation of MMFN are proposed and the emulation results are good. Tests on real signals have shown that better results can be obtained by using the improved methods. Some parameters of the MMFN are optimized by using the genetic algorithm, so that the performance of the net can be greatly improved.

Key Words Fuzzy Membership Function, Neural Network, Classifier, Genetic Algorithm