

# 基于模拟退火遗传算法的模糊分类器参数优化及其应用<sup>1</sup>

周 越 相敬林\* 杨 杰

(上海交通大学图像处理与模式识别研究所 上海 200030)

\*(西北工业大学航海工程学院 西安 710072)

**摘 要** 该文从结构和算法上研究了 Max-Min 模糊神经网络 (MMNN), 找出了其固有的局限性, 相应提出了一系列的改进措施形成改进 MMNN 算法。为了更好地提高网络的性能, 同时考虑到优化算法的收敛速度, 本文提出了基于模拟退火遗传算法的网络参数优化方法, 通过计算机仿真, 证明了该方法是可行的。最后, 运用它作为分类器对实际的船舶辐射噪声进行了分类实验, 与 BP 等算法进行了比较, 显示出其独特的优越性。

**关键词** 模糊隶属度函数, 神经网络, 模拟退火算法, 遗传算法, 分类器

**中图分类号** TP391.4, TN-052

## 1 引 言

神经网络之所以被广泛地应用于模式识别当中, 是因为它具有较好的自学习和拓展功能, 这是其它方法所不具备的。随着对神经网络的不断的深入研究, 新的理论不断地应用于其中。在提高网络的自学习和拓展功能的同时也带来了一定的挑战, 那就是网络规模逐渐增大, 结构越来越复杂, 算法也越来越冗长。与当今器件发展水平相比, 复杂神经网络的工程可实现性越来越小。

本文研究的 Max-Min 模糊神经网络 (MMNN), 实际上是一种模糊决策器, 每一类的训练样本在空间形成一系列的超盒, 同类超盒可以相互重叠, 不同类的超盒不能发生重叠, 否则该网络会按照一定准则自动调整超盒的大小, 消除重叠区。由于它具有神经网络的自学习和自拓展功能, 因此称之为 MMNN。虽然 MMNN 的自学习和自拓展功能方面不及神经网络强大, 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, 不受输入向量的维数影响。

本文在考虑到工程可实现性的前提下, 着重研究了 MMNN 的结构和算法, 提出了改进的 MMNN 算法 (IMMNN), 解决了原有算法中存在的不足之处, 使得在特征空间中形成超盒的过程更加合理。另外, 在 IMMNN 的基础上, 考虑到网络参数的优化, 本文提出了对网络的两个关键参数进行优化的算法 (op-IMMNN), 从而使得网络性能大大提高, 大幅度降低了错判概率, 提高了正确识别率。文中将通过仿真实验具体说明优化后的改进算法的特性。最后用表征实际船舶辐射噪声的 20 维特征向量来验证 op-IMMNN, 得到了满意的结果。

## 2 Max-Min 模糊神经网络

### 2.1 MMNN 结构

本文利用模糊神经网络作为分类器。它主要分为以下几个部分<sup>[1]</sup>(如图 1 所示): 第 1 层为输入层; 第 2 层为超盒的形成和调整层; 第 3 层为模糊决策层 (输出层)。它的基本工作机理及算法<sup>[2]</sup>如下所述:

#### (1) 超盒形成与扩展

若训练集  $T = \{X_h, d_h\}$ ,  $X_h = \{x_{h1}, x_{h2}, \dots, x_{hm}\} \in I^n$ ,  $d_h \in \{1, 2, \dots, m\}$ , 其中  $X_h$  是特征向量,  $d_h$  是每一个  $X_h$  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

<sup>1</sup> 1999-08-06 收到, 2000-03-18 定稿

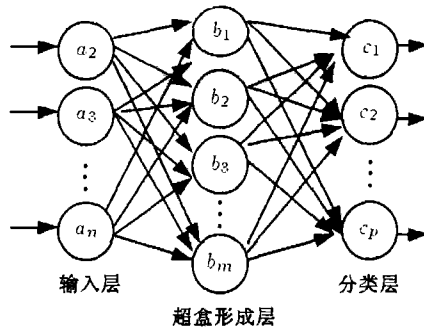


图 1 模糊神经网络结构

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  $V_{ij}$  和最大值  $W_{ij}$   $\{i = 1, \dots, n; j = 1, \dots, m\}$  构成。在建立超盒的过程中, 给定一个尺度量  $0 < \theta < 1$ , 形成新的超盒的条件是

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

若不满足条件, 超盒将自动按下面方式扩展:

$$\text{超盒的最小点修正为 } v_{ji}^{\text{new}} = \min(v_{ji}^{\text{old}}, x_{hi}), \quad \forall i = 1, 2, \dots, n$$

$$\text{超盒的最大点修正为 } w_{ji}^{\text{new}} = \max(w_{ji}^{\text{old}}, x_{hi}), \quad \forall i = 1, 2, \dots, n$$

(2) 检测交叠区, 对于不同类的超盒由以下 4 种情况来判断:

(a)  $v_{ji} < v_{ki} < w_{ji} < w_{ki}$ ,  $\theta^{\text{new}} = \min(w_{ji} - v_{ki}, \theta^{\text{old}})$ ;

(b)  $v_{ki} < v_{ji} < w_{ki} < w_{ji}$ ,  $\theta^{\text{new}} = \min(w_{ki} - v_{ji}, \theta^{\text{old}})$ ;

(c)  $v_{ji} < v_{ki} < w_{ki} < w_{ji}$ ,  $\theta^{\text{new}} = \min(\min(w_{ki} - v_{ji}, w_{ji} - v_{ki}), \theta^{\text{old}})$ ;

(d)  $v_{ki} < v_{ji} < w_{ji} < w_{ki}$ ,  $\theta^{\text{new}} = \min(\min(w_{ji} - v_{ki}, w_{ki} - v_{ji}), \theta^{\text{old}})$ ;

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。

(3) 消除重叠区, 按照以上 4 种情况对应作以下处理:

(a) 当  $v_{ji} < v_{ki} < w_{ji} < w_{ki}$  时,  $w_{js}^{\text{new}} = v_{ks}^{\text{new}} = \frac{w_{js}^{\text{old}} + v_{ks}^{\text{old}}}{2}$ ;

(b) 当  $v_{ki} < v_{ji} < w_{ki} < w_{ji}$  时,  $w_{ks}^{\text{new}} = v_{js}^{\text{new}} = \frac{w_{ks}^{\text{old}} + v_{js}^{\text{old}}}{2}$ ;

(c) 当  $v_{ji} < v_{ki} < w_{ki} < w_{ji}$  且  $(w_{ks} - v_{js}) < (w_{js} - v_{ks})$  时,  $v_{js}^{\text{new}} = w_{ks}^{\text{old}}$ ;

当  $v_{ji} < v_{ki} < w_{ki} < w_{ji}$  且  $(w_{ks} - v_{js}) > (w_{js} - v_{ks})$  时,  $v_{js}^{\text{new}} = w_{ks}^{\text{old}}$ ;

(d) 当  $v_{ki} < v_{ji} < w_{ji} < w_{ki}$  且  $(w_{ks} - v_{js}) > (w_{js} - v_{ks})$  时,  $v_{ks}^{\text{new}} = w_{js}^{\text{old}}$ ;

当  $v_{ki} < v_{ji} < w_{ji} < w_{ki}$  且  $(w_{ks} - v_{js}) < (w_{js} - v_{ks})$  时,  $v_{ks}^{\text{new}} = w_{js}^{\text{old}}$ ; 这样就形成了

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:

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

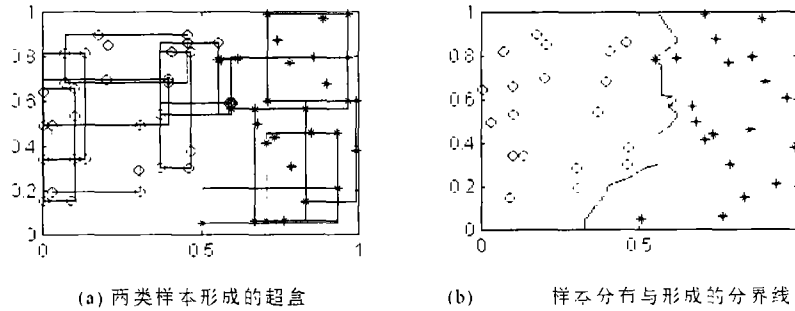


图 2 仿真实验结果 (图中坐标是归一化无量纲的)

其中  $\gamma$  为自定义常数, 它控制隶属函数的上下沿的斜率, 一般  $\gamma$  可以取为 4. 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.

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, 判该测试点归属于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  $c_k(x_p) = \max_{j=1}^m b_j(x_p) \cdot u_{jk}$ , 那么  $C = \max_{k=1}^K c_k(x_p)$ , 其中  $K$

表示总类别数,  $m$  表示对于某一类超盒的个数.

### 2.2 MMNN 的分析研究与改进

(1) 从 (1) 式可以看出训练集中样本排列的顺序也会直接影响超盒的形成, 因此在超盒的形成前, 本文将训练集样本随机排序. 第一个超盒的形成是随机的, 其后点是否产生超盒由 (1) 式决定, 若不能与前一点产生超盒, 它将与同集中最近的点产生超盒. 这一方法大大减少了不同类别超盒的重叠.

(2) 对 (2) 式的研究: 假设各类训练样本形成的子空间为  $\Omega_1^n, \Omega_2^n, \dots, \Omega_k^n$ , 由于训练样本有限, 因此在整个空间  $I^n$  中,  $\Omega_1^n \cup \Omega_2^n \cup \dots \cup \Omega_k^n \neq I^n$ , 也就是说, 存在不可分子空间为  $\Omega_{u1}^n, \Omega_{u2}^n, \dots, \Omega_{um}^n$  (即隶属度为一恒定值, 无法区分其类别). 在 (2) 式中当  $x - w > 1/\gamma$  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, 即将  $\Omega_{u1}^n, \Omega_{u2}^n, \dots, \Omega_{um}^n$  误判为  $K$  类中的某几类如图 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:

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

其中  $\alpha$  和  $\beta$  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得  $w$  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.

(3) 增加了“飞点”(即该点远离同类样本集中的其它点)的解决方法. 方法为: 第 1 步, 检测出“飞点”, 由于该点与同类中其它的点建立超盒时, 超盒会因此而产生误判区和死区(即某类中即将形成超盒的两个点在各个维上包含其它类中形成超盒的两个点, 这时超盒的重叠将无法去除); 第 2 步, 将“飞点”单独建立一个超盒.

(4) 由 (1) 式可知  $\theta$  的选取也影响超盒的形成, 实验研究发现,  $\theta$  选取的过小会使得判决区域很小,  $\theta$  选取的过大会使误判区域增大. 由于不同类别的样本在空间的散布是不一样的,

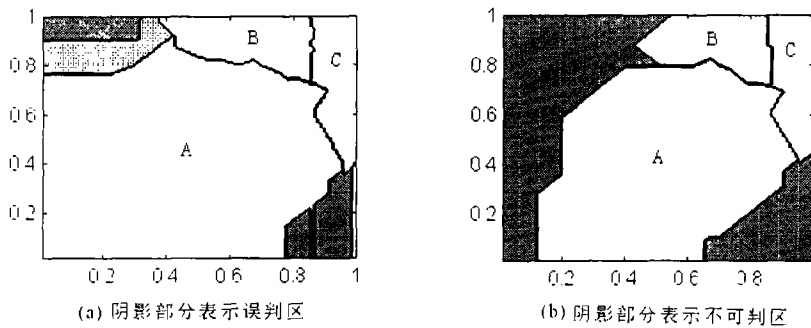


图3 算法改进前与修改后的对比 (图中坐标是归一化无量纲的)

那么对不同类样本用同一种  $\theta$  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  $\xi$  (即单位空间体积  $\Theta$  中包含的样本点个数  $N$ ), 制定  $\theta$  与  $\xi$  的关系  $\Gamma, \theta = \Gamma(\xi)$ , 使得对于不同的类利用不同的  $\theta$ 。

(5) 在训练样本形成超盒后对所有的训练样本进行检测, 把出现错判的样本点记录下来, 将错判点与其同类样本集中最近的的样本点建立一个超盒, 并判断该超盒会不会产生死区, 若会则放弃这个超盒, 寻找次最近点并建立超盒, 这样一直反复作到无错判点为止, 从而弥补判决区的缺陷, 也避免了死区的产生。

(6) 由 (1) 式和 (3) 式可以看出  $\theta$  和  $\gamma$  决定了最终的模糊隶属度函数。 $\theta$  直接影响超盒的大小, 从前面的公式知  $\theta$  还影响超盒的交叠和交叠的消除, 对于给定一个  $\theta$ , 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。在盒外的一定延拓是由  $\gamma$  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。 $\gamma$  过大会导致错判区域增大 (即错判点数  $M$  增大), 而  $\gamma$  过小使得可判区域过小, 即空间的自延拓能力低。由此可知,  $\theta$  和  $\gamma$  是具有一定的矛盾性。据此建立一个代价函数  $\Psi(\theta, \gamma)$  :

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

通过改变  $\theta$  和  $\gamma$  以寻找  $\Psi(\theta, \gamma)$  的全局最小值, 也就是  $\theta$  和  $\gamma$  的最优值。

(7) 需要说明的两点:

(a) 当  $K$  类的隶属度值相等时, 本文将该点定义为不可判点。

(b) 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判点, 这是由于训练集样本分布决定的。

通过以前的分析研究, 发现  $\theta$  和  $\gamma$  对网络的性能起到决定性作用, 因此对  $\theta$  和  $\gamma$  优化极为重要。下面我们将介绍基于模拟退火的遗传算法的网络参数优化问题。

### 3 IMMNN 算法的参数优化

这里仅以三类 2 维的分类问题为例介绍基于标准遗传算法 (GA) 的参数优化和基于模拟退火算法的遗传算法

#### 3.1 基于标准遗传算法 (GA) 的参数优化

(1) 编码方式: 由于优化参数为实数, 因此在编码的过程中必然存在量化误差, 考虑到实际中  $\theta$  的影响作用较  $\gamma$  的影响作用大, 故主要根据  $\theta$  的精度要求来选择二进制码长。因为  $0 < \theta \leq 1$ , 所以在编码长度为  $m$  时, 其量化误差为  $1/2^m$ 。 $\gamma$  的取值为  $[1, 10]$ , 所以  $\gamma$  的量化误差为  $9/2^m$ 。

在实际中, 我们选取  $m = 8$ , 将  $\theta, \gamma$  分别用二进制进行编码。为了在实际中易于操作, 将  $\theta, \gamma$  的二进制码组合成一个十六位的二进制码。

根据遗传规划理论中的模式定理, 遗传过程中存活的都是长度短、阶次低的模式; 因此在多变量编码时, 采用交叉编码方式。由于各参量编码的最高位对参量的特性起的作用最大, 把它们放在最前面, 不易破坏, 存活率较高。

若  $\theta$  的二进制码为  $D_{\theta 0} D_{\theta 1} \cdots D_{\theta 7}$ ,  $\gamma$  的二进制码为  $D_{\gamma 0} D_{\gamma 1} \cdots D_{\gamma 7}$ , 则组合后的二进制码为  $D_{\theta 0} D_{\gamma 0} D_{\theta 1} D_{\gamma 1} \cdots D_{\theta 7} D_{\gamma 7}$ , 其中  $\theta$  的基因占据奇数位,  $\gamma$  的基因占据偶数位。

(2) 解码: 将每个参量对应的十进制数值转变为对应其区间内的实数值。由于  $\theta, \gamma$  取值分别为  $0 \leq \theta \leq 1, 1 \leq \gamma < 10$ , 所以采用如下解码方式:  $\theta = (d_{\theta} + 1)/2^m, \gamma = [(d_{\gamma} \times 9)/2^m] + 1$ , 其中每个参数二进制码值对应的十进制数  $(d_{\theta}, d_{\gamma})$ 。

(3) 计算适应度: 对于适应度函数, 我们取两项之和, 即  $F(\theta, \gamma) = \lambda E_c + (1 - \lambda) E_s$ , 其中  $E_c$  为网络的结构复杂度, 即超盒的个数;  $E_s$  为错判的样本个数;  $\lambda$  是关系到结构复杂度和错误判决影响作用大小的比例值, 若  $\lambda$  取值接近于 0, 则产生较高的结构复杂度和较少的错判样本, 反之, 若  $\lambda$  的值接近于 1, 则会产生较低的结构复杂度和较多的错判样本。

(4) 复制及交换: 根据复制概率, 确定复制个体的个数  $cpn$ , 将具有最小适应度的  $cpn$  个个体用具有最大适应度的  $cpn$  个个体复制。

在选择进行交换的个体时, 采用了如下算法。具体如下:

设第  $i$  个个体对应的适应度值为  $f(i)$ , 按如下定义产生一新的数组  $ff$ :

$$ff(i) = \sum_{k=1}^i f(k), \quad F^S = \sum_{k=1}^n f(k), \quad (n \text{ 为种群中个体的数目})$$

在  $[0, F^S]$  上, 任取  $pn$  个随机数并排序,  $pn$  为进行交换的个体的个数。选择  $ff$  中大于或等于所产生的随机数的第一个个体为进行配对的个体。可见, 采用这种算法具有较大适应度的个体参与交换的概率大, 即进行繁殖的机会大, 这样, 优秀基因就能够遗传到下一代中。

(5) 突变处理: 根据突变概率, 确定进行突变的个体的数目  $tn$ , 在种群中随机选取  $tn$  个个体, 随机确定突变位, 进行突变处理。

(6) 确定是否满足结束条件。若是, 则操作结束; 否则, 回到第 2 步。

### 3.2 基于模拟退火算法的遗传算法

由于标准遗传算法在复制的过程中, 是用具有最大适应度的个体取代最小适应度的个体, 所以, 在迭代到一定次数时, 可能会出现种群中存在较多个具有最大适应度的个体, 这样就使得遗传过程中的交叉操作的结果 (即子代) 与配对池中的个体 (即父代) 不会有较大的差异, 造成收敛于局部最大值或者收敛的速度极为缓慢的现象。同时, 由于在适应度较差的个体中也可能存在适应度较高的个体所不具有的优秀基因, 而利用标准遗传算法则会使较差的个体的优秀基因被淘汰掉。所以, 在此我们利用一种优化的遗传算法——基于模拟退火算法的遗传算法。该方法是将模拟退火算法的思路用于遗传算法中种群的形成, 它比标准的遗传算法具有快得多的收敛速度及更强的逃脱局部极值和避免过早收敛的全局优化能力。由文献 [3] 证明, 该方法所产生的种群的每个个体均以概率 1 收敛到最优解 (一般 GA 方法只能保证群体中的最优个体以概率 1 收敛到最优解) 具体方法如下:

(1) 设定较高的初始温度  $T$  和终止温度  $T_0, i = 1$ 。

(2) 令  $t = T(i)$ , 若  $t > T_0$ , 则执行 (3), 否则, 算法结束。

(3) 设内部最大循环次数为  $stn$ , 若内部循环次数  $j > stn$ , 则  $i = i + 1$ ; 否则, 执行:

(a) 以一定的概率从种群中随机选择个体作为父代, 将其送入配对池中, 随机配对, 进行交叉操作。

(b) 配对产生的个体, 按照一定的概率进行突变处理后成为遗传的下一代 (即子代)。

(c) 计算子代的适应度  $f_s$ ，若子代的适应度大于父代的适应度  $f_f$ ，则用子代代替父代；否则，子代以概率  $\exp((f_s - f_f)/2)$  复制父代。

(d)  $j = j + 1$ ；算法结束；遗传过程中最大适应度曲线如图 4 所示。从图中可以看出，基于模拟退火的遗传算法具有快的收敛速度，由图 4(a) 和图 4(b) 可以看出，对所选的训练集和用于测试的样本集，标准的遗传算法和基于模拟退火的遗传算法均收敛于对应最大适应度为 400.2，参数值  $\theta=0.0547$ ， $\gamma$  取值可以有一定的变化。对如图 5(a) 所示的两类问题运用 MMNN 和 op-IMMNN 两种方法得到的分类边界如图 5(b) 和 5(c) 所示，显然 op-IMMNN 得到的边界线更趋于真值。

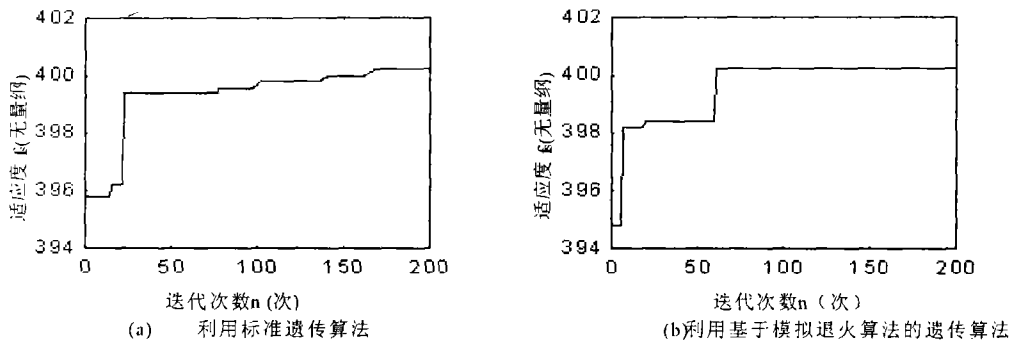


图 4 遗传过程中最大适应度曲线

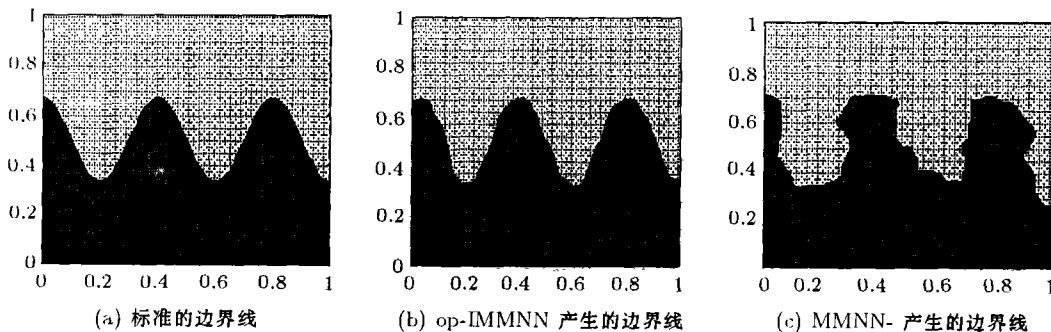


图 5 改进和优化后的算法与原始算法对于两类二维问题的处理结果比较 (图中坐标是归一化无量纲的)

## 4 实验样本空间的形成和分类实验

### 4.1 实验样本空间的形成

为了检验本文提出算法的可行性，和在高维特征空间中自拓展能力，将算法应用于实际的船舶类型识别的分类研究中，将船舶噪声作如下方法的特征提取：

由上图产生的特征空间是 20 维，为了直观考察图中特征提取方法的有效性和考察 MMNN 和 op-IMMNN 方法在二维空间的分类特性，将 20 维的特征向量通过降二维非线性算法降至二维。

### 4.2 分类实验

将由图 6 产生的 20 维、二维 3 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。在训练开始时, 将训练集中的样本也随机排序, 送入 MMNN 和 op-IMMNN 分别进行训练。

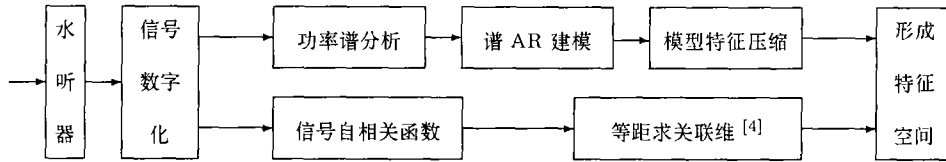
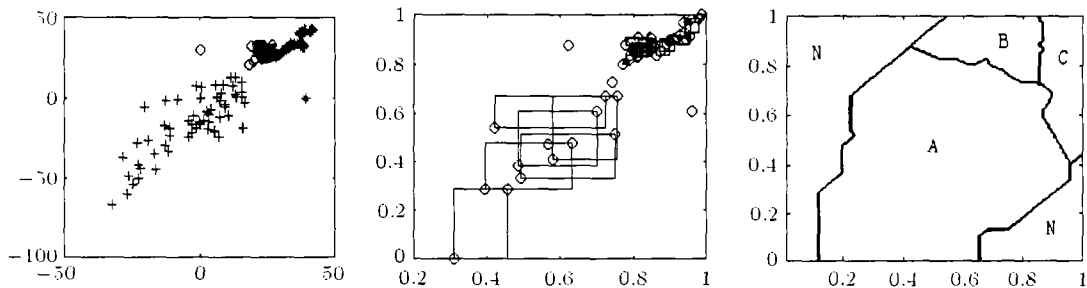


图 6 实验样品的形成

在二维情况下, 从图 7(a) 中可以看出 A 类散布较大, B 和 C 散布较小。根据实验统计结果可知, 在 3 类识别中, 对于 op-IMMNN 方法, 在形成超盒时应用 2.2 节中 (4) 式给出的变  $\theta$  的方法, 得到较好的分类效果, 使得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 97.5%; 而对于 MMNN 方法,  $\theta$  是固定的, 因此训练样本的正确识别率为 96.7%, 无法达到 100%; 检验样本集的正确识别率为 91.7%。

运用 MMNN, op-IMMNN 和 BP 算法进行对 3 类、20 维特征向量分类, 结果如表 1 表 2, 表 3 所示。



(a) 降二维特征向量分布 (b) 归一化训练样品产生的超盒 (c) op-IMMNN 类划分区域

图 7 降二维特征向量分布与超盒、类划分区域 ((b), (c) 中坐标是归一化无量纲的)

表 1 op-IMMNN 实验结果统计

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	95	92.5	90	92.5

( $R_{cor}$  表示总体正确识别率)

表 2 MMNN 实验结果统计

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	95	90	90	91.7	40	40	40	87.5	82.5	80	83.3

(固定  $\theta = 0.2, \gamma = 4$ )

表 3 BP 算法实验结果统计

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	95	92.5	92.5	93.3

从表 1, 表 2 可以看出, 原始的网络在作为分类器时由于其固有的局限性, 导致错误识别率较高, 而经过改进和参数优化后的网络正确识别率大大提高。本文应用多层感知机和改进的 BP 算法对同样的训练样本对其训练, 并且用同样的检验集进行测试, 得到的总体正确识别率为 93.3%。这表明 op-IMMNN 的自学习和自拓展能力是很强的, 但它的时间和规模代价是较小的, 这在工程实践中的应用提供了可能。

## 5 分析与讨论

本文采用 op-IMMNN 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, 这是优于 BP 算法的原因之一。

(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, 这是优于 BP 算法的原因之二。

(3) 经过研究发现, op-IMMNN 方法在处理类间嵌套和交叉同时存在的情况时, 能力强于基于最优 if-then 准则下的模糊分类器<sup>[5]</sup>, 特别是在特征向量维数大于 6 时, 其运算代价远远小于最优 if-then 准则下的模糊分类器。

但是这种方法与其他识别分类的方法存在着同样的不足, 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, 即训练样本具有完备性。这样才能使得超盒能覆盖整个特征子空间; 另外, 对  $\theta$  和  $\gamma$  的选择仍需要做进一步的研究。

## 6 结 论

本文详细研究了 Max-Min 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, 并且应用基于模拟退火的遗传算法对网络参数进行了优化, 通过仿真实验和对有限实际船舶辐射噪声的分类实验, 以及与 BP 算法和最优 if-then 准则下的模糊分类器的比较, 表明改进和优化措施明显提高了分类的正确识别率。另外, 研究发现, 由于这种算法相对于神经网络简单因此具有实际的应用前景, 为实际装置上开发模式识别系统提供了一种较好的途径。

## 参 考 文 献

- [1] P. K. Simpson, Fuzzy Min-Max networks-part 1, Classification, IEEE Trans. on Neural Networks, 1992, NN-3(5), 776-786.
- [2] P. K. Simpson, Fuzzy Min-Max networks-part 2, Clustering, IEEE Trans. on Fuzzy Systems, 1993, 1(1), 32-45.
- [3] 张讲社, 徐宗本等, 整体退火遗传算法及其收敛充要条件, 中国科学 E 辑, 1997, 27(2), 154-164.
- [4] 张涛, 文学章, 吸引子维数计算的几点改进, 浙江大学学报 (自然科学版), 1993, 27(5), 673-679.
- [5] Hisao Ishibuchi, Ken Nozahi, Selecting fuzzy if-then rules for classification problems using genetic algorithm, IEEE Trans. on Fuzzy Systems, 1995, 3(3), 260-270.



## THE PARAMETER OPTIMIZATION OF MMNN BASED ON GENETIC ALGORITHM COMBINED WITH SIMULATED ANNEALING AND ITS APPLICATION

Zhou Yue    Xiang Jinglin\*    Yang Jie

(*Inst. of Image Processing and Pattern Recognition, Jiaotong Univ., Shanghai 200030, China*)

(\**College of Marine Eng., Northwestern Polytechnical Univ., Xi'an 710072, China*)

**Abstract** In this paper, the structure and algorithm of Max-Min fuzzy neural network (MMNN) are studied in detail. In order to get rid of some intrinsic localization of the method and boost up the capability of the MMNN, a series of steps are presented and the improved project (IMMNN) is gained. With a view to making the capability even much better and compressing the time of the convergence, the op-IMMNN is put forward in which the parameters of IMMNN are optimized by genetic algorithm combined with simulated annealing. In the simulation, the result of op-IMMNN is superior over the conventional MMNN's. Finally, a satisfactory result is also obtained when op-IMMNN is regarded as a classifier to distinguish the types of the ships according to their actual radiated noise. Comparing with the neural network based on the back propagation algorithm, the advantages of the op-IMMNN are fully put up.

**Key words** Fuzzy membership function, Neural network, Simulated annealing algorithm, Genetic algorithm, Classifier

周 越: 男, 1969 年生, 博士生, 研究方向: 模式识别、时频分析和神经网络。在《电子科学学刊》、《信号处理》和《数据采集与处理》等刊物上发表多篇文章。

相敬林: 男, 1937 年生, 教授、博士生导师, 研究方向: 信号与信息处理、微弱信号检测和跟踪与定位。

杨 杰: 男, 教授, 博士生导师, 研究方向: 模式识别、数据挖掘和图像处理。