第17卷第1期 电子科学学刊 Vol. 17 No. 1
1995年1月 JOURNAL OF ELECTRONICS Jan. 1995

基于径向基函数网络的雷达目标一维像识别技术研究*

黄德双 保 铮
(西安电子科技大学雷达信号处理重点实验室 西安 710071)

摘要 本文研究了利用递推最小二乘(RLS)方法训练的径向基函数网络(RBFN)用于雷达目标一维像的识别与分类问题. 证明了 RBFN 与 Parzen 窗函数概率密度估计的等价性,指出 RBFN 隐层单元传输函数可以推广到一般的 Parzen 概率核函数或势函数形式。还就高斯、三角、双指数三种核函数讨论了径向基函数网络形状参数 α、递推最小二乘算法的遗忘因子 λ 对识别结果的影响以及 λ 与网络训练时间的关系。

关键词 模式识别,核函数,形状参数,遗忘因子,一维像,径向基函数,递推最小二乘法

1 引 言

自从 Rumelhart[1] 提出了反向传播(BP)学习算法以来,多层感知器网络的训练学习问题就基本上得到了解决. 然而,在实际应用中, BP 算法需要精心选择有效的学习常数 η、动量因子 α 和算法的初始值,这些因素往往很难事先确定,从而给其广泛应用带来了困难[2-5]. 此外, BP 算法本质上是 LMS 算法的推广,它不可避免 LMS 算法的局部极小值问题,且易受到输入模式协方差矩阵特征值散布的影响[6]. 近年来,尽管有人从优化的角度来寻找这些参数的最佳值[7],但也解决不了根本问题. 所以,需要寻找别的有效训练方法.

我们知道,成熟的自适应递推最小二乘(RLS)算法是在 LMS 算法基础上的更精确的递推方法,它不需要象 LMS 算法那样选择自适应步长,如果将其应用于前馈网络的训练学习上,也就相应于避免选择学习常数 η 和动量因子 α[6,8-10]. 此外, RLS 算法在每一时刻的迭代是由输入训练样本形成的协方差阵递推求得权值的精确解(误差为零时的解),因而其迭代的方向性更强,收敛速度也更快[11]. 所以,我们称之为快速 RLS-BP 学习算法[11].

本文主要讨论利用 RLS-BP 学习算法的 RBFN 来实现雷达目标一维像的识别和分类问题[11,12]. RLS 算法不是本文研究的主要内容,而是我们应用的工具. 我们从前人对

1993-06-14 收到, 1994-02-14 定稿
* 国家自然科学基金和高校博士点基金资助项目
黄德双 男, 1964年生,博士后,现在北京理工大学雷达技术研究所工作. 邮编 100081.
保 铮 男, 1927年生,教授,中国科学院院士,博士生导师,主要研究方向:阵列和多维信号处理,雷达系统与雷达成像,模式识别与神经网络.

有导师学习网络的结论中，证明了 RBFN 与 Parzen 窗核函数概率密度估计的等价性，指出 RBFN 隐层传输函数可以推广到一般的 Parzen 概率核函数或势函数形式。本文还给出了高斯、三角、双指数三种核形式的 RBFN 用于五种飞机目标的识别结果。此外，还讨论了 RBFN 的形状参数 a，RLS 算法的遗忘因子 λ 对识别率的影响，以及 λ 与网络训练时间的关系[11]。

2　递推最小二乘反向传播学习算法

在讨论 RBFN 的 RLS-BP 学习算法之前，先讨论单层神经网络加权 RLS 算法的基本原理，然后再讨论如何将此算法用于 RBFN 的训练学习上。

如图 1 所示的具有线性处理单元的单层神经网络，我们首先定义加权的误差目标函数[11]：

$$J(n)=\frac{1}{2}\sum_{t=1}^{n}\lambda^{n-t}\sum_{k=1}^{M}\varepsilon_k^2(t)=\frac{1}{2}\sum_{t=1}^{n}\lambda^{n-t}\sum_{k=1}^{M}(d_k(t)-y_k(t))^2, \tag{1}$$

式中 $\varepsilon_k(t),d_k(t),y_k(t),(1\leqslant k\leqslant M)$ 分别是输出层第 k 个节点的误差信号、期望输出和实际输出；λ 为加权遗忘因子，它的作用是将过去样本对当前估值的影响逐渐遗忘掉，使所估计的参数尽量反映当前时刻样本的特性。通常 λ 的取值范围为 $0<\lambda<1$。这里我们使用重复学习的概念[11]，所以，λ 的意义与自适应滤波里的概念还不完全相同，λ 取值越大，反映对过去的样本影响权重越大。对于样本重复学习的情况，等效到时域来看（如图 2 所示），相当于周期序列的跟踪，前一周期的影响，在后一周期又复现，所以，为了加快收敛速度，λ 取值可以取小一点，这一点在后面模拟中有说明。

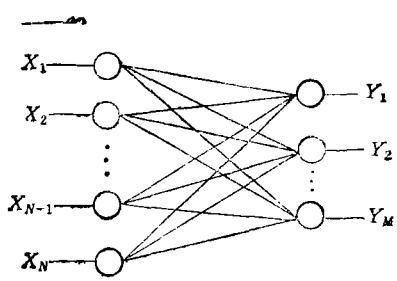


图 1　单层神经网络

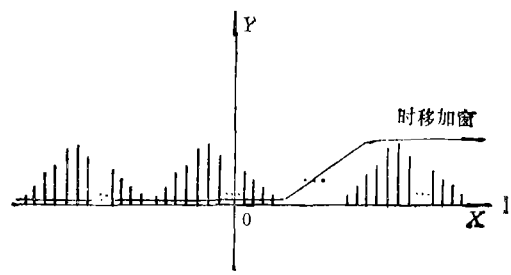


图 2　模式样本加窗学习

如图 1 所示，实际输出 $y_k(t)$ 可以表示为[11]

$$y_k(t)=\sum_{i=1}^{N}W_{ki}(n)x_i(t)=X(t)^TW_k(n),\quad 1\leqslant k\leqslant M, \tag{2}$$

其中 $W_k(n)=[W_{k1}(n),W_{k2}(n),\cdots,W_{kN}(n)]^T$，$X(t)=[x_1(t),x_2(t),\cdots,x_N(t)]^T$。将 $J(n)$ 对 $W_i(n)$ 微分得[11]

$$\frac{\partial J(n)}{\partial \hat{W}_i(n)} = -\sum_{t=1}^{n}\lambda^{n-t}\frac{\partial y_i(t)}{\partial \hat{W}_i(n)}\varepsilon_i(t) \triangleq 0. \tag{3}$$

记 $P(n) = \hat{R}^{-1}(n)$，利用矩阵求逆引理可得网络连接权矢量的 RLS 关系式如下[11]：

$$K(n) = \frac{P(n-1)X(n)}{\lambda + X(n)^T P(n-1)X(n)}, \tag{4}$$

$$P(n) = [P(n-1) - K(n)X(n)^T P(n-1)]/\lambda, \tag{5}$$

$$\hat{W}_i(n) = \hat{W}_i(n-1) + K(n)[d_i(n) - X(n)^T \hat{W}_i(n-1)], \tag{6}$$

$$J(n) = \lambda J(n-1) + \frac{1}{2}\sum_{k=1}^{N}[d_k(n) - X(n)^T \hat{W}_k(n-1)]^2. \tag{7}$$

(4)—(7) 式就是图 1 所示单层神经网络在已知输入模式 $X(t)$、期望输出 $d(t)$ 时所得的网络连接权矢量 RLS 迭代公式[11]. 显然，由(4)—(7)式可见，RLS 算法在每一时刻都能够得到网络权值矢量的精确解，且权值的迭代不需要任何外部控制参数，它比以 LMS 算法为基础的 BP 算法要优越得多.

3 径向基函数网络和 Parzen 核函数

RBFN 就是以一类径向辐射函数为基函数的神经网络[13,14]，它是一两层网络，如图 3 所示. 其中网络中隐层单元的作用相当于对输入模式进行一次变换，将低维的模式输入数据变换到高维空间内，以利于输出层进行分类判别[11,13]. 隐层单元的变换作用实际上也可以看作是对输入数据进行特征提取，因此，这个变换函数应该根据需要来选择[11].

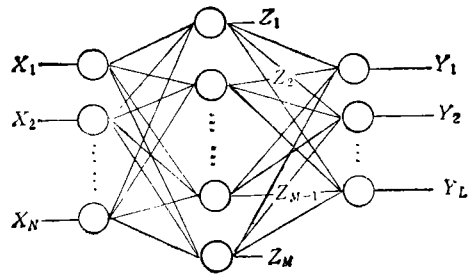


图 3 径向基函数网络结构示意图

根据 D. W. Ruck 等人[15]已经得出的结论，我们可以证明 RBFN 与 Parzen 核函数概率密度估计的等价性，从而将 RBFN 隐层传输函数推广到一般的 Parzen 概率核函数 (kernal function) 或势函数 (poiential function) 的形式. 下面给出 RBFN 概率分类的充分必要性定理[11]：

定理 对于 RBFN，在均方误差最小准则下，网络收敛后，其输出层的输出为模式样本后验概率估计的充分必要条件为隐层单元传输函数是 Parzen 窗密度核函数.

定理的证明可见文献[11].

由此定理我们可以推知，RBFN 的隐层单元传输函数并不局限于高斯核函数形式，它可以推广到满足 Parzen 核函数条件的所有函数形式. 假设 $K(\cdot)$ 为 Parzen 窗函

数，通常它应满足如下条件[11]：

$$(1)\ \sup_{-\infty<y<+\infty}|K(y)|<\infty, \tag{8}$$

$$(2)\ \int_{-\infty}^{+\infty}|K(y)|\mathrm{d}y<\infty, \tag{9}$$

$$(3)\ \int_{-\infty}^{+\infty}K(y)\mathrm{d}y=1, \tag{10}$$

$$(4)\ \lim_{y\to\infty}|yK(y)|=0. \tag{11}$$

因此，所有满足这 4 个条件的函数形式都可以作为 RBFN 隐层单元的传输函数. 这里我们主要讨论高斯、三角、双指数三种核函数形式(其它形式见文献[11]).它们对应的空间三维拓扑图如图 6(a),6(c),6(e) 所示[11].

这 3 种核函数形式为[11]

高斯核

$$z_i(t)=\Phi(|X(t)-S_i|)=\exp\left(-\frac{\sum_{j=1}^{N}(x_j(t)-S_{ij})^2}{2\alpha_i^2}\right),\ (1\leqslant i\leqslant M,\alpha_i>0); \tag{12}$$

三角核

$$z_i(t)=\Phi(|X(t)-S_i|)=\begin{cases}\alpha_i(1-\alpha_i|X(t)-S_i|), & |X(t)-S_i|\leqslant\alpha_i;\\ 0, & |X(t)-S_i|>\alpha_i;\end{cases}(1\leqslant i\leqslant M,\alpha_i>0), \tag{13}$$

双指数核

$$z_i(t)=\Phi(|X(t)-S_i|)=(\alpha_i/2)\exp(-\alpha_i|X(t)-S_i|),\ (1\leqslant i\leqslant M,\alpha_i>0); \tag{14}$$

式中的 $z_i(t)$ 为隐层单元的输出，即径向基函数；$X(t)$ 为第 t 个输入模式矢量；S_i 为隐层中第 i 个单元的变换中心矢量；α_i 为控制径向基函数空间拓扑分布的形状参数. 输出层是一线性处理单元，其第 k 个单元对应的输出为

$$y_k(t)=\sum_{i=1}^{M}W_{ki}(n)z_i(t)=z(t)^TW_k(n),\ (1\leqslant k\leqslant L). \tag{15}$$

假设输出层第 k 个节点对应于第 i 个模式时的期望输出为 $d_k(t)$，则我们按照(1)式定义加权输出误差目标函数. 将(4)—(7)式中的 $X(n)$ 换为 $Z(n)$，即可得到 RBFN 对应的 RLS-BP 学习算法.

4　雷达目标一维像识别结果

在模拟实验之前，我们还需要做如下假设[11]：(1)所有 RBFN 的希望输出为

$$d_i(x)=\begin{cases}1, & x\in\omega_i;\\ 0, & x\notin\omega_i;\end{cases}\ (i=1,2,\cdots,K) \tag{16}$$

其中 x 为输入样本；ω_i 为第 i 个模式类别集合. (2)在所有模拟实验中，除非特别说明，遗忘因子 λ 取 0.99，所有训练样本都作模归一化处理，且都被选作为隐层单元的中心.

4.1 5种飞机目标0°—100°范围一维横向像识别实验

我们对航空航天部二院207所微波暗室中测得的B-52、歼-6、歼-7、F-1 4种飞机目标模型数据和北京航空航天大学微波暗室中测得的强-5飞机目标模型数据进行了一维成像处理. 我们是利用所成一维像进行识别实验的. 假设以机头方向为0°方向，取0°—100°范围模型数据成出其对应的一维横向像（F-1飞机只有−16°～16°范围数据），从其中每隔3°（F-1每隔0.5°）抽取一个样本，形成30个训练样本；B-52、歼-6、歼-7、强-5、F-1剩下的测试样本分别为138、150、113、162、18. 图4和图5分别画出了B-52飞机对应的训练样本和测试样本一维横向像的分布情况. 如果我们根据(16)式的希望输出，选择最大值输出作为判决输出，就可得图6所示的3种核函数形式对B-52飞机在无噪情况下的识别判决置信度曲线[11]，其中判决曲线为输出节点对应所属目标期望值的实际输出值；而比较曲线为其它输出节点中对应的最大输出. 显然这两条曲线离得越远表明识别的结果越可信，当它们相交时，就发生了错判，置信度下降. 由图可见，3种核函数分别以99.38%，98.3%和100%的识别率分辨出B-52飞机目标.

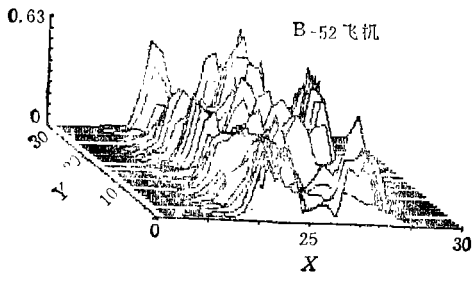


图4 B-52飞机训练样本分布情况

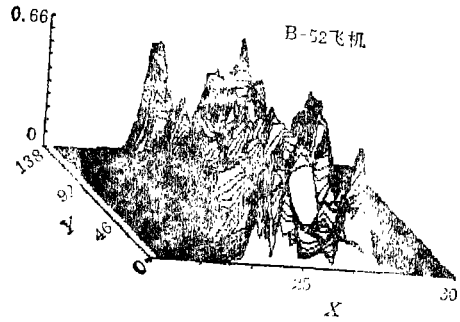
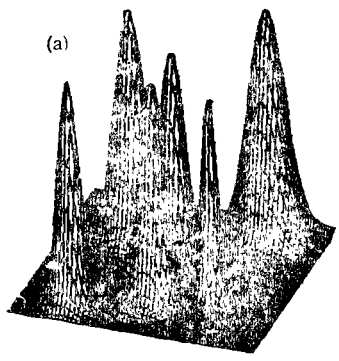
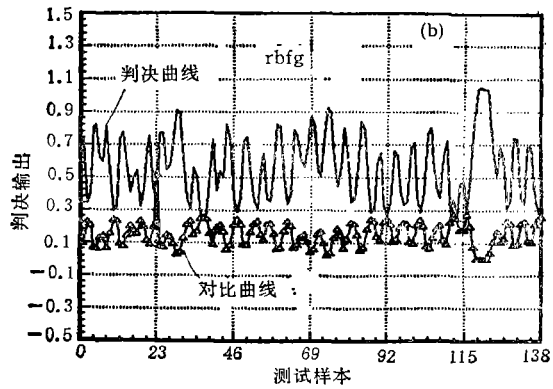


图5 B-52飞机测试样本分布情况

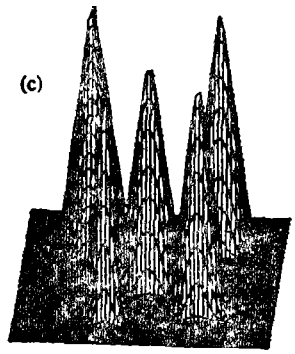
表1给出了3种核函数形式在不同信噪比测试样本情况下，5种飞机目标测试的统计识别结果. 由此表可以看出，一维横向像在信噪比较高时，其识别率是很高的. 这是由于横向像中的某一元素为目标相应窄纵条内各散射点回波之和，这些散射点回波间的相



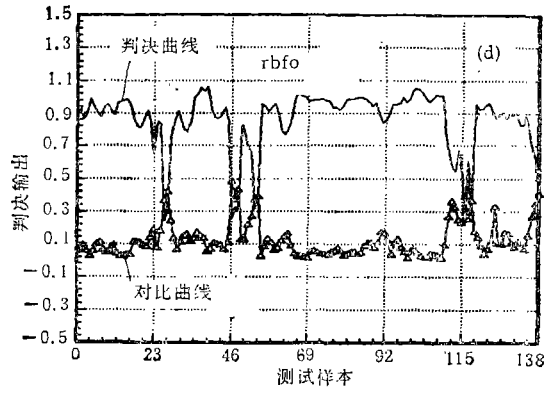
(a) 高斯核拓扑分布



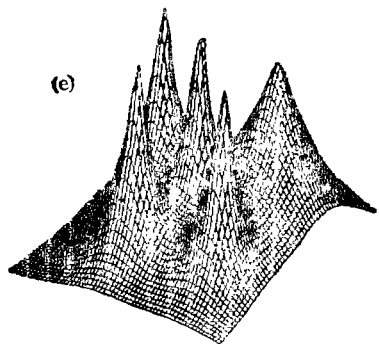
(b) 高斯径向基函数（rbfg）对应判决曲线



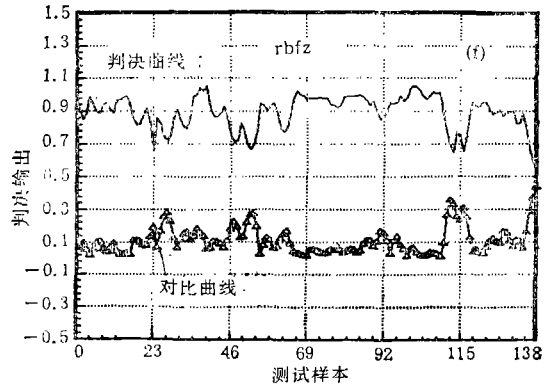
(c) 三角核拓扑分布



(d) 三角径向基函数（rbfo）对应判决曲线



(e) 双指数核拓扑分布



(f) 双指数径向基函数（rbfz）对应判决曲线

图 6　3 种核函数对 B-52 飞机一维横向像的判别置信度曲线

位关系随视角变化是较小的，即由于视角变化所呈现的模式稳定性比一维纵向像好得多，因而用它作为目标识别能获得较好的结果。

4.2　径向基函数形状参数 α 和遗忘因子 λ 对识别结果的影响

对于 RBFN 中的高斯非线性传输函数，其中的形状参数 α，R. P.W.Duin 等人[17]给出了其取值的经验公式

表 1　3 种径向基函数网络对 5 种飞机目标随噪声样本测试结果(%)

基函数 / 识别域 / 识别率 / 信噪比	高斯核					三角核					双指数核				
	B-52	歼-6	歼-7	强-5	F-1	B-52	歼-6	歼-7	强-5	F-1	B-52	歼-6	歼-7	强-5	F-1
无噪	98.38	95.33	99.11	100	100	98.3	97.33	99.11	99.38	100	100	96	99.11	99.38	100
30dB	100	95.33	99.11	100	100	100	95.33	99.11	99.38	100	100	95.33	99.38	99.38	100
20dB	100	95.33	99.11	100	88.88	100	94	99.11	99.38	100	100	94.66	99.11	99.38	100
15dB	99.27	91.33	98.23	100	72.22	100	89.33	99.11	99.38	100	100	89.33	99.11	99.38	100

$$\alpha_i = \left(\frac{1}{N}\,\mathrm{tr}\hat{\Sigma}\right)^{1/2} N_i^{-\beta/d}, \tag{17}$$

$$\alpha_i = \frac{\alpha^2}{N}(2N_i)^{-\beta/d}, \tag{18}$$

其中 $0<\beta<0.5$；N_i 为第 i 个模式对应的训练样本数；$\hat{\Sigma}$ 为样本协方差矩阵的估计；tr 为矩阵取迹运算；N 为输入训练样本的维数。

上面给出的经验公式只能作为实验中选择的参考，实际上，在进行 RBFN 训练时，它应该根据学习样本的时变特性自适应地变化。所以，我们讨论识别率对 α 的敏感性。图 7 给出了 5 种飞机目标在 20dB 信噪比测试样本下的识别率与 α 的对数关系曲线。

通过实验，我们发现[11]：影响识别性能好坏的关键因素是 RBFN 的形状参数。只要该参数在某一范围取值，通常可以保证得到很好的识别效果。此外，我们也同时观察了遗忘因子不同所引起的识别率的变化，发现在有效的形状参数前提下，λ 取值的大小对识别率的影响不大。

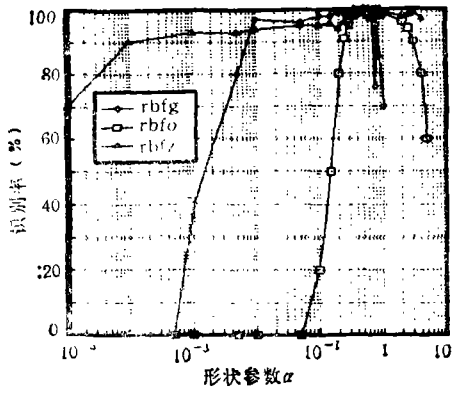


图 7 识别率与 α 的关系曲线

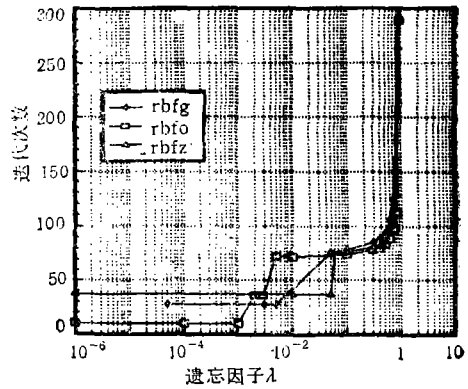


图 8 遗忘因子 λ 与收敛速度关系曲线

4.3 遗忘因子 λ 与收敛速度的关系

在实验中，我们还发现[11]遗忘因子 λ 是收敛速度的隐函数，λ 取值越小，收敛速度越快。但当 λ 取到一定下限时，收敛速度不再上升，而是呈现某一数值。图 8 给出了 3 种核函数形式对应的 λ 与收敛速度的对数关系曲线。大约在 λ 降到 0.05 后，双指数的迭代收敛次数不再下降，它的饱和值为 36；大约在 λ 降到 0.01 后，高斯核迭代收敛次数饱和值为 28；大约在 λ 降到 0.001 后，三角核迭代收敛次数饱和值为 9。

由此可以看出，这里的遗忘因子 λ 与自适应滤波中的 λ 有着本质的区别，这里 λ 取较小的值，不但不会影响性能，相反会加快收敛速度。

5 结 束 语

本文讨论了 RBFN 使用 RLS 训练方法来识别雷达目标一维像问题。主要就 5 种

飞机目标在 0°—100° 范围内一维横向像数据进行了若干实验.结果表明,RBFN 能够成功地用于目标一维像的识别. 文中证明了 RBFN 隐层单元传输函数可以选择一切满足 Parzen 窗函数的核函数形式,并着重讨论了高斯、三角、双指数 RBFN 的形状参数 α 对识别结果的影响,此外,还讨论了 RLS 算法的遗忘因子对识别结果的影响和与网络训练时间的关系. 这些工作对于我们进一步研究选用 RBFN 用于模式识别具有重要意义.

致谢　有关横向一维成像的工作是张庆文、叶炜同学完成的,特此致谢!

参　考　文　献

[1] Rumelhaart D B, Hintonn G B, Williams R J. Nature, 1986, 32(3): 533—536.
[2] Vogl T P. Biological Cybernics, 1988, 59(4): 257—263.
[3] Dahl B D. Accelerated Learning the Generalized Delta Rule. IJCNN, Vol.II, Washington DC: 1989, 523—530.
[4] Hush D R. Salas J M. Improving the Learning Rate of Back-Propagation Wigradient Reuse Algorithm, IJCNN, Vol.I, San Diego: 1988, 441—447.
[5] Parker D B. Optimal algorithms for adaptive networks, Second order back propagation, second order direct propagation, and second order Hebbian learning, IJCNN, Vol.II, San Diego: Nov. 1987, 593—600.
[6] Sadjadi M R A, Liou R J. IEEE Trans. on SP, 1992, SP-40(2): 446—450.
[7] Chan L W, Fallside F. Computer Speech and Language, 1987, 21(2): 234—240.
[8] Kollias S, Anastassiou D. Adaptive Training of Multilayer Neural Networks Using a Least Squares Estimation Technique. IJCNN, Vol.I, Washington DC: 1989, 383—390.
[9] Kollias S, Anastassiou D. IEEE Trans. on CAS, 1989, CAS-36(8): 1092—1101.
[10] Scalero R, Tepedelenlioglu N. IEEE Trans. on SP, 1992, SP-40(1): 202—210.
[11] 黄德双. 雷达目标一维象识别技术的研究: [博士论文]. 西安电子科技大学,1992.12.
[12] 西安电子科技大学电子工程研究所目标识别研究组. 雷达目标特性测式,识别与分类研究报告,1992.12.
[13] Renals S, Rohwer R. Phoneme Classification Experiments Using Radial Basis Functions, IJCNN, Vol.I, Washington DC: 1989, 461—467.
[14] Vrckovnik G, Carter C R, Haykin S. Radial Basis Function Classification of Impulse Radar Waveforms, IJCNN, Vol.I, San Diego: 1990, 45—50.
[15] Ruck D W, Rogers S K, Kabrisky, M, *et al*. IEEE Trans. on NN, 1990, NN-1(4): 296—298.
[16] Parzen E. Ann. Math. Statist., 1962, 33(6): 1065—1076.
[17] Duin R P W. IEEE Trans. on Compt., 1976, C-25(8): 1175—1179.

THE STUDY ON RECOGNITION TECHNIQUE OF RADAR TARGET'S ONE-DIMENSIONAL IMAGE BASED ON RADIAL BASIS FUNCTION NETWORK

Huang Deshuang Bao Zheng
(*Xidian University, Xi'an* 710071)

Abstract This paper studies the problem applying Radial Basis Function Network (RBFN) which is trained by the traditional Recursive Least Square Algorithm (RLSA) to the recognition of one dimensional image of radar targets. The equivalence between RBFN and the estimation of Parzen window probabilistic density is proved, it is pointed out that the I/O functions in RBFN hidden units can be extended to general Parzen window probabilistic kernel function or potential function, too. This paper discusses the effects of the shape parameter α in RBFN and the forgotten factor λ in RLSA on the results of the recognition of three kinds of kernel function such as Gaussian, Triangle, Double-exponential kernel functions, at the same time, and discusses also the relationship between λ and the training time in RBFN.

Key words Recognition, Kernel function, Shape parameter, Forgotten factor, One dimensional image, Radial basis function, Recursive least square