

一种快速全局优化的神经网络及其在数据融合中的应用¹

杨新星 焦李成

(西安电子科技大学雷达信号处理重点实验室 西安 710071)

摘要 本文将遗传算法的全局性和 EM 算法的快速性相结合, 提出了一种快速全局优化神经网络, 并将其应用于数据融合中. 理论与实验结果表明该算法在数据融合中具有很强的鲁棒性.

关键词 数据融合, 相关, 优化

中图分类号 TN-052, TP391

1 引言

由于神经网络在信息加工处理过程中具有很强的容错能力和鲁棒性, 所以自从神经网络发展成一种理论体系以来, 越来越受人们的关注. 它的训练算法一直是许多学者研究的热点. 本文提出了一种基于遗传算法全局性和 EM (Expectation Maximization) 算法^[1,2]快速性的快速的全局优化算法, 并且把具有这种快速的全局优化算法的神经网络用来解决多传感器数据融合中由航迹交叉产生的模糊相关问题. 仿真结果表明了本文方法的可行性和有效性.

2 神经网络的训练算法

2.1 快速的训练算法 我们选择单隐层神经网络, 并把 EM 算法用于神经网络的训练. 把神经网络的隐层 $\{z(n)\}_{n=1}^N$ 作为未知数据, $n = 1, 2, \dots, N$, 表示样本数, z 表示隐节点矢量. 它与输入 $\{x(n)\}_{n=1}^N$ 和输出 $t(n)$ 形成了完全数据. 此完全数据的对数概率的条件期望可表示如下:

$$\begin{aligned} Q(\theta/\theta^p) &= E\{\ln P(\{t\}, \{z\}, \{x\}/\theta)/\{t\}, \{x\}, \theta^p\} \\ &= \ln P(\{x\}/\theta) + \int P(\{z\}/\{t\}, \{x\}\theta^p) \ln P(\{z\}, \{t\}/\{x\}, \theta) d\{z\}. \end{aligned} \quad (1)$$

这里为了方便, 用 $\{\cdot\}$ 表示 $\{\cdot\}_{n=1}^N$. θ 用来表示网络权值的参数. 为了计算这个对数的条件期望, 需要建立几个概率模型. 对于输入 $\{x(n)\}$, 选取 $\{z\}$ 是均值为 $\{h(n)\}$ 的高斯分布, 同时输出 $t(n)$ 是在给定隐目标下具有均值为 $(w^{(2)})^T z(n)$ 高斯分布, 即

$$P(\{z\}/\{x\}, \theta) = B_1 \exp(-\lambda_1 E_1), \quad (2)$$

$$P(\{t\}/\{z\}, \theta) = B_2 \exp(-\lambda_2 E_2), \quad (3)$$

这里

$$E_1 = \sum_{n=1}^N \|z(n) - h(n)\|^2, \quad (4)$$

$$E_2 = \sum_{n=1}^N (t(n) - z(n)^T w^{(2)})^2, \quad (5)$$

¹ 1998-03-26 收到, 1998-12-02 定稿
国家“863”和国家教委夸世纪人才基金资助项目

B_1 和 B_2 是归一化参数, λ_1 、 λ_2 分别是 E_1 、 E_2 的权系数。 E_1 表示隐层的期望值和实际值的误差, 如图 1 所示。 E_2 表示当期望的隐节点值被输入时, 网络的期望输出和实际输出的误差, 如图 2 所示。 假设输出的概率只依赖于隐节点和参数, 即

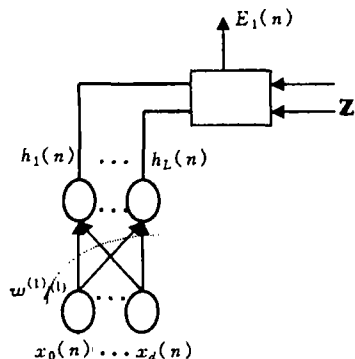


图 1 $E_1(n)$ 的计算

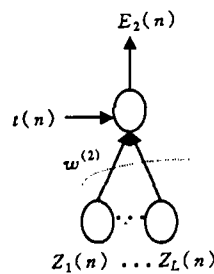


图 2 $E_2(n)$ 的计算

$$P(\{t\}/\{z\}, \{\mathbf{x}\}, \theta) = P(\{t\}/\{z\}, \theta). \quad (6)$$

根据 (2)、(3)、(6) 式可以得到联合分布

$$\begin{aligned} P(\{t\}, \{z\}/\{\mathbf{x}\}, \theta) &= P(\{t\}/\{z\}, \theta)P(\{z\}/\{\mathbf{x}\}, \theta) \\ &= A_{tz}^N \exp(-\lambda_1 E_1 - \lambda_2 E_2), \end{aligned} \quad (7)$$

这里

$$A_{tz} = \sqrt{(\lambda_2/\pi)(\lambda_1/\pi)^L}. \quad (8)$$

另外, 根据通常的单隐层神经网络模型可得

$$P(\{t\}/\{\mathbf{x}\}, \theta) = \left(\sqrt{\frac{\lambda_3}{\pi}} \right)^N \exp(-\lambda_3 E_3). \quad (9)$$

E_3 表示整个网络的训练误差, 它是一个平方能量函数, 即

$$E_3 = \sum_{n=1}^N (t(n) - y'(n))^2. \quad (10)$$

根据边缘概率与联合概率关系, 有

$$\lambda_3 = \lambda_1 \lambda_2 / (\lambda_1 + \lambda_2 \|\mathbf{w}^{(2)}\|^2). \quad (11)$$

根据 (1) 式可算出 EM 算法中第一步: 求期望表达式, 然后以期望表达式中的参数为变量进行最大化, 这是 EM 算法中的最大化步。 这样可以得到

$$\mathbf{w}_j^{(1)(p+1)} = \arg \min_{\mathbf{w}_j^{(1)}} \sum_n (\hat{z}_j(n) - g(\mathbf{w}_j^{(1)T} \mathbf{x}(n)))^2, \quad (12)$$

$$\mathbf{w}^{(2)} = \arg \min_{\mathbf{w}^{(2)}} E_3, \quad (13)$$

其中

$$\hat{z}_j(n) = h_j^{(p)}(n) + \frac{\lambda_2 \mathbf{w}_j^{(2)(p)}}{\lambda_1 + \lambda_2 \|\mathbf{w}^{(2)(p)}\|^2} e(n), \quad 1 \leq j \leq L,$$

$$e(n) = t(n) - y^{(p)}(n). \quad (14)$$

上式中的上标 p 表示第 p 次迭代。(13) 式的计算是利用了 (12) 式新产生的 $\mathbf{w}_j^{(1)}$, 即 (12) 式的计算一定在 (13) 式的前边。(12)、(13)、(14) 式是整个迭代过程中所涉及到的具体计算公式, 一直迭代到收敛, 即达到稳定值, 整个迭代过程就完成。这种算法把训练原来的两层神经网络分解成训练一组单一的神经原, 因此这种训练方法可大大提高训练速度。

2.2 快速的全局训练算法 在 2.1 节中所提出的训练算法虽然在速度上提高了很多, 但是 EM 算法在本质上类似于爬山最优化算法, 所以是一种局部寻优算法。如果初始解选择不好, 就有可能陷入局部解。所以需要初始解进行全局优化, 选出优秀的初始解。这就是本文提出的遗传算法^[3]所要完成的功能。当初始解被优化在一定的范围之内时, 转入 2.1 节中提出的快速的训练算法。这样既避免了单一遗传优化算法由于全局搜索而速度慢的缺点, 又避免了单一 2.1 节中提到的快速训练算法解的局部性的缺点。

在本文的遗传算法中, 人口选择二进制编码, 适应度函数选用网络的能量函数, 即 $E_3 = \sum_{n=1}^N (t(n) - y(n))^2$; 选择算子选用适应度比例的方法 (见文献 [3]); 交叉算子选择两点交叉; 变异算子选择位点变异。

3 全局快速的神经网络在数据融合中的应用

在本文中用到的传感器有: (1) 敌我识别器 (IFF), 此系统只用于敌我目标的分离; (2) 目标的 RCS 值 (u_1); (3) 目标的 IR 辐射特性 J (用于过滤) 或目标发动机类型 (用于相关), IR 的取值 u_2 表示; (4) ESM 信息 (u_3); (5) 目标航速 (u_4); (6) 点迹对于航迹目标预测点的 Mahalanours 距离 (u_5)^[4], 且

$$M_{ij}(k) = d_{ij}^T(k) S(k)^{-1} d_{ij}(k), \quad (15)$$

其中 $d_{ij}(k)$ 是 k 时刻第 i 个点迹相对于第 j 条航迹预测点之间的残差, $S(k)$ 是其协方差矩阵。

下面我们把目标特性的论域限定为

$$U = \{u_1, u_2, u_3, u_4, u_5\} \quad (\text{相关时}); \quad (16)$$

$$U' = \{u'_1, u'_2, u'_3\} \quad (\text{过滤时}). \quad (17)$$

同时, 可以建立一目标特征数据库, 其中包括各种类型飞机, 如: 歼击机 (v_1), 轰炸机 (v_2), 运输机 (v_3), 民航机 (v_4), 直升机 (v_5), \dots , v_l . 因此, 目标类型的论域可限定为

$$V = \{v_1, v_2, v_3, v_4, v_5, \dots, v_l\}. \quad (18)$$

相关就是通过将点迹的各个特征参数与航迹目标的相应的特征参数逐一进行比较, 根据它们之间的相似程度, 综合作出判决。在本文中特征参数之间的相似程度是用神经网络来完成。首先用上述的方法对神经网络进行训练, 把飞机的各个特征参数作为网络的输入。然后在航迹跟踪和相关时, 把接收到的航迹目标相应的特征参数输入训练好的网络就可以得出相应的飞机类型。所以在跟踪多个目标时, 利用神经网络进行航迹相关, 就不会因为多个航迹交叉而发生丢失目标的现象。

4 仿真结果和结论

如图 3 所示, 在仿真中采用了两种不同类型的飞机: 歼击机、轰炸机。虚线代表一类飞机, 实线代表另一类飞机, 彼此相互作用接近和机动飞行, 并且这两条航迹在一段时间内互相交叉。背景噪声仍设定为在空中作均匀分布, 密度为 1:5。其中的 RCS 数值和沿方位角的变化特性取自文献 [5] 且据有关资料设定的, 并取轰炸机 RCS 的均值为歼击机的 5 倍。目标的红外辐射特性是参考文献 [6] 设定的。目标航速设定为轰炸机 (300 ~ 800)km/h, 歼击机 (800 ~ 2400)km/h。

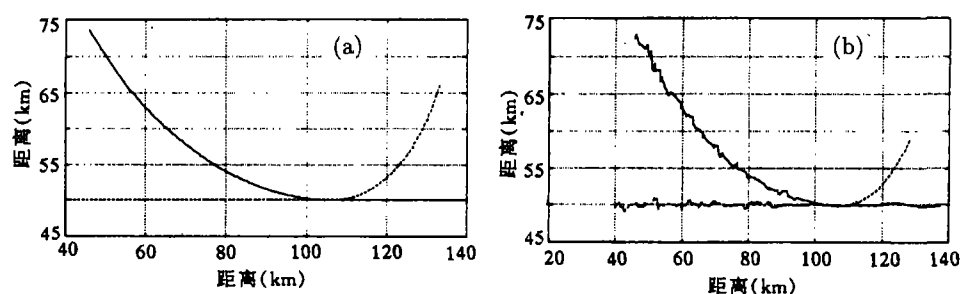


图 3 (a) 原航迹 (b) 航迹跟踪图

用上面对所提出的神经网络算法进行了 50 次 Monto Carlo 仿真, 得出了此算法在航迹跟踪中的误相关率和成功相关率, 并列举在表 1 中:

误相关率 (%)			成功相关率 (%)		
目标 1	目标 2	平均	目标 1	目标 2	平均
1.02	4.86	2.94	98.98	95.14	97.06

仿真结果表明, 本文算法由于引入了属性参数, 大大地提高了成功的相关率, 对于航迹交叉处的模糊相关问题提出了一种行之有效的方法。

参 考 文 献

- [1] Moon T K. The expectation maximization algorithm. IEEE Signal Processing Magazine, 1996, 13(6): 47-60.
- [2] Jamshidian M, Jennrich R. Acceleration of the EM algorithm by using quasi-Newton methods. J. R. Statist. Soc. B, 1997, 59(3): 569-587.
- [3] 陈国良, 等. 遗传算法及其应用. 北京: 人民邮电出版社, 1996, 69-77.
- [4] Long M W. Airborne Early Warning System Concepts. Boston, London: Artech House, 1992, 293-321.
- [5] Sarma V V S. Multisensor data fusion and decision support for airborne target identification. IEEE Trans.on SMC., 1991, SMC-21(5): 969-978.
- [6] 王功伯. 喷气式飞机的红外辐射特性. 红外与激光技术, 1989, 2(1): 26-29.

- [7] 杨新星. 模糊神经网络数据融合算法的研究: [硕士论文]. 西安: 西安电子科技大学, 1997.

A FAST GLOBAL OPTIMIZATION NEURAL NETWORK AND ITS APPLICATION TO DATA FUSION

Yang Xinxing Jiao Licheng

(Key Lab. for Radar Signal Processing, Xidian University, Xi'an 710071)

Abstract This paper presents a fast global optimization neural network and applies it to the data fusion. This neural network is based on the global property of genetic algorithm and the high speed property of expectation maximization (EM) algorithm. The simulation results show that this neural network is robust in the data fusion.

Key words Data fusion, Correlation, Optimization

杨新星: 男, 1971 年生, 博士生, 研究兴趣为智能信号处理, 子波理论及其应用、图像处理.

焦李成: 男, 1959 年生, 教授, 博士生导师, IEEE 高级会员, 主要研究方向为智能信号处理和非线性科学.