文章编号：1000-8152(2003)01-0066-04

最佳拟合与神经网络相结合实现传感器特性线性化

朱庆保
(南京师范大学 计算机系，江苏 南京 210097)

摘要：提出了一种传感器特性线性化的方法.该方法把传感器特性分为线性和非线性段，用一种改进的 BP 神经网络映射传感器特性非线性段的反函数作为校正环节，用最佳拟合方法得到线性段的直线方程，从而实现传感器特性的线性化. 经过仿真试用表明，这种方法可使传感器的非线性误差减小近十倍. 最后，给出了一些仿真实验和仿真结果.

关键词：神经网络；最佳拟合；传感器；线性化；仿真
中图分类号：TP212.6　　**文献标识码**：A

Sensor's characteristics linearization using artificial neural networks and best-fitting

ZHU Qing-bao
(Department of Computer Technology, Nanjing Normal University, Jiangsu Nanjing 210097, China)

Abstract: A method of the sensor's characteristics linearization is proposed. The sensor characteristics are compartmentalized into linear and nonlinear segments, the inverse functions of nonlinear segment are mapped with the improved BP neural networks as correction links, and a straight line equation is produced with best-fitting in the linear interval. The results of simulations show that the nonlinear errors of the sensor are decreased approximately tenfold. Finally, an experiment is given.

Key words: neural networks; best-fitting; sensor; linearization; simulation

1 引言(Introduction)

传感器的误差大小直接影响到测控系统的性能，而非线性误差是传感器的主要误差，因此，传感器特性线性化的研究成为国内外研究的一项重要课题.

为了消除或补偿传感系统的非线性误差或者为了使用非线性传感特性，国外不少学者提出了非线性传感特性线性化的神经网络方法[1~3]，这些方法大多数是用一个 BP 神经网络去映射传感特性的反函数作为校正环节，从而实现线性化.但是，许多传感器都有一段线性段，往往在两端或一端呈非线性，具有明显的分段特征，其反函数是很难用一个式子来表示的.而神经网络是通过权的重叠并通过在训练过程中对这些重叠的权寻求折衷，从而实现对输入数据的泛化，达到映射非线性函数的目的.由于传感器特性的分段性，这种寻求折衷的过程即网络训练过程是相当困难的，即使训练成功，其精度也是很低的.换言之，用一个神经网络同时映射几个函数是十分困难的，即使数学意义上存在这种网络，由于不同结构的网络有无穷多个，要找到这样一种结构的网络犹如大海捞针.很显然，国外目前研究的一些非线性传感器特性线性化的神经网络方法还有很大的局限性.为此，作者研究了一种最佳逼近拟合与神经网络相结合实现传感器特性分区线性化的方法，取得了较好的效果.

2 传感器特性线性化的基本原理(Fundamental principle of sensor's characteristics linearization)

根据许多传感器特性曲线中间呈线性，两端呈非线性的特点，本文把传感器的特性曲线分为 3 段，即：不考虑传感器的重复误差和迟滞误差，可认为传感器的特性函数是在输入量程 $x \in [\zeta_0, \zeta_h]$ 上呈单调变化的连续函数.设已获得 n 组观测值 (x_i, y_i)，$i = 1,2,3,\cdots,n$，其中 $x_1 < x_2 < x_3 < \cdots < x_n$，从这 n 组数据及其分布，把其分为 3 个区间，即令

收稿日期：2001-02-05；收修改稿日期：2002-04-22.

$$y = f(x) = \begin{cases} f_1(x), \text{当 } \zeta_0 \leqslant x < \zeta_1, \\ f_2(x), \text{当 } \zeta_1 \leqslant x < \zeta_2, \\ f_3(x), \text{当 } \zeta_2 \leqslant x < \zeta_h. \end{cases}$$

其中，$f_2(x)$ 为一维的直线方程，可用切比雪夫最佳逼近方法求出，其区间大小则根据精度要求自适应确定. 当该区间小于某一值时，则判该直线不存在，这时，仅分为 $f_1(x)$ 和 $f_3(x)$ 两个函数. $f_1(x)$ 和 $f_3(x)$ 是非线性函数，为了消除或补偿它的非线性误差，可使其输出分别通过一个补偿环节，以下以 $f_1(x)$ 的补偿环节为例，$f_3(x)$ 原理相同. 设补偿环节的特性函数为 $p = F_1(y)$，p 为补偿后的输出. 若令 $F_1(y) = f_1^{-1}(y)$，则有

$$p = f_1^{-1}(y) = x,\ x \in [\zeta_0, \zeta_1]. \tag{1}$$

很显然，$F_1(y)$ 也是一个非线性函数. 从式(1) 可知，补偿后的输出 p 与被测量值 x 一致，使补偿后的传感器特性具有理想特性.

实际应用中，非线性函数 $F_i(y)(i = 1,3)$ 的表达式难以准确求出，为此，用两个多层的 BP 神经网络分别映射它们，这样不必求出 $f_i(x)$ 及 $F_i(y)(i = 1,3)$ 的参数，经神经网络学习训练后，由神经网络直接输出线性化后的结果.

3 最佳直线拟合与最佳线性区间确定 (Straight line best-fitting and definition of the best straight line interval)

为了得到最大的线性区间，令 $\zeta_1 = \zeta_0, \zeta_2 = \zeta_h$. 设样本点 (x_i, y_i) 有 n 个，且有 $\zeta_1 < x_1 < x_2 < \cdots < x_n < \zeta_2$. 将这 n 个离散点按 x_i 的大小顺序依次用直线连接起来，构成在区间 $[\zeta_1, \zeta_2]$ 上的单值连续函数 $y = f(x)$. 令 s 为平面 xoy 上直线的全体，$P(x) = kx + C$ 是 s 中的任意一条直线；设拟合直线 $P^*(x) = k^* x + C^*$，构造误差函数 $e(x) = f(x) - P^*(x)$，$x \in [\zeta_1, \zeta_2]$. 根据切比雪夫最佳逼近原理，最佳逼近直线的充要条件是在 $x_i(i = 1,2,\cdots,n)$ 中至少有一个 $x_{k0}, x_{k1}, x_{k2}(1 \leqslant k_0 < k_1 < k_2 \leqslant n)$ 的交错点组. 一旦获得交错点组 x_{k0}, x_{k1}, x_{k2}，就可以列出联立方程

$$\begin{cases} f(x_{k0}) - p^*(x_{k0}) = e^*(x_{k0}), \\ f(x_{k1}) - p^*(x_{k1}) = e^*(x_{k1}), \\ f(x_{k2}) - p^*(x_{k2}) = e^*(x_{k2}). \end{cases} \tag{2}$$

且有 $e^*(x_{k0}) = -e^*(x_{k1}) = e^*(x_{k2})$，记 $e^*(x_{k0})$ 为 E^*，因 $x_{ki}(i = 1,2,3)$ 取自样本点，故对应样本点输出为 $y_{k0} = f(x_{k0}), y_{k1} = f(x_{k1}), y_{k2} = f(x_{k2})$. 将 $P^*(x) = k^* x + C^*$ 代入式(2)

$$\begin{cases} y_{k0} - (k^* x_{k0} + C^*) = E^*, \\ y_{k1} - (k^* x_{k1} + C^*) = -E^*, \\ y_{k2} - (k^* x_{k2} + C^*) = E^*. \end{cases} \tag{3}$$

为了在最大的区间内得到一条满足要求的最佳直线，求取交错点组和自适应确定区间大小的算法步骤如下：

Step 1　设定系统允许误差 ε. 为得到最大的区间，故初始化 $a = \zeta_0, b = \zeta_h$，样本数 $= n$.

Step 2　在样本点集 $x_i, i = 1,2,\cdots,n, x_i \in [a,b]$ 中选取按大小顺序排列的初始点组 x_{k0}, x_{k1}, x_{k2}，且令 $k_0 = 1, k_1 = (n+1)/2, k_2 = n$.

Step 3　将 x_{k0}, x_{k1}, x_{k2} 及对应的 y_{k0}, y_{k1}, y_{k2} 代入式(3)，确定出 k^*, E^* 和 C^*.

Step 4　对任意的样本点 (x_i, y_i)，计算

$$E_{\max} = \max_{1 \leqslant i \leqslant n} | y_i - (k^* x_i + C^*) |.$$

结果有如下 3 种情况，相应的处理步骤为

Case 1　$E_{\max} > E^*$，必有样本点 (x_k, y_k) 使得

$$E_{\max} = | y_k - (k^* x_k + C^*) | = \max_{1 \leqslant i \leqslant n} | y_i - (k^* x_i + C^*) |$$

成立，则用 x_k 取代 x_{k0}, x_{k1}, x_{k2} 中的一个，得到新的交错点组 x_{k0}, x_{k1}, x_{k2}，使新的点组仍依次分布于直线 $P^*(x) = k^* x + C^*$ 两侧. 重复 Step 3 和 Step 4.

Case 2　$\varepsilon < E_{\max} < E^*$，当前区间内期望的直线不存在，则变更区间，令 $a = a + ih, b = b - jh, q = (b-a)/h$，$h$ 为样本间隔. 为了得到最大直线区间，区间变更时，依次令 $i = 1, j = 0,1,2,\cdots,q-2, i = 2, j = 0,1,2,\cdots,q-2,\cdots, i = q-3, j = 0,1,2,\cdots,q-2$，分别重复重复 Step 2 ~ Step 4，若 $b - a$ 小于某一值，则期望的直线不存在，这时，记忆区间中点，以区间中点为界，分两段用神经网络映射.

Case 3　$E_{\max} \leqslant E^* \leqslant \varepsilon$，记忆上述直线方程及区间 a，b，学习过程结束.

4 BP 神经网络映射方法及其算法 (The mapping methods of BP neural networks and its algorithm)

4.1 BP 神经网络传统算法及对本应用的缺陷 (The defect of conventional algorithm of BP neural networks in this application)

为了分别用一个 BP 神经网络映射 $f_1(x)$ 和 $f_3(x)$，本文设计出的 BP 网络如图 1 所示，因为是用于单输入单输出的系统，令图 1 中的所有输入都

等于 x，等价于输入端再加一个输入节点，令该节点与 I 层的连接权 $w_{0i} = 1(i = 1,2,3)$ 时的情况.在传统的基本算法中，大多都使用了 Sigmoid 函数作为激励函数，把网络的输入输出数据都压缩到 0 至 1 之间，这种方法使网络的收敛性得到了很大的改善，对于模式识别和逻辑运算及所学习的函数的数值较小时较为适合，并取得了十分成功的应用.但在本应用中，由于所学函数的数值较大，因而采用传统算法时有两个缺陷：其一，学习速度慢；其二，学习精度低.为此，针对本应用，本文采用了一种不用 Sigmoid 函数的改进算法，取得了较好的效果.

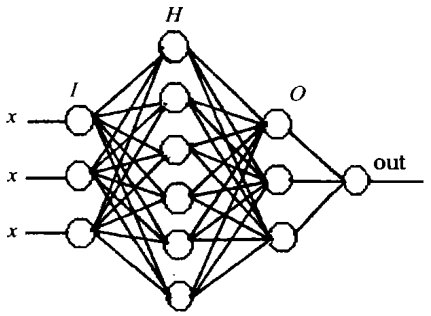

图 1 BP 神经网络模型

Fig. 1 Model BP neural networks

4.2 针对本应用采用的改进 BP 算法(An improvement BP algorithm in this application)

有关的研究表明：神经元的作用函数不必非取 Sigmoid 函数，而只需满足有界单调递增或有界连续即可逼近非线性连续函数[4].根据这一研究结果和本应用的特点，并考虑到传感器特性一般是单调递增的连续函数，因此，神经元不使用 Sigmoid 特性，而是采用了下面的 $g(x)$ 特性函数，该函数为

$$g(x) = \begin{cases} \alpha x, & \text{当 } x \in [\zeta_l, \zeta_h], \\ 0, & \text{当 } x \notin [\zeta_l, \zeta_h]. \end{cases}$$

式中，α 为一很小的常系数(例如 0.00001).$[\zeta_l, \zeta_h]$ 为量程上下限. 仿真实验证明，对于本应用，这种方法收敛快，映射精度高，达到了实用要求.

根据图 1，每个 BP 网络的基本训练练算法步骤如下(适于计算机实现的基本算法)：

Step 1 初始化.把各层的连接权 w_{ij} 初化为一个很小的随机值.

Step 2 前向计算

$$\text{net}_{ih} = \sum_i w_{ih} x_i,\ \text{out}_{ho} = g(\text{net}_{ih}) = \alpha \text{net}_{hj},$$

$$\text{net}_{ho} = \sum_h w_{ho} \text{out}_{ho},\ \text{out}_o = g(\text{net}_{ho}) = \alpha \text{net}_{ho},$$

$$\text{net}_{\text{out}} = \sum_o w_o \text{out}_o,\ \text{out} = g(\text{net}_{\text{out}}) = \alpha \text{net}_{\text{out}}.$$

其中，out_{ho}，out_o 和 out 分别为 H 层、O 层和输出层的输出，w_{ih}，w_{ho}，w_o 分别为 I 到 H 层、H 到 O 层和 O 到输出的连接权，$i = 1,2,3$，$h = 1,2,\cdots,6$，$o = 1,2,3$，x_i 为输入，且 $x_1 = x_2 = x_3$.

Step 3 反向误差计算

$$\delta = (y - \text{out})g'(\text{net}_{\text{out}}) = (y - \text{out})\alpha,$$

$$\delta_o = \delta w_o g'(\text{net}_{ho}) = \delta w_o \alpha,$$

$$\delta_h = \sum_{o=1}^{3} \delta_o w_{ho} g'(\text{net}_{ih}) = \alpha \sum_{o=1}^{3} \delta_o w_{ho}.$$

式中 $o = 1,2,3$，$h = 1,2,\cdots,6$，$i = 1,2,3$. δ 为第 k 次训练后的输出误差，在此忽略了下标 k. δ_o，δ_h 是根据 δ 反传到 O 层、H 层的误差.

Step 4 修正各层权：若有 $E_{\max} = \max\{\delta_k\} < \epsilon$，$k = 1,2,\cdots,n$，$n$ 为样本数，ϵ 为设定的学习控制误差，则训练完成，否则按(4)式调整权，并重复 Step 2 ~ Step 4.

$$\begin{cases} \Delta w_o = \delta \text{out}_o, & w_{3(n+1)} = w_{o(n)} + \eta \Delta w_o, \\ \Delta w_{ho} = \delta_o \text{out}_h, & w_{ho(n+1)} = w_{ho(n)} + \eta \Delta w_{ho}, \\ \Delta w_{ih} = \delta_h x_i, & w_{1ij(n+1)} = w_{1ij(n)} + \eta \Delta w_{1ij}. \end{cases} \tag{4}$$

式中，η 是学习步长.

为了得到最优结果，可用自适应方法确定 ϵ.以上算法中 x_i 的区间为 $[\zeta_l, \zeta_h]$. 另外需要说明的是，该改进算法是针对本应用中学习函数呈单调、低维、小区间等特点研究的，不一定适于其它情况.

5 仿真与应用结果(Simulation and the results of application)

作者用本文所提出的算法校正传感系统的非线性误差，取得了较满意的效果.下面给出一个仿真应用实例，以说明应用方法和效果.

智能气动测微仪中的传感器特性曲线是较典型的非线性特性，它中间一小段呈现较好的线性，两端为非线性.设定拟合误差 $\epsilon < 0.5$ 时，用最佳拟合法拟合出的直线段区间约为 60 μm.在两端，其最大非线性误差高达 6.7 μm.为说明问题方便并不失一般性，本文仅说明一端的线性化方法，其区间设为 $[c, d]$，且令 $c = 100$，$d = 200$. 另一端方法完全相同.

由于传感器非线性段的反函数较难准确求出，因此，本文用实验法给出教师信号.设给定训练样本为 (X^*, Y^*)，其中 Y^* 为神经网络期望的输出，也就是对应的传感系统的输入位移量 S，X^* 为给定的 BP 网络的输入，亦即 S 经传感系统后的输出再经

A/D转换后的值，由于传感系统的非线性作用，$X \neq kS$(k 为常系数). 为了建立一组训练样本，设样本间隔为 h，并使微机处于连续采样状态，输入端用一个标准的微动台架微动进给，以微变传感器的输入，当微机的采样值为 $x_i^* = i * h$ 时，记录并键入输入值，即为 y_i^*. 其中 $i \in [0, M-1]$，$M = (d-c)/h + 1$. 规定的样本键入完毕后，网络即进入训练过程，这一过程即为学习. 为了在脱离硬件的情况下，即能检查网络的工作情况，本文对该传感器的学习和应用过程进行了仿真. 在区间 $[c, d]$ 范围，根据一组实测值，用拟合方法得出的反函数的曲线方程近似为

$$s = F(x) = ax + bx^2. \tag{5}$$

其中，$a = 0.000167$，$b = 1$；s 为输入到传感器的实际位移量，x 是模拟传感系统的输出再经 A/D 转换后的值. 令 $X^* = \{x_0, x_1, \cdots, x_i, \cdots, x_M\}$，$x_{i+1} = x_i + h$，由式(5)得到 $Y^* = \{s_0, s_1, \cdots, s_i, \cdots, s_M\}$，作为训练样本，经训练后，输入任意的 s，$s \in [c, d]$，用式(5)的反函数求出 x(模拟 A/D 实际采样值)作为网络的输入，网络即可输出正确的解 s.

对本例的仿真结果如表 1 示(表 1 中误差均为非线性段的非线性误差绝对值).

表 1　误差分析与比较

Table 1　Errors analysis and comparison

样本间隔	无网络时最大误差	无网络时最小误差	无网络时平均误差	加网络后最大误差	加网络后最小误差	加网络后平均误差
20	6.79	1.75	3.85	0.86	0.0	0.38
30	6.79	1.75	3.85	0.97	0.0	0.46

6　结语(Conclusion)

仿真实验和试用证明，对于有一段线性段的传感器，本文方法是一种较好的方法之一，能取得较好的效果，可使传感器的非线性误差减小十倍左右. 与传统多项式拟合比，本文用神经网络取代非线性段的高次曲线拟合，达到同样精度的情况下，本文算法简单得多，易于实现，所需样本数也少得多. 与近年国内外研究的用一个神经网络校正法比，精度高得多，训练速度也快得多，具有较好的应用前景和推广价值.

参考文献(References)：

[1] TAIB M N, NARAYANASWAMY R. Multichannel calibration technique for optical-fibre chemical sensor using artificial neural network [J]. *Sensors and Actuators B (Chemical)*, 1997, 39(1/3): 365 - 370.

[2] DEMPSEY G L, ALT N J, OLSON B A, et al. Control sensor linearization using artificial neural networks [J]. *Analog Integrated Circuits and Signal Processing*, 1997, 13(3): 321 - 322.

[3] ATTARI M, BOUDJEMA F, HENICHE M. An artificial neural network to linearize a G (tungsten vs. tungsten 26% rhenium) thermocouple characteristic in the range of zero to 2000 degrees C [A]. *Proceedings of the IEEE International Symposium on Industrial Electronics* [C]. New York: IEEE Press, 1995, 1: 176 - 180.

[4] HORNIK K, STIMCHCOMBE M, WHITE H. Universal approximation of an unknown mapping and its derivatives using multilayer feedwood networks [J]. *Neural Networks*, 1990, 3(3): 551 - 560.

作者简介：

朱庆保　(1955 —)，男，南京师范大学计算机系教授. 主要研究方向：人工智能与智能测控.　Email: qbzhu@email.njnu.edu.cn