

基于支持向量机的传感器非线性误差校正

周鸣争, 汪 军

(安徽工程科技学院计算机科学与工程系 安徽 芜湖 241000)

【摘要】提出了一种用支持向量机校正传感器非线性误差的原理和方法。该算法只依据样本就可以正确辨识传感器逆模型特征, 而不需关于逆模型函数形式的任何先验知识, 并将原问题转化为一个凸二次优化问题, 能够保证找到的极值解就是全局最优解, 具有较好的泛化能力。通过对电容式湿敏传感器误差校正的应用表明: 该算法可取得较好的效果。

关键词 支持向量机; 传感器; 非线性误差; 校正
中图分类号 TP18 **文献标识码** A

Nonlinear Errors Correction of Sensors Based on Support Vector Machine

ZHOU Ming-zheng, WANG Jun

(Dept of Comp Sci. & Eng., Anhui University of Technology and Science Wuhu Anhui 241000)

Abstract Principle and method of correcting nonlinear error of sensors based on the support vector machine is given. The algorithm can identify and know the contrary model characteristic of sensor correctly only according to the sample, having no use for any priori knowledge about contrary model function. It also convert original problem into protruding quadratic optimizing of question. The algorithm can ensure that extreme solution is optimal and have ability of common. In the end, through using in application of wet-capacitance sensor error correction, the algorithm can make better result.

Key words support vector machine; sensors; nonlinear errors; correct

近年来, 有些学者通过采用人工神经网络来解决这种非线性误差校正问题^[1], 但由于神经网络算法存在有局部极小点, 泛化非线性映射的能力不强, 有时较难达到测控系统精度的要求。

支持向量机(Support Vector Machine, SVM)是一种新的机器学习方法^[2], 它建立在统计学习理论的基础上, 根据文献[3-4]的结构风险最小化原则, 具有较好的泛化能力。对于有限样本学习问题已经表现出很多优于已有方法的性能。同时SVM算法是一个凸二次优化问题, 能够保证找到的极值解就是全局最优解, 能较好地解决小样本、非线性和高维数的问题。本文针对传感器非线性误差校正的需求, 提出了一种基于支持向量机的传感器非线性误差的校正算法。

1 校正原理及算法

众所周知, 在测控系统中, 通常采用传感器来完成各种物理量的转换, 再经过变送器, 输出能反映与所测物理量大小相对应的电信号, 通过A/D后, 可以获得其数字量。其过程如图1所示。大多为一单输入、单输出系统, 一般可表示为:

$$y = f(x) \quad (1)$$

式中 y 表示传感器系统的输出; x 表示传感器系统的输入。

在实际应用中, 由于式(1)中的 $f(x)$ 是一非线性函数, 其表达式难以精确表达, 这种非线性关系在检测过程中会造成输入和输出间总存在有一定的误差, 若该误差超过系统允许的范围, 则必须进行校正。一般常用校正的方法有两种:

收稿日期: 2003-10-02

基金项目: 安徽省自然科学基金资助项目(03042306); 安徽省科技厅国际合作资助项目(0208800)

作者简介: 周鸣争(1958-), 男, 工学学士, 教授, 主要从事智能控制、模式识别等方面的研究。

- (1) 采用其一段近似线性段;
- (2) 用硬件补偿的方法将其中的一段补偿成近似线性。

这些方法难以做到全程补偿, 应用受到一定限制。对式(1)为了消除传感器的非线性误差, 可使其输出 y 通过一个补偿环节, 若该环节的特征函数为: $z = F(y)$ 。而且使: $F(y) = f^{-1}(y)$ 。则有:

$$z = f^{-1}(y) = x \quad (2)$$

这样非线性误差就可以被有效消除。很显然 $F(*)$ 也是一个非线性函数, 这样补偿后的输出 z 与被测量值 x 一致, 使传感器具有理想特征, 其过程如图2所示。在实际中一般非线性函数 $F(y)$ 的表示难以准确求出, 但根据支持向量机理论, 若已知 k 个输入样本:

$$(y_1, z_1), (y_2, z_2), \dots, (y_k, z_k) \quad (3)$$

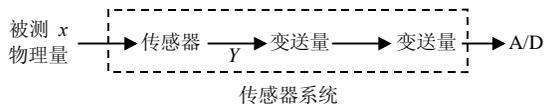


图1 传感器系统结构组成

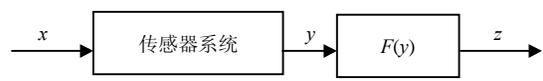


图2 非线性校正过程

机器学习的问题就是根据 k 个独立同分布观测样本, 求一个最优的函数 $f^*(y)$, 使得通过样本训练后, 对某样本以外的 y , 通过 f^* 找出对应的 z , 这是一个非线性函数回归问题。首先将 y 域用一个非线性变换 $y \rightarrow \phi(y)$, 将输入空间映射到一个高维的特征空间, 然后在特征空间中进行线性回归, 即:

$$f^*(y) = \mathbf{w} \cdot \phi(y) + b \quad (4)$$

式中 “ \cdot ” 表示向量的点积。根据结构风险最小化准则^[5], 在允许误差的情况下, 引入松弛因子 $\xi_i \geq 0$ 和 $\xi_i^* \geq 0$, 则最优化问题是转化为下式问题。其目标函数为:

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + c \sum_{i=1}^k (\xi_i + \xi_i^*) \quad (5)$$

约束条件为:

$$f^*(y_i) - z_i \leq \xi_i^* + \varepsilon \quad i=1, 2, \dots, k \quad (6)$$

$$z_i - f^*(y_i) \leq \xi_i + \varepsilon \quad i=1, 2, \dots, k \quad (7)$$

$$\xi_i, \xi_i^* \geq 0 \quad i=1, 2, \dots, k \quad (8)$$

式(5)中, 第一项使函数更为平坦, 从而提高泛化能力; 第二项则减少误差, 其中 $c(c>0)$ 表示函数 f^* 的平滑度和允许误差大于 ε 数值之间的折衷, ε 为允许误差, 从式(6)、(7)可看出, 误差函数为:

$$E = \begin{cases} 0 & \text{如果 } |f^*(y_i) + b - z_i| < \varepsilon \\ |f^*(y_i) + b - z_i| - \varepsilon & \text{其他} \end{cases} \quad (9)$$

对式(5)求解是一个凸二次优化问题, 引入lagrange函数:

$$L(\mathbf{w}, \xi_i, \xi_i^*) = \frac{1}{2} \|\mathbf{w}\|^2 - c \sum_{i=1}^k (\xi_i + \xi_i^*) - \sum_{i=1}^k \alpha_i (\varepsilon + \xi_i + z_i - f^*(y_i)) - \sum_{i=1}^k \alpha_i^* (\varepsilon + \xi_i^* + z_i - f^*(y_i)) - \sum_{i=1}^k (r_i \xi_i + r_i^* \xi_i^*) \quad (10)$$

式中 $\alpha_i, \alpha_i^*, r_i, r_i^* \geq 0$, 对式(10) $\mathbf{w}, b, \xi_i, \xi_i^*$ 的偏导数都应等于零, 将偏导数结果代入式(10), 得到对偶优化问题。其 α_i, α_i^* 最大化目标函数为:

$$\mathbf{w}(\alpha_i, \alpha_i^*) = \sum_{i=1}^k z_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^k (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) [\phi(y_i) \phi(y_j)] - \sum_{i=1}^k (\alpha_i + \alpha_i^*) \varepsilon \quad (11)$$

其约束条件为:

$$\sum_{i=1}^k (\alpha_i - \alpha_i^*) = 0 \quad (12)$$

$$\alpha_i, \alpha_i^* \in [0, c] \quad (13)$$

由此, 函数的逼近问题就归结为二次规划问题。式(11)中 $\phi(y_i) \cdot \phi(y_j)$ 为高维特征空间的点积运算, 而且函数 ϕ 是未知的、高维的。支持向量机理论只考虑高维特征空间的点积运算 $\mathbf{K}(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$, 而不直接使用

函数 ϕ ,从而有效地解决了这个问题。一般称 $\mathbf{K}(x_i, x_j)$ 为核函数。根据泛函的有关理论,只要一种核函数 $\mathbf{K}(x_i, x_j)$ 满足Mercer条件,它就对应某一变换空间的内积。常用的核函数有:多项式核函数、Radial Basis核函数、Sigmoid核函数和样条(Spline)核函数^[6]。因此只要采用适当的内积函数 $\mathbf{K}(x_i, x_j)$,就可使式(11)中高维空间的内积运算通过原空间的运算来实现,这时式(11)表示为:

$$\mathbf{w}(\alpha, \alpha^*) = \sum_{i=1}^k y_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^k (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \mathbf{K}(x_i, x_j) - \sum_{i=1}^k (\alpha_i + \alpha_i^*) \varepsilon \quad (14)$$

解式(14)在约束条件式(12)、(13)下的二项规划问题得:

$$\mathbf{w} = \sum_{i=1}^k (\alpha_i - \alpha_i^*) \phi(y_i) \quad (15)$$

其逼近函数为:

$$f^*(y) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) \mathbf{K}(y, y_i) + b^* \quad (16)$$

此时 \mathbf{w} 不能显示表示,而核函数的引入,使逼近函数求解绕过特征空间,直接在输入空间上求取,按KKT条件,在鞍点有 $\alpha_i \cdot \alpha_i^* = 0$,即任何一点 α_i 和 α_i^* 都不会同时为非零,由此可以得出,位于不灵敏区内得样本点其对应都等于零,外部的点对应有 $\alpha_i = c$ 或 $\alpha_i^* = c$,而在边界上 ξ_i 和 ξ_i^* 均为零,因而有:

$$\alpha_i, \alpha_i^* \in (0, c) \quad (17)$$

从而得:

$$\varepsilon - z_i + f^*(y_i) = 0 \quad (18)$$

$$\varepsilon + z_i - f^*(y_i) = 0 \quad (19)$$

由以上两式求解得 b ,一般使 $\alpha_i \neq 0$ 和 $\alpha_i^* \neq 0$ 相对应的样本 y_i ,即在不灵敏区边界上或外界样本,称为支持向量(svs),这样式(15)和(16)可表示为:

$$\mathbf{w} = \sum_{i=1}^k (\alpha_i - \alpha_i^*) \phi(y_i) = \sum_{svs} (\alpha_i - \alpha_i^*) \phi(y_i) \quad (20)$$

$$f^*(y) = \sum_{i=1}^k (\alpha_i - \alpha_i^*) k(y, y_i) + b = \sum_{svs} (\alpha_i - \alpha_i^*) k(y, y_i) + b \quad (21)$$

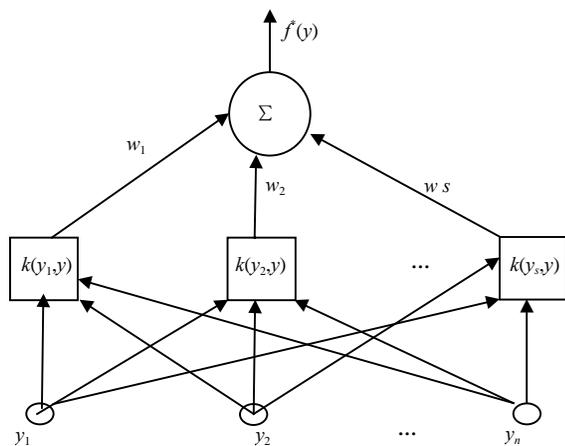


图3 基于SVM非线性校正过程

在样本集中删除非支持向量后重新训练所得结果与原结果相同。也就是说回归它的过程中,只有那些支持向量才起作用。因此,基于SVM非线性校正过程如图3所示,其算法步骤为:

1) 实测(或按理论)采集一组对输入 x_i 传感器输出值 y_i 及理想特性应输出值 $z_i (i=1, 2, \dots, k)$ 组成训练样本和测试样本集。

2) 选择某一核函数作为算法的核函数 $\mathbf{K}(y_i, y_j)$,确定精度误差 ε 和核函数的相应参数,其步骤为:

(1) 选择某一 ε 值固定不变;

(2) 对所选的核函数,通过不断修改其参数,用训练样本进行训练学习,以获得在固定 ε 值下的最优的核函数参数;

(3) 将所获得的最佳参数作为核函数参数,再通过不断修改 ε 的值,使用训练的样本进行学习,以获得最佳的 ε 值;

(4) 用所获得的最佳参数和 ε 值作为校正系统的精度误差和核函数参数。

3) 通过学习后用式(21)作为补偿环节函数 $F(y)$;用测试样本集进行测试;若满足校正要求,则将 $F(y)$ 应用于实际系统,完成对非线性误差的校正,否则,再选择另一核函数,重复2)。

3 应用举例

利用本文所提出的校正算法,对电容式湿敏传感器在5%~95%RH范围内进行非线性校正,传感器经变

换电路和A/D转换后, 每间隔5%RH为一采样点, 所得的特性曲线如图4所示, 显而易见传感器传输特性的两端非线性比较明显; 其中最大误差为5.366 1。平均误差为3.687 6。

采用本文算法校正时, 首先对每一湿度RH值, 在测得 $f(\text{RH})$ 值(即采样所得值)的基础上, 确定期望的输出值 P , 建立一组训练样本集。同时取高斯基RBF函数 $k(y_i, y_j) = \exp(-\frac{(y_i - y_j)^2}{2\delta^2})$ 为系统算法的核函数, 取 $c = \infty$; $\varepsilon = 0.001$, 使用训练样本分别在 $\delta = 0.2$ 和 $\delta = 0.5$ 下, 通过Matlab SVM toolbox进行仿真, 结果如图4b和表1所示, 最大非线性误差基本在给出样本误差(即系统校正精度)0.02之内, 非线性误差大大减少, 经在实际“分布式温、湿度测控系统”中使用, 取得了很好的效果。

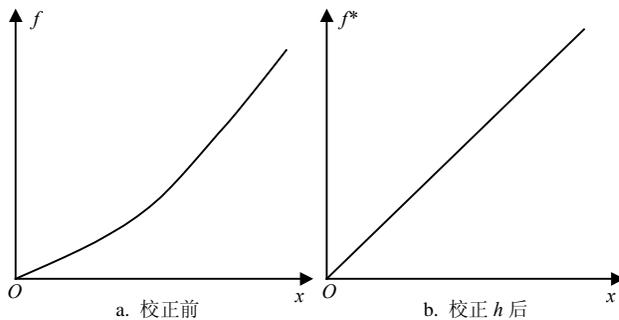


图4 校正前后的结果

表1 仿真校正后的结果

δ	平均误差	最大误差
0.2	0.002	0.006
0.5	0.005	0.012

4 结束语

本文所提出的校正算法, 只依据样本学习就可以正确辨识传感器逆模型特性, 实现非线性校正, 而不需要关于逆模型函数形式的任何先验知识。同时该算法是一个凸二次优化问题, 能够保证找到的极值解就是全局最优解, 并具有良好的泛化能力, 为解决传感器非线性误差校正提供了一种较有效的途径。应用结果表明, 与目前采用的其他算法法想比较, 在校正精度和算法的推广性上都具有一定的优越性。但算法中核函数的选择与其参数的确定, 对校正精度有一定的影响, 如何优化, 有待作进一步研究。

参 考 文 献

- [1] 朱庆保. 用于传感器非线性误差校正的新颖神经网络[J]. 软件学报, 1999, 10(12):1 298-1 303.
- [2] Vapnik C V. Support vector networks[J]. Machine Learning, 1995, 20(1): 1-25.
- [3] Vapnik V N. Statistical learning theory [M]. New York: Jone wileg, 1998.
- [4] Vapnik V. An overview of statistical learning theory[J]. IEEE Transaction on Neural Networks, 1999, 10(5): 988-999.
- [5] Gunn S. Support vector machine for classification and regression[R]. University of Southampton: Image Speck & Intelligent System Group, 1998.

编 辑 孙晓丹