

基于循环神经网络的传感器漂移补偿方法 An RNN-Based Approach to Compensating Sensor Drift

沈文炜, 施惠昌

SHENG Wen-wei, SHI Hui-chang

(上海大学通信与信息学院, 上海 200072)

(School of Communications and Information Engineering, Shanghai University, Shanghai 200072, China)

摘要: 本文讨论了一种基于循环神经网络的传感器补偿新方法。该方法利用循环神经网络对时序信号的捕捉能力, 对传感器的漂移进行预测, 从而减少传感器的校正次数, 提高测量的精度。在训练和预测期间, 充分融合传感器的“同类数据”完成对传感器的训练和漂移预测。

Abstract: An approach to compensating sensor drift based on recurrent neural networks is discussed in the paper. The features of recurrent neural networks will be used to improve the measurement accuracy of sensors and to reduce sensor calibration times by drift prediction. Lots of “congener data” will be integrated into the “real data” to train RNN and predict sensor drift.

关键词: 循环神经网络; 传感器漂移; 非线性补偿; 误差预测

Key words: recurrent neural network; sensor drift; nonlinear compensation; error prediction

中图分类号: TP212

文献标识码: A

1 引言

传感器的漂移是影响传感器测量精度和稳定性的重要因素, 而传感器漂移的诱因复杂, 难以消除。现代控制系统要求传感器具有线性工作特性。实际工程应用中采用标准器对工作的传感器进行定期校正是一种有效手段, 但并非在所有场合都行之有效。例如, 在一些大型设备或需要连续工作的系统中, 这种检测费时费力, 成本很高。

2 传感器漂移的预测校正

神经网络具有很强的非线性映射能力, 一个三层的前馈神经网络具有映射任意函数关系的能力, 且不需要对时间序列作任何假设。RNN(Recurrent Neural Networks, 简称 RNN) 是一种既有前馈通路, 又有反馈通路的神经网络。反馈通路的引入, 使得网络能够有效地处理时间序列的上下文信息。这种特性在语音识别中已经得到应用^[1,2]。同样, 这种特性也可以用于传感器漂移的预测校正。校正框图如图 1 所示^[3], 传感器信号 S 经过预处理模块的

去噪、归一化等处理后, 得到信号 S_c ; 预测模块根据 S_c 对误差进行预测得到 d_p , 最后由 S_c 接受 d_p 的校正。本文在利用递归神经网络进行校正的同时, 对 RNN 网络的训练进行了简化, 略去附加的训练模块 (见图 2)。经实验验证, 可以达到提高其测量精度减少传感器校正次数的目的。

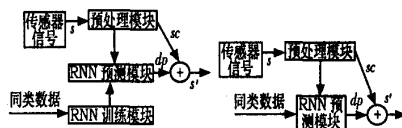


图 1 校正框图

图 2 简化的校正框图

传感器的漂移多受偶然因素的影响, 其行为有随机性, 难于归纳和表述。但是, 大多数传感器的漂移都有其惯性, 这又使得对其进行预测成为可能。利用神经网络的非线性映射特性, 并融合传感器的“同类数据” (同类数据是指同一类型的传感器在近似的工作环境下, 由校正测量取得的漂移数据) 来训练神经网络^[4], 可以实现对传感器漂移的预测。如图 3 所示, d_k 为同类传感器 K 的漂移误差曲线 (实线), d_p 为目标传感器 P 的漂移误差曲线 (虚线)。假设各同类传感器的漂移误差 d_k 已知, 而目标传感器 P 的漂移误差 d_p 未知。我们将依据 $d_1^i, d_2^i, \dots, d_{i-1}^i, d_p^1$ 进行多步预测, 依次

• 收稿日期: 2004-09-01; 修订日期: 2004-12-22

作者简介: 沈文炜 (1976-), 男, 安徽安庆人, 硕士生, 研究方向为信号与信息处理和数据融合; 施惠昌, 教授, 博士生导师, 研究方向为自动检测与控制 and 信号与信息处理。

通讯地址: 200072 上海市延长路 149 号上海大学 M7-222 室; Tel: (021)56332314; E-mail: shen_wenwei@sina100.com

Address: School of Communications and Information Engineering, Shanghai University, Shanghai 200072, P. R. China

预测出 $d_p^2, d_p^3, d_p^4, \dots$ 。这些预测值有一定的校正时间间隔。预测过程中同类数据的作用相当于预测前取得的经验值。这里之所以只使用了 $k-1$ 个同类数据,是因为被排除的一个同类数据序列只参加网络的训练,不参加预测。

具体预测过程如下:在传感器所在系统开始工作时,需要完成对 P 的首次校正,取得真实的漂移数据 d_p^1 ,同时得到同类数据 $d_1^1, d_2^1, \dots, d_{k-1}^1$ 。预测由图 4 所示的循环神经网络进行:由 $d_1^1, d_2^1, \dots, d_{k-1}^1, d_p^1$ 预测得到 d_p^2 ,第二步预测由 $d_1^2, d_2^2, \dots, d_{k-1}^2, d_p^2$ 得到 d_p^3 ,依次类推。这样,我们可以得到与“同类数据”等长的漂移序列。利用所得到的预测序列可以对传感器进行实时校正。

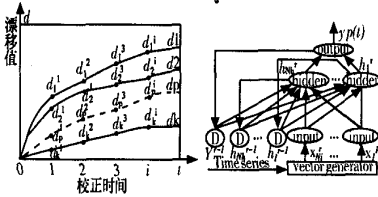


图 3 传感器漂移曲线 图 4 循环神经网络结构

3 基于 RNN 神经网络数据融合的预测

3.1 RNN 的结构

本文采用如图 4 所示三层的循环神经网络结构。循环神经网络中的反馈通路可将某些神经元的输出经过一个或几个时间节拍延迟之后送到其他神经元或自身(图中的 D 节点为延迟单元)。反馈通路的引入,使得网络能够有效地处理漂移时间序列的上下文信息。

3.2 RNN 的训练

网络的训练过程采用方向传播(Back Propagation)算法。该算法的学习过程为:输入信息从输入层单元传到隐层单元,经隐层单元逐层处理后再传到输出层单元和延迟节点,最后由输出层单元处理后产生一个输出模式。如果输出模式与期望模式的误差不能满足要求,则通过修正各层神经元的连接权值,使误差逐步趋于最小值。上述算法的计算步骤如下:

(1)计算隐层。输出神经元的连接权值的上下边界,并初始化权值与阈值。

(2)提交训练向量 p ,由式(1)、(2)计算网络输出值 Y^r 及误差函数值(式(5))。

$$Y^r = \sum_{i=1}^{N_h} w_{hi} h_i^r - s_o \quad (1)$$

$$h_i^r = g(\sum_{i=1}^{N_1} w_{xi} x_i^r + \sum_{i=1}^{N_h} w_{hi} h_i^{r-1} + w_{iy} y^{r-1} - s_j) \quad (2)$$

其中, w_{hi} 为隐层单元 i 与输出单元的连接权值, h_i^r 为隐层单元的输出值, r 为当前的输入向量序号, s_o 为输出单元的阈值, x_i^r 为输入矢量 x^r 的第 i 个分量, Y^{r-1} 为网络在前一时间节拍($r-1$)的输出值, N_1 为输入向量的大小, h_k^{r-1} 为隐层单元 k 在($r-1$)时间节拍的输出值, s_j 为隐层单元的阈值, w_{hj} 为隐层单元 k 与隐层单元 j 的连接权值, w_{ij} 为隐层单元 j 与输出层的连接权值, w_{xi} 为输入单元 i 与隐层 j 的连接权值。

对于激活函数 $g(x)$,选择对数-s 型函数:

$$g(x) = 1/(1 + e^{-x}) \quad (3)$$

(3)更新输出神经元的权值和阈值。网络的训练采用基于梯度下降的反向传播算法^[5]。为提高收敛速度,将使用自适应的训练速率 $\alpha^p(t)$ 。权值的更新由下式得到:

$$\Delta w_{ij}(t) = -\alpha^p(t) \frac{\partial E^p(t)}{\partial w_{ij}(t)}, \Delta s_j(t) = -\alpha^p(t) \frac{\partial E^p(t)}{\partial s_j(t)} \quad (4)$$

其中, $E^p(t)$ 为误差函数,用于计算梯度:

$$E^p(t) = \frac{1}{2} (Y_o^p(t) - D_o^p)^2 \quad (5)$$

$Y_o^p(t)$ 为输入训练向量 p 时,第 t 次迭代的网络的输出值; D_o^p 为网络的理想输出值。

(4)计算隐层神经元的误差值,并更新其权值和阈值。训练的总误差 $E(t)$ 为:

$$E(t) = \sum_{p=1}^P E^p(t) \quad (6)$$

式(3)所示的隐层激活函数的自适应训练速率,以及输出层线性激活函数的自适应学习速率对于训练的速度和稳定都很重要。本文中自适应速率的选取参考了文献[4]。

(5)重复(2)~(4)步,直至网络收敛至 $E(t)$ 小于设定误差值。

预测网络工作得到的预测结果是否可靠、准确,在很大程度上取决于网络的训练质量。然而,实际中取得足够、高质量的 RNN 训练数据是一个很棘手的问题。一般是对现场工作的传感器进行实地校正测量来得到。这有两个困难:(1)系统运行的初期,难以取得足够的训练数据;(2)更为重要的是,正是由于某些设备本身或其工作的环境难以进行传统的校正或者代价高昂,才必须减少或避免必需的校正。这样一来,网络本身又将得不到真实的、足够的训练数据进行训练。文献[4]提出利用“同类数据”对附加的训练网络进行训练,很好地克服了这个矛盾,经过训练的训练网络产生足够多的数据,这些数据可以用于 RNN 预测网络的训练。在此过程中,训练网络提取该类传感器漂移的特性,并对目标传感器的漂移特性进行局部的逼近和再现,从而获得足够的训练样本,以完成对 RNN 的训练。但是,这种方法同时也增加了网络的复杂性。本文将同类数据直接用于 RNN 的训练,略去训练模块;同时也免除了对训练模块的训练,大大简化了预测过程。

如图 3 所示, d_1, d_2, \dots, d_k 为已取得的同类传感器漂移数据曲线,在校正点 i 的漂移值为 $d_1^i, d_2^i, \dots, d_k^i$ 。为得到训练向量,假设 d_1^i 对应的漂移值为预测对象, $d_2^i, d_3^i, \dots, d_k^i$ 对应的漂移值为“同类数据”。我们要求网络可以根据“已知”数据 $d_2^i, d_3^i, \dots, d_k^i$ 预测得到 d_1^i ,以此完成对网络的第一组训练。这组训练集中共包含 i 个训练向量。同样的训练可以假想 d_2^i 的各漂移值为目标值继续进行。这样的训练最多可以有 k 组。全部训练过程共使用了 $i * k$ 个训练向量。由于在训练和预测的过程中网络的结构不变,因而训练向量和预测时的输入向量长度应相同,并且应与输入层节点的个数保持一致。当网络进入预测工作状态时,输入层的一个节点需要接受真实数据,因而此时参加预测的同类数据应该剔除一个。

(下转第 100 页)

必要的时候,可由一个处理器发起仲裁 ID 初始化广播消息,即以 IPI 形式令各模块的仲裁 ID 值恢复到它们的 APIC ID 值的初始情况。

6 结束语

高级可编程中断控制器系统由于其优异的可编程性,使得微处理器对中断的控制更加灵活。通过对本地中断向量和 I/O APIC 重定向表中各个入口的编程,可以把接收到的中断以指定的中断方式发送给处理器去处理。而 IPI 使得在处理器之间传递初始化、启动以及各种中断请求变得更加方便。无论是单处理器系统还是多处理器系统,它都能够对中断的管理提供良好的支持。由于这诸多的优点,高级可编程中断控制器系统才得以在英特尔的主流微处理器中被广泛使用,并且其功能和结构仍在不断趋于完善。相信在今后的微处理器中仍将会看到功能更强大、结构更完善的高级可编程中断控制器系统的身影。

参考文献:

- [1] IA-32 Intel Architecture Software Developer's Manual Volume 3: System Programming Guide[M]. Intel Corp, 2002.
- [2] The Importance of Implementing APIC-Based Interrupt Subsystems on Uniprocessor PC[M]. Intel Corp, 2003.
- [3] Intel 82093AA I/O Advanced Programmable Interrupt Controller (I/O APIC)[M]. Intel Corp, 1996.

(上接第 96 页)

4 示例

本例中选择使用的 RNN 如前所述。具体为三层:输入层含八个节点,隐层六个节点,输出层含一个节点。网络的训练需要大量的漂移数据。由于大量的传感器漂移数据表现为二阶、非平稳的序列,并伴随着随机性。为此,我们将采用式(7)所示的数学模型来产生训练数据。

$$y = a + (b + c) \sqrt{x/c} + d \sin(ex + c) + fR_1 + g \sin(x - c) + hR_2 \quad (7)$$

其中, $a \sim h$ 为常数, R_1 为均匀分布的随机数, R_2 为正态分布的随机数。

将 $data1 \sim data4$ 、 $data6 \sim data9$ 曲线作为“同类数据”,从中等间隔取得 8×30 个训练数据,对网络进行训练。完成训练后,网络将 $data5$ 作为预测对象,接收初始校正正值以及同类数据本进行预测。 $data10$ 为网络工作的结果;预测的间隔与取得训练数据的间隔一致。各校正点相对于实际漂移的预测误差如图 5 所示。

由图 5 可见,网络的预测在初期误差偏大,随着网络对漂移的捕捉以及漂移趋向饱和,误差将趋于减小。在 30 步预测中只有初期两次超过和接近 10%,总的平均误差为 3.15%。由于该误差率是以传感器的漂移为基准,因此经过该网络的校正后,一般来说传感器信号精度可提高 1 至 2 个数量级。

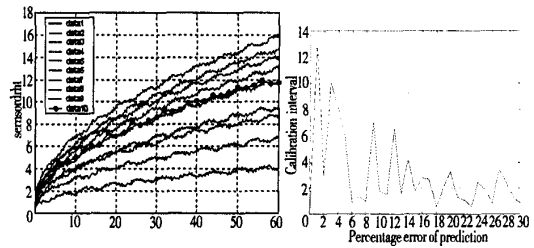


图 4 传感器漂移校正

图 5 传感器漂移预测误差

5 结束语

本文将神经网络作为传感器校正漂移误差的基本模块,经充分训练的神经网络在融合“同类数据”的基础上,可以预测和跟踪误差的发展趋势。这种方法可以在各种测量仪器的误差校正中方便地实现,并且只占用较少的系统资源。由于网络的预测质量与“同类数据”的质量和数量紧密相关,所选取的训练数据必须具有典型性。同时,可以在多大程度上减少传感器的校正次数也依赖于所取的“同类数据”的长度。

递归神经网络本身具有动态映射能力,进行设计时,不需要知道被补偿传感器的结构特性等先验知识,简化了补偿系统的结构设计。神经元的响应函数反映了神经网络模型逼近非线性函数的能力。增加递归神经网络模型的隐层节点数有助于提高动态补偿的精度,但会增加网络的复杂程度和训练的时间^[6]。在实际应用中,应根据实际情况选取合适的神经元响应函数,以便得到最佳的神经网络模型结构。

参考文献:

- [1] 朱小燕,王昱,徐伟.基于循环神经网络的语音识别模型[J].中文信息学报,2001,15(2):45-50.
- [2] Tebelskis J. Speech Using Neural Networks[R]. Technical Report CMU-CS-95-142, Carnegie Mellon University, 1995.
- [3] 施惠昌.一种利用函数链神经网络进行传感器建模的新方法[J].传感器技术,2000,19(3):21-24.
- [4] Turchenko V, Kochan V, Sachenko A, et al. Advanced Sensor Data Integration Using Neural Networks[A]. IEEE 2002 28th Annual Conf on Industrial Electronics Society (IECON'02). Vol 3[C]. 2002. 1876-1880.
- [5] Martin T Hagan, Howard B Demuth, Mark H Beale. Neural Network Design[M]. Beijing: China Machine Press, 1996.
- [6] 田社平.基于递归神经网络模型的传感器非线性动态补偿[J].上海交通大学学报,2003,37(1):13-16.