

基于优化 BP 神经网络的物理量回归方法

潘俊虹^{1,2,3} 王宜怀^{2,4} 吴 薇^{1,2,3}

(武夷学院数学与计算机学院 福建 武夷山 354300)¹

(苏州大学计算机科学与技术学院 江苏 苏州 215006)²

(认知计算与智能信息处理福建省高校重点实验室 福建 武夷山 354300)³

(软件新技术与产业化协同创新中心 江苏 苏州 215006)⁴

摘要 在物联网实际应用系统的开发中,传统回归方法面对 A/D 转换物理量回归时存在表达方式不统一、非线性校正能力及动态适应性弱等问题。文中在分析 A/D 转换物理量回归要素的基础上,依据 BP 神经网络的非线性映射能力,提出了利用布谷鸟算法进行优化的 BP 神经网络,并利用其实现统一数学表达的 A/D 转换物理量回归方法。实践表明,该方法具有数学公式统一、非线性校正能力及动态适应性强等特点。该方法既适用于利用通信方式将 A/D 采集的数据直接送至 PC 机处理的物联网系统,也适用于利用 PC 机进行学习,将神经网络结构参数存储于 MCU 内的 Flash 中,在物联网终端直接将 A/D 值转为实际物理量的环境。

关键词 A/D 转换, BP 神经网络, 布谷鸟搜索算法, 物理量回归, 动态校正

中图分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.12.027

Physical Quantity Regression Method Based on Optimized BP Neural Network

PAN Jun-hong^{1,2,3} WANG Yi-huai^{2,4} WU Wei^{1,2,3}

(School of Mathematics and Computer, Wuyi University, Wuyishan, Fujian 354300, China)¹

(Department of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006, China)²

(Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian

Education Institutions, Wuyishan, Fujian 354300, China)³

(Collaborative Innovation Center of Novel Software Technology and Industrialization, Suzhou, Jiangsu 215006, China)⁴

Abstract In the development of practical application system of Internet of Things (IoT), traditional regression methods have the problems of non-uniform expression, poor non-linear correction ability and dynamic adaptability for physical quantity regression of A/D conversion. Based on the analysis of the physical quantity regression elements of A/D conversion, and according to the non-linear mapping ability in BP artificial neural network, this paper proposed a BP neural network optimized by cuckoo search algorithm, and utilized it to realize physical quantity regression method of A/D conversion with unified mathematical expression. Practice shows that this method has the characteristics of uniform mathematical formula, strong nonlinear correction ability and dynamic adaptability. This method is not only suitable for IoT system using communication methods to send A/D acquisition data to PC directly, but also suitable for the environment in which PC is used to learn. The neural network structure parameters are stored in the Flash of MCU, and the A/D value is directly converted to the actual physical quantity at the terminal IoT.

Keywords A/D conversion, BP neural network, Cuckoo search algorithm, Physical quantity regression, Dynamical correction

1 引言

在物联网系统实际应用中,采集终端实际模拟信号通常经过模数转换模块(Analog To Digital Convert Module, A/D 转换模块)将模拟信号转化为数字信号,供上位机进行处理。

得到稳定的 A/D 采样值后,还需要在人机交互终端把 A/D 值回归成实际的物理量形式,这个过程称作物理量回归^[1]。由于在实际应用中,从传感器到信号放大器再到 A/D 转换模块的一系列信号传输过程中,任何电路的变化都会导致信号的非线性变化(见图 1),并最终导致实际物理量输出 $F(x)$ 的

到稿日期:2017-11-21 返修日期:2018-01-04 本文受国家自然科学基金(61672369),福建省教育厅科研基金(JAT160521),武夷学院校科研基金(XD201506)资助。

潘俊虹(1980—),男,硕士,高级工程师,主要研究方向为嵌入式技术、物联网, E-mail: pjhtj1980@126.com(通信作者);王宜怀(1962—),男,博士,教授,主要研究方向为嵌入式系统、神经网络;吴 薇(1978—),女,硕士,副教授,主要研究方向为嵌入式系统、计算机应用技术。

非线性变化,因此物理量回归准确与否直接决定系统性能。对于非线性的物理量回归,通常都是采用各种方法确定一个数学公式,近似地逼近 A/D 数字量和物理量的映射曲线进行回归。

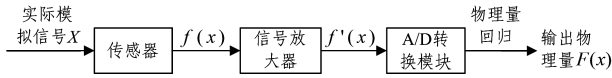


图 1 物理量转换示意图

Fig. 1 Schematic diagram of physical quantity conversion

传统的物理量回归方法有公式法、查表法、基于最小二乘的分段直线、多项式回归法以及三次样条插值法等,不同的方法适用于不同的场景。低端系统中由于硬件资源有限,适合耗费资源少且快速的回归方法,如公式法通过固定的转换公式来直接换算,虽然速度快,但精确度较差。采用查表法进行物理量回归时,对于转换精度较高的 16 位 A/D 转换模块来说,由于数据量庞大,查表法显得无能为力。应用在高端系统中时,有强大的硬件计算资源的支持,可采用更加灵活的技术方案,如基于最小二乘的分段直线方式动态校正物理量回归,但分段直线将非线性问题分割成若干个线性区间,回归准确度不足。无论哪种回归方法,其目的都在于尽量减小或消除非线性问题产生的误差,确定一个带有参数的经验模型来逼近真实物理量。另一方面,实际应用中不同类型的传感器、电路特性也会影响回归算法的选择。因此,以上传统方法较难在不同系统、不同阶段的物理量回归问题上实现统一的回归表达方式。

实验表明,温度、湿度、光照、噪音等环境物理量的变化和 MCU 数字转换 A/D 值之间的映射关系,在测量区间内基本上为一条单调非线性连续曲线,而人工神经网络是一个强大的非线性逼近的动态自学习系统,其在复杂问题的非线性预测方面表现得尤为突出^[2]。Hecht-Nielsen^[3]于 1989 年用理论证明了单隐层的三层神经网络具有逼近任意连续函数的能力。Ito^[4]也于 1991 年证明了神经网络隐层函数采用阶跃函数或 Sigmoid 型函数对任意连续函数的逼近能力。基于此,人工神经网络在很多领域都得到了广泛应用。如文献[5]利用人工神经网络进行语音质量预测;文献[6]将人工神经网络应用于传感器等硬件进行非线性校正;文献[7]采用人工神经网络进行水文预测,得到了较好的效果。本文首先分析了 BP 神经网络的基本原理,针对 BP 神经网络过渡依赖网络权值和阈值初始值等问题,采用布谷鸟搜索算法对其进行优化,以提高网络收敛速度和精度,并用优化的 BP 神经网络实现对 A/D 值非线性物理量的回归和校正。与其他常用的物理量回归方法相比,本文所提方法不仅有较高的精度,而且能适应不同的物联网应用阶段和环境要求,并能实现物理量回归的统一数学表达和更加灵活的动态在线校正功能。

2 基于 CS 优化 BP 神经网络

2.1 BP 人工神经网络的原理

Rumelhart 等于 1986 年提出了反向传播(Back Propagation, BP)神经网络的概念。BP 神经网络是一种按误差反向

传播训练的多层前馈网络。BP 神经网络采用梯度下降法来计算目标函数的最小值,具有任意复杂的模式分类能力和优良的多维函数映射能力。基本的 BP 算法包括信号的前向传播和误差的反向传播两个过程,属于有教师信号学习网络,其原理如图 2 所示。

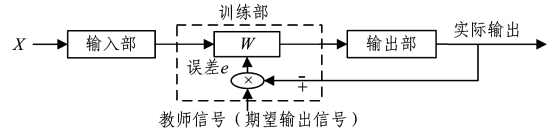


图 2 BP 神经网络原理图

Fig. 2 Principle diagram of BP neural network

设有一个 n 层结构的 BP 神经网络,输入样本 $x = \{x_i, i = 1, 2, \dots, n\}$,网络预测输出为 $y = \{y_i, i = 1, 2, \dots, n\}$,隐层第 i 层到第 j 层神经元的连接权值为 w_{ij} ,神经元的激活函数一般使用 sigmoid 函数:

$$f(x) = 1 / (1 + e^{-x_i}) \quad (1)$$

前导输出公式为:

$$U_i = \sum_{j=1}^n (w_{ij} f(x_j) - \theta_j) \quad (2)$$

其中, θ_j 为第 j 层的阈值,用来调整第 j 层的单元活性。

神经网络的误差计算、传递从输出层开始,一直递归传递到输入层,在这个反向传播的过程中,经过多个输入模式样本的反复训练学习,不断修正网络权值和阈值,使得误差逐渐减小,以至最终消除误差,该过程就是 BP 神经网络的逆向反馈。对于输出层,有:

$$E_j = O_j(1 - O_j)(T_j - O_j) \quad (3)$$

处于网络中间的隐层通过下一层的误差按权重累加,不与数据直接关联,计算式如下:

$$E_j = O_j(1 - O_j) \sum_{k=1}^n E_k W_{jk} \quad (4)$$

其中, W_{jk} 表示当前层节点 j 到下一层节点 k 的网络连接权值, E_k 为下一层节点 k 的误差率。通过误差率对网络权值和阈值进行更新操作。

更新网络连接权值:

$$\Delta W_{ij} = \lambda E_j O_i \quad (5)$$

$$W_{ij} = W_{ij} + \Delta W_{ij} \quad (6)$$

更新阈值:

$$\Delta \theta_j = \lambda E_j \quad (7)$$

$$\theta_j = \theta_j + \Delta \theta_j \quad (8)$$

其中, λ 表示学习速率,速率越大,收敛就越快,但网络容易陷入局部最优解;学习速率越小,收敛速度越就慢,但可以逐步逼近全局最优解。在实际应用中,学习速率通常在 0~1 之间。

2.2 BP 神经网络回归的非线性数学表达式提取

一个 BP 神经网络学习结束后,就可以确定一组网络的权值和阈值,也就是说可以给出一个数学模型来描述这个最终确定的网络。当需要解决类似实际问题时,可以根据这个确定的数学表达式直接恢复出 BP 神经网络,将输入提交给网络,并由其运行产生输出。

设提交给网络的输入为 $x = \{x_i, i = 1, 2, \dots, n\}$, n 为输入维数。根据输入层与中间层的连接权值以及中间层各节点的

阈值,可计算得到中间层各节点的输入,计算式如下:

$$s_i = \sum_{j=1}^m w_{ij} x_j - \theta_j, j=1, 2, \dots, m \quad (9)$$

其中, m 为中间层节点个数。通过中间层各个节点的输出 $b_j = f(s_j)$ (函数 f 由激活函数确定), 以及中间层到输出层的网络权值和阈值, 可计算得到输出层的输入:

$$g_t = \sum_{i=1}^m w_{it} b_i - \theta_t, t=1, 2, \dots, p \quad (10)$$

其中, p 为输出层节点个数。再由输出层激活函数即可得到输出层各节点的输出:

$$F(x) = f(g_t) \quad (11)$$

BP神经网络在实际应用过程中存在一定缺陷: 较为依赖初始权值和阈值; 均方误差梯度曲线特性易陷入局部极小值, 从而导致训练失败; 网络权值、阈值随机初始化, 可能导致每次训练的结果不尽相同等^[8]。下文将利用布谷鸟搜索算法对其进行优化。

2.3 基于布谷鸟算法优化BP神经网络(CS-BP网络)

剑桥大学的Yang等于2009年提出了一种新式搜索算法——布谷鸟搜索算法CS(Cuckoo Search)^[9]。CS算法通过模拟布谷鸟的寄生育雏行为, 结合莱维(Levy)飞行搜索求解最优化问题。CS算法是一种性能优异的全局寻优算法, 能够在复杂的多维向量空间快速逼近全局最优解。实验证明, 布谷鸟算法在收敛速度和精度方面都优于粒子群优化算法(Particle Swarm Optimization, PSO)和遗传算法(Genetic Algorithm, GA)等传统搜索算法^[10-13]。

CS搜索算法的基本假设如下:

- (1) 每只布谷鸟一次只产一枚卵, 并随机选择寄生鸟巢对卵进行孵化;
- (2) 在随机选择的一组寄生鸟巢中, 位置最好的寄生鸟巢会被保留到下一代孵化使用;
- (3) 布谷鸟可利用的寄生鸟巢数量固定, 鸟巢原主人发现外来陌生鸟蛋的概率是 P (即新位置产生概率)。

CS算法中鸟巢位置的更新规则:

$$x_{g+1,i} = x_{g,i} + \alpha \oplus L(\beta), i=1, 2, \dots, n \quad (12)$$

其中, $x_{g,i}$ 表示迭代到 g 代鸟巢时, 第 i 个鸟巢所处的位置; α 表示步长控制量; \oplus 表示点对点乘法; $L(\beta)$ 表示莱维飞行随机搜索路径, 服从莱维概率分布:

$$L(\beta) \sim u = t^{-\beta}, 1 < \beta \leq 3 \quad (13)$$

在标准布谷鸟优化算法中, 利用莱维飞行随机产生步长, 步长较小会降低搜索速度, 步长较大降低搜索精度, 通常采用的步长值为 1。

CS搜索算法的基本流程如下:

- (1) 随机初始化 n 个鸟巢的位置 (x_1, x_2, \dots, x_n) ;
- (2) 计算每个鸟巢位置的适应度函数值 $f(x_i)$, 并记录当前最优鸟巢位置;
- (3) 保留最优鸟巢的位置不变, 利用鸟巢位置更新计算式(12)来更新其余鸟巢的位置;
- (4) 将更新后的鸟巢位置与更新前的鸟巢位置进行对比, 若步骤(3)中的新位置较好, 则接受该位置;
- (5) 以宿主发现外来陌生卵的概率随机产生新鸟巢来替

代丢弃的鸟巢, 至此得到新一代鸟巢;

(6) 未达到结束条件, 返回步骤(2)继续执行;

(7) 符合结束条件, 算法结束, 输出最优解。

为了消除BP网络对初始值的敏感性, 提升BP网络的训练速度和精度, 利用布谷鸟算法来寻优网络初始参数, 具体步骤如下: 首先将随机产生的网络参数编码成鸟巢位置向量, 再把BP算法的目标函数(均方误差函数计算式(14))作为鸟巢的适应度函数, 在预设的迭代次数或者误差限定内搜索最优的网络连接权值和阈值, 利用布谷鸟算法找到近似的全局最优解, 最后用得到的最优初始权值和阈值来初始化BP网络, 对BP网络进行训练。

BP网络适应度函数:

$$f(i) = \sum_{k=1}^Q (y_k - \bar{y}_k)^2 \quad (14)$$

其中, Q 为样本数, y_k 为实测值, \bar{y}_k 为网络预测输出值。

基于布谷鸟算法CS优化的BP神经网络算法的流程如图3所示。

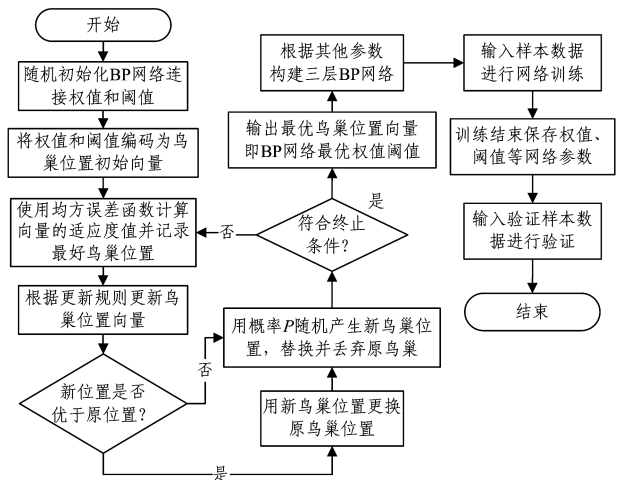


图3 布谷鸟算法优化的BP网络的流程

Fig. 3 Flowchart of BP network optimized by cuckoo algorithm

3 基于CS-BP神经网络的A/D值物理量回归算法

3.1 BP网络的结构设计

根据上文所述, 建立单隐层的三层BP神经网络模型, 如图4所示。首先进行BP网络各项参数的设置: 输入层节点个数为 1, 输出层节点个数为 1, 隐层节点个数根据经验计算式(15)^[14]可得到一个初步范围。然后使用原始样本测试一组隐层节点数据对网络训练效率的影响, 测试结果如表1所列。在固定学习 5000 次的情况下, 隐层节点数为 5 时, BP网络的训练效果最佳。因此, 本文将隐层节点个数设为 5, 学习次数设为 5000, 由于本文实验对象 A/D 值物理量回归曲线相对平缓, 网络学习速率宜设较大值, 本文设为 0.6, 误差目标为 1×10^{-6} 。隐层激活函数采用 Sigmoid 函数。

隐层节点个数经验公式为:

$$H_{\text{num}} = \sqrt{in_{\text{num}} + out_{\text{num}}} + a \quad (15)$$

其中, H_{num} 为隐层节点数, in_{num} 为输入层节点数, out_{num} 为输出

层节点数, a 为 1~10 之间的常数。

需要布谷鸟算法进行选优的权值和阈值总数可采用下式确定:

$$n = (n_{in} + 1) \times n_h + (n_h + 1) \times n_{out} \quad (16)$$

其中, n_{in} 为输入层节点数, n_h 为隐层节点数, n_{out} 为输出层节点数。本文设计的 BP 神经网络结构的输入数和输出数均为 1, 隐层节点个数为 5, 由计算式可得需要优化的权值、阈值总数为 16。

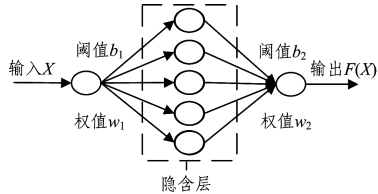


图 4 单隐层 5 节点的三层 BP 网络

Fig. 4 Triple-pane BP network of 5 nodes in single hidden

表 1 隐层节点数训练效率对比

Table 1 Training efficiency comparison for number of hidden nodes

隐层节点数	均方误差 (MSE)	确定系数 (R-square)
3	3.26094e-4	0.9937
4	1.81734e-4	0.9968
5	1.47391e-4	0.9998
6	4.06566e-4	0.9961
7	3.2607e-4	0.9967
8	5.83357e-4	0.9857
9	6.01016e-4	0.9833
10	1.75065e-3	0.9723
11	1.46814e-3	0.9769
12	1.85362e-3	0.9632

3.2 A/D 值物理量回归算法

BP 网络结构设计完成后, 就可以结合布谷鸟算法形成完整的 A/D 值物理量回归算法。基于 CS-BP 神经网络的物理量回归算法的步骤如下:

- (1) 根据上文的 BP 网络结构设计进行网络参数的初始化。
- (2) 随机初始化 BP 网络各层的权值和阈值, 并将其编码成布谷鸟算法中一个鸟巢的空间位置向量; 将计算式(14)均方误差函数作为布谷鸟算法的适应度函数; 利用布谷鸟算法进行全局寻优, 得到 BP 网络的最优初始化权值和阈值。
- (3) 将采集的温度 A/D 值和实测值样本数据进行归一化处理, 提供给 BP 网络进行训练, 当训练到给定的次数或者误差限值时, 训练终止。保存最终的权值和阈值及其相关参数。
- (4) 用测试数据对 BP 网络进行验证, 将网络输出结果反归一化后得到最终温度值。

4 实验过程及分析

4.1 实验过程

本文以温度值的物理量回归为例进行实验。温度采集终端由以恩智浦公司的 KL25Z128V1K4^[15] (简称 KL25) 芯片为核心的评估板连接 AD590IH 温度传感器构成, KL25 运行速率可达 48 MHz, 拥有 128KB Flash 和 16KB SRAM, 内置的 16 位 A/D 转换模块将传感器的模拟电信号转换成数字信号, 并通过集成的 UART 串口通信模块发送至高端的计算机 (实际应用中还可采用无线射频传输等其他通信手段)。

A/D 模块的采样精度为 16 位, 按照顺序在最低 0 °C 到最高 125 °C 区间共取 20 个样本点, 测得实际温度和对应的 A/D 数字量。将样本中的 14 份用来训练神经网络, 3 份用来验证网络输出, 剩余 3 份用来测试。

为了方便存取数据, 本文采用文本文件形式存储样本数据, 数据分为两列, 第一列为 A/D 采样数字量, 第二列为对应的实测温度值, 如表 2 所列。

表 2 温度 A/D 值样本存储表

Table 2 Storage table of temperature A/D value sample

A/D 数字量值 (十六进制)	实测温度值 /°C	A/D 数字量值 (十六进制)	实测温度值 /°C
2DE6	1.1	34E3	47.2
2E6E	3.5	35B2	54.3
2EE3	5.2	3625	58.6
2F91	8.5	36D0	66.7
30B3	12.2	374F	75.6
316F	18.6	37E5	85.5
3243	23.3	386F	96.4
32FF	26.3	38C1	105.8
3363	31.6	398B	113.1
3417	38.2	3AE3	123.8

CS-BP 神经网络训练误差的变化趋势如图 5 所示。由图 5 可知, 在 BP 网络训练迭代 32 次时, 验证误差达到最小值, 网络训练终止, 此时网络各层的最佳权值和阈值如表 3 所列, 由这些参数即可确定本文物理量回归的 CS-BP 神经网络结构。

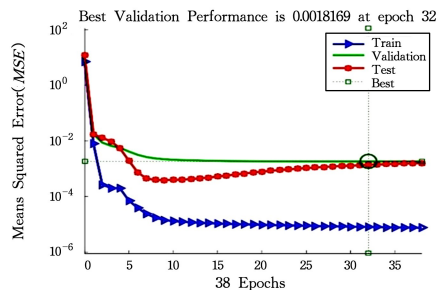


图 5 CS-BP 神经网络训练误差的变化趋势

Fig. 5 Changing trend of training error in CS-BP neural network

表 3 CS-BP 神经网络的各层参数

Table 3 Parameters of each layer in CS-BP neural network

输入层到中间隐层的连接权值 $w_i^{(1,2)}$		输入层到中间隐层的阈值 $b_i^{(1,2)}$		中间隐层到输出层的连接权值 $w_i^{(2,3)}$		中间隐层到输出层的阈值 $b_i^{(2,3)}$	
w_1	7.144165602	b_1	-6.704270140	w_1	-0.11023973512		
w_2	-2.31374844	b_2	3.8537185676	w_2	0.10301946435		
w_3	-5.94881878	b_3	0.5964983573	w_3	0.231431336453	b	0.18951646753
w_4	-1.12739589	b_4	-0.676522409	w_4	0.845432678067		
w_5	6.585021038	b_5	6.1906915533	w_5	-0.02618357038		

4.2 A/D值物理量回归算法的对比

常用的物理量回归法中公式法根据热敏电阻与温度之间的函数关系给出了一个固定的公式,一般由厂家的技术手册提供,通过 A/D 值直接计算得出物理量。而查表法则建立在公式法的基础上,通过公式计算出一定范围内的(如 0~100℃)每 0.1℃对应的 A/D 值并建立有序表,使用时根据 A/D 值查找相对应的物理量(四舍五入)。最小二乘法和三次样条插值法则分别是利用最小二乘原理和样条插值原理在给定区间内对原始曲线进行逼近的方式进行物理量的回归。

表 4 列出了常用物理量回归方法与本文设计的 CS-BP 神经网络物理量回归方法的性能指标对比。图 6 给出了各物理量回归方式的回归曲线。

表 4 物理量回归方法的性能对比

Table 4 Performance comparison of physical quantity regression methods

回归方法	均方差(MSE)	相关系数(R)	运行时间/s
公式法	7.6312e-3	0.9465	0.056
查表法			0.183
最小二乘	3.64453e-3	0.9727	0.552
三次样条插值	1.0596e-4	0.9993	0.638
CS-BP 神经网络	1.0479e-4	0.9998	0.891

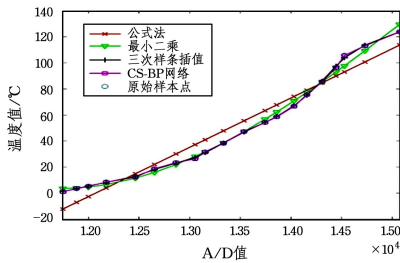


图 6 物理量回归方法的对比曲线

Fig. 6 Comparison curves of physical quantity regression methods

4.3 结果分析

从表 4 和图 6 可以看出,以线性映射回归的公式法和查表法的速度较快,但因为是固定公式运算,无法考虑运行过程中的非线性因素,所以精度不足、灵活性不够,适用于低端系统。而人工神经网络的学习的预测曲线比最小二乘拟合和三次样条插值拟合更贴近真实曲线,效果更好,但耗时相对更长。基于最小二乘原理和样条插值原理的回归在本质上依然是线性回归模型,虽然形式简单、易于建模,但对于非线性映射问题无法进行统一的数学表达。而人工神经网络具有层级结构以及高维映射,拥有强大的动态自学习调整能力,在任何阶段都能由训练学习最终得到统一的数学表达方式。理论上,它是一种更加适合 A/D 值物理量回归的方法。

虽然人工神经网络不是通过确定的数学公式来描述真实的映射关系,但是它经过学习后得到的网络权值、阈值参数是确定的,由此得到的神经网络模型也是确定的。当一些特定的问题无法用准确的数学公式表征时,人工神经网络可以提供对实际映射更加接近的表述,这也是人工神经网络的优势。当然,从运算复杂度来看,人工神经网络的运算量要大于最小二乘和三次样条插值的运算量,时间消耗相对较长,但本文认为对于硬件资源丰富的高端计算机来说,这一缺陷并不明显。

在物理量回归这类实际问题的处理上,用这些时间的消耗来换取回归准确率的有效提升以及回归方法的统一是值得的,也是有意义的。

5 基于 CS-BP 的 A/D 值物理量回归及其在线动态校正中的应用

5.1 基于 CS-BP 的 A/D 值物理量回归应用

由于神经网络训练学习通常需要大量的运算,在物联网系统高端计算机等硬件资源丰富的平台上能够正常运行,但在资源有限的低端 MCU 上运行时效率相对较低。因此,可以在系统高端使用计算能力强大的计算机来进行网络学习训练,训练完毕后最终的网络状态是确定的,可以根据设计的网络结构和训练参数提取其数学表达式。

如 3.1 节设计的 BP 神经网络结构为 3 层(1,5,1)型神经网络,即输入节点数 1,单隐层节点数为 5,输出节点数也为 1 的神经网络模型。该网络模型根据式(9)一式(11)以及表 3 中确定的各层权值、阈值参数以及隐层激活函数提取其对应的非线性表达式,如式(17)所示:

$$F(x) = \sum_{i=1}^5 (\omega_i^{(2,3)} \times \text{sig}(\omega_i^{(1,2)} \times x + b_i^{(1,2)})) + b^{(2,3)} \quad (17)$$

其中, sig 为本文使用的 sigmoid 隐层激活函数 $f(x) = 1/(1 + e^{-x})$, $\omega_i^{(2,3)}$ 为隐层第 i 个节点到输出层节点(网络中第 2 层到第 3 层)的连接权值(见表 3,下同), $\omega_i^{(1,2)}$ 为输入层到隐层第 i 个节点(网络中第 1 层到第 2 层)的连接权值, $b_i^{(1,2)}$ 为输入层到隐层第 i 个节点的阈值, $b_i^{(2,3)}$ 为隐层第 i 个节点到输出层的阈值。

将训练所得权值和阈值及其相关参数都以变量的形式固化在低端 MCU 的 flash 存储器中,MCU 在需要时就可根据 flash 存储器中的参数以及数学表达式快速复原网络回归数学模型,这样就可实现物理量在低端系统中的快速物理量回归,系统测试界面如图 7 所示。

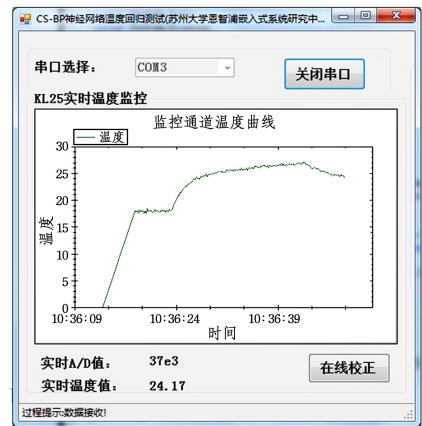


图 7 基于 CS-BP 神经网络的物理量回归测试

Fig. 7 Physical regression test based on CS-BP neural network

表 5 列出了 0~120℃ 区间内系统回归的 10 个物理量和实测结果,结果显示误差均不超过 0.3℃,说明 CS-BP 神经网络回归算法达到了设计的要求。为了避免 flash 擦写时的误操作,本文在 flash 中单独定义 F_BPdta 的参数段,用以保存 BP 网络的相关参数。这样在操作时就不需要移动其他存储在该块的数据。

表 5 CS-BP 网络输出测试
Table 5 CS-BP network output test

(单位:℃)			
A/D 数字量值 (十六进制)	实测温度值	网络输出值	误差
2E6E	3.5	3.5	0
2F91	8.5	8.6	0.1
30B3	12.2	12.3	0.1
3243	23.3	23.5	0.2
3417	38.2	38.2	0
34E3	47.2	47.2	0
3625	58.6	58.4	0.2
37E5	85.5	85.4	0.1
38C1	105.8	106.0	0.2
398B	113.1	112.8	0.3

事实上,近年来 MCU 性能已经得到大幅提高,一些简单的神经网络的学习训练可移入 MCU 进行。本文提出的基于布谷鸟优化 BP 神经网络的物理量回归及校正方法已应用于课题组研发的“工业室内空气净化”“无线气象监测”“智能路灯监控系统”等项目中,到目前为止运行情况良好。

5.2 基于 CS-BP 的 A/D 值物理量在线动态校正

在实际应用过程中,物联网应用系统所处环境的变化、元器件以及通信电路的老化等原因,都会导致原先物理量回归算法逐渐偏离实际值。为了更加灵活且准确地进行物理量回归,本文设计了一种基于 A/D 值物理量回归算法的在线动态校正技术方案,基本方法如下。

系统运行过程中,如果发现输出值不够准确,通过计算机人机交互端把准确的样本加入到训练数据集中,在达到设定的样本数量或者周期后,程序自动根据设定的参数构建神经网络进行训练。训练完成后,记录神经网络学习最佳权值、阈值以及状态参数,以确定新的神经网络模型,以后的物理量输出采用最新确定的网络进行回归,从而实现物理量回归的自动在线校正功能。获得新的权值、阈值及相关参数后,也可将其写入到终端 MCU 的 Flash 存储器当中,以对原来的权值、阈值等参数进行更新,在低端实现快速且动态的在线校正功能。物理量回归的动态校正流程如图 8 所示。

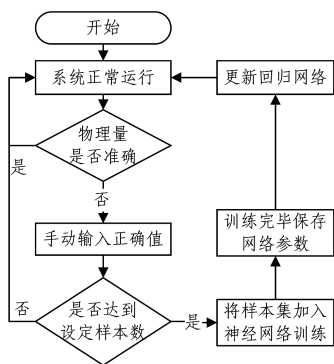


图 8 物理量回归的在线动态校正流程

Fig. 8 Online dynamic correction flow of physical quantity regression

结束语 神经网络拥有强大的非线性映射能力,适用于 A/D 数字量的非线性物理量回归。文中针对 BP 神经网络的缺陷,利用布谷鸟算法优化初始参数,并将优化后的参数提供给 BP 网络进行训练,提高了 BP 网络的训练效率。实验和实际应用表明,使用优化后的神经网络实现物理量回归是一种较为精确可靠的方法。同时,由于人工神经网络的构建

只与当前输入的样本有关,与系统硬件、所处环境均无关,在不同的应用场景中,都可通过神经网络建立统一的数学模型表达。因此,由人工神经网络建立的物理量回归和校正模型可用于各种物联网应用系统及其不同阶段,具有广泛的适用性。

需要注意的是,在设计 BP 网络结构时,与网络相关的各种参数的设置(如中间隐层神经节点个数、学习步长、迭代次数等)都会影响训练的效率和精度,有些参数能够根据理论分析给出,有些参数则必须根据实际情况进行多次调整。由于增加了网络优化步骤,客观上使得网络学习训练的耗时随之增加,下一步将研究如何进一步优化网络,改善网络结构设计,节省网络运行耗时,进一步提高算法的整体性能。

参考文献

[1] CAO J H, HE L X, WANG Y H. AD Conversion Technique of Online Dynamic Reviser Based on KL25[J]. Research and Exploration in Laboratory, 2013, 32(10): 250-252. (in Chinese)
曹金华, 贺黎潇, 王宜怀. 基于 KL25 的 A/D 转换动态在线校正技术[J]. 实验室研究与探索, 2013, 32(10): 250-252.

[2] ZHANG G C, GAO Q, DU J Q, et al. Rockburst criterion based on artificial neural networks and nonlinear regression[J]. Journal of Central South University (Science and Technology), 2013, 44(7): 2977-2981. (in Chinese)
张光存, 高谦, 杜聚强, 等. 基于神经网络及非线性回归的岩爆判据[J]. 中南大学学报(自然科学版), 2013, 44(7): 2977-2981.

[3] HECHT-NIELSEN R. KOLMOGOROV's Mapping Neural Network Existence Theorem[C]// International 1989 Joint Conference on Neural Networks. 1989: 593-605.

[4] ITO Y. Approximate representation of a continuous function by a neural network with scaled or unscaled sigmoid units[C]// IJCNN-91-Seattle International Joint Conference on Neural Networks. 1991: 889.

[5] ORCIK L, VOZNAKM, ROZHON J, et al. Prediction of Speech Quality Based On Resilient Backpropagation Artificial Neural Network[J]. Wireless Personal Communications, 2016, 96(4): 1-15.

[6] PENG J W, LV W H, XING H Y, et al. Temperature compensation for humidity sensor based on improved AG-BP neural network[J]. Chinese Journal of Scientific Instrument, 2013, 34(1): 156-159. (in Chinese)
彭基伟, 吕文华, 行鸿彦, 等. 基于改进 GA_BP 神经网络的湿度传感器的温度补偿[J]. 仪器仪表学报, 2013, 34(1): 156-159.

[7] CHEN Y H, CHANG F J. Evolutionary artificial neural networks for hydrological systems forecasting[J]. Hydrol, 2009, 367(1): 125-137.

[8] LI Y, WU M Y, ZOU L. BP neural network for image restoration based on cuckoo algorithm[J]. Journal of Computer Applications, 2017, 37(S1): 173-175. (in Chinese)
李扬, 吴敏渊, 邹炼. 基于布谷鸟算法的 BP 神经网络图像复原[J]. 计算机应用, 2017, 37(S1): 173-175.

[9] YANG X S, DEB S. Cuckoo Search via Lévy flights[C]// Nature & Biologically Inspired Computing, 2009. IEEE, 2010: 210-214.

[10] YANG H H, ZHANG X F, XIE P M, et al. Multiprocessor Task Scheduling Method Based on Cuckoo Search Algorithm[J]. Computer Science, 2015, 42(1): 86-89. (in Chinese)

杨辉华, 张晓凤, 谢谱模, 等. 基于布谷鸟搜索的多处理器任务调度算法[J]. 计算机科学, 2015, 42(1): 86-89.

- [11] NAWI N M, KHAN A, REHMAN M Z. A new back-propagation neural network optimized with cuckoo search algorithm[M]// Computational Science and Its Applications — ICCSA2013. Springer Berlin Heidelberg, 2013: 413-426.
- [12] VALIAN E, MOHANNA S, TAVAKOLI S. Improved cuckoo search algorithm for feedforward neural network training[J]. International Journal of Artificial Intelligence & Applications, 2011, 2(3): 36-43.
- (上接第 165 页)
- [3] ELOUEDI Z, MELLOULI K, SMETS P. Assessing Sensor Reliability for Multisensor Data Fusion Within the Transferable Belief Model [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2004, 34(4): 782-787.
- [4] GUO H W, SHI W K, DENG Y. Evaluating Sensor Reliability in Classification Problems Based on Evidence Theory [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2006, 36(5): 970-981.
- [5] SCHUBERT J. Conflict management in Dempster-Shafer theory using the degree of falsity [J]. International Journal of Approximate Reasoning, 2011, 52(3): 449-460.
- [6] JOUSSELME A L, GRENIER D, BOSSE E. A new distance between two bodies of evidence [J]. Information Fusion, 2001, 2(2): 91-101.
- [7] KLEIN J, COLOT O. Automatic discounting rate computation using a dissent criterion [C]// Proceedings of the Workshop on the Theory of Belief Functions, Brest, France, 2010: 1-6.
- [8] YANG Y, HAN D, HAN C. Discounted combination of unreliable evidence using degree of disagreement [J]. International Journal of Approximate Reasoning, 2013, 54(8): 1197-1216.
- [9] LIU Z G, DEZERT J, PAN Q, et al. Combination of sources of evidence with different discounting factors based on a new dissimilarity measure [J]. Decision Support Systems, 2011, 52(1): 133-141.
- [10] YAGER R R. An intuitionistic view of the Dempster-Shafer belief structure [J]. Soft Computing, 2014, 18(11): 2091-2099.
- [11] SONG Y F, WANG X D, LEI L, et al. An evidential view of similarity measure for Atanassov's intuitionistic fuzzy sets [J]. Journal of Intelligent & Fuzzy Systems, 2016, 31(6): 1653-1668.
- [12] SONG Y F, WANG X D, LEI L, et al. Combination of interval-valued belief structures based on intuitionistic fuzzy set [J]. Knowledge-Based Systems, 2014, 67: 61-70.
- [13] LI X N, SONG Y F, QUAN W. Evaluating evidence reliability based on intuitionistic fuzzy MCDM model [J]. Journal of Intelligent & Fuzzy Systems, 2016, 31(3): 1167-1182.
- [14] SMETS P. Data fusion in the transferable belief model [C]// Proceedings of the 3rd International Conference on Information Fusion, Paris, France, 2000: 21-33.
- [15] ZADEH L A. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination [J]. AI Magazine, 1986, 2(1): 85-90.
- [13] WANG H S, WANG Y N, WANG Y C. Cost Estimation of Plastic Injection Molding Parts through Integration of PSO and BP Neural Network [J]. Expert Systems with Applications, 2013, 40(8): 418-428.
- [14] NURY A H, HASAN K, ALAM M J B. Comparative study of wavelet-ARIMA and wavelet — ANN models on temperature time series data in northeastern Bangladesh [J]. Journal of King Saud University-Science, 2017, 29(1): 47-61.
- [15] NXP. KL25 Sub-Family Reference Manual [DB/OL]. [2014-10-12]. <http://www.nxp.com>.
- [16] FLOREA M C, JOUSSELME A L, BOSSE E. Robust combination rules for evidence theory [J]. Information Fusion, 2009, 10(2): 183-197.
- [17] SHAFER G. A Mathematical Theory of Evidence [M]. Princeton, NJ: Princeton University Press, 1976.
- [18] ZHU H, BASIR O. Extended Discounting Scheme for Evidential Reasoning as Applied to MS Lesion Detection [C]// 7th International Conference on Information Fusion, Stockholm, Sweden, 2004: 280-287.
- [19] ZADEH L A. Fuzzy sets [J]. Information and Control, 1965, 8(3): 338-353.
- [20] ATANASSOV K T. Intuitionistic fuzzy sets [J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.
- [21] BUSTINCE H, BURILLO P. Vague sets are intuitionistic fuzzy sets [J]. Fuzzy Sets and Systems, 1996, 79(3): 403-405.
- [22] HONG D H, KIM C. A note on similarity measures between vague sets and between elements [J]. Information Sciences, 1999, 115(1): 83-96.
- [23] LI J P, YANG Q B, YANG B. Dempster-Shafer theory is a special case of Vague sets theory [C]// Proceedings of the 2004 International Conference on Information Acquisition, Hefei, 2004: 50-53.
- [24] DYMOVA L, SEVASTJANOV P. An interpretation of intuitionistic fuzzy sets in terms of evidence theory: Decision making aspect [J]. Knowledge-Based Systems, 2010, 23(8): 772-782.
- [25] DYMOVA L, SEVASTJANOV P. The operations on intuitionistic fuzzy values in the framework of Dempster-Shafer theory [J]. Knowledge-Based Systems, 2012, 35(9): 132-143.
- [26] XING Q H, LIU F X. Method of determining membership and nonmembership function in intuitionistic fuzzy sets [J]. Control and Decision, 2009, 24(3): 393-397. (in Chinese)
邢清华, 刘付显. 直觉模糊集隶属度与非隶属度函数的确定方法 [J]. 控制与决策, 2009, 24(3): 393-397.
- [27] DENG Y, SHI W K, ZHU Z F, et al. Combining belief functions based on distance of evidence [J]. Decision Support Systems, 2004, 38(3): 489-493.
- [28] BACCOUR L, ALIMI A M, JOHN R I. Similarity measures for intuitionistic fuzzy sets: State of the art [J]. Journal of Intelligent & Fuzzy Systems, 2013, 24(1): 37-49.
- [29] SONG Y F, WANG X D, QUAN W, et al. A new approach to construct similarity measure for intuitionistic fuzzy sets [J/OL]. Soft Computing. <https://doi.org/10.1007/100500-017-2912-0>.