

DOI: 10.13382/j.jemi.B1902628

基于深度卷积神经网络的轴承故障诊断方法^{*}

唐 波^{1,2} 陈慎慎¹

(1.中国计量大学 计量测试工程学院 杭州 310018;
2.浙江大学 浙江省先进制造技术重点实验室 杭州 310058)

摘 要:针对传统浅层轴承故障诊断方法依赖于人工特征提取和诊断专业知识从而缺乏自适应性问题,结合卷积神经网络善于识别二维形状的特点,提出一种基于深度卷积神经网络的故障诊断方法(DCNN)。首先,为充分展现滚动轴承故障特征信息,利用短时傅里叶变换得到滚动轴承振动时间序列的二维时频谱;其次,通过卷积神经网络自适应提取时频谱中不同故障特征;最后,将提取的轴承故障特征利用 Softmax 分类器输出诊断结果,实现轴承故障诊断。通过实测故障轴承数据对该方法进行验证,结果表明 DCNN 在多故障、变负载的轴承故障诊断准确率高达 99.9%,证明了所提方法具有良好的泛化性能和可行性。

关键词:滚动轴承;故障诊断;卷积神经网络

中图分类号: TH212;TH213 **文献标识码:** A **国家标准学科分类代码:** 510.40

Method of bearing fault diagnosis based on deep convolutional neural network

Tang Bo^{1,2} Chen Shenshen¹

(1.College of Metrology and Measurement Engineering, China Jiliang University, Hangzhou 310018, China; 2. Zhejiang Province Key Laboratory of Advanced Manufacturing Technology, Zhejiang University, Hangzhou 310058, China)

Abstract: Aiming at the lack of adaptability as the traditional shallow bearing fault diagnosis method relying on artificial feature extraction and diagnosis expertise, a fault diagnosis method based on deep convolutional neural network is proposed to recognize two-dimensional shapes. Firstly, in order to fully display the fault characteristic information of rolling bearing, the two-dimensional time spectrum of rolling bearing vibration time series is obtained by using the short-time Fourier transform. Secondly, different fault features are extracted by convolutional neural network adaptively. Finally, Softmax classifier is used to output the diagnosis results to realize bearing fault diagnosis. The results show that the accuracy of the measured bearing fault diagnosis is up to 99.9%, proves that the proposed method has a good generalization performance and feasibility.

Keywords: rolling bearing; fault diagnosis; convolutional neural network

0 引 言

滚动轴承作为旋转机械一种重要的基础零部件,广泛的应用于高铁、风力发电机和航空发动机等机械装备中,轴承运行正常与否直接影响整个旋转机械精度、可靠性及使用寿命。但滚动轴承在服役期间受外界冲击、磨损、腐蚀等因素影响,易出现点蚀、剥落及裂纹等缺陷,导致密切联系的机械部件之间爆发连锁反应,造成整个设备安全事故甚至威胁人身性命。故对轴承进行状态监测分析和研究轴承故障诊断技术具有重要意义。

振动信号分析作为轴承故障诊断的主要方法,包括快速傅里叶变换(FFT)、小波变换(WT)^[1]、经验模态分解(EMD)^[2]及短时傅里叶变换^[3](short-time Fourier transform, STFT)等,通过信号处理方法人为提取时域、频域或时频域特征诊断轴承故障具有一定的准确率,但也存在着依赖人工经验的缺点。随后研究者提出基于知识的故障诊断方法,包括人工神经网络^[4]、支持向量机^[5](support vector machine, SVM)、聚类算法、遗传算法^[6]和模糊推断^[7]等可部分满足故障诊断识别的要求,然而机器设备工作时负载变化等因素影响模型的泛化能力。近年来,深度学习作为机器学习的一种主要算法,以应用于

收稿日期:2019-09-27 Received Date: 2019-09-27

^{*} 基金项目:国家自然科学基金(11872061)、国家市场监督管理总局科技计划(2019MK103)资助项目

计算机视觉、语音识别及自然语言处理等模式识别任务中^[8]。深度卷积神经网络 (deep convolutional neural networks, DCNN) 作为深度学习模型之一, 自适应提取数据基础特征, 实现数据特征识别。文献[8]将每个训练样本以某一尺度的“时间步”进行划分作为一维卷积神经网络模型的输入, 通过实验表明算法的有效性。文献[9]提取时域指标构成特征矩阵输入卷积神经网络 (CNN) 进行训练, 实现小样本下的齿轮故障智能定量诊断。文献[10]利用 s 变换处理轴承信号得到时频图像, 输入 CNN 实现轴承多工况的有效识别。上述研究使用人工提取的时域指标、原始振动信号及二维时频图输入 CNN 进行故障诊断, 而轴承振动信号一般混叠有强噪声信号, 使得振动信号不能够提取出有效的时域指标, 且 CNN 是为识别二维形状而设计的, 则一维卷积神经网络模型的识别性能受到不利影响。

为充分展现轴承故障时频信息及利用卷积神经网络自提取特征的优势, 本文提出基于 STFT 变换的 CNN 故障诊断方法。首先利用 STFT 变换对轴承振动信号进行时频分析; 然后使用主成分分析 (principal component analysis, PCA) 将高维时频谱转换到低维的特征空间; 最后通过超参数实验来构建 CNN, 以变负载工况及注入噪声实验条件下验证模型的鲁棒性, 并将其应用到实测故障轴承诊断中。

1 基于 STFT 和 CNN 的轴承故障诊断方法

针对“浅层学习”算法和基于一维卷积神经网络轴承故障诊断的局限, 本文提出基于 STFT 和 CNN 的轴承故障诊断方法, 利用 PCA 对轴承振动信号 STFT 后的时频特征降维, 构成 CNN 输入数据集, 并构建深度卷积神经网络, 实现故障特征自提取和轴承故障诊断结果, 则基于 STFT 和 CNN 的故障诊断方法流程如图 1 所示, 分为 4 个步骤, 振动信号采集与扩充、信号 STFT 分析与 PCA 降维、CNN 故障诊断模型设计与训练、轴承故障诊断。

1) 振动信号采集与扩充, 使用 n 点长度窗口截取轴承振动信号作为样本信号, 再窗口移动步长小于单个样本的信号长度得到下一个样本数据。

2) 样本信号 STFT 分析与 PCA 降维, 对样本信号进行 STFT 分析, 得到样本信号的时频谱, STFT 的运算公式为:

$$STFT_f(t, \omega) = \int_{-\infty}^{\infty} f(t) g^*(t - \tau) e^{-j\omega\tau} d\tau \quad (1)$$

式中: $f(t)$ 为时域信号; $*$ 表示复共轭; $g(t - \tau)$ 为中心位于 τ 时刻的时间窗口。

PCA 是一种数据压缩的方法, 通过少数几个主成分

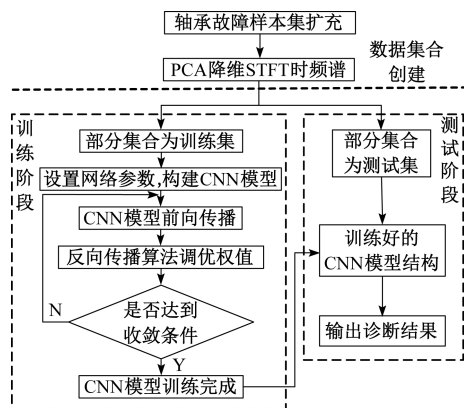


图1 STFT 和 CNN 的故障诊断流程

Fig.1 Fault diagnosis flow chart based on STFT-CNN

解释多变量的方差, 使其尽可能完整地保留原始变量的信息, 且彼此间不相关, 以达到简化数据的目的。使用 PCA 对样本信号时频谱进行降维, 创建数据集。

3) 卷积神经网络故障诊断模型设计与训练, 使用训练集进行 CNN 模型训练, 经调试相关卷积核大小, 卷积核数目和学习率等超参数, 获得性能较好的 CNN 模型。

卷积神经网络将卷积层和池化层相结合作为网络的基本结构, 在最后输出特征图后接上全连接层和分类器。CNN 中卷积层通过多个卷积核检测输入特征图所有位置上的特定特征, 实现同一输入特征图上的权值共享。卷积操作的一般形式为:

$$x_j^l = f(\sum_{i \in M_j} x_i^{l-1} w_{ij}^l + b_j^l) \quad (2)$$

式中: x_j^l 为第 l 层第 j 个元素; M_j 为 $l-1$ 层特征图的第 j 个卷积区域; x_i^{l-1} 为 $l-1$ 层的元素; w_{ij}^l 为对应卷积核的权重矩阵; b_j^l 为偏置项。 $f(\cdot)$ 为激活函数, 常用 ReLU 函数。

池化层对通过卷积层特征图的不同位置的特征进行聚合统计, 实现图像从高层次到低层次的特征表达, 减少了数据维度与网络参数。最大池是最常用的池类型, 得到输入的最大值。

$$x_j^l = \max_{(i-1)l+1 \leq i \leq il} x_i^{l-1}(t) \quad (3)$$

式中: l 为池化局部区域的长度。

4) 轴承故障诊断, 通过测试集和故障轴承实验验证基于 STFT 和 CNN 的轴承故障诊断模型的有效性。

2 实验结果与分析

2.1 数据集描述

本文使用美国西储大学 (Case Western Reserve University, CWRU) 的轴承数据进行分析, 轴承故障实验台如图 2 所示。CWRU 轴承实验装置包括电机、扭矩传

感器、功率计和电子控制设备,实验滚动轴承电蚀出单点故障,共外圈、内圈和滚动体 3 种故障,故障直径有 0.18、0.36 及 0.54 mm。故障轴承安装到测试电机中,采用不同的负载 0~3 HP (1 HP=746 W) 进行测试。

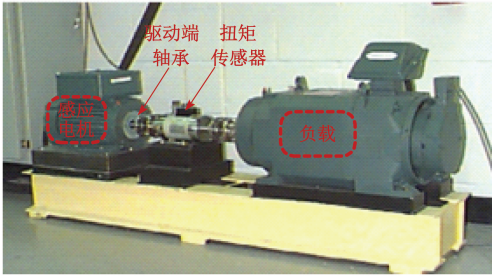


图 2 西楚大学轴承实验装置

Fig.2 Bearing test equipment of case western reserve university

CNN 模型训练与测试数据集 1 由 1 HP 下转速 ω 为 1 797 r/min 的 10 种不同故障类型的轴承数据组成,分别为正常状态、3 种不同损伤程度的滚动体故障、内圈故障和外圈故障,每类样本以 50% 的重叠率进行分割实现样本集扩容,如图 3 所示。为简化 CNN 网络结构参数,采用 PCA 对短时傅里叶变换后的时频矩阵进行降维,经计算累计贡献率,取前 8 个主成分即可保留原时频矩阵的 95% 的信息,为方便 CNN 输入和训练,将时频矩阵 reshape 为 64 阶矩阵。数据集 1 选择负载 1 HP 下每种故障类型 400 个样本,共计 4 000 样本,随机选择 80% 的数据进行模型训练,20% 进行模型测试。

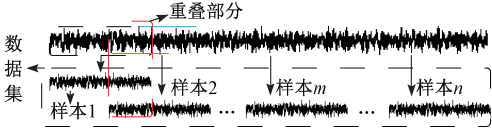


图 3 数据集增强技术

Fig.3 Data set enhancement techniques

2.2 卷积神经网络模型

构建 CNN 模型时,选择合适的模型参数可以在保证较高分类正确率的前提下加快模型的训练速度,CNN 模型通过试验确定的超参数有每个卷积层含有的卷积核数目及卷积核大小、优化器类型及学习率。

1) 卷积层参数

CNN 模型中的卷积层可改变的超参数为卷积层数目、卷积核的数目及大小。若增加卷积层的数目,会增加模型的训练时间和模型需要训练的参数,且本文训练样本数目与维数均较少的情况下,可能会引起过拟合问题。因此,在固定卷积层数目为 2 层的前提下,考察不同卷积核配置方式和训练时间对样本分类准确率的影响。CNN 为更好的得到图像中心信息和卷积时图像填充后的对

称,卷积核大小一般设为奇数,卷积核数目设置按照 16 的倍数倍增。将卷积层 C1 和 C2 的卷积层数目、卷积核大小作为变量,并定义取值范围如下:

{‘C1’:16,32,64,128,256;‘C2’:16,32,64,128,256;‘卷积核大小’:3×3,5×5,7×7;}

这些参数以遍历的方式组合,使用指标 $Score = Accuracy / (Loss \times Time)$ 评估 CNN 模型,其中 $Accuracy$ 为样本识别准确率, $Loss$ 为损失函数值, $Time$ 为 CNN 训练及识别时间。在识别正确率高、训练时间短和损失函数小的原则下,考虑 $Score$ 指标确定最优的网络结构,这里选取几条 $Score$ 值较大的曲线进行展示,如图 4 所示。图 4 中横坐标(16,32)为第 1 层卷积核数目为 16,第 2 层卷积核数目为 32;纵坐标(5×5,3×3)为第 1 层卷积核大小为 5×5,第 2 层卷积核大小为 3×3。可确定卷积核配置方式为(32,64)和(7×7,5×5)。

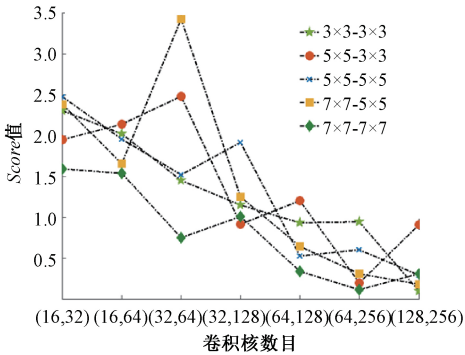


图 4 卷积核配置方式

Fig.4 Convolution kernel configuration

2) 优化器及学习率

CNN 模型选择合适优化器及学习率对提高模型训练速度和分类准确率具有重要作用。CNN 常用的优化器有 RMSprop、Adam 和 Sgdm,同时考虑不同学习率对于实际收敛速度的影响,以测试集的平均正确率作为评价指标,实验结果如图 5 所示。由图 5 可以看出,Sgdm 和 Adam 优化器的实验结果相近,在学习率为 0.001 时可以达到最佳的分类结果,通过对比各自优化器的最佳分类结果,采用 Sgdm 优化器,学习率设置为 0.001。

3) 鲁棒性分析

基于 STFT-CNN 算法确定的模型参数如表 1 所示。

在实际轴承工况中,其工作负载经常出现变化,导致模型的泛化性能受到影响。上述卷积神经网络模型训练测试均采用负载为 1 HP 的不同故障轴承及不同缺陷直径数据。现采用不同负载不同缺陷轴承数据及正常轴承数据共计 2 000 个作为训练测试数据集 2,如表 2 所示。模型参数设置与表 1 一致,在 1 HP 训练集下的变负载测试集准确率为 26.5%,说明 1 HP 训练集下 CNN

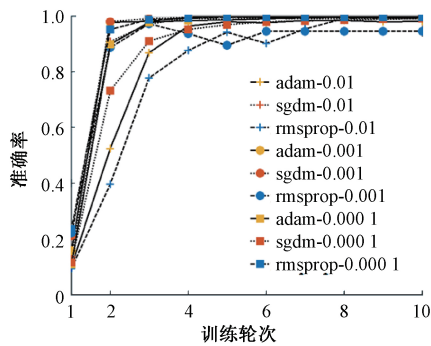


图 5 优化器及学习率调参

Fig.5 Optimizer and learning rate adjustment parameters

表 1 STFT-CNN 参数设置

Table 1 STFT-CNN parameter settings

参数名称	参数值
卷积层数	2
池化层数	2
卷积核大小	7×7,5×5
池化尺寸	2
优化器	Sgdm
优化器学习率	0.001
卷积层神经元数目	(32,64)

提取的特征不能排除负载变化的影响;在将变负载数据集 2 输入至 CNN 学习,得到对轴承变负载工况不敏感的故障特征,变负载测试集 CNN 分类准确率为 100%,可有效识别不同转速下轴承故障形式,可见轴承故障类型越全面则模型故障诊断效果及泛化性能越好。

表 2 泛化性测试集组成

Table 2 Generalization test set composition

负载/HP	内圈缺陷/mm	滚动体缺陷/mm	外圈缺陷/mm
1	0.18	0.18	0.18
2	0.36	0.36	0.36
3	0.54	0.54	0.54

4) 与其他算法对比

本文算法与目前主流智能故障诊断算法故障轴识别准确率对比结果表 3 所示。本文提出的 STFT-CNN 算法准确率为 100%,由于 DBN、SAE 和 RNN 等深度学习方法没有有效学习出在故障非线性特征,因此其识别率较本文算法有一定差距;而 SVM 由于具有浅层结构,限制了识别率的进一步提升。

2.3 滚动轴承故障诊断实例

以轴承型号为 N/NU 205EM 的故障轴承验证所提方

表 3 算法准确率对比

Table 3 Comparison of algorithm accuracy

算法	准确率/%	算法	准确率/%
CNN-STFT	100	DBN ^[11]	99.3
SAE ^[12]	95.72	RNN ^[13]	99.85
MSCNN ^[14]	99.59	AFDCNN ^[9]	99.4
CNN-DWT ^[15]	99.8	SVM ^[16]	0.996

法的有效性,实验采集轴承的加速度信号,采样频率为 80 kHz,采用运行转速分别为 900 r/min 的滚动体故障、1 200 r/min 的外圈故障及 1 500 r/min 的内圈故障数据,共计 4 800 个训练样本和 1 200 测试样本。试验结果如图 6 所示,故障轴承 CNN 诊断损失函数经 7 次达到收敛结果,最后识别率稳定在 99.92%。

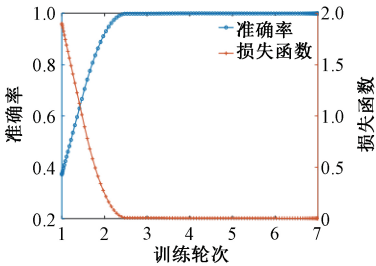


图 6 实测轴承故障诊断

Fig.6 Measured bearing fault diagnosis

为评估本模型提取特征的能力,利用 t 分布随机近邻嵌入 (t-distributed stochastic neighbor embedding, t-SNE) 流形学习方法将各层提取的特征降至 2 维进行特征可视化,如图 7 所示,图中同种颜色的散点表示同一损伤类型的样本,其中 1 为滚动体故障、2 为外圈故障、3 为内圈故障。可以看出 CNN 模型逐层提取时频故障非线性特征的能力,使得各类轴承故障明显可分。

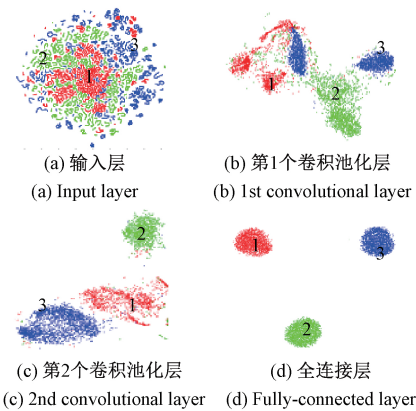


图 7 各层特征二维可视化

Fig.7 Visualization of different layers

3 结 论

本文提出一种基于 STFT 变换和 DCNN 网络的滚动轴承故障诊断方法,通过 CNN 模型对二维时频特征的自适应学习,确定网络参数实现轴承故障诊断,克服传统特征提取方法对专业知识和经验的依赖。通过理论分析、鲁棒性测试及实际应用,得到了如下结论:1)通过 CNN 自适应学习降维时频故障特征,避免“浅层学习”算法中故障特征提取环节,是一种智能化的诊断过程;2)CNN 模型通过逐层提取时频故障非线性特征,且通过多负载变化试验,证明该算法具有良好的鲁棒性和泛化性能;3)虽然该算法模型优于目前主流的智能化故障诊断算法,下一步应针对实际工况下的强轴承振动噪声信号进行训练学习,进一步提高 CNN 算法的泛化性能,凸显模型处理工业“大数据”的优势。

参考文献

- [1] 赵光权,姜泽东,胡聪. 基于小波包能量熵和 DBN 的轴承故障诊断 [J]. 电子测量与仪器学报, 2019, 33(2): 32-38.
ZHAO G Q, JIANG Z D, HU C. Bearing fault diagnosis based on wavelet packet energy entropy and DBN [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(2): 32-38.
- [2] 贾亚超,李国龙,何坤. 基于灰色准则与 EEMD 的滚刀振动信号降噪方法 [J]. 仪器仪表学报, 2019, 40(7): 187-194.
JIA Y C, LI G L, HE K. Denoising method for vibration signal of hob based on grey criterion and EEMD [J]. Chinese Journal of Scientific Instrument 2019, 40(7): 187-194.
- [3] KHODJA MEA, AIMER A F. Bearing fault diagnosis of a PWM inverter fed-induction motor using an improved short time fourier transform [J]. Journal of Electrical Engineering & Technology, 2019, 14(3): 1201-1210.
- [4] GUNERKAR R S, JALAN A K. Fault diagnosis of rolling element bearing based on artificial neural network [J]. Journal of Mechanical Science and Technology, 2019, 33(2): 505-511.
- [5] KUMAR T P, SAIMURUGAN M. A multi-sensor information fusion for fault diagnosis of a gearbox utilizing discrete wavelet features [J]. Measurement Science and Technology, 2019 30(8): 085101.
- [6] 王波,刘树林. 基于 QGA 优化广义 S 变换的滚动轴承故障特征提取 [J]. 振动与冲击, 2017, 36(5): 108-113,119.
WANG B, LIU SH L. Fault feature extraction of rolling bearing based on QGA generalized S transform [J]. Vibration and Shock, 2017, 36(05): 108-113,119.
- [7] 郑近德,潘海洋. 基于多变量多尺度模糊熵的行星齿轮箱故障诊断 [J]. 振动与冲击, 2019, 38(6): 187-193.
ZHENG J D, PAN H Y. Fault diagnosis of planetary gear box based on multivariable multiscale fuzzy entropy [J]. Vibration and Shock, 2019, 38(6): 187-193.
- [8] 曲建岭,余路. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法 [J]. 仪器仪表学报, 2018, 39(7): 134-143.
QU J L, YU L. An adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 134-143.
- [9] 罗鹏,杨宇. AFDCNN 方法及其在齿轮故障定量诊断中的应用 [J]. 电子测量与仪器学报, 2018, 32(1): 33-39.
LUO P, YANG Y. AFDCNN method and its application in quantitative diagnosis of gear fault [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(1): 33-39.
- [10] 杜小磊,陈志刚,张楠. 基于同步挤压 S 变换和深度学习的轴承故障诊断 [J]. 组合机床与自动化加工技术, 2019(5): 90-93,97.
DU X L, CHEN ZH G, ZHANG N. Faults diagnosis of bearings based on synchro squeezed S transform and deep learning [J]. Modular Machine Tool&Automatic Manufacturing Technique, 2019(5): 90-93,97.
- [11] MA S, CHU F. Ensemble deep learning-based fault diagnosis of rotor bearing systems [J]. Computers in Industry, 2019(105): 143-152.
- [12] XIANG Z, ZHANG X. Fault diagnosis of rolling bearing under fluctuating speed and variable load based on TCO Spectrum and Stacking Auto-encoder [J]. Measurement, 2019(138): 162-174.
- [13] LIU H, ZHOU J. Fault diagnosis of rolling bearings with recurrent neural network based autoencoders [J]. Isa Transactions, 2018(77): 167-178.
- [14] JIANG G, HE H. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox [J]. IEEE Transactions on Industrial Electronics, 2019, 66(4): 3196-3207.
- [15] 陈仁祥,黄鑫. 基于卷积神经网络和离散小波变换的滚动轴承故障诊断 [J]. 振动工程学报, 2018, 31(5): 883-891.

CHEN R X, HUANG X. Fault diagnosis of rolling bearing based on convolutional neural network and discrete wavelet transform [J]. Journal of Vibration Engineering, 2018, 31(5): 883-891.

- [16] ONG P, TIEH T H C. Efficient gear fault feature selection based on moth-flame optimisation in discrete wavelet packet analysis domain [J]. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2019, 41(6): 266.

作者简介



唐波, 2015 年于浙江大学获博士学位, 现为中国计量大学讲师, 主要研究方向为振动故障诊断。

E-mail: tang-bo001@163.com

Tang Bo received Ph. D. from Zhejiang University in 2015. He is currently a lecturer at China Jiliang University. His main research interest includes vibration fault diagnosis.