

融合多小波分解的深度卷积神经网络 轴承故障诊断方法

陶唐飞^{1,2}, 周文洁², 况佳臣², 徐光华²

(1. 西安交通大学现代设计及转子轴承系统教育部重点实验室, 710049, 西安;

2. 西安交通大学机械工程学院, 710049, 西安)

摘要: 针对卷积神经网络及其与信号降噪预处理集成方法面临高噪声环境和低质量数据挑战时难以有效地提取信号有用特征的问题,提出了一种融合 Geronimo-Hardin-Massopust 多小波分解的深度卷积神经网络模型(GHMMD-DCNN)。该模型思想是将多小波包分解与卷积神经网络深度融合,即设计多个一级多小波分解层以提取信号的低频分量和高频分量,再将多个一级多小波分解层与卷积层交替联接,使模型能够多尺度地提取并学习信号有用的时频域信息,信号分解和特征学习交替执行,进而实现强噪声鲁棒特征提取。在不同工况下的航空高速轴承振动数据上进行测试,结果表明:所提模型训练时能够快速达到稳定收敛,并且识别准确率均能达到 99.9% 以上;提出的方法在强噪声干扰下的故障辨识准确度和识别稳定性均优于对比方法,验证了其优秀的抗噪声干扰能力;在少训练样本测试中,提出的方法在单类训练样本数量为 60 时的平均诊断准确率达 91.19%,相比于其他方法最低提升了 13.19%,验证了 GHMMD-DCNN 模型具有更优的低样本泛化能力。

关键词: 多小波分解;卷积神经网络;深度学习;轴承故障诊断

中图分类号: TH17 **文献标志码:** A

DOI: 10.7652/xjtub202405004 **文章编号:** 0253-987X(2024)05-0031-11

Bearing Fault Diagnosis Method of Deep Convolutional Neural Network Based on Multiwavelet Decomposition

TAO Tangfei^{1,2}, ZHOU Wenjie², KUANG Jiachen², XU Guanghua²

(1. Key Laboratory of Education Ministry for Modern Design & Rotor-Bearing System, Xi'an Jiaotong University,

Xi'an 710049, China; 2. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: To tackle the challenge of convolutional neural network and its integration methods with denoising preprocessing methods struggling to effectively extract useful signal features amidst high noise environments and low-quality data, a deep convolutional neural network model based on the Geronimo-Hardin-Massopust multiwavelet decomposition (GHMMD-DCNN) is proposed. The model's concept revolves around deeply integrating the multiwavelet packet decomposition with the convolutional neural network. In other words, this involves the creation of multiple first-level multiwavelet decomposition layers to extract the low-frequency and high-frequency signal components, and these layers are lined alternately with the convolutional layer. This approach enables the model to extract and learn the

useful time-frequency information of the signal on a multiscale basis. The signal decomposition and the feature learning are executed alternately, and robust feature extraction is realized even under strong noise conditions. Tests are carried out using aerospace high-speed bearing vibration data under different working conditions. The results show that the proposed model is able to reach stable convergence quickly and the recognition accuracy surpasses 99.9%. The proposed method showcases superior fault recognition accuracy and stability in the presence of significant noise interference compared to contrast methods, which demonstrates its excellent anti-noise ability. In the test of fewer training samples, the proposed method achieves an impressive average diagnosis accuracy of 91.19% with only 60 training samples per class. This represents a 13.19% enhancement over alternative methods, verifying the GHMMD-DCNN's exceptional low-sample generalization ability.

Keywords: multiwavelet decomposition; convolutional neural network; deep learning; bearing fault diagnosis

2021年,国家工业和信息化部先后发布了《“十四五”大数据产业发展规划》和《“十四五”智能制造发展规划》,强化以数据驱动发展新兴产业并加快产业变革的战略导向,力争未来15年内基本实现规模以上制造业企业全面普及数字化^[1-2]。随着数字化产业的推进,机械装备逐渐朝着精密化、复杂化的方向发展,设备间的关联性愈加紧密,机械故障诊断技术成为稳步推进制造业发展的重要保障^[3-4]。滚动轴承作为机械设备关键部件之一,广泛应用于兵器装备、航空航天、风力发电等领域^[5]。统计表明,轴承故障约占达旋转类机械总故障的30%~40%^[6]。因此,对轴承进行及时准确地故障诊断能够有效保障现代化机械设备的安全运行。

目前,多数基于数据驱动的轴承故障诊断方法建立于分析振动信号的基础上,传统信号分析方法、浅层机器学习方法、深度学习方法是目前广为应用的故障诊断方法^[7]。传统信号分析方法基于时频域分析等方法进行特征提取,通过专家知识进行特征频率分析从而辨识故障,因此该方法依赖于操作人员的专业性。浅层机器学习方法通常采用核主成分分析^[8]、K-L变换^[9]等方法,对所提取的特征向量进行特征选择,再设计分类器进行分类,如人工神经网络^[10]、隐马尔可夫模型^[11]和支持向量机^[12]等。浅层机器学习方法在故障特征选择与辨识上减少了人工的参与,但仍然依赖人工经验优选特征向量。深度学习具有自动提取故障特征的能力^[13],避免了对专家知识的依赖,其强大的特征学习能力显著提高了故障识别准确率,是目前“大数据”时代国内外各领域广为关注的技术之一。

2012年,卷积神经网络(CNN)^[14]在图像识别展现出来的巨大优势迅速成为学术界研究的热点。该

网络结构主体一般由卷积层、激活层、池化层等组成,整个特征提取过程具有权值共享和空间不变性的特性^[15]。近年来,CNN因其强大的自适应特征提取能力在机械故障诊断领域中发挥出惊人的作用。Zhang等^[16]提出了基于宽卷积核的深度卷积神经网络(WDCNN)模型,在第一个卷积层引入了大卷积核,有效捕捉了轴承振动信号的中低频特征。康涛等^[17]将通道和时间注意力机制与多尺度CNN相融合,有力地提升了模型对轴承关键故障特征的提取能力。

滚动轴承在实际运行环境中常常受到机器系统和周围环境的噪声干扰,轴承发生故障时产生的瞬态故障脉冲成分通常淹没在这些强背景噪声中,导致CNN算法过度拟合无效特征而难以准确识别故障。针对该问题的常见解决方案有两种:①集成CNN模型与独立的信号预处理步骤;②将传统信号处理思想与CNN模型进行融合。前者通常独立进行信号降噪预处理,再将处理过的信号输送至CNN中进行特征提取,实现难度低。但是,若信号降噪不足,模型仍将受到噪声影响。反之,若信号降噪过度,有用信息也随之减少,模型将过度拟合无效特征。因此,CNN与信号降噪预处理集成方法存在难以合理保留有用信息而导致有效特征提取不足的问题。后者能避免这一缺陷,并能充分发挥出传统信号处理方法和CNN网络各自的优势。例如,深度残差收缩网络(DRSN-CW)^[18]融合了小波阈值去噪思想,添加了残差模块和自适应阈值门限,使得学习到的高级特征可以变得更具鉴别性。Li等^[19]将CNN第一层卷积核核函数更改为Laplace小波核函数,使CNN融合了连续小波变换的优势。Wang等^[20]受小波包分解机制启发,提出了一种融合多层小波注意力机制的卷积神经网络(MWA-CNN)框

架,其策略是将小波包分解与 CNN 深度融合,将多个单小波分解层嵌入 CNN 模型,使信号能够逐层分解和特征学习,有效地提升了 CNN 模型的抗噪声能力。然而,目前将传统信号处理方法和深度学习网络深度融合的方式并不普遍,这是因为二者在特征提取方面可能存在冲突,模型的开发难度较高。因此,这种融合方法目前仍处于探索阶段,但是该方法可以充分发挥两者的优势,对于提高信号处理和模式识别的性能仍具有很大的潜力。

受 MWA-CNN 启发,将有明确物理意义的传统信号处理方法与 CNN 深度融合,是提升 CNN 特征提取性能的有效手段。相比于单小波,多小波兼具对称性、正交性、短支撑性、高阶消失矩的独特优势^[21],且多小波变换是一个多输入-多输出系统,这为无额外数据增强条件下实现少标签样本及强背景噪声的轴承故障诊断提供了一种可能。

针对卷积神经网络易受噪声干扰、识别性能严重依赖于大规模和高质量数据的问题,以及 CNN 与信号降噪预处理集成方法存在难以合理地保留有用信息而导致有效特征提取不足的问题,本文基于机械故障诊断领域常用的 Geronimo-Hardin-Massopust 多小波,即 GHM 多小波,提出了一种融合 GHM 多小波分解的深度卷积神经网络(GHMMD-DCNN)轴承故障诊断模型。该模型将多个一级 GHM 多小波分解层与卷积层交替联接,使模型能够多尺度地学习信号有用的时频域信息。模型将在航空发动机高速轴承振动信号数据上进行实验测试,并与 WDCNN、集成经验模态分解方法的 WDCNN(WDCNN-EMD)、DRSN-CW 和以单小波为基础的 MWA-CNN 进行性能对比。

1 多小波变换理论

1.1 多小波理论

多小波理论的基本思想是,由多尺度函数生成多分辨率分析空间,在正交多分辨率分析空间中,其 r 重多尺度函数 $\Phi(t)$ 和 r 重多小波函数 $\Psi(t)$ 满足以下两尺度关系^[22]

$$\Phi(t) = \sqrt{2} \sum_{k=0}^l H_k \Phi(2t-k), \quad k \in \mathbf{Z} \quad (1)$$

$$\Psi(t) = \sqrt{2} \sum_{k=0}^l G_k \Phi(2t-k), \quad k \in \mathbf{Z} \quad (2)$$

式中: H_k 与 G_k 为 $r \times r$ 的低通滤波器系数矩阵和高通滤波器系数矩阵; l 为系数矩阵总数。

本文采用机械故障诊断领域中常用的 GHM 多

小波^[23],该小波为二重多小波,因此后文的分析均是在 $r=2$ 的基础上讨论的,将不再重复说明。与单小波相比,GHM 多小波在相同逼近阶下具备更短的支撑,可以保证信号特征提取过程中能量分布更加集中;此外,GHM 多小波具有两个不同频率性质的基函数,能够多尺度地提取信号特征的不同部分^[24]。GHM 多小波两尺度矩阵序列^[25]分别为

$$\begin{cases} H_0 = \begin{bmatrix} \frac{3\sqrt{2}}{10} & \frac{4}{5} \\ -\frac{1}{20} & -\frac{3\sqrt{2}}{20} \end{bmatrix}; H_1 = \begin{bmatrix} \frac{3\sqrt{2}}{10} & 0 \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ H_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{20} & -\frac{3\sqrt{2}}{20} \end{bmatrix}; H_3 = \begin{bmatrix} 0 & 0 \\ -\frac{1}{20} & 0 \end{bmatrix} \end{cases} \quad (3)$$

$$\begin{cases} G_0 = \begin{bmatrix} -\frac{1}{20} & -\frac{3\sqrt{2}}{20} \\ \frac{\sqrt{2}}{20} & \frac{3}{10} \end{bmatrix}; G_1 = \begin{bmatrix} \frac{9}{20} & -\frac{\sqrt{2}}{2} \\ -\frac{9\sqrt{2}}{20} & 0 \end{bmatrix} \\ G_2 = \begin{bmatrix} \frac{9}{20} & -\frac{3\sqrt{2}}{20} \\ \frac{9\sqrt{2}}{20} & -\frac{3}{10} \end{bmatrix}; G_3 = \begin{bmatrix} -\frac{1}{20} & 0 \\ -\frac{\sqrt{2}}{20} & 0 \end{bmatrix} \end{cases} \quad (4)$$

GHM 多小波的多尺度函数 ϕ_1 、 ϕ_2 和多小波函数 ψ_1 、 ψ_2 如图 1 所示。可以从图中明确看出, ϕ_1 的支撑区间为 $[0, 1]$, ϕ_2 、 ψ_1 和 ψ_2 的支撑区间为 $[0, 2]$, GHM 多小波的最大支撑区间为 2, 具有短支撑性。

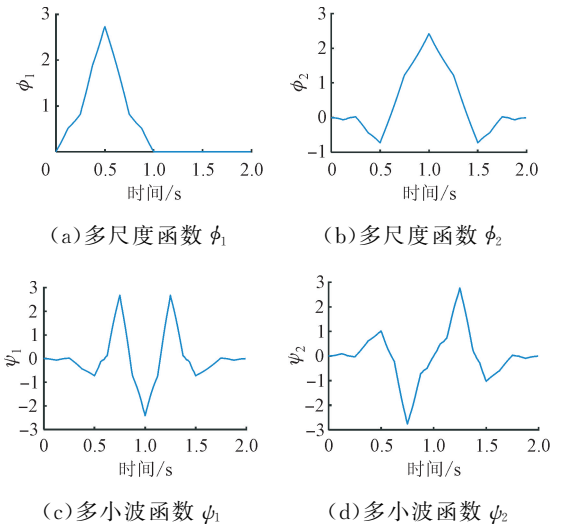


图 1 GHM 多尺度函数和多小波函数

Fig. 1 GHM multi-scale function and multi-wavelet functions

1.2 离散多小波分解

离散多小波分解算法与单小波的 Mallat 塔形算法相似,不同之处在于多小波为 r 重矢量输入。在进行 GHM 多小波分解前,需要将一维信号转为 2 维矢量输入。重复预处理的前处理方式在信号降噪中已被证明效果较好^[26],因此本文采用该处理方式对信号进行预处理,即

$$\mathbf{c}_{0,k} = \mathbf{Q} \begin{bmatrix} x_k \\ x_k \end{bmatrix} \quad (5)$$

式中: x_k 为一维信号输入; $\mathbf{c}_{0,k}$ 为前处理后的二维矢量输出; $\mathbf{Q}$ 为重复预处理矩阵,具体为

$$\mathbf{Q} = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} \quad (6)$$

离散 GHM 多小波分解过程如图 2 所示,其中 $\mathbf{H}$ 和 $\mathbf{G}$ 分别为低通滤波器和高通滤波器的矩阵表示, $2\downarrow$ 表示隔 2 抽样,离散多小波分解算法表达式^[27]如下

$$\mathbf{c}_{j-1,n} = \sum_k \mathbf{H}_{k-2n} \mathbf{c}_{j,k} \quad (7)$$

$$\mathbf{d}_{j-1,n} = \sum_k \mathbf{G}_{k-2n} \mathbf{c}_{j,k} \quad (8)$$

式中: $\mathbf{c}_{j-1,n}$ 为 2 维低频分量; $\mathbf{d}_{j-1,n}$ 为 2 维高频分量。

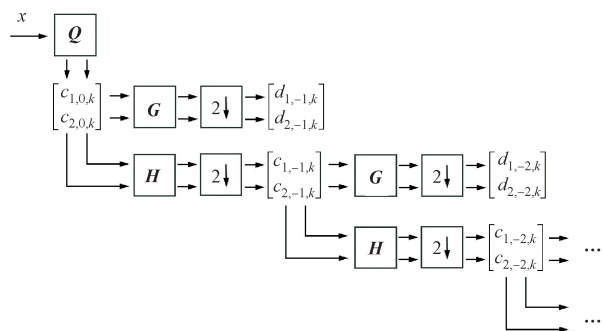


图 2 离散 GHM 多小波分解原理

Fig. 2 Principle of GHM multiwavelet decomposition

1.3 离散多小波包分解

离散多小波包分解同经典小波包分解一样,可以将细节分量进行再分解,进而提高细节分量的频率分辨率^[28]。轴承损伤所引发的振动冲击会激发系统固有振荡,在高频出现调制峰,因此对高频部分进行细分能够更好地诊断故障,引入小波包分解是有必要的。多小波包分解过程如图 3 所示,可以看出,多小波包分解对细节分量也按照树结构进行了分解,每次分解也是矩阵运算。

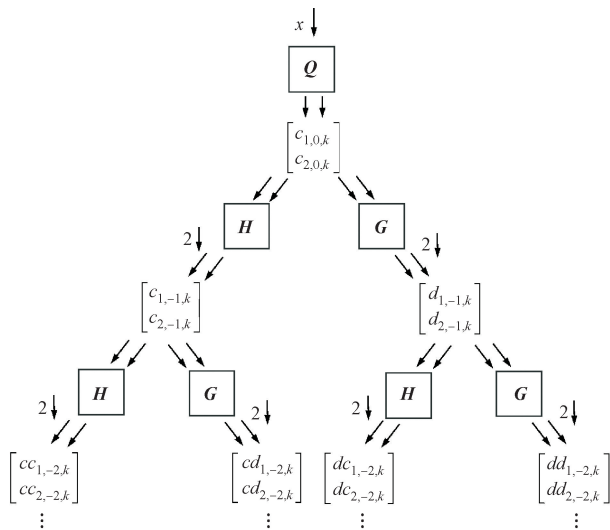


图 3 多小波包分解原理

Fig. 3 Principle of multiwavelet packet decomposition

2 GHMMD-DCNN 模型

本文结合 GHM 多小波包分解能够多尺度提取出更丰富的频率信息和 CNN 卷积采样自动提取局部特征的优势,将多个 GHM 多小波分解层与卷积层深度集成,构建了融合 GHM 多小波分解的深度卷积神经网络模型,整体架构如图 4 所示。整个网络由 1 个重复预处理层、5 个一级多小波分解模块、6 个卷积模块、1 个全局平均池化层和 1 个全连接层组成。

首先,将原始一维轴承振动信号按照固定样本长度分割,得到足够数量的一维样本序列,将归一化后的样本序列输送到预处理层中按式(5)和式(6)进行重复预处理,转换为二维矢量输入,每一维向量长度与初始长度一致。然后,将二维矢量输送到多小波分解模块,按式(7)和式(8)进行一级 GHM 多小波分解,得到低频分量和高频分量,将两分量拼接组合后输送到卷积模块中进行自动特征提取。卷积模块由卷积层、批归一化层(BN)、ReLU 激活层组成,卷积层提取时域和频域的重要特征,BN 层能够提升网络训练效率,ReLU 层将所有的负值设置为 0,网络的稀疏性得以提高,能够节省计算资源。由于多小波分解层中包含降采样操作,因此本模型不采用最大池化层进行特征降维。

卷积模块从各频段分量学习有用的故障特征信息,将所提取特征分量再输送到多小波分解层,进行一级 GHM 多小波分解,得到多个特征分量的多频段信息,再次将这些分量组合拼接输送至卷积模块。

一级多小波分解层与卷积层再循环 3 次交替执行,该循环执行完成后,使用卷积模块进行重复特征提取。信号的时域特征和频率特征能够被逐层分解和

学习,在深层信号分解和特征学习的过程中,模型能够更加有效地区分出有效信号与噪声信号的信息差异,从而学习到有价值的信息。

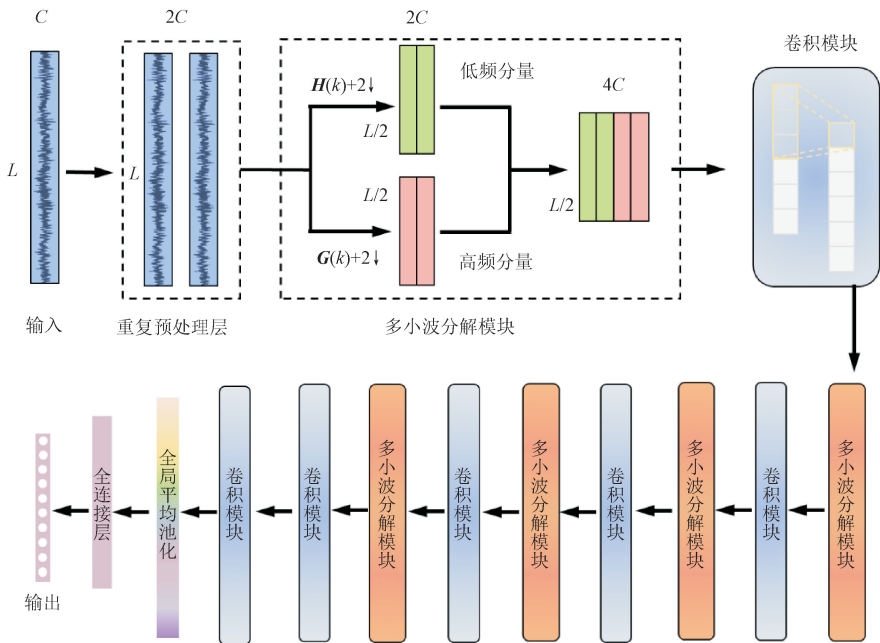


图 4 GHMMD-DCNN 模型的整体结构

Fig. 4 The overall structure of model GHMMD-DCNN

为了防止过拟合,在一级多小波分解层后引入随机失活层,随机失活率设为 0.2。同样地,为防止全连接输出层中权重参数过多而导致网络过拟合,将全局平均池化层置于最后一个卷积模块后,用于计算特征分量每个通道平均值,最后利用全连接层和 Softmax 函数量衡出每个类别节点对应的概率值。模型采用交叉熵损失函数衡量输出类别与真实类别的差距,公式为

$$\mathcal{L} = - \sum_{j=1}^J p(x_j) \log(q(x_j)) \tag{9}$$

式中: J 为轴承健康状态总类别数量; $q(x_j)$ 为样本属于第 j 类的预测概率; $p(x_j)$ 为样本属于第 j 类的真实概率。

在模型结构参数设置上,信号的输入长度为 4 096。多小波分解层中对输入信号进行周期延拓以保证所分解得到的信号长度为原始信号长度的一半,其输出通道为输入通道的 2 倍。本文模型具体参数如表 1 所示,其中输出维度“ $2 \times 4\,096$ ”表示通道数为 2,长度为 4 096,卷积核参数“ $4 \times 16 \times 3 \times 2$ ”为输入通道 $\times$ 输出通道 $\times$ 核大小 $\times$ 步长。

表 1 GHMMD-DCNN 模型结构参数		
Table 1 Structural parameters of GHMMD-DCNN model		
网络层	卷积核参数	输出维度
输入		$1 \times 4\,096$
重复预处理层		$2 \times 4\,096$
多小波分解层 1		$4 \times 2\,048$
卷积层 1	$4 \times 16 \times 3 \times 2$	$16 \times 1\,024$
多小波分解层 2		32×512
卷积层 2	$32 \times 32 \times 3 \times 1$	32×512
多小波分解层 3		64×256
卷积层 3	$64 \times 64 \times 3 \times 1$	64×256
多小波分解层 4		128×128
卷积层 4	$128 \times 128 \times 3 \times 1$	128×128
多小波分解层 5		256×64
卷积层 5	$256 \times 256 \times 3 \times 1$	256×64
卷积层 6	$256 \times 256 \times 3 \times 1$	256×62
全局平均池化层		256×1
全连接层		J

模型的训练流程如图 5 所示,采用具有自适应学习率特性的 Adam 优化器更新网络参数,初始学习率设定为 0.001,最大轮次设置为 100。为防止学习率过大导致模型无法稳定地收敛到最优解,采用固定步长学习率衰减,在第 40 轮、第 80 轮学习率以 0.1 的衰减率进行衰减。记录并保存 100 轮内验证效果最优的网络参数,最后在测试集上进行故障诊断测试,实现轴承故障辨识任务。

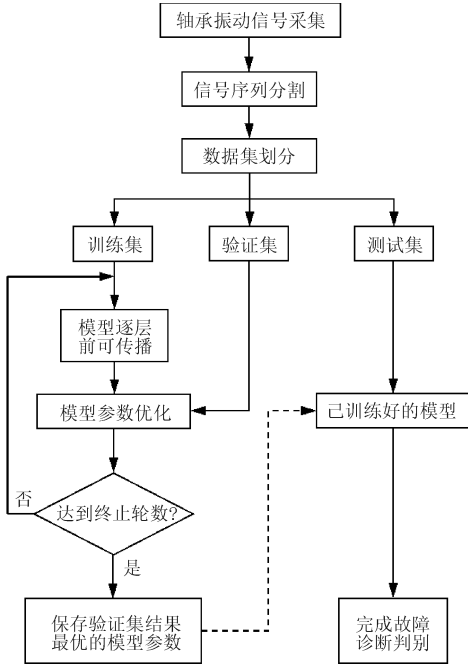


图 5 GHMMD-DCNN 模型的训练流程
Fig. 5 Training process of model GHMMD-DCNN

3 实验验证与分析

3.1 数据集准备

本文使用的轴承振动信号数据集来源于意大利都灵理工大学机械和航天工程系搭建的航空发动机高速轴承实验台^[29],其全局构造如图 6(a)所示,轴承与加速度传感器安装位置如图 6(b)所示。其中 B1、B2 和 B3 处分别安装 3 个同轴连接的航空发动机高速轴承,轴承如图 6(c)所示,轴承详细信息可参见文献[29],其中 B1 为被试轴承。图 6(b)中 A1、A2 处均采用加速度传感器分别记录 x 、 y 和 z 方向的输出,输出共计为 6 个通道的数据。6 个通道的数据均在采样频率 51 200 Hz 下收集,采样时长为 10 s。轴承存在无损伤、压痕直径为 450、250、150 μm 的内圈损伤、压痕直径为 450、250、150 μm 的滚动体损伤,共计 7 种状态,7 种状态依次记为

0~6。

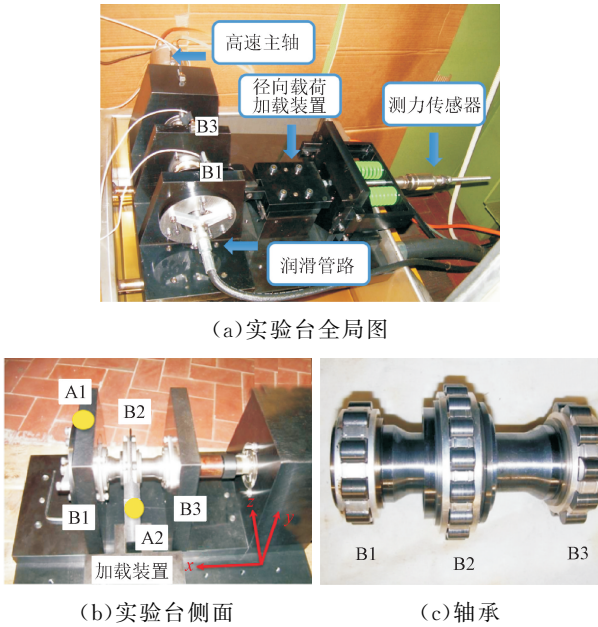


图 6 本文轴承实验台
Fig. 6 Bearing test rig

本文在 3 个工况下测试所提出模型的性能,为了简洁表达,分别标记为 WC1、WC2 和 WC3,具体信息如表 2 所示。实验采用 4 096 个数据点作为一段信号输入,并做最大最小值归一化处理。每一工况下每一类轴承故障振动信号有 $6 \times 512\,000$ 个样本点,以相邻两段信号的重叠率为 26% 的方式进行数据分割,可以准备 1 000 个样本。其中,训练、验证和测试样本数量分别为 500、200 和 300,即对于每一工况下的数据集,共计训练样本 3 500 个、验证样本 1 400 个和测试样本 2 100 个,测试、验证与训练样本之间无样本点重叠。

表 2 本文实验工况参数

Table 2 Parameters of working condition		
工作条件	主轴转速/Hz	径向载荷/N
WC1	100	1 000
WC2	100	1 400
WC3	200	1 400

3.2 GHMMD-DCNN 模型故障诊断结果分析

本文基于 Pytorch 框架搭建 GHMMD-DCNN 模型,计算机显卡为 NVIDIA GeForce RTX 2080 GPU。为降低深度网络训练学习随机性的影响,进行 5 次重复实验,3 个工况 WC1、WC2、WC3 下的准确率分别为 $(99.990\,5 \pm 0.0261)\%$ 、100%、100%,可以发现模型的稳定性很好,在不同工况下也能够

保证识别准确率达到 99.9% 以上。为进一步测试模型的训练性能,在工况 WC2 下观测模型在训练集和验证集上的损失函数与识别准确率随训练轮次变化的曲线,分别如图 7 和图 8 所示。可以看出,模型在训练集与验证集上的曲线变化高度一致,表明模型没有发生过拟合现象,而且模型的收敛速度很快,训练集约在第 11 轮训练时识别准确率达到 99%,约在第 41 轮训练后训练集和验证集的识别准确率均达到 100%。模型保持稳定收敛,表明所设计的 GHMMD-DCNN 模型是合理可行的。

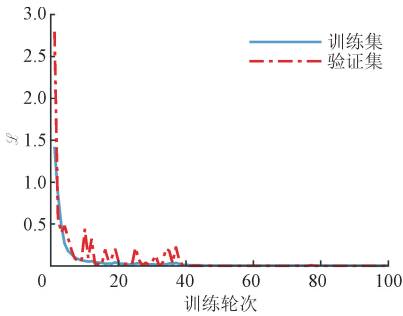


图 7 在 WC2 下的误差损失变化曲线
Fig. 7 Error loss variation curve under WC2

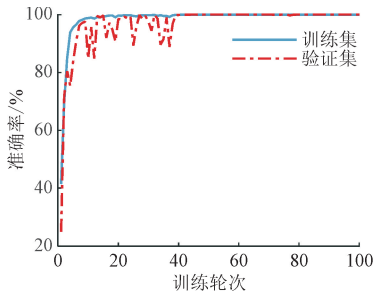


图 8 在 WC2 下的识别准确率变化曲线
Fig. 8 Accuracy variation curve under WC2

3.3 不同方法诊断结果对比分析

为更全面地测试 GHMMD-DCNN 模型的性能,与 WDCNN、WDCNN-EMD、DRSN-CW、MWA-CNN 方法进行对比测试。WDCNN 为普通的卷积神经网络,特别之处在于第一个卷积层选取大卷积核,此处卷积核大小设为 64;WDCNN-EMD 方法先将振动信号进行经验模态分解和重构,再将处理后的信号输送至 WDCNN 训练;DRSN-CW 为残差网络的改进模型,能逐层实现信号自适应降噪;本文所提出 GHMMD-DCNN 模型是在 MWA-CNN 的基础上改进的,区别在于 MWA-CNN 使用的是单小波,此处选用文献[20]中优选的“db16”小波。

图 9 为工况 WC3 下各方法的诊断结果,可明确看出,本文提出方法 GHMMD-DCNN 在识别准确性及诊断稳定性均表现为最优;MWA-CNN 和 DRSN-CW 的识别性能次之,准确率分别为 $(99.8952 \pm 0.1186)\%$ 和 $(99.6667 \pm 0.1347)\%$,诊断稳定性较高;WDCNN-EMD 方法的识别准确率最低,为 $(94.2381 \pm 1.105)\%$,实验结果波动较大,表明该方法的性能易受数据扰动。与 WDCNN 方法相比,WDCNN-EMD 的平均准确率降低了 2.0381%,表明卷积神经网络对经过 EMD 分解重构预处理后的信号特征提取效果并不如直接对原始信号进行特征提取,说明信号在预处理过程中虽去除了部分噪声信息,但也去除了部分关键故障信息,CNN 存在有效特征提取不足的情况。

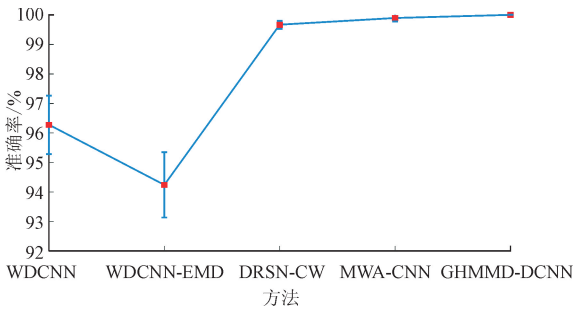


图 9 不同方法在 WC3 下的识别准确率对比
Fig. 9 Comparison of recognition accuracy of different methods under WC3

为进一步验证所提出方法在特征提取能力方面的优越性,使用 t 分布随机近邻嵌入(t -SNE)方法对各模型输出的特征降维,可视化结果如图 10 所示。在图 10(a)和 10(b)中可发现,各类特征类内聚合程度较好,但标签为 2、3、5 的三类故障特征在降维空间中存在重叠现象,分类边界混淆,即 WDCNN 和 WDCNN-EMD 难以完全区分压痕直径为 250、150 μm 的内圈损伤和压痕直径为 250 μm 的滚动体损伤。根据图 10(c),DRSN-CW 能够较好地地区分出 7 种轴承健康状态,但分类边界冗余复杂,部分类内部存在多个聚类簇,类内数据点聚合程度较差,表明模型没有学习到足够泛化的特征;MWA-CNN 的特征降维结果仍存在同样的问题,从图 10(d)可以看出,标签为 0、2、3、4 的故障特征类内聚合程度较差,即无损伤、压痕直径为 250、150 μm 的内圈损伤和压痕直径为 450 μm 的滚动体损伤的类内特征差异被过度学习,存在过拟合现象。根据图 10(e),GHMMD-DCNN 的 t -SNE 结果最优,各类的分类边界清晰分明,不同类别的数据点距离较大,类内数

据点聚合程度高,表明模型的具有更优的数据聚类和分离能力。对 MWA-CNN 和 DRSN-CW 的特征表征效果不佳问题进行分析,可能的原因是数据多样性不足或数据量不足,模型过度适应训练数据的细节特点,导致泛化性能较差。这表明所提出模型 GHMMD-DCNN 能够在少量带标签训练样本上训练出泛化性能较强的模型,从而高效率且高准确率地实现轴承故障诊断。

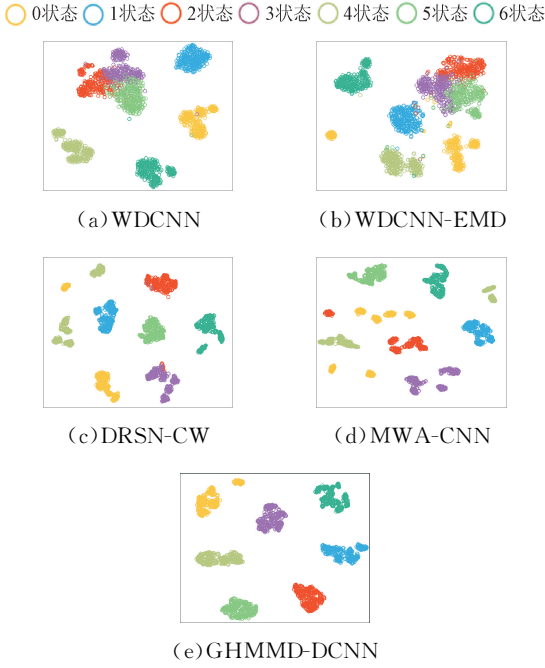


图 10 不同方法在 WC3 下的特征提取可视化结果对比
Fig. 10 Comparison of feature extraction visualization results of different methods under WC3

为了进行噪声鲁棒性故障诊断测试,对原始振动数据添加不同强弱的高斯白噪声,采用信噪比(R)来评价噪声强弱,公式为

$$R = 10\lg\left(\frac{P_s}{P_n}\right) \quad (10)$$

式中: P_s 和 P_n 分别表示原始振动信号与额外添加高斯白噪声信号的能量。

本次实验将在信噪比分别为 -4 、 -2 、 0 、 2 、 4 dB 的 5 种噪声条件下进行故障识别测试,每次测试仍然进行 5 次重复实验。图 11 和图 12 分别为工况 WC1、WC2 下的准确率结果,可以看出,WDCNN 和 WDCNN-EMD 的性能严重受噪声干扰,经过经验模态分解预处理的信号并没有起到提高模型性能的作用,甚至降低了模型的分类准确度;而 GHMMD-DCNN 在两个工况的不同噪声条件下的诊断结果均表现出色。

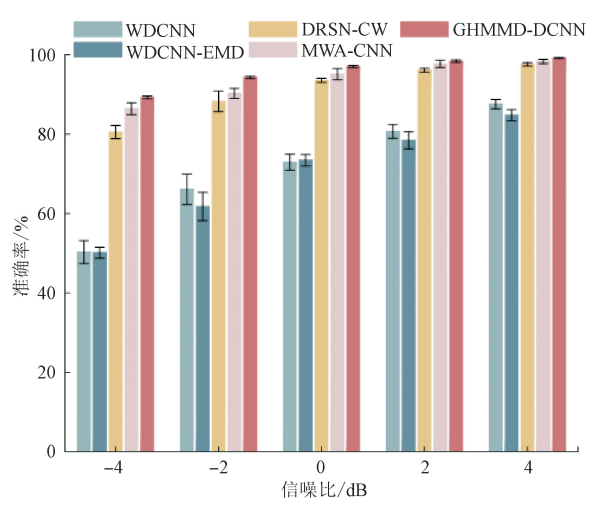


图 11 不同方法在 WC1 下的抗噪声能力对比
Fig. 11 Comparison of anti-noise ability of different methods under WC1

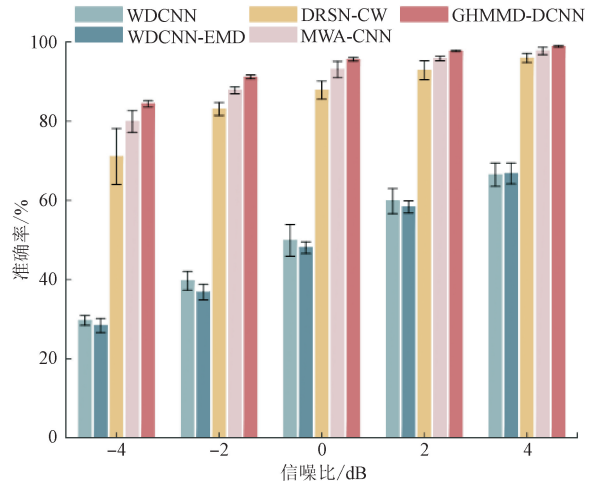


图 12 不同方法在 WC2 下的抗噪声能力对比
Fig. 12 Comparison of anti-noise ability of different methods under WC2

在 WC1 下, WDCNN、WDCNN-EMD、DRSN-CW、MWA-CNN 和 GHMMD-DCNN 方法在 $R=4$ dB 下的准确率分别为 $(87.581 \pm 1.1361)\%$ 、 $(84.762 \pm 1.4078)\%$ 、 $(97.619 \pm 0.4364)\%$ 、 $(98.3524 \pm 0.5245)\%$ 、 $(99.21 \pm 0.1147)\%$ 。在 $R=-4$ dB 下, 5 种方法的准确率分别为 $(50.305 \pm 2.873)\%$ 、 $(50.1524 \pm 1.374)\%$ 、 $(80.5524 \pm 1.6614)\%$ 、 $(86.4095 \pm 1.4641)\%$ 、 $(89.2762 \pm 0.3112)\%$, 平均识别准确率相比于 $R=4$ dB 下的诊断结果分别下降了 37.276% 、 34.6095% 、 17.0667% 、 11.9429% 和 9.9333% 。所提出 GHMMD-DCNN 模型的识别准确率显著高于 WDCNN、WDCNN-EMD 和 DRSN-CW, 比 WDCNN 的准确率提升了 38.9712% , 而相

比于 MWA-CNN 方法,GHMMD-DCNN 的准确率约提升 2.866 7%。对比该工况下 WDCNN、WDCNN-EMD 结果可发现,在这 5 种不同强弱的噪声环境中,WDCNN-EMD 的识别准确率均不如 WDCNN,表明在强噪声背景下使用信号降噪预处理方法容易去除有用的故障信息。

相比于 WC1,WC2 为负载高转速较低的工作环境,因此故障特征更加微弱隐蔽,对比图 11 和图 12,可以发现各方法的识别准确率均有下降。两个普通的 CNN 方法的平均分类准确率在 $R=-4$ dB 时仅为 29.666 7%和 28.4%,严重受噪声干扰,显然模型过拟合噪声细节,很难提取出有用的故障信息。根据图 12 可知,GHMMD-DCNN 模型的识别准确性和识别稳定性显著优于 WDCNN、WDCNN-EMD 和 DRSN-CW;与 MWA-CNN 相比,GHMMD-DCNN 模型在信噪比为 -4 、 -2 、 0 、 2 、 4 dB 情况下的平均准确率分别提升了 4.485 7%、3.4%、2.552 4%、1.876 2%和 1.333 3%,且 GHMMD-DCNN 的稳定性仍优于

MWA-CNN,表明模型具有较优的噪声鲁棒性。

为进一步考察 GHMMD-DCNN 模型的泛化能力,在工况 WC1 下对每类带标签训练样本数量分别为 30、60、90、120 的数据进行测试,每类测试集样本数量为 300,对比方法仍然为前文所述的 4 种方法,实验结果如表 3 所示。可以看出,GHMMD-DCNN 模型的识别准确率显著高于其他 4 类方法,在单类训练样本数量为 30 时平均诊断准确率比 WDCNN、WDCNN-EMD、DRSN-CW、MWA-CNN 分别提升了 29.31%、29.44%、17.2%、8.57%;所提方法在单类训练样本数量为 60 时的平均识别准确率达 91.19%,比其他 4 类方法分别提升了 28.84%、30.74%、14.84%、13.19%。GHMMD-DCNN 模型为 MWA-CNN 的改进模型,相比于 MWA-CNN,本文模型在单类训练样本数量为 90 和 120 条件下的识别准确率仍然有所提升,分别提升了 7.96%和 7.86%。以上结果表明,本文模型具有更优的低样本泛化能力。

表 3 较少训练样本数时的测试结果对比
Table 3 Comparison of test results with fewer training samples

单类带标签 训练样本数	识别准确率/%				
	WDCNN	WDCNN-EMD	DRSN-CW	MWA-CNN	GHMMD-DCNN
30	49.533 3±1.866 4	49.409 5±1.322 5	61.647 6±3.316 4	70.276 2±1.981 3	78.847 6±1.968 7
60	62.352 4±3.074 8	60.447 6±1.171 6	76.352 4±2.922 4	78.000 0±3.195 1	91.190 5±1.578 6
90	68.714 3±2.189 7	66.400 0±2.195 3	88.476 2±3.476 3	88.438 1±2.898 6	96.400 0±0.486 3
120	72.190 5±1.328 2	72.628 5±1.311 6	92.085 7±1.279 3	89.790 5±0.999 2	97.647 6±0.410 5

4 结 论

本文从信号处理方法与卷积神经网络深度融合的角度出发,提出了一种能够多尺度地提取并学习信号有用信息的 GHMMD-DCNN 模型,在航空发动机高速轴承上进行了故障诊断测试,得到了以下结论。

(1)所提出模型 GHMMD-DCNN 交替执行信号分解和特征学习两过程,深层次并多尺度地从多个频带分量中学习故障特征信息与噪声细节的差别,进而提高模型的抗噪声能力。实验结果表明,提出的模型能够快速收敛达到稳定,在无额外噪声添加的情况下识别准确率均能达到 99.9%,所提取特征的 t -SNE 结果中,类间边界分明、类内数据点聚合程度高,表明所提出模型具有良好的泛化能力。

(2)抗噪性能测试结果表明,GHMMD-DCNN

模型的识别准确性和识别稳定性均显著优于普通 CNN 方法和集成经验模态分解信号预处理的 CNN 方法,且该模型在各个噪声条件下的诊断结果均优于 MWA-CNN,表明将多小波分解方法与卷积神经网络深度融合的 GHMMD-DCNN 模型对于抗噪声干扰是有效的。

(3)低训练样本测试结果表明,所提出方法显著优于其他对比方法。与原基础模型 MWA-CNN 相比,所提出的模型在单类训练样本数量为 30、60、90 和 120 的平均识别准确率分别提升了 8.57%、13.19%、7.96%和 7.86%,表明该模型在训练样本量少的条件下仍能学习到数据的真实分布和特征,模型存在更优的低样本泛化能力。

参考文献:

[1] 赵碧.“十四五”大数据产业发展规划发布 [J]. 中国

- 产经新闻, 2021-12-02(2).
- [2] 董凯. “十四五”智能制造发展规划解读及趋势研判[J]. 中国工业和信息化, 2022(1): 24-29.
DONG Kai. Interpretation and trend analysis of the 14th Five Year Plan for the development of intelligent manufacturing China industry and information technology [J]. China Industry & Information Technology, 2022(1): 24-29.
 - [3] 陈雪峰, 菅艳阳. 智能运维与健康管理的[M]. 北京: 机械工业出版社, 2018.
 - [4] 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战[J]. 机械工程学报, 2018, 54(5): 94-104.
LEI Yaguo, JIA Feng, KONG Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era [J]. Journal of Mechanical Engineering, 2018, 54(5): 94-104.
 - [5] 何正嘉, 陈进, 王太勇, 等. 机械故障诊断理论及应用[M]. 北京: 高等教育出版社, 2010.
 - [6] 时献江, 王桂荣, 司俊山. 机械故障诊断及典型案例解析[M]. 北京: 化学工业出版社, 2013.
 - [7] LEI Yaguo, YANG Bin, JIANG Xinwei, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap [J]. Mechanical Systems and Signal Processing, 2020, 138: 106587.
 - [8] LIU Xiaozhi, XIE Jie, LUO Yanhong, et al. A novel power transformer fault diagnosis method based on data augmentation for KPCA and deep residual network [J]. Energy Reports, 2023, 9(8): 620-627.
 - [9] 徐统, 王红军, 宋智勇, 等. 基于K-L散度的VMD瞬时能量与PNN的滚动轴承故障诊断[J]. 电子测量与仪器学报, 2019, 33(8): 117-123.
XU Tong, WANG Hongjun, SONG Zhiyong, et al. Rolling bearing fault diagnosis using VMD energy feature and PNN based on Kullback-Leibler divergence [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(8): 117-123.
 - [10] UNAL M, ONAT M, DEMETGUL M, et al. Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network [J]. Measurement, 2014, 58: 187-196.
 - [11] LÓPEZ C, NARANJO Á, LU Siliang, et al. Hidden Markov model based stochastic resonance and its application to bearing fault diagnosis [J]. Journal of Sound and Vibration, 2022, 528: 116890.
 - [12] 张鑫, 赵建民, 李海平, 等. 基于NIC-DWT-WOAS-VM的齿轮箱混合故障诊断[J]. 振动与冲击, 2020, 39(11): 146-151, 164.
ZHANG Xin, ZHAO Jianmin, LI Haiping, et al. Compound fault diagnosis for gearbox based on NIC-DWT-WOASVM [J]. Journal of Vibration and Shock, 2020, 39(11): 146-151, 164.
 - [13] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.
 - [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
 - [15] MCAULEY J J, RAMISA A, CAETANO T S. Optimization of robust loss functions for weakly-labeled image taxonomies [J]. International Journal of Computer Vision, 2013, 104(3): 343-361.
 - [16] ZHANG Wei, PENG Gaoliang, LI Chuanhao, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals [J]. Sensors, 2017, 17(2): 425.
 - [17] 康涛, 段蓉凯, 杨磊, 等. 融合多注意力机制的卷积神经网络轴承故障诊断方法[J]. 西安交通大学学报, 2022, 56(12): 68-77.
KANG Tao, DUAN Rongkai, YANG Lei, et al. Bearing fault diagnosis using convolutional neural network based on a multi-attention mechanism [J]. Journal of Xi'an Jiaotong University, 2022, 56(12): 68-77.
 - [18] ZHAO Minghang, ZHONG Shisheng, FU Xuyun, et al. Deep residual shrinkage networks for fault diagnosis [J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4681-4690.
 - [19] LI Tianfu, ZHAO Zhibin, SUN Chuang, et al. Wavelet KernelNet: an interpretable deep neural network for industrial intelligent diagnosis [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(4): 2302-2312.
 - [20] WANG Huan, LIU Zhiliang, PENG Dandan, et al. Interpretable convolutional neural network with multi-layer wavelet for noise-robust machinery fault diagnosis [J]. Mechanical Systems and Signal Processing, 2023, 195: 110314.
 - [21] 刘志刚, 钱清泉. 自适应阈值多小波故障暂态信号去噪方法[J]. 系统工程与电子技术, 2004, 26(7): 878-880.
LIU Zhigang, QIAN Qingquan. Adaptive shrinkage value de-noising method of fault transient signals with multiwavelets [J]. Systems Engineering and Electronics, 2004, 26(7): 878-880.
 - [22] 袁静, 何正嘉, 王晓东, 等. 平移不变多小波相邻系

数降噪方法及其在监测诊断中的应用 [J]. 机械工程学报, 2009, 45(4): 155-160.

YUAN Jing, HE Zhengjia, WANG Xiaodong, et al. Translation-invariant multiwavelets denoising using neighboring coefficients and its application to monitoring and diagnosis [J]. Journal of Mechanical Engineering, 2009, 45(4): 155-160.

[23] 洪连环. 面向航空发动机转子系统故障识别的多小波方法研究 [D]. 南京: 南京航空航天大学, 2020.

[24] 陈景龙. 自适应多小波的时频域优化构造与机械故障特征提取方法 [D]. 西安: 西安交通大学, 2014.

[25] GERONIMO J S, HARDIN D P, MASSOPUST P R. Fractal functions and wavelet expansions based on several scaling functions [J]. Journal of Approximation Theory, 1994, 78(3): 373-401.

[26] BUI T D, CHEN Guangyi. Translation-invariant denoising using multiwavelets [J]. IEEE Transactions on Signal Processing, 1998, 46(12): 3414-3420.

[27] SUN Hailiang, HE Zhengjia, ZI Yanyang, et al. Multiwavelet transform and its applications in mechanical fault diagnosis: a review [J]. Mechanical Systems and Signal Processing, 2014, 43(1/2): 1-24.

[28] 李东敏. 基于多小波包和人工神经网络的电力系统故障类型识别研究 [D]. 成都: 西南交通大学, 2008.

[29] DAGA A P, FASANA A, MARCHESIELLO S, et al. The Politecnico di Torino rolling bearing test rig: description and analysis of open access data [J]. Mechanical Systems and Signal Processing, 2019, 120: 252-273.

(编辑 杜秀杰)