

# 基于多小波包样本熵的轴承损伤程度识别方法<sup>\*</sup>

张建宇<sup>1</sup>, 张随征<sup>1</sup>, 管磊<sup>2</sup>, 杨洋<sup>1</sup>

(1. 北京工业大学北京市先进制造技术重点实验室 北京, 100124)

(2. 江苏迈安德食品机械有限公司 扬州, 225127)

**摘要** 为了自动辨识不同尺度下的轴承故障,建立了一种基于多小波包系数样本熵和 BP 神经网络的模式判别方法。针对 5 种尺度下的轴承外圈故障信号,分别采用 GHM 多小波包完成三层分解。为了充分利用多小波包的分析优势,将分解后的 16 个频段信号分别求系数样本熵,并将其作为神经网络的输入向量。通过三层 BP 神经网络的训练、学习,并与 dB10 小波包神经网络做了对比研究。结果表明,多小波包样本熵可以区别不同损伤程度的故障信号,且多小波包样本熵与神经网络结合,其辨识精度更高,分类效果明显优于传统单小波,便于轴承损伤程度的自动识别。

**关键词** 故障程度;多小波包;样本熵;BP 神经网络;自动识别

**中图分类号** TP306+.3; TH17

## 引言

轴承的运行状态直接影响设备的工作性能,有效识别滚动轴承的损伤程度对于延长设备使用寿命、实现真正的预知维修具有重要意义。目前,大部分故障诊断的定量分析还处于起步阶段。姚红良等<sup>[1]</sup>提出一种利用碰摩转子系统响应中的高次谐波分量进行碰摩位置和碰摩力定量诊断的方法。朱忠奎<sup>[2]</sup>在基于支持向量回归的轴承故障定量诊断方法的研究中,提出将支持向量分类机和支持向量回归机用于故障特征分类和回归描述,为故障诊断模型的建立提供了理论基础。文献[3]针对不同损伤程度的滚动轴承内外圈故障,提出了一种基于经验模式分解(empirical mode decomposition,简称 EMD)和 Lempel-Ziv 指标的评估方法,给出了评估故障程度的 Lempel-Ziv 指标取值区间。文献[4]研究了基于小波分析的旋转机械振动信号定量特征,取得了较好效果。由于设备的复杂性和故障表现的多样性,一般的衡量指标很难有较广泛的适用性,这是定量诊断的难点所在。当轴承存在故障时,表现出非线性、非平稳特征,传统的基于线性系统的特征提取方法不能有效提取故障的特征。近年来,非线性方法引起了人们广泛的关注。文献[5]引入近似

熵概念并将其用于轴承状态监测。Pincus<sup>[6]</sup>提出的样本熵是近似熵的改进算法,它的优越性在于可以较少地依赖时间序列长度,已广泛用于振动信号处理<sup>[7-8]</sup>。苏文胜等<sup>[9]</sup>将样本熵引入故障诊断领域,讨论了样本熵的性能和计算参数的选择,结合小波包分解和样本熵,提出了一种滚动轴承故障特征的提取方法。多小波是在小波分析的基础上发展起来的,能同时满足对称性、正交性、紧支撑性和高阶消失矩等特性。文献[10]通过自适应方法构造多小波,对滚动轴承的复合故障进行分离,取得了较好效果。

笔者针对轴承故障信号的特点,充分利用多小波包能够对信号精细分解的性质,计算出分解后的样本熵。基于 BP 神经网络建立了轴承故障程度识别网络模型,实现对轴承故障的定量识别,为维修提供决策依据;并与 dB10 小波包神经网络模型进行比对,验证多小波包神经网络的准确性。

## 1 基于多小波包分解能量比的神经网络模型

### 1.1 多小波包分解的基本原理

多小波同单小波一样,也是基于多分辨率分析(multiple resolution analysis,简称 MRA)小波,区

<sup>\*</sup> 北京市教委科技计划资助项目(KM201410005027)  
收稿日期:2013-01-17;修回日期:2013-03-18

别在于多小波有多个尺度函数和小波函数,即  $\phi(x)$  由  $r$  个尺度函数的平移  $\phi_0(t-k), \phi_1(t-k), \dots, \phi_{r-1}(t-k)$  生成。于是,  $r \times r$  的矩阵  $\mathbf{H}_k$  和  $\mathbf{G}_k$  满足矩阵的尺度方程<sup>[11]</sup>为

$$\phi(x) = \sum_{k \in Z} \mathbf{H}_k \phi(2x - k) \quad (1)$$

$$\psi(x) = \sum_{k \in Z} \mathbf{G}_k \psi(2x - k) \quad (2)$$

其中:  $\phi(x)$  为尺度函数;  $\psi(x)$  为  $\phi(x)$  对应的小波函数。

根据多小波的多分辨分析,得到多小波的分解和重构公式为

$$\begin{cases} s_{j-1,n} = \sum_k \mathbf{H}_{k-2n} s_{j,k} \\ d_{j-1,n} = \sum_k \mathbf{G}_{k-2n} s_{j,k} \\ s_{j,k} = \sum_n \mathbf{H}_{k-2n}^* s_{j-1,n} + \sum_n \mathbf{G}_{k-2n}^* d_{j-1,n} \end{cases} \quad (3)$$

其中:  $s_{j,k}$  代表  $r$  维低频分量;  $d_{j,k}$  代表  $r$  维高频分量。

令  $\mathbf{U}_0(t) = \Phi(t), \mathbf{U}_1(t) = \Psi(t)$ , 将尺度子空间  $V_j$  和小波子空间  $W_j$  用一个新的  $\mathbf{U}_j^n$  统一起来表征, 定义<sup>[8]</sup>为

$$\mathbf{U}_{j-1,2n} = \sum_{k=-\infty}^{\infty} \mathbf{H}_k \mathbf{U}_{j,n}(2t - k) \quad (4)$$

$$\mathbf{U}_{j-1,2n+1} = \sum_{k=-\infty}^{\infty} \mathbf{G}_k \mathbf{U}_{j,n}(2t - k) \quad (5)$$

根据多小波包的空间分解得到多小波包重构公式为

$$\mathbf{U}_{j,n} = \sum_k (\mathbf{H}_k^* \mathbf{U}_{j-1,2n} + \mathbf{G}_k^* \mathbf{U}_{j-1,2n+1}) \quad (6)$$

其中:  $\mathbf{H}_k^*$  和  $\mathbf{G}_k^*$  分别为  $\mathbf{H}_k$  和  $\mathbf{G}_k$  的共轭转置矩阵。

$$\phi(x) = \sum_{k \in Z} \mathbf{H}_k \phi(2x - k) \quad (7)$$

$$\psi(x) = \sum_{k \in Z} \mathbf{G}_k \psi(2x - k) \quad (8)$$

根据多小波的多分辨分析,得到多小波包的分解和重构公式为

$$\begin{cases} s_{j-1,n} = \sum_k \mathbf{H}_{k-2n} s_{j,k} \\ d_{j-1,n} = \sum_k \mathbf{G}_{k-2n} s_{j,k} \\ s_{j,k} = \sum_n \mathbf{H}_{k-2n}^* s_{j-1,n} + \sum_n \mathbf{G}_{k-2n}^* d_{j-1,n} \end{cases} \quad (9)$$

其中:  $s_{j-1,n}$  代表  $r$  维低频分量;  $d_{j-1,n}$  代表  $r$  维高频分量。

多小波包分解优于小波包的主要原因是由于多小波包分解的各个系数均为多维矩阵。笔者利用维

数  $r=2$  的 GHM 多小波<sup>[12]</sup>对信号进行多小波包分解,预处理方法选择重复过采样,并与 dB 10 小波包分解结果进行对比。多小波包的 3 层分解示意图如图 1 所示。图中:  $\mathbf{S}_1$  和  $\mathbf{S}_2$  为原始信号经过预处理后的信号矩阵;  $\mathbf{L}_1$  和  $\mathbf{L}_2$  为  $\mathbf{S}_1$  和  $\mathbf{S}_2$  多小波包 1 层分解后的低频部分,  $\mathbf{H}_1$  和  $\mathbf{H}_2$  为高频部分。分别对 1 层分解后的低、高频部分进行 2 层分解得到 4 个频带。同理,3 层分解得到 8 个频带,每一个频带均为 2 维矩阵,共 16 个频段。

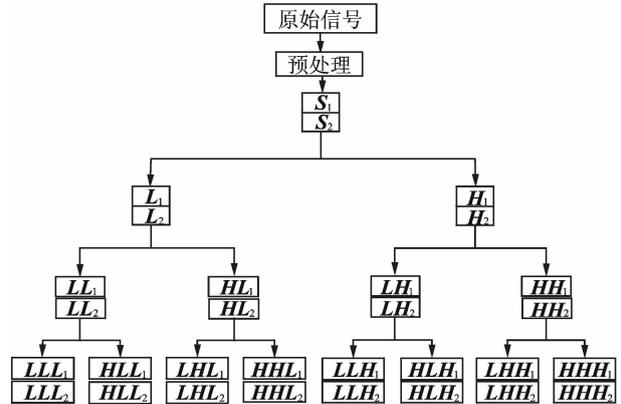


图 1 多小波包三层分解示意图

Fig. 1 Schematic diagram of three level decomposition by multiwavelet packet

### 1.2 样本熵

一般的,对于由  $N$  个数据组成的时间序列  $\{x(n)\} = x(1), x(2), \dots, x(N)$ , 样本熵的计算方法<sup>[13]</sup>为

1) 按序号组成一组维数为  $m$  的向量序列  $\mathbf{X}_m(1), \dots, \mathbf{X}_m(N-m+1)$ , 其中:  $\mathbf{X}_m(i) = \{x(i), x(i+1), \dots, x(i+m-1)\}, 1 \leq i \leq N-m+1$ 。这些向量代表从第  $i$  点开始的  $m$  个连续的  $x$  值。

2) 第 1 向量  $\mathbf{X}_m(i)$  与  $\mathbf{X}_m(j)$  之间的距离  $d[\mathbf{X}_m(i), \mathbf{X}_m(j)]$  为两者对应元素中最大差值的绝对值, 即  $d[\mathbf{X}_m(i), \mathbf{X}_m(j)] = \max_{k=0, \dots, m-1} (|x(i+k) - x(j+k)|)$ 。

3) 对于给定的  $\mathbf{X}_m(i)$ , 统计  $\mathbf{X}_m(i)$  与  $\mathbf{X}_m(j)$  之间距离小于等于  $r$  的  $j(1 \leq j \leq N-m, j \neq i)$  的数目, 并记做  $B_i$ 。对于  $1 \leq i \leq N-m+1$ , 定义

$$B_i^m(r) = \frac{1}{N-m-1} B_i \quad (10)$$

4) 定义  $B^{(m)}(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r)$

5) 增加维数到  $m+1$ , 计算  $\mathbf{X}_m(i)$  与  $\mathbf{X}_m(j)(1 \leq$

$j \leq N-m, j \neq i$ ) 距离小于等于  $r$  的个数, 记为  $A_i$ 。

$A_i^m(r)$  的定义为

$$A^{< m >}(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \quad (11)$$

这样,  $B_i^m(r)$  为两个序列在相似容限  $r$  下匹配  $m$  个点的概率, 而  $A_i^m(r)$  为两个序列匹配  $m+1$  个点的概率。样本熵定义为

$$\text{SampEn}(m, r) = \lim \left\{ -\ln \left[ \frac{A_m(r)}{B_m(r)} \right] \right\} \quad (12)$$

当  $N$  为有限值时, 可以用下式估计

$$\text{SampEn}(m, r, N) = \left\{ -\ln \left[ \frac{A_m(r)}{B_m(r)} \right] \right\} \quad (13)$$

可以看出, 样本熵的值与  $m, r$  的取值有关。因此, 确定  $m, r$  两个参数的值对于样本熵的计算非常重要。根据 Pincus 的研究结果,  $m=1, 2, r=0.1\text{std} \sim 0.25\text{std}$  (std 为原始数据  $\{x(n)\} = x(1), x(2), \dots, x(N)$  的标准差) 计算得到的样本熵具有较为合理的统计特性。在本研究计算中取  $m=2, r=0.2\text{std}$ 。

## 2 轴承损伤程度的神经网络识别

### 2.1 不同尺度的外圈故障实验

不同尺度的外圈故障轴承依次安装在轴承实验台上, 实验台所用轴承型号为 6307, 传感器为加速度传感器, 完成实验数据的采集。实验分 5 次进行, 分别对应于不同尺度的外圈点蚀故障。点蚀采用电火花加工, 为了形象表明轴承从故障早期一直发展到严重故障的过程, 笔者设计了多尺度损伤实验, 尺度分别为 0.2, 0.5, 2.0, 3.5 和 5.0 mm。轴承实验台结构如图 2 所示。电机转速  $r=1496\text{r/min}$ , 轴承外径  $D=80\text{ mm}$ , 内径  $d=35\text{ mm}$ , 滚动体个数  $z=8$ , 接触角  $\alpha=0$ 。实验每次采集 40 组数据, 采样频率为 15360 Hz, 采样点数为 8192。每次采集 40 组数据, 随机选择 20 组用于训练网络, 剩下的 20 组用于检验网络。图 3 为不同故障尺度的轴承振动信号时域及频谱图, (a)~(e) 分别代表 0.2, 0.5, 2.0, 3.5 和 5.0 mm 五种工况。由振动信号的时域和频谱图可见, 不同的故障尺度对应的能量集中带不同, 随着故障尺度的不断增大, 其能量集中趋向于 2 kHz 和 4 kHz; 因此, 根据能量集中差异能够实现故障程度的识别。

对信号进行 3 层 GHM 多小波分解, 得到 8 个

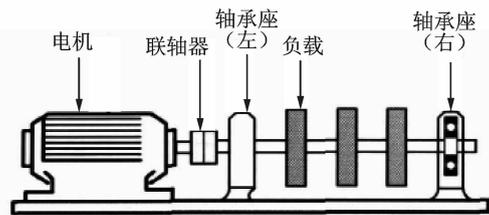


图 2 滚动轴承故障模拟实验台

Fig. 2 Test rig for seeded fault experiments of rolling bearing

频段、16 行信号。分别计算这 16 行信号的样本熵值, 表 1 为 5 种不同程度故障信号的训练样本在各个频段样本熵的平均值。可以看出, 不同的故障尺度其系数样本熵值的大小也不同, 可以根据这个特点对故障尺度进行区别。

### 2.2 多小波神经网络模型训练

采用 dB10 小波包和 GHM 多小波包对上述实验信号进行三层分解, 以分解后的系数样本熵作为神经网络输入向量, 样本熵值构成原始信号的特征向量  $[C_1, C_2, \dots]$ 。为了便于网络输入和提高特征向量的聚类性, 对特征向量进行归一化, 归一化后的特征向量的各元素为 0~1 之间的数。归一化的方法是对每一模式的相同频段样本熵值组成的序列进行归一化处理, 公式为

$$C_{i,j} = \frac{C_{i,j} - C_{\min}(j)}{C_{\max}(i) - C_{\min}(j)} \quad (14)$$

其中:  $C_{\max}(j), C_{\min}(j)$  对应  $j$  个序列的最大值和最小值,  $j=1, 2, \dots, N$  ( $N$  为 100)。

以此作为特征向量构造训练样本集, 并对 BP 神经网络进行训练, 使其达到要求的输出精度, 并且能够识别故障程度。

综合考虑, 笔者建立了三层 BP 网络。网络的输入向量范围为  $[0-1]$ , 隐层神经元的传递函数采用正切函数 tansig, 输出层神经元的传递函数采用 logsig, 输出模式为 0-1。训练函数为 trainlm, 学习函数为 learnngdm, 性能函数为 mse; 设定训练次数为 50000, 训练目标为 0.01。根据多小波包分解特性, 设定轴承模式识别的 BP 网络模型输入节点数为 16 (dB 10 小波包为 8), 分别对应轴承振动信号 16 (dB 10 小波包为 8) 个频段的系数样本熵, 输出层节点数为 5, 对应轴承的 5 类模式。建立小波包和多小波包神经网络综合模型, 每个模型网络训练的样本数为  $20 \times 5 = 100$  组, 网络输出的对应关系如表 2 所示。

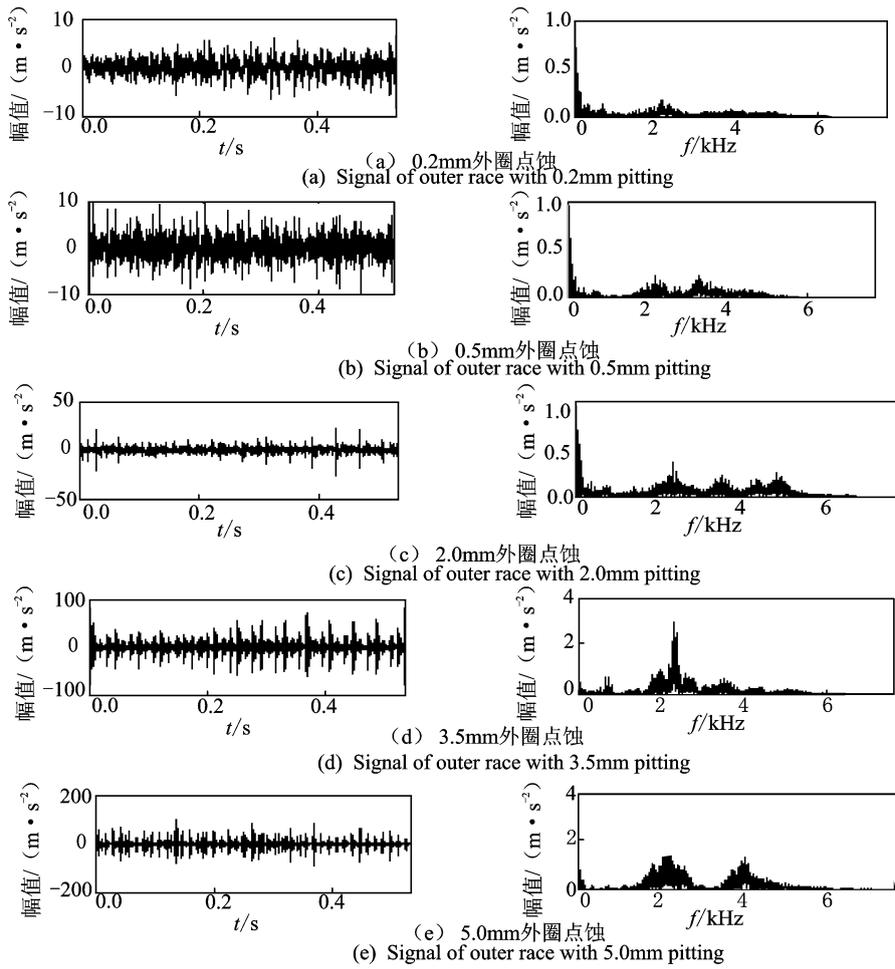


图 3 不同尺度的轴承故障信号时域及频谱图

Fig. 3 Waveform and spectrum of rolling element bearing with different level defects

表 1 故障信号经 GHM 多小波分解后的系数样本熵

Tab. 1 Sample entropy of faulty signals after GHM multiwavelet decomposition

故障尺度/mm	信号															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0.2	2.96	2.86	2.85	2.20	3.26	3.44	3.04	3.03	2.25	2.20	2.06	2.04	2.34	2.35	2.11	2.21
0.5	3.21	4.14	2.79	2.31	2.62	2.15	2.79	2.31	1.47	1.45	1.74	1.51	1.66	1.67	1.63	1.55
2.0	3.54	3.24	2.63	2.58	2.98	2.57	3.26	3.33	2.08	1.94	1.97	2.39	2.14	2.30	2.15	2.21
3.5	3.95	3.76	3.64	2.99	3.05	3.61	2.36	2.87	2.93	2.30	3.19	2.86	2.82	3.33	2.74	2.98
5.0	4.17	3.42	2.70	2.70	3.43	4.11	2.30	3.49	2.49	2.45	3.06	3.10	2.73	2.39	2.60	2.73

表 2 网络输出对应关系

Tab. 2 Mapping results between network outputs and fault severities

序号	理想输出	轴承故障程度/mm
1	1 0 0 0 0	0.2
2	0 1 0 0 0	0.5
3	0 0 1 0 0	2.0
4	0 0 0 1 0	3.5
5	0 0 0 0 1	5.0

### 3 模式识别网络检验

通过对神经网络的训练,得到对不同尺度故障进行分类的网络模型。为了研究网络的性能,使用实验中采集的后 20 组信号进行验证。测试中,为了提高模式识别的准确性,当输出向量中 5 个元素的最大值超过 0.6 时,即认定该输出有效,最大元素在向量中的位置与表 1 轴承模式中的序号对应,否则认为输出无效。表 3 为网络检验结果。由表 3 数据

可知:a.当缺陷为0.2mm时,网络的识别准确率最低,其原因在于故障尺寸小,使得激励产生的振动小,能量集中不明显,从而导致网络的识别结果有误,随着故障程度加大,识别误差减小,识别准确率较高;b.多小波包神经网络的识别效果好于dB10小波包神经网络的识别结果,原因由于多小波包三层分解后各个频段包含两行数据,共有16个特征量,能更好地将故障信息包含在分解后的16个频段中。因此,多小波包能够实现了对故障信号精细的逼近,提供丰富的故障信息,与传统的小波包相比存在更大的优势。

表3 网络检验结果

Tab.3 Verification results of network

输入样本类型	样本个数/组	误判个数/组		准确率/%	
		dB10小波包	GHM多小波包	dB10小波包	GHM多小波包
1	20	6	4	70	80
2	20	7	2	65	90
3	20	5	1	75	95
4	20	4	2	80	90
5	20	5	1	75	95

## 4 结论

1) 通过对比不同故障尺度的频谱图和和多小波系数样本熵可以看出,不同的故障尺度其能量的集中区段也不同。故障尺度较小时其能量分布较均匀,随着故障程度的增加,其能量分布越集中;同样,对于不同故障尺度的多小波系数样本熵,其值大小也不同。因此,可以根据以上两个指标来判断故障,并识别故障的损伤程度。

2) 多小波包能够实现了对故障信号精细的逼近,提供丰富的故障信息,与传统小波包相比存在更大的优势。从检测结果可以看出,多小波包神经网络比传统dB10小波包神经网络具有更高的可靠性和识别精度。

## 参 考 文 献

[1] 姚红良,韩清凯.基于谐波分量的转子系统碰摩故障定量诊断方法[J].机械工程学报,2012,48(5):43-48. Yao Hongliang, Han Qingkai. Method for detecting rubbing fault in rotor system based on harmonic components[J]. Journal of Mechanical Engineering, 2012, 48(5): 43-48. (in Chinese)

[2] 朱忠奎.基于支持向量回归的轴承故障定量诊断方法研究[D].苏州:苏州大学,2012.

[3] 窦东阳,赵英凯.基于EMD和Lempel-Ziv指标的滚动轴承损伤程度识别研究[J].振动与冲击,2010,29(3):5-8. Dou Dongyang, Zhao Yingkai. Fault severity assessment for rolling element bearings based on EMD and Lempel-Ziv index[J]. Journal of Vibration and Shock,

2010, 29(3): 5-8. (in Chinese)

[4] 侯敬宏,黄树红,申韬,等.基于小波分析的旋转机械振动信号定量特征研究[J].机械工程学报,2004,40(1):131-135. Hou Jinghong, Huang Shuhong, Shen Tao, et al. Wavelet-based quantitative analysis of vibration signal of rotary machines[J]. Journal of Mechanical Engineering, 2004, 40(1):131-135. (in Chinese)

[5] Yan Ruqiang, Gao R X. Approximate entropy as a diagnostic tool for machine health monitoring[J]. Mechanical Systems and Signal Processing, 2007, 21(2): 824-839.

[6] Pincus S M. Approximate entropy as a measure of system complexity[J]. Proceedings of the National Academy of Science of the United States of America, 1991, 88(6): 2297-2301.

[7] Lake D E, Richman J S, Griffin M P, et al. Sample entropy analysis of neonatal heart rate variability[J]. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2002, 283(3): 789-797.

[8] Pincus S M. Assessing serial irregularity and its implications for health[J]. Annals of the New York Academy of Sciences, 2002, 954:245-267.

[9] 苏文胜,王奉涛.基于小波包样本熵的滚动轴承故障特征提取[J].振动、测试与诊断,2011,31(2):162-166. Su Wensheng, Wang Fengtao. Feature extraction of rolling element bearing fault using wavelet packet sample entropy[J]. Journal of Vibration, Measurement & Diagnosis, 2011,31(2):162-166. (in Chinese)

[10] 王晓冬,何正嘉,訾艳阳.多小波自适应构造方法及滚动轴承复合故障诊断研究[J].振动工程学报,2010,23(4):438-444. Wang Xiaodong, He Zhengjia, Zi Yanyang. Adaptive construction of multiwavelet and research on composite fault diagnosis of rolling bearing[J]. Journal of Vibration Engineering, 2010,23(4):438-444. (in Chinese)

[11] Strela V, Strang G. The application of multiwavelet filterbanks to image processing[J]. IEEE Transaction on Image Processing, 1999, 8(4): 548-563.

[12] 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: 373-401.

[13] 周鹏,葛家怡,曹红宝.基于样本熵的运动想象分类研究[J].信息与控制,2008,37(2):191-196. Zhou Peng, Ge Jiayi, Cao Hongbao. Classification of motor imagery based on sample entropy [J]. Information and Control, 2008,37(2):191-196. (in Chinese)



第一作者简介:张建宇,男,1975年8月生,副教授。主要研究方向为机电设备故障诊断。

E-mail: zhjy\_1999@bjut.edu.cn

