

基于 VMD 的自适应复合多尺度模糊熵及其在滚动轴承故障诊断中的应用

郑近德, 姜战伟, 代俊习, 潘紫微

(安徽工业大学 机械工程学院, 安徽 马鞍山 243032)

摘 要: 提出了一种基于自适应多尺度模糊熵、ILS(迭代拉普拉斯得分)特征选择和粒子群优化支持向量机(PSO-SVM)的滚动轴承故障诊断方法。该方法采用变分模态分解对振动信号进行分解和重构,并计算重构信号的复合多尺度模糊熵;同时采用迭代拉普拉斯得分选择敏感故障特征,并将特征选择结果输入到基于粒子群优化支持向量机的多故障分类器进行识别。将提出的方法应用于滚动轴承试验数据分析。结果表明:该方法对试验数据的故障识别率为 100%。并将基于 ILS 特征选择方法与基于 SFS(sequential forward selection)特征选择进行了对比,表明基于 SFS 特征选择的最高识别率为 92.86%,而基于 ILS 特征选择的故障识别率达到 100%。

关 键 词: 滚动轴承; 故障诊断; 多尺度模糊熵; 变分模态分解; 特征选择

中图分类号: V231.92; TH165.3

文献标志码: A

VMD based adaptive composite multiscale fuzzy entropy and its application to fault diagnosis of rolling bearing

ZHENG Jinde, JIANG Zhanwei, DAI Junxi, PAN Ziwei

(School of Mechanical Engineering, Anhui University of Technology,
Maanshan Anhui 243032, China)

Abstract: A rolling bearing fault diagnosis approach was proposed based on the adaptive multiscale fuzzy entropy, ILS (iterative Laplacian score) and PSO-SVM (particle swarm algorithm optimization support vector machine). In the proposed method, the variational mode decomposition was used for the decomposition and reconstruction. Then composite multiscale entropy fuzzy of the reconstructed signals was calculated. Besides, the iterative Laplacian score algorithm was used for the sensitive fault feature selection and the selected features were input to the PSO-SVM based classifier for training and recognition. Finally, the proposed method was applied to experiment data of rolling bearing. Results showed that the identifying rate of proposed method was 100%. Also, the ILS based feature selection was compared with the SFS (sequential forward selection) method; and the result indicated that the highest identifying fault rate of SFS based method was 92.86% while the identifying

收稿日期: 2016-10-19

基金项目: 国家自然科学基金(51505002); 安徽省高校自然科学研究重点资助项目(KJ2015A080);
安徽工业大学研究生创新研究基金(2016062)

作者简介: 郑近德(1986—),男,副教授、硕士生导师,博士,主要从事非线性动力学、动态信号处理和机械故障诊断等方面的研究。
E-mail: lqdlzheng@126.com

通讯作者: 姜战伟(1990—),男,硕士生,主要研究方向为机械系统状态监控和故障诊断等。E-mail: jzwptt666888@gmail.com

引用格式: 郑近德,姜战伟,代俊习,等. 基于 VMD 的自适应复合多尺度模糊熵及其在滚动轴承故障诊断中的应用[J]. 航空动力学报, 2017, 32(7): 1683-1689. ZHENG Jinde, JIANG Zhanwei, DAI Junxi, et al. VMD based adaptive composite multiscale fuzzy entropy and its application to fault diagnosis of rolling bearing[J]. Journal of Aerospace Power, 2017, 32(7): 1683-1689.

fault rate of the ILS based fault diagnosis method reached to 100%.

Key words: rolling bearing; fault diagnosis; multiscale fuzzy entropy; variational mode decomposition; feature selection

由于受到负载、摩擦和冲击等非线性因素的影响,滚动轴承的故障振动信号表现为非平稳和非线性特性。许多非线性动力学方法,如分形,信息熵和近似熵等被广泛地应用在机械故障诊断中。如文献[1]将分形应用于离心压缩机叶轮旋转失效中;文献[2]将近似熵应用于齿轮箱的故障信号诊断。但近似熵存在自身匹配的不足^[3],对此,Richman 等^[4]提出了样本熵(sample entropy, 记为 SampEn,量符号记为 E_s)。虽然近似熵和样本熵都可作为量化信号复杂度的指标,但由于向量的相似性由原始数据的绝对值差确定,所以它们在评估两种模式是否相似时,当原始数据存在轻微波动或基线漂移时,造成统计的不稳定性,很难得到准确的分析结果。为此,Chen 等^[5-6]提出了模糊熵(fuzzy entropy, 记为 FuzzyEn,量符号记为 E_f),FuzzyEn 通过去均值操作,解决了基线漂移的干扰,从而减少了对参数的依赖性和敏感性,得到更好的统计稳定性。但 SampEn 和 FuzzyEn 都是衡量时间序列单一尺度的复杂性。为了弥补单一尺度的不足,Costa 等^[7-9]提出了衡量时间序列在不同尺度上复杂性的方法,即多尺度熵(multiscale entropy, MSE)。但是,MSE 粗粒化序列的计算中,熵值的偏差会随着粗粒化序列长度减小而增大,多尺度时间序列的熵值估计误差反而会随着粗粒化尺度因子的增大而增大^[10]。

最近,郑近德等^[11]针对多尺度熵中粗粒化方式的不足,提出了复合多尺度模糊熵(composite multiscale fuzzy entropy, CMFE),并将其与 MSE 进行了对比分析,结果表明,CMFE 考虑同一尺度下的多个粗粒化时间序列的熵值,能够有效地抑制由粗粒化时间序列变短而导致的熵值突变,CMFE 能够得到更好的稳定性熵值。尽管 CMFE 弥补了 MSE 的不足,但在 CMFE 中,数据的粗粒化过程本质上是线性光滑化和原始时间序列的提取,这导致仅仅捕获粗尺度上的低频成分,损失了细尺度上的高频成分。而且由于线性操作,在处理非平稳、非线性振动信号情况下,其算法提取不同尺度信息时不具有自适应性。

Dragomiretskiy 和 Zosso^[12]在传统维纳滤波的基础上,提出了一种非递归的自适应的信号分解方法——变分模态分解(variational mode de-

composition, VMD)。VMD 可以将信号分解问题转化为约束优化问题,将一个复杂的非平稳振动信号自适应地分解为若干个不同粗粒尺度的本征模态函数(intrinsic mode function, IMF)之和。VMD 具有精度高和收敛速度快的特点,消除了指数衰减直流偏移,非常适合处理滚动轴承故障振动信号^[13]。结合 VMD 方法的优势,本文提出了一种基于 VMD 的自适应复合多尺度模糊熵(adaptive composite multiscale fuzzy entropy, ACMFE)算法,并将 ACMFE 与粒子群优化-支持向量机(particle swarm optimization-support vector machine, PSO-SVM)^[14]结合,提出了一种滚动轴承故障诊断方法。最后,将提出的方法应用于滚动轴承故障试验数据分析。结果表明:论文提出的方法能够有效地诊断出不同的故障状态,其是一种有效的滚动轴承故障诊断方法。

1 基于 VMD 的自适应复合多尺度模糊熵算法

1.1 变分模态分解

VMD 是一种自适应信号分解方法,该方法利用迭代的方法求解变分模型的最优解,从而获取各个 IMF 的频率中心和带宽,能够自适应地将各 IMF 分量有效分离。

为了估计每一个模态的带宽,构建的受约束的变分问题如下:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\},$$
$$\text{s. t.} \quad \sum_k u_k = f \tag{1}$$

其中 $\{u_k\} := \{u_1, \dots, u_k\}$ 和 $\{\omega_k\} := \{\omega_1, \dots, \omega_k\}$ 分别代表一系列模态函数及其中心频率。

引入二次惩罚因子 α 与拉格朗日乘法算子 λ ,上述约束问题简化为非约束变分问题。增广拉格朗日乘子表达式如下:

$$L(\{u_k\}, \{\omega_k\}, \lambda) :=$$
$$\alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 +$$
$$\left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \tag{2}$$

采用交替方向乘子法求解增广拉格朗日式子的鞍点,即通过交替更新 $u_k^{n+1}, \tau_k^{n+1}, \lambda^{n+1}$ 寻找约束变分问题的最优解。

为了更新 u_k ,给出非负频率区间上的半空间积分形式

$$\hat{u}_k^{n+1} = \operatorname{argmin}_{\hat{u}_k, u_k \in X} \left[\int_0^\infty 4\alpha(\omega - \omega_k)^2 |\hat{u}_k(\omega)|^2 + 2 \left| f(\omega) - \sum_i \hat{u}_i(\omega) + \frac{\lambda(\omega)}{2} \right|^2 d\omega \right] \quad (3)$$

在正频率区间,解得上式二次优化问题

$$\hat{u}_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

式(4)可以认为是一个带有 $1/(\omega - \omega_k)^2$ 剩余趋势项的维纳滤波。对这个滤波解析信号 $\hat{u}_k(\omega)$ 进行傅里叶逆变换后的实部就是时域模式 $u_k(t)$ 。

中心频率 ω_k 不在重构保真度项中,而仅仅出现在带宽项中,即

$$\omega_k^{n+1} = \operatorname{arg\,min}_{\omega_k} \left[\left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \cdot u_k(t) \right) e^{-j\omega_k t} \right] \right\|_2^2 \right] \quad (5)$$

同理,将其变换到频域

$$\omega_k^{n+1} = \operatorname{arg\,min}_{\omega_k} \left[\int_0^\infty (\omega - \omega_k)^2 |\hat{u}_k(\omega)|^2 d\omega \right] \quad (6)$$

解得中心频率的更新方法

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (7)$$

VMD 算法具体过程如下:

- 1) 初始化 $\{\hat{u}_k^1\}, \{\omega_k^1\}, \lambda^1, n \leftarrow 0$;
- 2) $n = n + 1$, 执行整个循环;
- 3) $k = k + 1$, 直到 $k = K$, 更新 u_k

$$\begin{aligned} &\hat{u}_k^{n+1}(\omega) \leftarrow \\ &\frac{f(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\lambda^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \end{aligned} \quad (8)$$

更新 ω_k

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (9)$$

4) 更新 λ

$$\lambda^{n+1}(\omega) \leftarrow \lambda^n(\omega) + \tau \left(f(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (10)$$

5) 重复步骤 2)~步骤 4),直至满足收敛条件 $\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < \epsilon$, 整个循环结束,输出 K 个分量。

VMD 是一种自适应与准正交信号分解方法,将复杂信号非递归地分解成一系列不同尺度有限带宽的 IMF 之和,非常适合处理非平稳、非线性复杂信号。

1.2 复合多尺度模糊熵

CMFE 计算步骤如下:

1) 对于给定的时间序列 $\{x(i), i = 1, 2, \dots, N\}$,通过式(11)来定义粗粒化序列 $y_k^{(\tau)} = \{y_{k,1}^{(\tau)}, y_{k,2}^{(\tau)}, \dots, y_{k,p}^{(\tau)}\}$,即

$$\begin{aligned} y_{k,j}^{(\tau)} &= \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau+k-1} x_i \\ 1 \leq j \leq N/\tau, \quad 1 \leq k \leq \tau \end{aligned} \quad (11)$$

2) 对于每个粗粒尺度因子 τ ,首先计算各个粗粒时间序列 $y_k^{(\tau)} (1 \leq k \leq \tau)$ 的模糊熵,然后对 τ 个熵值求均值,从而得到该粗粒尺度因子下的 CMFE(量符号记为 E_{cmf})

$$E_{\text{cmf}}(X, \tau, m, n, r) = \frac{1}{\tau} \sum_{k=1}^{\tau} E_f(y_k^{(\tau)}, m, n, r) \quad (12)$$

最后将得到的模糊熵值表达成粗粒尺度因子的函数,这个过程被称为复合多尺度模糊熵分析(CMFE)。

CMFE 的粗粒化过程本质上是一种线性和平稳的平滑算法,由于原始振动信号是非平稳、非线性复杂振动信号,直接采用 CMFE 不可避免有一定的局限。因此,论文考虑首先采用 VMD 分解为一系列不同尺度 IMF 之和,各个 IMF 分量近似平稳,再通过从粗到精尺度或从精到粗尺度对原始信号进行自适应尺度的重构;然后分别估计重构信号的 CMFE,称之为自适应复合多尺度模糊熵(ACMFE,量符号记为 E_{acmf})。

1.3 自适应复合多尺度模糊熵

ACMFE 的主要步骤如下:

步骤一 原始复杂振动信号经 VMD 自适应地分解为一系列不同尺度的 IMF 之和;

步骤二 计算经依次剔除低频或高频信号后叠加信号的 CMFE。步骤二中的不同尺度化方式决定了 ACMFE 有两种算法,分别为从精到粗 ACMFE 和从粗到精 ACMFE 算法。

1) 算法 1(从精到粗)

(1)采用 VMD 算法对原始非线性振动信号

进行自适应分解,从而得到一系列不同粗粒尺度的 IMF 分量之和。

(2)通过依次剔除高频 IMF (量符号记为 F_{im}),再将剩余 IMFs 分量叠加,则得到自适应粗粒化序列 $y(s)$,即

$$y(s) = \sum_{i=s}^k F_{im,i}, 1 \leq s \leq k \quad (13)$$

式中 s 指自适应尺度因子, k 是经 VMD 分解后 IMFs 总数。

(3) 计算 ACMFE。

$$E_{acmf}(\tau) = E_{cmf}(y(s), \tau, m, n, r) \quad (14)$$

式中 τ 表示尺度因子, m 表示嵌入维数, n 表示指数函数梯度参数, r 表示相似容限。

2) 算法 2(从粗到精)

(1)利用 VMD 算法对原始非线性振动信号进行分解,从而得到一系列不同尺度的 IMF 分量之和。

(2)通过依次剔除低频 IMF,再将剩余 IMFs 分量叠加,则得到自适应粗粒化序列 $y(s)$,即

$$y(s) = \sum_{i=s}^k F_{im,i}, 1 \leq s \leq k \quad (15)$$

(3)计算 ACMFE。

$$E_{acmf}(\tau) = E_{cmf}(y(s), \tau, m, n, r) \quad (16)$$

从 CMFE 和 ACMFE 算法过程可以看出, CMFE 由于线性操作,在处理非平稳、非线性复杂振动信号时,其算法提取不同尺度信息时不具有自适应性;而对于这类信号,ACMFE 方法由于 VMD 自适应地分解信号,能够得到自适应粗粒化尺度。所以在处理非平稳、非线性复杂振动信号时,ACMFE 方法优于 CMFE 和 MSE。

2 基于 ACMFE 的滚动轴承故障诊断方法

2.1 方法步骤

在得到振动信号的 ACMFE 故障特征之后,得到的特征向量的维数一般比较高,若直接采用分类器进行识别会降低识别效率。因此,有必要采用特征选择的方法消除无关或冗余特征。拉普拉斯得分(Laplacian score, LS)^[15]是一种新的挖掘特征信息效果显著的特征选择算法。其综合了拉斯拉斯特征映射^[16]与局部保持投影^[17]利用局部保持能力来衡量特征信息,将高维特征空间转化为低维特征空间,保存了特征集内含的整体几何结构信息。而 ILS^[18]通过其局部保持能力,在每次迭代中剔除最不相关特征来逐渐更新近邻图,选择的特征子集比 LS 更能体现数据潜在的

流行结构。因此,论文采用 ILS 算法对特征向量矩阵进行降维。最后,为了实现智能诊断,采用训练速度快、适合小样本分类的 PSO-SVM 多类模式分类器对滚动轴承的不同故障类型进行模式识别。

基于上述分析,论文提出的基于 ACMFE, ILS 和 PSO-SVM 的滚动轴承故障诊断方法步骤如下:

1) 假定滚动轴承的状态包含 K 类,每类状态采集 N 组样本;

2) 提取每类原始振动信号的 ACMFE,每组样本得到 r 个特征值,组成特征向量矩阵 $R^{N \times r}$,其中, $r = k \times \tau_{\max}$, k 是 VMD 算法中的自适应尺度数,需要预先设定; $\tau_{\max}$ 是 CMFE 算法中最大尺度因子,一般小于等于 20;

3) 每种状态各取 $N/2$ 组组成原始训练样本,其余作为原始测试样本。采用 ILS 对原始训练样本特征值进行排序,将得分较高的前 p 个序列对应的特征值组成敏感状态低维特征向量矩阵,得到一个反映振动信号内在信息的低维训练样本集。依据原始训练样本特征值的 ILS 排序,原始测试样本同样取 ILS 序列中得分较高的前 p 个序列对应的原始测试样本特征值作为低维测试样本集,一般 p 取 6;

4) 将低维训练样本集输入到 PSO-SVM 进行训练,用训练好的 PSO-SVM 对低维测试样本集进行分类预测。根据 PSO-SVM 多模式分类器的输出结果来判断滚动轴承的状态和故障类型。

2.2 实验验证

为了验证提出方法的有效性,将该方法应用于轴承试验数据分析。试验数据采用美国 Case Western Reserve University 的滚动轴承试验测试数据^[19]。测试轴承型号为 6205-2RSJEM SKF 深沟球轴承,通过电火花加工技术在滚动轴承上布置单点故障,故障直径大小与深度分别为 0.5334mm 和 0.2794mm。在转速为 1730 r/min、负载为 2238kW,采样频率为 12kHz 条件下,采集到正常(norm)以及具有局部单点点蚀的内圈(inner race, IR)、滚动体故障(ball element, BE)和外圈(outer race, OR)4 种状态的轴承振动信号。每种状态取 29 组数据,每组数据长度为 4096 个数据点。

将提出的方法应用于上述实验数据分析,具体步骤如下:

1) 提取每类原始振动信号的 ACMFE,每种状态取 29 组样本,每组样本得到 80 个特征值,4

种状态共得到 116 组样本,组成特征向量矩阵 $\mathbf{R}^{116 \times 80}$,其中 $r = k \times \tau_{\max} = 80, k$ 取 8, $\tau_{\max}$ 取 10;

2) 每种状态取 15 组组成原始训练样本集 $\mathbf{R}^{60 \times 80}$,其余作为原始测试样本集 $\mathbf{R}^{56 \times 80}$ 。采用

ILS 对原始训练样本特征向量矩阵的所有特征值依照重要性进行排序,其排序前如图 1(a)所示,排序后的特征值重要性顺序如图 1(b)所示。

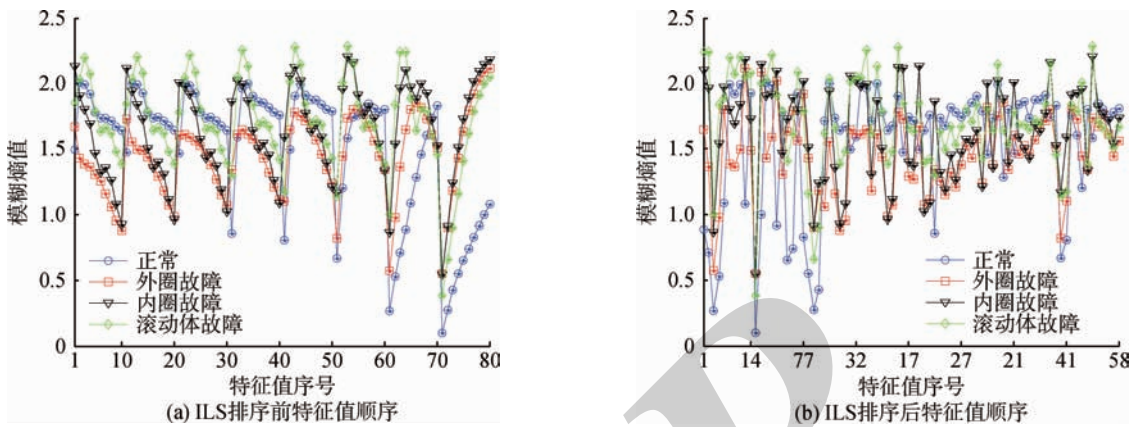


图 1 特征值排序前后对比

Fig. 1 Feature value comparisons before and after sorting

由图 1 可以看出,ILS 排序前四类样本的所有特征值几乎都非常靠近,而采用 ILS 依据特征重要性进行排序后,4 类样本的特征值在前 20 个

维度上分的很开,这说明了采用 ILS 可以选取出区分不同状态内在信息的特征值。在图 1(b)中,所有特征值按照得分从高到底排序结果如下:

$$\begin{aligned} &S_{64} > S_{63} > S_{61} > S_{62} > S_{65} > S_3 > S_4 > S_{13} > S_{80} > S_{14} > S_{71} > S_{79} > S_2 > S_{23} > S_{78} > S_5 > \\ &S_{75} > S_{76} > S_{24} > S_{77} > S_{74} > S_{72} > S_{73} > S_8 > S_{12} > S_7 > S_{10} > S_9 > S_{42} > S_{32} > S_{22} > S_{33} > \\ &S_{18} > S_{34} > S_{15} > S_{20} > S_{19} > S_{43} > S_{11} > S_{17} > S_{28} > S_1 > S_{30} > S_{40} > S_{31} > S_6 > S_{29} > S_{38} > \\ &S_{39} > S_{27} > S_{25} > S_{37} > S_{35} > S_{50} > S_{67} > S_{16} > S_{41} > S_{66} > S_{49} > S_{21} > S_{48} > S_{36} > S_{26} > S_{47} > \\ &S_{46} > S_{45} > S_{54} > S_{70} > S_{51} > S_{41} > S_{55} > S_{68} > S_{52} > S_{60} > S_{53} > S_{57} > S_{56} > S_{69} > S_{59} > S_{58} \end{aligned}$$

其中 S_j 表示第 j 个特征值的得分。得分值越大,与其对应的特征越重要。

将得分较高的前 6 个序列对应的特征值作为低维敏感故障训练样本特征集 $\mathbf{R}^{60 \times 6}$ 。从原始测试样本集中选择与训练样本特征集中特征序列相同的特征值作为测试样本特征集 $\mathbf{R}^{56 \times 6}$ 。

3) 采用训练样本特征集对基于 PSO-SVM 多类模式分类器进行训练,由于 SVM 是二分类器,在处理多分类时,就需要构造合适的多类分类器,常见的方法有一对多和一对一两种,本文采用基于一对一法的 SVM 多模式分类器。其中用粒子群优化算法优化 SVM 中惩罚参数 c 和核函数参数 g ,对训练样本特征集进行交叉验证(cross validation, CV)意义下的识别率作为 PSO 中的适应度函数值,设定 PSO 算法中的局部搜索能力 c_1 为 1.5,全局搜索能力 c_2 为 1.7,种群数量为 20,迭代次数为 200。搜索到的最佳惩罚因子 c 为 0.1,核函数参数 g 为 12.1675。

4) 用训练好的 PSO-SVM 多模式分类器对测试样本特征集进行智能识别预测。根据 PSO-SVM 多模式分类器的输出结果来判断滚动轴承的工作条件和故障类型,预测的输出结果如图 2 所示,图中纵坐标 1,2,3,4 分别表示正常、外圈故障、内圈故障、滚动体故障。

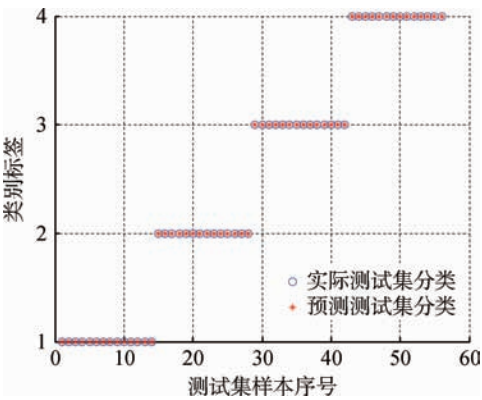


图 2 测试样本预测结果

Fig. 2 Prediction results of test samples

由图 2 可知,所有测试样本类别都得到了准确的分类,论文提出的方法对测试样本集的识别率达到 100%,这充分说明了论文方法的有效性。

为了说明多尺度分析的必要性,首先,采用 VMD 分别对上述正常、外圈、内圈和滚动体故障 4 类状态振动信号进行自适应尺度分解,得到 8 个不同尺度的 IMFs,再按照算法 1 步骤(2)方式重构自适应粗粒化尺度,最后对各个粗粒化尺度直接模糊熵分析,即令 $\tau=1$,得到 8 个特征值,其诊断结果如图 3 所示。虽然其分类结果达到了 100%,但是这里提取的特征值是 ILS 分值序列为 $S_1, S_{11}, S_{21}, S_{31}, S_{41}, S_{51}, S_{61}$ 和 S_{71} 的特征值,其分值并不是上述特征分值重要性序列的前几个,没有充分挖掘内在区分不同状态的特征信息,很难保证其诊断应用上的一般性,而且特征过多会造成信息冗余,降低计算性能。

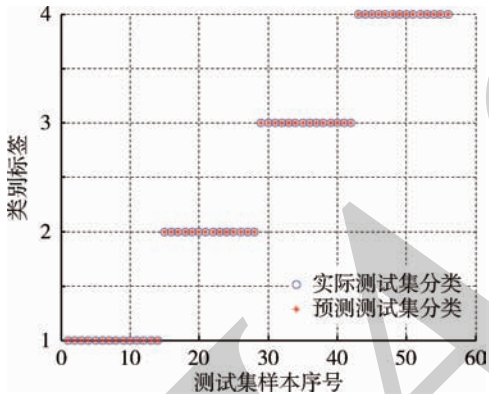


图 3 单一尺度特征值的分类结果

Fig. 3 Classification results by using single scale feature value

为了研究选取的敏感特征个数对诊断率的影响,采用 ILS 方法依照特征重要性排序后,分别选取特征最重要的前 1~10 个敏感特征,在其余相同参数下分别对基于 PSO-SVM 的多故障分类器进行训练与测试,故障识别率如图 4 所示。由图 4 可知,采用 ILS 方法选取的敏感特征数在 1~5 个时,正确识别率都高于 94.64%,而且当敏感特征数大于 5 时,识别率都能达到 100%,这说明敏感特征个数过少不能够完全反映振动信号内在的故障特征信息,而敏感特征个数太多会造成信息冗余,降低计算性能。

最后,为了说明 ILS 特征选择方法的优势,采用序列前向选择(sequential forward selection, SFS)^[20]算法分别选取 1~10 个敏感故障特征,结果如图 4 所示,从图中可以看出,随着选取特征个

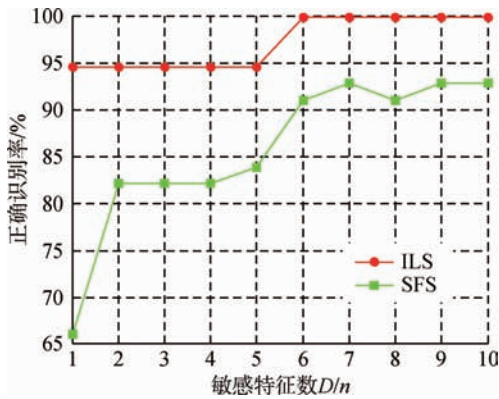


图 4 ILS 和 SFS 算法选择不同敏感特征数的识别率

Fig. 4 Identification rate by using ILS and SFS algorithms with different number of sensitive features

数的增加,采用 ILS 特征选择方法的诊断率都高于 SFS 选择方法,而且当特征个数大于 1 时,采用 ILS 方法的最低诊断率为 94.64%大于 SFS 选择方式的最高诊断率(92.86%),这说明了 ILS 算法在挖掘特征信息方面的优势。

3 结 论

1) 提出了一种衡量时间序列复杂性的方法——自适应复合多尺度模糊熵(ACMFE)。ACMFE 能够自适应地得到粗粒化时间序列,在一定程度上解决了 CMFE 粗粒化方式的不足。

2) 提出了一种基于 ACMFE, ILS 和 PSO-SVM 的滚动轴承故障诊断方法。试验数据分析结果表明,论文方法对试验数据的故障识别率为 100%,高于现有方法的识别率。

3) 同时将基于 ILS 特征选择方法与随机选择和基于 SFS 特征选择进行了对比,结果表明进行 ILS 特征选择后的故障识别率达到 100%,而基于 SFS 特征选择的最高识别率为 92.86%,这说明了 ILS 特征选择的必要性和优越性。

尽管如此,方法也有不足之处,如 ACMFE 中参数的选择需要人为设定等,笔者正在对这些问题进行研究和完善。

参考文献:

[1] 王乐,张家忠,周成武,等. 离心压缩机叶轮旋转失速的相空间重构及分形特征[J]. 航空动力学报, 2014, 29(10): 2424-2433.
WANG Le, ZHANG Jiazhong, ZHOU Chengwu, et al. Phase space reconstruction and fractal characteristic of rotating stall in impeller of centrifugal compressor[J]. Journal of Aerospace Power, 2014, 29(10): 2424-2433. (in Chi-

nese)

[2] 张淑清,李盼,胡永涛,等.多重分形近似熵与减法 FCM 聚类的研究及应用[J].振动与冲击,2015,34(18):205-209.
ZHANG Shuqing,LI Pan,HU Yongtao,et al. Application of multifractal approximate entropy and subtractive FCM clustering in gearbox fault diagnosis[J]. Journal of Vibration and Shock,2015,34(18):205-209. (in Chinese)

[3] PINCUS S M. Approximate entropy;a complexity measure for biological time series data[C]// Proceedings of the IEEE Seventeenth Annual Northeast. Hartford,CT: IEEE,1991: 35-36.

[4] RICHMAN J S,MOORMAN J R. Physiological time-series analysis using approximate entropy and sample entropy[J]. American Journal of Physiology-Heart and Circulatory Physiology,2000,278(6):2039-2049.

[5] CHEN W,WANG Z,XIE H,et al. Characterization of surface EMG signal based on fuzzy entropy[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2007,15(2):266-272.

[6] CHEN W,ZHUANG J,YU W,et al. Measuring complexity using FuzzyEn, ApEn, and SampEn[J]. Medical Engineering and Physics,2009,31(1):61-68.

[7] COSTA M,GOLDBERGER A L,PENG C K. Multiscale entropy analysis of biological signals[J]. Physical Review: E,2005,71(2):021906. 1-021906. 18.

[8] COSTA M,GOLDBERGER A L,PENG C K. Multiscale entropy analysis of complex physiologic time series[J]. Physical Review Letters,2002,89(6):068102. 1-068102. 4.

[9] COSTA M,PENG C K,GOLDBERGER A L,et al. Multiscale entropy analysis of human gait dynamics[J]. Physica A: Statistical Mechanics and Its Applications, 2003, 330 (1):53-60.

[10] WU S D,WU C W,LIN S G,et al. Time series analysis using composite multi-scale entropy[J]. Entropy, 2013, 15 (3):1069-1084.

[11] 郑近德,潘海洋,程军圣,等. 基于复合多尺度模糊熵的滚动轴承故障诊断方法[J]. 振动与冲击,2016,35(8):116-123.
ZHENG Jinde,PAN Haiyang,CHENG Junsheng,et al. Composite multiscale fuzzy entropy based rolling bearing fault diagnosis method [J]. Journal of Vibration and Shock,2016,35(8):116-123. (in Chinese)

[12] DRAGOMIRETSKIY K,ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014,62(3):531-544.

[13] 姜战伟,郑近德,潘海洋,等. POVMD 与包络阶次谱的变工况滚动轴承故障诊断[J]. 振动、测试与诊断,2017,37 (3):609-617.
JIANG Zhanwei,ZHENG Jinde,PAN Haiyang,et al. Fault diagnosis under variable conditions based on parameter optimized variational mode decomposition and envelope order spectrum[J]. Journal of Vibration, Measurement and Diagnosis,2017,37(3):609-617. (in Chinese)

[14] GARCÍA-NIETO P J, GARCÍA-GONZALO E, VILÁN J A V,et al. A new predictive model based on the PSO-optimized support vector machine approach for predicting the milling tool wear from milling runs experimental data[J]. International Journal of Advanced Manufacturing Technology,2015,86(1):1-12.

[15] HE X,CAI D,NIYOGI P. Laplacian score for feature selection[R]. Vancouver,Canada:Advances in Neural Information Processing Systems,2005.

[16] BELKIN M, NIYOGI P. Laplacianeigenmaps for dimensionality reduction and data representation [J]. Neural Computation,2003,15(6):1373-1396.

[17] NIYOGI X. Locality preserving projections[J]. Neural Information Processing Systems,2004,16(1):186-197.

[18] ZHU L,MIAO L,ZHANG D. Iterative Laplacian score for feature selection[M]. Berlin:Springer,2012.

[19] 姜战伟,郑近德,潘海洋,等. 基于多尺度时不可逆与 t-SNE 流形学习的滚动轴承故障诊断[J]. 振动与冲击, 2017,36(17):61-68.
JIANG Zhanwei,ZHENG Jinde,PAN Haiyang,et al. Rolling bearing fault diagnosis method based on multiscale time irreversibility and t-SNE manifold learning[J]. Journal of Vibration and Shock,2017,36(17):61-68. (in Chinese)

[20] MARCANO-CEDEÑO A,QUINTANILLA-DOMÍNGUEZ J, Cortina-Januchs M G,et al. Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network [C] // Proceedings of 36th Annual Conference on IEEE Industrial Electronics Society. Glendale,AZ:IEEE,2010:2845-2850.

(编辑:王碧珺)