

基于 ICEEMDAN 和 PSO-LSSVM 的石油机械滚动轴承故障诊断方法研究

郑立朝¹, 宋宏志¹, 顾启林¹, 章宝玲¹, 安宏鑫¹, 张瀚阳², 别锋锋²

(1. 中海油田服务股份有限公司 油田生产事业部, 天津 300459;

2. 常州大学 机械与轨道交通学院, 江苏 常州 213164)

摘要: 针对滚动轴承疲劳故障振动信号具有能量弱、特征稀疏等特点, 提出了一种通过改进自适应噪声完备经验模态分解方法与粒子群优化的最小二乘支持向量机结合的故障识别方法; 对轴承不同故障信号利用改进的自适应噪声完备经验模态算法分解为一系列固有模态函数分量; 根据相关系数-方差贡献率准则筛选出最能表征原始信号状态的分量, 并计算重构分量的奇异谱熵值构成特征向量; 将提取的特征向量集合输入到基于粒子群优化的最小二乘支持向量机分类器中, 进行模型的训练和故障模式的识别, 与 SVM 和 LSSVM 分类器模型进行准确率和效率比较; 试验结果表明, 该方法在滚动轴承故障信号中能有效提取故障特征, 准确率达 98.75%, 具有一定可靠性和实用性。

关键词: 滚动轴承; ICEEMDAN 分解; 奇异谱熵; PSO-LSSVM; 模式识别

Research on Rolling Bearing Fault Diagnosis Method in Petroleum Machinery Based on ICEEMDAN and PSO-LSSVM

ZHENG Lizhao¹, SONG Hongzhi¹, GU Qilin¹, ZHANG Baoling¹,

AN Hongxin¹, ZHANG Hanyang², BIE Fengfeng²

(1. Production Optimization, China Oilfield Services Limited, Tianjin 300459, China;

2. School of Mechanical and Rail Transit, Changzhou University, Changzhou 213164, China)

Abstract: In view of the weak energy and sparse features of fatigue fault vibration signals of rolling bearings, a fault identification method combining improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) and particle swarm optimization least-squares support vector machine (PSO-LSSVM) is proposed. The improved adaptive noise complete empirical mode algorithm is used to decompose different bearing fault signals into a series of inherent modal function (IMF) components; The component that can best represent the original signal state is selected according to the correlation core-variance contribution ratio criterion, and the singular spectrum entropy of the reconstructed component is calculated to form the feature vector; The extracted feature vector set is input into the least square support vector machine classifier based on particle swarm optimization, which trains the model and identifies the fault mode. The accuracy and efficiency of the model are compared with that of the support vector machine (SVM) and least-squares support vector machine (LSSVM) classifier. The test results show that the method can effectively extract fault characteristics from rolling bearing fault signals, with an accuracy of 98.75%, which has certain reliability and practicability.

Keywords: rolling bearing; ICEEMDAN decomposition; singular spectrum entropy; PSO-LSSVM; pattern recognition

0 引言

滚动轴承是旋转机械的重要部分, 其状态影响整个设备的运行^[1], 一旦发生故障容易产生巨大的经济损失。滚动轴承的故障特征频率可用于检测故障, 但在强噪声环境下, 故障信号可能被淹没, 导致信噪比低, 难以区分某些频率特征。当滚动轴承发生故障时, 提取非平稳非线性信号的特征是困难的。因此, 提高机械设备的可靠性, 保证设备的安全运行, 对滚动轴承健康状态的准确判断至关

重要^[2]。

20 世纪 60 年代, 国外就逐步对滚动轴承进行振动检测并进行后续的状态诊断。1962 年, 文献 [3] 首先将加速度传感器应用到滚动轴承上, 并通过波形的峰值变化趋势来判断故障类型。1965 年, 随着傅里叶变换算法的提出, 信号可以从传统的时域分析变化进入频域分析^[4]。到 70 年代, 日本 NSSC 公司研发的测振仪在采集过程中消除了低频段的噪声干扰, 这使得处理的信号准确率得到很大的提升。

收稿日期:2023-12-04; 修回日期:2023-12-20。

基金项目:国家重点研发计划项目(2021YFB3302104)。

作者简介:郑立朝(1984-),男,硕士,工程师。

引用格式:郑立朝,宋宏志,顾启林,等.基于 ICEEMDAN 和 PSO-LSSVM 的石油机械滚动轴承故障诊断方法研究[J].计算机测量与控制,2024,32(8):129-137.

到 90 年代,意大利卡西诺大学针对早期故障,利用谱技术实现了对滚动轴承的诊断^[5]。文献 [6] 针对滚动轴承声发射信号提出了一种三角包络故障频率方法,这对滚动轴承故障源定位是非常有价值的工具。文献 [7] 提出了一种基于 AE 熵来区分滚动轴承不同损伤阶段,尤其对于临界损伤的鉴别十分有效。针对滚动轴承信号的实际特点,文献 [8] 在 1998 年提出了经验模态分解 (EMD, empirical mode decomposition) 算法; Smith 提出了局部均值分解算法。这些时频分析方法的诞生,对滚动轴承实际检测信号的处理有着重大意义,为频谱分析领域开创了新思路。

传统的利用频谱分析方法虽然可以找到滚动轴承故障特征频率,但是不能确定损伤程度,具有一定的局限性。近年来一些基于自适应信号分解的方法得到了发展,如 EMD 可以将信号自适应地分解为不同尺度的固有模态函数 (IMF, intrinsic mode function) 以及残余分量^[9],可以对不同分量进行独立分析,然而,该算法缺少严谨的数学基础,计算效率低下,容易导致模态混叠和虚假分量等诸多问题。为了应对这一难题,文献 [10] 提出了集合经验模态分解算法 (EEMD, ensemble empirical mode decomposition), EEMD 在一定程度上能够抑制模态混叠现象,但是该算法的执行时间较长,从而导致效率较低。在此基础上,文献 [11] 提出自适应完全集合经验模态分解方法 (CEEMDAN, complete ensemble empirical mode decomposition with adaptive noise), CEEMDAN 可有效降低重构误差,提高分解效率,然而,在信号分解过程中,模态混叠与产生虚假分量的问题是无法完全解决的。在此基础上 Colominas 等提出改进型 CEEMDAN 算法一种方法以提高信号分解完备性。

滚动轴承故障诊断的关键在于模式识别。支持向量机 (SVM, support vector machine) 最早由 Vapnik 提出,其核心思想是通过使用核函数对坐标进行变换,使低维线性不可分问题转化为高维线性可分问题,然后在高维空间中构造出超平面实现非线性分类和回归。由于其具有最小化结构风险与经验风险等优点,在处理轴承状态监测等小样本数据问题时,泛化能力比人工神经网络更强,因此在机械状态监测与故障机器识别中得到了广泛应用^[12]。为解决 SVM 在求解对偶问题时计算过程繁杂、效率较低等问题, Suykens 等人的研究成果是将复杂的二次规划问题转化为线性方程组求解问题的最小二乘支持向量机 (LSSVM, least squares support vector machines) 方法,这一方法的出现大大提高了计算效率。滚动轴承故障的诊断实际上涉及模式识别过程,最小二乘支持向量机的分类技术发展迅速^[13]。基于粒子群 (PSO, particle swarm optimization) 优化的算法具有搜索速度快、效率高和收敛能力强等优点^[14]。

综上所述,提出基于 ICEEMDAN-PSO-LSSVM 模型的滚动轴承故障识别诊断方法。首先,将原始信号进行 ICEEMDAN (Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) 分解,得到固有模态函数 (IMF, intrinsic mode function)。然后,通过奇异谱

熵的提取方法得到特征向量集合。随后,利用基于 PSO-LSSVM 模型的框架,输入包含特征向量的集合进行模型的训练,以此来达到对滚动轴承故障的识别。

1 基本原理

1.1 ICEEMDAN 算法

ICEEMDAN 是 EEMD 的一种改进算法,在重构过程中加入正负对互补的白噪声来有效抑制噪声问题,提高分解效率^[15]。Colominas 等人提出了改良的 CEEMDAN 算法,而 ICEEMDAN 方法则在第 K 个 IMF 分量中加入了特殊白噪声,计算出每个模态分量信号加噪声的局部平均值,并将分解得到的残差信号和局部平均值的差值,这种特殊的白噪声 $E_k(\mathbf{w}^{(i)})$,降低分量当中的残余噪声,改善了容易产生虚假成分和模态混合的缺点^[16]。其分解的具体流程是:

1) 向初始信号 x 中添加白噪声:

$$X_1^{(i)} = x + e_1 E_1(\mathbf{w}^{(i)}), (i = 1, 2, \dots, n) \quad (1)$$

式中, x 是原始信号, e_1 是第一次分解信号的噪声标准差, $\mathbf{w}^{(i)}$ 是一系列高斯白噪声。

2) 第一个残差:

$$r_1 = (X_1^{(i)} - E_1(X_1^{(i)})) \quad (2)$$

其中: (\cdot) 表示求平均值。

3) 求得第一个分量,原始信号 x 减去第一次计算得到残差 r_1 :

$$\mathbf{IMF}_1 = x - r_1 \quad (3)$$

4) 将第二残差估计为一系列的 $r_1 + e_2 E_2(\mathbf{w}^{(i)})$ 均值并得到第二个分量:

$$\mathbf{IMF}_2 = r_1 - r_2 = r_1 - \{[r_1 + e_2 E_2(\mathbf{w}^{(i)})]\} \quad (4)$$

5) 第 k 阶模态的残差 r_k :

$$r_k = [X_k^{(i)} - E_k(X_k^{(i)})] \quad (5)$$

6) 计算第 k 阶模态分量 \mathbf{IMF}_k :

$$\mathbf{IMF}_k = r_{k-1} - r_k \quad (6)$$

式中, k 为 \mathbf{IMF} 的总数。

7) 返回 5) 计算 r_{k+1} 。

ICEEMDAN 分解方法可以通过向信号中注入经过 EMD 分解后的独特白噪声,有效减少 IMF 分量中的残余噪声。并且 IMF 的获取是由添加白噪声后的信号局部均值与残余分量的差值得得,不仅能够减小分解结果内部的残留噪声,还可以有效地防止由于分量叠加而产生的伪模态,增强了信噪比,能够更精确地将信号分解为不同的 IMF 分量,提高了信号分析能力。

1.2 LSSVM 原理与模型优化

1.2.1 LSSVM 原理

支持向量机理论^[17]学习速度快,受到输入数据维数影响小,参数选取简单等优点,为实现滚动轴承故障诊断提供了关键条件。各种种群算法对 SVM 的超参数进行寻优,但并没有考虑到种群迭代过程中移动步长的调整,会导致前期收敛缓慢,有的后期无法收敛,使最终结果陷入局部最优值。针对这一问题,本文采用 LSSVM 作为一种基于统计理论的改进型 SVM,可以解决线性和非线性多变量校准

并相对快速地解决多变量校准问题。二者之间的主要区别在于 LSSVM 的损失函数为平方误差而不是非负误差。并且与传统的 SVM 不同, 使用等式约束而不是不等式约束, 因此比 SVM 具有更好地解决更高计算问题的能力。此外, 使用 LSSVM 可以更轻松、更有效地处理大规模问题。因此, 它已成功地应用于函数逼近和模式识别的许多领域, 具有很高的准确性和更高的泛化能力。LSSVM 试图最小化泛化误差的上限而不是经验误差, 并且与人工神经网络方法相比, 它可以在相同的训练条件下提供更可靠和改进的泛化性能。其理论体系先进完备, 可以将二次优化问题的解决转化为求解线性方程组^[18], 从而极大地简化问题, 目前已在数据回归、模式识别、时间序列预测等方面得到成功应用。

具体的 LSSVM 建模过程如下:

给定样本 $\{(x_i, y_i), i=1, 2, \dots, m\}$ 给定样本 $x_i \in \mathbf{R}^d$ 为输入向量, $y_i \in \mathbf{R}$ 为输出值, m 表示样本数量。LSSVM 可以通过样本训练后得到函数 f , 通过 f 即可将输入转化为输出, 对数据进行分类或者回归。首先通过非线性映射 $x \rightarrow \varphi$ 将输入转换到高维特征空间, 然后在特征空间实现分类, 得到线性回归函数表示为式 (7):

$$f(x) = \mathbf{W}^T \varphi(x) + b \quad (7)$$

式中, \mathbf{W} 为权值向量; $\varphi(x)$ 是非线性映射函数; b 是偏差量。

由 LSSVM 优化后的目标见式 (8):

$$\begin{cases} \min J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{1}{2} c \sum_{i=1}^m \xi_i^2 \\ s. t. y_i = \omega^T \varphi(x) + b + \xi_i, i = 1, 2, \dots, m \end{cases} \quad (8)$$

其中: ξ_i 为误差变量; c 为正则化参数, 引入 Lagrange 函数, 由式 (8) 可得 (9):

$$L(\omega, a, b) = \frac{1}{2} \|\omega^T\|^2 + \frac{1}{2} c \sum_{i=1}^m \xi_i - \sum_{i=1}^m a_i [\omega^T \varphi(x) + b + \xi_i - y_i] \quad (9)$$

满足 KKT 条件是, 由此得到式 (10):

$$\begin{cases} \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^m a_i y_i = 0 \\ \frac{\partial L}{\partial a_i} = 0 \Rightarrow y_i = \omega^T \varphi(x) + b + \xi_i \\ \frac{\partial L}{\partial \omega} = 0 \Rightarrow \omega = \sum_{i=1}^m a_i \varphi(x) \\ \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow a_i = c \xi_i, i = 1, 2, \dots, m \end{cases} \quad (10)$$

消去 ω 和 ξ_i , 式 (10) 可以转化为 (11):

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & K(x_1, x_1) & \cdots & K(x_1, x_m) \\ \vdots & \vdots & \cdots & \vdots \\ 1 & K(x_m, x_1) & \cdots & K(x_m, x_m) + \frac{1}{c} \end{bmatrix} \times \begin{bmatrix} b \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix} \quad (11)$$

LSSVM 回归估计在引入核函数 $\mathbf{K}(x_i, x)$ 后可最终表示为式 (12):

$$f(x) = \sum_{i=1}^m \alpha_i \mathbf{K}(x_i, x) + b \quad (12)$$

式中, α_i 和 b 通过式 (16) 得出; $\mathbf{K}(x_i, x)$ 表示核函数。为了保证模型的运算效率, 我们选择了径向基函数 (RBF) 作为误差模型的核函数。RBF 核函数具有训练速度快、计算精度高等优点。误差模型如式 (13) 所示:

$$\mathbf{K}(x_i, x) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right) \quad (13)$$

式中, σ 表示核函数的宽度系数。

LSSVM 模型在使用过程中极易受到正则化参数 c 和核函数参数 σ 的影响, 正则化参数 c 的大小会影响 SVM 回归曲线的效果, 若 c 较小, 易发生欠学习的现象, 回归曲线趋于平坦, 训练误差变大; 若 c 较大, 易发生过学习的现象, 训练误差变小, 导致推广能力变差。核函数的宽度系数 σ 代表了支持向量的相关程度, 若 σ 较小, 支持向量的相关性较低, 易发生局部优化的情况, 造成 SVM 过训练的现象; 若 σ 较大, 支持向量间的联系较深, 使预测难以达到精度要求, 造成 SVM 欠训练的现象。因此, 建立 LSSVM 模型时有必要对这两个参数进行优化。所以, 本文使用 PSO 算法对 LSSVM 核函数的宽度系数 σ 和正则化参数 c 进行优化。

1.2.2 基于 PSO 的 LSSVM 参数优化

粒子群算法 (PSO) 算法模拟鸟群的觅食行为, 通过群体中个体之间的协作和信息共享, 寻找离食物最近的鸟的邻近区域, 是一种典型的随机搜索策略^[19]。与遗传算法相比, PSO 在解决各种优化问题方面表现出一些较为突出的特点, 因具有实现容易、精度高、收敛快等优点, 被广泛应用于多目标优化问题。在遗传算法中, 一旦种群变化很可能会破坏之前学习到的问题特性, 而 PSO 保留了种群中所有粒子所学习到的良好特性^[20]。PSO 的基本原理是通过模拟生物的本能行为, 可以解释为一个 G 维空间的优化问题, $\mathbf{U} = (U_1, U_2, \dots, U_m)$ 为 m 个粒子组成的种群, 其中第 i 个粒子的位置为 $U_i = (U_{i1}, U_{i2}, \dots, U_{iG})$ 。每个粒子的适应值是由目标函数根据当前位置和速度计算出来, $\mathbf{V}_i = (V_{i1}, V_{i2}, \dots, V_{iG})$ 是第 i 个粒子的速度, $\mathbf{P}_i = (P_{i1}, P_{i2}, \dots, P_{iG})$ 为每个粒子的极值, 群体极值为 $\mathbf{P}_a = (P_{a1}, P_{a2}, \dots, P_{aG})$, 表示粒子速度和位置的迭代公式如下:

$$\mathbf{V}_{id}^{k+1} = \omega \mathbf{V}_{id}^k + c_1 r_1 (P_{id}^k - U_{id}^k) + c_2 r_2 (P_{ad}^k - U_{id}^k) \quad (14)$$

$$U_{id}^{k+1} = U_{id}^k + V_{id}^{k+1} \quad (15)$$

式中, ω 为惯性权重, c_1 、 c_2 为加速常数, r_1 、 r_2 为在 $0 \sim 1$ 之间的随机数。

惯性权重 ω 对 PSO 算法的收敛性有关键性的作用, 动态的惯性权重 ω 对全局寻优的效果更好, 通常采用线性递减方法, 其计算公式如下:

$$\omega = \omega_k - \frac{\omega_k - \omega_{ek}}{k} \quad (16)$$

式中, k_m 为最大迭代数, ω_r 为初始惯性权重, ω_k 为最大惯性权重。

PSO 优化超参数的步骤如下:

1) 设置 PSO 的种群大小、迭代次数、优化参数个数和优化范围等各个参数, 并选择适应度函数为训练集的诊断精度;

2) 随机初始化每个粒子的速度和位置, 并计算出初始的适应度值;

3) 计算每个粒子的适应度值以确定全局最优, 并检查是否达到精度要求。如果达到要求, 则停止并输出相关的超参数。如果未达到要求, 则继续进行迭代循环;

4) 对于每个粒子, 更新其速度和位置, 并根据边界条件限制其位置和速度的范围;

5) 对每个粒子进行适应度值的计算, 并更新它们各自的历史最优位置;

6) 更新整体的最佳位置, 直到达到精度要求, 然后停止 PSO 算法, 并输出相应的超参数值。

1.2.3 PSO-LSSVM 预测模型

本文使用 PSO 算法对 LSSVM 核函数的宽度系数 σ 和正则化参数 c 进行优化, 选择最优的参数, 达到最优的预测结果。PSO 优化 LSSVM 的基本流程如下。

1) 对历史数据进行归一化和标准化处理, 形成训练样本矩阵。

2) 初始化 PSO 参数, 将 LSSVM 参数 σ 和 c 映射成粒子。

3) 根据种群的预测结果来评估每个粒子的适应度, 找到个体的最佳位置和种群的最佳位置, 然后更新粒子的速度和位置, 生成新的种群。

4) 将新种群的个体最优位置和种群最优位置与上一代进行比较, 若适应值更小则更新最优位置和速度; 反之则保持不变。

5) 判断是否满足寻优结束条件, 满足则停止迭代, 否则转到 4)。

6) 将最优参数 (σ, c) 赋值给 LSSVM 预测模型, 输入测试样本, 得到基于 PSO-LSSVM 模型预测的结果。

2 分量筛选与特征选取

ICEEMDAN 方法将原始信号分解为多个 IMF 成分, 但并非所有 IMF 成分都能准确反映原始信号的特性。为了准确识别故障类型, 需要去除这些不准确的成分。在概率统计中, 相关系数可以用来衡量信号与 IMF 之间的相似程度, 较高的相关系数表示信号与 IMF 更为相似, 而与虚假分量和噪声成分的相关系数较低。因此, 在决定权重时, 可以根据信号和 IMF 权重之间的相关性来进行选择^[21]。另外, 方差贡献率可以表示因子的重要性, 根据统计分析中的相关理论, 其数值越大, 对应因子的重要性就越大。本文采用相关系数-方差贡献率作为筛选准则。

2.1 相关系数

相关系数是表征两者之间的相关联程度的统计量^[22],

相关系数越大表示二者相关性越强。将原信号与经过 ICEEMDAN 分解得到的各 IMF 分量做相关性分析, 计算结果越大表示该 IMF 分量包含的原始信号成分越多^[23]。

首先计算原信号与各个 IMF 分量的自相关系数:

$$R_x(m) = \frac{1}{N} \sum_{i=0}^{N-1} x(i)x(i+m) \quad (17)$$

自相关函数进行归一化, 求 $R_{IMF_1}(m)$ 、 $R_{IMF_2}(m)$ 、 \dots 、 $R_{IMF_n}(m)$ 与原始信号的自相关函数 $R_x(m)$ 的相关系数, 其定义为:

$$\rho(j) = \frac{\sum_{i=1}^{2N-1} R_{IMF_j}(i)R_x(i)}{\sqrt{\sum_{i=1}^{2N-1} R_{IMF_j}^2(i) \sum_{i=1}^{2N-1} R_x^2(i)}} \quad (18)$$

式中, N 为信号采样点数, j 为第 j 个 IMF 分量。

2.2 方差贡献率

各信号成分的相对重要程度用方差贡献率来表示。本文提出通过计算各阶 IMF 方差贡献率来说明各阶 IMF 份量的重要程度, 因子分析法的统计意义表明, 方差贡献率可以确定因子的相对重要性^[24]。方差贡献率是一个统计量, 用来表示相对重要的因素。信号经过 ICEEMDAN 分解为 $IMF_1, IMF_2, \dots, IMF_n$, 定义各分量与原信号的方差贡献率计算如下:

$$mseb(i) = \frac{\frac{1}{N} \sum_{j=1}^N h_i^2(j) - \left[\frac{1}{N} \sum_{j=1}^N h_i(j) \right]^2}{\sum_{i=1}^n \left\{ \frac{1}{N} \sum_{j=1}^N h_i^2(j) - \left[\frac{1}{N} \sum_{j=1}^N h_i(j) \right]^2 \right\}} \quad (19)$$

式中, N 为信号的长度。

2.3 奇异谱熵

信号的复杂性与其包含的信息量成正比, 同时带来的不确定性和随机性也会增加^[25]。奇异谱熵反映了时域信号在奇异谱划分下各模式的不确定程度^[26], 它可以较好地地从有限时间序列中获取有用的信息。通过 ICEEMDAN 对信号进行分解, 得到 i 阶 IMF 分量, 再对分量进行奇异值分解, 记为 μ_i , 则信号的奇异谱熵表示为:

$$E_{se} = - \sum_{i=1}^M q_i \log q_i \quad (20)$$

式中, $q_i = \frac{\mu_i}{\sum_{i=1}^M \mu_i}$ 为第 i 个 IMF 分量的奇异值在整个奇异值谱中的比重。

2.4 故障诊断方法流程

图 1 展示了一种滚动轴承故障诊断方法的流程, 该方法基于 ICEEMDAN 分解算法、提取奇异谱熵特征算法以及结合 PSO-LSSVM 分类器模型算法。图 2 为特征处理与模式识别过程图。

1) 信号分解。利用 ICEEMDAN 方法分别对降噪后的滚动轴承振动信号和声发射信号进行分解, 得到一系列的 IMF 分量;

2) 特征提取。根据相关系数与方差贡献率的变化趋势分别筛选能表征原信号特征的 IMF 分量并计算奇异谱熵值;

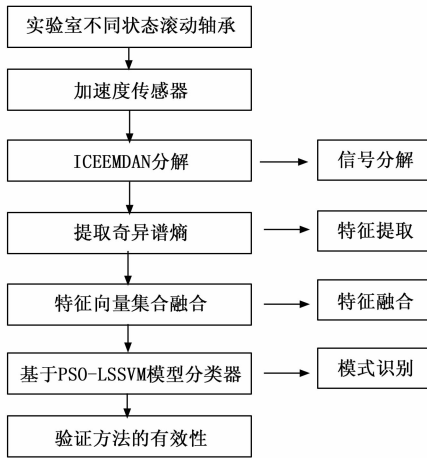


图 1 故障诊断流程图

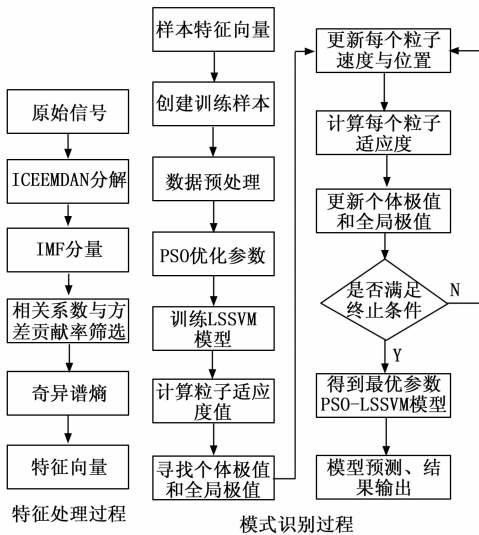


图 2 特征处理与模式识别过程图

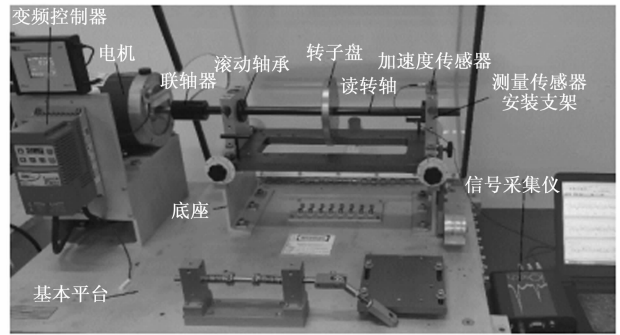
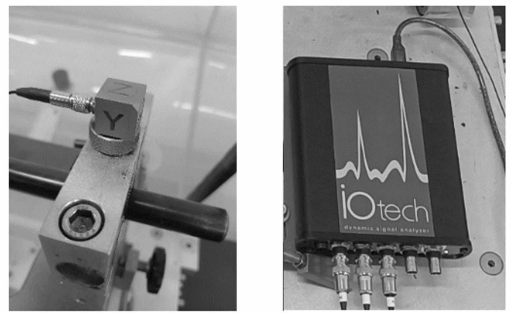


图 3 滚动轴承故障检测试验平台



(a) 加速度传感器 (b) IOtech640U型动态信号分析仪

图 4 振动信号采集系统

ICEEMDAN 分解算法分解降噪处理后的滚动轴承振动信号得到如图 5 所示, 不同状态下振动信号的 ICEEMDAN 分解图 (前 10 阶分量)。

从图 5 中观察到, 随着阶数的增加, 频率逐渐降低, 所包含的原始信号特征信息变得更少, 需要选择那些含有原始信号特征的 IMF 分量, 如果直接分析 IMF, 则难以得到与原信号有关的有效性质。通常信号在 ICEEMDAN 分解过程中得到一系列 IMF 分量中存在不能表征原始信号特征信息的虚假分量, 若直接进行特征提取会造成工作量繁重且会影响分析结果。通过计算每个分量的相关系数, 如表 1 所示。

表 1 不同故障类型振动信号 IMF 分量的相关系数值

状态类型	IMF ₁	IMF ₂	IMF ₃	IMF ₄	IMF ₅
外圈故障	0.17	0.22	0.245	0.14	0.1
内圈故障	0.2	0.36	0.35	0.085	0.03
滚动体故障	0.21	0.25	0.192	0.12	0.094
状态类型	IMF ₆	IMF ₇	IMF ₈	IMF ₉	IMF ₁₀
外圈故障	0.09	0.004	0.003	0.003	0.001
内圈故障	0.027	0.025	0.001	0.002	0.001
滚动体故障	0.008	0.01	0.045	0.003	0.001

将原始信号与相关系数最大的 IMF 分量分别进行包络谱分析, 对比分析两者图像的冲击特征。图 6 为振动信号外圈故障包络谱, 图 7 为振动信号内圈故障包络谱, 图 8 为振动信号滚动体故障包络谱。

3) 特征融合。将不同信号提取的特征值集合融合在一起构造一个新的特征向量集合;

4) 模式识别。通过 PSO-LSSVM 模型, 将特征向量输入其中进行训练和模式识别, 同时与 SVM 和 LSSVM 模型的识别准确率进行对比, 以此验证模型的有效性。

3 实验研究

由于滚动轴承在运行过程中会产生振动, 其中会包含故障类型的特征信息。通过设计数据采集系统, 采集不同状态类别下的实验信号, 为后续判别滚动轴承在运行过程中是否存在缺陷提供了数据支持。本文实验如图 3 所示, 选择 SQI-MFS 机械故障模拟试验台进行。试验采集轴承不同状态下振动信号。振动信号测试仪器如图 4 所示, 由 1A314E 型加速度传感器、IOtech640U 型动态信号分析仪所组成, 采样频率设置为 12 800 Hz, 采样点数为 12 800。

3.1 数据处理与特征提取

通过试验平台, 在轴承正常状态、外圈故障、内圈故障和滚动体故障 4 种状态下各采集 50 组振动信号。利用

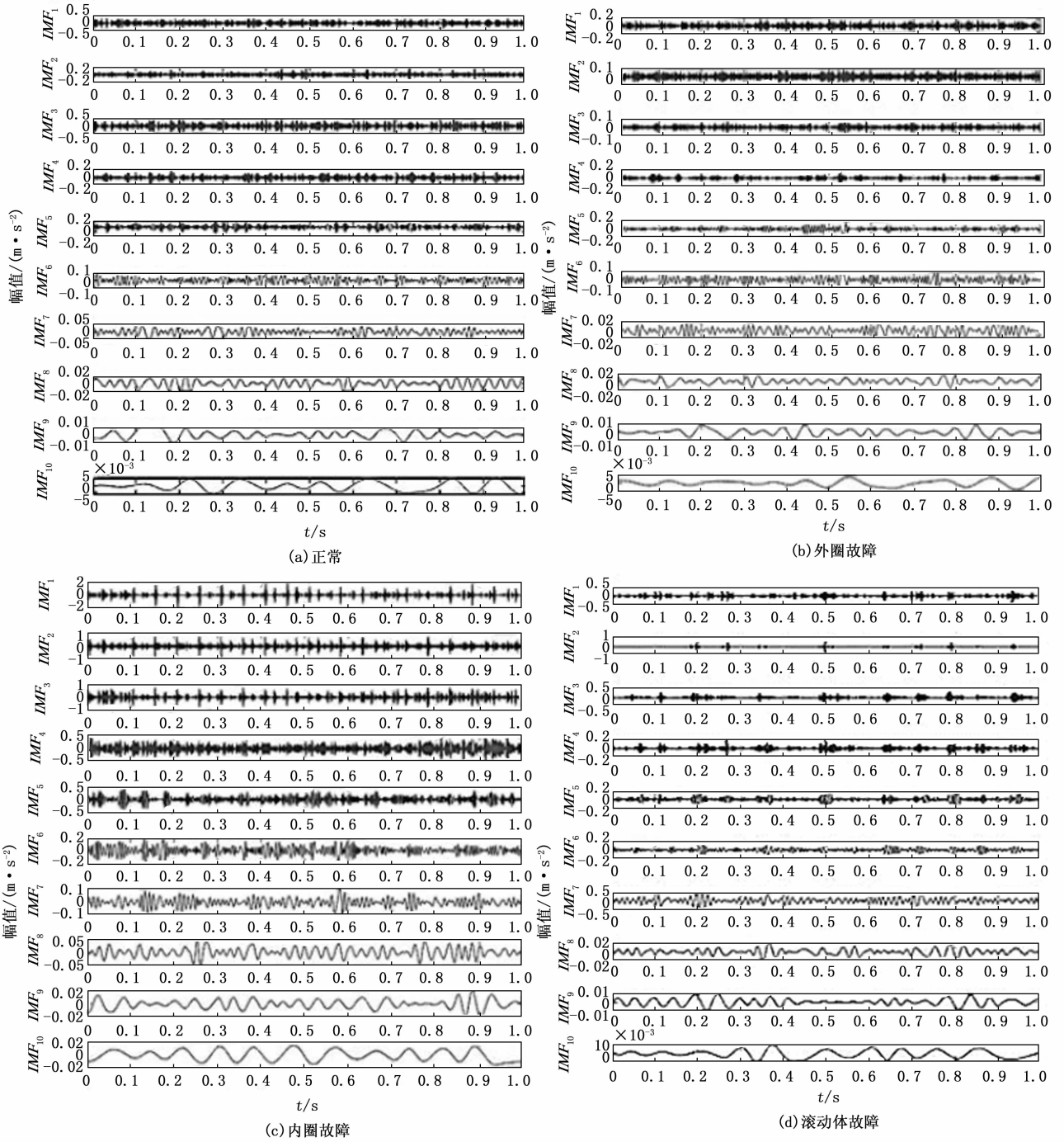
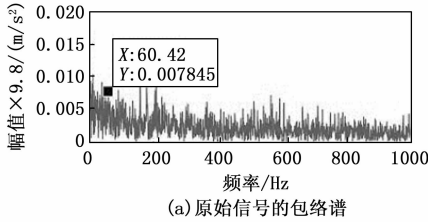


图 5 振动信号 ICEEMDAN 分解

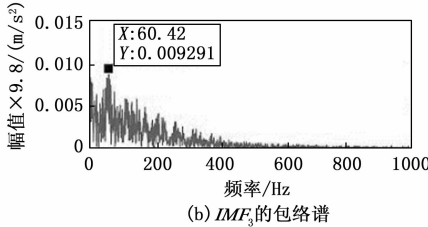
由上述不同类型信号的包络谱图可以看出，有原始信号的故障特征表现不明显，干扰分量较多，只能勉强找到故障特征频率。经过处理后的信号再进行包络谱分析环境背景干扰分量得到了很好的抑制，故障冲击表现更加直观。此外从内圈故障分析图中可以看出故障频率特征有转频干扰，这可能是内圈的结构所致，它不停地随着旋转轴一起转动导致位置不固定，且中间隔着滚动体和外圈，这就可能导致振动传感器无法准确地对故障信息进行采集。

综合对比相关系数与方差贡献率两参数，可甄别 IMF 有效分量。通过对相关系数及其各自方差贡献率的甄别，把各 IMF 分量相关系数及方差贡献率绘制如图 9 所示。

图 9 体现出 IMF 分量在各条件下对原信号相关系数及方差贡献率之间的变化关系，其中保留值较高的 IMF 分量。大多数冲击信号都集中分布在 IMF₁~IMF₅ 考虑选择 IMF₁~IMF₅ 个分量，将剩余 IMF 个分量剔除。从中可以看出前五阶 IMF 的相关系数-方差贡献率的变化波动较大，其余

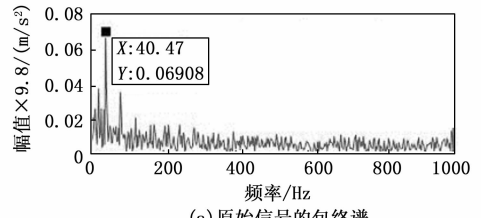


(a) 原始信号的包络谱

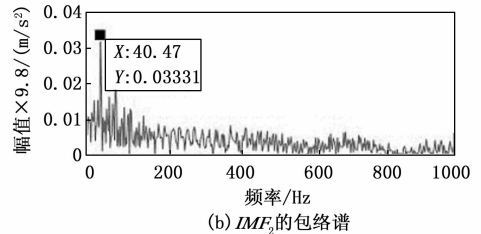


(b) IMF_3 的包络谱

图 6 振动信号外圈故障包络谱

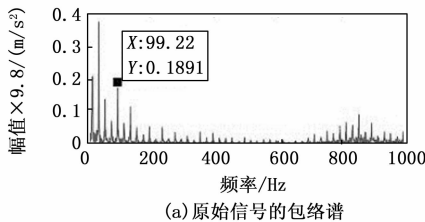


(a) 原始信号的包络谱

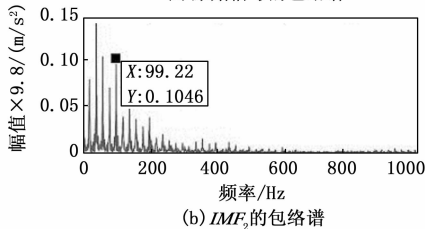


(b) IMF_2 的包络谱

图 8 振动信号滚动体故障包络谱



(a) 原始信号的包络谱



(b) IMF_2 的包络谱

图 7 振动信号内圈故障包络谱

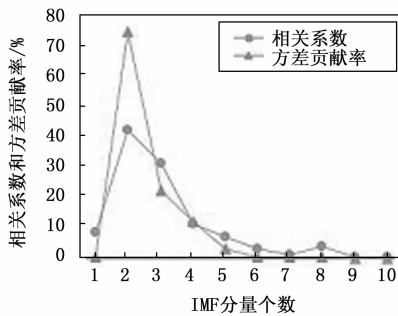
分量无太大波动趋势甚至接近于零, 因此对于振动信号提取前五阶分量的奇异谱熵值作为特征值, 其部分结果分别如表 2 所示。

表 2 4 种状态振动信号 $IMF_1 - IMF_5$ 奇异谱熵

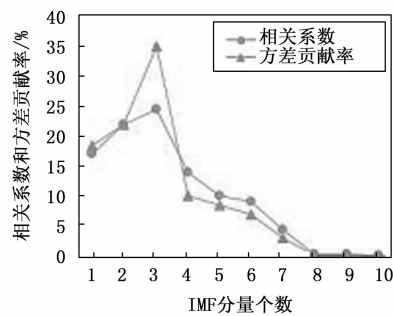
状态类型	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5
正常	7.477 2	6.690 3	6.075 4	5.408 2	2.206 2
外圈故障	6.960 8	5.368 3	4.156 4	2.499 4	1.451 2
内圈故障	6.315 4	4.944 8	3.699 7	2.037 1	1.244 7
滚动体故障	6.970 2	6.035 8	5.686 9	4.782 1	4.738 7

3.2 故障模式识别

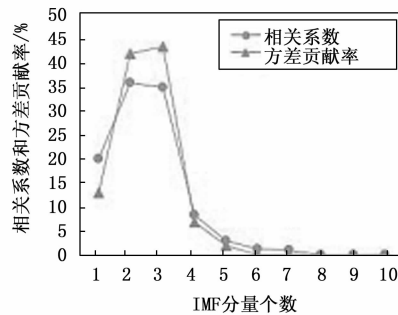
将振动信号 $IMF_1 \sim IMF_5$ 分量的奇异谱熵构建特征向量集合。为了实现滚动轴承的故障智能识别, 采用抽取不



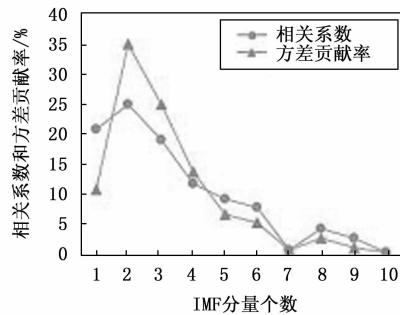
(a) 正常



(b) 外圈故障



(c) 内圈故障



(d) 滚动体故障

图 9 振动信号相关系数及方差贡献率

同状态下的特征向量集合的方法，其中 70% 被选为训练样本，作为基于 PSO-LSSVM 模型的输入。剩余的 30% 作为测试样本。在训练阶段，我们将 1 用于表示正常状态的特征标号，将 2 用于表示外圈故障特征的标号，将 3 用于表示内圈故障特征的标号，将 4 用于表示滚动体故障特征的标号。为了评估本文提出方法的优势，使用了基于 SVM 和 LSSVM 的两种模式识别方法。

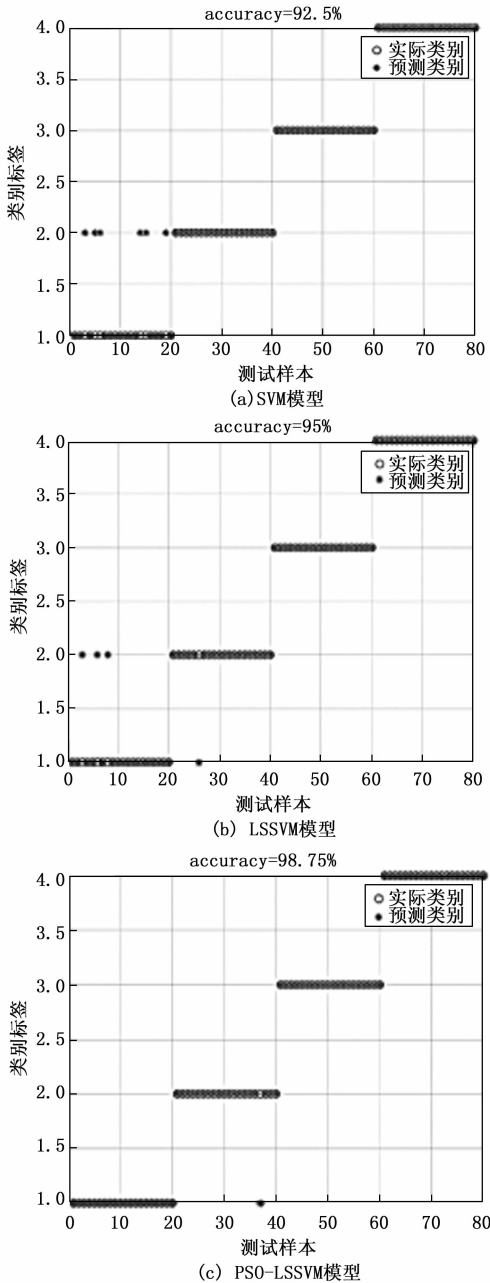


图 10 模型分类器识别图

表 3 模型分类器结果对比

分类模型	识别准确率/%	耗时/s
SVM	92.5	41.7
LSSVM	95.0	29.3
PSO-LSSVM	98.75	18.4

根据图 10，SVM 模型没有识别出 6 组样本，并且算法诊断耗时为 41.7 s。而 LSSVM 模型也没有识别出 4 组样本，算法诊断耗时为 29.3 s。在 PSO-LSSVM 模型中，仅有 1 组测试对象未被正确识别，整体的识别准确率达到 98.75%。此外，该算法的运行时间只需 18.4 s。与之相比，其他模型的识别结果详见表 3。与 SVM 和 LSSVM 模型相比，本文的 PSO-LSSVM 模型在识别准确率和算法运行时间上都表现出更好的性能，显示出了令人满意的识别效果。

为进一步验证该算法的有效性，与文献 [9, 15-17, 25] 特征提取方法的诊断准确率进行对比，7 种方法的诊断准确率如表 4 所示。由表 4 知，本文方法在考虑多种故障类型时，在保证算法较简单的优势下，仍然得到更优的准确率。

表 4 模型分类器结果对比

方法	特征参数	分类模型	识别准确率/%
文献 [9]	EMD-SVD	SVM	94
文献 [17]	CEEMDAN 排列熵	SVM	93.1
文献 [25]	CEEMDAN 能量矩	PSO-SVM	96
文献 [15]	EEMD	MFOA-SVM	98.44
本文方法	ICEEMDAN 奇异谱熵	LSSVM	95.0
本文方法	ICEEMDAN 奇异谱熵	PSO-LSSVM	98.75

4 结束语

针对滚动轴承疲劳故障振动信号具有能量弱、特征稀疏的特点，提出基于 ICEEMDAN-PSO-LSSVM 模型的滚动轴承故障识别诊断方法，通过前期故障机理研究、国内外研究现状，确定研究方法原理，利用 ICEEMDAN 进行信号分解，通过分量筛选后的奇异谱熵作为特征值，输入 PSO-LSSVM 模型中进行训练与识别，最终得到最优模型。本文通过故障模拟实验研究进行对比论证，现将其结论整理如下：

- 1) 通过使用 ICEEMDAN 分解技术，对采集信号进行分解，得到一系列 IMF 分量。然后，通过相关系数-方差贡献率准则来筛选出最能够表征原始信号状态的分量。这样一来，就能够实现选择那些包含原始信号特征的 IMF 分量。
- 2) 将满足标准分量的奇异谱熵作为特征向量集合，分别输入到 SVM 模型、LSSVM 模型和 PSO-LSSVM 模型中进行识别。结果显示，基于 SVM 模型的准确率为 92.5%，耗时 41.7 s，基于 LSSVM 模型的准确率为 95%，耗时 29.3 s，基于 PSO-LSSVM 模型的准确率为 98.75%，耗时 18.4 s，此外，PSO-LSSVM 模型还能在更短的时间内完成分类任务。这说明 PSO-LSSVM 模型在模式识别方面表现更出色。

参考文献；

[1] 赵 云, 宿 磊, 李 可, 等. 基于改进 MPE 与 KELM 的滚动轴承故障诊断 [J]. 噪声与振动控制, 2022, 42 (1): 125-131.

[2] PEETERS C, GUILLAUME P. A comparison of cepstral edi-

- ting methods as signal pre-processing techniques for vibration-based bearing fault detection [J]. *Mechanical Systems and Signal Processing*, 2017, 9 (1): 354 - 381.
- [3] MOHSEN N, MICHAEL D. Dynamic modeling of rolling element bearings with surface contact defects using bond graphs [J]. *ASME, Journal of Tribology*, 2011, 133 (1): 1 - 12.
- [4] 尹文哲, 夏虹, 彭彬森, 等. 基于 FFT 和 CNN 的滚动轴承故障诊断方法 [J]. *应用科技*, 2021, 48 (6): 97 - 101.
- [5] CRISTIÁN M V, CHRISTOPH H. A method for reduction of acoustic emission (AE) data with application in machine failure detection and diagnosis [J]. *Mechanical Systems and Signal Processing*, 2017, 97: 44 - 58.
- [6] CHAI M, ZHANG Z, DUAN Q. A new qualitative acoustic emission parameter based on Shannon's entropy for damage monitoring [J]. *Mechanical Systems and Signal Processing*, 2018, 100: 617 - 629.
- [7] CAVACECE M, INTROINI A. Analysis of damage of ball bearings of aeronautical transmissions by auto-power spectrum and cross-power spectrum [J]. *Journal of Vibration and Acoustics*, 2002, 124 (2): 180 - 185.
- [8] HUANG N E, SHEN Z, LONG S R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. *Royal Society of London Proceedings*, 1998, A (454): 903 - 995.
- [9] 吴虎胜, 吕建新, 来凌红, 等. 基于 EMD-SVD 模型和 SVM 滚动轴承故障模式识别 [J]. *噪声与振动控制*, 2011, 31 (2): 89 - 93.
- [10] WU Z, HUANG N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. *Advances in Adaptive Data Analysis*, 2009, 1 (1): 1 - 41.
- [11] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G. A complete ensemble empirical mode decomposition with adaptive noise [C] // *IEEE International Conference on Acoustics Speech and Signal Processing*. IEEE, 2011: 4144 - 4147.
- [12] 王新, 闫文源. 基于变分模态分解和 SVM 的滚动轴承故障诊断 [J]. *振动与冲击*, 2017, 36 (18): 252 - 256.
- [13] 方必武, 刘涤尘, 王波, 等. 基于小波变换和改进萤火虫算法优化 LSSVM 的短期风速预测 [J]. *电力系统保护与控制*, 2016, 44 (8): 37 - 43.
- [14] 杨武帮, 高丙朋, 陈飞, 等. 基于变分模态分解和 PSO-SVM 的起重机齿轮箱故障诊断 [J]. *机械传动*, 2021, 45 (4): 105 - 111.
- [15] 何青, 褚东亮, 毛新华. 基于 EEMD 和 MFFOA-SVM 滚动轴承故障诊断 [J]. *中国机械工程*, 2016, 27 (9): 1191 - 1197.
- [16] 李铭, 何毅斌, 马东等. 基于 ICEEMDAN-MPE 和 AO-LSSVM 的滚动轴承故障诊断 [J]. *电子测量技术*, 2022, 45 (23): 66 - 71.
- [17] 胡超, 杨妍, 王松涛, 等. 基于 QAPSO-SVM 的航空发动机故障诊断 [J]. *热能动力工程*, 2020, 35 (12): 40 - 46.
- [18] 张利彪, 周春光, 马铭, 等. 基于粒子群算法求解多目标优化问题 [J]. *计算机研究与发展*, 2004 (7): 1286 - 1291.
- [19] 刘福政, 高军伟. 应用 CEEMDAN 能量矩和 PSO-PNN 的轴承故障诊断 [J]. *现代制造工程*, 2020 (11): 126 - 130.
- [20] 张霓, 曾乐襄, 何熊熊, 等. 基于滚动时域粒子群优化的视频去雾算法 [J]. *控制与决策*, 2021 (9): 2218 - 2224.
- [21] 张俐, 袁玉宇, 王枫. 基于最大相关信息系数的 FCBF 特征选择算法 [J]. *北京邮电大学学报*, 2018, 41 (4): 86 - 90.
- [22] 李清泉, 秦冰阳, 司雯, 等. 混合粒子群优化小波自适应阈值估计算法及用于局部放电去噪 [J]. *高电压技术*, 2017, 43 (5): 1485 - 1492.
- [23] 李火坤, 刘世立, 魏博文, 等. 基于方差贡献率的泄流结构多测点动态响应融合方法研究 [J]. *振动与冲击*, 2015, 34 (19): 181 - 191.
- [24] 殷红, 陈强, 彭珍瑜. 传感器优化布置的齿轮箱轴承故障特征提取 [J]. *噪声与振动控制*, 2020, 40 (4): 67 - 72.
- [25] 周建民, 王发令, 张臣臣, 等. 基于特征优选和 GA-SVM 的滚动轴承智能评估方法 [J]. *振动与冲击*, 2021, 40 (4): 227 - 234.
- [26] 巩孟林, 陈卫, 钟也磐, 等. 基于稀疏表示和 SVM 的航空发动机故障诊断 [J]. *噪声与振动控制*, 2017, 37 (3): 162 - 167.
- [20] NAM J, PAN S J, KIM S. Transfer defect learning [C] // *2013 35th International Conference on Software Engineering (ICSE)*. IEEE, 2013: 382 - 391.
- [21] BERNHARD S, JOHN P, THOMAS H, et al. Analysis of representations for domain adaptation [C] // *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, 2007: 137 - 144.
- [22] ZHANG J, LI W, OGUNBONA P. Joint geometrical and statistical alignment for visual domain adaptation [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1859 - 1867.
- [23] KARIM S, WARNARS H L H S, GAOL F L, et al. Software metrics for fault prediction using machine learning approaches: A literature review with PROMISE repository dataset [C] // *2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*. IEEE, 2017: 19 - 23.
- [24] D'AMBROSIO M, LANZA M, ROBBES R. Evaluating defect prediction approaches: a benchmark and an extensive comparison [J]. *Empirical Software Engineering*, 2012, 17: 531 - 577.
- [25] LI Z, JING X Y, WU F, et al. Cost-sensitive transfer kernel canonical correlation analysis for heterogeneous defect prediction [J]. *Automated Software Engineering*, 2018, 25: 201 - 245.
- [26] YUAN Z, CHEN X, CUI Z, et al. ALTRA: Cross-project software defect prediction via active learning and tradaboost [J]. *IEEE Access*, 2020, 8: 30037 - 30049.