

DOI:10.3969/j.issn.1001-4551.2019.06.013

基于 EEMD 和 CS-SVM 的 滚动轴承故障诊断研究*

梁治华¹, 曹江涛¹, 姬晓飞^{2*}

(1. 辽宁石油化工大学 信息与控制工程学院, 辽宁 抚顺 113001;
2. 沈阳航空航天大学 自动化学院, 辽宁 沈阳 110136)

摘要:针对数据驱动的滚动轴承故障诊断大多采用支持向量机进行分类,而传统支持向量机的分类方法容易陷入局部最优,无法准确进行故障诊断的问题,对滚动轴承振动信号的特征选择和支持向量机的优化方法进行了研究。分析了粒子群算法优化支持向量机和遗传算法优化支持向量机的不足;基于莱维飞行的布谷鸟搜索算法,引入了一种对支持向量机的参数进行寻优的方法,用于提高滚动轴承故障诊断的识别准确率;该方法首先使用集合经验模态分解对信号数据进行了处理,然后计算本征模态函数的均方根作为特征向量,输入布谷鸟搜索算法优化的支持向量机;最后进行了训练和测试。研究表明:利用该方法对实测信号进行分析和诊断,可以准确地识别故障发生的位置以及严重程度;通过与传统优化方法进行对比,验证了该算法的优越性。

关键词:集合经验模态分解;布谷鸟搜索;支持向量机;故障诊断

中图分类号:TH133.33;TP29

文献标志码:A

文章编号:1001-4551(2019)06-0622-06

Fault diagnosis of rolling bearing based on EEMD and CS-SVM

LIANG Zhi-hua¹, CAO Jiang-tao¹, JI Xiao-fei²

(1. School of Information and Control Engineering, Liaoning Shihua University, Fushun 113001, China;
2. School of Automation, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: Aiming at the problem that in the optimization of support vector machines, data-driven rolling bearing fault diagnosis, most of them are classified by support vector machine, however, the traditional support vector machine classification method is easy to fall into local optimum, and it is impossible to carry out fault diagnosis accurately, the feature selection of rolling bearing vibration signal and the optimization method of support vector machine were studied. The disadvantages of the support vector machine optimization by using genetic algorithm and particle swarm optimization were pointed out. In order to improve the accuracy of rolling bearing fault diagnosis, a cuckoo search algorithm based on Levy flight was introduced to find the optimal parameters of support vector machine. First, ensemble empirical mode decomposition was used to process signal data, and then the root mean square of the intrinsic mode function was put into the support vector machine optimized by cuckoo search algorithm to train and test this model. The results indicate that the proposed method can be used to analyze and diagnose the measured signals, and the location and severity of the faults can be accurately identified. The superiority of the algorithm is verified by comparison with traditional optimization methods.

Key words: ensemble empirical mode decomposition(EEMD); cuckoo search(CS); support vector machine(SVM); fault diagnosis

收稿日期:2018-11-15

基金项目:辽宁省科技公益研究基金资助项目(2016002006);辽宁省自然科学基金资助项目(201602557);辽宁省教育厅科学研究服务地方项目(L201708)

作者简介:梁治华(1987-),男,辽宁抚顺人,硕士研究生,主要从事机械故障诊断方面的研究。E-mail:67897057@qq.com

通信联系人:姬晓飞,女,副教授,硕士生导师。E-mail:jixiaofei7804@126.com

0 引言

旋转机械作为一种十分重要的动力装置,广泛应用于水处理、采矿、采石等行业。当今的旋转机组的复杂度越来越高,建立智能化的故障诊断系统,是故障诊断研究的主要方向^[1-2]。基于数据驱动的轴承故障诊断主要包括3部分:信号处理、故障特征提取和故障识别。

信号处理和故障特征提取的方法有傅里叶变换(FT)、小波变换(WT)和经验模态分解(EMD)等^[3]。FT是对全信号的频域分析方法,无法确定频率所对应的时刻;WT是对傅里叶变换的改进版本,可以对信号同时进行时域和频域的分析,但是小波函数的选取适应度较差;EMD是一种自适应的数据处理方法,可以将任何类型的信号分解成若干个包含原信号不同时间尺度局部特征的本征模态函数(IMF),可以挑选相关的IMF进行信号分析或重新组合^[4],然而EMD方法存在模态混叠现象^[5]。2009年,吴等人提出了一种改进的EMD方法—集合经验模态分解(EEMD),EEMD可以消除EMD的模态混叠现象,在分析旋转机械振动信号方面优于EMD^[6-8]。

故障识别的方法有人工神经网络(ANN)、决策树(DT)、支持向量机(SVM)等^[9-11]。然而,基于ANN的轴承故障诊断不仅要求样本的数量要足够多,而且学习的速度慢;DT的缺点是如果有很多不相关的变量,DT表现的不好,而且会潜在的引进偏差等;SVM能解决小样本、非线性和高维模式识别的情况,较好地解决了ANN的过拟合和局部最优问题。近十多年来,SVM已经广泛地应用于旋转机械故障诊断和状态监测^[12-15]。

然而,SVM中核函数的选取以及核函数参数的取值会影响SVM对故障分类的准确率,为了提高SVM对故障分类的准确率,李源等^[16]提出了一种基于小波包能量特征与PSO-SVM的轴承故障诊断方法;陈园艺等^[17]提出了一种基于EEMD和PSO-SVM的轴承故障诊断方法;秦波等^[18]提出了一种基于小波包、奇异值分解和GA-SVM的轴承故障诊断方法;蒋恩超等^[19]提出了一种基于小波包和GA-SVM的轴承故障诊断方法。然而使用粒子群算法(PSO)和遗传算法(GA)都会陷入局部最优,因此故障分类准确率不够稳定。

本文将引入一种利用布谷鸟算法优化支持向量机(CS-SVM)的轴承故障诊断方法。

1 故障信号处理及特征提取

1.1 集合经验模态分解

集合经验模态分解的具体步骤如下:

(1) 确定噪声集合 M 的数量和噪声振幅大小。

(2) 将给定振幅数值生成的白噪声 $n_i(t)$ 添加到原始信号 $x(t)$,以生成新的信号,即:

$$x_i(t) = x(t) + n_i(t), \quad i = 1, 2, \dots, m \quad (1)$$

式中: $n_i(t)$ —第 i 次增加的白噪声序列; $x_i(t)$ —第 i 次加入白噪声之后的信号。

(3) 对添加白噪声后的信号进行 EMD 分解,得到 IMF 分量:

$$x_i(t) = \sum_{s=1}^S c_{i,s}(t) + r_{i,s}(t) \quad (2)$$

式中: S —IMF 分量的数量; $r_{i,s}(t)$ —最终残差,是信号的平均趋势; $c_{i,s}(t)$ —IMF($c_{i,1}, c_{i,2}, \dots, c_{i,s}, \dots, c_{i,S}$),是信号从高到低的不同频带的成分信息。

(4) 用不同的白噪声序列重复步骤(2) M 次,获得 IMF 的集合,即:

$$[\{c_{1,s}(t)\}, \{c_{2,s}(t)\}, \dots, \{c_{M,s}(t)\}], s = 1, 2, \dots, S \quad (3)$$

(5) 计算分解的相应 IMF 的集合的均值作为最终结果:

$$c_s(t) = \frac{1}{M} \sum_{i=1}^M c_{i,s}(t), i = 1, 2, \dots, M; s = 1, 2, \dots, S \quad (4)$$

式中: $c_s(t)$ —EEMD 分解的 IMF 分量。

1.2 特征提取

基于 EEMD 轴承故障诊断的特征提取主要依赖于不同故障位置,以及不同故障程度产生的振动信号有所不同,因此,通过提取振动信号的特征,可以较为准确地对不同的故障进行分类。

轴承故障诊断通常选取的特征有:峭度、裕度、方差、均方根、偏度、能量等,文献[20]采用了很多特征组合,本文只使用单一的特征,以此来验证识别方法的准确性。

轴承振动信号的均方根值大多与轴承表面磨损而产生的无规则振动的波形有比较好的相关性,因此,本文通过对不同的特征进行测试总结,选取了不同特征的测试结果,如表1所示。

表1 选取不同特征的测试结果

特征	CS-SVM/(%)	PSO-SVM/(%)	GA-SVM/(%)
峭度	87.727 3	84.545 5	79.545 5
均值	59.545 5	58.181 8	54.090 9
方差	97.727 3	95.454 5	89.090 9
偏度	72.272 7	71.818 2	67.272 7
均方根	100	98.636 4	97.272 8

表1说明:选取均方根特征的识别准确率最高。

因此,本文对 IMF 提取均方根,作为特征用于故障诊断,即:

$$X_{rms} = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}} \quad (5)$$

式中： x_i —样本值； N —样本个数。

2 故障特征分类

2.1 支持向量机

该实验用的 SVM 是 Matlab 下的 Libsvm 模块。SVM 算法最初是为二值分类的问题设计的,当处理多类问题时就需要构造合适的多类分类器^[21]。目前,构造 SVM 多类分类器的方法主要有:

(1)一对多法(1-v-r SVMs)。训练时依次把某个类别的样本归为一类,其他剩余的样本归为另一类,这样 k 个类别的样本就构造出了 k 个 SVM。分类时将未知样本分类为具有最大分类函数值的那类;

(2)一对一法(1-v-1 SVMs)。其做法是在任意两类样本之间设计一个 SVM,因此 k 个类别的样本就需要设计 $k(k-1)/2$ 个 SVM;当对一个未知样本进行分类时,最后得票最多的类别即为该未知样本的类别。Libsvm 中的多类分类就是根据这个方法实现的。

通过选取合适的核函数以及合适的惩罚因子 C 和核函数参数 g ,SVM 可以解决大多数的分类问题,基于径向基(RBF)核函数参数少、适用度广的特点,轴承故障诊断分类大多用 SVM 的 RBF 核函数,所以通过训练集寻找最优的 C 和 g 尤为关键。传统训练方法有:网格法、PSO、GA 等方法。

但是,传统的这几种方法会陷入局部最优,或者存在计算速度慢等缺点,因此,本文引入布谷鸟搜索优化支持向量机的分类方法,在保证计算速度的前提下,更准确地进行轴承故障的分类。

2.2 布谷鸟搜索算法

布谷鸟搜索(CS)是 YANG Xin-she 和 DEB S^[22]于 2009 年开发的自然启发式算法。该算法可以通过 Levy 飞行来增强全局搜索能力。

2.2.1 Levy 飞行

Levy 飞行是一种步长符合重尾分布的一种随机游走的模式^[23],可以增加种群多样性,避免陷入局部最优。用 Levy 飞行生成随机数应包括两个步骤:随机方向的选择和服从 Levy 分布的步长的生成。方向的生成服从均匀分布,步长的生成最有效最直接的方法就是用 Mantegna 算法实现对称的 Levy 分布。

在 Mantegna 算法中,步长 s 可以使用两个服从高斯分布的变量 U 和 V 来计算,即:

$$s = \frac{U}{|V|^{1/\lambda}}, U \sim N(0, \sigma^2), V \sim N(0, 1) \quad (6)$$

其中,方差的计算为:

$$\sigma^2 = \left[\frac{\Gamma(1 + \lambda)}{\lambda \Gamma((1 + \lambda)/2)} \cdot \frac{\sin(\pi\lambda/2)}{2^{(\lambda-1)/2}} \right]^{1/\lambda} \quad (7)$$

2.2.2 布谷鸟搜索

布谷鸟搜索算法是开关参数 p_a 控制的局部随机游走和全局探索随机游走的平衡组合。局部随机游走可以表示为:

$$x_i^{t+1} = x_i^t + \alpha s \odot H(p_a - \varepsilon) \odot (x_j^t - x_i^t) \quad (8)$$

式中: \odot —点乘; $H(*)$ —单位阶跃函数; ε —均匀分布的随机数; α —步长缩放系数; p_a —开关参数; s —步长; x_j^t, x_i^t —随机的两个不同的解。

全局探索随机游走使用 Levy 飞行,即:

$$x_i^{t+1} = x_i^t + \alpha L(s, \lambda) \quad (9)$$

$$L(s, \lambda) = \frac{\lambda \Gamma(\lambda) \sin(\pi\lambda/2)}{\pi} \cdot \frac{1}{s^{1+\lambda}} \quad (10)$$

式中: α —步长缩放系数; s —步长; $\Gamma(\lambda)$ —对于给定的 λ 是个常数。

通过局部随机游走和全局探索随机游走,可以有效地避免陷入局部最优。

2.3 布谷鸟搜索支持向量机(CS-SVM)

通过 CS 算法优化 SVM,就是利用 CS 算法训练 RBF 核函数的参数 g 和惩罚因子 C ,利用训练集以 SVM 的预测的错误率为适应度函数(错误率越低,对应的 C 和 g 越优),自适应地寻找一个最优的 C 和 g 组合,然后用最优的 C 和 g 组合作为测试集 SVM 的输入参数进行轴承故障的分类。

具体步骤如下:

- (1) 设置 SVM 参数 g 和惩罚因子 C 的范围;
- (2) 设置 CS 的迭代次数 $time$ 、巢穴数量 n 、被发现概率 p_a 、需要寻优参数个数 dim ;
- (3) 对于每个巢穴,随机初始化 g 和 C 的值,作为 SVM 的参数,求得准确率,并且以预测错误率作为适应度值,找到最优的巢穴(即当前最优 g 和 C 的值):

$$fnew = 1 - \frac{fitness}{100} \quad (11)$$

式中: $fitness$ —通过 g 和 C 为参数的 SVM 分类的准确率; $fnew$ —此时的适应度(目的是求最小 $fnew$ 下对应的 g 和 C)。

(4) 使用 Levy 飞行更新巢穴,将更新后的 g 和 C 的值作为 SVM 的参数,重复步骤(3),得到一组新的巢穴;

(5) 以 p_a 为概率随机淘汰掉一部分巢穴,并且根据公式(8)更新巢穴,重复步骤(3),得到一组新的巢穴,将新的适应度值与步骤(4)得到的适应度值进行

比较,得到最优的巢穴;

(6) 从步骤(5)中找到最佳巢穴,如果优化目标函数值满足结束条件,则输出全局最好的巢穴和最好的适应度值,否则返回步骤(4)继续优化。

由于 CS 算法是局部最优搜索与全局最优搜索相结合的搜索算法,不易陷入局部最优的情况,分类准确度会比传统方法更高。

3 实验及结果分析

3.1 实验准备

该实验采用的滚动轴承故障数据来源于美国凯斯西储大学轴承数据中心,该数据中心实验平台包括电机、转矩传感器、功率计和电子控制设备,测试的是支撑电机的轴承,使用电火花加工技术在轴承上模拟了点蚀故障,故障直径分别为 0.18 mm、0.36 mm、0.53 mm、0.71 mm(模拟故障从微弱到严重),采样频率是 12 000 Hz。

该实验采集正常数据、内圈缺陷 0.18 mm 数据、内圈缺陷 0.36 mm 数据、内圈缺陷直径 0.53 mm 数据、内圈缺陷 0.71 mm 数据、滚动体缺陷 0.18 mm 数据、外圈缺陷 0.18 mm 数据(3 点位置、6 点位置、12 点位置)、外圈缺陷 0.36 mm 数据、外圈缺陷 0.53 mm 数据共 11 类,每类 40 组数据,其中每类前 20 组数据作为训练集,后 20 组数据作为测试集。

3.2 实验结果

本研究使用 EEMD 方法进行数据处理,并且提取变换后的 14 个 IMF(不足 14 个则用 0 补位)的均方根作为特征输入 CS-SVM、PSO-SVM、GA-SVM 进行对比试验(迭代 100 次,种群规模为 20)。

由于 3 种算法种群最优准确率辨识度较差,本文选择平均准确率进行对比,可以直观地了解 3 种算法的差距。

3 种算法每次迭代的种群平均准确率如图 1 所示。

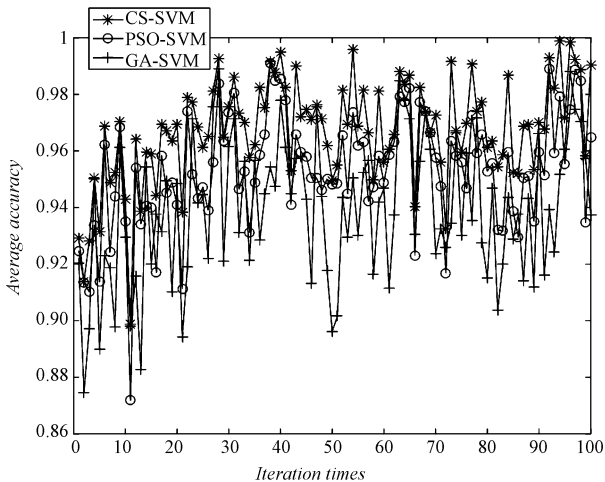


图 1 3 种算法每次迭代的种群平均准确率

由图 1 可知:总体上 CS-SVM 的准确率最高,PSO-SVM 其次,GA-SVM 最低。

CS-SVM 的单次测试准确率如图 2 所示。

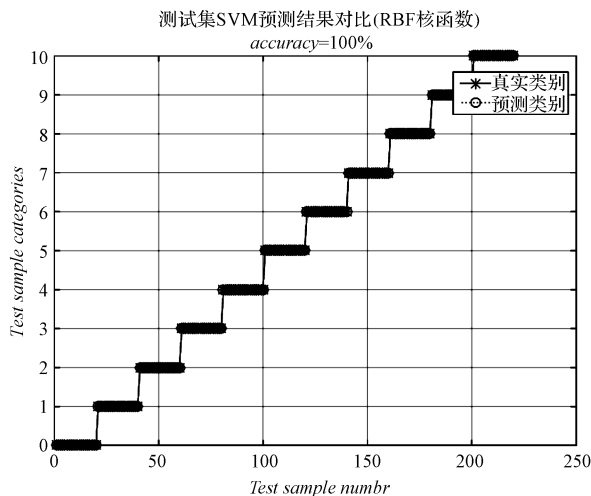


图 2 CS-SVM 的单次测试准确率

PSO-SVM 的单次测试准确率如图 3 所示。

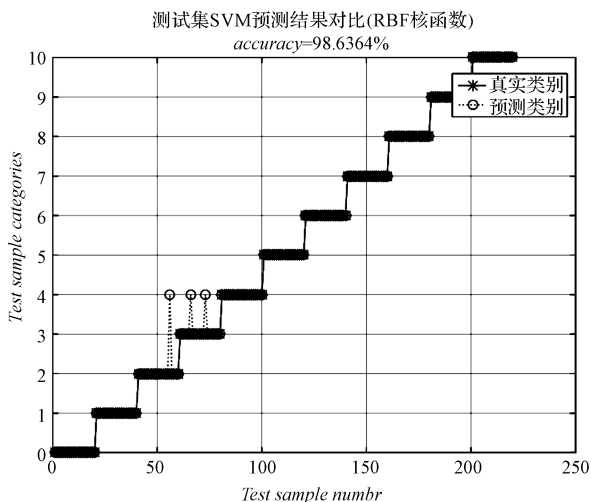


图 3 PSO-SVM 的单次测试准确率

GA-SVM 的单次测试准确率如图 4 所示。

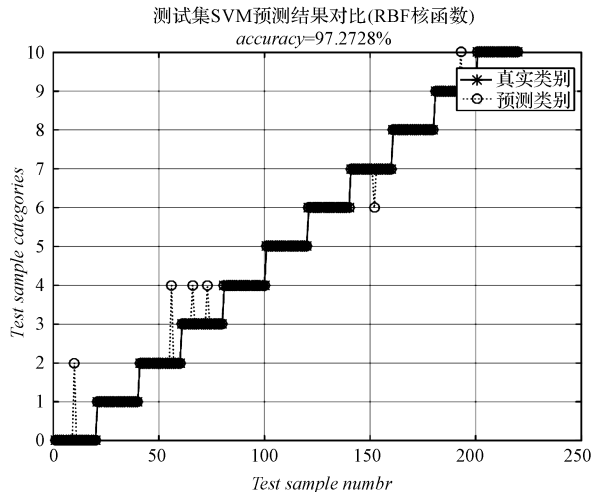


图 4 GA-SVM 的单次测试准确率

由图(2~4)可见:

(1)GA-SVM 会将无缺陷错分类成有缺陷,而且内圈的不同严重程度缺陷、滚动体不同缺陷位置、外圈的不同严重程度缺陷之间会误分类;

(2)PSO-SVM 对于无缺陷、滚动体的不同位置缺陷以及外圈不同严重程度的缺陷之间的分类很准确,但是对于内圈不同严重程度的缺陷之间会产生误分类;

(3)CS-SVM 对不同位置的缺陷以及不同严重程度的缺陷分类都十分准确。

为了使结果更有说服力,本研究采取实验 50 次取平均值的方法进行结果对比,不同分类方法的比较如表 2 所示。

表 2 不同分类方法比较

分类方法	分类准确率/(%)	50 次平均时间/s
CS-SVM	100	23.62
PSO-SVM	99.89	16.42
GA-SVM	99.63	23.84

通过实验可知:

虽然 CS-SVM 的分类时间不是最短,但是在可接受范围之内,在分类时间不增加太多的条件下,CS-SVM 的分类准确率优于传统的 PSO-SVM 和 GA-SVM。

为了验证实验的准确性,使实验结果更加有说服力,本研究再次针对美国辛辛那提大学轴承数据进行实验,其数据预处理方法与凯斯西储大学数据操作方法类似(辛辛那提大学轴承数据只分正常数据、内圈缺陷数据、外圈缺陷数据、滚动体缺陷数据 4 类)。

不同分类方法比较如表 3 所示。

表 3 不同分类方法比较

分类方法	分类准确率/(%)	50 次平均时间/s
CS-SVM	98.87	22.54
PSO-SVM	96.84	18.36
GA-SVM	93.98	26.11

针对辛辛那提大学轴承数据的实验表明:CS-SVM 在不同数据平台上的效果也优于 PSO-SVM 和 GA-SVM。

4 结束语

本文利用 EEMD 方法对信号数据进行了分解,提

取了 IMF 分量的均方根作为特征,利用布谷鸟搜索算法优化支持向量机对此特征进行了分类,并且通过美国凯斯西储大学的轴承数据和美国辛辛那提大学的轴承数据进行了验证,不仅准确地判断了轴承故障发生的位置(内圈、滚动体、外圈的 3 个位置),而且从准确率上优于传统的 PSO-SVM 和 GA-SVM。

在下一阶段,笔者将通过现场实测数据进行实验,以进一步验证此方法的泛化能力。由于现场实测数据与实验室模拟数据的故障类型有些许差异,在今后的研究中,还需要进一步改进该方法,以减少实验误差。

参考文献(References):

- [1] 刘冬. 基于振动信号处理的旋转机械故障诊断[D]. 上海:上海交通大学信息安全工程学院,2010.
- [2] 司景萍,马继昌,牛家骅,等. 基于模糊神经网络的智能故障诊断专家系统[J]. 振动与冲击,2017,36(4):164-171.
- [3] 李舜酩,郭海东,李殿荣. 振动信号处理方法综述[J]. 仪器仪表学报,2013,34(8):1907-1915.
- [4] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis[J]. **Proceedings Mathematical Physical & Engineering Sciences**, 1998, 454(1971):903-995.
- [5] WU Zhao-hua, HUANG N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. **Advances in Adaptive Data Analysis**, 2011,1(1):1-41.
- [6] ŽVOKELJ M, ZUPAN S, PREBIL I. Non-linear multivariate and multiscale monitoring and signal denoising strategy using kernel principal component analysis combined with ensemble empirical mode decomposition method[J]. **Mechanical Systems & Signal Processing**, 2011, 25(7):2631-2653.
- [7] LEI Y, HE Z, ZI Y. EEMD method and WNN for fault diagnosis of locomotive roller bearings[J]. **Expert Systems with Applications**,2011,38(6):7334-7341.
- [8] ZHOU Y, TAO T, MEI X, et al. Feed-axis gearbox condition monitoring using built-in position sensors and EEMD method[J]. **Robotics & Computer Integrated Manufacturing**, 2011,27(4):785-793.
- [9] WANG Y, LI Zhi-hua, WU Chun-hua, et al. A survey of online fault diagnosis for PV module based on BP neural network[J]. **Power System Technology**, 2013,37(8):483-486.

- [10] MURALIDHARAN V, SUGUMARAN V. Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump [J]. **Measurement**, 2013, 46(1):353-359.
- [11] VAPNIK V N. The nature of statistical learning theory[J]. **Technometrics**, 2002, 8(6):1564-1564.
- [12] KONAR P, CHATTOPADHYAY P. Bearing fault detection of induction motor using wavelet and support vector machines (SVMs) [J]. **Applied Soft Computing**, 2011, 11(6):4203-4211.
- [13] HUANG J, HU X, GENG X. An intelligent fault diagnosis method of high voltage circuit breaker based on improved EMD energy entropy and multi-class support vector machine [J]. **Electric Power Systems Research**, 2011, 81(2):400-407.
- [14] SAMANTA B. Gear fault detection using artificial neural networks and support vector machines with genetic algorithms [J]. **Mechanical Systems & Signal Processing**, 2004, 18(3):625-644.
- [15] WIDODO A, YANG B S. Support vector machine in machine condition monitoring and fault diagnosis [J]. **Mechanical Systems & Signal Processing**, 2008, 21(6):2560-2574.
- [16] 李娜源,袁梅,王瑶,等. SVM 与 PSO 相结合的电机轴承故障诊断[J]. 重庆大学学报, 2018, 41(1):99-107.
- [17] 蒋恩超,傅攀,张思聪. 小波包和 GA-SVM 在轴承故障诊断中的应用[J]. 计算机测量与控制, 2017, 25(10):7-10.
- [18] 陈园艺,孙建平. 基于 EEMD 和 PSO-SVM 的滚动轴承故障诊断[J]. 电力科学与工程, 2016, 32(10):47-52.
- [19] 秦波,刘永亮,王建国,等. 基于小波包与奇异值分解的 GA-SVM 滚动轴承故障诊断方法研究[J]. 机械设计与制造, 2016(6):38-42.
- [20] 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.
- [21] 刘志刚,李德仁,秦前清,等. 支持向量机在多类分类问题中的推广[J]. 计算机工程与应用, 2004, 40(7):10-13.
- [22] YANG X S, DEB S. Cuckoo search via lévy flights [C]. NaBIC World Congress on Nature & Biologically Inspired Computing, Coimbatore: IEEE, 2010.
- [23] VISWANATHAN G M, BARTUMEUS F, BULDYREV S V, et al. Lévy flight random searches in biological phenomena [J]. **Physica A Statistical Mechanics & Its Applications**, 2002, 314(1):208-213.

[编辑:李辉]

本文引用格式:

梁治华,曹江涛,姬晓飞. 基于 EEMD 和 CS-SVM 的滚动轴承故障诊断研究[J]. 机电工程, 2019, 36(6):622-627.

LIANG Zhi-hua, CAO Jiang-tao, JI Xiao-fei. Fault diagnosis of rolling bearing based on EEMD and CS-SVM [J]. *Journal of Mechanical & Electrical Engineering*, 2019, 36(6):622-627.

《机电工程》杂志: <http://www.meem.com.cn>