

DOI:10.3969/j.issn.1001-4551.2019.11.008

基于 GG 模糊聚类的滚动 轴承退化阶段划分研究*

孙德建¹, 胡 雄^{1*}, 王 冰¹, 王 微¹, 林积昶²

(1. 上海海事大学 物流工程学院, 上海 201306; 2. 32145 部队 81 分队, 河南 新乡 453000)

摘要:针对滚动轴承退化特征提取以及性能退化阶段准确划分的问题,采用 Logistic 混沌映射,对谱熵在复杂度演化中的变化规律进行了研究。提出了一种基于均方根、谱熵、“弯曲时间参数”特征以及 GG 模糊聚类的滚动轴承退化阶段划分方法,并采用 IMS 轴承实验中心的滚动轴承全寿命试验数据进行了实例分析。研究表明:谱熵参数能够有效描述性能退化过程中的复杂度变化规律,对复杂度变化十分敏感,计算速度快;引入的 Curved Time 参数能够反映退化状态在时间尺度上的集聚特性,更符合机械设备的性能退化规律,因此 GG 模糊聚类方法能够实现滚动轴承等机械设备性能退化阶段的准确划分。

关键词:谱熵; GG 模糊聚类; 滚动轴承; 特征提取

中图分类号:TH133.33; TP806.3

文献标识码:A

文章编号:1001-4551(2019)11-1166-06

Degradation condition division of rolling bearing based on gath-geva fuzzy clustering

SUN De-jian¹, HU Xiong¹, WANG Bing¹, WANG Wei¹, LIN Ji-chang²

(1. Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China;

2. 81 team, 32145 troops, Xinxiang 453000, China)

Abstract: Aiming at the problem of degradation degree description performance degradation stage division for rolling bearings, the variation pattern of spectrum entropy in complexity evolution was studied by using the Logistic chaos mapping sequence. A division method of rolling bearing degradation stages based on root mean square, spectral entropy, “bending time parameter” and gath-geva(GG) fuzzy clustering was proposed. The example analysis was carried out and the life test data from the IMS bearing test center. The results show that reflect the complexity evolution tendency is able to be reflected by spectrum entropy which has a advantage of sensitive to variation and fast calculation speed. The continuity of the same state on the time scale is able to be described by introduced Curved Time parameter which is more according to performance degradation pattern for mechanical equipments. The degradation conditions of mechanical equipment such as bearings can be divided accurately by GG fuzzy clustering.

Key words: spectrum entropy(SE); gath-geva(GG)fuzzy clustering; rolling bearing; feature extraction

0 引 言

港口起重机械是码头重要的物资装卸设备,滚动轴

承是其运行机构中的关键旋转支撑部件。在恶劣的工作环境以及循环大冲击作业下,一旦发生故障,轻则带来经济损失,重则导致人员伤亡。采集并分析旋转支撑

收稿日期:2019-03-12

基金项目:国家高技术研究发展计划(“863 计划”)资助项目(2013AA041106)

作者简介:孙德建(1982-),男,福建宁德人,博士研究生,工程师,主要从事港口机械状态监测与故障预测方法方面的研究。E-mail:djsun@shmtu.edu.cn

通信联系人:胡雄,男,教授,博士生导师。E-mail:huxiong@shmtu.edu.cn

部件的运行监测信号,准确识别其性能退化阶段,能够有效降低故障发生概率。在实际工况下,滚动轴承振动信号一般表现出非线性、非平稳、非周期的特点,需要针对性地分析其中蕴含的退化规律以及状态演化规律。

退化特征提取是准确识别健康状态的基础^[1]。目前该领域主要基于时域、频域以及时频域分析方法而展开研究。其中,时域和频域统计特征因计算简便快速而得到一定的应用,如信号的有效值、方差、平均频率等^[2-3],但该类方法本质上基于线性信号分析方法,无法全面反映性能退化过程中的非线性规律。而针对工程振动信号非线性、非平稳性的特点,以信息熵为基础的复杂度分析方法为非线性退化规律挖掘提供了一条有效的思路。包括以模糊熵^[4]、样本熵^[5]、近似熵^[6]为代表的行为复杂度和以谱熵、 C_0 复杂度为代表的结构复杂度。研究表明:行为复杂度一般涉及嵌入维数的选取,结果稳定性不足,运算速度较慢;结构复杂度方法参数少,计算速度快。其中的谱熵以 FFT 变换为基础,不涉及中间参数,结果稳定,能够有效地分析短时、非平稳、有噪声干扰的数据^[7]。文献[8]证明了 SE 和 C_0 复杂度曲线能够正确有效地描述连续混沌系统的动力学特征。当前,谱熵算法主要应用在混沌系统的复杂性分析上,利用该算法分析机械设备性能退化规律的研究很少,开展基于谱熵的退化特征分析具有一定的研究空间。

退化阶段的划分是健康状态评估中的难点问题。采用主观划分方法缺乏一定的科学性。而实际的特征数据演化进程中,同类状态的数据中蕴含着一定的集聚性。挖掘特征数据的内禀的关联性和集聚性,能够提高退化阶段划分的科学性。为促进现场在线状态评估奠定基础,以 GG (gustafaon-kessel clustering) 为代表的无监督聚类引入了模糊最大似然估计距离范数^[9],能够更精确地挖掘数据集中的聚类特性。目前关于 GG 算法的研究集中在故障模式的诊断,对性能退化状态的聚类则相对较少^[10-12]。

综上,本文提出一种基于 GG 模糊聚类的滚动轴承退化阶段划分方法,提取振动信号的均方根、谱熵以及“弯曲时间参数”作为三维特征向量,以模糊聚类方法对性能退化状态进行无监督聚类,采用 IMS 的轴承全寿命试验数据对方法进行验证。

1 基于谱熵的退化特征分析

1.1 谱熵定义

谱熵的主要思想是以傅里叶变换为基础,分析傅

里叶变换得到的频域内的能量分布,并基于香农熵理论而得到描述信号复杂度的指标。

该算法的基本流程如下^[13]:

(1) 直流部分去除。假设 $x(n)$ 为长度 N 的时间序列,利用下式去除掉信号的直流成分,从而可使频谱更准确地表征信号的能量,即:

$$x(n) = x(n) - \bar{x} \quad (1)$$

式中: \bar{x} —信号的均值, $\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$ 。

(2) 对去除直流分量的信号进行离散傅里叶变换,即:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk} = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad (2)$$

(3) 相对功率谱计算。对经过离散处理后的频谱序列,取其前半部分进行计算,并利用 Parseval 算法,得到其中一个特定频率的功率谱大小,即:

$$p(k) = \frac{1}{N} |X(k)|^2, k = 0, 1, 2, \dots, N/2 - 1 \quad (3)$$

信号的总功率可定义为:

$$P_{\text{tot}} = \frac{1}{N} \sum_{k=0}^{N/2-1} |X(k)|^2 \quad (4)$$

相对功率谱的概率可表示为:

$$P_k = \frac{P(k)}{P_{\text{tot}}} = \frac{\frac{1}{N} |X(k)|^2}{\frac{1}{N} \sum_{k=0}^{N/2-1} |X(k)|^2} = \frac{|X(k)|^2}{\sum_{k=0}^{N/2-1} |X(k)|^2} \quad (5)$$

其中, $\sum_{k=0}^{N/2-1} P_k = 1$ 。

(4) 以香农熵理论为基础,计算信号的谱熵 se , 即:

$$se = - \sum_{k=0}^{N/2-1} P_k \ln P_k \quad (6)$$

一般情况下,由于 se 的最大值为 $\ln(N/2)$, 一般会对谱熵进行归一化操作,得到归一化的谱熵,即:

$$SE(N) = \frac{se}{\ln\left(\frac{N}{2}\right)} \quad (7)$$

通过以上定义可以看出,谱熵能够描述信号的结构稳定性:功率谱变化情况越不稳定,则信号的结构组成越简单,其序列振幅越不明显,相应的得到的测量值也较小;反之,信号的结构组成越复杂,参数的取值越大。

本文以谱熵对滚动轴承的复杂性退化趋势进行描述。

1.2 退化特征性能分析

笔者以 Logistic 混沌映射序列作为仿真信号,分析

谱熵的演化规律,信号表达式如下^[14]:

$$x(t + 1) = \lambda x(t)(1 - x(t)) \quad (8)$$

式中: $x(t) \in (0,1)$; t — 迭代次数; λ — 非线性参数(不同的参数取值使信号呈现不同的动力学行为)。

设置初始值为 $x = 0.4$, 分别计算每个 λ 取值下的谱熵。

Logistic 混沌映射谱熵参数演化曲线如图 1 所示。

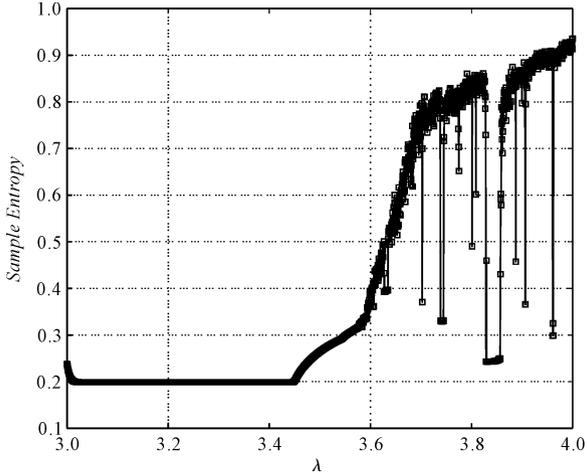


图 1 Logistic 混沌映射谱熵参数演化曲线

由图 1 可以看出:

当 $\lambda \in [3, 3.57]$ 时,系统处于周期变化的状态, SE 取值维持在 0.2 附近,并随着周期个数的增加在 3.45 处出现谱熵取值的增大;当 $\lambda \in [3.57, 4.0]$ 时,系统处于混沌状态,同时一些狭小的区间混杂有周期性循环,使序列的复杂性下降。谱熵参数能够清晰反映混沌态复杂度不断增大的趋势,并在其中混杂的周期状态时出现取值的下降。

与此同时,谱熵的运算过程中不涉及中间参数,运算速度快,适合在线退化特征分析。

2 退化阶段划分流程与评价

2.1 退化特征选取

为了提高特征向量的全面性,本研究提取信号的三维特征向量 $[RMS, SE, Curved\ time]$ 。其中,均方根 RMS 能够表征信号的能量累积变化; SE 复杂度能够反映性能退化进程中的复杂度变化; $Curved\ time (CT)$ 能够反映性能退化进程中的时间集聚特性。该指标的提出是考虑到同一种运行状态在退化时间上具有内在的集聚特性,并且机械设备在性能退化初期的变化平缓,在退化进程中后期的变化会比较剧烈。

结合指数函数的特性,本文将全寿命时间参数 T 归一化并映射到函数 $CT = e^T - 1$ 中获得初期平缓、后期剧烈的“弯曲时间维度”,从而更准确地反映机械设

备性能退化的时间分布特性。

2.2 退化阶段划分效果评价

本文采用 GG 模糊聚类^[15] 对退化阶段进行划分。模糊聚类的效果采用基于隶属度矩阵 U 的分类系数 (*classification coefficient, CC*) 和平均模糊熵 (*average fuzzy entropy, AFE*)^[16]。 CC 指标越接近 1, AFE 指标越接近 0,模糊聚类的效果越好。此外,本文提出并采用序列离散度指标 (*sequence dispersion, SD*),对同一退化状态在时间尺度的集聚性进行评价。

具体的计算方法如下:

对于某个聚类,假设 I 为该集合的样本标签序列, n 为该聚类的样本个数, $m - 1$ 为该标签序列最大值与最小值之差,定义该聚类的序列离散度如下:

$$b = \frac{(m - n)}{m} \quad (9)$$

显然,如果 I 为连续序列,则 $b = 0$; I 越不连续,序列中存在“空位”越多,序列离散度越大。

假设整个样本集合被划分为 c 类,则此次聚类的

$$\gamma = \sum_{i=1}^c b_i \quad (10)$$

由式(10)可知:该指标越接近于 0,代表时间集聚度越高,退化状态聚类效果越好;取值越大,时间集聚度越低,聚类效果越差。

2.3 特征向量选取

本文提出一种基于 GG 模糊聚类的滚动轴承退化阶段划分方法。

该退化状态划分流程如图 2 所示。

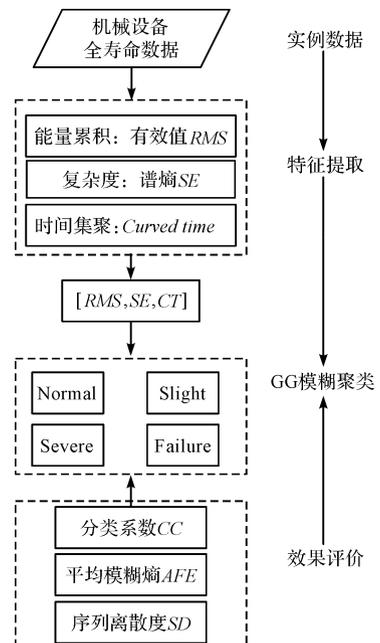


图 2 退化状态划分流程

由图 2 可知:在获得滚动轴承全寿命数据之后,按照退化特征提取、GG 模糊聚类、聚类效果评价 3 个阶段,可以实现轴承全寿命退化阶段的无监督划分。

3 实例分析

本节采用的全寿命数据集来自辛辛那提大学 IMS 中心^[17]。实验中采用的轴承类型为 Rexnord ZA-2115 双列滚子轴承,滚子数量为 16,滚子组节圆直径为 75.501 mm,滚子直径为 8.407 4 mm,接触角为 15.17°。

本研究选取其中一组数据集进行分析。该组试验的负载为 5 000 N、转速 1 500 r/min、采样频率 20 kHz,每组采样时间为 1 s,组间采样间隔为 10 min。数据集采样组数为 984。

试验台停机时,检查发现 1#轴承出现故障,失效形式为外圈故障。其余,2#~4#轴承完好。

3.1 退化特征提取

本文分别对每组采样数据进行退化特征分析,计算谱熵、RMS 以及 *Curved Time*。

轴承全寿命数据集下的性能退化特征趋势如图 3 所示。

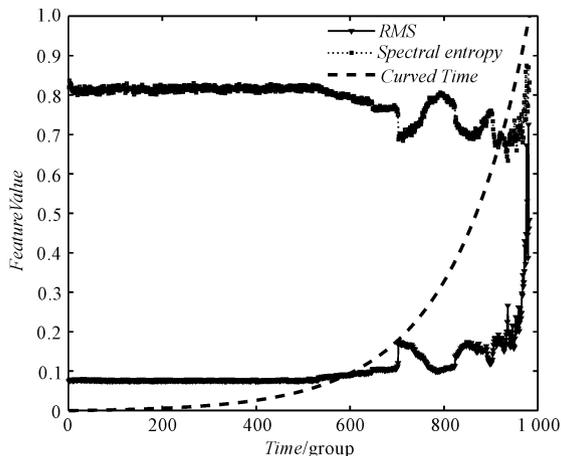


图 3 轴承全寿命数据集下的性能退化特征趋势图

由图 3 可以看出:

整体趋势上,信号的复杂程度随着性能退化程度加深而逐渐降低,有效值 RMS 的趋势则与之相反。这说明随着性能退化程度的增加,信号中的随机成分逐渐减少,信号的复杂度随之降低;从能量累积观点上看,信号的能量随着退化程度的增加而不断增大,有效值 RMS 也随之增大。从细节上看,两个特征参数均呈现出一定的阶段性,反映了轴承性能退化的不同状态。

另外,*Curved Time* 参数的变化率在一定程度上反映了退化状态的演化速率。

3.2 退化阶段划分

参考文献研究中常用的 4 种不同退化阶段的划分方法^[18-19]将全寿命退化过程划分为:正常状态(Normal)、轻微退化状态(Slight)、严重退化状态(Severe)、失效状态(failure)。本文采用 GG 模糊聚类算法对 4 种退化阶段进行无监督聚类。设置参数为 $c = 4, m = 2$,容差为 $\varepsilon = 0.000\ 01$ 。

滚动轴承的状态退化阶段划分结果如图 4 所示。

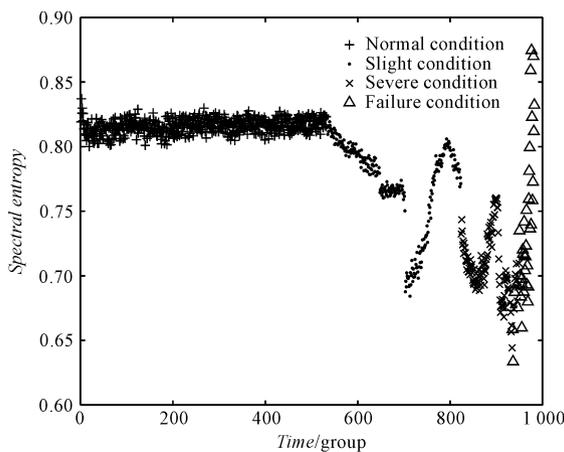


图 4 退化状态划分结果图

由图 4 可以看出:

在约第 520 组采样点之前,轴承处于正常状态,谱熵取值维持在 0.7 附近。当轴承性能轻微退化时,谱熵曲线敏感度快速下降,并出现明显的波动现象,直到第 820 组采样点左右。该阶段自集聚为轻微退化阶段。之后认为轴承性能严重退化,谱熵曲线基本维持在 0.4 左右,数值反弹不大;当进入到失效状态时,退化特征参数总体较低,且出现一些数值异常的离散点,此时认为轴承已经完全失效。

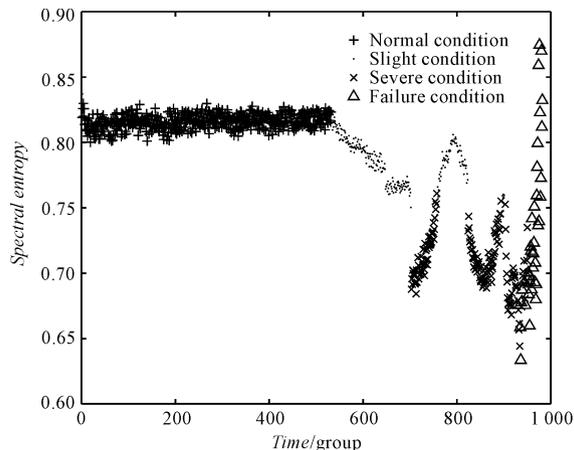
按照聚类的评价方法,4 个分组的序列离散度分别为 0.001 9、0.065 1、0.050 0、0.410 3,本次分类的总的序列离散度为 0.527 2。

3.3 对比分析

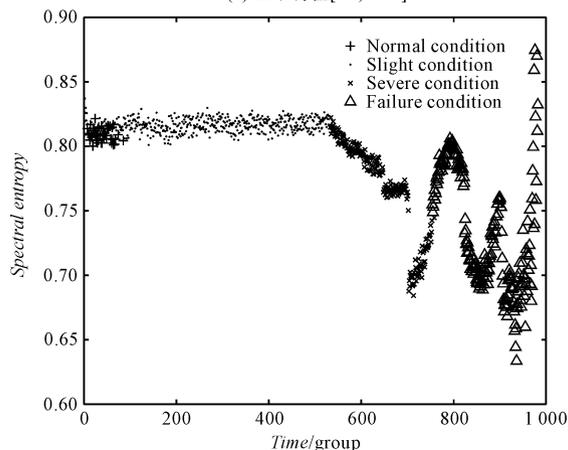
本文设计并采用了 *Curved Time* 参数作为时间集聚度的特征指标,为了分析弯曲时间特征参数对于状态划分的影响,分别采用二维特征 $[SE, RMS]$ 和三维特征 $[SE, RMS, T]$ 进行对比分析,其中, T 为未经映射的时间特征参数。

不同特征参数下的聚类效果如图 5 所示。

不同参数下的聚类定量评价结果如表 1 所示。



(a) 二维特征[SE,RMS]



(b) 三维特征[SE,RMS,T]

图 5 不同特征参数的聚类效果图

表 1 不同特征参数的定量评价结果

特征选取	分类系数	平均模糊熵	序列离散度
[SE,RMS]	0.981 0	0.023 0	3.595 9
[SE,RMS,T]	0.983 9	0.025 1	0.957 7
[SE,RMS,CT]	0.982 6	0.013 6	0.527 2

对比表 1 可以看出:

3 种方法在分类系数和平均模糊熵参数上取值相近,但序列离散度取值较大,说明不同退化阶段的时间集聚性较差。图 5(a)由于未采用时间特征维度,使得同一状态的样本的评判标准缺乏时间上的考量,导致轻微退化状态出现不连续性;图 5(b)尽管引入了时间 T 参数,但由于未做映射处理,使得性能退化在时间尺度上的“变化速率”相同,进而导致了状态分类的“过早集聚”,出现一定程度上的状态边界误判。

本研究保持三维特征向量[SE,RMS,CT]不变,分别采用 GK 聚类、FCM 聚类进行对比分析。

定量对比结果如表 2 所示。

表 2 不同聚类算法的定量评价结果

聚类算法	分类系数	平均模糊熵	序列离散度
GK 聚类算法	0.871 9	0.266 6	0.016 1
FCM 聚类算法	0.845 5	0.301 2	0.025 6
GG 聚类算法	0.982 6	0.013 6	0.527 2

对比表 2 可以看出:

由于引入了 *Curved Time* 参数,3 种方法的序列离散度 SD 均较低,说明分类的时间集聚度较高。但 GK 和 FCM 两种聚类算法的平均模糊熵较高,说明模糊矩阵 U 中的隶属度取值相近,容易造成状态误判。而 GG 聚类算法由于引入了模糊最大似然估计距离范数,在分类效果上较优。

4 结束语

本文提出了一种 GG 模糊聚类的滚动轴承退化阶段划分方法,并采用仿真和实例数据进行了分析验证,得到了以下结论:

(1) 谱熵参数能够反映信号中不规则成分的比例,有效描述性能退化过程中的规律性,并且对复杂度变化十分敏感,计算速度快,对 Logistics 方程和实例信号的分析验证了该方法的有效性;

(2) 引入的 *Curved Time* 参数能够反映退化状态在时间尺度上的集聚特性,更符合机械设备的性能退化规律;

(3) GG 聚类方法能够对任意形状的数据进行聚类,将时间约束加入到特征向量中,能够在保持聚类精度的同时,提高退化状态的时间集聚度。

以该结论为基础,下一步有必要深入研究基于 GG 聚类的在线状态评估方法。

参考文献 (References):

- [1] 年夫顺. 关于故障预测与健康管理的几点认识[J]. 仪器仪表学报,2018,39(8):1-14.
- [2] ARANEO R, ATTOLINI G, CELOZZI S, et al. Time-domain shielding performance of enclosures: a comparison of different global approaches[J]. *IEEE Transactions on Electromagnetic Compatibility*, 2016,58(2):434-441.
- [3] 肖顺根,马善红,宋萌萌,等. 基于 EEMD 和 PCA 滚动轴承性能退化指标的提取方法[J]. 江南大学学报:自然科学版,2015,14(5):572-579.
- [4] 王付广,李 伟,郑近德,等. 基于多频率尺度模糊熵和 ELM 的滚动轴承剩余寿命预测[J]. 噪声与振动控制,2018,38(1):188-192.
- [5] 杨大为,赵永东,冯辅周,等. 基于参数优化变分模态分解和多尺度熵偏均值的行星变速箱故障特征提取[J]. 兵工学报,2018,39(9):1683-1691.
- [6] 李学军,何能胜,何宽芳,等. 基于小波包近似熵和 SVM 的圆柱滚子轴承诊断[J]. 振动、测试与诊断,2015(6):1031-1036.
- [7] 冉 杰,刘衍民,王常春,等. 二维离散 Lorenz 混沌系统的复

- 杂度分析[J]. 遵义师范学院学报,2018,20(4):81-82,99.
- [8] 叶晓林,牟俊,王智森,等. 基于 SE 和 C_0 算法的连续混沌系统复杂度分析[J]. 大连工业大学学报,2018,37(1):67-72.
- [9] 王冰,王微,胡雄,等. 基于 GG 模糊聚类的退化状态识别方法[J]. 仪器仪表学报,2018,39(3):21-28.
- [10] 张立国,李盼,李梅梅,等. 基于 ITD 模糊熵和 GG 聚类的滚动轴承故障诊断[J]. 仪器仪表学报,2014,35(11):2624-2632.
- [11] WANG N, LIU X, YIN J. Improved Gath-Geva clustering for fuzzy segmentation of hydrometeorological time series[J]. *Stochastic Environmental Research and Risk Assessment*, 2012,26(1):139-155.
- [12] YU K, LIN T R, TAN J W. A bearing fault diagnosis technique based on singular values of EEMD spatial condition matrix and Gath-Geva clustering[J]. *Applied Acoustics*, 2017(121):33-45.
- [13] 孙克辉,贺少波,何毅,等. 混沌伪随机序列的谱熵复杂性分析[J]. 物理学报,2013,62(1):35-42.
- [14] 王涛,杨越,顾雪平,等. 基于小波模糊熵 GG 聚类的同调机群识别[J]. 电力自动化设备,2018,38(7):140-147.
- [15] 张淑清,包红燕,李盼,等. 基于 RQA 与 GG 聚类的滚动轴承故障识别[J]. 中国机械工程,2015,26(10):1385-1390.
- [16] 陈东宁,张运东,姚成玉,等. 基于 FVMD 多尺度排列熵和 GK 模糊聚类的故障诊断[J]. 机械工程学报,2018,54(14):16-27.
- [17] LEE J, QIU H, YU G, et al. Bearing dataset from IMS of university of Cincinnati and NASA ames prognostics data repository[EB/OL]. [2010-12-10]. <http://ti.arc.nasa.gov/tech/dashpcoe/prognostic-data-repository>.
- [18] 田再克,李洪儒,孙健,等. 基于改进 MF-DFA 和 SSM-FCM 的液压泵退化状态识别方法[J]. 仪器仪表学报,2016,37(8):1851-1860.
- [19] 谭晓栋,邱静,罗建禄,等. 基于 HSGT 的装备健康状态评估技术[J]. 振动、测试与诊断,2017,37(5):886-891.

[编辑:方越婷]

本文引用格式:

孙德建,胡雄,王冰,等. 基于 GG 模糊聚类的滚动轴承退化阶段划分研究[J]. 机电工程,2019,36(11):1166-1171.

SUN De-jian, HU Xiong, WANG Bing, et al. Degradation condition division of rolling bearing based on gath-geva fuzzy clustering[J]. *Journal of Mechanical & Electrical Engineering*, 2019,36(11):1166-1171. 《机电工程》杂志:<http://www.meem.com.cn>

CNKI中国知网
www.cnki.net

万方数据
WANFANG DATA

VIP
维普资讯

全国邮发代号:14-53 中国连续出版物号:

ISSN 1002-2333

CN23-1196/TH

机械工程师®

杂志
月刊

(每月10号出版)

创刊于 1969

MECHANICAL
ENGINEER

中国期刊方阵双效期刊;《中国学术期刊网络出版总库》全文收录;《中国学术期刊综合评价数据库》(CAJ-CDE)来源期刊;《中文科技期刊数据库CSTJ》全文收录;《中国学术期刊数据库CSPD》全文收录;《中国核心期刊(遴选)数据库》全文收录;《中国学术期刊》(光盘版)全文收录;《CAJ-CD规范》执行优秀期刊。

地址:哈尔滨市香坊区文治头道街30号 邮编:150040 电话:0451-82120966 82127726 网址:www.jxgcs.com 电邮:hrbengineer@163.com jixie888@126.com

2020 征订中 邮发代号:14-53 20元/册 240元/年(12期)