

# 精细复合多尺度波动散布熵在液压泵故障诊断中的应用

姜万录<sup>1,2</sup>, 赵亚鹏<sup>1,2</sup>, 张淑清<sup>3</sup>, 李 满<sup>1,2</sup>

(1. 燕山大学 河北省重型机械流体动力传输与控制重点实验室, 河北 秦皇岛 066004;

2. 燕山大学 先进锻压成形技术与科学教育部重点实验室, 河北 秦皇岛 066004;

3. 燕山大学 电气工程学院, 河北 秦皇岛 066004)

**摘要:** 液压泵振动信号具有非线性、非平稳性的特点, 熵算法在该类信号分析方面有着独到的优势, 但传统的熵算法在液压泵振动信号特征提取中有计算速度慢、熵值不准确、不稳定等不足, 为了更有效地提取故障特征信息并提高故障诊断准确性, 将精细复合多尺度波动散布熵 (refined composite multiscale fluctuation dispersion entropy, RCMFDE) 引入到液压泵的故障特征提取中, 提出了一种基于 RCMFDE 和粒子群优化支持向量机结合的液压泵故障诊断方法。计算不同故障振动信号的 RCMFDE, 并选取合适尺度下的多个 RCMFDE 值作为特征向量形成特征样本, 输入粒子群优化支持向量机中进行故障分类识别。通过仿真信号和液压泵故障实测信号进行分析, 并将所提出的方法与基于多尺度样本熵 (multiscale sample entropy, MSE)、多尺度排列熵 (multiscale permutation entropy, MPE)、多尺度符号动态熵 (multiscale symbolic dynamic entropy, MSDE)、多尺度散布熵 (multiscale dispersion entropy, MDE)、精细复合多尺度散布熵 (refined composite multiscale dispersion entropy, RCMDE)、多尺度波动散布熵 (multiscale fluctuation dispersion entropy, MFDE) 的故障特征提取方法进行对比。试验结果表明, 该方法能够更加准确地识别多类液压泵故障并能对液压泵性能退化程度进行有效评估。

**关键词:** 波动散布熵; 精细复合多尺度波动散布熵 (RCMFDE); 粒子群优化支持向量机; 故障诊断; 液压泵

中图分类号: TH165 + .3

文献标志码: A

DOI:10.13465/j.cnki.jvs.2022.08.002

## Application of refined composite multiscale fluctuation dispersion entropy in hydraulic pumps fault diagnosis

JIANG Wanlu<sup>1,2</sup>, ZHAO Yapeng<sup>1,2</sup>, ZHANG Shuqing<sup>3</sup>, LI Man<sup>1,2</sup>

(1. Hebei Provincial Key Laboratory of Heavy Machinery Fluid Power Transmission and Control, Yanshan University,

Qinhuangdao 066004, China; 2. Key Laboratory of Advanced Forging & Stamping Technology and Science,

Ministry of Education of China, Yanshan University, Qinhuangdao 066004, China;

3. Institute of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China)

**Abstract:** The vibration signal of hydraulic pump has the characteristics of non-linearity and non-stationarity. Entropy algorithms have a unique advantage in this kind of signal analysis. However, the traditional entropy algorithms still have shortcomings of slow calculation speed, inaccurate entropy value and unstable entropy value in hydraulic pump vibration signal feature extraction. To extract fault feature information more effectively and improve fault diagnosis accuracy, the refined composite multiscale fluctuation dispersion entropy (RCMFDE) was introduced into the fault feature extraction of hydraulic pumps. A hydraulic pump fault diagnosis method based on RCMFDE and the particle swarm optimization support vector machine (PSO - SVM) algorithm was proposed. Firstly, the RCMFDE values of different fault vibration signals were calculated and the multi-RCMFDE values were selected at appropriate scales as feature vectors to form feature samples. Then the feature samples were input to PSO - SVM for fault diagnosis. Through analyzing the simulation signals and hydraulic pump experiments signals, the proposed method was compared with the fault diagnosis methods based on multiscale sample entropy (MSE), multiscale permutation entropy (MPE), multiscale symbolic dynamic entropy (MSDE), multiscale dispersion entropy (MDE), refined composite multiscale dispersion entropy (RCMDE), and multiscale fluctuation dispersion entropy (MFDE). Experimental results show that the proposed

method can accurately identify multiple types of hydraulic pump faults and effectively evaluate the performance degradation degree of hydraulic pump.

**Key words:** fluctuation dispersion entropy; refined composite multiscale fluctuation dispersion entropy (RCMFDE); particle swarm optimization support vector machine; fault diagnosis; hydraulic pump

液压泵作为液压系统的动力源,经常工作在高转速高负载状态,其健康状态将直接影响整套设备的正常工作,因此对液压泵的故障诊断研究具有非常重要的意义。当液压泵出现故障时,采集到的振动信号通常会表现出复杂的非线性、非平稳特征并且容易受到环境噪声的干扰,由于传统的时域和频域分析方法存在一定的局限性,所以无法准确地提取特征信息。

基于各类熵值的非线性动力学分析方法已经广泛地应用在机械故障诊断领域。熵值反映了时间序列复杂度的大小,常用的有近似熵 (approximate entropy, AE)<sup>[1]</sup>、样本熵 (sample entropy, SE)<sup>[2]</sup>、排列熵 (permutation entropy, PE)<sup>[3]</sup>、符号动态熵 (symbolic dynamic entropy, SDE)<sup>[4]</sup>、散布熵 (dispersion entropy, DE)<sup>[5]</sup> 等。这些熵被用来量化单一时间尺度上信号的不规律性,但无法分析多个时间尺度上的信息。为了克服这一局限性,人们将这些熵值算法与粗粒化多尺度方法和精细复合多尺度化方法结合,推广到了多尺度分析领域。

这些多尺度熵算法广泛地应用在旋转机械故障诊断中。文献[6]将多尺度样本熵 (multiscale sample entropy, MSE)<sup>[7]</sup>应用在生理信号和旋转机械特征提取中,但是由于样本熵算法的局限性,在处理短信号时不够稳定,且可能没有模板向量彼此匹配而无法定义熵值,在实际工程应用时计算速度也不够快。多尺度排列熵 (multiscale permutation entropy, MPE)<sup>[8]</sup>基于排列熵算法,考虑了时间序列信号幅值排列模式或顺序关系,计算速度快,文献[9-10]将 MPE 应用在滚动轴承和液压泵的故障诊断中。但 MPE 仅考虑了幅值的顺序而忽略了幅度的信息,也没有考虑嵌入向量中出现相等幅值时的影响,并且对噪声敏感<sup>[11]</sup>。

散布熵的提出克服了上述缺点,计算速度快,稳定性高,并且抗噪能力有大幅提升,而且振动信号微小的变化不会改变其映射后的类别标签<sup>[12]</sup>。文献[13-18]将多尺度散布熵 (multiscale dispersion entropy, MDE)<sup>[19]</sup>和精细复合多尺度散布熵 (refined composite multiscale dispersion entropy, RCMDE)应用在滚动轴承、滑动轴承、齿轮箱等旋转机械的故障诊断中,取得了良好的诊断效果。但散布熵仅考虑了幅值的绝对性而没有考虑相对性,无法对信号的波动性进行评估。符号动态熵的概念与散布熵相似,文献[20-21]已将其用于滚动轴承和齿轮箱的故障诊断中。但符号动态熵只能提取信号的状态模式和状态转移概率,没有考

虑到信号的波动性。

Azami 等<sup>[22]</sup>提出了基于波动散布熵 (fluctuation-based dispersion entropy, FDE),并推广为多尺度波动散布熵 (multiscale fluctuation-based dispersion entropy, MFDE)和精细复合多尺度波动散布熵 (refined composite multiscale fluctuation dispersion entropy, RCMFDE)<sup>[23]</sup>,波动散布熵考虑了时间序列的波动性,对基线漂移有更强的鲁棒性,并且在相同参数下,由于所有可能散布模式的减少使熵的计算速度更快。文献[24]将波动散布熵的复合多尺度形式应用在滚动轴承的故障诊断中。复合多尺度方法和精细复合多尺度方法的结果相似,但后者效果更好。

支持向量机 (support vector machines, SVM)<sup>[25]</sup>作为一种常用的机器学习工具,非常适合解决小样本、非线性、高维度的分类问题。但实际应用中核函数类型、核函数参数、惩罚因子的选取对 SVM 的分类效果有很大影响。粒子群优化算法 (particle swarm optimization, PSO)<sup>[26]</sup>是受到鸟群或鱼群觅食的社会行为启发的群体智能优化方法,具有很好的全局寻优能力,因此可用于对支持向量机的超参数进行优化。

鉴于 RCMFDE 在非线性动力学复杂性刻画方面的综合优势,本文将将其引入液压泵的故障诊断中,提出了一种基于 RCMFDE 和粒子群优化的支持向量机 (RCMFDE-PSO-SVM)相结合的故障诊断方法。首先,通过 RCMFDE 对液压泵不同故障状态的振动信号提取多尺度熵值。然后,选取合适尺度的多个熵值组成的特征向量构成特征样本输入到 PSO-SVM 分类器中进行故障诊断。将所提出的故障诊断方法分别与基于 MSE, MPE, MSDE, MDE, RCMDE, MFDE 的特征提取方法以及 SVM, KNN 分类器进行对比,结果表明 RCMFDE-PSO-SVM 方法具有更高的诊断效率和故障识别正确率。

## 1 精细复合多尺度波动散布熵

### 1.1 波动散布熵算法

FDE 与 DE 不同, FDE 考虑了信号之间的相对关系而不是绝对关系,其计算步骤如下:

**步骤 1** 对于给定的一维时间序列  $x = \{x_1, x_2, \dots, x_N\}$ , 将  $x$  映射到  $\{1, 2, \dots, c\}$ 。为此,首先利用正态累积分布函数 (normal cumulative distribution function, NCDF) 将时间序列  $x$  映射到  $y = \{y_1, y_2, \dots, y_N\}$ ,  $y_j \in (0,$

1),即

$$y_j = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{x_j} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (1)$$

式中,  $\mu$  和  $\sigma^2$  分别为原始序列的期望和方差,再通过线性变换

$$z_j^c = \text{round}(cy_j + 0.5) \quad (2)$$

将  $y$  映射到  $\{1, 2, \dots, c\}$  的范围内。式中:  $\text{round}(\cdot)$  为四舍五入取整函数;  $c$  为类别个数。

**步骤2** 根据嵌入维数  $m$  和时延  $d$  计算嵌入向量  $z_i^{m,c}$

$$z_i^{m,c} = [z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c], \quad i = 1, 2, \dots, N - (m-1)d \quad (3)$$

将每个嵌入向量  $z_i^{m,c}$  映射到基于波动的散布模式  $\pi_{v_0 v_1 \dots v_{m-1}}$

$$z_i^c = v_0, z_{i+d}^c = v_1, \dots, z_{i+(m-1)d}^c = v_{m-1} \quad (4)$$

分配给每个时间序列  $z_i^{m,c}$  的所有可能散布模式数量为  $(2c-1)^{m-1}$ 。

**步骤3** 计算每个可能的散布模式  $\pi_{v_0 v_1 \dots v_{m-1}}$  的概率  $p(\pi_{v_0 v_1 \dots v_{m-1}})$

$$p(\pi_{v_0 v_1 \dots v_{m-1}}) = \frac{\text{count}\left\{ \begin{array}{l} i \mid i \leq N - (m-1)d, \\ z_i^{m,c} \text{ 具有模式 } \pi_{v_0 v_1 \dots v_{m-1}} \end{array} \right\}}{N - (m-1)d} \quad (5)$$

式中,  $\text{count}(\cdot)$  为  $z_i^{m,c}$  映射到  $\pi_{v_0 v_1 \dots v_{m-1}}$  的个数,即概率  $p(\pi_{v_0 v_1 \dots v_{m-1}})$  等于  $z_i^{m,c}$  映射到  $\pi_{v_0 v_1 \dots v_{m-1}}$  的个数除以嵌入维数为  $m$  对应的嵌入向量总数。

**步骤4** 最后根据香农熵的定义,计算基于波动的散布熵

$$E_{\text{FD}}(x, m, c, d) = - \sum_{\pi=1}^{(2c-1)^{m-1}} p(\pi_{v_0 v_1 \dots v_{m-1}}) \cdot \ln[p(\pi_{v_0 v_1 \dots v_{m-1}})] \quad (6)$$

FDE 计算中考虑散布模式相邻元素之间的差异,称为基于波动的散布模式。在这种算法中,能得到维数为  $m-1$  的模式向量,模式向量的每个元素范围为  $-c+1 \sim c-1$ ,因此,总共有  $(2c-1)^{m-1}$  种基于波动的散布模式。当所有散布模式具有相等的概率值时,熵值最大,为  $\ln[(2c-1)^{m-1}]$ ,此时信号是完全随机的。

## 1.2 多尺度波动散布熵

MFDE 是基于粗粒化时间序列和 FDE 的结合,计算过程如下:

对于给定的时间序列  $u = \{u_1, u_2, \dots, u_L\}$ ,首先,利用向下取整将其分成长度为  $\lfloor L/\tau \rfloor$  的非重叠片段,  $\tau$  为尺度因子,  $\lfloor \cdot \rfloor$  为下取整运算。然后,计算出每个片段的平均值,得到时间序列

$$x_j^{(\tau)} = \frac{1}{\tau} \sum_{b=(j-1)\tau+1}^{j\tau} u_b, \quad 1 \leq j \leq \left\lfloor \frac{L}{\tau} \right\rfloor \quad (7)$$

然后计算尺度因子  $\tau$  下的 FDE,得到尺度因子  $\tau$  的函数 MFDE

$$E_{\text{MFD}}(x, m, c, d, \tau) = E_{\text{FD}}(x_j^{(\tau)}, m, c, d) \quad (8)$$

## 1.3 精细复合多尺度波动散布熵

传统的粗粒化多尺度方法由于截取的是非重叠片

段,没有充分考虑每个片段两端相邻元素之间的关系,随着比例因子的增加,计算结果的稳定性变差。所以用精细复合多尺度化方法进行替代,其计算过程如下:

对于给定的时间序列  $u = \{u_1, u_2, \dots, u_L\}$ ,对于尺度因子  $\tau$  粗粒化过程中不同的起点,创建不同的时间序列。原始序列  $u$  的第  $k$  个粗粒化时间序列  $x_k^{(\tau)} = \{x_{k,1}^{(\tau)}, x_{k,2}^{(\tau)}, \dots\}$  为

$$x_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{b=k+(j-1)\tau}^{k+j\tau-1} u_b, \quad 1 \leq j < \left\lfloor \frac{L}{\tau} \right\rfloor, \quad 1 \leq k \leq \tau \quad (9)$$

然后计算  $x_k^{(\tau)}$  中散布模式  $\pi$  的平均概率  $\bar{p}^{(\tau)}(\pi_{v_0 v_1 \dots v_{m-1}})$

$$\bar{p}^{(\tau)}(\pi_{v_0 v_1 \dots v_{m-1}}) = \frac{1}{\tau} \sum_{k=1}^{\tau} p_k^{(\tau)} \quad (10)$$

式中,  $p_k^{(\tau)}$  为不同起始点的每个可能的散布模式  $\pi_{v_0 v_1 \dots v_{m-1}}$  的概率。RCMFDE 定义为这些位移序列散布模式概率平均值的香农熵

$$E_{\text{RCMFD}}(x, m, c, d, \tau) = - \sum_{\pi=1}^{(2c-1)^{m-1}} \bar{p}^{(\tau)}(\pi_{v_0 v_1 \dots v_{m-1}}) \cdot \ln[\bar{p}^{(\tau)}(\pi_{v_0 v_1 \dots v_{m-1}})] \quad (11)$$

在上述两种多尺度化方法中,不仅仅是粗粒化与 FDE 的结合,而且在所有尺度的 NCDF 映射中都要保持原始信号的  $\mu$  和  $\sigma^2$  不变,这类似于 MSE 算法中要保持相似容限  $r$  恒定,因为保持  $r$  恒定比独立计算每个尺度下的  $r$  效果更好<sup>[27-28]</sup>。

## 1.4 参数选择

MFDE 和 RCMFDE 的参数有时延  $d$ 、嵌入维数  $m$ 、类别个数  $c$  和尺度因子  $\tau$ 。对于时延  $d$  通常设置为 1,因为当  $d > 1$  时可能会丢掉一些重要的频率信息导致混叠。只有采样频率远大于信号的最高频率分量时,可以选择其他的时延。嵌入维数  $m$  太小可能检测不到信号中的动态变化,而较大的  $m$  无法观察到微小的变化,应用中通常选择  $2 \leq m \leq 5$ 。为了避免噪声对信号的影响,应该使用较小的类别个数  $c$ ,但过小的  $c$  值会丢失数据的详细信息导致错误的概率估计,因此通常取  $3 \leq c \leq 8$ 。

## 2 仿真信号分析

在复杂性评估算法研究中,通常都利用白噪声和粉红噪声作为评估对象来验证算法性能。为了评估 RCMFDE 的性能,下面将其与 MSE, MPE, MSDE, MDE, RCMDE, MFDE 6 种算法进行比较。分别选取 50 组长度为 4 096 点的高斯白噪声 (white Gaussian noise, WGN) 和粉红噪声 (1/f noise) 信号,分别计算 20 个尺度下上述 7 种熵的均值和标准差,绘制误差棒图,如图 1 所示。MSE 的参数设置为  $m=3, r=0.15\sigma$  ( $\sigma$  为标准差); MPE 的参数设置为  $m=3$ ; MSDE, MDE, RCMDE, MFDE, RCMFDE 的参数设置为  $m=3, c=6, d=1$ 。

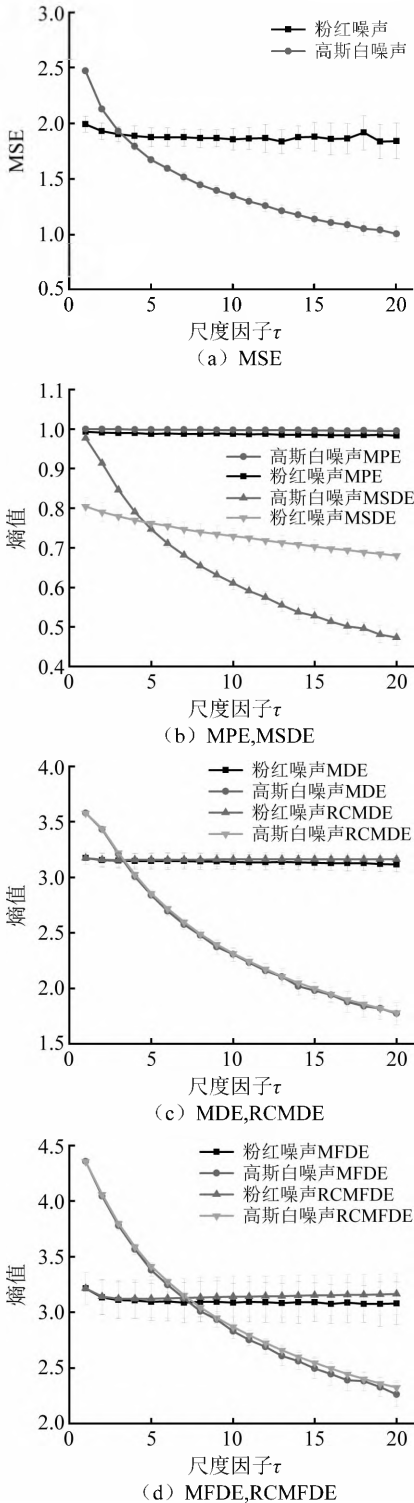


图1 粉红噪声和高斯白噪声的7种熵值的误差棒图  
Fig.1 Error bars of seven entropies of 1/f noise and WGN

从图1中可以看出,除MPE外,其余熵算法均能明显地区分出两种噪声,且熵值曲线趋势相同。在较低尺度上高斯白噪声的熵值更高,随着尺度因子的增加,高斯白噪声熵值单调下降,而粉红噪声的熵值基本保持不变。这是由于粉红噪声具有长程相关性,其1/f衰减在时间上产生分形结构。而高斯白噪声结构简单,各成分相关性小。仿真结果与两种噪声的理论分析一致,即粉红噪声有比高斯白噪声更复杂的结构,而高斯白噪声更不规则。

由于标准差可能会随着平均值的变化而变化,变异系数(coefficient of variation, CV)可以消除测量尺度和量纲不统一对数据离散程度大小的影响。分别计算上述7种熵值的变异系数,其结果如图2所示。

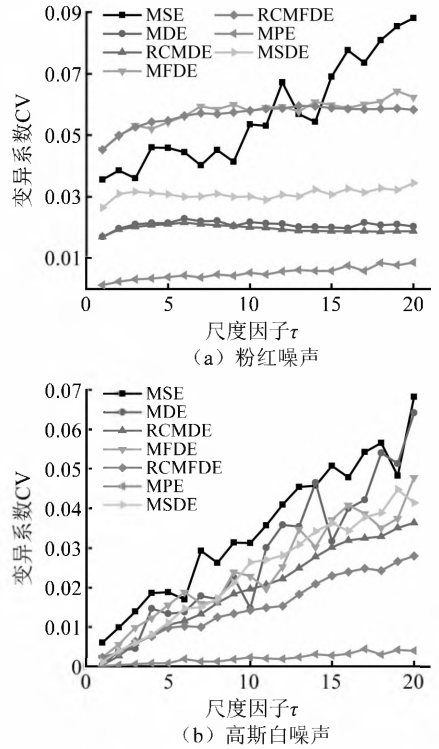


图2 粉红噪声和高斯白噪声7种熵的变异系数  
Fig.2 CVs of the seven entropies of 1/f noise and WGN

从图2中可以看出,在高斯白噪声中MPE的变异系数最小,但无法区分两种噪声。基于DE和FDE算法的变异系数比MSE的小,基于FDE算法的变异系数比基于DE算法的小。两种噪声中精细复合多尺度方法的变异系数都最小,稳定性最高。

为了评估7种熵算法的计算复杂性,使用不同点数的高斯白噪声信号各50组( $N_i = 1024i, i = 1, 2, \dots, 10$ ),记录其计算时间的平均值 $\bar{t}_i$ ,结果如图3所示。计算机环境为CPU: Inter Core i5-9300H 2.4 GHz; RAM:8 G; MATLAB R 2019b。

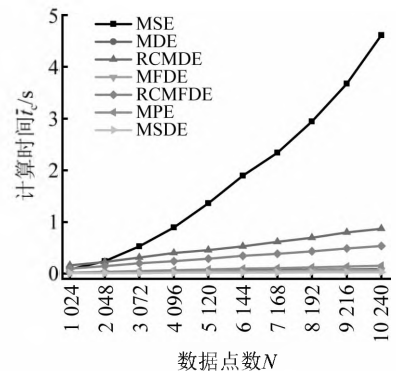


图3 7种熵算法的平均计算时间随高斯白噪声信号长度的变化  
Fig.3 Average calculation times of the seven entropies with different lengths of WGN

从图3中可以看出,7种熵的计算时间随着数据长度的增加而增加。MSE算法的时间复杂度为 $O(N^2)$ 。由于SDE,DE和FDE算法的时间复杂度为 $O(N)$ ,MSDE,MDE,RCMDE,MFDE,RCMFDE的时间复杂度也为 $O(N)$ 。在相同的参数下,FDE中所有可能的散布模式数量更少,从而计算速度比DE更快。对比RCMFDE和RCMDE可以看出较明显的差异。

### 3 RCMFDE - PSO - SVM 故障诊断算法

#### 3.1 SVM 原理

对于多分类问题,使用多个“一对一”分类SVM,在各个类别之间构造分类函数,对于 $n$ 个类别共需要构造 $n(n-1)/2$ 个决策超平面。设训练样本集为 $(x_i, y_i), i=1, 2, \dots, N, x_i \in Rd, y_i \in \{\text{故障模式标签}(n\text{个})\}$ ,对第 $r$ 类和 $s$ 类,寻找如下问题的解

$$\begin{cases} \max_{a^{rs}} Q(a^{rs}) = \\ \sum_{i=1}^N a_i^{rs} - \frac{1}{2} \sum_{i,j=1}^N a_i^{rs} a_j^{rs} y_i^{rs} y_j^{rs} K(x_i^{rs}, x_j^{rs}) \\ \text{s. t. } \sum_{i=1}^N y_i^{rs} a_i^{rs} = 0, \quad 0 \leq a_i \leq C, i = 1, \dots, N \end{cases} \quad (12)$$

式中: $C > 0$ 为惩罚因子; $K(x_i^{rs}, x_j^{rs}) = e^{-\frac{\|x_i^{rs} - x_j^{rs}\|}{\sigma^2}}$ 为核函数。对第 $r$ 类和第 $s$ 类之间的决策函数为

$$F(x) = \text{sign} \left[ \sum_{i=1}^N a_i^{rs} y_i^{rs} K(x_i^{rs}, x^{rs}) + b^{rs} \right] \quad (13)$$

分类时使用投票法,判断 $x_i$ 属于第 $r$ 类故障,则第 $r$ 类故障的票数加1,否则第 $s$ 类故障的票数加1。所有决策函数都判断完成之后,最终得票最多的类别为样本点 $x_i$ 所属的故障类别。

#### 3.2 PSO 原理

在解空间中随机初始化 $m$ 个粒子构成初始种群,记第 $i$ 个粒子当前位置为 $x_i$ ,每个粒子都对应于优化问题的一个解。为每个粒子随机初始化一个速度 $v_i$ ,速度用来决定粒子在解空间中的运动。由目标函数为之确定一个适应值,在算法的每次迭代中,粒子将跟踪自身和种群当前找到的最优解,其中每个粒子当前找到的极值为 $p_i$ ,种群当前找到的全局极值为 $p_g$ ,逐代搜索,直到最后得到最优解。粒子的速度和位置更新公式为

$$v_{i+1} = wv_i + c_1 r_1 (p_i - x_i) + c_2 r_2 (p_g - x_i) \quad (14)$$

$$x_{i+1} = x_i + v_{i+1} \quad (15)$$

式中: $c_1$ 为局部学习因子; $c_2$ 为全局学习因子; $w$ 为惯性因子; $r_1, r_2$ 为 $[0,1]$ 均匀分布的随机数。

### 3.3 RCMFDE - PSO - SVM 故障诊断方法

采用PSO优化的SVM作为分类器,实现对液压泵故障的识别。本文提出的RCMFDE - PSO - SVM液压泵故障诊断方法的流程图如图4所示,诊断步骤如下:

**步骤1** 首先,使用加速度传感器采集不同故障类型的振动信号;

**步骤2** 对采集的信号计算RCMFDE,分析多尺度故障信息;

**步骤3** 选取合适尺度下的多个RCMFDE值作为特征向量,获得故障特征样本集;

**步骤4** 将故障特征样本集按比例随机划分为训练样本集和测试样本集,将训练样本集输入PSO - SVM训练模型;

**步骤5** 将测试样本集输入训练好的模型进行故障识别。



图4 液压泵故障诊断流程图

Fig. 4 Flow chart of hydraulic pump fault diagnosis

## 4 基于RCMFDE - PSO - SVM的液压泵故障诊断

### 4.1 振动信号采集

液压泵振动信号采自液压泵故障模拟试验台,如图5所示。液压泵为MCY14-1B型斜盘式轴向柱塞泵,柱塞数目为7,理论排量为10 mL/r,额定工作压力为31.5 MPa。驱动电机型号Y132M-4,额定转速为1480 r/min。加速度传感型号YD72D,频率范围为1 Hz ~ 18 kHz。对液压泵端盖的振动信号进行采集。试验时调定主溢流阀将系统压力设置为10 MPa,采样频率设为10 kHz,每次采样时长为10 s。

### 4.2 试验1:液压泵不同类型故障的诊断试验

本试验分别采集了液压泵的正常状态、中心弹簧失效、滑靴磨损、单柱塞松靴和斜盘磨损5种状态下的振动信号,故障是用故障元件替换正常元件注入的。将每种信号不重叠地截取48组数据,每组数据中含有

2 048 个数据点,5 种状态共获得 240 组样本。每种故障各选一组典型样本的时域波形,如图 6 所示。

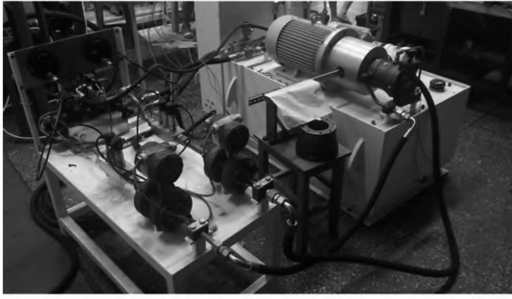


图 5 液压泵故障模拟试验台

Fig. 5 Hydraulic pump failure simulation test bench

为了比较各种熵值算法在不同尺度下的性能,计算 240 个样本的 RCMFDE 值,同时计算出 MSE, MPE, MSDE, MDE, RCMDE, MFDE 值。MSE 的参数选择为  $m=3, r=0.15\sigma, d=1, \tau=20$ 。MPE 的参数设置为  $m=3, d=1, \tau=20$ 。MSDE 和基于 DE, FDE 的各种算法的参数均选取为  $m=3, c=6, d=1, \tau=20$ 。计算 7 种熵的均值和标准差,绘制误差棒图,如图 7 所示。

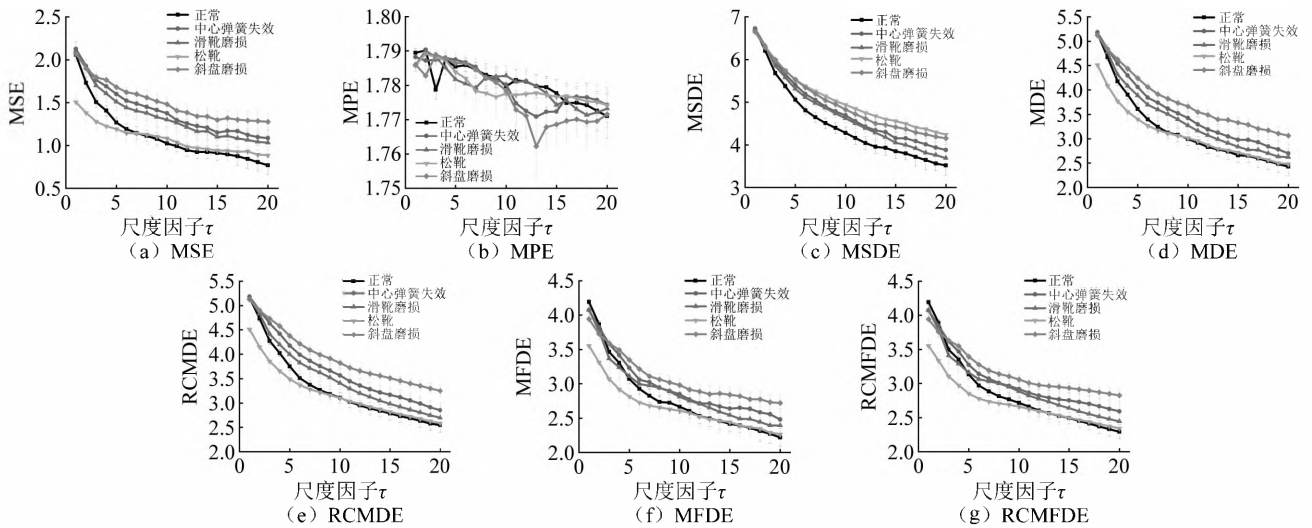


图 7 不同故障振动信号的 7 种熵的误差棒图

Fig. 7 Error bars of the seven entropies of different fault vibration signals

从图 7 可以看出,正常和 4 种故障状态的多尺度熵值明显不同,精细复合多尺度方法比传统多尺度方法曲线更为平滑(见图 7(e)和图 7(g)),误差更小,稳定性更高。熵值曲线在较小和较大的尺度上有交叉或重叠,都不利于分类,所以需要选择合适的特征向量维数。

依次将 MSE, MPE, MSDE, MDE, RCMDE, MFDE, RCMFDE 的前 2 个~20 个尺度下的熵值作为特征向量构造故障特征样本集。训练集和测试集的比例划分为 7:3,按这个比例随机划分训练集和测试集,即从每类信号中选取 33 个特征样本为训练集,15 个特征样本为

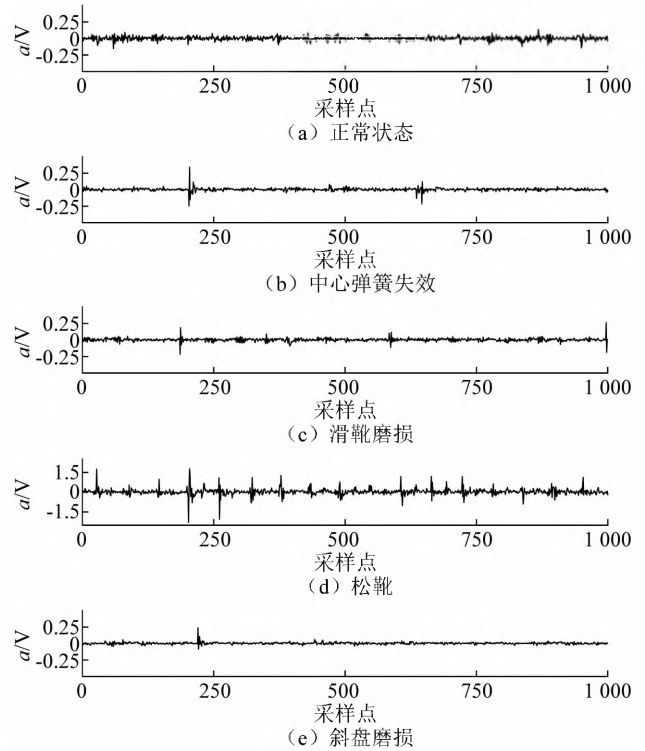


图 6 振动信号时域波形

Fig. 6 Time-domain waveforms of vibration signal

测试集。计算 10 次分类结果的平均准确率作为最终准确率,比较这 7 种算法的性能以及特征向量维数对分类结果的影响。

采用基于粒子群优化的支持向量机对这些特征向量进行分类,以实现液压泵故障的识别。SVM 的核函数选择高斯核函数,通过 PSO 算法寻找最优的两个超参数,即惩罚因子  $C$  和核函数宽度系数  $\sigma$ 。PSO 算法参数设置为粒子数量 25、迭代次数 50、局部学习因子 2、全局学习因子 2、惯性因子 0.8。将正常、中心弹簧失效、滑靴磨损、单柱塞松靴、斜盘磨损分别标记为类别 0~4。然后,将训练集输入分类器进行训练,选择

3 次交叉验证的最高准确率的参数作为分类器超参数, 再将测试集输入到分类器对训练后的模型进行测试。结果如图 8 所示。

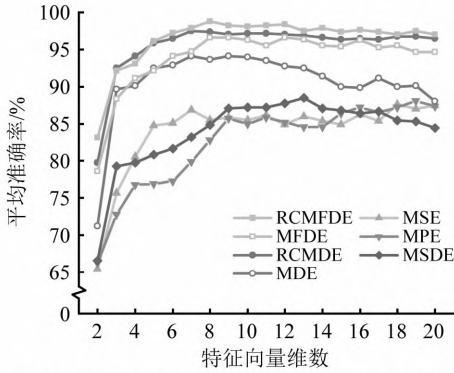


图 8 不同特征向量维数的平均识别准确率  
Fig. 8 Average recognition accuracies of different eigenvector dimensions

在图 8 中,通过对比可以看出,RCMFDE 取得了最好的效果,当特征向量维数为 8 时平均准确率最高,达到了 0.99。4 种基于 DE 和 FDE 算法的准确率都高于 MSE, MPE 和 MSDE 算法,其中基于 FDE 算法的分类效果优于基于 DE 的算法,精细复合多尺度化方法比传统多尺度化方法准确率更高。

为了研究训练集和测试集划分比例对分类准确率的影响,选择前 8 个尺度下的熵值作为特征向量构造样本集,将训练集、测试集比例分为 5 种情况 (1 : 9, 3 : 7, 5 : 5, 7 : 3, 9 : 1), 分别计算 20 次的平均准确率。结果如图 9 所示。

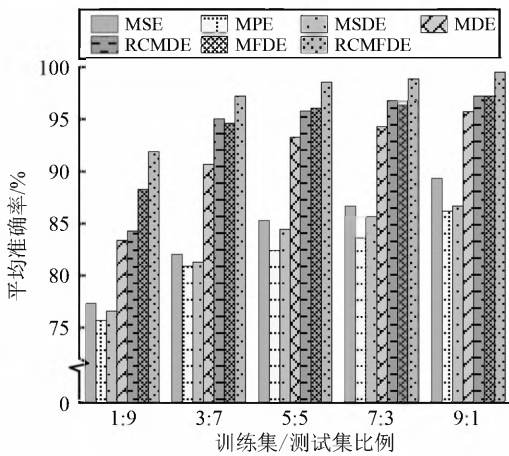


图 9 不同训练/测试集比例下平均识别准确率  
Fig. 9 Average recognition accuracies of different training and testing set ratio

RCMFDE 在不同的训练集、测试集比例下仍具有最高的故障识别准确率,其余趋势也与图 8 的分析结果相同。所有算法的准确率都随着比例的增加而增加,如果训练样本足够多,可以获得更高的正确率,然而计算成本也越高。为了在准确率和计算效率之间取得平衡,本文训练集和测试集的比例选为 7 : 3。

为了研究不同分类器对准确率和计算时间的影

响,验证 RCMFDE - PSO - SVM 算法的突出优势。对 MPE, MSDE, RCMDE, RCMFDE 4 种算法分别使用 PSO - SVM, SVM (惩罚因子  $C = 1$ , 宽度系数  $\sigma = 1.581$ ), KNN (临近系数  $k = 5$ ) 3 种分类器诊断 20 次,记录平均识别准确率和计算时间(该时间包含构造特征样本集所用时间),结果如表 1 所示。

表 1 3 种分类器平均识别准确率和计算时间  
Tab. 1 Average recognition accuracies and calculation times of the three classifiers

算法	平均准确率	训练时间/s	测试时间/s
MSE - PSO - SVM	0.866	55.72	21.65
MPE - PSO - SVM	0.836	24.26	8.85
MSDE - PSO - SVM	0.856	20.26	7.46
RCMDE - PSO - SVM	0.967	35.33	12.93
RCMFDE - PSO - SVM	0.988	32.34	11.54
MSE - SVM	0.817	50.40	21.63
MPE - SVM	0.730	12.25	8.82
MSDE - SVM	0.791	9.56	7.32
RCMDE - SVM	0.839	30.10	12.90
RCMFDE - SVM	0.949	26.95	11.52
MSE - KNN	0.848	50.40	21.59
MPE - KNN	0.650	10.89	8.79
MSDE - KNN	0.834	8.03	7.51
RCMDE - KNN	0.928	30.10	12.92
RCMFDE - KNN	0.973	26.95	11.55

由表 1 对比可以看出,RCMFDE - PSO - SVM 算法具有最高的分类识别准确率并且兼顾了较快的计算效率。

### 4.3 试验 2: 液压泵性能退化程度识别试验

滑靴磨损是液压泵主要的故障形式之一,对液压泵进行状态监测和滑靴性能退化评估非常重要<sup>[29]</sup>。使用 3 种不同磨损程度的柱塞滑靴组件依次替换正常柱塞滑靴组件(如图 10 所示),来模拟液压泵滑靴磨损故障的退化过程。将 3 种滑靴磨损程度分别命名为轻度、中度、重度滑靴磨损。对液压泵端盖的振动信号进行采集,采样频率为 10 kHz,每次采样时长为 10 s。将正常信号与 3 种故障信号不重叠地截取 48 组数据,每组数据中含有 2 048 个数据点,共获取 196 组样本。



图 10 滑靴磨损柱塞与正常柱塞对比  
Fig. 10 The comparison of loose shoe between wear and normal

为了进一步验证精细复合多尺度方法的优越性,使用了 MSE, MPE, MSDE, RCMDE, RCMFDE 5 种算法

计算 4 种状态下泵端盖振动信号的熵值,绘制误差棒图,结果如图 11 所示。

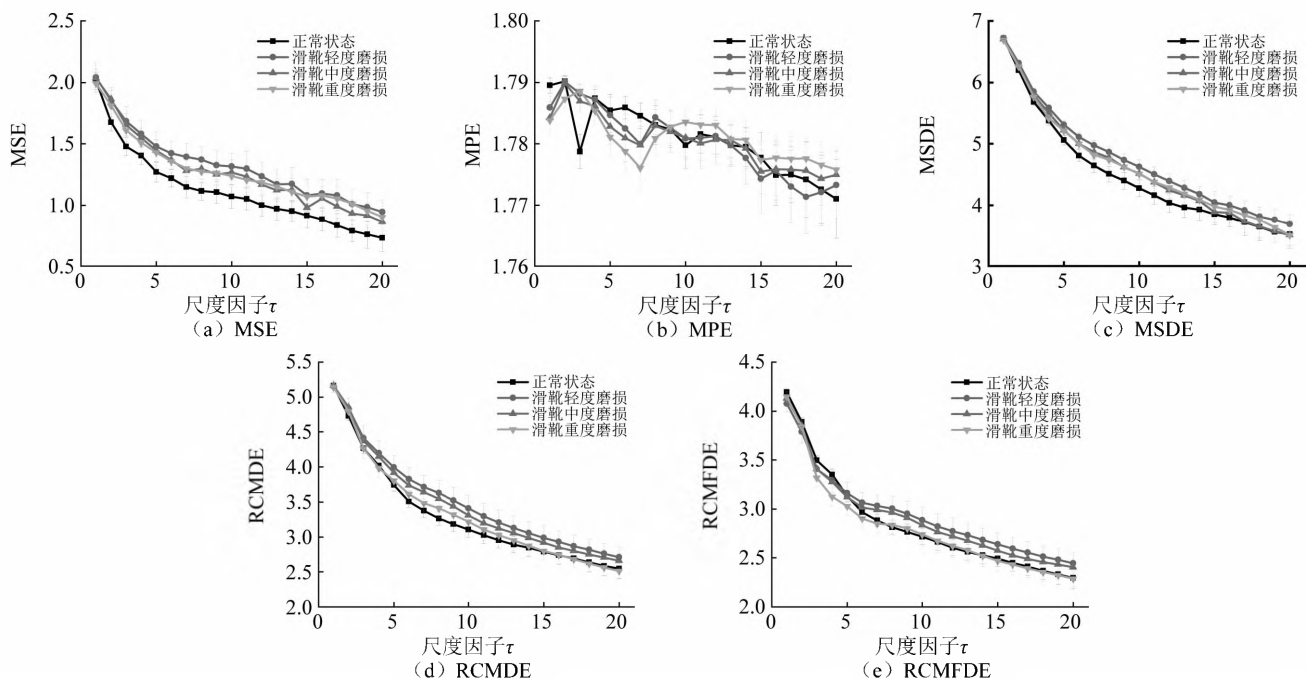


图 11 4 种磨损程度下 5 种熵的误差棒图

Fig. 11 Error bars of the five entropies for four types of wear degrees

从图 11 中可以看出: MSE 的正常状态的曲线和故障状态曲线明显不同,但 3 种故障程度曲线交叉重叠在一起,不利于区分; MPE 的 4 条曲线交叉重叠且误差很大; MSDE 的中、重度磨损曲线几乎完全重叠; RCMDE 的 3 种故障程度曲线有一定的区分度,但正常状态和滑靴重度磨损曲线在较低和较高的尺度都有重叠; RCMFDE 的 3 种故障程度曲线在较高尺度区分度大,正常状态曲线与故障状态曲线在低尺度也有较好的区分性。

依次将 MSE, MPE, MSDE, RCMDE, RCMFDE 的前 2 个 ~ 20 个尺度下的熵值作为特征向量构造特征样本集。训练集和测试集的比例划分为 7 : 3, 划分方法为随机划分,即每类信号中 33 个特征样本为训练集, 15 个特征样本为测试集。选择 10 次分类结果的平均准确率作为最终识别准确率,比较这 5 种算法的识别性能以及特征向量维数对分类结果的影响,采用 PSO - SVM 方法对这些特征样本进行分类识别,结果如图 12 所示。

由图 12 可以看出, RCMFDE 和 RCMDE 算法的分类效果明显高于其余算法, RCMFDE 算法取得了最好的效果,当特征向量维数为 9 时磨损程度平均识别准确率最高,达到了 0.91。MSE, MPE, MSDE, RCMDE, RCMFDE 平均诊断时间分别为 20.2 s, 8.2 s, 7.3 s, 12.6 s 和 10.5 s (计算环境与上文相同), RCMFDE - PSO - SVM 算法在获得最高诊断准确率的同时,耗时

较短。

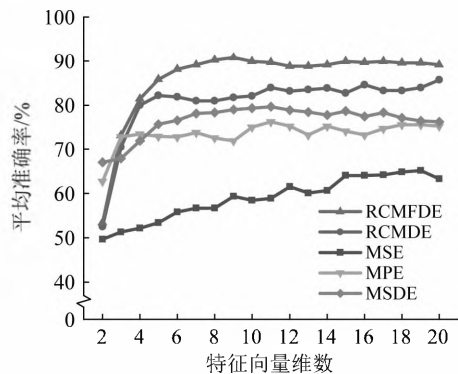


图 12 5 种熵不同特征向量维数的平均识别准确率  
Fig. 12 The average recognition accuracies of different eigenvector dimensions of the five entropies

## 5 结论

将 RCMFDE 特征提取算法引入到液压泵故障诊断领域,提出了一种基于 RCMFDE - PSO - SVM 的液压泵故障诊断方法。

(1) 利用仿真信号对 7 种熵算法 MSE, MPE, MSDE, MDE, RCMDE, MFDE, RCMFDE 的性能进行了比较。结果表明 RCMFDE 能够捕捉到时间序列的波动性,对基线漂移有更强的鲁棒性,并且计算速度更快,稳定性更高。

(2) 对液压泵不同类型故障进行试验,模拟了正常、中心弹簧失效、滑靴磨损、单柱塞松靴和斜盘磨损 5 种状态。利用上述 7 种熵算法进行特征提取并与

SVM, KNN 分类器进行比较, 结果表明 RCMFDE - PSO - SVM 方法诊断准确率最高。

(3) 将该方法用于液压泵性能退化程度的识别, 进行了轻、中、重 3 种滑靴磨损程度的试验, 利用 5 种熵算法 MSE, MPE, MSDE, RCMDE, RCMFDE 进行特征提取, 结果再次证明 RCMFDE - PSO - SVM 方法的诊断准确率最高, 计算时间短。

(4) 基于 RCMFDE - PSO - SVM 的故障诊断方法能有效地提取液压泵的故障特征, 并成功地识别不同类型、不同程度的故障。该方法在液压泵故障诊断和性能退化评估上都具有显著优势。

#### 参 考 文 献

- [ 1 ] PINCUS S M. Approximate entropy as a measure of system complexity [J]. Proceedings of the National Academy of Sciences of the United States of America, 1991, 88 ( 6 ) : 2297 - 2301.
- [ 2 ] 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 ) : H2039 - H2049.
- [ 3 ] BANDT C, POMPE B. Permutation entropy: a natural complexity measure for time series [J]. Physical Review Letters, 2002, 88 ( 17 ) : 174102.
- [ 4 ] LI Y B, YANG Y T, LI G Y, et al. A fault diagnosis scheme for planetary gearboxes using modified multi-scale symbolic dynamic entropy and mRMR feature selection [J]. Mechanical Systems and Signal Processing, 2017, 91 : 295 - 312.
- [ 5 ] ROSTAGHI M, AZAMI H. Dispersion entropy: a measure for time-series analysis [J]. IEEE Signal Processing Letters, 2016, 23 ( 5 ) : 610 - 614.
- [ 6 ] 姜万录, 董克岩, 朱勇, 等. 基于多尺度熵偏均值的液压泵故障特征识别 [J]. 液压与气动, 2016 ( 7 ) : 12 - 17.  
JIANG Wanlu, DONG Keyan, ZHU Yong, et al. Fault feature identification based on partial mean of multiscale entropy for hydraulic pump [J]. Chinese Hydraulics & Pneumatics, 2016 ( 7 ) : 12 - 17.
- [ 7 ] COSTA M, GOLDBERGER A L, PENG C K. Multiscale entropy analysis of complex physiologic time series [J]. Physical Review Letters, 2002, 89 : 068102.
- [ 8 ] MORABITO F C, LABATE D, LA FORESTA F, et al. Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG [J]. Entropy, 2012, 14 ( 7 ) : 1186 - 1202.
- [ 9 ] 王余奎, 李洪儒, 叶鹏. 基于多尺度排列熵的液压泵故障识别 [J]. 中国机械工程, 2015, 26 ( 4 ) : 518 - 522.  
WANG Yukui, LI Hongru, YE Peng. Fault identification of hydraulic pump based on multi-scale permutation entropy [J]. China Mechanical Engineering, 2015, 26 ( 4 ) : 518 - 522.
- [ 10 ] WU S D, WU P H, WU C W, et al. Bearing fault diagnosis based on multiscale permutation entropy and support vector machine [J]. Entropy, 2012, 14 ( 8 ) : 1343 - 1356.
- [ 11 ] AZAMI H. Refined composite multiscale dispersion entropy and its application to biomedical signals [J]. IEEE Transactions on Biomedical Engineering, 2017, 64 ( 12 ) : 2872 - 2879.
- [ 12 ] JAVIER E, HAMED A. Coarse-graining approaches in univariate multiscale sample and dispersion entropy [J]. Entropy, 2018, 20 ( 2 ) : 138.
- [ 13 ] 乔新勇, 顾程, 韩立军. 基于 VMD 多尺度散布熵的柴油机故障诊断方法 [J]. 汽车工程, 2020, 42 ( 8 ) : 1139 - 1144.  
QIAO Xinyong, GU Cheng, HAN Lijun. Diesel engine fault diagnosis method based on VMD and multi-scale dispersion entropy [J]. Automotive Engineering, 2020, 42 ( 8 ) : 1139 - 1144.
- [ 14 ] 吴守军, 冯辅周, 吴春志, 等. 基于 VMD - DE 的坦克行星变速箱故障诊断方法研究 [J]. 振动与冲击, 2020, 39 ( 10 ) : 170 - 179.  
WU Shoujun, FENG Fuzhou, WU Chunzhi, et al. Research on fault diagnosis method of tank planetary gearbox based on VMD - DE [J]. Journal of Vibration and Shock, 2020, 39 ( 10 ) : 170 - 179.
- [ 15 ] 付文龙, 谭佳文, 王凯. 基于 VMD 散布熵与改进灰狼优化 SVDD 的轴承半监督故障诊断研究 [J]. 振动与冲击, 2019, 38 ( 22 ) : 190 - 197.  
FU Wenlong, TAN Jiawen, WANG Kai. Semi-supervised fault diagnosis of bearings based on the VMD dispersion entropy and improved SVDD with modified grey wolf optimizer [J]. Journal of Vibration and Shock, 2019, 38 ( 22 ) : 190 - 197.
- [ 16 ] 柯赞, 宋恩哲, 姚崇, 等. 层次离散熵及其在高压共轨喷油器故障诊断中的应用 [J]. 振动与冲击, 2021, 40 ( 2 ) : 72 - 80.  
KE Yun, SONG Enzhe, YAO Chong, et al. Hierarchical dispersion entropy and its application in fault diagnosis of high pressure common rail injectors [J]. Journal of Vibration and Shock, 2021, 40 ( 2 ) : 72 - 80.
- [ 17 ] 李颖, 王金东, 赵海洋, 等. 基于参数优化 VMD 和 MDE 的往复压缩机轴承故障诊断方法 [J]. 组合机床与自动化加工技术, 2019 ( 4 ) : 120 - 123.  
LI Ying, WANG Jindong, ZHAO Haiyang, et al. Fault diagnosis method of reciprocating compressor bearing based on parameter optimization VMD and MDE [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2019 ( 4 ) : 120 - 123.
- [ 18 ] 李从志, 郑近德, 潘海洋, 等. 基于精细复合多尺度散布熵与支持向量机的滚动轴承故障诊断方法 [J]. 中国机械工程, 2019, 30 ( 14 ) : 1713 - 1719.  
LI Congzhi, ZHENG Jinde, PAN Haiyang, et al. Fault diagnosis method of rolling bearings based on refined composite multiscale dispersion entropy and support vector machine [J]. China Mechanical Engineering, 2019, 30 ( 14 ) : 1713 - 1719.
- [ 19 ] AZAMI H, KINNEY-LANG E, EBIED A, et al. Multiscale dispersion entropy for the regional analysis of resting-state

- magnetoencephalogram complexity in Alzheimer's disease [C]//39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Seogwipo: IEEE, 2017.
- [20] LI Y B, LIANG X H, WEI Y, et al. A method based on refined composite multi-scale symbolic dynamic entropy and ISVM-BT for rotating machinery fault diagnosis [J]. *Neurocomputing (Amsterdam)*, 2018,315: 246-260.
- [21] 丁闯,冯辅周,张兵志,等. 改进多尺度符号动力学信息熵及其在行星变速箱特征提取中的应用[J]. *振动与冲击*, 2020,39(13): 97-102.  
DING Chuang, FENG Fuzhou, ZHANG Bingzhi, et al. MMSDE and its application in feature extraction of a planetary gearbox[J]. *Journal of Vibration and Shock*, 2020,39(13): 97-102.
- [22] AZAMI H, ESCUDERO J. Amplitude-and fluctuation-based dispersion entropy[J]. *Entropy*, 2018,20(3): 210.
- [23] AZAMI H, ARNOLD S E, SANEI S, et al. Multiscale fluctuation-based dispersion entropy and its applications to neurological diseases [J]. *IEEE Access*, 2019, 7: 68718-68733.
- [24] GAN X, LU H, YANG G Y. Fault diagnosis method for rolling bearings based on composite multiscale fluctuation dispersion entropy[J]. *Entropy*, 2019,21(3): 290.
- [25] 姜万录,吴胜强. 基于SVM和证据理论的多数据融合故障诊断方法[J]. *仪器仪表学报*, 2010, 31(8): 1738-1743.
- JIANG Wanlu, WU Shengqiang. Multi-data fusion fault diagnosis method based on SVM and evidence theory [J]. *Chinese Journal of Scientific Instrument*, 2010, 31(8): 1738-1743.
- [26] KENNEDY J, EBERHART R. Particle swarm optimization [C]//Proceedings of ICNN'95-International Conference on Neural Networks. Perth: IEEE, 1995.
- [27] COSTA M, GOLDBERGER A L, PENG C K. Multiscale entropy analysis of biological signals[J]. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 2005,71(2): 21906.
- [28] CASTIGLIONI P, CORUZZI P, BINI M, et al. Multiscale sample entropy of cardiovascular signals: does the choice between fixed-or varying-tolerance among scales influence its evaluation and interpretation? [J]. *Entropy*, 2017,19(11): 590.
- [29] 姜万录,孔德田,李振宝,等. 基于完备总体经验模态分解和模糊熵结合的液压泵退化特征提取方法[J]. *计量学报*, 2020,41(2): 202-209.  
JIANG Wanlu, KONG Detian, LI Zhenbao, et al. Degradation feature extraction method of hydraulic pump based on integrated complete ensemble empirical mode decomposition and fuzzy entropy [J]. *Acta Metrologica Sinica*, 2020,41(2): 202-209.