

# 基于改进时域多尺度散布熵与支持向量机的 转辙机故障诊断

曹 源<sup>1</sup>, 宋 迪<sup>1</sup>, 胡小溪<sup>2</sup>, 孙永奎<sup>1</sup>

(1. 北京交通大学轨道交通运行控制系统国家工程研究中心, 北京 100044;  
2. 北京交通大学轨道交通控制与安全国家重点实验室, 北京 100044)

**摘 要:** 为充分挖掘转辙机振动信号的有效故障信息, 提高故障诊断准确率, 提出了一种基于集合经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)、改进时域多尺度散布熵(Improved Time-domain Multiscale Dispersion Entropy, TMDE)与粒子群优化算法(Particle Swarm Optimization algorithm, PSO)优化支持向量机(Support Vector Machine, SVM)的故障诊断方法. 首先, 通过EEMD方法将不同故障类型的振动信号分解成若干个模态函数(Intrinsic Mode Functions, IMFs); 其次, 采用相关系数与峭度的混合筛选准则筛选IMFs并重构信号; 再次, 应用所提ITMDE算法提取重构信号的多尺度故障特征; 最后将得到的特征向量输入经PSO搜索最优参数后的SVM进行训练和测试. 实验分类准确率为100%, 分析表明所提方法优于传统的多尺度排列熵、多尺度散布熵的故障诊断方法, 能精确地识别转辙机故障类型.

**关键词:** 故障诊断; 转辙机; 集合经验模态分解; 改进时域多尺度散布熵; 支持向量机

**基金项目:** 国家自然科学基金重点项目(No.U1934219); 国家自然科学基金优秀青年基金项目(No.52022010)

**中图分类号:** U284.92

**文献标识码:** A

**文章编号:** 0372-2112(2023)01-0117-11

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20210773

## Fault Diagnosis of Railway Point Machine Based on Improved Time-Domain Multiscale Dispersion Entropy and Support Vector Machine

CAO Yuan<sup>1</sup>, SONG Di<sup>1</sup>, HU Xiao-xi<sup>2</sup>, SUN Yong-kui<sup>1</sup>

(1. National Engineering Research Center of Rail Transportation Operation and Control System, Beijing Jiaotong University, Beijing 100044, China; 2. State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** To fully mine the effective fault information and improve the fault diagnosis accuracy, a fault diagnosis approach for railway point machine is proposed by ensemble empirical mode decomposition (EEMD), improved time-domain multiscale dispersion entropy (ITMDE) and support vector machine (SVM) optimized by particle swarm optimization algorithm (PSO). Firstly, vibration signals with different fault types are decomposed into several intrinsic mode functions (IMFs) by EEMD. Subsequently, the hybrid screening criteria of correlation coefficient and kurtosis (CCKC) are used to screen IMFs and reconstruct the signal. Afterwards, the proposed ITMDE algorithm is employed to extract multiscale fault features from the reconstructed signal. Finally, the PSO is devoted to search the optimal parameters of SVM, with which the obtained feature vectors are trained and tested. The experimental classification accuracy reaches 100%. The results show that the proposed method is superior to the traditional multiscale permutation entropy and multiscale dispersive entropy fault diagnosis methods, and can accurately identify the fault type of the railway point machine.

**Key words:** fault diagnosis; railway point machine; ensemble empirical mode decomposition; improved time-domain multiscale dispersion entropy; support vector machine

**Foundation Item(s):** National Natural Science Foundation of China (No.U1934219, No.52022010)

## 1 引言

转辙机是室外信号设备三大主要设备之一,其故障直接影响列车的安全可靠运行.为了提高其工作的安全性与稳定性,对转辙机进行状态监测和故障诊断十分必要.但传统的基于故障树、专家系统<sup>[1]</sup>的故障诊断方法存在着智能化程度低、主观性影响大的缺陷,因此应用智能故障诊断技术<sup>[2]</sup>,实现对转辙机工作状态的实时感知和故障预警具有十分重要的意义.

目前对转辙机故障诊断的主流方向是基于电压、电流、功率信号的故障诊断方法<sup>[3]</sup>,此方向优点是目前数据积累量大,研究比较成熟.但电信号的诊断方法只能反映电机的输出力即道岔系统的转换阻力情况.因此本文借鉴广泛应用于滚动轴承<sup>[4]</sup>、风电机组等机械设备领域的振动信号故障诊断方法,采集转辙机动作时的振动信号,从机械角度对转辙机进行分析,填补了基于振动信息的转辙机故障诊断领域的空白.

振动信号的分析方法目前已有较为成熟的相关研究,其故障诊断流程主要包括信号获取、信号预处理、特征提取、和故障分类四个步骤.在信号数据预处理阶段,经验模态分解(Empirical Mode Decomposition, EMD)是一种时频分析方法<sup>[5]</sup>,广泛应用于如转辙机振动信号这种非线性、非平稳信号<sup>[6]</sup>.但其端点效应和模态混叠的缺陷会影响后续特征提取和分类准确率.因此Wu Z<sup>[7]</sup>等在EMD的基础上加入高斯白噪声改进后提出了集合经验模态分解法(Ensemble Empirical Mode Decomposition, EEMD),该方法能够有效消除或缓解EMD存在的缺陷问题.振动信号经过EEMD分解,可得到若干个平稳的本征模函数(Intrinsic Mode Functions, IMFs)分量和余项,但不是所有的IMFs都能反映故障特征,其中不乏有含噪声项和残差.文献[8]用相关系数法筛选关键IMFs并对其进行信号重构,但是单纯的相关系数法会遗漏一些能量小但冲击信息高的IMFs分量.

为更加有效地提取振动信号的特征,熵算法被引入一维时间序列信号的故障诊断中,并取得了不错的效果.文献[9]将样本熵(Sample Entropy, SE)用于脑电信号特征提取,文献[10]将排列熵(Permutation Entropy, PE)应用于齿轮箱故障诊断等,都取得了很高的诊断准确率.但是SE计算过程复杂,不适合处理长时间序列信号;PE虽然算法复杂度低计算快,但没有考虑幅值之间的差异.最近Rostaghi<sup>[11]</sup>提出一种新的衡量时间序列复杂度的方法——散布熵(Dispersion Entropy, DE),该算法通过建立映射过程和类别数量收集信号波动信息,保证算法复杂度要求的同时将信号幅值信息计算在内,解决了SE和PE方法的问题.在应用于脑电信号<sup>[12]</sup>、机械

振动信号<sup>[13]</sup>中时取得了稳定的结果.但单纯的DE算法只能提取单一时间尺度特征,而获取信号多尺度特征能更全面提取信号的冲击与随机信息<sup>[14]</sup>.因此,文献[15]在DE基础上开发出了多尺度散布熵(Multiscale Dispersion Entropy, MDE),结果表明MDE可以取得比DE更高的分类准确率.然而MDE在提取多尺度特征过程的粗粒化处理会随着尺度因子的增加出现曲线波动和下采样现象,影响计算精度和准确率.

在智能分类方面应用较多的有支持向量机(Support Vector Machine, SVM)、k邻近算法(k-Nearest Neighbor, k-NN)、人工神经网络(Artificial Neural Network, ANN)、深度学习(Deep Learning, DL)等.文献[16]比较了ANN、SVM、DL应用在旋转机械故障诊断中的优缺点:ANN在复杂非线性序列上有着不错的拟合度,但是它有较多的参数需要设定且需要大量样本进行训练;DL不需要特征提取,但是同样需要大量的样本训练神经网络.k-NN虽然算法简单但在复杂特征的分类准确率上表现不佳.而SVM有着高分类准确率的同时,其基于结构风险最小化的原理机制<sup>[17]</sup>使得它在小样本、非平稳性、非线性和局部极小等的问题上往往能够表现出强大的优越性.考虑到转辙机故障诊断样本量较小、训练复杂度低的要求,本文采用SVM方法用于分类.应用广泛的SVM为具有高斯核的SVM,但是其两个对分类精度影响很大的超参数:惩罚因子 $c$ 和核参数 $g$ 的选择缺乏指导标准<sup>[18]</sup>.因此有学者引入一些优化算法用来对 $c$ 和 $g$ 进行最佳寻优,如遗传算法(Genetic Algorithm, GA)、粒子群优化(Particle Swarm Optimization Algorithm, PSO)等.文献[19]对PSO和GA进行了比较,并通过统计分析和形式假设检验指出,GA在有效性上和PSO上性能接近,但PSO有着非常显著的计算效率的优势.

基于以上研究内容启发,本文提出了一种基于振动信号分析的转辙机故障诊断方法:EEMD方法用于分解降噪,相关系数和峭度双重筛选准则用于IMFs筛选并对信号重构;针对MDE粗粒化过程的弊端,本文通过对其进行尺度平移并求取DE熵值的均值与标准差特征的改进,提出一种改进时域多尺度散布熵(Improved Time-domain Multiscale Dispersion Entropy, ITMDE)算法用于特征提取;最后采用PSO优化SVM分类器用于故障分类.通过实验结果证明了所提方法的实用性和优越性,为转辙机故障诊断领域开辟了一条新的路线.

## 2 基于ITMDE特征的故障诊断方法

针对转辙机振动信号的非线性、非平稳性特征,本文所提转辙机故障诊断方法具体步骤如下所示:

(1)信号采集. 在一定采样频率 $f_s$ 下,利用振动传感器分别采集转辙机不同工况下的振动信号.

(2)数据预处理. 对振动信号进行EEMD分解,得到本征模函数IMFs.

(3)信号重构. 依据相关系数和峭度的双重筛选准则对IMFs进行筛选,挑选出信息量大的IMF后进行信号重构.

(4)特征提取. 对重构信号利用ITMDE算法计算信号熵值,生成特征向量.

(5)故障诊断. 将所得特征数据分为训练集和测试集,用训练集训练PSO-SVM模型后对测试集进行故障诊断分类.

故障诊断方法流程如图1所示.

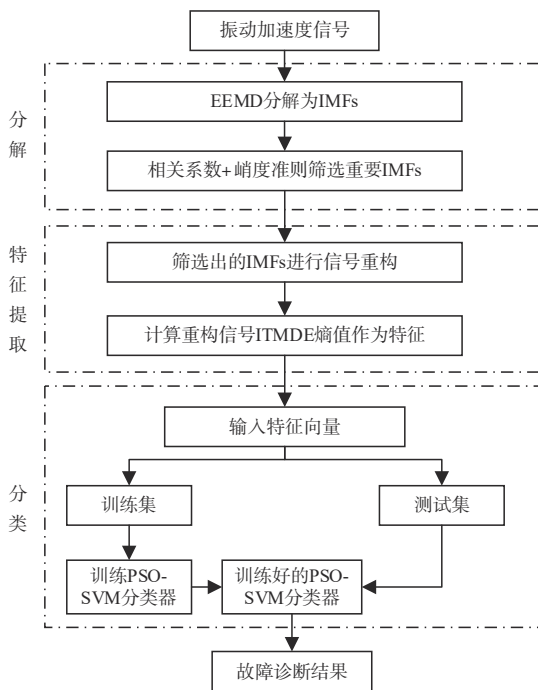


图1 故障诊断方法流程

### 2.1 集合经验模态分解EEMD

EEMD在EMD基础上加入高斯白噪声进行改进来解决EMD存在的问题. 主要步骤如下<sup>[20]</sup>:

步骤1:随机生成均值 $\mu$ 为0,标准差 $\sigma$ 为 $e_n$ 的高斯白噪声 $n_m(t)$ , $t$ 表示时间, $m$ 表示第 $m$ 次EMD分解, $1 \leq m \leq M$ .

步骤2:将 $n_m(t)$ 加入原始信号 $y(t)$ 中,得到新信号 $y_m(t)$ ,即:

$$y_m(t) = y(t) + n_m(t) \quad (1)$$

步骤3:对加入高斯白噪声后的新信号 $y_m(t)$ 进行EMD分解,得到 $S$ 个IMFs分量 $c_{s,m}(t)$ :

$$y_m(t) = \sum_{s=1}^S c_{s,m}(t) + r_m(t) \quad (2)$$

式中, $r_m(t)$ 为分解筛除 $m$ 个IMF后的信号残余分量; $c_{s,m}(t)$ 为进行第 $m$ 次EMD分解后的第 $s$ 个IMFs分量,由频率从高到低排列.

步骤4:若 $m < M$ ,则继续转回步骤2, $m = m + 1$ ,直到达到第 $M$ 次结束.

步骤5:计算 $M$ 次分解所得全部IMFs的均值

$$\bar{c}_s(t) = \frac{1}{M} \sum_{m=1}^M c_{s,m}(t) \quad (3)$$

步骤6: $\bar{c}_s(t)$ 为EEMD分解所得的IMF分量.

### 2.2 相关系数与峭度的IMFs双重筛选准则

皮尔逊相关系数是一种度量两个向量间线性相关度的方法. 峭度是反映信号中冲击含量大小的重要参数. 本文提出一种相关系数与峭度的IMFs双重筛选准则(Correlation Coefficient and Kurtosis Criterion, CCKC), 具体步骤如下.

(1)使用EEMD对转辙机振动信号进行分解,得到IMFs.

(2)计算每个IMF的相关系数和峭度,相关系数的计算公式为:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (4)$$

式中 $X_i, Y_i$ 为时间序列信号, $n$ 为两者的长度, $\bar{X}, \bar{Y}$ 分别为 $X_i$ 和 $Y_i$ 的均值.

峭度计算公式为

$$\text{kur}_i = \frac{1}{M} \sum_{m=1}^M \left( \frac{X_i(m) - \text{mean}(X_i)}{\sigma} \right)^4 \quad (5)$$

式中 $\text{kur}_i$ 为 $X_i$ 的峭度, $M$ 为 $X_i$ 的长度, $\sigma$ 为标准差.

(3)设定相关系数阈值为 $\text{thr}_1$ ,选定峭度阈值为 $\text{thr}_2$ ,各IMF满足两个条件之一即被挑选:(a)相关系数值大于 $\text{thr}_1$ , (b)峭度值大于 $\text{thr}_2$ .

(4)使用挑选出的IMFs对信号进行重构.

### 2.3 改进时域多尺度散布熵

#### 2.3.1 散布熵

作为一种特征提取方法,散布熵DE有着高计算效率和对复杂信号典型特征的高敏感性的特点,其具体计算步骤如文献[21]所描述,DE表达式为

$$\text{DE}(x, m, c, d) = - \sum_{\pi=1}^{c^m} P(\pi_{v_0 v_1 \dots v_{m-1}}) \ln(P(\pi_{v_0 v_1 \dots v_{m-1}})) \quad (6)$$

#### 2.3.2 多尺度散布熵

多尺度散布熵MDE通过对原信号进行粗粒化处理,得到 $\tau$ 个多尺度信号,并分别计算其DE值得到. 粗

粒化过程即对长度为  $N$  的原始时域信号  $x_j (1 \leq j \leq N)$ , 通过式(7)构造  $\tau$  个尺度信号序列:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq \frac{N}{\tau} \quad (7)$$

其中  $\tau=1, 2, \dots$  为尺度因子, 若  $\tau=1$ , 则为原始时域信号; 当  $\tau>1$  时, 原始时域信号被分割为  $\tau$  个长度为  $N/\tau$  的粗粒化序列. 以尺度因子  $\tau=3$  为例, 信号粗粒化过程的原理如图 2 所示.

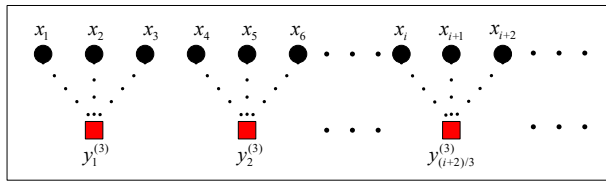


图2 尺度为3的粗粒化过程

对于上述所得  $\tau$  个粗粒序列  $y_j^{(\tau)}$ , 分别计算其散布熵值, 则 MDE 表达式为

$$\text{MDE}(x, m, c, d, \tau) = \text{DE}(y_j^{(\tau)}, m, c, d) \quad (8)$$

### 2.3.3 改进多尺度散布熵

MDE 的粗粒化过程会随着尺度因子  $\tau$  的增大, 使得所得新序列中离散点数大大减少, 从而可能导致获取的熵值所表征的特征信息丢失. 为此改进多尺度散布熵 (Improved Multiscale Dispersion Entropy, IMDE) 采用了连续平移取平均的粗粒化方法<sup>[22]</sup>:

$$y_{k,j}^{\tau} = \frac{\sum_{f=0}^{\tau-1} x_{f+k+\tau(j-1)}}{\tau}, 1 \leq j \leq \frac{N}{\tau}; 1 \leq k \leq \tau \quad (9)$$

求取尺度为  $\tau$  的 IMDE 过程中, 通过宽度为  $\tau$  的“窗”在原始信号上平移滑动的过程, 使得对应每个尺度因子  $\tau$  可以获得  $\tau$  个不同的粗粒化序列  $z_k^{(\tau)} = y_{k,1}^{\tau}, y_{k,2}^{\tau}, \dots, y_{k,\frac{N}{\tau}}^{\tau} (k=1, 2, \dots, \tau)$ . 对于每个尺度因子  $\tau$ , 计算各粗粒化序列  $z_k^{(\tau)}$  的散布熵值, 然后求取所得散布熵的均值即为  $\tau$  尺度散布熵:

$$\text{IMDE}(x, m, c, d, \tau) = \frac{1}{\tau} \sum_{k=1}^{\tau} \text{DE}(z_k^{(\tau)}, m, c, d) \quad (10)$$

以尺度因子  $\tau=3$  为例, IMDE 的粗粒化计算过程如图 3 所示.

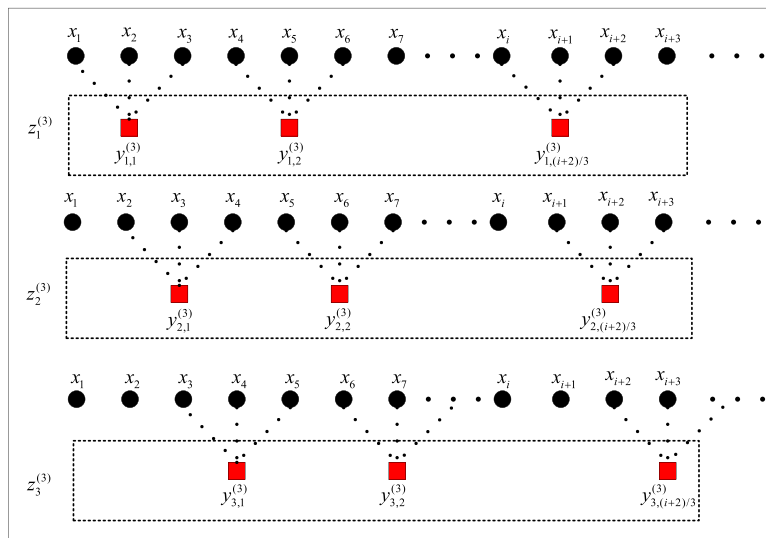


图3 尺度为3的改进粗粒化过程

### 2.3.4 改进时域多尺度散布熵

IMDE 算法中对粗粒化过程所得的多尺度序列的散布熵值进行求取均值, 一定程度上提高了熵值的精确性和计算稳定性. 但是均值操作会掩盖粗粒化平移过程中熵值的突变性. 鉴于此笔者提出在 IMDE 基础上, 除计算平移滑动熵值的均值以外, 加入其他时域特征如标准差的改进时域多尺度散布熵 ITMDE, 从而可以反映出多尺度熵值的离散程度. ITMDE 的标准差部分计算公式如下:

$$\sqrt{\frac{1}{\tau} \sum_{k=1}^{\tau} (\text{DE}(z_k^{(\tau)}) - \text{IMDE}(x, m, c, d, \tau))^2} \quad (11)$$

其中 IDME 为式(14)计算的尺度为  $\tau$  时所得散布熵的均值.

## 2.4 粒子群优化支持向量机

### 2.4.1 支持向量机

SVM 是基于 Vapnik-Chervonenkis (VC) 维数理论和结构风险最小化原则的一种分类方法. 它将输入向量转换到更高维的特征空间, 构造出边际最大化的最佳分离超平面来实现分类, 其原理如图 4 所示. 最常用的是以径向基函数 (Radical Basis Function, RBF) 为核函数的 SVM, 但该方法中的两个超参数  $g$  和  $c$  的确定是分类的关键, 它们的选择决定了误差宽容度与数据映射后的分布, 进而直接影响分类性能.

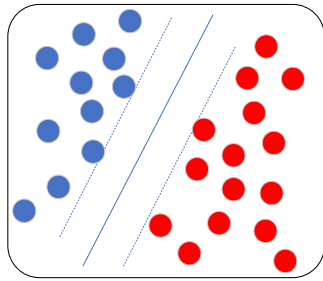


图4 SVM分类原理

2.4.2 粒子群优化算法

粒子群优化算法 PSO 是模仿鸟类移动和寻找食物行为的算法,每个鸟或粒子都有自己的位置和速度,根据距离最佳位置的距离不断更新.其原理如下<sup>[23]</sup>:

$$v_{i+1} = wv_i + c_1r_1(p_{best} - x_i) + c_2r_2(g_{best} - x_i) \quad (12)$$

$$x_{i+1} = x_i + v_{i+1} \quad (13)$$

式中,  $x_i$  是每个粒子的位置,  $v_i$  是每个粒子的速度,  $p_{best}$  是粒子的最优值,  $g_{best}$  是全局最优值,  $r_1$  和  $r_2$  是 0 到 1 间的随机数,  $c_1$  和  $c_2$  是加速系数.

2.4.3 粒子群优化支持向量机

采用 PSO 算法对 SVM 的两个超参数  $g$  和  $c$  进行优化的步骤如图 5 所示.

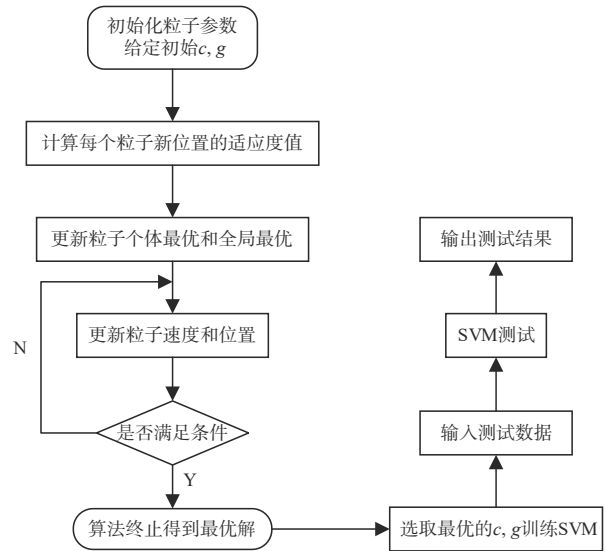


图5 PSO-SVM算法流程

表 1 八种工况数据描述

编号	工况描述	数量	时间长度/s
a	正常工作状态(负载力 4 kN)	58	7.1~8.1
b	负载力过小(约 3 kN)	60	6.8~7.9
c	负载力过大,一级程度(约 5 kN)	60	6.7~8.2
d	负载力过大,二级程度(约 6 kN)	40	7.1~8.3
e	尖轨卡异物,出现摩擦振动现象	40	7.6~8.5
f	卡缺口,表示电路未接通	60	6.5~7.8
g	空载(杆件断裂)	58	7.5~8.3
h	摩擦连接器摩擦力不足,打滑振动	60	6.7~7.7
共计		436	

3 实验与结果分析

3.1 实验设置与数据描述

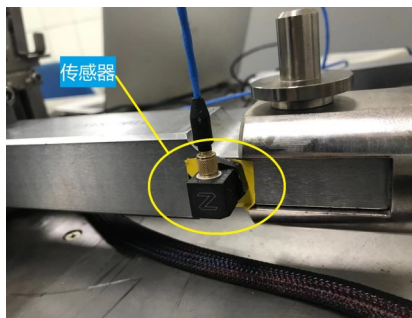
本文所用振动数据采集自西安铁路信号公司的一台 ZDJ9 型转辙机. ZDJ9 型转辙机是铁路系统中最常用的转辙机型号之一,其内部视图如图 6(a)所示.传感器选用 PCB 的 356A16 型号三轴加速度振动传感器,安装于转辙机的动作杆处,如图 6(b)所示.利用 NI 动态数据采集卡连接计算机进行数据采集,采集频率 5.12 kHz,实验装置如图 6(c)所示.

实验采集八种工况共 436 组数据,每种工况数据约 50 组,八种工况的数据描述如表 1 所示,每种工况对应时

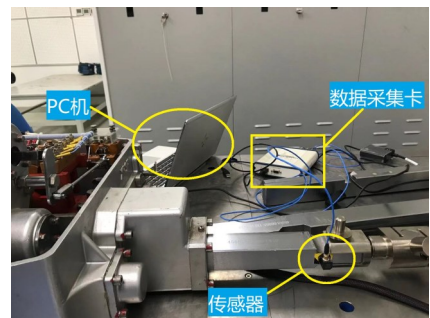
域波形如图 7 所示.从图中可看出 a、b、c、d 四个工况为不同负载力工况,时域波形十分相近:第一个较小幅值的冲击为表示电路断开,动节点打入另一侧时产生;中间部分为电机带动减速器再到摩擦连接器,最后带动滚珠丝杠将旋转运动转为直线运动带动动作杆进行转换;最后的大幅值冲击为转换完成,另一侧的自动开闭器打入到静



(a) ZDJ9 转辙机内部结构



(b) 振动传感器安装位置



(c) 振动信号采集系统

图6 ZDJ9转辙机数据采集平台

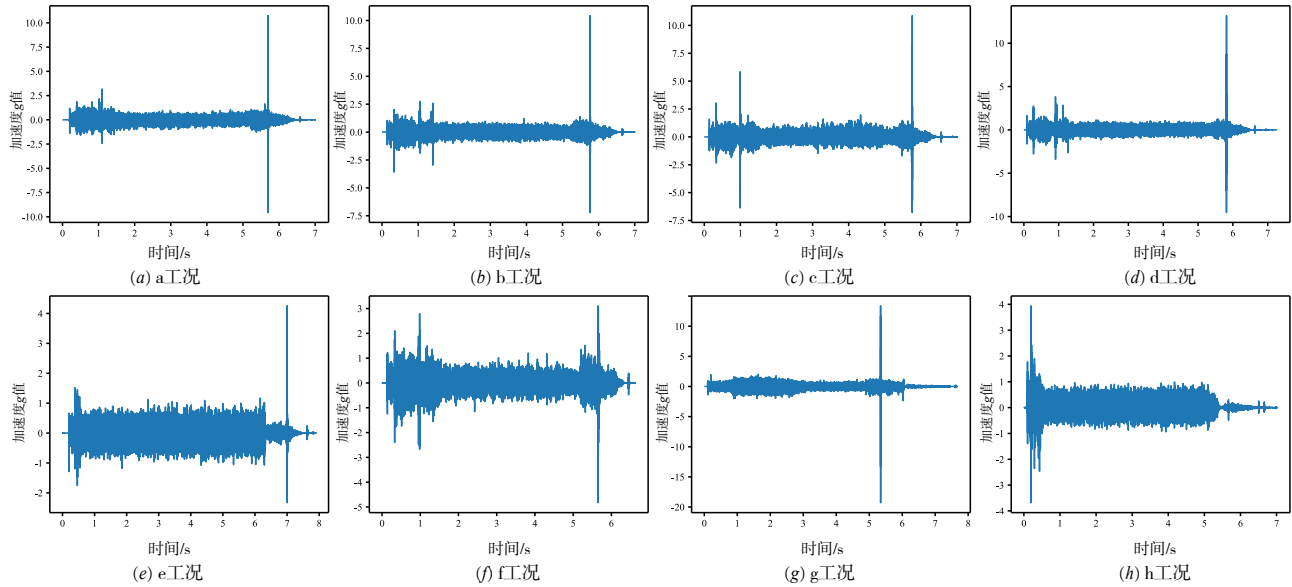


图7 八种工况下的振动信号时域波形

节点,接通表示电路.当尖轨卡异物时(工况e),时域波形会呈现出幅值相近、高频的振动现象.工况f与工况c十分相近,唯一的区别为在工况f下,转换完成后动节点未打入静节点,达到中间位置,其冲击幅值小.空载工况g与正常负载a工况相比,主要特征为两个冲击部分动作时间更短,启动冲击更小.而工况h由于摩擦连接器摩擦力不足,导致初始波形冲击幅度高,随后是平稳的打滑波形,且最终由于输出扭矩不足开闭器未能闭合.

### 3.2 数据预处理

按照2.1、2.2节描述的方法对信号进行EEMD分解、IMFs筛选与信号重构.以一个工况c数据为例进行EEMD分解,分解结果如图8所示.随后计算各IMFs与原始信号的相关系数,如图9所示.由结果可以看出随着IMF阶数增大,其相关系数呈现出越来越低趋势,即信号频率和能量越来越低,IMF5以后相关系数已经在0.1以下.但观察图8可知在IMF5以后仍有部分IMF有着明显的冲击特征,计算各IMF峭度值如图10所示.

如结果所示,IMF6~IMF9虽然相关系数低,但仍具有较高的峭度值,说明其含有反映信号冲击的信息,如果仅考虑相关系数而将其舍弃的话将会丢失信号关键特征,从而影响后续的特征提取和分类.按照2.2节描述的方法,本文选取 $\text{thr}_1=0.3$ , $\text{thr}_2=5$ 对EEMD分解后的IMFs进行挑选并重构.

### 3.3 特征提取

为验证本文所提方法的优势,将ITMDE与机械故障诊断中广泛应用的多尺度排列熵(Multiscale Permutation Entropy, MPE)算法、未做改进的MDE算法、初步

改进的IMDE算法进行比较.

提取多尺度散布熵时需要4个参数选择:类别个数 $c$ 、嵌入维数 $m$ 、时间延迟 $d$ 以及尺度因子 $\tau$ .文献[11]建议嵌入维数 $m$ 和类别 $c$ 取值不宜过小或过大, $m$ 通常取2或3, $c$ 取4~8间的整数,时延 $d$ 一般取1,处理的数据长度应大于2000.由于MDE,IMDE,ITMDE主要区别在于粗粒化过程的不同,求取DE熵值过程是一致的,故 $c$ 、 $m$ 、 $d$ 的选取相同.至于尺度因子 $\tau$ 的选取,当 $\tau$ 过小时,难以从拾取的转辙机振动信号中完整地提取出状态特征信息;而当 $\tau$ 取值较大时,由于信号粗粒化处理所得新序列中有效数据点数大幅减少,容易导致计算结果不稳定且获取的熵值不准确.经过实验对比以及数据长度考虑(30000~40000),尺度 $\tau$ 选择20.全部参数设置为: $m=2$ , $c=6$ , $d=1$ , $\tau=20$ .

在8个工况中分别随机选取数据计算MPE、MDE以及IMDE熵值曲线,如图11所示.由图11(a)和(b)可以看到在MPE与MDE的熵值曲线对比中,MPE的熵值混叠性较为严重,而MDE的熵值分布则较为散布,更易区分.图11(b)和(c)的对比中可以发现IMDE与MDE相比,熵值曲线更为平滑和收敛,在几个混叠区域的区分度也较之MDE有了提升,鲁棒性更强.

但是由图11(c)观察可知,IMDE的熵值曲线中,仍有高度重合的工况,如工况c和f.随机取工况c和f中两个数据样本进行比较,计算其IMDE熵值如图12所示.可见两者熵值曲线的幅值和走势高度一致,很难区分.因此计算其ITMDE即熵值标准差曲线,结果如图13所示,可见其熵值曲线有明显的区分度.即在尺度因子 $\tau$ 平移过程中,尽管两者熵值均值近似,

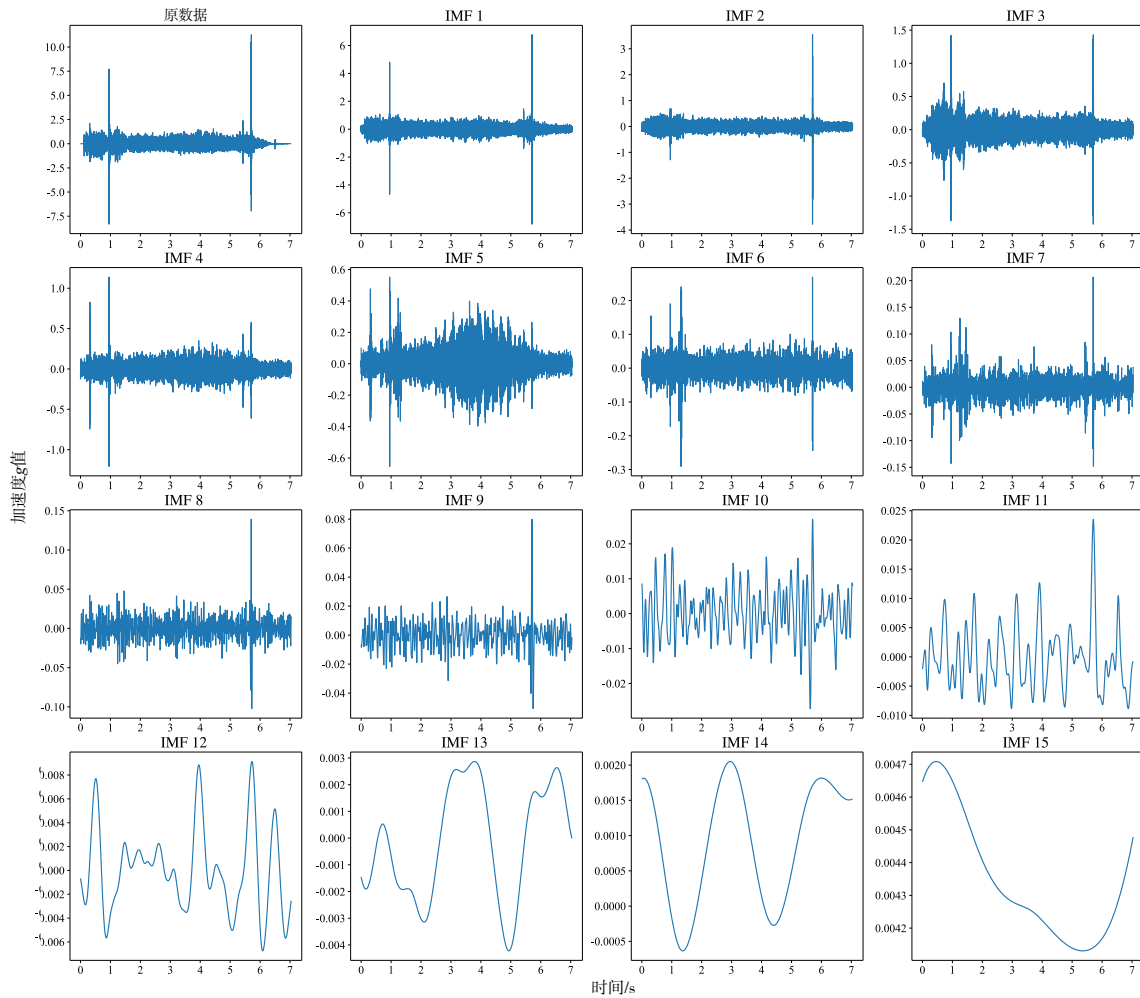


图8 EEMD分解结果

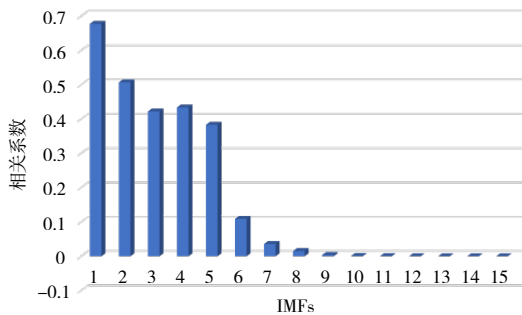


图9 各IMF相关系数

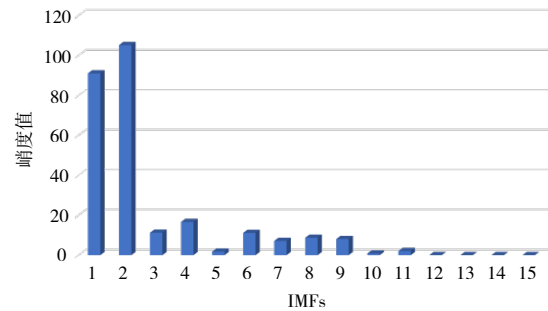


图10 各IMF峭度值

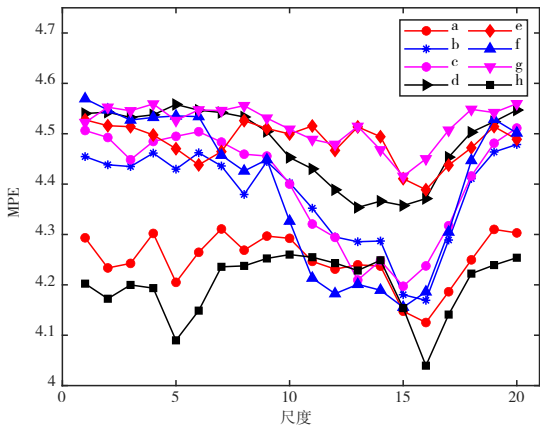
但其离散度有着显著区别. 证明以熵值标准差特征补充IMDE熵值均值特征区分度的可行性和必要性.

### 3.4 故障分类结果对比分析

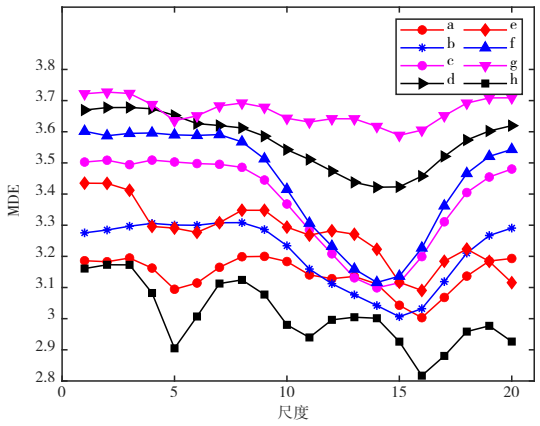
对所有8种工况的436组数据按照前文所述方法进行EEMD分解,相关系数与峭度准则筛选IMF并重

构,然后根据所提ITMDE算法生成特征向量后使用PSO-SVM进行诊断分类. 其中60%即262组数据用于训练集,40%即174组数据用于测试集.

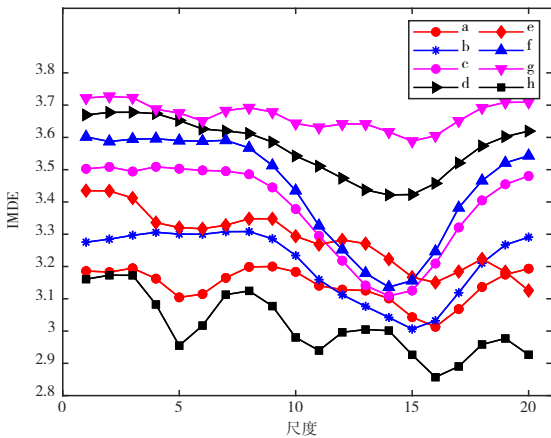
首先将训练集特征向量输入PSO-SVM分类器中进行训练. PSO-SVM参数选择为:PSO粒子维数2;种群数量100;种群进化代数40;学习因子 $c_1$ 和 $c_2$ 分别为0.2和



(a) MPE



(b) MDE



(c) IMDE

图 11 不同熵值曲线比较

0.5; SVM 的参数  $c$  和  $g$  寻优范围为 0~10. 得到最优核函数宽度  $g$  和惩罚参数  $c$  后对测试集进行测试, 测试结果分类准确率达到 100%. 为了更好验证本文方法的优越性, 对数据预处理方法、特征提取算法、分类模型与其他算法进行结果对比分析.

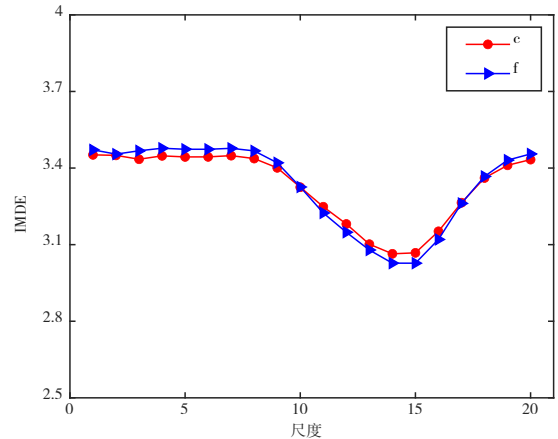


图 12 c 和 f 工况 IMDE 熵值曲线

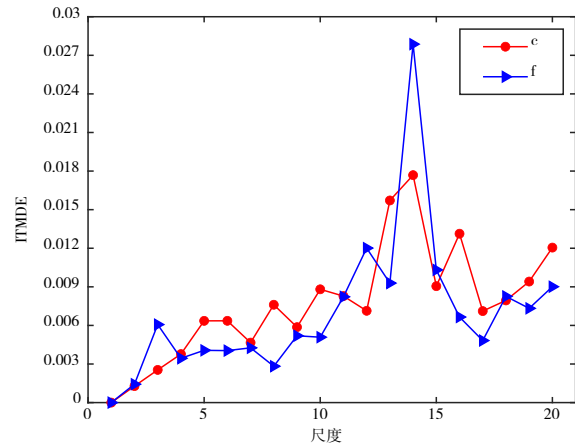


图 13 c 和 f 工况 ITMDE 熵值曲线

### 3.4.1 特征提取方法结果比较

分别以 MPE、MDE、IMDE、ITMDE 算法作为特征提取方法, 得到特征向量后输入 PSO-SVM 分类器, 诊断结果的混淆矩阵如图 14 所示.

从图 14(a) 和 (b) 可以看出, MPE 仅在工况 a 和 g 上全部识别正确, 其他工况均有识别混淆的情况. 如将工况 e 和 h 识别为正常工况的情况, 主要原因为排列熵算法未考虑幅值特性, 对 e 和 h 工况中非典型数据未能提取其前后冲击部分的幅值特征. 而 MDE 则对工况 e 和 h 实现了准确识别, 总体准确率比 MPE 高出 1.15%, 证明了散布熵算法在考虑了信号幅值特征后的优越性.

在图 14(b) 和 (c) 的对比中发现, MDE 在不同负载力故障工况的识别上仍有不足, 如工况 b 和 d. 混淆原因为不同负载力工况本身波形与特征相似, 而 MDE 算法在进行粗粒化处理时会随着尺度因子  $\tau$  的增大使得熵值精确性降低. IMDE 则通过尺度平移操作能够更精确地提取相似数据的特征, 从而实现对工况 b 和 d 的准确识别.

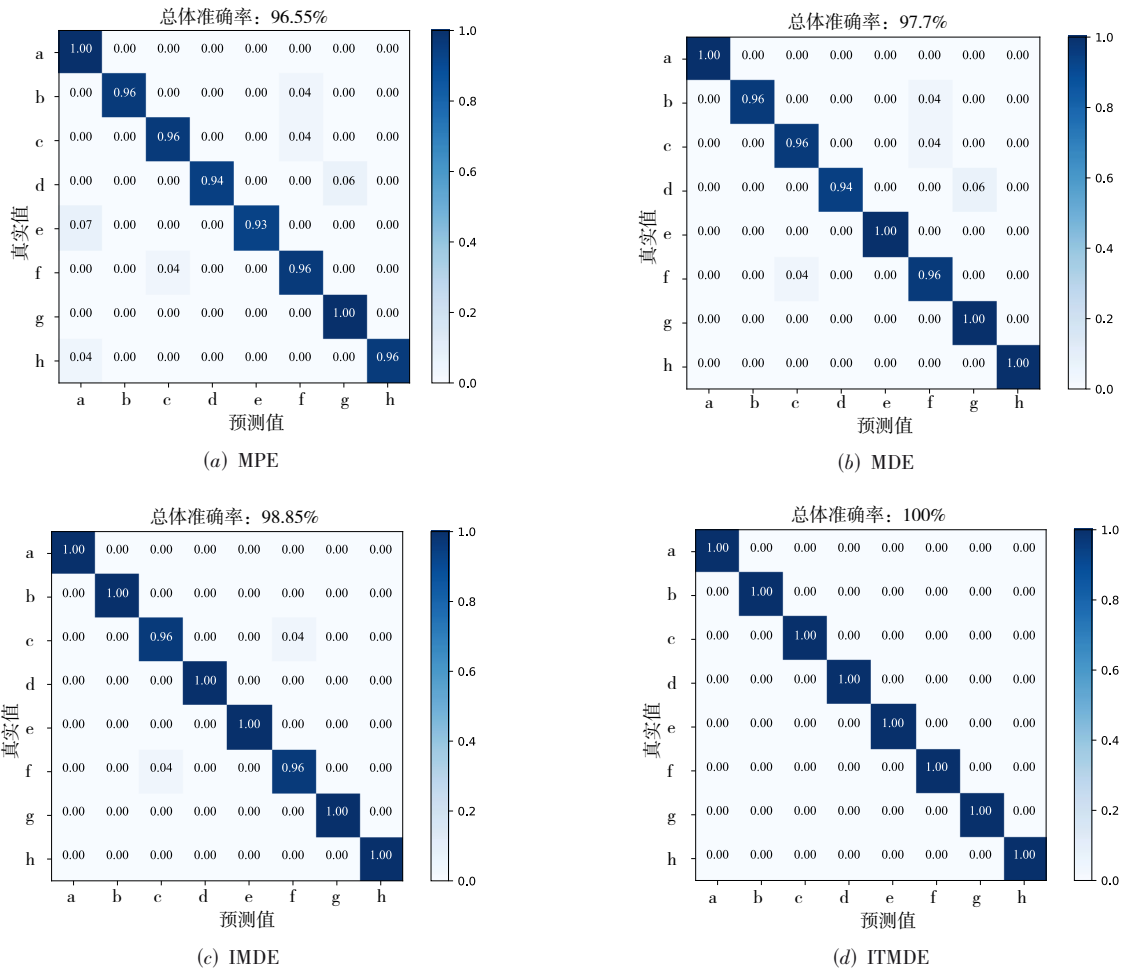


图 14 不同特征提取方法故障分类结果

但在某些工况如 c 和 f 工况上 IMDE 仍有识别混淆的现象,由 3.3 节分析可知这是由于两者部分工况 IMDE 特征幅值与走势高度一致导致. 而 ITMDE 通过加入 IMDE 熵值离散度的获取改进后,达到了 100% 的识别准确率. 证明了本文所提改进时域多尺度散布熵方法在特征提取上的优势.

3.4.2 分类方法结果比较

为了验证 SVM 在分类方面的优越性, k-NN、BP (Back Propagation neural networks, BP) 神经网络被用于对比分析, BP 神经网络为 ANN 的一种典型神经网络. 为控制变量, 特征向量均采用所提 EEMD+相关系数与峭度筛选准则(CCKC)+ITMDE 的方法获取. 不同分类器的分类结果比较如表 2 所示.

表 2 不同分类器诊断分类结果

分类器	预处理与特征提取方法	分类准确率/%
BP	EEMD+CCKC+ITMDE	90.2
k-NN	EEMD+CCKC+ITMDE	91.4
SVM	EEMD+CCKC+ITMDE	93.1

可以看到在分类器选择上 SVM 在分类准确度上明显优于 ANN 和 k-NN. 但如前文所述, SVM 中 g 和 c 超参数的选择对分类效果的影响非常大, 随机选择参数的 SVM 分类结果只有 93% 左右, 证明了参数优化的必要性.

而为了验证 PSO 在优化 SVM 参数的优越性, 两种不同的 SVM 优化算法: 网格搜索优化 SVM (Grid-SVM) 和遗传算法优化 SVM (GA-SVM) 被用于对比分析. 不同优化方法的分类结果如表 3 所示. 可以发现, Grid-SVM 和 GA-SVM 算法的分类准确率虽均达到了 98% 以上, 显著提高了 SVM 的分类效果, 但两者在运行效率上均逊于 PSO-SVM 算法, 且只有 PSO-SVM 方法达到了 100% 的准确率. 证明了 PSO-SVM 分类器在分类上的优越性.

表 3 不同优化方法诊断分类结果

分类器	分类准确率/%	运行时间/s
Grid-SVM	98.3	15.3
GA-SVM	98.85	12.5
PSO-SVM	100	9.9

### 3.4.3 预处理方法结果比较

最后为了验证所提EEMD+CCKC筛选IMFs并重构预处理方法的必要性,将原始信号不做处理、EMD分解+CCKC以及EEMD+仅相关系数(CC)筛选IMFs重构的不同预处理方法,按照相同的ITMDE算法提取特征、PSO-SVM分类器分类步骤进行试验,试验结果如表4所示。

可以看出不做分解降噪处理的原始信号进行分类的结果只有86.78%,证明EEMD分解处理效果明显优于EMD分解。而采用CCKC方法筛选IMFs重构后的分类结果比仅用CC方法进行筛选后的结果高1.14%。证明了EEMD+CCKC预处理方法的必要性。

表4 不同数据预处理方法诊断分类结果

预处理方法	IMFs筛选方法	特征提取与分类方法	整体准确率/%
原始信号	—	ITMDE+PSO-SVM	86.78
EMD	CCKC	ITMDE+PSO-SVM	95.98
EEMD	CC	ITMDE+PSO-SVM	97.12
EEMD	CCKC	ITMDE+PSO-SVM	100

通过以上对比分析可以看到,本文提出的基于EEMD分解、相关系数与峭度双重IMFs筛选准则、ITMDE算法特征提取、PSO-SVM分类的转辙机故障诊断方法,即使在训练样本数量相对较小的情况下仍然达到了100%的准确率,优于其他传统故障诊断方法。表明所提方法能够有效识别转辙机不同复杂故障工况,可以更好地适用于现场实际采集故障样本数量不足的情况,具有广阔的应用前景。

## 4 结论与展望

本文提出一种基于EEMD分解、ITMDE与PSO-SVM的转辙机工况识别方法。实验表明,ITMDE算法弥补了IMDE算法均值化过程中忽略熵值离散性的缺点,所得计算结果更精确,稳定性更好。利用PSO优化算法对SVM进行优化可以避免人为选择关键参数带来的主观误差,能够获得更准确的分类识别结果。利用本文所述的故障分类方法,可以实现对转辙机不同故障状态的分类,对转辙机的状态监测和故障诊断具有一定的参考借鉴价值。

### 参考文献

[1] 杜洁. 基于故障树技术的铁路信号设备故障诊断专家系统的实现方法研究[D]. 北京: 北京交通大学, 2009.  
DU J. Research on Expert System Implementation of Railway Signal Equipment Fault Diagnosis Based on Fault Tree Analysis[D]. Beijing: Beijing Jiaotong University, 2009. (in Chinese)

[2] CAO Y, LI P, ZHANG Y Z. Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing[J]. Future Generation Computer Systems, 2018, 88: 279-283.

[3] LI Z, YIN Z, TANG T, et al. Fault diagnosis of railway point machines using the locally connected autoencoder[J]. Applied Sciences, 2019, 9(23): 5139.

[4] SHAO K X, FU W L, TAN J W, et al. Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris Hawks optimization-based SVM for fault diagnosis of rolling bearing[J]. Measurement, 2021, 173: 108580.

[5] HUANG N, 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 of the Royal Society A: Mathematical, Physical and Engineering Sciences, 1998, 454(1971): 903-998.

[6] CAO Y, ZHANG Y Z, WEN T, et al. Research on dynamic nonlinear input prediction of fault diagnosis based on fractional differential operator equation in high-speed train control system[J]. Robotics & Maching Learning, 2019, 29(1): 340-346.

[7] WU Z H, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.

[8] 张安安, 黄晋英, 冀树伟, 等. 基于卷积神经网络图像分类的轴承故障模式识别[J]. 振动与冲击, 2020, 39(4): 165-171.  
ZHANG A A, HUANG J Y, JI S W, et al. Bearing fault pattern recognition based on image classification with CNN[J]. Journal of Vibration and Shock, 2020, 39(4): 165-171. (in Chinese)

[9] 高军峰, 司慧芳, 余彬, 等. 基于脑电样本熵的测谎分析[J]. 电子学报, 2017, 45(8): 1836-1841.  
GAO J F, SI H F, YU B, et al. Lie detection analysis based on the sample entropy of EEG[J]. Acta Electronica Sinica, 2017, 45(8): 1836-1841. (in Chinese)

[10] 王志坚, 常雪, 王俊元, 等. 排列熵优化改进变模态分解算法诊断齿轮箱故障[J]. 农业工程学报, 2018, 34(23): 59-66.  
WANG Z J, CHANG X, WANG J Y, et al. Gearbox fault diagnosis based on permutation entropy optimized variational mode decomposition[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(23): 59-66. (in Chinese)

[11] ROSTAGHI M, AZAMI H. Dispersion entropy: A measure for time-series analysis[J]. IEEE Signal Processing Letters, 2016, 23(5): 610-614.

[12] 张学军, 景鹏, 何涛, 等. 基于变分模态分解的癫痫脑电

- 信号分类方法[J]. 电子学报, 2020, 48(12): 2469-2475.
- ZHANG X J, JING P, HE T, et al. An epileptic electroencephalogram signal classification method based on variational mode decomposition[J]. Acta Electronica Sinica, 2020, 48(12): 2469-2475. (in Chinese)
- [13] YAN X A, JIA M P. Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection[J]. Knowledge-Based Systems, 2019, 163: 450-471.
- [14] COSTA M, GOLDBERGER A L, PENG C K. Multiscale entropy analysis of complex physiologic time series[J]. Physical Review Letters, 2002, 89(6): 068102.
- [15] AZAMI H, ESCUDERO J. Coarse-graining approaches in univariate multiscale sample and dispersion entropy[J]. Entropy (Basel, Switzerland), 2018, 20(2): 138.
- [16] LIU R N, YANG B Y, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: A review[J]. Mechanical Systems and Signal Processing, 2018, 108: 33-47.
- [17] ZHANG X Y, LIANG Y T, ZHOU J Z, et al. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM[J]. Measurement, 2015, 69: 164-179.
- [18] KRAMAR V A, ALCHAKOV V V, DUSHKO V R, et al. Application of support vector machine for prediction and classification[J]. Journal of Physics: Conference Series, 2018, 1015: 032070.
- [19] DWIVEDI R, DIKSHIT O. A comparison of particle swarm optimization (PSO) and genetic algorithm (GA) in second order design (SOD) of GPS networks[J]. Journal of Applied Geodesy, 2013, 7(2): 135-145.
- [20] 张敏, 蔡振宇, 包珊珊. 基于 EEMD-Hilbert 和 FWA-SVM 的滚动轴承故障诊断方法[J]. 西南交通大学学报, 2019, 54(3): 633-639, 662.
- ZHANG M, CAI Z Y, BAO S S. Fault diagnosis of rolling bearing based on EEMD-Hilbert and FWA-SVM[J]. Journal of Southwest Jiaotong University, 2019, 54(3): 633-639, 662. (in Chinese)
- [21] 车一鸣, 王冬梅, 王国兴, 等. 基于 IMDE 和 ORF 模型的断路器工况识别[J]. 组合机床与自动化加工技术, 2019 (12): 85-90.
- CHE Y M, WANG D M, WANG G X, et al. Working condition identification of circuit breaker based on IMDE and ORF model[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2019(12): 85-90. (in Chinese)
- [22] LIU B, WANG R M, PENG Z Q, et al. Identification of denatured biological tissues based on compressed sensing and improved multiscale dispersion entropy during HIFU treatment[J]. Entropy (Basel, Switzerland), 2020, 22(9): 944.
- [23] 任强, 官晟, 王凤军, 等. 基于 EEMD 和 PSO-SVM 的电机气隙偏心故障诊断[J]. 组合机床与自动化加工技术, 2021(2): 73-76, 85.
- REN Q, GUAN S, WANG F J, et al. Motor air-gap eccentricity fault diagnosis based on EEMD and PSO-SVM[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(2): 73-76, 85. (in Chinese)

### 作者简介



曹 源 男, 1982 年出生于河南省开封市. 现为北京交通大学电子信息工程学院教授、博士生导师. 主要研究方向为列车运行控制系统健康管理.

E-mail: ycao@bjtu.edu.cn



宋 迪 (通讯作者) 男, 1992 年出生于河北省衡水市. 现为北京交通大学电子信息工程学院硕士研究生. 主要研究方向为转辙机故障诊断.

E-mail: 19120243@bjtu.edu.cn



胡小溪 男, 1995 年出生于湖北省襄阳市. 现为北京交通大学电子信息工程学院博士研究生. 主要研究方向是道岔转换系统的状态监测与寿命预测.



孙永奎 男, 1993 年出生于河南省永城市. 2021 年毕业于北京交通大学电子信息工程学院, 获得交通信息工程及控制专业博士学位. 现为北京交通大学电子信息工程学院副教授. 主要研究方向为列车运行控制系统故障诊断.