

DOI: 10.19650/j.cnki.cjsi.J1904768

基于 CEEMDAN-SQI-SVD 的齿轮箱局部故障特征提取*

古莹奎, 曾磊, 张敏, 李文飞

(江西理工大学机电工程学院 赣州 341000)

摘要:经验模态分解(EMD)及以其为基础发展而来的方法在故障诊断领域中得到广泛应用,对于分解后固有模态函数(IMF)的有效选择及基于有效 IMF 故障特征的准确提取至关重要。为更高效地解决此类问题,提出一种基于具有自适应白噪声的完整集成经验模态分解(CEEMDAN)结合信号质量指数(SQI)算法与奇异值分解(SVD)的齿轮箱局部故障最优特征提取算法。以具有不同故障级别的齿轮局部裂纹进行试验验证方法的有效性,通过试验获取原始数据并进行 CEEMDAN 分解,利用 SQI 进行有效 IMF 选取,再结合 SVD 对有效 IMF 进行分解以获取最优特征向量,并输入至 BP 神经网络进行训练与测试,最后将测试结果与数种常规方法进行比较。结果表明,针对齿轮箱的局部故障,提出的 CEEMDAN-SQI-SVD 算法识别精度高,并优于数种常规方法。

关键词:故障诊断;固有模态函数;自适应白噪声的完整集成经验模态分解;特征提取;BP 神经网络

中图分类号: TH165.3 TH132 **文献标识码:** A **国家标准学科分类代码:** 460.4099

Feature extraction method for gearbox local fault based on CEEMDAN-SQI-SVD

Gu Yingkui, Zeng Lei, Zhang Min, Li Wenfei

(School of Mechanical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: Empirical Modal Decomposition (EMD) and the methods based on EMD have been widely used in the field of fault diagnosis. The selection of Intrinsic Mode Function (IMF) after decomposition is important for accurate extraction of fault features. To solve such problem more effectively, the gearbox local fault optimal feature extraction algorithm based on Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) combined with Signal Quality Index (SQI) algorithm and Singular Value Decomposition (SVD) is proposed in this study. The method is evaluated by experiments on local crack of gear with different fault levels. Firstly, the original data are obtained by experiment. Then, they are decomposed by CEEMDAN. The effective IMF is decomposed by SVD to obtain the optimal feature vector, which is the input of BP neural network for training and test. Finally, the test results are compared with several common methods. Experimental results show that the proposed CEEMDAN-SQI-SVD algorithm has high recognition accuracy and is better than several conventional methods for local fault of gearbox.

Keywords: fault diagnosis; intrinsic mode function; complete ensemble empirical mode decomposition with adaptive noise; feature extraction; BP neural network

0 引言

齿轮箱是机械动力传输中的重要组成部分,广泛应用于工业领域。由于齿轮箱通常在复杂的环境下工作,齿轮及轴承易受疲劳点蚀、剥落和断裂等损伤模式的影响,

导致传动效率下降并带来安全隐患。为保证齿轮箱正常运行,多种新的状态监测与故障诊断技术被应用于齿轮箱的健康管理。

傅里叶变换在齿轮箱故障诊断中应用广泛,但傅里叶变换是基于线性系统和稳定信号的假设前提下应用的方法,存在一定的局限性。在实际中,发生故障的齿轮箱

收稿日期:2019-02-25 Received Date:2019-02-25

* 基金项目:国家自然科学基金(61463021)、江西省青年科学家培养对象计划(20144BCB23037)、江西省自然科学基金(20181BAB202020)项目资助

通常伴有冲击和强噪音等干扰因素,使振动信号表现出非线性与非稳定的特性。针对传统傅里叶变换的不足,提出了小波变换时频分析方法。但小波变换本身也存在一些不可避免的缺陷,如小波基函数的选择和降噪时阈值的选择通常根据实际情况及技术人员经验确定,具有较强的主观性且存在能量泄露等问题^[1]。

一些自适应信号分解方法也被用于处理非线性非稳定性信号,如变分模态分解(variational mode decomposition, VMD)、经验小波变换(empirical wavelet transform, EWT)和经验模态分解(empirical mode decomposition, EMD)等。VMD通过建立对应的约束变分模型表达式,引入二次惩罚因子和拉格朗日乘法算子,将约束性变分问题转换为非约束性变分问题,从而获得具有数学意义并且能有效抑制模态混叠现象的带限固有模态函数^[2]。但VMD方法也存在一些缺陷。首先,在用VMD方法分解信号之前,必须预先设定模态函数与惩罚函数;其次,在用VMD方法对信号分解后,进行滤波或提取特征值时,需选择合适的分量和排序,但VMD方法分解出来的分量排序不规律,从而给分量的选择和排序带来困难^[3-5]。EWT方法通过对信号的傅里叶频谱进行自适应分割构造正交的小波滤波器组,从而提取出具有紧支撑傅里叶频谱的调幅-调频单分量成分^[6]。EWT具有良好的自适应性,但由于其粗糙的频谱分割问题,在处理噪声及非平稳信号方面仍有所欠缺^[7-8]。EMD方法依据输入振动信号自身的特点,自适应地将振动信号分解成若干个固有模态函数(IMF)之和^[9]。该方法被认为是傅立叶变换和小波变换等传统时频分析方法的重大突破。EMD方法与其他方法相结合,如神经网络^[10-11]、贝叶斯分类器^[12]、模糊神经网络^[13]、小波降噪^[14]等,有效地提高了故障诊断的效率和精度。

近年来,以EMD为基础,提出了集成经验模态分解(ensemble empirical mode decomposition, EEMD)和具有自适应白噪声的完整集成经验模态分解(complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN)等方法,均是对EMD出现的模态混叠现象进行改进与优化的方法,且在齿轮箱的故障诊断中得到了应用^[15-17]。

CEEMDAN方法是以EMD与EEMD方法为基础发展而来的一种新方法^[18]。在分解的每一阶段都添加自适应的白噪声,通过计算唯一的余量信号以获取各个模态分量。与EEMD方法相比,无论集成次数为多少,重构误差几乎为0,其分解过程具有完整性,克服了EEMD分解效率低的问题以及EMD的模态混叠问题,分解效果优于EMD及EEMD^[19-21]。

针对EMD及以EMD为基础发展而来的方法,对不同信号分解得到的IMF个数是不同的,通常只有一部分

IMF中包含故障信息,而其他IMF则是与故障无关的分量,或是由于受边界效应、分解迭代误差及实际环境噪声等因素影响而产生的虚假分量。因此,在进行齿轮箱故障诊断中需要选取含故障信息的有效IMF,然后对有效IMF进行不同方法的处理以达到故障诊断的目的。在选取有效IMF的方法上,目前并没有统一的标准,通常简单的以相关系数^[11,22-24]、频率高低^[25-26]以及峰度值大小^[17,27]等单个指标作为选取有效IMF的依据。在有效IMF个数的确定上同样没有统一的标准,通常根据实际情况及专业技术人员经验确定。而有效IMF的选取是进行故障诊断至关重要的一步,其选取质量的高低直接影响后续处理结果的精度。为解决有效IMF选取的问题,本文以齿轮箱齿轮不同失效程度的局部故障为对象,以分解效果较优的CEEMDAN为基础,提出一种最优有效IMF选取的SQI算法,并结合有效IMF的奇异值阵构成最优故障特征向量,输入至BP神经网络进行训练并识别,通过试验将该方法与常规方法进行比较。

1 算法模型

1.1 EMD与EEMD理论

EMD算法是一种自适应性的数据分析方法,它将一个复杂的信号分解为有限个表征信号特征时间尺度的IMF的和,每个IMF均满足以下2个条件^[9]:1)信号中的极值点数目和零点数目相等或最多相差1;2)在信号任意一点,由局部极大值点和局部极小值点构成的2条包络线平均值为0。

依据此2个EMD的IMF分解条件,给定一个原始信号数据序列,其EMD分解结果可表示为:

$$S(t) = \sum_{k=1}^m IMF_k(t) + r_m(t) \quad (1)$$

式中: $S(t)$ 为一个非线性和非平稳的原始信号数据序列; $IMF_m(t)$ 为第 m 个IMF; $r_m(t)$ 为余项,是一个可以表征信号均值变化趋势的项。

EMD方法对于含有间歇信号、脉冲信号和噪声信号的信号进行分解时,其分解结果存在较为严重的模态混叠现象。EEMD方法将噪声辅助分析应用于经验模式分解中,以促进抗混分解,在一定程度上有效地抑制了模态混叠现象^[28]。EEMD算法如下。

1) 令 $S(t)$ 为一个非线性和非平稳的原始信号数据序列, $V^i(t)$ 代表第 i 次试验中添加的具有标准正态分布的白噪声序列。则第 i 次的信号数据序列可表示为:

$$S^i(t) = S(t) + V^i(t) \quad (2)$$

式中: $i = 1, 2, \dots, I$ 为试验次数。

2) 将每一次试验产生的信号数据序列 $S^i(t)$ 进行EMD分解,得到 $IMF_k^i(t)$, $k = 1, \dots, K$ 为分解的IMF个数。

3) 定义 $S(t)$ 的第 k 个固有模态分量为 IMF_k , 相应的对 IMF_k 进行平均得到 $\overline{IMF_k}$, 即:

$$\overline{IMF_k} = \frac{1}{I} \sum_{i=1}^I IMF_k^i(t) \quad (3)$$

1.2 CEEMDAN 理论

虽然 EEMD 可有效地缓解 EMD 的模态混叠现象, 但 EEMD 不能保持 EMD 的完整性。由于各种噪声信号的 EMD 分解结果不同, 会使 EEMD 产生重构误差, 影响重构原始信号的准确性。在 EEMD 的基础上, 提出 CEEMDAN 方法^[29]。

定义算子 $E_k(x)$ 为信号 x 通过 EMD 方法所产生的第 k 个模态分量, CEEMDAN 所产生的第 k 个模态分量记为 $\widehat{IMF}_k$ 。CEEMDAN 算法如下。

1) 令 $S(t)$ 为一个非线性和非平稳的原始信号数据序列, $V^i(t)$ 表示第 i 次试验中添加的具有标准正态分布的白噪声序列, ε_k 表示第 k 个模态分量的信噪比系数, 对信号 $S(t) + \varepsilon_0 V^i(t)$ 进行 I 次试验, 并通过 EMD 分解得到第 1 个模态分量如下:

$$\widehat{IMF}_1(t) = \frac{1}{I} \sum_{i=1}^I IMF_1^i(t) = \overline{IMF}_1(t) \quad (4)$$

2) 在此阶段, 即 $k=1$ 时, 计算第 1 个余量信号:

$$r_1(t) = S(t) - \widehat{IMF}_1(t) \quad (5)$$

3) 进行 i 次试验, 每次试验中对信号 $r_1(t) + \varepsilon_1 E_1(V^i(t))$ 进行 EMD 分解, 直至得到每次第 1 个模态分量为止。此时, 第 2 个模态分量计算如下。

$$\widehat{IMF}_2(t) = \frac{1}{I} \sum_{i=1}^I E_1(r_1(t) + \varepsilon_1 E_1(V^i(t))) \quad (6)$$

4) 对之后的每一阶段, 即 $k=2, \dots, K$, 同步步骤 3) 一致, 计算第 k 个余量信号, 第 $k+1$ 个模态分量计算如下:

$$r_k(t) = r_{k-1}(t) - \widehat{IMF}_k(t) \quad (7)$$

$$\widehat{IMF}_{k+1}(t) = \frac{1}{I} \sum_{i=1}^I E_1(r_k(t) + \varepsilon_k E_k(V^i(t))) \quad (8)$$

5) 继续执行步骤 4), 直至所得余量信号的极值点个数最多不超过 2 个时停止分解, 则此时得最终余量信号, 即余项信号为:

$$R(t) = S(t) - \sum_{k=1}^K \widehat{IMF}_k \quad (9)$$

则原始信号数据序列 $S(t)$ 经 CEEMDAN 分解结果可表示为:

$$S(t) = \sum_{k=1}^K \widehat{IMF}_k + R(t) \quad (10)$$

1.3 奇异值分解理论

奇异值分解(singular value decomposition, SVD)理论的本质是正交变换, 是谱分析理论在任意矩阵的推广, 在信号处理方面主要体现为数据降维、压缩、弱信号特征提

取和分离等, 其实际意义为反映信号能量集中和信号包含情况, 奇异值越大, 说明其所对应的子矩阵所表达的独立成分在整个矩阵中越突出且包含信息越多。SVD 反映矩阵的固有特征, 得到的子矩阵相互正交, 所以具有良好的平稳性, 能最大限度地减少特征量的冗余^[30]。

在故障诊断领域中, SVD 的应用主要体现在对振动信号进行降噪及故障信号特征提取两方面。在故障信号特征提取方面, 通常通过 SVD 得到的奇异值作为故障特征提取指标, 结合分类器算法如 BP 神经网络来进行故障识别。

对任一信号 $\mathbf{X}$, 其离散时间序列为 $\mathbf{X} = [x_1, x_2, \dots, x_n]$, 对 $\mathbf{X}$ 序列以长度 n 进行分段, 并按式(11)构造特征矩阵 $\mathbf{A}$ ^[31]:

$$\mathbf{A} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_{n+1} & x_{n+2} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{(m-1)n+1} & x_{(m-1)n+2} & \cdots & x_{mn} \end{bmatrix} \quad (11)$$

式中: $N = mn$ 表示采样点数。

对矩阵 $\mathbf{A}$ 进行奇异值分解,

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{A}' \quad (12)$$

式中: $\mathbf{U}$ 和 $\mathbf{V}$ 分别为 $m \times n$ 和 $n \times n$ 阶矩阵, 且 $\mathbf{U} \mathbf{U}' = \mathbf{I}$, $\mathbf{V} \mathbf{V}' = \mathbf{I}$; $\mathbf{S}$ 是 $m \times n$ 对角矩阵, 对角元素为 $s_1, s_2, \dots, s_p$, $p = \min(m, n)$, $s_1 \geq s_2 \geq \dots \geq s_p$ 。 $s_1, s_2, \dots, s_p$ 即为矩阵 $\mathbf{A}$ 的奇异值, $\mathbf{U}$ 和 $\mathbf{V}$ 分别为左右奇异阵。

1.4 基于 CEEMDAN-SQI-SVD 的齿轮箱局部故障最优特征提取算法

针对齿轮的局部故障特征, 结合皮尔逊相关系数与峰度 2 个指标进行 IMF 组合选取, 并以重构信号的均方误差与原始信号的相关系数及重定义信噪比 3 个指标定义 SQI, 构建成 SQI 算法。SQI 值能够体现出该重构信号在噪声抑制和信号保真两方面平衡度的好坏, 通过比较各个重构信号的 SQI 值大小来评判重构信号的 IMF 选取质量高低, 从而实现最优有效 IMF 的选取, 并进一步应用 SVD 方法完成最优故障特征提取。

CEEMDAN-SQI-SVD 算法具体步骤如下。

1) CEEMDAN 分解 $IMF_k(t)$

将原始信号数据序列 $S(t)$ 导入 CEEMDAN 进行分解, 得到一系列 $IMF_k(t)$ 及余项 $R(t)$, 其中 $k=1, 2, \dots, K$ 为 IMF 分量序号, $t=1, 2, \dots, T$ 为时间序列。

2) 皮尔逊相关系数 R_k 与峰度 K_k 计算

皮尔逊相关系数可表征各个 IMF 与原始信号的相关性大小, 通常认为与原始信号相关系数较大的 IMF 为真实分量, 与原始信号相关系数较小的 IMF 为虚假噪声分量; 峰度指标具有检测淹没在强烈干扰背景中瞬态信号的能力。

当齿轮无故障时, 其振动信号可视为平稳信号, 此时

峰值值很小;当齿轮存在局部故障时,其振动信号会出现冲击响应的脉冲型信号,此时峰值值会较大。峰值值越大,表明信号含冲击成分越多、故障信息越多。为确定有效 IMF,剔除无关分量及虚假噪声分量对后续分析的影响,采用皮尔逊相关系数与峰度 2 个指标进行有效 IMF 筛选。

计算 $IMF_k(t)$ 与原始信号数据序列 $S(t)$ 的相关系数:

$$R_k = \frac{\sum_{t=1}^T (IMF_k(t) - \overline{IMF_k(t)}) (S(t) - \overline{S(t)})}{\sqrt{\sum_{t=1}^T (IMF_k(t) - \overline{IMF_k(t)})^2 (S(t) - \overline{S(t)})^2}} \quad (13)$$

计算 $IMF_k(t)$ 的峰度:

$$K_k = \frac{T \sum_{t=1}^T (IMF_k(t) - \overline{IMF_k(t)})^4}{\left(\sum_{t=1}^T (IMF_k(t) - \overline{IMF_k(t)})^2 \right)^2} \quad (14)$$

3) 确定有效 IMF 最大个数 L

根据原始信号数据序列 $S(t)$ 本身复杂程度及分解的

IMF 个数,确定有效 IMF 最大个数为 L ,通常 $1 \leq L \leq \frac{K}{2}$ 。

4) 有效 IMF 选取规则及重构信号 $S_{ij}^L(t)$

根据 L 值进行有效 IMF 的选取,并对所选取的 IMF 进行信号重构,得到所有的重构信号 $S_{ij}^L(t)$ 以及重构信号 $S_{ij}^L(t)$ 的 IMF 分量序号记录矩阵 SN_{ij}^L ,其中 l 表示此重构信号的 IMF 个数, z 表示当重构信号 IMF 个数为 l 时的重构信号个数, i 表示以相关系数大小为指标选取的前 i 个 IMF, j 表示以峰度大小为指标选取的前 j 个 IMF, i 与 j 所代表的 IMF 不重复,即 $l = i + j$ 。相关系数为优先选取指标。以 $L = 1, 2, 3$ 为例,IMF 的选取规则如图 1 所示。

由图 1 可知,当 $L = 1$ 时,返回重构信号 S_{11}^1 与分量序号记录矩阵 SN_{11}^1 ;当 $L = 2$ 时,返回重构信号 S_{20}^{21}, S_{11}^{22} 以及与其相对应的分量序号记录矩阵 $SN_{20}^{21}, SN_{11}^{22}$;当 $L = 3$ 时,返回重构信号 $S_{30}^{31}, S_{21}^{32}, S_{21}^{33}, S_{12}^{34}$ 以及与其相对应的分量序号记录矩阵 $SN_{30}^{31}, SN_{21}^{32}, SN_{21}^{33}, SN_{12}^{34}$ 。

计算每个重构信号 $S_{ij}^L(t)$ 与原始信号数据序列 $S(t)$ 的均方误差 $MSE_{S_{ij}^L}$,相关系数 $R_{S_{ij}^L}$ 及重定义信噪比 $MSNR_{S_{ij}^L}$ 。均方误差的计算如下。

$$MSE_{S_{ij}^L} = \frac{1}{T} \sum_{t=1}^T (S(t) - S_{ij}^L(t))^2 \quad (15)$$

相关系数的计算如下。

$$R_{S_{ij}^L} = \frac{\sum_{t=1}^T (S_{ij}^L(t) - \overline{S_{ij}^L(t)}) (S(t) - \overline{S(t)})}{\sqrt{\sum_{t=1}^T (S_{ij}^L(t) - \overline{S_{ij}^L(t)})^2 (S(t) - \overline{S(t)})^2}} \quad (16)$$

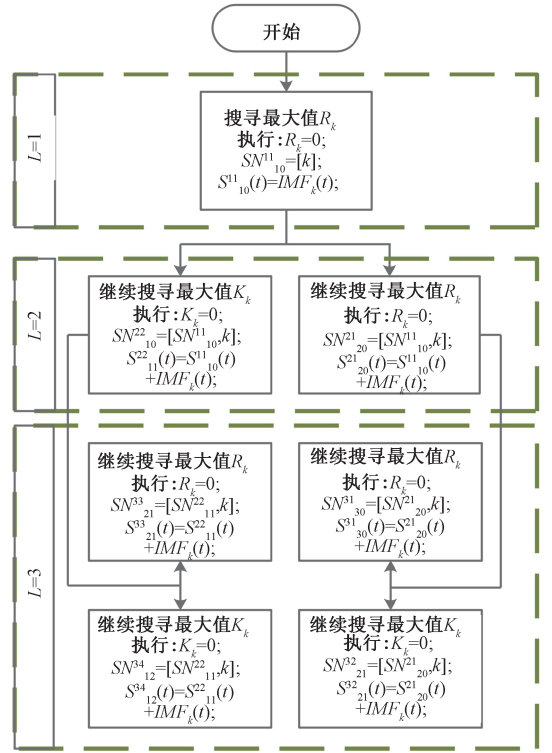


图 1 IMF 选取规则

Fig.1 IMF selection rule

信噪比的计算如下。

$$SNR = 10 \times \lg \left(\frac{P_s}{P_n} \right) \quad (17)$$

式中: P_s 与 P_n 分别代表信号和噪声的有效功率。

因在实测信号中,信号和噪声是混合在一起的,无法得到各自的有效功率,故对信噪比重定义如下。

$$MSNR_{S_{ij}^L} = 10 \times \lg \left(\frac{\sum_{t=1}^T (S(t) - S_{ij}^L(t))^2}{\sum_{t=1}^T (S(t))^2} \right) \quad (18)$$

均方误差是一个总体精度的衡量指标,其值越小,表明重构信号与原始信号的逼近程度越好,降噪效果越好;重构信号与原始信号相关系数越大,表明重构信号所包含的故障信息越多;重定义信噪比越大,表明重构信号有用信号能量与噪声信号能量的比值越大。故本文采用均方误差、相关系数及重定义信噪比 3 个指标对重构信号的质量进行评价。

5) 定义信号质量指数 $SQI_{S_{ij}^L}$

为避免赋权等人因素干扰,对步骤 4) 的均方误差,相关系数及重定义信噪比计算所得结果分别进行离差标准化,即无量纲化,使各个元素均处于 $[0, 1]$ 区间,对序列 $x_1, x_2, \dots, x_n$ 进行变换,离差标准化公式如下。

$$y_i = \frac{x_i - \min_{1 \leq i \leq n} \{x_j\}}{\max_{1 \leq j \leq n} \{x_j\} - \min_{1 \leq i \leq n} \{x_j\}}$$

(19)

通过离差标准化得到各个重构信号 $S_{ij}^L(t)$ 的均方误差、相关系数及重定义信噪比分别为 $\widehat{MSE}_{S_{ij}^L}$, $\widehat{R}_{S_{ij}^L}$, $\widehat{MSNR}_{S_{ij}^L}$, 则定义信号质量指数 $SQI_{S_{ij}^L}$ 如下。

$$SQI_{S_{ij}^L} = \widehat{MSNR}_{S_{ij}^L} + \widehat{R}_{S_{ij}^L} - \widehat{MSE}_{S_{ij}^L}$$

(20)

信号质量指数 $SQI_{S_{ij}^L}$ 值越大,说明重构信号 $S_{ij}^L(t)$ 的质量越高,则根据其所对应的 IMF 分量序号记录矩阵 SN_{ij}^L 便可确定最优有效 IMF 。

6) 最优故障特征提取

将记录矩阵 SN_{ij}^L 中最优有效 IMF 按照 CEEMDAN 分解顺序进行排列,利用 SVD 分别对每个 IMF 进行分解并获取每个 IMF 的奇异值,将所有有效 IMF 的奇异值组合构建最优故障特征向量以完成最优故障特征的提取。

至此完成 CEEMDAN-SQI-SVD 算法,得到每个样本的最优故障特征。

2 试验分析

2.1 试验设备及数据获取

搭建图 2 所示的齿轮箱试验台,齿轮箱由电机通过传动带驱动,电机转速由控制台调频控制。齿轮箱为两级传动,各齿轮参数如表 1 所示。

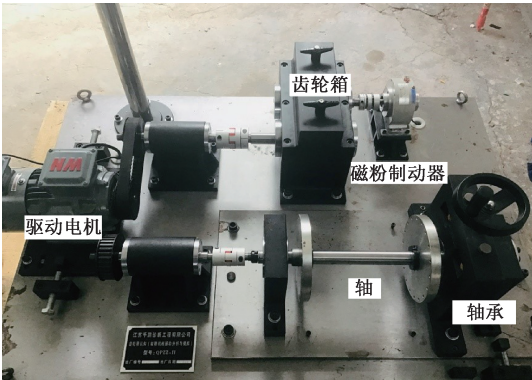


图 2 齿轮箱试验台
Fig.2 Gear box test-bed

表 1 齿轮参数
Table 1 Gear parameters

齿轮	齿数	模数	齿宽/mm	压力角/(°)	分度圆直径/mm	材料
正常	55	2.5	25	20	137.5	低碳钢
故障	75	2.5	25	20	187.5	低碳钢

选取齿轮局部裂纹故障作为失效模式进行研究。齿轮常见裂纹厚度约为 0.2~0.5 mm;常见的裂纹深度常

小于弦齿厚度的 1/2,即当裂纹深度大于弦齿厚度的 1/2 时,轮齿会迅速断裂。裂纹通常在最大应力点处产生,即在分度圆节线附近处出现,裂纹与分度圆节线的垂线通常呈 40°~50°。加工 4 种不同级别的齿轮裂纹作为测试齿轮,如图 3 所示,其相关参数如表 2 所示。

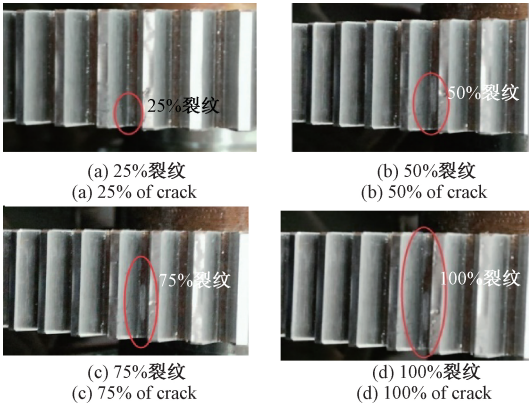


图 3 4 种裂纹故障齿轮
Fig.3 Four kinds of crack fault gear

表 2 裂纹齿轮故障参数
Table 2 Fault parameters of cracked gear

故障程度/%	裂纹宽度/mm	裂纹深度/mm	裂纹厚度/mm	裂纹角度/(°)
25	6	0.6	0.4	45
50	12	1.2	0.4	45
75	18	1.8	0.4	45
100	25	2.5	0.4	45

实验中,加速度传感器安装在齿轮箱输出轴轴承端盖的径向方向上。电机转速设置为 1 200 r/min,设置采样频率为 5.12 kHz。为保证单个样本的数据包含数个故障周期,设置单个样本采样时间为 0.5 s,即单个样本所含数据点数为 2 060 个。

2.2 算法应用与分析

基于 CEEMDAN-SQI-SVD 的齿轮箱局部故障最优特征提取算法及 BP 神经网络训练与识别的处理过程如图 4 所示。

首先获取 4 种故障级别的齿轮振动信号,每种故障级别的齿轮振动信号收集 100 个样本,50 个样本用于训练 BP 神经网络,另 50 个样本用于验证所提方法的有效性,共计 400 个样本。对每种故障级别的齿轮各取其中一个样本进行原始振动信号波形图绘制,如图 5 所示。由于外在环境等噪声影响以及故障频率调制现象,从 4 种故障级别的原始波形图无法找出规律性变化或特征,很难判断其故障级别。

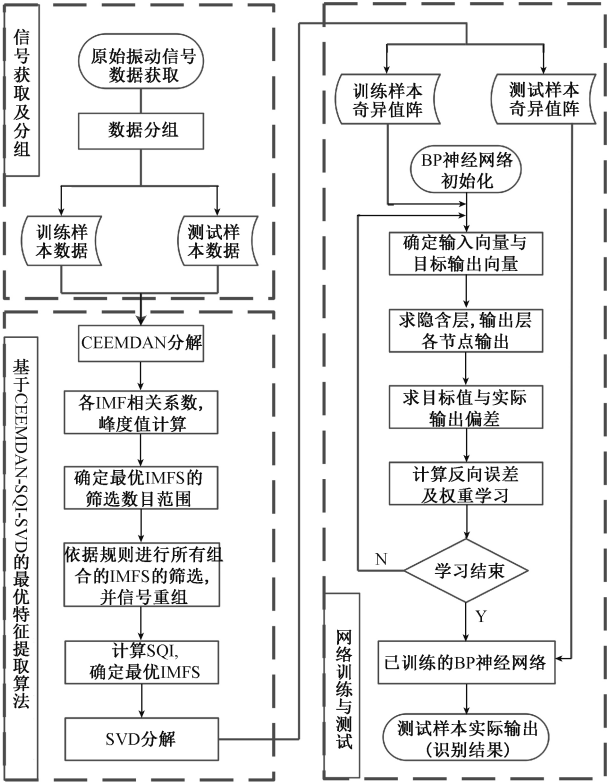


图4 最优特征提取算法及BP神经网络训练与识别流程

Fig.4 Optimal feature extraction algorithm and BP neural network training and recognition processing

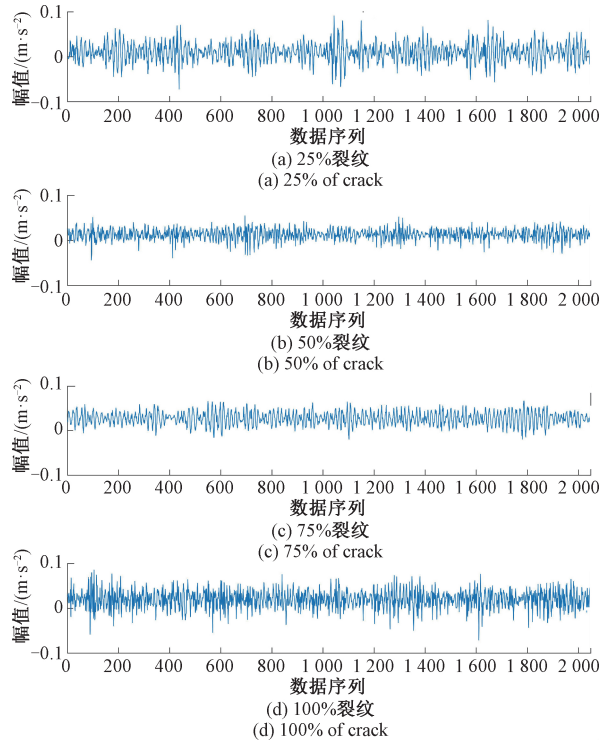


图5 故障齿轮原始信号波形

Fig.5 Original signal waveform of the fault gear

利用 CEEMDAN 对所有样本进行信号分解。将 CEEMDAN 添加的噪声标准偏差设置为原始振动信号的 0.2 倍,每次筛选最大迭代次数设置为 100 次,平均次数设置为 30 次时,CEEMDAN 分解效率及分解结果均较为出色。在此选择裂纹故障级别为 25% 的第 1 个样本进行方法说明。此样本经过 CEEMDAN 分解后得到图 6 所示的 12 个 IMF 及 1 个余项。

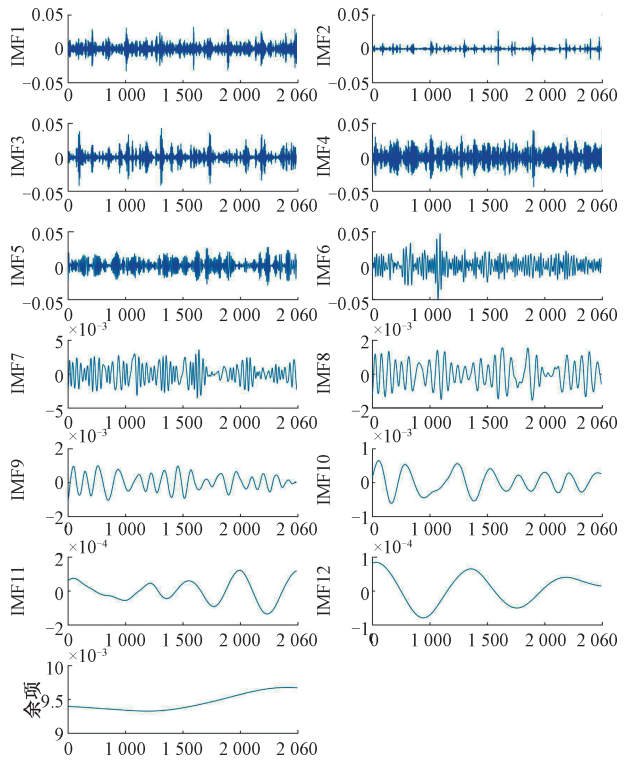


图6 CEEMDAN 分解结果

Fig.6 The decomposed results by CEEMDAN

由图 6 可知,经过 CEEMDAN 分解后 12 个 IMF 依次由高频至低频排列,在此可根据预测故障及齿轮箱参数大致确定故障频率所处范围。将 IMF 从时域波形图转为频域波形图观察每个 IMF 的频率带,仅选取处于故障频率范围的 IMF 作为有效 IMF 进行进一步的频域分析或时频分析。但此做法需提前对故障所处频率带范围进行计算,且需专业技术人员的经验对图形做出诊断,主观因素较大且效率低,故不做深究。

分别计算此 12 个 IMF 与原始振动信号的相关系数及峰值值,计算结果如图 7 所示。

由图 7 可知,相关系数最大的分量为 IMF4,峰值最大的分量为 IMF2,即不同指标所得的含有故障信息最多的 IMF 并不相同。若仅依据单一指标选择一个分量,则无法确定结果的精确度;若依据单一指标选择多个分量,一方面分量选择的个数目前没有一个统一的标准,另一方面,不同指标选取的多个分量可能仍然不同。如若选

取3个分量,依据相关系数选取的分别为IMF1、IMF3、IMF4;而依据峰度选取的分别为IMF1、IMF2、IMF3。两者并不相同,同样会无法确定结果的精确度。因此,本文在指标的选取上结合相关系数及对局部故障敏感的峰度2个指标进行有效IMF选取,以避免单一指标可能造成的片面性,且在选取的过程中不限定有效IMF的具体个数,而是给定一个有效IMF个数的范围,通过比较不同组合的SQI值大小来确定最优的有效IMF。

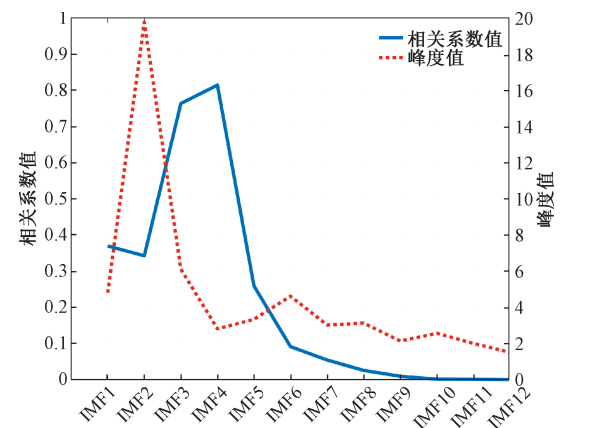


图7 各IMF相关系数与峰度值

Fig.7 IMF correlation coefficients and kurtosis values

根据此范例信号分解的IMF个数,确定最大有效IMF个数为7,将分解的IMF依据第2节的方法进行处理,共得到127个重构信号,即127种有效IMF组合。在此需指出,因为算法本身的选择规则只能确保每种组合选取的IMF顺序不同,无法判断其筛选结果是否相同,即在127种有效IMF组合中,会出现多组相同的组合筛选结果,但每个组合中的IMF被筛选的顺序不同的现象。因所求结果仅为有效IMF,与其被筛选顺序无关,故对重复现象不做考虑。去除重复组合后,SQI排在前3的重构信号信息如表3所示。

表3 SQI前三重构信号信息

Table 3 SQI information of first three reconstruction signal			
排序	重组信号	SQI	有效分量
1	S_{42}^{63}	0.975 8	IMF1、IMF2、IMF3、IMF4、IMF5、IMF6
2	S_{41}^{52}	0.973 3	IMF1、IMF2、IMF3、IMF4、IMF5
3	S_{33}^{65}	0.956 7	IMF1、IMF2、IMF3、IMF4、IMF6、IMF8

由表3可知,排序1与2的SQI值仅相差0.002 5,即此2个组合所包含的故障信息量非常接近;而排序1有效分量为6个,排序2的有效分量为5个,若考虑后续处理的效率,可选取有效分量较少的排序2组合。排序1所对应的分量恰好为前6个高频分量,在此仅为一个巧合。通过试验发现在其他样本中也多次出现排序1所对应的分量与

此范例排序2所对应的5个分量或排序3所对应的6个分量相同,甚至还有部分排序1的仅有4个分量。从SQI算法结果可知,经SQI算法所筛选的有效IMF中,均具有高频分量,高峰度值分量以及高相关系数分量,平衡了以高频,高相关系数,高峰度值等指标进行有效IMF选取的方法,兼顾了上述方法的优点,具有一定的代表性。

在此范例信号中,以此排序1结果为准,排序1的SQI值为127种组合最大值,其代表的有效分量分别为IMF1、IMF2、IMF3、IMF4、IMF5、IMF6。选择具有代表性的前3个有效分量做频谱分析,得到傅里叶频谱图如图8所示。该样本采样频率 $f_s=5\ 120\text{ Hz}$,故障齿轮所在轴转频率 $f_r=18.3\text{ Hz}$,故障齿轮齿数 $Z=75$,计算得故障齿轮啮合频率 $f_z=1\ 375\text{ Hz}$ 。

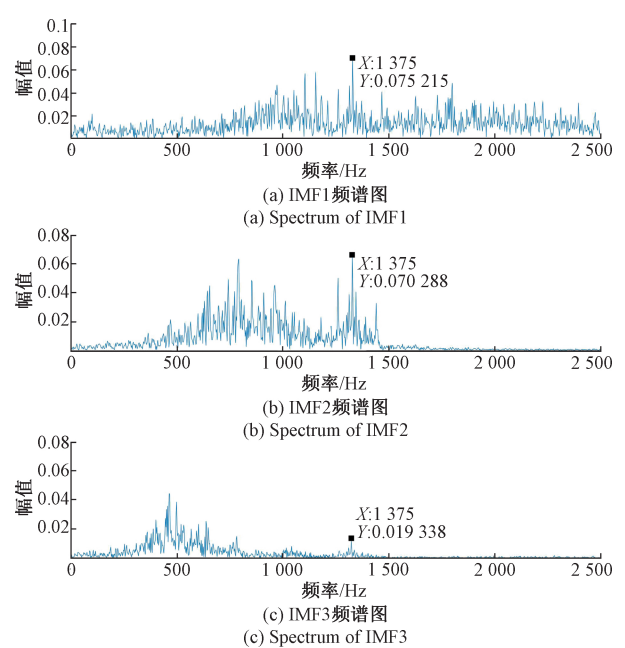


图8 IMF1-IMF3 频谱图

Fig.8 Spectrum of IMF1-IMF3

由图8可知,前3个有效分量谱峰在故障齿轮啮合频率 $f_z=1\ 375\text{ Hz}$ 处均较为明显,并且在啮合频率处的谱线两边存在有非常明显的边频带,边频带的幅值变化清晰可见,符合该样本实际故障特征,体现出SQI算法所筛选的有效IMF在有效去噪的基础上保留了较为完整的故障信息。

以排序1结果所示6个分量作为最优有效IMF进行奇异值分解,得到6个分量的奇异值如表4所示。

表4 IMF 奇异值

Table 4 The singular value

IMF 序号	1	2	3	4	5	6
奇异值	0.334 5	0.186 5	0.163 6	0.145 7	0.063 7	0.050 6

对剩余 399 个样本均做以上处理,得到各个样本的有效 IMF 奇异值。为验证所提出方法的有效性,利用 BP 神经网络进行不同故障级别的识别。

BP 神经网络输入层节点数 m 设置为每个样本的有效 IMF 数目,根据 SVD 分解结果显示,400 个样本中有 355 个样本有效 IMF 数目为 6,39 个样本有效 IMF 数目为 5,6 个样本有效 IMF 数目为 4;因神经网络输入层节点数需一致,有效 IMF 数目为 6 的样本数占据总样本数的 88.75%,故确定节点数 m 为 6。对于其余 45 个最优 IMF 个数非 6 个的样本取其次优的 6 个 IMF 作为有效 IMF 以统一输入层节点数。输出节点数 n 为故障级别数,即 $n=4$ 。最佳隐含层节点数根据多次试验确定为 8 个时最佳。综上,确定 BP 神经网络结构为 $6\times 8\times 4$,且输出结果为 1×4 的矩阵,如表 5 所示。选择对数 S 型激励函数作为隐含层激励函数,选择线性激励函数作为输出层激励函数。学习速率为 0.01,训练次数为 100 次,其余参数均为默认参数。

表 5 BP 神经网络输出矩阵结构

Table 5 BP neural network output matrix structure	
故障级别	输出矩阵
25% 裂纹	[1 0 0 0]
50% 裂纹	[0 1 0 0]
75% 裂纹	[0 0 1 0]
100% 裂纹	[0 0 0 1]

依据上述参数设置对 BP 神经网络进行初始化以及对 200 个训练样本进行训练,并对 200 个测试样本进行测试,识别结果如图 9 所示。

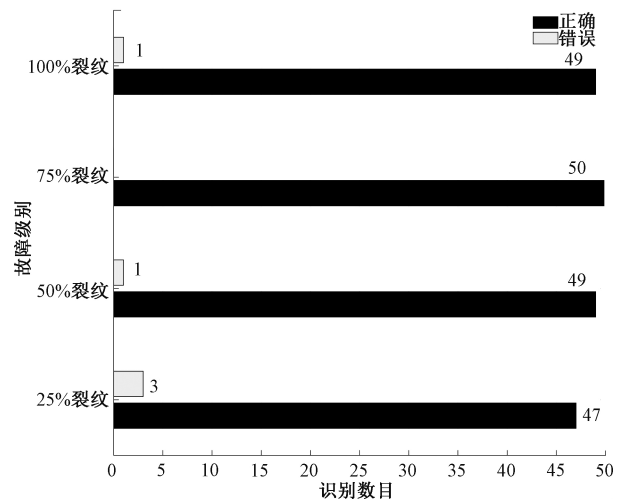


图 9 测试样本识别结果

Fig.9 Test sample identification results

由图 9 可知,25% 裂纹的故障级别识别正确 47 个,50% 与 100% 裂纹的故障级别均识别正确 49 个,75% 裂纹故障级别的 50 个测试样本均识别正确。由此结果可知,经过 CEEMDAN-SQI-SVD 算法所提取的故障特征向量结合 BP 神经网络能够较高精度识别齿局部故障不同级别的齿轮。

为进一步验证本文所提算法的有效性,将本文方法与常见的几种方法进行比较。在故障程度识别上均采用 BP 神经网络进行识别,且网络参数设置均相同。因每种故障程度的训练样本量与测试样本量均仅为 50 个,为避免巧合及误差,每种方法单次试验均在每种故障级别的 100 个样本中随机抽取 50 个作为训练样本,剩余 50 个作为测试样本,共做 10 次试验,并取 10 次试验结果的平均值作为最终识别结果,最终识别结果如表 6 所示。

表 6 4 种方法识别结果

Table 6 Identification results of four methods

方法	有效 IMF 个数	故障级别(裂纹大小)/%				平均识别率/%
		25	50	75	100	
EMD-SVD ^[12]	5	82.2	94.2	92.6	77.2	86.55
EMD-Energy Entropy ^[32]	9	67.2	36.8	99.2	60.0	65.80
VMD-ApEn ^[33]	4	61.0	95.8	74.4	96.2	81.85
CEEMDAN-SQI-SVD	6	92.4	98.0	99.8	98.2	97.1

由表 6 可知,基于 EMD-SVD 算法的有效 IMF 个数为 5,其对不同故障级别齿轮的识别率高低不均,原因可能为 EMD 分解的 IMF 具有较为严重的模态混叠现象;而其平均识别率为 86.55%,识别效果在 4 种方法中较为出色,一方面是因为基于 EMD 系列的方法比较适合处理故障诊断问题,另一方面在于 SVD 能够较有效地提取故障特征。基于 EMD-Energy Entropy 算法的有效 IMF 个数达到 9 个,但其平均识别率最低,可能的原因为有效 IMF 个数较多,所提取的特征包含较多的噪声等干扰信息。基于 VMD-ApEn 算法的有效 IMF 个数为 4 个,其在故障级别为 50% 与 100% 时识别率较高,在为 25% 与 75% 时识别率较低,可能的原因是通过 VMD 分解得到的有效 IMF 个数较少导致所提取的特征包含的故障信息较少,另一方面 ApEn 不能有效地区分不同级别的故障,从而造成不同故障识别率差别较大。本文所提方法有效 IMF 个数为 6 个,有效 IMF 个数较为均衡,在有效去噪的前提下保留了较多的故障信息,SVD 又能够较好的提取故障

特征,故本文方法在 4 种故障级别的识别率及平均识别率均最高,体现了 CEEMDAN-SQI-SVD 算法的有效性。由表 6 还可知,在故障级别为 25% 裂纹时,4 种方法的识别率相对来说均较低,体现出故障初期阶段的诊断较为困难这一普遍性问题,但本文所提方法仍然具有 92.4% 的识别率,说明本文所提方法在诊断初期故障具有较好的效果。

3 结 论

本文提出了一种基于 CEEMDAN-SQI-SVD 的齿轮箱局部故障最优特征提取算法,经试验获取故障齿轮原始振动信号数据,通过 CEEMDAN 对原始振动信号数据进行分解得到一系列 IMF,利用本文定义的 SQI 算法确定最优有效 IMF,并应用 SVD 对最优有效 IMF 进行分解,将有效 IMF 的奇异值构建奇异值阵作为最优故障特征向量输入至 BP 神经网络进行训练及测试,并与其他常规特征提取方法进行比较,结果表明 CEEMDAN-SQI-SVD 对裂纹分别为 25%、50%、75% 和 100% 的测试故障齿轮的识别率分别为 92.4%、98.0%、99.8% 和 98.2%,平均识别率达到 97.1%,均优于现有的几种常规方法,体现出本方法的有效性。同时也需指出,本方法适应的故障诊断范围为齿轮箱旋转器件的某一局部故障,具有较强的针对性,但也能够对不同类故障模式的识别,以及对于同类均匀性不同级别的故障识别的研究提供一定的参考,例如齿轮不同故障程度的均匀性磨损识别,可尝试通过更改或增减 SQI 算法中的筛选指标,重新定义 SQI 来实现。

参考文献

- [1] PENG Z, CHU F, HE Y. Vibration signal analysis and feature extraction based on reassigned wavelet scalogram[J]. Journal of Sound and Vibration, 2002, 253(5): 1087-1100.
- [2] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition [J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [3] MAHGOUN H. Detection of gear faults in variable rotating speed using variational mode decomposition (VMD) [J]. Mechanics & Industry, 2016, 17(2): 207-213.
- [4] 王建国, 陈帅, 张超. 基于 VMD 与多特征融合的齿轮故障诊断方法[J]. 机械传动, 2017, 41(3): 160-165.
- [5] WANG J G, CHEN SH, ZHANG CH. Fault diagnosis method of gear based on VMD and multi-feature fusion[J]. Journal of Mechanical Transmission, 2017, 41(3): 160-165.
- [5] 程军圣, 李梦君, 欧龙辉, 等. FA-PMA-VMD 方法及其在齿根裂纹故障诊断中的应用[J]. 振动与冲击, 2018, 37(15): 27-32.
- [6] CHENG J SH, LI M J, OU L H, et al. FA-PMA-VMD method and its application in gear tooth root crack fault diagnosis[J]. Journal of Vibration and Shock, 2018, 37(15): 27-32.
- [6] GILLES J. Empirical wavelet transform [J]. IEEE Transactions on Signal Processing, 2013, 61(16): 3999-4010.
- [7] 辛玉, 李舜酩, 王金瑞, 等. 基于迭代经验小波变换的齿轮故障诊断方法[J]. 仪器仪表学报, 2018, 39(11): 79-86.
- [8] XIN Y, LI SH M, WANG J R, et al. Gear fault diagnosis method based on iterative empirical wavelet transform[J]. Chinese Journal of Scientific Instrument, 2018, 39(11): 79-86.
- [8] MERAINANI B, BENAZZOUZ D, RAHMOUNE C. Early detection of tooth crack damage in gearbox using empirical wavelet transform combined by Hilbert transform[J]. Journal of Vibration and Control, 2017, 23(10): 1623-1634.
- [9] HUANG N E, SHEN Z, LONG S, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [C]. Proceedings of the Royal Society A, 1998, 454: 903-995.
- [10] CHENG G, CHENG Y L, SHEN L H, et al. Gear fault identification based on Hilbert-Huang transform and SOM neural network [J]. Measurement: Journal of the International Measurement Confederation, 2013, 46(3): 1137-1146.
- [11] JAOUHER B A, NADER F, LOTFI N, et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals[J]. Applied Acoustics, 2015, 89:

- 16-27.
- [12] ASR M Y, ETTEFAGH M M, HASSANNEJAD R, et al. Diagnosis of combined faults in rotary machinery by non-naive Bayesian approach [J]. *Mechanical Systems and Signal Processing*, 2017, 85: 56-70.
- [13] GAI J B, HU Y F. Research on fault diagnosis based on singular value decomposition and fuzzy neural network [J]. *Shock and Vibration*, 2018, doi: 10.1155/2018/8218657.
- [14] 张刚, 李红威. 小波-EMD 和随机共振级联微弱信号检测 [J]. *电子测量与仪器学报*, 2018, 32(1): 57-65.
ZHANG G, LI H W. Wavelet-EMD and stochastic resonance cascade weak signal detection [J]. *Journal of Electronic Measurement and Instrumentation*, 2018, 32(1): 57-65.
- [15] 田晶, 王英杰, 王志, 等. 基于 EEMD 与空域相关降噪的滚动轴承故障诊断方法 [J]. *仪器仪表学报*, 2018, 39(7): 144-151.
TIAN J, WANG Y J, WANG ZH, et al. Fault diagnosis for rolling bearing based on EEMD and spatial correlation denoising [J]. *Chinese Journal of Scientific Instrument*, 2018, 39(7): 144-151.
- [16] FANG L, SUN H C. Study on EEMD-Based KICA and its application in fault-feature extraction of rotating machinery [J]. *Applied Sciences-Basel*, 2018, 8(9): 1441.
- [17] CHEN H G, CHEN P, CHEN W H, et al. Wind turbine gearbox fault diagnosis based on improved EEMD and Hilbert square demodulation [J]. *Applied Sciences-Basel*, 2017, 7(2): 128.
- [18] KUAI M, CHENG G, PANG Y, et al. Research of planetary gear fault diagnosis based on permutation Entropy of CEEMDAN and ANFIS [J]. *Sensors*, 2018, 18(3): 782.
- [19] VANRAJ S S D, PABLA B S. Non-contact incipient fault diagnosis method of fixed-axis gearbox based on CEEMDAN [J]. *Royal Society Open Science*, 2017, 4(8): 170616.
- [20] LYU Y, YUAN R, WANG T, et al. Health degradation monitoring and early fault diagnosis of a rolling bearing based on CEEMDAN and improved MMSE [J]. *Materials*, 2018, 11(6): 1009.
- [21] BAI L L, HAN Z N. A hybrid de-noising algorithm for the gear transmission system based on CEEMDAN-PE-TFPF [J]. *Entropy*, 2018, 20(5): 361.
- [22] KEDADOUCHE M, THOMAS M, TAHAN A. A hybrid method combining Teager Kaiser energy operator, empirical mode decomposition and minimum entropy deconvolution for monitoring gears damages [J]. *Mechanics & Industry*, 2015, 16(6): 610.
- [23] LYU Y, YUAN R, SONG G. Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing [J]. *Mechanical Systems and Signal Processing*, 2016(81): 219-234.
- [24] LIANG J Y, ZHONG J H, YANG Z X. Correlated EEMD and effective feature extraction for both periodic and irregular faults diagnosis in rotating machinery [J]. *Energies*, 2017, 10, 10: 1652.
- [25] SHI X J, LI W T, GAO Q K. Research on fault classification of wind turbine based on IMF Kurtosis and PSO-SOM-LVQ [C]. *Proceedings of the IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference*, 2017, 191-196.
- [26] AMARNATH M, KRISHNA I, R PRAVEEN. Local fault detection in helical gears via vibration and acoustic signals using EMD based statistical parameter analysis [J]. *Measurement: Journal of the International Measurement Confederation*, 2014, 58: 154-164.
- [27] ZHAO D Z, LI J Y, CHENG W D. Feature extraction of faulty rolling element bearing under variable rotational speed and gear interferences conditions [J]. *Shock and Vibration*, 2015, doi: 10.1155/2015/425989.
- [28] 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.
- [29] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise [C]. *Proceedings of the IEEE International Conference on Speech and Signal Processing*, 2011: 4144-4147.
- [30] 张安安, 黄晋英, 卫洁洁, 等. 基于 EMD-SVD 与 PNN 的行星齿轮箱故障诊断研究 [J]. *机械传动*, 2018,

42(12): 160-165.

ZHANG AN AN, HUANG J Y, WEI J J, et al. Research of fault diagnosis of planetary gearbox based on EMD-SVD and PNN[J]. Journal of Mechanical Transmission, 2018, 42(12): 160-165.

- [31] 曾鸣, 杨宇, 郑近德, 等. μ -SVD 降噪算法及其在齿轮故障诊断中的应用[J]. 机械工程学报, 2015, 51(3): 95-103.

ZENG M, YANG Y, ZHENG J D, et al. μ -SVD based denoising method and its application to gear fault diagnosis[J]. Journal of Mechanical Engineering, 2015, 51(3): 95-103.

- [32] GAO S X, CHEN X H, DING Y. Fault recognition of gear pump based on EMD neural network [C]. Proceedings of the 2nd International Conference on Measurement, Instrumentation and Automation, 2013: 1635-1639.

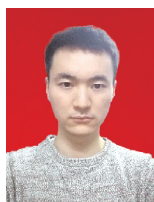
- [33] LI K, SU L, WU J J, et al. A rolling bearing fault diagnosis method based on variational mode decomposition and an improved kernel extreme learning machine[J]. Applied Sciences-Basel, 2017(7): 1004.

作者简介



E-mail: guyingkui@163.com

Gu Yingkui (Corresponding author) received his B. Sc. and M. Sc. degrees both from Jiangxi University of Science and Technology in 1998 and 2002, and Ph. D. degree from Dalian University of Technology in 2005, respectively. He is currently a professor at Jiangxi University of Science and Technology. His main research interests include reliability and system safety.



古莹奎 (通信作者), 分别在 1998 年和 2002 年于江西理工大学获得学士学位和硕士学位, 2005 年于大连理工大学获得博士学位, 现为江西理工大学教授, 博士生导师, 主要研究方向为可靠性与系统安全。

曾磊, 2016 年毕业于四川农业大学获得学士学位, 现为江西理工大学硕士研究生, 主要研究方向为可靠性与故障诊断。

E-mail: 13698079602@163.com

Zeng Lei received his B. Sc. degree from Sichuan Agricultural University in 2016. He is currently a M. Sc. candidate at Jiangxi University of Science and Technology. His main research interests include reliability engineering and fault diagnosis.