

改进 VMD 和非凸重叠组收缩降噪的行星齿轮箱 早期故障特征提取

王朝阁¹, 李宏坤², 曹顺心², 周强², 刘艾强², 任学平³

(1. 上海海事大学物流工程学院, 上海 201306; 2. 大连理工大学机械工程学院, 辽宁 大连 116024;
3. 内蒙古科技大学机械工程学院, 内蒙古 包头 014010)

摘要: 针对强背景噪声下行星齿轮箱早期微弱故障特征难以被识别且变分模态分解算法中模态个数 K 和惩罚因子 α 需要依赖人为经验反复尝试而不能自适应确定的问题, 提出了改进变分模态分解 (Modified Variational Mode Decomposition, MVMD) 方法, 通过故障信号尺度空间谱的自适应分割来确定所需模态个数 K , 同时建立峭度最大判定准则自动选取最佳惩罚因子 α 。在此基础上, 将 MVMD 与自适应非凸重叠组收缩降噪算法 (Adaptive Non-convex Overlap Group Shrinkage, ANOGS) 相融合, 提出了基于 MVMD-ANOGS 的行星齿轮箱早期故障诊断方法。对故障振动信号进行 MVMD 最佳分解, 获取多个模态分量; 利用 ANOGS 算法对峭度最大的敏感模态进行稀疏降噪, 从而突出信号中的故障冲击特征; 对降噪模态进行包络解调处理, 以提取明显的故障特征频率来判别故障。通过仿真信号和工程实验数据分析表明, 相比传统 VMD 方法、EEMD 方法和快速谱峭度方法, 该方法能成功地提取微弱故障冲击特征且更加清晰, 提高了行星齿轮箱早期故障的表征能力与诊断精度。

关键词: 故障诊断; 行星齿轮箱; 变分模态分解; 重叠组收缩算法; 特征提取

中图分类号: TH165⁺.3; TH132.425 **文献标志码:** A **文章编号:** 1004-4523(2021)06-1293-12

DOI: 10.16385/j.cnki.issn.1004-4523.2021.06.022

引 言

行星齿轮箱具有传动平稳、减速比大、效率高和承载能力强等优点, 已被广泛应用于风力发电机、直升机、高速列车和舰船等大型复杂机械设备中^[1]。然而, 行星齿轮箱通常在高负荷、强冲击和高污染等恶劣工况下运行, 同时受到温度、润滑等物理化学条件的影响, 极易导致齿轮出现点蚀、剥落和疲劳裂纹等局部故障^[2]。齿轮箱一旦发生故障, 将加速设备性能恶化, 轻则引起整个动力传输系统失效, 重则导致设备毁坏。因此, 若能在齿轮箱失效早期实现故障溯源并及时消除隐患, 对保证设备稳定安全运行, 避免灾难性事故发生意义重大。

由于恶劣的环境、复杂的结构和独特的运行方式, 使得从行星齿轮箱壳体上获取的故障振动信号表现为强背景噪声、非平稳、非线性和耦合调制等特点。近年来, 针对非平稳非线性信号的自适应信号分析方法得到了广泛的研究, 它为周期性脉冲提取和齿轮箱故障诊断提供了有力工具。例如, 经验模态分解 (Empirical Mode Decomposition, EMD) 能够根据信号在时间尺度上的局部特性自适应地分解为多个内禀模

态分量^[3]。然而, EMD 具有过包络、欠包络、端点效应和模态混叠等不足, 限制了其在故障诊断领域中的应用。随后, Smith^[4]提出了局部均值分解 (Local Mean Decomposition, LMD) 方法, 它克服了 EMD 中包络、欠包络的问题, 具有端点效应不明显、迭代次数少等优点。但同 EMD 方法一样, LMD 仍然存在模态混叠的问题。为了抑制模态混叠现象, Yang 等^[5]在集合经验模式分解 (Ensemble Empirical Mode Decomposition, EEMD)^[6]基础上提出了一种噪声辅助分析的总体局部平均分解 (Ensemble Local Mean Decomposition, ELMD) 方法。然而, ELMD 中添加噪声后在信号重构时存在残余噪声污染和运算量大等问题。此外, 不同于 EMD 及其扩展算法, Gilles 提出的经验小波变换 (Empirical Wavelet Transform, EWT)^[7]通过对信号的 Fourier 频谱进行划分来确定不同频带下的模态分量, 但该方法鲁棒性差, 在 Fourier 频谱区间划分时易受噪声干扰而导致频谱划分过于密集, 使结果出现过分解现象。

为克服上述方法缺陷, Dragomiretskiy 等^[8]提出了变分模态分解 (Variational Mode Decomposition, VMD), 作为一种新颖的自适应信号分析方法, VMD 把信号分离过程转化到变分框架中, 并采用

非递归的方式获取若干个具有稀疏特性的模态分量。VMD算法自提出以来,因其良好的噪声鲁棒性、较低的计算复杂度和完备的数学理论基础而受到学者们的青睐。Wang等^[9]采用VMD对转子系统碰摩故障分析,并研究了其等效滤波特性,结果证明VMD在提取瞬态冲击方面比EWT,EMD和EEMD更加有效。然而,VMD的优越性依赖于模态数目 K 和惩罚因子 α 的准确选取,这极大地制约了其应用。目前,文献中主要是采用如下三种途径来获得适当的参数:第一种根据先验知识或中心频率观察法^[10-11],显然,在复杂的情况下这种方法缺乏适应性且降低了解析的精度;第二种采用优化算法,如粒子群算法、蚱蜢算法和灰狼算法等^[12-14],虽然这些优化算法可搜寻到适当的参数,但需要大量的迭代试验,计算效率较低;第三种通过若干评价指标来选取最佳参数。例如,Zhang等^[15]通过能量和相关系数搜索 K 值,但这种方法选取的 K 值并不适用于所有的信号,同时未考虑惩罚因子 α 。因此,仍然需要对VMD算法做进一步优化,通过自适应判据来指导参数的合理选择。此外,VMD得到的模态分量中仍含有大量的噪声,特别是周期脉冲对噪声很敏感,影响故障特征的有效识别。

针对上述问题,本文提出一种融合改进VMD和自适应非凸重叠组稀疏收缩降噪的早期故障特征提取方法。首先,MVMD算法中通过故障信号尺度空间谱的自适应分割来确定分解所需模态个数 K ,同时建立峭度最大值判定准则自动选取最佳的惩罚因子 α 。然后,利用ANOGS算法对峭度最大的敏感模态进行自适应稀疏降噪,从而增强微弱的周期性故障冲击特征。数值仿真和工程实验数据分析表明所提方法能够有效、准确地提取早期微弱故障特征,为行星齿轮箱故障预测与诊断提供了一种新技术。

1 改进的变分模态分解

1.1 VMD算法

VMD算法将信号的分离转换为变分模型最优化的求解问题,在分解过程不断更新各个模态的频率中心与带宽,从而完成信号频域自适应分割,最终获取到若干个具有稀疏特性的模态分量^[8-12]。因此,VMD实质是构造变分问题和求解变分问题。

首先,构建约束变分问题为

$$\begin{cases} \min_{\{u_k, \omega_k\}} \left\{ \sum_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{S.t.} \sum_k u_k(t) = x(t) \end{cases} \quad (1)$$

式中 $\{u_k\} = \{u_1, u_2, u_3, \dots, u_k\}$ 为分解的 K 个模态分量; ω_k 为对应模态 u_k 的中心频率; $\partial_t(\cdot)$ 表示对时间 t 的偏导数; $\delta(t)$ 为脉冲函数; $x(t)$ 代表原始信号。

其次,为获取如上约束变分问题最优解,引入Lagrange乘法算子 $\lambda(t)$ 与二次惩罚因子 α ,将其转变为无约束的变分问题^[13-14]。扩展的Lagrange表达式如下

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \underbrace{\alpha \sum_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2}_{\text{移频信号的L2范数梯度平方}} + \\ & \underbrace{\left\| x(t) - \sum_k u_k(t) \right\|_2^2}_{\text{二次惩罚项}} + \\ & \underbrace{\left\langle \lambda(t), x(t) - \sum_k u_k(t) \right\rangle}_{\text{拉格朗日乘数项}} \end{aligned} \quad (2)$$

利用乘法算子交替方向法(Alternate Direction Method of Multipliers, ADMM)求解式(2)的最优解^[15]。首先,预设模态个数;同时,初始化模态函数频域表达式,及对应的中心频率和Lagrange乘法算子;然后,各模态分量及其中心频率分别按照以下两式不断更新

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{x}(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \quad (3)$$

式中 $\hat{u}_k^{n+1}(\omega)$, $\hat{x}(\omega)$, $\hat{\lambda}(\omega)$ 分别表示 $u_k^{n+1}(t)$, $x(t)$, $\lambda(t)$ 的傅里叶变换。

$$\omega_k^{n+1}(\omega) = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (4)$$

模态和中心频率每次迭代更新后, Lagrangian乘数根据下式也被更新

$$\hat{\lambda}^{n+1} \leftarrow \hat{\lambda}^n + \tau \left(\hat{x} - \sum_k \hat{u}_k^{n+1} \right) \quad (5)$$

式中 τ 为噪声容忍度,为了获得较好的降噪效果,设 $\tau = 0$ 。继续上述迭代,直到满足收敛条件,即

$$\sum_{k=1}^K \|u_k^{n+1} - u_k^n\|_2^2 / \|u_k^n\|_2^2 < \epsilon \quad (6)$$

式中 ϵ 为收敛性判定准则,本文取 $\epsilon = 1 \times 10^{-7}$ 。

1.2 参数的自适应选取

由上述理论可知,在VMD算法中模态个数 K 、惩罚因子 α 、噪声容忍度 τ 和收敛准则 ϵ 这四个主要参数需要提前确定。与前两个参数相比, τ 和 ϵ 对分解结果影响较小,因此通常采用算法中默认值($\tau = 0$ 和 $\epsilon = 1 \times 10^{-7}$)。然而,模态个数 K 决定着分解的准确性和效率,惩罚因子 α 影响着模态分量的频带宽度与噪声抑

制性能。因此,寻找与待分析信号相匹配的最佳参数组合 $[K, \alpha]$ 对于旋转机械故障诊断的成功至关重要。

1.2.1 分解模态参数 K 的确定

信号傅里叶谱的局部信息经常受到噪声严重影响,而无法找到有意义的模态。基于频谱的尺度空间表示,可以很好地抑制噪声干扰,能以一种简单快速的方式寻找到频谱中有意义的模态,具有无参数化和自适应性。对于待分析信号 $x(t)$,其离散傅里叶变换为

$$X(f) = \sum_{t=-\infty}^{+\infty} x(t)e^{-2j\pi ft} \quad (7)$$

信号傅里叶频谱的离散尺度空间可表示为^[16]

$$L(f, s) = g(f, s) * X(f) = \sum_{\tau=-P}^{+P} g(\tau, s) X(f - \tau) \quad (8)$$

式中 $g(f, s) = \frac{1}{\sqrt{2\pi f}} e^{-\frac{f^2}{2s}}$ 代表核函数; s 为尺度参数; P 为高斯核函数长度,这里取 $P = 6\sqrt{s} + 1$ 。

一旦设定尺度参数 s_0 值,则尺度空间 $L(f, s_0)$ 的每个局部极大值对应一个单分量成分^[17],该单分量的带宽由相邻的两局部极小值确定。因此,通过检测 $L(f, s_0)$ 尺度曲线上的所有局部极大值,在获取的极值两侧搜索局部极小值来确定所有单分量的划分边界点。如此可以找到尺度空间谱分割的所有边界点 $\omega_n (n = 1, 2, 3, \dots, M - 1)$,将其划分为连续的 M 个频带,即自动找到尺度谱中 M 个有意义的模态分量。上述模态数目成为 MVMD 算法中最佳模态个数确定的依据(即 $K^* = M$)。此外,为保证每个分解的单分量中都能尽可能多地携带与故障相关的信息,尺度参数 n 应该取较大的值,然而过大的尺度参数将会忽略掉与故障相关的微弱特征。因此,本文尺度参数取值 $s = 3 \times f_{ch}^2, f_{ch}$ 为故障特征频率。

1.2.2 惩罚因子 α 的确定

惩罚因子影响着模态频带宽度,带宽较大可能导致模态中混有较多的噪声,而较小的带宽则会使模态中遗漏掉一些重要的故障信息。峭度对故障瞬态冲击尤其敏感,通常与故障相关的模态分量为非高斯信号,其峭度值大于 3。峭度值越大,说明模态分量中含有越丰富的故障冲击成分。其定义为

$$Ku = \frac{E[u(t) - \mu]^4}{\sigma^4} \quad (9)$$

式中 $u(t)$ 代表模态分量, μ 为 $u(t)$ 均值, σ 为标准差。

在确定模态个数 K 后,利用峭度判据构造最佳惩罚因子的选取策略,从而分解出包含故障冲击成分最丰富的模态。本文设置惩罚因子取值范围 $\alpha \in [1500, 4000]$ (参考文献[12-13]),步长为 100,通过迭代搜索方式选取最佳参数。在迭代分解过程

中,当模态的峭度值达到最大或保持不变时,对应的 α 即为分解所需最佳惩罚因子 α^* 。

2 自适应非凸重叠组收缩降噪算法

2.1 NOGS 降噪原理

设一维信号 $x \in R^N$, 在高斯白噪声 w 的污染下,得到的实际观测信号 y 可表达为

$$y(i) = x(i) + w(i) \quad (10)$$

假设 x 具有重叠组稀疏特性,求解 x (即从观测信号 y 中恢复 x),可转化为如下的稀疏优化问题

$$x^*(i) = \arg \min_{x(i)} \left\{ F(x(i)) = \frac{1}{2} \|y(i) - x(i)\|_2^2 + \gamma R(x(i)) \right\} \quad (11)$$

式中 $x^*(i)$ 为降噪后的信号; $\|y(i) - x(i)\|_2^2$ 为目标函数; γ 为正则化参数; $R(x(i))$ 为相应群组稀疏惩罚函数,其定义为

$$R(x(i)) = \sum_{i \in I} \left[\sum_{j \in J} |x(i+j)|^2 \right]^{\frac{1}{2}} \quad (12)$$

式中 i 为组索引, j 表示组大小,且每组具有相同尺寸 $|J|$, 本文取 $J = 5$ ^[18]。

所以,根据重叠组稀疏惩罚函数的定义,损失函数可转化为

$$F(x(i)) = \frac{1}{2} \|y(i) - x(i)\|_2^2 + \gamma \sum_{i \in I} \left[\sum_{j \in J} |x(i+j)|^2 \right]^{\frac{1}{2}} \quad (13)$$

式(13)中,目标函数与惩罚函数都是典型的凸函数。因此,本文利用优控最小化算法(Majorization-Minimization, MM)迭代求取上述最优解,得到

$$x(i) = \frac{y(i)}{1 + \lambda r(i)}, i \in I (I' = \{i \in I, x(i) \neq 0\}) \quad (14)$$

$$\text{式中 } r(i) = \sum_{j \in I} \left[\sum_{m \in J} |x(i-j+m)|^2 \right]^{\frac{1}{2}}, i \in I'.$$

重复迭代,直到算法收敛,最终输出信号 x 。NOGS 算法的详细介绍见文献[18]。

2.2 正则化参数的自适应选取

正则化参数 γ 在模型中起着平衡损失函数和正则化项的作用,因此,选择合适的参数 γ 对于提高周期脉冲的提取精度具有重要意义。 γ 需要正比于 σ , 这里 σ 是噪声的标准差,可表示为

$$\sigma = \frac{1}{N} \sum_{i=1}^N |u_{l_{\max}}(i) - \bar{u}_{l_{\max}}| \quad (15)$$

式中 $u_{i_{\max}}$ 为从 MVMD 分解结果中筛选出的敏感模态; N 表示 $u_{i_{\max}}$ 的长度; $\bar{u}_{i_{\max}}$ 为 $u_{i_{\max}}$ 的均值。因此, 正则化参数 γ 可写为

$$\gamma = p\sigma \quad (16)$$

γ 的选取依赖于合适的乘子 p 。本文构造评价指标 (Measure Index, MI) 来指导参数 γ 的选择, MI 定义为模态信号与残差信号的包络负熵^[19]之比:

$$MI = \frac{EE(\hat{u}_{i_{\max}})}{EE(r)} \quad (17)$$

$$EE(u_i) = \left\langle \frac{SE^2(u_i)}{\langle SE^2(u_i) \rangle} \ln \left(\frac{SE^2(u_i)}{\langle SE^2(u_i) \rangle} \right) \right\rangle \quad (18)$$

式中 $\hat{u}_{i_{\max}}$ 为降噪后的模态; r 为剩余信号, 是模态分量 $u_{i_{\max}}$ 与 $\hat{u}_{i_{\max}}$ 的差值; $SE(u_i) = |u_i + \text{Hilbert}(u_i)|^2$ 为模态 u_i 的平方包络; $\langle \cdot \rangle$ 表示求均值。EE 越大, 说明模态 u_i 中所包含周期性冲击特征越突出。在 ANOGS 算法中, 选取最大 MI 对应的 γ 作为最佳正则化参数, 从而得到故障冲击信号最终提取结果。

3 应用研究

MVMD-ANOGS 算法在行星齿轮箱早期故障特征提取中的具体实现步骤如下:

步骤 1: 获取齿轮箱故障振动信号 $x(t)$ 。

步骤 2: 计算信号 $x(t)$ 的尺度空间谱表示 $L(f, n)$, 根据 1.2.1 节尺度空间谱分割原理, 对 $L(f, n)$ 进行自适应分割, 获取尺度谱中 M 个有意义的模态, 令 VMD 中分解的最佳模态个数 $K^* = M$ 。

步骤 3: 在 $K^* = M$ 情况下, 设置惩罚因子的取值范围 $\alpha \in [1500, 4000]$, 步长为 100, 进行迭代分解, 比较每个惩罚因子下获取模态的峭度最大值大小, 将最大峭度对应的 α 判定为最优惩罚因子 α^* 。

步骤 4: 利用最佳参数组合 $[K^*, \alpha^*]$, 对 $x(t)$ 进行 MVMD 最优模态分解, 获得一系列的窄带模态分量 u_i , 选取峭度值最大的模态作为敏感模态 $u_{i_{\max}}$ 。

步骤 5: 使用 ANOGS 算法对敏感模态进行稀疏降噪, 且在评价指标 MI 最大值下获取最佳降噪结果 $\hat{u}_{i_{\max}}$, 从而显著提升模态中的周期性冲击特征。

步骤 6: 对降噪后的模态 $\hat{u}_{i_{\max}}$ 做包络解调处理, 提取故障特征频率并与理论值对比, 从而准确判行星齿轮箱故障的类型和位置。诊断方法流程图如图 1 所示。

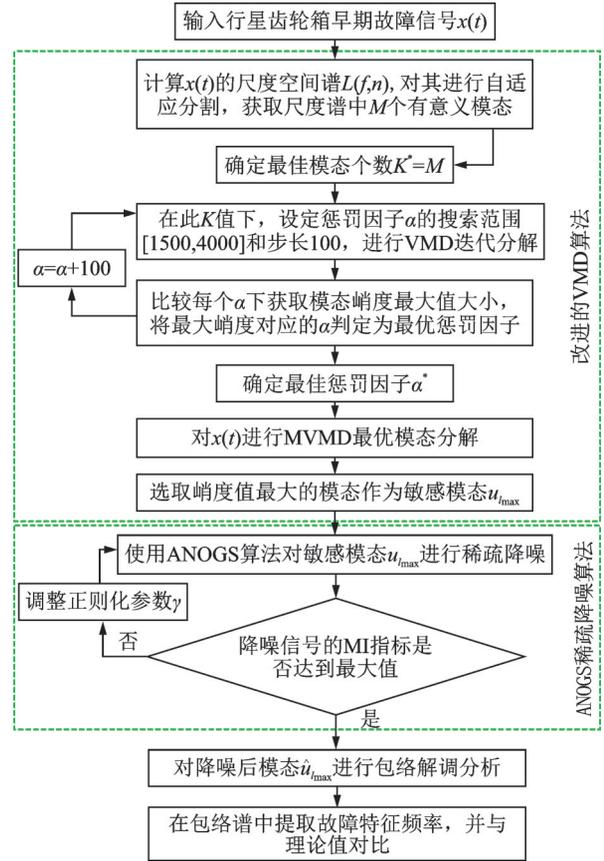


图1 基于 MVMD-ANOGS 算法的诊断流程图

Fig. 1 The flowchart of fault diagnosis based on MVMD-ANOGS

4 仿真信号分析

4.1 太阳轮局部故障模型

为验证所提方法, 构建太阳轮局部故障仿真模型。该信号模型 $y(t)$ 由太阳轮故障冲击振动 $g(t)$ 、其他零部件的正常旋转振动 $r(t)$ 、随机冲击 $h(t)$ 和随机噪声 $n(t)$ 这 4 部分构成, 其方程如下

$$y(t) = g(t) + r(t) + h(t) + n(t) \quad (19)$$

当太阳轮发生局部故障时, 传感器拾取到的故障振动信号可以用调幅调频的形式来表示^[20]。因此, 建立的太阳轮故障冲击模型为

$$\begin{cases} g(t) = s(t) \sum_{k=0}^K a_k(t) \cos [2\pi k f_m t + b_k(t) + \phi_k] \\ s(t) = 1 - \cos(2\pi f_{sr} t) \\ a_k(t) = \sum_{n=1}^N A_{kn} \cos(2\pi n f_s t + \alpha_{kn}) \\ b_k(t) = \sum_{n=1}^N B_{kn} \cos(2\pi n f_s t + \beta_{kn}) \end{cases} \quad (20)$$

式中 $s(t)$ 表示太阳轮旋转引起的调幅效应; $a_k(t)$ 和 $b_k(t)$ 分别表示太阳轮故障的调幅 (AM) 与调频 (FM) 函数; f_m 为齿轮副啮合频率; k 代表齿轮啮合倍

频的数量; f_{sr} 为太阳轮的绝对旋转频率; f_s 为太阳轮故障特征频率; ϕ_k 为第 k 次齿轮啮合谐波的初相位; A_{kn} 和 B_{kn} 分别为调幅和调频函数的幅值; α_{kn} 和 β_{kn} 分别是第 k 次齿轮啮合谐波的的第 n 个边带相位。

通常行星齿轮箱呈现周期性运转,所以用高、低谐波分量分别表示不同零部件旋转振动

$$r(t) = \sum_i C_i \sin(2\pi f_i t + \theta_i) \quad (21)$$

式中 i 为谐波分量的个数; C_i , f_i 和 θ_i 分别代表谐波分量的幅值、频率与相位。

此外,在故障数据采集过程中,可能由于电磁干扰或偶然撞击机壳产生随机冲击

$$\begin{cases} h(t) = \sum_j R_j S(t - Tr_j) \\ S(t) = e^{-\beta t} \sin(2\pi f_{RE} t) \end{cases} \quad (22)$$

式中 随机变量 R_j 与 Tr_j 分别为随机脉冲幅值和发生时刻; $S(t)$ 为单位脉冲分量, f_{RE} 为脉冲激发的共振频率; β 为阻尼系数。 $n(t)$ 为添加到仿真信号中信噪比 $SNR = -8$ dB 的高斯白噪声。仿真信号参数值如表 1 所示。

表 1 太阳轮局部故障信号仿真数值

Tab. 1 Simulation values of the sun gear fault signal

参数	数值	参数	数值
$K = N$	3	f_1/Hz	3
f_m/Hz	500	f_2/Hz	20
f_{sr}/Hz	20	C_1	0.6
$\alpha_{kn} = \beta_{kn}$	0	C_2	0.3
$A_1 = B_1$	0.5	θ_i	0
$A_2 = B_2$	0.4	f_{RE}/Hz	2000
$A_3 = B_3$	0.3	β	125
f_s/Hz	50		

4.2 太阳轮故障信号仿真实验

仿真分析时设置信号采样频率为 5120 Hz, 数据长度取 10240 点。将表 1 中的参数值分别代入式 (20)-(22) 中, 得到太阳轮故障冲击、其他零部件的旋转振动、随机冲击与高斯白噪声如图 2 所示。图 3 为混合故障仿真信号及其 FFT 频谱和包络谱。由图 3 可知, 时域波形中太阳轮故障激起的微弱周期性冲击已被强噪声完全掩盖, 频谱和包络谱中均无法提取与故障相关的特征频率。因此, 在强噪声干扰下, 传统的时、频域和包络解调方法失去对太阳轮早期微弱故障的检测能力。

为了准确提取太阳轮早期微弱故障特征, 利用本文方法对图 3(a) 中信号进行处理。首先, 计算混合故障仿真信号的尺度空间谱表示, 如图 4(a) 所示。可以看到, 尺度空间谱根据信号特征自适应确定有

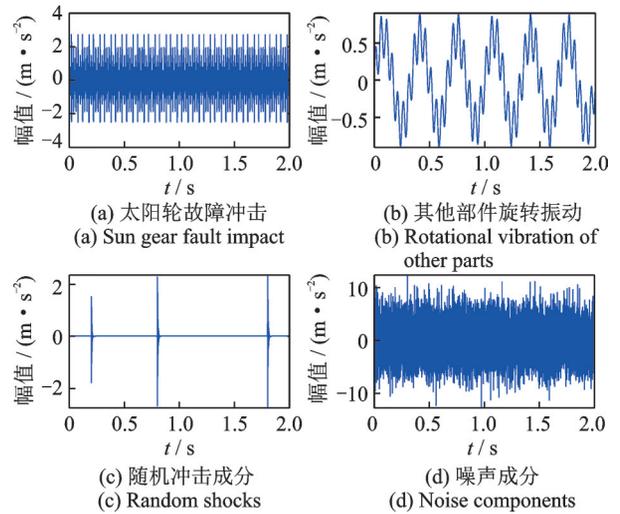


图 2 仿真信号各分量

Fig. 2 Each simulation component of sun gear fault

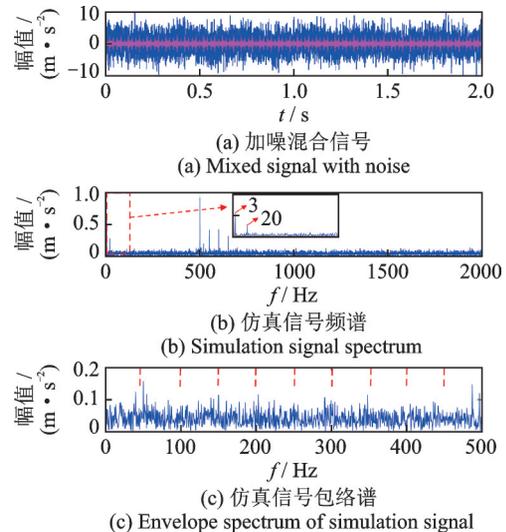


图 3 混合故障仿真信号及傅里叶频谱和包络谱

Fig. 3 Mixed fault simulation signal and its FFT and envelope spectrum

意义模式的分割边界(如图 4(a) 中红色点划线所示)及模式数目为 7。因此, 确定 MVMD 算法中最佳分解模式个数为 $K^* = 7$ 。然后, 设定惩罚因子 α 搜索范围为 $[1500, 4000]$, 采用峭度最大化原则选取最优参数 α^* 。图 4(b) 为惩罚因子与峭度最大值之间的关系。由图 4(b) 可知, 当 $\alpha = 3000$ 时获取峭度最大值, 此时判定 MVMD 中最优惩罚因子为 $\alpha^* = 3000$ 。确定最佳参数组合 $[K^*, \alpha^*] = [7, 3000]$ 后, 将混合故障仿真信号进行 MVMD 最优模式分解, 获取的 7 个模式分量如图 5 所示。根据敏感模式筛选准则, 选取峭度最大的模式 u_2 作为敏感模式。最后, 利用提出的 ANOGS 算法对 u_2 进行自适应稀疏降噪, 降噪后模式及其包络谱如图 6 所示。由图 6 可知, 降噪模式的时域波形中故障冲击特征得到增强且被清晰地呈现出来, 噪声与其他干扰成分被有效抑制; 同时, 对应包

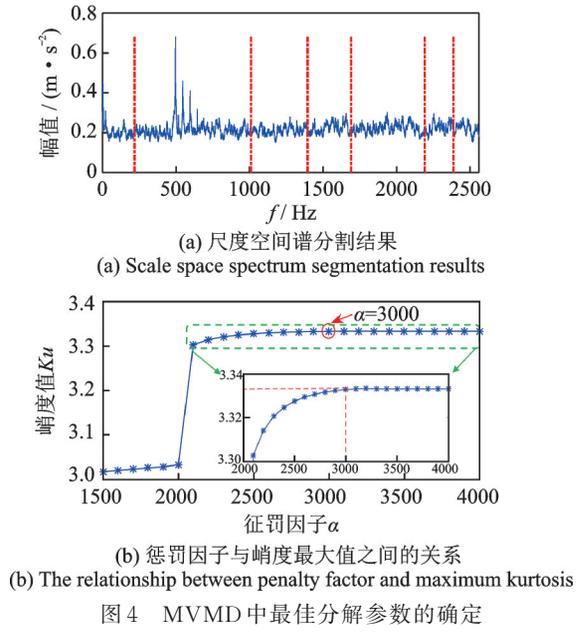


图4 MVMD中最佳分解参数的确定
Fig. 4 Determining the optimal decomposition parameters in MVMD algorithm

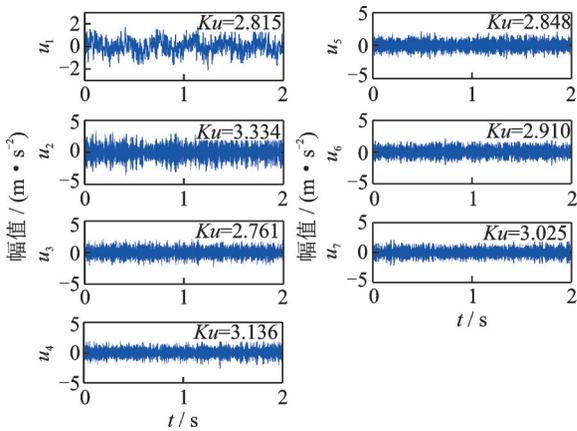


图5 混合故障仿真信号的MVMD分解结果
Fig. 5 The MVMD decomposition results of simulation signal

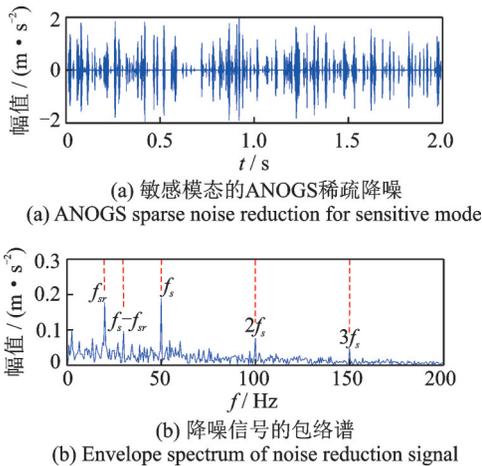


图6 本文MVMD-ANOGS方法对仿真信号处理结果($p=0.18, \gamma=0.13$)
Fig. 6 The processing results of simulation signal obtained by MVMD-ANOGS method ($p=0.18, \gamma=0.13$)

络谱中在太阳轮故障特征频率 f_s 及其倍频 nf_s ($n=1, 2, 3$)、太阳轮绝对转频 f_{sr} 处出现明显谱峰。据此可判断太阳轮发生了故障,这与仿真结果相吻合。

为验证本文方法获取VMD最佳参数组合 $[K^*, \alpha^*]=[7, 3000]$ 的可靠性,人为随机选择一组分解参数 $[10, 2500]$,使用更改参数后VMD和ANOGS算法对混合故障仿真信号进行处理。图7为非最优参数下获取的分析结果。由图7可知,模态信号经ANOGS降噪后,虽然时域信号中故障冲击成分被提取出来,但其包络谱中无关成分的干扰仍然较为严重,仅能看到不突出的太阳轮故障特征频率及其2倍频,处理效果不如本文方法。上述结果表明,人为主观参数选取会对VMD分解产生严重影响,导致分解的模态中不能提取清晰丰富的故障信息。而本文方法能有效地避免这些现象,自适应准确地实现最佳模态分解,便于后续的分析与处理。

作为对比,分别利用EEMD方法和FSK方法对混合故障仿真信号进行处理。图8为EEMD分解

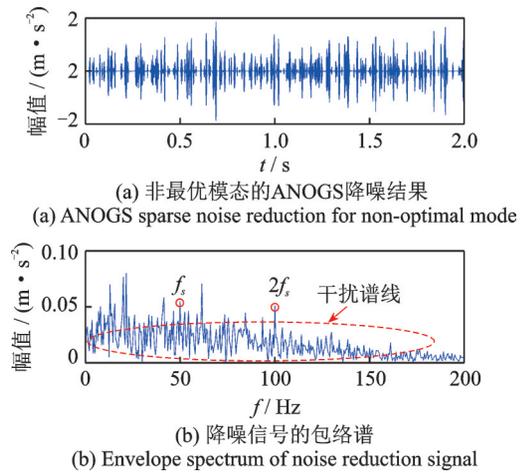


图7 更改参数后的分析结果($p=0.49, \gamma=0.22$)
Fig. 7 Analysis results of simulation signal after changing the optimized parameters ($p=0.49, \gamma=0.22$)

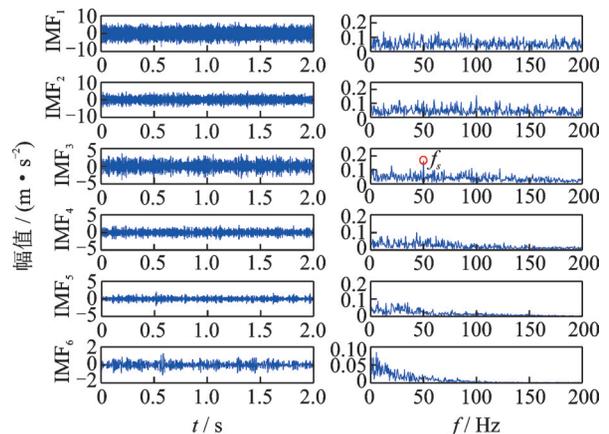


图8 EEMD对仿真信号分析结果(前6个IMF分量)
Fig. 8 Analysis results of simulation signal by EEMD (The first six IMF)

的前 6 个 IMF 分量。可以观察到,经 EEMD 分解后,仅在 IMF3 的包络谱中出现了太阳轮故障特征频率成分,但这未能提供足够的证据来准确判别太阳轮故障。图 9 为 FSK 方法处理后的结果。由图 9 可以看到,滤波后信号中并没有呈现出明显的故障冲击特征,对应包络谱中频率成分较为复杂,很难辨识太阳轮故障特征频率。对比图 6 中 MVMD-ANOGS 分析结果,上述 2 种方法都难以提取到明显的太阳轮故障特征信息,处理结果不理想。

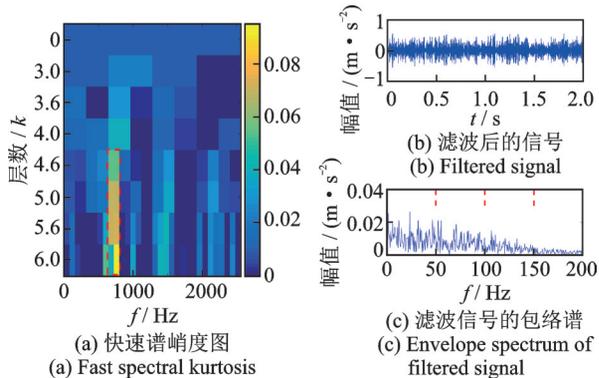


图 9 快速谱峭度方法对仿真信号分析结果

Fig. 9 Analysis results of simulation signal by FSK method

5 实验验证

5.1 实验说明

为检验所提方法在实际行星齿轮箱故障诊断中的有效性,在 Spectra Quest 公司研发的工业动力传动故障诊断综合实验台上开展行星齿轮箱故障实验,实验系统如图 10 所示。该实验台主要由交流驱动电机、行星齿轮箱、2 级平行轴齿轮箱与磁力加载器等组成。行星齿轮箱参数如表 2 所示。为模拟齿轮局部损伤,通过线切割技术在太阳轮和行星轮的某个轮齿上沿着齿根且垂直于轮齿中心线的方向加工宽为 0.15 mm,深为 1 mm 的微小贯通裂纹作为故障,故障齿轮如图 11 所示。此次实验使用



图 10 行星齿轮箱实验系统

Fig. 10 Planetary gearbox test bench

PCB352C33 型加速度传感器收集故障振动信号。传感器固定在行星齿轮箱壳体顶部的垂直、水平和轴向测点上。选用 DT9837 数据转换仪和一台安装 DAQ 软件的笔记本电脑作为本次实验的数据采集系统。实验时,行星齿轮箱恒定输入转速为 1200 r/min (即 $f_{sr} = 20$ Hz),数据采样频率设置为 5120 Hz,所用信号长度为 10240 点。在该实验条件下,计算行星齿轮箱中各齿轮故障特征频率如表 3 所示。

表 2 行星齿轮箱结构参数

Tab. 2 Planetary gearbox configuration parameters

齿轮	太阳轮	行星轮(数量)	齿圈
齿数	20	40(3)	100



图 11 齿轮故障件

Fig. 11 Tested gear faulty parts

表 3 行星齿轮箱中齿轮特征频率

Tab. 3 Characteristic frequencies of gears in planetary gearbox

参数	数值/Hz
啮合频率 f_m	333.33
太阳轮绝对转频 f_{sr}	20
行星架转频 f_c	3.33
太阳轮故障特征频率 f_s	50
行星轮故障特征频率 f_p	8.33
齿圈故障特征频率 f_r	9.99

5.2 实测信号分析

5.2.1 太阳轮故障信号

图 12 为太阳轮齿根裂纹故障信号时域波形及其 FFT 频谱和包络谱。由图 12 可知,时域波形中出现一些杂乱的峰值,但未见局部故障激发的明显周期性冲击。频谱中在 2 倍啮合频率 $2f_m$ 处峰值比较凸出,但由于强背景噪声的影响,其两侧的调制频率已基本被淹没。包络谱中除了可识别到太阳轮绝对转频 f_{sr} 和行星架转频的倍频 nf_c ($n = 7, 10, 13$) 成分外,在太阳轮故障特征频率处无突出的峰值。因此,采用传统的时、频域分析和包络谱方法对太阳轮早期故障信号进行分析难以取得满意的结果。

为了揭示隐藏在动态信号中的太阳轮早期微弱故障征兆,利用本文方法对太阳轮故障信号进行分析。首先,计算太阳轮故障信号的尺度空间谱,如图

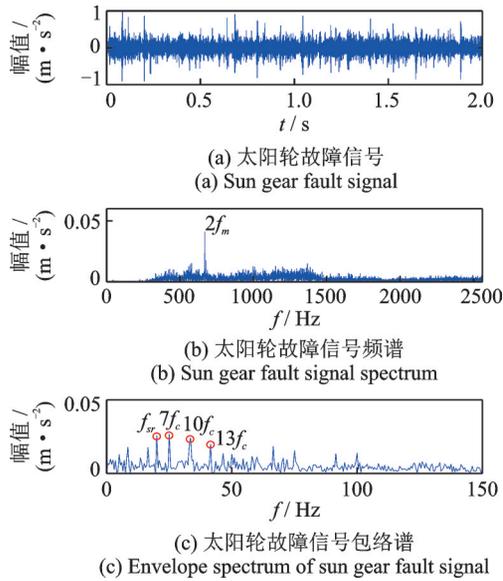
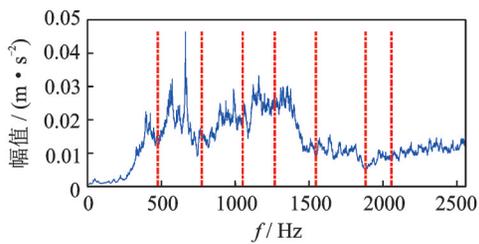


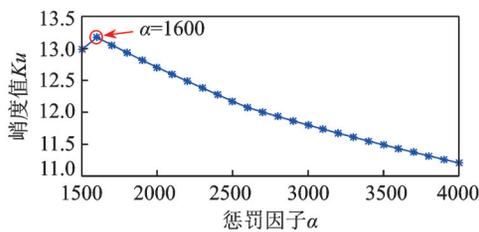
图12 太阳轮齿根裂纹故障信号

Fig. 12 Fault signal of sun gear tooth root crack

13(a)所示。可以看到,尺度空间谱根据太阳轮故障信号特征自适应划分的模态分割边界(如图13(a)中红色点划线所示)及模态数目为8。因此, MVMD算法中最佳模态个数取值 $K^* = 8$ 。然后, 设置惩罚因子 α 的搜索区间为 $[1500, 4000]$, 利用峭度最大化原则选取最佳参数。惩罚因子与峭度最大值之间的关系如图13(b)所示, 当 $\alpha = 1600$ 时取得峭度最大值, 所以获取最佳惩罚因子为 $\alpha^* = 1600$ 。确定最佳参数组合 $[K^*, \alpha^*] = [8, 1600]$ 后, 采用MVMD对太阳轮故障信号进行最优模态分解, 获得的8个模态分量如图14所示。根据敏感模态确定准则, 选取峭度最大的 u_8 作为敏感模态。最后, 利用ANOGS算法对敏感模态 u_8 进行自适应稀疏降噪处



(a) 尺度空间谱分割结果
(a) Scale space spectrum segmentation results



(b) 惩罚因子与峭度最大值之间的关系
(b) The relationship between penalty factor and maximum kurtosis

图13 MVMD中最佳分解参数的确定

Fig. 13 Determining the optimal parameters of MVMD

理, 降噪模态信号及其包络谱如图15所示。可以看到, 降噪后隐藏在动态信号中的微弱故障冲击特征得到增强且被清晰地揭示出来, 噪声及其他干扰成分得到明显抑制。同时, 在其对应包络谱中太阳轮故障特征频率的 $n/(3f_s)$ ($n = 1, 2, \dots, 6$)倍频处出现明显的谱线(实际齿轮箱中, 3个行星轮不可能完全相同, 这种差异使3个行星轮与太阳轮啮合时产生的故障冲击被视作3个不同的冲击, 因此会出现 $n/3$ 倍的太阳轮故障频率成分)。因此, 上述分析结果表明太阳轮发生了局部损伤, 这与实验设置相一致。

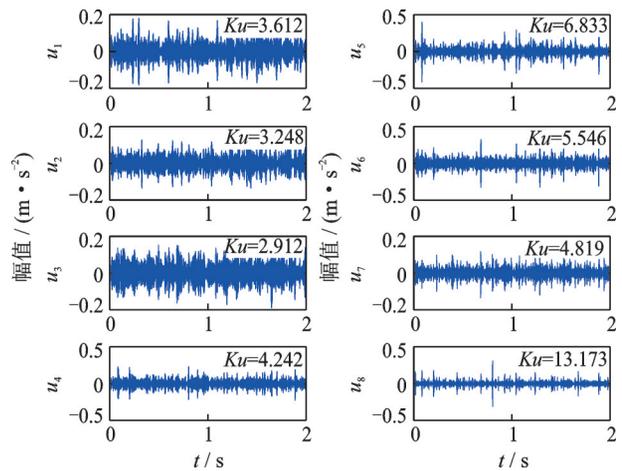
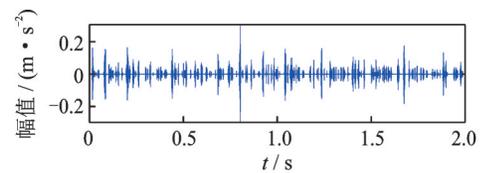
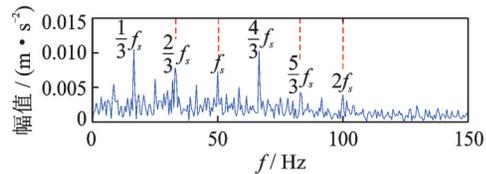


图14 太阳轮故障信号的MVMD分解结果

Fig. 14 MVMD decomposition results of sun gear fault signal



(a) 敏感模态的ANOGS稀疏降噪
(a) ANOGS sparse noise reduction for sensitive mode



(b) 降噪信号的包络谱
(b) Envelope spectrum of noise reduction signal

图15 本文MVMD-ANOGS方法太阳轮故障信号处理结果($p=0.50, \gamma=0.01$)

Fig. 15 The processing results of sun gear fault signal obtained by MVMD-ANOGS method ($p=0.50, \gamma=0.01$)

为验证获取的最优参数组合 $[K^*, \alpha^*] = [8, 1600]$ 在实际行太阳轮早期故障信号中的可靠性, 随机选取一组分解参数 $[10, 2200]$, 利用更改参数的VMD和ANOGS算法对太阳轮故障信号进行分析。图16为非最优参数下取得的分析结果。虽然降噪模态的时域波形中出现了冲击特征, 但在其包络谱中仅能提取到 $n/(3f_s)$ ($n = 2, 3, 4$)倍的太阳

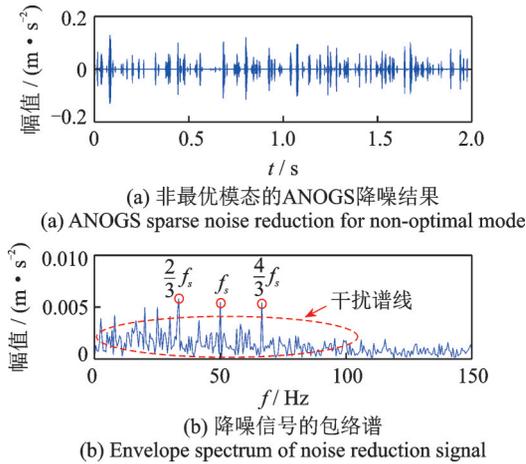


图 16 更改参数后的分析结果($p=0.61, \gamma=0.02$)

Fig. 16 Analysis results of sun gear fault signal after changing the optimized parameters ($p=0.61, \gamma=0.02$)

轮故障特征频率成分,且存在许多未知谱线,这对准确判别故障非常不利。分析表明,随机参数选取会对VMD分解产生严重影响,导致分解模态中不能有效提取太阳轮早期故障特征信息。而本文方法能避免这些现象,自适应准确地实现最佳模态分解,从而提取更加清晰丰富的故障特征。

作为对比,分别采用EEMD方法和FSK方法对太阳轮早期故障信号进行处理。图17为EEMD获取的前7个IMF分量。可以看出,EEMD得到的模态信号时域波形中并未提取出明显的故障冲击特征;同时,在对应的包络谱中仅提取到行星架转频和太阳轮绝对转频的倍频成分,太阳轮故障特征频率处无明显谱线。图18为FSK方法的分析结果。由图18可知,经FSK滤波后信号中出现了强弱不等的冲击特征,但这些冲击的周期性并不清晰。在滤波

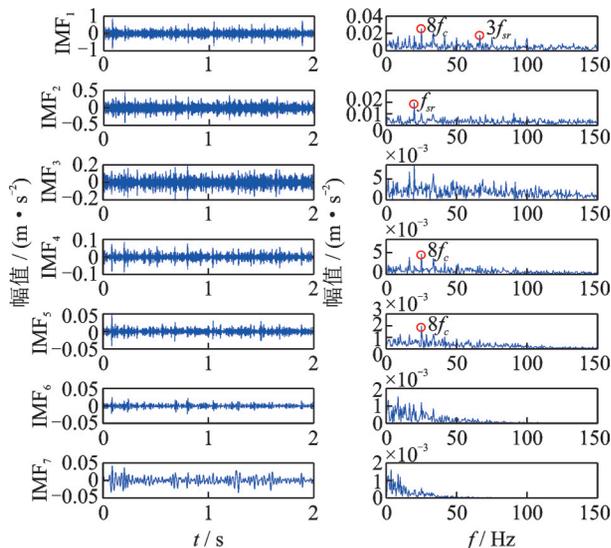


图 17 EEMD对太阳轮故障信号分析结果(前7个IMF)

Fig. 17 Analysis results of sun gear fault signal by EEMD (the first seven IMF)

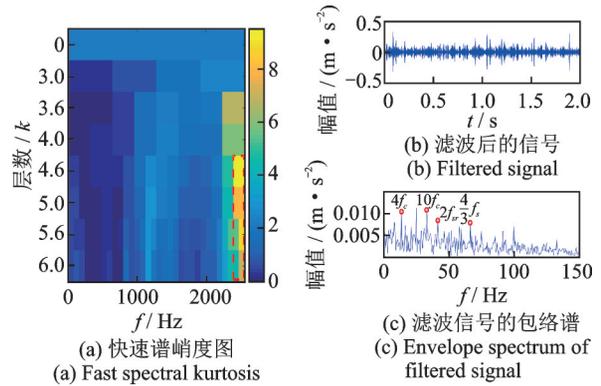


图 18 快速谱峭度方法对太阳轮故障信号的分析结果

Fig. 18 Analysis results of sun gear fault signal by FSK method

信号的包络谱中,仅可以看到行星架转频的倍频及 $4/3f_s$ 倍的太阳轮故障特征频率。对比可知,EEMD方法和FSK方法都难以将太阳轮故障激发的微弱周期性冲击特征清晰地提取出来,处理结果不如图15中本文方法理想。

5.2.2 行星轮故障信号

图19为行星轮齿根裂纹故障信号时域波形及其FFT频谱和包络谱。由图19可知,时域信号中行星轮局部故障引起的微弱周期性冲击已被强噪声所掩盖,从其频谱和包络谱中均无法识别到表征行星轮故障的频率信息。因此,传统的时、频域分析及包络谱方法无法对行星轮早期故障进行有效合理的诊断。

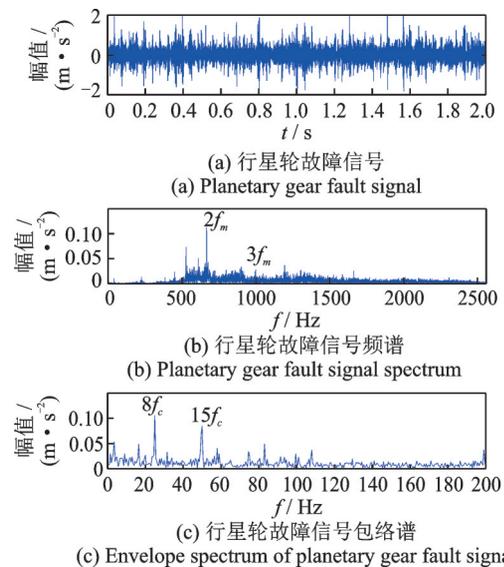


图 19 行星轮齿根裂纹故障信号

Fig. 19 Fault signal of planetary gear tooth root crack

对行星轮故障信号应用本文方法进行特征提取。图20(a)为行星轮故障信号的空间尺度谱表示。由图20(a)可知,根据行星轮故障信号特征,空间尺度谱自适应划分模态的分割边界并确定出8个有意义的模态分量。因此,设定MVMD中最佳模

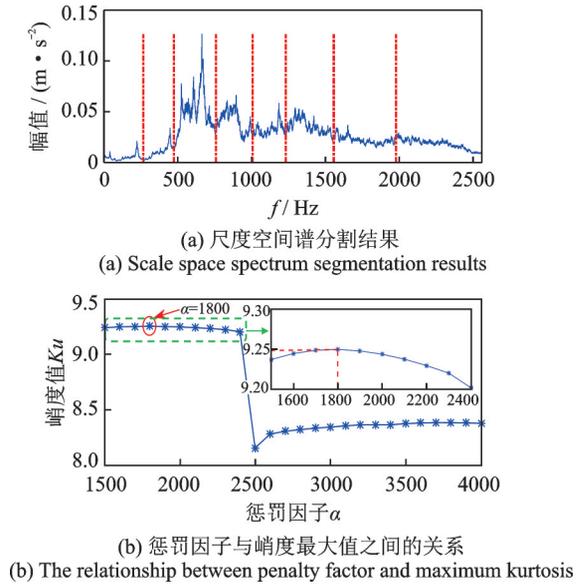


图20 MVMD中最佳分解参数的确定
Fig. 20 Determining the optimal parameters of MVMD

态个数为 $K^* = 8$ 。图20(b)为惩罚因子与峭度最大值之间的关系,当 $\alpha = 1800$ 时取得峭度最大值,所以确定最佳惩罚因子为 $\alpha^* = 1800$ 。在最佳参数组合 $[K^*, \alpha^*] = [8, 1800]$ 下,利用MVMD对行星轮故障信号进行最优模态分解,获取的8个模态分量如图21所示。根据敏感模态筛选准则,选取峭度最大的模态 u_8 作为敏感模态。进一步地,利用提出的ASCSD算法对 u_8 进行稀疏降噪,降噪模态信号及其包络谱如图22所示。可以观察到,降噪后表征行星轮故障的周期性冲击特征得到突显且被清晰地提取出来,噪声及其他干扰被有效抑制;同时,在相应包络谱中行星轮故障特征频率的倍频 nf_p ($n = 2, 3, 6, 8, 9, 13, 15$) 处呈现出较明显的谱峰。上述分析结果表明行星轮出现了故障,这与实验设置相符。

为验证最佳参数组合 $[K^*, \alpha^*] = [8, 1800]$ 在实际行星轮早期故障诊断中的可靠性,随机选取参数

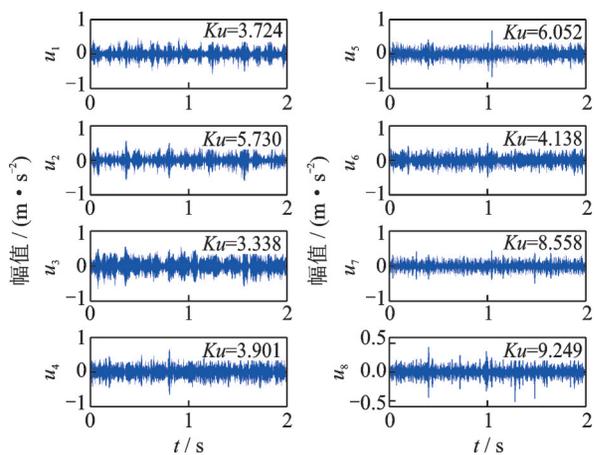


图21 行星轮故障信号的MVMD分解结果
Fig. 21 MVMD decomposition results of planetary gear fault signal

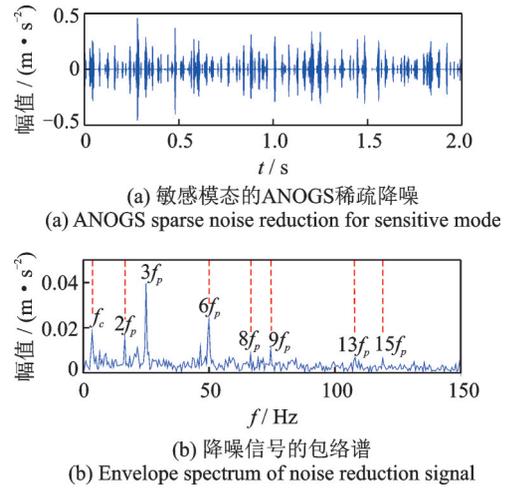


图22 本文MVMD-ANOGS方法行星轮故障信号处理结果 ($p = 0.23, \gamma = 0.01$)

Fig. 22 The processing results of planetary gear fault signal by MVMD-ANOGS method ($p = 0.23, \gamma = 0.01$)

组合 $[5, 1200]$, 利用更改参数后的VMD和ANOGS方法对行星轮故障信号进行处理。图23为非最优参数下获取的处理结果。可以看到,虽然降噪模式的时域波形中出现幅值较大的冲击,但其对应的包络谱中存在许多无关干扰频率,仅能识别出3倍和6倍的行星轮故障特征频率成分,且谱线的整体幅值水平低于图22(b)。分析表明,随机参数选取导致VMD分解结果无法得到保障,从而不能在其分解的模态中提取清晰丰富的行星轮故障特征信息。而本文方法能避免这些现象,自适应准确地实现最佳模态分解,从而更有利于早期故障特征的识别。

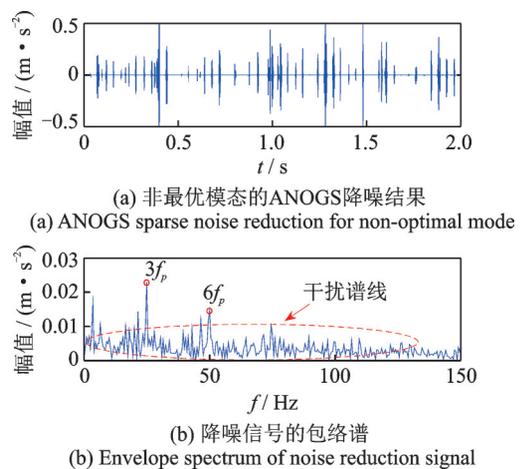


图23 更改参数后的分析结果 ($p = 0.81, \gamma = 0.04$)

Fig. 23 Analysis results of simulation signal after changing the optimized parameters ($p = 0.81, \gamma = 0.04$)

作为对比,分别采用EEMD方法和FSK方法对行星轮故障振动信号进行处理。图24为EEMD方法分解的前7个IMF分量。由图24可知,获取分量的时域波形中冲击特性并不明显,对应IMF1和IMF3分

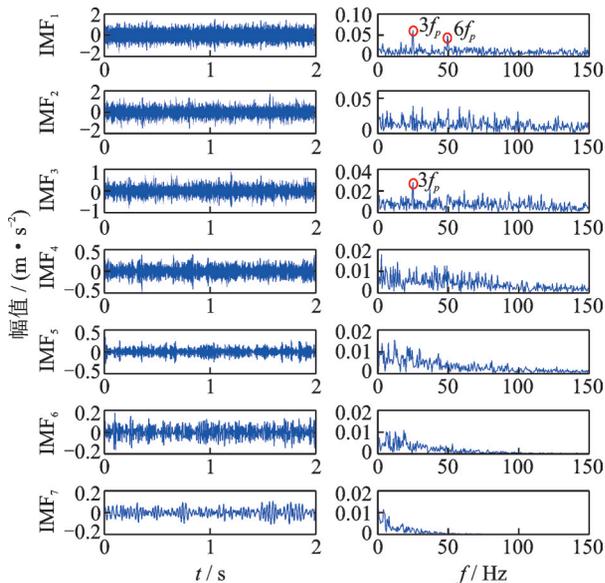


图 24 EEMD 对行星轮故障信号分析结果(前 7 个 IMF)
Fig. 24 Analysis results of planetary gear fault signal by EEMD (the first seven IMF)

量的包络谱在行星轮故障特征频率的倍频处出现微小谱峰,但其他无关频率的幅值较为突出,不利于故障特征准确辨识。图 25 为 FSK 方法分析结果。可以看到,滤波后时域波形中出现了部分幅值较大的冲击成分,在滤波信号的包络谱中,虽然可以提取到行星轮故障特征频率的倍频,但整个谱图中杂乱谱线较多,干扰特征频率的识别。与图 22 中本文方法对比可知,EEMD 和 FSK 方法虽可以提取部分故障特征,但其在噪声抑制和抗干扰方面表现不足,无法将微弱的行星轮故障特征清晰地提取出来。

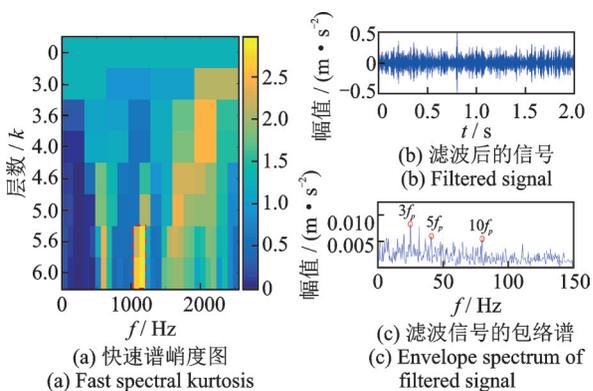


图 25 快速谱峭度方法对行星轮故障信号的分析结果
Fig. 25 Analysis results of planetary gear signal by FSK method

6 结 论

本文提出了一种基于 MVMD-ANGOS 的早期故障诊断方法,并将其应用于强背景噪声下行星齿轮箱微弱故障特征提取中。通过信号仿真和工程实

验数据的分析可知,该方法能够显著提升微弱的周期性故障冲击特征,在行星齿轮箱早期故障诊断和预测中具有一定的优势。

(1) MVMD 方法能够自适应获取分解所需的最佳模态个数 K 和惩罚因子 α 且具有理论依据,有效解决了 VMD 中主要影响参数选取需要依赖人为经验与先验准则的不足,实现了模态最优分解。

(2) 提出的 ANGOS 算法能够对敏感模态进行自适应稀疏降噪,最大限度地消除噪声和其他无关成分的干扰,从而突显模态中微弱的故障冲击特征,提高了故障诊断的精确度和鲁棒性。

(3) 通过与传统的 VMD, EEMD 和快速谱峭度方法对比,本文方法能提取到更加清晰明显的故障特征频率成分且故障表征能力得到提高,实现了行星齿轮箱早期故障的准确识别与诊断,为实际工程应用提供了一种思路。

参考文献:

- [1] Lei Y, Lin J, Zuo M J, et al. Condition monitoring and fault diagnosis of planetary gearboxes: A review[J]. Measurement, 2014, 48: 292-305.
- [2] Guo Y, Zhao L, Wu X, et al. Tooth root crack detection of planet and sun gears based on resonance demodulation and vibration separation[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69 (1): 65-75.
- [3] Dybala J, Zimroz R. Rolling bearing diagnosing method based on empirical mode decomposition of machine vibration signal[J]. Applied Acoustics, 2014, 77: 195-203.
- [4] Smith J S. The local mean decomposition and its application to EEG perception data[J]. Journal of the Royal Society Interface, 2005, 2(5): 443-454.
- [5] Yang Y, Cheng J, Zhang K. An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems[J]. Measurement, 2012, 45: 561-570.
- [6] 雷亚国,孔德同,李乃鹏,等. 自适应总体平均经验模式分解及其在行星齿轮箱故障检测中的应用[J]. 机械工程学报, 2014, 50(3): 64-70.
- Lei Y, Kong D, Li N, et al. Adaptive ensemble empirical mode decomposition and its application to fault detection of planetary gearboxes[J]. Journal of Mechanical Engineering, 2014, 50(3): 64-70.
- [7] Gilles J. Empirical wavelet transform[J]. IEEE Transactions on Signal Processing, 2013, 61(16): 3999-4010.
- [8] Dragomiretskiy K, Zosso D. Variational Mode Decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [9] Wang Y, Markert R, Xiang J, et al. Research on varia-

- tional mode decomposition and its application in detecting rub-impact fault of the rotor system[J]. *Mechanical Systems and Signal Processing*, 2015, 60: 243-251.
- [10] Li Z, Jiang Y, Guo Q, et al. Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving-speed variations[J]. *Renewable Energy*, 2018, 116: 55-73.
- [11] Wang C, Li H, Huang G, et al. Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum [J]. *IEEE Access*, 2019, 7: 31501-31516.
- [12] 唐贵基, 王晓龙. 参数优化变分模态分解方法在滚动轴承早期故障诊断中的应用[J]. *西安交通大学学报*, 2015, 49(5): 73-81.
- Tang G, Wang X. Parameter optimized variational mode decomposition method with application to incipient fault diagnosis of rolling bearing[J]. *Journal of Xi'an Jiaotong University*, 2015, 49(5): 73-81.
- [13] Zhang X, Miao Q, Zhang H, et al. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery[J]. *Mechanical Systems and Signal Processing*, 2018, 108: 58-72.
- [14] Gu R, Chen J, Hong R, et al. Incipient fault diagnosis of rolling bearings based on adaptive variational mode decomposition and Teager energy operator[J]. *Measurement*, 2020, 149: 106941.
- [15] Zhang M, Jiang Z, Feng K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump[J]. *Mechanical Systems and Signal Processing*, 2017, 93: 460-493.
- [16] Pan J, Chen J, Zi Y, et al. Mono-component feature extraction for mechanical fault diagnosis using modified empirical wavelet transform via data-driven adaptive Fourier spectrum segment[J]. *Mechanical Systems and Signal Processing*, 2016, 72-73: 160-183.
- [17] 丁显, 徐进, 滕伟, 等. 基于无参数经验小波变换的风电齿轮箱故障特征提取[J]. *振动与冲击*, 2020, 39(8): 99-105.
- Ding X, Xu J, Teng W, et al. Fault feature extraction of a wind turbine gearbox using adaptive parameterless empirical wavelet transform[J]. *Journal of Vibration and Shock*, 2020, 39(8): 99-105.
- [18] Chen P, Selesnick I W. Translation-invariant shrinkage/thresholding of group sparse signals[J]. *Signal Processing*, 2014, 94(1): 476-489.
- [19] Antoni J. The infogram: Entropic evidence of the signature of repetitive transients[J]. *Mechanical Systems and Signal Processing*, 2016, 74: 73-94.
- [20] Zhang M, Wang K, Wei D, et al. Amplitudes of characteristic frequencies for fault diagnosis of planetary gearbox[J]. *Journal of Sound and Vibration*, 2018, 432: 119-132.
- [21] Feng Z, Lin J, Zuo M J. Joint amplitude and frequency demodulation analysis based on intrinsic time-scale decomposition for planetary gearbox fault diagnosis[J]. *Mechanical Systems and Signal Processing*, 2016, 72-73: 223-240.

Incipient fault feature extraction of planetary gearbox based on modified VMD and non-convex overlap group shrinkage denoising

WANG Chao-ge¹, LI Hong-kun², CAO Shun-xin², ZHOU Qiang², LIU Ai-qiang², REN Xue-ping³

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

2.School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China;

3.Institute of Mechanical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China)

Abstract: Aiming at the problems that the incipient weak fault characteristics of the planetary gearbox are difficult to be identified under strong background noise, and the mode number K and penalty factor α of variational mode decomposition (VMD) must be set in advance and cannot be adaptively determined, a modified variational mode decomposition (MVMD) algorithm is proposed. The MVMD method determines the mode number K by adaptive scale space spectrum segmentation of the fault signal, and establishes the maximum kurtosis criterion to automatically select the penalty factor α . On this basis, a novel method for incipient fault diagnosis of planetary gearbox based on the MVMD and adaptive non-convex overlap group shrinkage (MVMD-ANOGS) is proposed. The collected vibration signal of the planetary gearbox is decomposed by MVMD, and then a series of narrow-band modal components are obtained. The proposed ANOGS algorithm performs sparse denoising on the sensitive mode with the maximum kurtosis, thereby highlighting the fault impact characteristics of the mode signal. The envelope demodulation analysis is performed on the denoised mode, and the obvious fault characteristic frequencies are extracted from its envelope spectrum. The feasibility of the proposed method is validated using both the numerical simulation and practical experimental dates of planetary gearbox. Moreover, compared with the traditional VMD, EEMD and fast spectral kurtosis methods, the proposed method can extract clearly weak fault impact features and improve the incipient fault identification accuracy of planetary gearbox.

Key words: fault diagnosis; planetary gearbox; variational mode decomposition; overlap group shrinkage algorithm; feature extraction

作者简介: 王朝阁(1992-),男,博士,讲师。电话:18342236929;E-mail: cgwang@shmtu.edu.cn

通讯作者: 李宏坤(1984-),男,教授。电话:13084158910;E-mail: lihk@dlut.edu.cn