

采用双小波字典的齿轮箱复合故障特征提取方法

汪向东^{1,2}, 刘强^{1,2}, 向蕾蕾³, 苏振³, 陈保家^{1,2}, 陈述⁴, 章文忠^{1,2}

(1. 三峡大学水电机械设备设计与维护湖北省重点实验室, 443002, 湖北宜昌; 2. 三峡大学机械与动力学院, 443002, 湖北宜昌; 3. 中国长江电力股份有限公司, 430000, 武汉; 4. 三峡大学水利与环境学院, 443002, 湖北宜昌)

摘要: 针对齿轮箱复合故障诊断中齿轮与轴承特征相互干扰、现有方法难以有效分离高频谐波与冲击成分的难题,提出一种基于双小波字典的稀疏正则化特征提取方法。通过构建 Morlet 小波-齿轮振动匹配模型与 Laplace 小波-轴承冲击响应模型的联合字典,结合广义极小极大凹惩罚函数优化稀疏分解过程,在保证目标函数凸性的同时克服了传统 l_1 范数导致的幅值衰减问题。仿真与实验结果表明:所提方法实现了齿轮局部损伤与轴承外圈故障的同步解耦,相比 MOMEDA 及 IESFOgram 方法,特征频率谐波提取完整度高;对于前 5 倍故障特征频率的信噪比提升了 1~2 dB。

关键词: 齿轮箱故障;稀疏分解;复合故障;故障诊断

中图分类号: TH17 **文献标志码:** A

DOI: 10.7652/xjtuxb202509010 **文章编号:** 0253-987X(2025)09-0099-11

Dual-Wavelet Dictionary Based Feature Extraction Method for Compound Faults in Gearboxes

WANG Xiangdong^{1,2}, LIU Qiang^{1,2}, XIANG Leilei³, SU Zhen³, CHEN Baojia^{1,2}, CHEN Shu⁴, ZHANG Wenzhong^{1,2}

(1. Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance, China Three Gorges University, Yichang, Hubei 443002, China; 2. College of Mechanical and Power Engineering, China Three Gorges University, Yichang, Hubei 443002, China; 3. China Yangtze Power Co., Ltd., Wuhan 430000, China; 4. College of Hydraulic & Environmental Engineering, China Three Gorges University, Yichang, Hubei 443002, China)

Abstract: To address the challenges of mutual interference between gear and bearing features in gearbox compound fault diagnosis and the difficulty of existing methods in effectively separating high-frequency harmonics and impact components, a sparse regularization feature extraction method based on a dual wavelet dictionary is proposed. By constructing a joint dictionary comprising a Morlet wavelet-gear vibration matching model and a Laplace wavelet-bearing impact response model, combined with the generalized minimax concave penalty function to optimize the sparse decomposition process, the convexity of the objective function is preserved while overcoming the amplitude attenuation issue caused by traditional norms. Simulations and experiments demonstrate that the proposed method achieves synchronous decoupling of gear localized damage and bearing outer ring faults. Compared to the MOMEDA and IESFOgram

收稿日期: 2025-02-17。 作者简介: 汪向东(2000—),男,硕士生;陈保家(通信作者),男,教授,博士生导师。 基金项目: 水资源高效利用与工程安全国家工程研究中心开放研究基金资助项目(1524020003);湖北省重点研发计划资助项目(2023BAB129);湖北省中央引导地方科技发展专项资助项目(2024CSA095)。

网络出版时间: 2025-06-09

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20250607.1714.002>

methods, it exhibits higher completeness in extracting harmonic features of fault characteristic frequencies, with a 1—2 dB improvement in the signal-to-noise ratio for the first five multiples of the fault characteristic frequency.

Keywords: gearbox fault; sparse decomposition; compound fault; fault diagnosis

齿轮箱在汽车、航空、能源和工程机械等关键领域中扮演着动力传递和结构支撑的重要角色,其能否正常运行对机械设备的安全服役至关重要^[1]。如果长期在高负载、变工况和高速运转等恶劣环境下运行,齿轮箱零部件容易发生失效^[2]。齿轮箱内部结构紧凑、零件众多,当关键部件发生故障时,受环境噪声和系统内非故障成分的干扰,故障信号将变得极为复杂^[3]。在实际操作中,轻微损伤不会立即触发停机检修,但是随着故障的发展及传动部件的相互耦合作用,后期极易引发多部件复合故障^[4]。

目前的研究多集中于单一故障诊断,对于复合故障的诊断因服役环境的复杂性和多部件失效的随机性而更具挑战。国内外学者开展广泛研究,提出了一系列方法应用于轴承、齿轮的复合故障特征提取中,如谱峭度(SK)类与改进包络谱类方法、解卷积方法、稀疏表示方法等。接下将围绕这些方法展开介绍。

SK类与改进包络谱类方法因其在捕捉信号冲击性方面的有效性而受到广泛关注。SK将峭度概念引入频谱,通过幅值的变化来反映信号的冲击性。Antoni等对SK方法进行了重新定义,将其视为随机非平稳过程,并提出了谱峭度和快速峭度图^[5-6]。Mauricio等^[7]通过在1/3二叉树上映射诊断特征值来确定最优积分频带,并提出了特征优化图改进包络谱(IESFOgram),该方法在行星齿轮箱和发动机齿轮箱的故障诊断中显示出更优越的性能。当谱峭度类与改进包络谱类方法在解调故障频带时,常通过固定带宽和中心频率的滤波器,并依据特定指标来锁定故障频带。这些方法可能存在滤波器不够灵活、过分解或欠分解的问题。

解卷积方法假设故障信号是故障脉冲与系统传递函数的卷积,通过迭代调整滤波器系数,构建逆滤波器以提取故障特征。Wiggins^[8]提出了最小熵解卷积(MED),通过优化峭度目标函数,迭代寻找最优滤波器,以突出故障信号并降低噪声影响。McDonald等在引入周期性概念后,有学者发展了两种新的解卷积方法:多点最优最小熵解卷积(MOMEDA)^[9]和最大二阶循环平稳盲解卷积^[10]。解卷积类算法在面对复合故障问题时,单一的解卷积滤波器只能提取局部特征而无法适应不同故障类型。

稀疏表示方法是一种在信号处理中广泛应用的技术,旨在通过选择合适的过完备字典,使得信号能够以简洁的方式进行表示,进而提取出信号中的关键特征成分。稀疏表示相对于谱峭度类和改进包络谱类方法能更全面捕捉复合故障信号中的局部特征和全局特征;相较于解卷积类方法能够提取出更明显的特征,具有一定的抗干扰能力。因而稀疏表示被认为是上述方法的改进与升华^[11],其核心思路是利用信号本身的稀疏性质,即信号可以用较少的非零系数来表示,从而实现了对信号特征的精炼提取。

稀疏表示字典分为两类:解析字典和学习字典。解析字典基于预设的波形函数来匹配特征,如小波变换^[12]和调Q因子小波变换^[13]等。学习字典则通过训练信号样例来学习对象的特征,包括K-奇异值分解字典(K-SVD)^[14]、移不变字典^[15]和广义主成分分析字典等^[16]。

稀疏分解算法直接影响分解的收敛性和运算速度。常用的算法包括基追踪(MP)^[17]、正交匹配追踪(OMP)^[18]和前向后向分裂算法(FBS)^[19]等。

稀疏惩罚函数可促进稀疏目标函数最小化,有助于提取信号的关键特征成分,平衡重构信号的稀疏性和保真性。He等^[20]通过反正切惩罚函数和分裂增强拉格朗日收缩算法,有效提取了齿轮箱中齿轮的复合故障特征。Selesnick等^[19]提出的广义极小极大凹(GMC)惩罚函数,能够有效促进信号的稀疏性,同时保持整体目标函数的凸性。

齿轮箱中齿轮与轴承如果同时发生故障,尽管会共同引起周期性冲击,但两者故障信号并不完全具有相似性,同时二者通常都会被强背景噪声掩盖,而轴承的故障冲击信号通常会弱于齿轮故障冲击信号,这一系列原因使得从测得信号中有效分离出它们的故障特征变得极具挑战性。针对此,本文为匹配齿轮与轴承的特征分别进行Morlet小波和Laplace小波的字典构造,并引入GMC惩罚函数以促进稀疏表示模型更准确的估计信号成分。

1 双小波字典复合故障特征提取

双小波字典复合故障特征提取的方法主要分为两个阶段:双字典构建和故障特征提取。在双字典

构建阶段,先获取齿轮箱的故障信号,再利用奇异值分解(SVD)技术从这些信号中提取出齿轮故障的特征字典。随后,通过OMP算法得到残差信号。对残差信号进行SVD分解,以提取轴承故障的特征字典。最终,构建出两个特征字典: $\mathbf{A}_1$ 代表齿轮故障特征字典, $\mathbf{A}_2$ 代表轴承故障特征字典。在故障特征提取阶段,首先使用FBS算法对信号进行处理。然后,利用构建的双字典 $\mathbf{A}_1$ 和 $\mathbf{A}_2$, 分别提取齿轮和轴承的故障特征。最后,通过重构信号的包络谱分析,进一步提取故障特征。详细的算法如图1所示,图中, $\mathbf{a}_1$ 和 $\mathbf{a}_2$ 分别为轴承、齿轮故障稀疏系数。

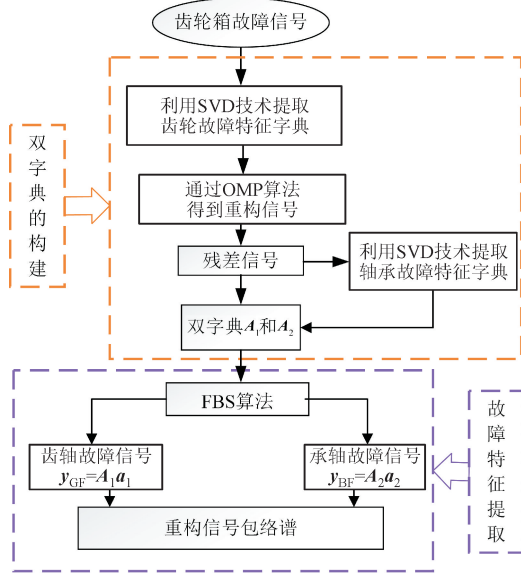


图1 非凸正则化双字典提取齿轮箱复合故障特征流程

Fig.1 Flowchart for extracting compound fault features of gearbox using non-convex regularized dual dictionary

1.1 齿轮箱信号稀疏表示

针对齿轮箱中的复合振动信号,可以建立如下模型

$$y(t) = y_1^o(t) + y_2^*(t) + n(t) \quad (1)$$

式中: $y_1^o(t)$ 表示轴承故障成分; $y_2^*(t)$ 表示齿轮故障成分; $n(t)$ 代表噪声。

通过求解目标函数的最优稀疏系数,可以得到

$$(\mathbf{a}_1^{\text{opt}}, \mathbf{a}_2^{\text{opt}}) = \arg \min \frac{1}{2} \|\mathbf{y} - \mathbf{A}_1^{\text{opt}} \mathbf{a}_1^{\text{opt}} - \mathbf{A}_2^{\text{opt}} \mathbf{a}_2^{\text{opt}}\|_2^2 +$$

$$\lambda_1 \varphi(\mathbf{a}_1^{\text{opt}}) + \lambda_2 \varphi(\mathbf{a}_2^{\text{opt}}) \quad (2)$$

式中: $\mathbf{A}_1^{\text{opt}}$ 和 $\mathbf{A}_2^{\text{opt}}$ 分别是针对轴承和齿轮故障优化后的字典矩阵; λ_1 和 λ_2 是相应的正则化惩罚参数; $\mathbf{a}_1^{\text{opt}}$ 和 $\mathbf{a}_2^{\text{opt}}$ 分别表示轴承故障和齿轮故障最优稀疏系数; $\varphi(\cdot)$ 表示稀疏正则化函数。

利用求得的最优稀疏系数 $\mathbf{a}_1^{\text{opt}}$ 和 $\mathbf{a}_2^{\text{opt}}$, 可得轴承

和齿轮原始信号表示为

$$\hat{\mathbf{y}}_1^o = \mathbf{A}_1^{\text{opt}} \mathbf{a}_1^{\text{opt}}; \hat{\mathbf{y}}_2^* = \mathbf{A}_2^{\text{opt}} \mathbf{a}_2^{\text{opt}} \quad (3)$$

1.2 齿轮箱复合故障特征提取的字典构建

轴承局部缺陷引发的周期性冲击呈单边指数衰减振荡特点^[21],而调幅调频齿轮振动信号呈现周期性冲击呈双边指数衰减振荡特点^[4]。根据小波内积变换原理,分别采用 Laplace 小波函数 $A_L(t)$ 和 Morlet 小波函数 $A_M(t)$ 匹配轴承故障冲击特征和齿轮故障特征字典,其数学表达为

$$A_L(t) = \exp \frac{-\zeta_1}{\sqrt{1-\zeta_1^2}} 2\pi f_{d1}(t - \tau_{d1}) \cdot \sin(2\pi f_{d1}(t - \tau_{d1})) \quad (4)$$

$$A_M(t) = \exp \frac{-\zeta_2}{\sqrt{1-\zeta_2^2}} [2\pi f_{d2}(t - \tau_{d2})]^2 \cdot \cos(2\pi f_{d2}(t - \tau_{d2})) \quad (5)$$

式中: ζ_i 为阻尼因子; f_{di} 为振荡频率; τ_{di} 为时移因子 ($i = 1, 2$, 分别对应轴承和齿轮),固有振荡频率和相对阻尼因子影响着原子波形。这些参数通常会采用相关滤波^[22]方法从振动信号中识别。文献[23]中提出了基于SVD理论确定Laplace字典参数的方法,受其启发,本文同样采用同样思路,分步骤确定Morlet字典参数和Laplace字典。

(1)对振动信号进行奇异值分解后有矩阵 $\mathbf{U}_z \mathbf{S}_z \mathbf{V}_z^T$ 。构造矩阵 $\mathbf{F}_z$

$$\mathbf{F}_z = (\mathbf{U}_{z/2} \mathbf{S}_{z/2}^{1/2})^\dagger (\mathbf{U}'_{z/2} \mathbf{S}'_{z/2}) \quad (6)$$

式中: $\mathbf{U}_{z/2}$ 和 $\mathbf{U}'_{z/2}$ 分别为 $\mathbf{U}_z$ 的前 $z/2$ 行和后 $z/2$ 行; $\mathbf{S}_{z/2}^{1/2}$ 表示奇异值矩阵的平方根; $\dagger$ 表示伪逆运算。进而得出矩阵 $\mathbf{F}_z$ 的特征值 $\lambda_{z1} = (\lambda_1, \lambda_2, \dots, \lambda_z)$,通过这些特征值确定Morlet字典中参数组合 ζ_1 和 f_{d1} 分别为

$$\zeta_1 = \sqrt{1 + \frac{1}{\text{Re}[\ln(\lambda_{z1})/\Delta t]}} \quad (7)$$

$$f_{d1} = \frac{1}{2\pi} \text{Im} \left[\frac{\ln(\lambda_{z1})}{\Delta t} \right] \quad (8)$$

式中: Re 和 Im 分别表示取复数的实部和虚部; Δt 为采样间隔。

(2)通过参数组合 ζ_1 、 f_{d1} 和 $\tau_{d1} \in [0: \Delta\tau_{d1}: T_C]$ 下的 Morlet 字典函数集 $A_M(t)$ 与原信号 $y_{\text{COM}}(t)$ 进行相关系数最大值索引,保留响应相对阻尼因子 ζ_1 、振荡频率 f_{d1} 和时移因子 τ_{d1} ,构造出匹配于齿轮特征的 Morlet 字典 $A_M(t)$

$$\lambda_M = \frac{\langle y_{\text{COM}}(t), A_M(t) \rangle}{\|y_{\text{COM}}(t)\|_2 \|A_M(t)\|_2} \quad (9)$$

(3)采用OMP算法对字典 $A_M(t)$ 进行计算,获得重构信号、残差信号 y_{re} ,同流程(1)和(2),从残差

信号 y_{re} 中获得匹配轴承特征的 Laplace 字典 $A_L(t)$ 。

1.3 极小极大凹惩罚函数

惩罚函数的设计直接影响齿轮箱复合故障特征提取中稀疏表示系数的稀疏性和信号特征的保真度。传统的 l_1 范数正则化虽然能有效诱导稀疏性, 但会系统性低估故障特征幅值。广义的极小极大凹 (GMC) 惩罚函数通过保持目标函数凸性的同时, 避免了周期性故障脉冲幅值的低估问题。

定义尺度参数为 b (b 为压缩尺度参数) 的极小极大凹惩罚函数为

$$\varphi_G(\alpha) = |\alpha| - h(\alpha) \quad (10)$$

$$h(\alpha) = \begin{cases} |\alpha| - \frac{1}{2}b\alpha^2, & |\alpha| \leq \frac{1}{b} \\ \frac{1}{2b}, & |\alpha| > \frac{1}{b} \end{cases} \quad (11)$$

式中: α 为信号在某个字典原子上的投影权重。

引入辅助变量 $\mathbf{v}$, 构建广义 GMC 惩罚函数的最终表达式

$$G_B(\alpha) =$$

$$\|\alpha\|_1 - \min_{\mathbf{v}} \left\{ \|\mathbf{v}\|_1 + \frac{1}{2} \|\mathbf{B}(\alpha - \mathbf{v})\|_2^2 \right\} \quad (12)$$

式中: $\mathbf{B}$ 为线性变换矩阵; $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 分别表示 l_1 范数和 l_2 范数。

由式(2)加入 GMC 惩罚函数后可得目标函数为

$$F(\alpha_i^{\text{opt}}, \mathbf{v}_i) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}_i \alpha_i^{\text{opt}}\|_2^2 + \lambda G_B(\alpha_i^{\text{opt}}) \quad (13)$$

将式(13)可进行如下变化

$$\begin{aligned} F(\alpha_i^{\text{opt}}, \mathbf{v}_i) &= \max_{\mathbf{v}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A}_i \alpha_i^{\text{opt}}\|_2^2 + \right. \\ &\quad \left. \lambda_i \|\alpha_i^{\text{opt}}\|_1 - \lambda_i \|\mathbf{v}_i\|_1 - \frac{\lambda_i}{2} \|\mathbf{B}(\alpha - \mathbf{v}_i)\|_2^2 \right\} \\ &= \frac{1}{2} (\alpha_i^{\text{opt}})^T (\mathbf{A}_i^T \mathbf{A}_i - \lambda_i \mathbf{B}^T \mathbf{B}) \alpha_i^{\text{opt}} + \\ &\quad \lambda_i \|\alpha_i^{\text{opt}}\|_1 + \max_{\mathbf{v}_i} \{G(\alpha_i, \mathbf{v}_i)\} \end{aligned} \quad (14)$$

进而简化为

$$\begin{aligned} G(\alpha_i^{\text{opt}}, \mathbf{v}_i) &= (\lambda_i \mathbf{v}_i^T \mathbf{B}^T \mathbf{B} - \mathbf{y}^T \mathbf{A}_i) \alpha_i^{\text{opt}} - \\ &\quad \frac{1}{2} \mathbf{v}_i^T \mathbf{B}^T \mathbf{B} \mathbf{v}_i - \lambda_i \|\mathbf{v}_i\|_1 + \frac{1}{2} \mathbf{y}^T \mathbf{y} \end{aligned} \quad (15)$$

式中: T 表示矩阵的转置。

根(14)和式(15)中可知, 为获得目标函数 $F(\alpha_i^{\text{opt}}, \mathbf{v}_i)$ 的最优稀疏系数解 α_i^{opt} , 即需获得 $F(\alpha_i^{\text{opt}}, \mathbf{v}_i)$ 、 $G(\alpha_i^{\text{opt}}, \mathbf{v}_i)$ 为凸函数时, α_i^{opt} 点向的极大值。由于

式(15)中 $\|\alpha_i^{\text{opt}}\|_1$ 作为 l_1 范数是凸函数, 因而为了使目标函数为凸函数, 矩阵 $\mathbf{A}_i^T \mathbf{A}_i - \lambda_i \mathbf{B}^T \mathbf{B}$ 必须是半正定, 其中, $\mathbf{A}_i^T \mathbf{A}_i$ 是字典矩阵, 满足

$$\mathbf{A}_i^T \mathbf{A}_i - \lambda_i \mathbf{B}^T \mathbf{B} \geq 0 \quad (16)$$

式中, 应存在 $0 \leq \gamma_i \leq 1$ 来控制函数凸性使得 $\mathbf{B} = \sqrt{\gamma_i / \lambda_i} \mathbf{A}_i$ 。故式(13)的目标函数可进一步变化为

$$F(\alpha_i^{\text{opt}}, \mathbf{v}_i) =$$

$$\arg \min_{\alpha_i^{\text{opt}}} \max_{\mathbf{v}_i} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A}_i \alpha_i^{\text{opt}}\|_2^2 + \lambda_i \|\alpha_i^{\text{opt}}\|_1 - \right. \\ \left. \lambda_i \|\mathbf{v}_i\|_1 - \frac{\gamma_i}{2} \|\mathbf{A}_i(\alpha_i^{\text{opt}} - \mathbf{v}_i)\|_2^2 \right\} \quad (17)$$

这里采用 FBS 算法^[24]对式(17)进行凸优化求解, 即在满足式(16)的情况下, 通过该算法求出目标函数的全局最小值, 这里将式(17)分裂为两项

$$F(\alpha_i^{\text{opt}}, \mathbf{v}_i^{\text{opt}}) = f(\alpha_i, \mathbf{v}_i) + g(\alpha_i, \mathbf{v}_i) \quad (18)$$

$$\begin{aligned} f(\alpha_i, \mathbf{v}_i) &= \frac{1}{2} \|\mathbf{y} - \mathbf{A}_i \alpha_i^{\text{opt}}\|_2^2 - \lambda_i \|\mathbf{v}_i\|_1 - \\ &\quad \frac{\gamma_i}{2} \|\mathbf{A}_i(\alpha_i^{\text{opt}} - \mathbf{v}_i)\|_2^2 \end{aligned} \quad (19)$$

$$g(\alpha_i, \mathbf{v}_i) = \lambda_i \|\alpha_i^{\text{opt}}\|_1 \quad (20)$$

FBS 算法对式(19)和(20)通过梯度下降和计算近似算子求出, 即

$$\mathbf{z}^{(k)} = \alpha_i^{(k)} - \mu_{\text{step}} [\nabla f(\alpha_i, \mathbf{v}_i)] \quad (21)$$

$$\begin{aligned} \alpha_i^{(k+1)} &= \min_{\alpha_i} \left\{ \frac{1}{2} \|\mathbf{z}^{(k)} - \alpha_i\|_2^2 + \mu_{\text{step}} g(\alpha_i, \mathbf{v}_i) \right\} = \\ &\quad \min_{\alpha_i} \left\{ \frac{1}{2} \|\mathbf{z}^{(k)} - \alpha_i\|_2^2 + \mu_{\text{step}} \lambda_i \|\alpha_i\|_1 \right\} \end{aligned} \quad (22)$$

式中: μ_{step} 为迭代步长, 用于控制每次迭代的更新幅度, 以确保算法的收敛性和稳定性。通过矩阵计算即可求解 α_i^{opt} 。

首先, 定义软阈值算子为

$$\text{soft}(\chi, \rho) = \chi \frac{\max\{|\chi| - \rho, 0\}}{\max\{|\chi|, \rho\}} \quad (23)$$

式中: ρ 是阈值参数, 用于控制稀疏性。

然后, 根据 FBS 算法更新变量 $\mathbf{z}^{(k)}$ 和 $\alpha_i^{(k+1)}$

$$\mathbf{z}^{(k)} = \alpha_i^{(k)} - \mu_{\text{step}} \mathbf{A}_i^T (\mathbf{A}_i (\alpha_i^{(k)} + \gamma_i (\mathbf{v}_i^{(k)} - \alpha_i^{(k)})) - \mathbf{y}) \quad (24)$$

$$\alpha_i^{(k+1)} = \text{soft}(\mathbf{z}^{(k)}, \lambda_i \mu_{\text{step}}) \quad (25)$$

同理, 通过 FBS 算法也可以求出 $\mathbf{v}_i^{\text{opt}}$

$$\tilde{\omega}^{(k)} = \mathbf{v}_i^{(k)} - \mu_{\text{step}} \gamma_i \mathbf{A}_i^T \mathbf{A}_i (\mathbf{v}_i^{(k)} - \alpha_i^{(k)}) \quad (26)$$

$$\mathbf{v}_i^{(k+1)} = \text{soft}(\tilde{\omega}^{(k)}, \mu_{\text{step}} \lambda_i) \quad (27)$$

式中: γ_i 是控制函数凸性的参数 (控制函数凸性的参数), 用于确保目标函数的凸性。

为了比较各类惩罚函数的差异, 将 GMC 惩罚函数与对数惩罚函数、反正切和指数惩罚函数、平滑

剪切绝对偏差惩罚函数、最小最大凹惩罚范数进行对比。在图 2 中, l_0 范数的正则化是理想的惩罚函数,但该惩罚函数约束后的目标函数高度非凸,无法输出全局最优结果,因而引入的惩罚函数都是为了逼近 l_0 范数的正则化。从图 2 各惩罚函数图像可见, l_1 范数的正则项会明显偏离于 l_0 范数,这便造成了 l_1 正则化低估了稀疏系数,重构精度因此降低。其他保证目标函数凸性的惩罚函数如对数惩罚函数、反正切、平滑剪切绝对偏差惩罚函数、最小最大凹惩罚范数和指数惩罚函数,尽管相对于 l_1 范数正则化能够更接近理想的 l_0 范数,但它们仍然存在一定的偏差,而 GMC 函数能完全理想地逼近 l_0 范数。

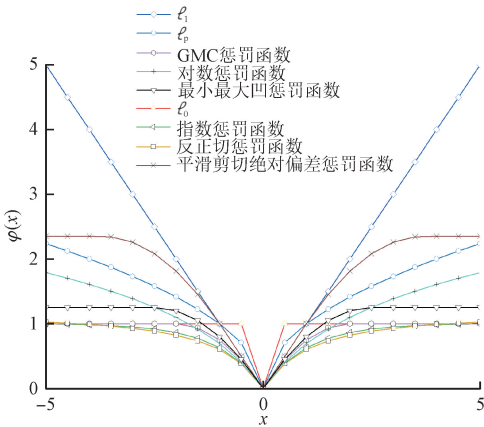


图 2 各惩罚函数图像

Fig. 2 Graphs of various penalty functions

2 仿真信号分析

为了验证本文方法对复合故障特征提取的有效性,构建了一个时间长度为 1 s、采样频率为 12 kHz 的齿轮箱复合故障仿真信号^[25]。信号由 3 部分构成:齿轮啮合信号、齿轮故障信号和滚动轴承故障信号,其表达式为

$$x_{gb}(t) = x_{gm}(t) + x_{gf}(t) + x_{bingf}(t) \tag{28}$$

$$\begin{aligned} x_{gm}(t) = & [1 + A_G \cos(2\pi f_{r,dng}t)] \cos[2\pi f_m t + \\ & B_G \sin(2\pi f_{r,den}t)] + [1 + A_G \cos(2\pi f_{r,den}t)] \cdot \\ & \cos[2\pi f_m t + B_G \sin(2\pi f_{r,dng}t)] + \\ & \sum_{n=1}^N A_{sh} \exp[-\beta_m(t - nT_m)] \cdot \\ & \sin[\varphi_m(t - nT_m)] u(t - nT_m) \end{aligned} \tag{29}$$

$$\begin{aligned} x_{gf}(t) = & [1 + A_{GF} \cos(2\pi f_{GF}t)] \cos(2\pi f_m t + \\ & B_{GF} \sin(2\pi f_{GF}t)) + \sum_{m=1}^M [A_{m,gf} \exp(-\beta_{GF}(t - nT_{GF})) \cdot \\ & \sin(2\pi \omega_{GF}(t - nT_{GF})) u(t - nT_{GF})] \end{aligned} \tag{30}$$

$$\begin{aligned} x_{bingf}(t) = & \sum_{i=1}^I [A_{BF} \exp(-\beta_{BF}(t - i/f_{BF} - \tau)) \cdot \\ & \sin(2\pi \omega_{BF}(t - i/f_{BF} - \tau))] \end{aligned} \tag{31}$$

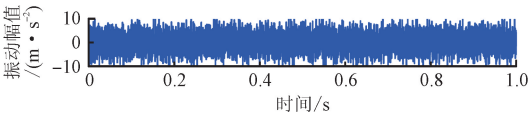
式(29)中, $x_{gm}(t)$ 为齿轮啮合信号,前两项分别描述了主动轮齿轮和从动轮齿轮的调幅调频现象,最后一项则为齿轮啮合产生的脉冲。式(30)表示了齿轮局部故障信号 $x_{gf}(t)$, 第一项表示齿轮故障引起的调幅调频现象,第二项为故障齿轮的脉冲。式(31)表示了轴承故障的故障脉冲仿真信号。

仿真信号模型参数见表 1。为了模拟现场实测环境,在仿真信号中加入 -10 dB 的高斯白噪声。图 3 为仿真信号的时域波形、频谱和包络谱,由于故障间的相互耦合和噪声干扰作用,仅能在包络谱中显示齿轮故障特征频率 f_{GF} 。

表 1 仿真信号模型参数

Table 1 Simulation signal model parameters

变量	数值
啮合齿轮的调制信号幅值 $A_G / (\text{m} \cdot \text{s}^{-2})$	0.2
主动齿轮转动频率 $f_{r,dng} / \text{Hz}$	16.67
齿轮啮合频率 f_m / Hz	367
齿轮啮合处调频 B_G / Hz	0.3
从动齿轮的转动频率 $f_{r,den} / \text{Hz}$	6.67
齿轮啮合引起的脉冲振幅 $A_{sh} / (\text{m} \cdot \text{s}^{-2})$	0.1
齿轮啮合结构阻尼频率 β_m / Hz	1 000
齿轮啮合引起共振频率 φ_m / Hz	2 000
齿轮故障分量调制信号幅值 $A_{GF} / (\text{m} \cdot \text{s}^{-2})$	0.1
故障齿轮的转动频率 f_{GF} / Hz	16.67
齿轮故障分量的调频 B_{GF} / Hz	0.3
齿轮故障引起的脉冲振幅 $A_{m,gf} / (\text{m} \cdot \text{s}^{-2})$	7
故障齿轮的阻尼特征频率 β_{GF} / Hz	1 000
故障齿轮引起的共振频率 ω_{GF} / Hz	5 000
轴承故障引起的脉冲振幅 $A_{BF} / (\text{m} \cdot \text{s}^{-2})$	5
轴承的故障特征频率 f_{BF} / Hz	60
轴承滚动体的随机滑动距离 τ_{BF}	$0.01/f_{bf} \sim 0.02f_{bf}$
轴承故障引起的共振频率 ω_{BF} / Hz	7 000
轴承的结构阻尼特征频率 β_{BF} / Hz	1 200



(a) 时域信号

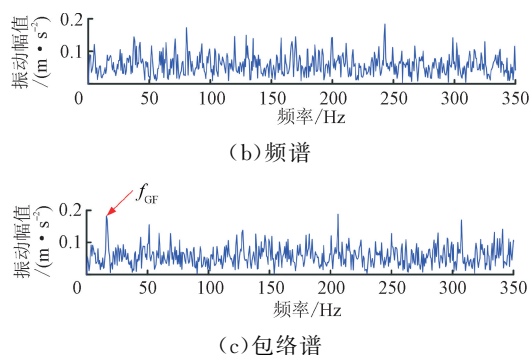


图3 加噪声复合故障仿真信号

Fig. 3 Simulated composite fault signal with added noise

利用 MOMEDA 方法对仿真信号进行分析。图 4 中 MOMEDA 方法的窗函数值为 100, 滤波器尺寸为 1 000 个采样点, 信号长度和理论故障周期值 $L_{s3}/f_{GF} = 616$ 个采样点, 确定故障区间范围为 $[400 : 1 : 900]$ 个采样点。图 4 表明, 尽管能够提取出 f_{GF} , 但故障特征频率周边也存在较多的干扰频率。在图 5 中 MOMEDA 方法的窗函数值为 80 个采样点, 滤波器尺寸为 720 个采样点, 信号长度和理论故障周期为 $L_{s3}/f_{BF} = 171$ 个采样点, 确定故障区间范围为 $[100 : 1 : 350]$ 个采样点。图 5(b) 的包络谱分析有效地提取了轴承的故障特征频率 f_{BF} 及其倍频成分。综合图 4 和图 5 的信号分析结果, 从它们的时域信号中不难看出, MOMEDA 方法存在边缘效应, 未能准确还原原始信号长度, 因而同样存在信息丢失的问题; MOMEDA 方法虽然可以每次设置不同参数提取出复合故障信号中的每个特征, 但因为故障周期范围中同样存在噪声干扰, 所以在提取的故障特征频率周围往往会存在噪声频率。

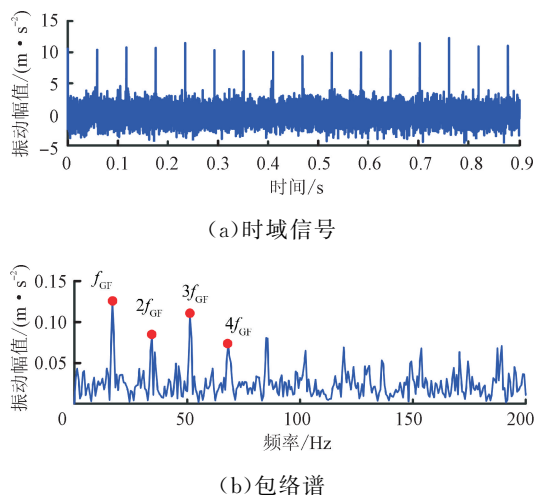


图4 MOMEDA 方法齿轮箱仿真信号齿轮故障特征

Fig. 4 Gearbox fault feature extraction results from MOMEDA on simulated signals

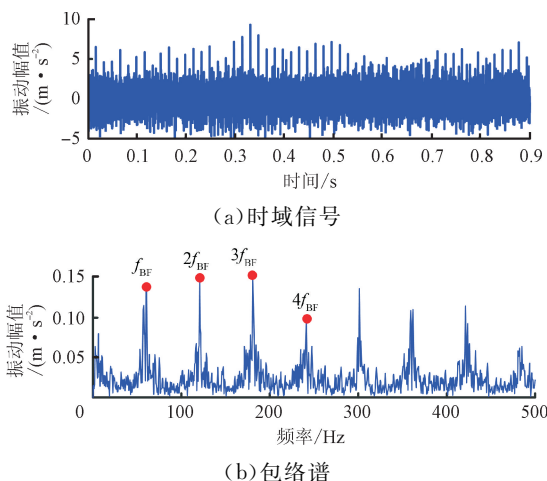


图5 MOMEDA 方法齿轮箱仿真信号轴承故障特征

Fig. 5 Results of MOMEDA for extracting bearing fault features from gearbox simulation signals

IESFOgram 方法对齿轮箱复合故障的分析结果如图 6 所示, 图 6(a) 是基于二叉树循环谱相关 (CSC) 优化的改进版增强包络 (EES) 分析结果, 用于提取 f_{GF} , 图中可观察到齿轮故障特征的 2 倍频与 3 倍频成分, 但幅值较弱, 特征显著性不足。图 6(b) 为从 IESFOgram 方法中提取的瞬时能量谱 (IES)。可以从 IES 中获取 f_{BF} 及其 2 倍频、3 倍频成分, 但同样存在特征提取不明显的局限性。

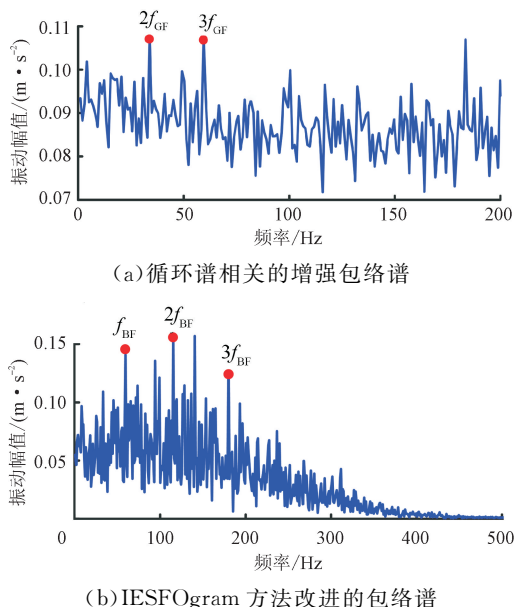


图6 EES 方法和 IESFOgram 方法的齿轮箱仿真信号谱图

Fig. 6 Spectrum of gearbox simulation signals from EES and IESFOgram

利用本文提出的方法进行故障特征提取, 通过构建 Morlet 小波和 Laplace 小波字典, 有效匹配齿

轮和轴承的故障特征,并通过引入 GMC 惩罚函数,优化稀疏分解过程,从而在不牺牲凸性的前提下改善了基于 l_1 范数正则化的幅值低估问题。双小波字典方法提取仿真齿轮、轴承故障特征的结果如图 7 和图 8 所示。可以看出:重构信号的包络谱准确地分离出了齿轮故障频率 f_{GF} 和轴承故障频率 f_{BF} 及其倍频;与 MOMEDA 和 IESFOgram 方法相比,本文方法故障特征频率的幅度更大,噪声干扰更少,说明本文方法在有效提取故障特征的同时提高了信噪比,其在识别齿轮箱内的复合故障特征方面表现出更高的准确性。

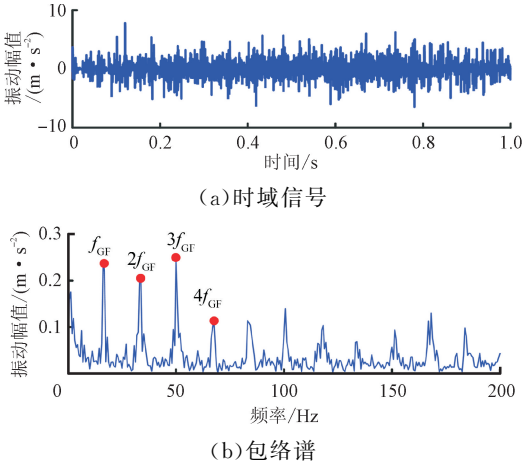


图 7 双小波字典方法提取仿真齿轮故障特征
Fig. 7 Simulation gear fault feature extraction using dual wavelet dictionary method

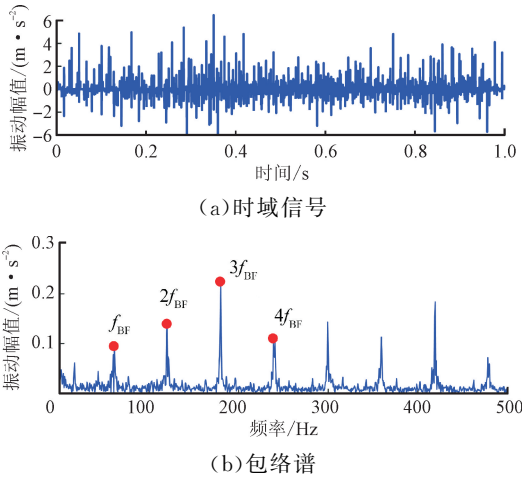


图 8 双小波字典方法提取仿真轴承故障特征
Fig. 8 Simulation bearing fault feature extraction using dual wavelet dictionary method

为了评估各分解方法在提取齿轮和轴承故障特征信息方面的效能,本研究采用文献[26]中的信噪比计算公式

$$S_n = 10 \lg \frac{P_{\text{signal}}}{P_{\text{noise}}} = 10 \lg \frac{\sum_{i=1}^n P[i f_{ds}]}{\sum_{l=\text{round}(0, 5 f_{ds})} P[l] - \sum_{i=1}^n P[i f_{ds}]} \quad (32)$$
$$f_{ds} = \text{round}\left(\frac{f}{\Delta f}\right) + 1 \quad (33)$$

式中: $P[\cdot]$ 表示用于计算 SNR 分量的包络谱幅度; f 是轴承的理论故障特征频率; Δf 是频率分辨率; f_{ds} 为包络谱幅值序列中 f 的阶次; n 为第 n 个故障, 本文取 n 的值为 5, 即计算前 5 个故障特征频率的信噪比。该公式在计算过程中考虑了不同方法带宽划分差异对信号噪声特性的影响, 从而使得所得信噪比更为准确和可靠。通过这一量化指标, 可以更客观地评估各方法在提取故障特征信息方面的性能。

3 种方法分解的齿轮和轴承前 5 个故障特征频率的信噪比如图 9 所示。可以看出, 本文双小波字典方法所得到的信噪比最高, 本文方法在包络谱中可以得到明显的故障特征频率, 且所占有的能量较高。其次是 MOMEDA 方法, 能较为完整地提取出故障特征频率, 但重构能量有损失。对于 EES 和 IESFOgram 方法未完整地提取出故障特征频率, 且信噪比最低。

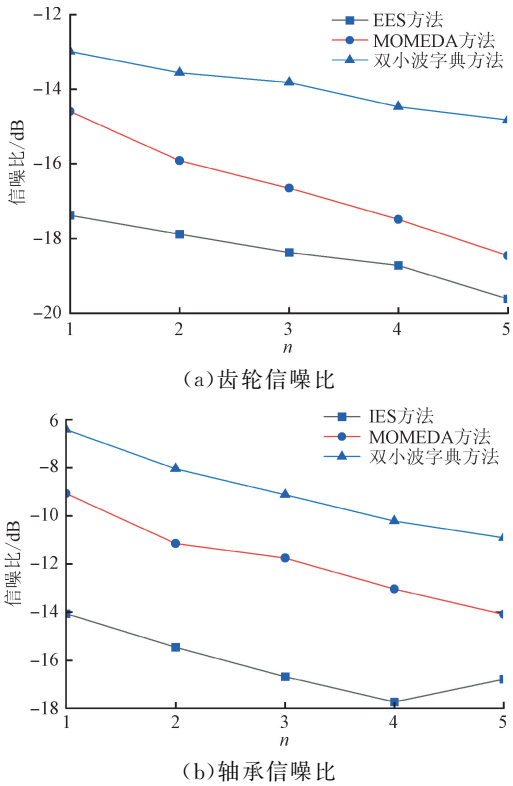


图 9 3 种方法得到的仿真信号每个分量的信噪比
Fig. 9 SNR of each component of the simulated signal obtained by three methods

3 实测信号分析

为了进一步验证该方法对于实测信号处理的有效性,将其应用于2009年PHM挑战赛数据^[27]分析。数据文件为“helical_5_45hz_High_2.txt”,采频为200 kHz,输入端转频率为 $f_1=45$ Hz。在齿轮箱的Ⅱ轴上的齿发生了断齿故障,同时此轴靠近输出端的轴承存在内圈故障。根据减速器齿轮箱结构的传动比可知Ⅱ轴转动频率 $f_2=1/3f_1$,由此可推算出齿轮故障频率 $f_{GF1}=15$ Hz,轴承内圈故障特征频率为 $f_{BF1}=4.95f_2=74.25$ Hz,Ⅱ轴和Ⅲ轴齿轮啮合频率 $f_{GM2}=8f_1=360$ Hz。读取长度为0.5 s的数据,其时域波形、频谱和包络谱如图10所示。从图10(b)中可观察到输入轴转频 f_1 以及Ⅱ轴齿轮和Ⅲ轴齿轮啮合处的啮合频率 f_{GM2} 。从图10(c)只能观察到 f_1 及其倍频成分,未能有效提取齿轮箱中的故障特征,进而同仿真信号一样,通过不同方法比较分析以提取出故障特征。

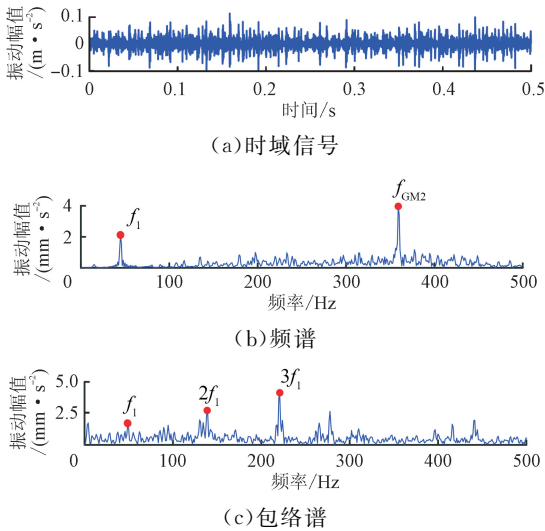


图10 实测信号时域波形、频谱及其包络谱
Fig. 10 Simulation signal time-domain waveform, spectrum, and envelope spectrum

MOMEDA方法提取的齿轮、轴承故障特征结果如图11、图12所示。图11中MOMEDA方法的窗函数值为70个采样点,滤波器尺寸为2300个采样点,信号长度和理论故障周期 $L/f_{GF1}=4474.27$ 个采样点,设置故障区间范围为 $[4400:1:6000]$ 个采样点。从图11(b)的包络谱可有效识别出齿轮故障特征频率的部分倍频成分。图12中MOMEDA方法的窗函数值为100个采样点,滤波器尺寸为3700个采样点,信号长度和理论故障周期 $L/f_{BF1}=913.24$ 个采样点,确定故障区间范围为 $[850:1:1200]$ 个采样

点。图12(b)的包络谱中看似提取出了轴承内圈故障的故障特征频率 $f'_{BF1}=79.34$ Hz及其倍频成分,而根据实际转速推导出故障特征频率 $f_{BF1}=73.82$ Hz,可见由MOMEDA方法提取出的故障特征频率及其倍频实际与理论值之间存在一定偏差。这是由于MOMEDA方法以多点峭度的最大值作为目标函数求得最优滤波器,因而滤波器的更新会受目标矢量的影响,在多点峭度值达到最大值时可能涵盖了多个故障的故障周期,故MOMEDA方法会在多种周期性脉冲同时存在的情况下造成误诊。

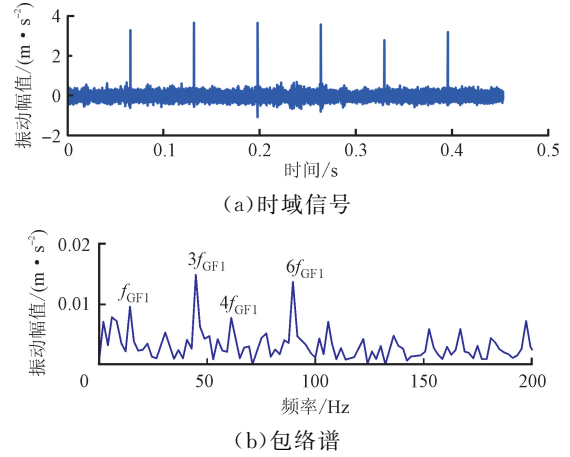


图11 MOMEDA方法提取齿轮故障特征
Fig. 11 Results of gear fault feature extraction using MOMEDA

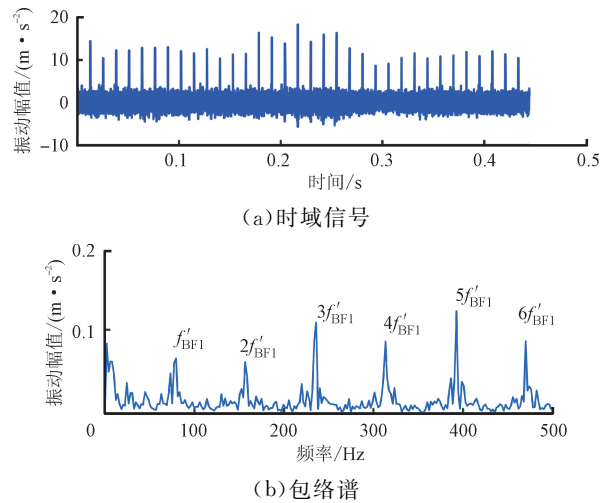
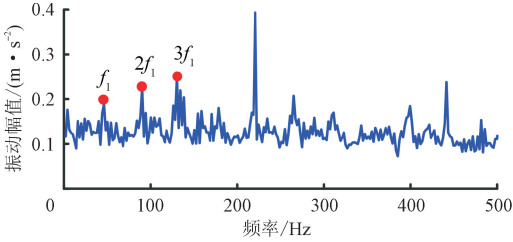


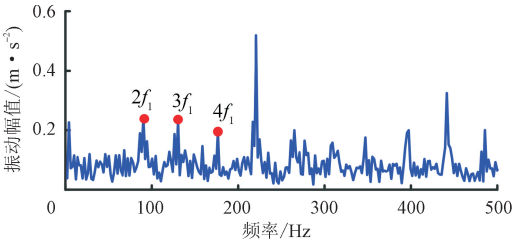
图12 MOMEDA方法提取轴承故障特征
Fig. 12 Results of bearing fault feature extraction using MOMEDA

IESFOgram方法对齿轮箱复合故障的分析结果如图13所示。图13(b)所示的增强包络谱(EES)中仅能识别出输入轴转频 f_1 及其倍频成分,而图13(d)的瞬时能量谱同样主要呈现 f_1 的倍频特征。分析表明,这是因为IESFOgram方法考虑全谱频带对多个

共振频率周围的故障分量进行积分,这其中存在小权重的故障无关量会影响故障分量的识别。在这里进行分析的数据,输入端的转动频率为故障无关量,最后影响了 IESFOgram 方法提取故障特征。



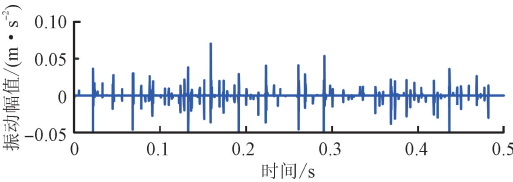
(a)EES 的增强包络谱



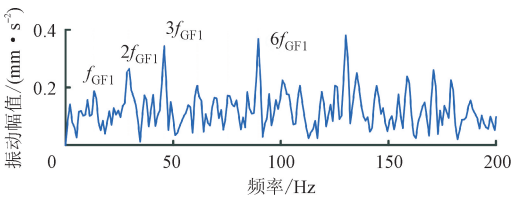
(b)IESFOgram 改进的包络谱

图 13 EES 和 IESFOgram 的齿轮箱实测信号谱图
Fig. 13 Spectrum of gearbox measured signals from EES and IESFOgram

应用本文提出方法对齿轮箱的复合故障信号进行分析,结果如图 14、图 15 所示。与 MOMEDA 方法相比,本文方法也有效地提取了齿轮故障,而轴承故障特征频率更为准确,并且在包络谱中的噪声更少。与 IESFOgram 方法相比,本文方法在识别齿轮箱内的故障特征方面表现出更高的准确性:它不仅有效识别了齿轮故障的特征频率 f_{GF1} 及其倍频,还获取了轴承内圈故障 f_{BF1} 及其倍频成分。

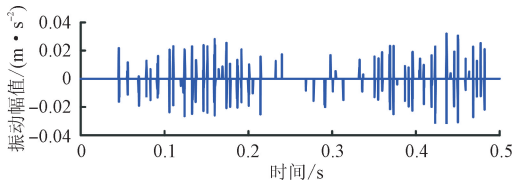


(a)时域信号

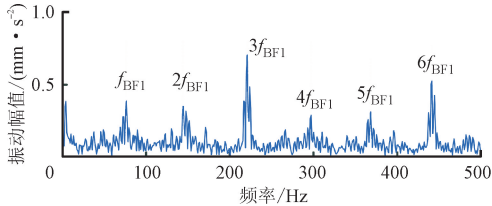


(b)包络谱

图 14 双小波字典方法提取齿轮故障特征结果
Fig. 14 Simulation gear fault feature extraction using dual wavelet dictionary method



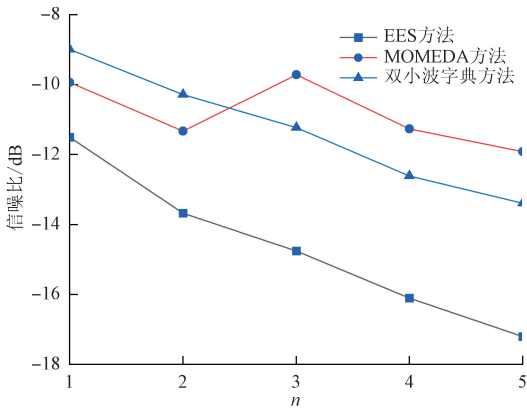
(a)时域信号



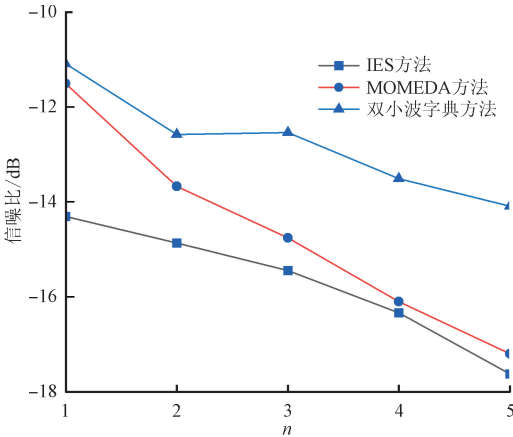
(b)包络谱

图 15 双小波字典方法提取轴承故障特征结果
Fig. 15 Simulation bearing fault feature extraction using dual wavelet dictionary method

为了客观评估分解效果,计算了 3 种方法分解的齿轮和轴承前 5 个故障特征频率的信噪比,结果如图 16 所示。



(a)齿轮信噪比



(b)轴承信噪比

图 16 3 种方法得到的实测信号每个分量的信噪比
Fig. 16 SNR of each component of the measured signal obtained by three methods

由图 16(a)可见,本文方法前 2 倍频的故障特征频率的信噪比最优,3~5 倍频的故障特征频率的信噪比最优的是 MOMEDA 方法。EES 中仅能提取出输入轴转频 f_1 及其倍频信噪比很低。由图 16(b)可见,本文方法得到的信噪比都是最优的,这是由于 MOMEDA 方法所提取的故障特征频率有偏移,导致信噪比较低,而 IES 仅能提取出输入轴转频 f_1 及其倍频,其信噪比也较低。

4 结 论

针对于齿轮箱齿轮和轴承复合故障,本文提出了一种的双小波稀疏字典诊断方法,通过对仿真和实测信号分析,得到如下结论。

(1)利用 SVD 的原理,从振动信号中提取 Laplace 小波和 Morlet 小波参数,构建了齿轮和轴承故障的字典,有效匹配齿轮箱中发生的故障特征。

(2)GMC 惩罚函数在保持目标函数凸性的同时,有效解决了基于 l_1 范数惩罚函数可能导致的重构信号幅值低估问题。

(3)所提方法在前 5 倍故障特征频率的信噪比提升了 1~2 dB,显著优于 MOMEDA 及 IESFOgram 方法,验证了其在复合故障特征提取中的优越性。

参考文献:

- [1] 朱沁玥,何海昊,李锋,等. 基于短时傅里叶变换和深度卷积神经网络的直升机齿轮箱故障诊断方法[J]. 失效分析与预防, 2022, 17(1): 1-8.
ZHU Qinyue, HE Haihao, LI Feng, et al. A fault diagnosis method of helicopter gearbox based on short-time Fourier transform and deep convolutional neural network [J]. Failure Analysis and Prevention, 2022, 17(1): 1-8.
- [2] 陈雪峰,郭艳婕,许才彬,等. 风电装备故障诊断与健康监测研究综述[J]. 中国机械工程, 2020, 31(2): 175-189.
CHEN Xuefeng, GUO Yanjie, XU Caibin, et al. Review of fault diagnosis and health monitoring for wind power equipment [J]. China Mechanical Engineering, 2020, 31(2): 175-189.
- [3] GOSWAMI P, NANDAN RAI R. A systematic review on failure modes and proposed methodology to artificially seed faults for promoting PHM studies in laboratory environment for an industrial gearbox [J]. Engineering Failure Analysis, 2023, 146: 107076.
- [4] 李蓉. 齿轮箱复合故障诊断方法研究[D]. 长沙: 湖南大学, 2013.
- [5] ANTONI J, RANDALL R B. The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines [J]. Mechanical Systems and Signal Processing, 2006, 20(2): 308-331.
- [6] ANTONI J. Fast computation of the kurtogram for the detection of transient faults [J]. Mechanical Systems and Signal Processing, 2007, 21(1): 108-124.
- [7] MAURICIO A, SMITH W A, RANDALL R B, et al. Improved envelope spectrum via feature optimisation-gram (IESFOgram): a novel tool for rolling element bearing diagnostics under non-stationary operating conditions [J]. Mechanical Systems and Signal Processing, 2020, 144: 106891.
- [8] WIGGINS R A. Minimum entropy deconvolution [J]. Geop exploration, 1978, 16(1/2): 21-35.
- [9] 王富珂,高丙朋,蔡鑫. 基于迭代 SGMD 与改进 MOMEDA 的滚动轴承微弱故障诊断[J]. 组合机床与自动化加工技术, 2024(12): 145-150.
WANG Fuke, GAO Bingpeng, CAI Xin. Weak fault diagnosis of rolling bearing based on iterative SGMD and improved MOMEDA [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2024(12): 145-150.
- [10] 林云,郭瑜,陈鑫. 参数优化最大二阶循环平稳盲解卷积行星轮轴承故障提取[J]. 振动与冲击, 2023, 42(2): 321-328.
LIN Yun, GUO Yu, CHEN Xin. Planet bearing fault extraction based on parameter optimized maximum second-order cyclostationary blind deconvolution [J]. Journal of Vibration and Shock, 2023, 42(2): 321-328.
- [11] 孙寅. 齿轮箱中轴承与齿轮复合故障的分离与特征提取方法研究[D]. 北京: 北京工业大学, 2021.
- [12] WANG Shibin, HUANG Weiguo, ZHU Z K. Transient modeling and parameter identification based on wavelet and correlation filtering for rotating machine fault diagnosis [J]. Mechanical Systems and Signal Processing, 2011, 25(4): 1299-1320.
- [13] 余发军,周凤星. 基于可调 Q 因子小波变换和谱峭度的轴承早期故障诊断方法[J]. 中南大学学报(自然科学版), 2015, 46(11): 4122-4128.
YU Fajun, ZHOU Fengxing. Bearing early faults diagnosis based on tunable Q-factor wavelet transform and spectral kurtosis [J]. Journal of Central South University(Science and Technology), 2015, 46(11): 4122-4128.
- [14] 李继猛,李铭,姚希峰,等. 基于集合经验模式分解

- 和 K-奇异值分解字典学习的滚动轴承故障诊断 [J]. 计量学报, 2020, 41(10): 1260-1266.
- LI Jimeng, LI Ming, YAO Xifeng, et al. Rolling bearing fault diagnosis based on ensemble empirical mode decomposition and K-singular value decomposition dictionary learning [J]. Acta Metrologica Sinica, 2020, 41(10): 1260-1266.
- [15] 邢展, 罗玉珺, 礼莹, 等. 基于目标优化的移不变字典匹配追踪轴承故障诊断算法研究 [J]. 机械, 2024, 51(7): 1-8.
- XING Zhan, LUO Yujun, LI Ying, et al. Research on bearing fault diagnosis algorithm of shift invariant dictionary matching pursuit based on objective optimization [J]. Machinery, 2024, 51(7): 1-8.
- [16] VIDAL R, MA Yi, SASTRY S. Generalized principal component analysis (GPCA) [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(12): 1945-1959.
- [17] YANG Hongyu, MATHEW J, MA Lin. Fault diagnosis of rolling element bearings using basis pursuit [J]. Mechanical Systems and Signal Processing, 2005, 19(2): 341-356.
- [18] LI Yanjun, CHEN Wendong. A correlation coefficient sparsity adaptive matching pursuit algorithm [J]. IEEE Signal Processing Letters, 2023, 30: 190-194.
- [19] SELESNICK I, FARSHCHIAN M. Sparse signal approximation via nonseparable regularization [J]. IEEE Transactions on Signal Processing, 2017, 65(10): 2561-2575.
- [20] HE Wangpeng, CHEN Binqiang, ZENG Nianyin, et al. Sparsity-based signal extraction using dual Q-factors for gearbox fault detection [J]. ISA Transactions, 2018, 79: 147-160.
- [21] 李景乐, 谢馨, 王华庆. 改进 Laplace 小波字典在轴承故障诊断中的应用 [J]. 轴承, 2018(11): 57-60.
- LI Jingle, XIE Xin, WANG Huaqing. Application of improved Laplace wavelet dictionary in fault diagnosis for bearings [J]. Bearing, 2018(11): 57-60.
- [22] HUANG Weiguo, SONG Zeshu, ZHANG Cheng, et al. Multi-source fidelity sparse representation via convex optimization for gearbox compound fault diagnosis [J]. Journal of Sound and Vibration, 2021, 496: 115879.
- [23] 霍永超. 加工中心自动换刀系统典型故障稀疏诊断方法研究 [D]. 长春: 吉林大学, 2021.
- [24] WANG Shibin, SELESNICK I, CAI Gaigai, et al. Nonconvex sparse regularization and convex optimization for bearing fault diagnosis [J]. IEEE Transactions on Industrial Electronics, 2018, 65(9): 7332-7342.
- [25] WANG Tianyang, CHU Fulei, HAN Qinkai, et al. Compound faults detection in gearbox via meshing resonance and spectral kurtosis methods [J]. Journal of Sound and Vibration, 2017, 392: 367-381.
- [26] 郑近德, 潘海洋, 程军圣, 等. 基于自适应经验傅里叶分解的机械故障诊断方法 [J]. 机械工程学报, 2020, 56(9): 125-136.
- ZHENG Jinde, PAN Haiyang, CHENG Junsheng, et al. Adaptive empirical Fourier decomposition based mechanical fault diagnosis method [J]. Journal of Mechanical Engineering, 2020, 56(9): 125-136.
- [27] The Prognostics and Health Management Society. Apparatus [EB/OL]. [2025-02-17]. <https://phmsociety.org/data-analysis-competition/apparatus/>.

(编辑 亢列梅)