

基于多指标优化 TQWT 稀疏表示的轴承故障诊断

强煜栋¹ 万方义¹ 张春林¹ 张昌兴²

(1. 西北工业大学 航空学院, 西安 710072)

(2. 西安电子科技大学 先进材料与纳米科技学院, 西安 710126)

摘要:【目的】针对轴承振动故障特征提取精度偏低、故障信息易被强噪声淹没、信号自适应构建缺乏合理评价指标以及数据量冗余等问题,开展轴承故障特征增强与故障诊断方法研究。提出了一种引入非凸正则项系数的稀疏分解方法,并基于多指标特征矩阵实现信号自适应重构。【方法】首先,通过 tunable Q-factor 小波变换 (Tunable Q-factor Wavelet Transform, TQWT) 对信号进行稀疏分解,获得信号的最简表达,构造非凸惩罚函数以提升重构信号的稀疏性,同时保留关键故障特征信息;其次,综合谐波与噪声能量比、峭度偏斜度比及稀疏度构建融合指标参数矩阵,用于优选可调小波参数与稀疏惩罚项阈值;最后,利用优选参数完成信号去噪与重构,再对重构信号进行解调分析,提取轴承故障特征信息。【结果】仿真信号与试验结果表明,所提方法能够有效提取轴承早期故障冲击特征,实现轴承振动故障的准确诊断。

关键词: 多指标优化; 可调品质因子小波; 稀疏分解; 故障诊断; 信号处理

中图分类号: TH212; TH213.3

DOI: 10.16579/j.issn.1001.9669.2026.05.014

0 引言

滚动轴承广泛应用汽车、航空航天、电力工程、重型机械等各类机械设备领域。滚动轴承主要起着支撑转轴、传递轴向载荷等作用。在大型机械如发动机传动系统的故障中,滚动轴承运行故障占比近30%^[1],准确判断滚动轴承的故障类型并加以维修,可降低维修成本,延长机械的使用寿命,提高生产过程的可靠性与安全性,具有重大的科学意义和工程应用价值^[2]。

当今监测滚动轴承运行状态的普遍方法是使用加速度传感器或振动测量仪监测工况下轴承振动状态。其中使用小波分析对信号进行微观分析逐渐成为研究的热点^[3-4]。SELESNICK^[5]基于离散小波变换 (Discrete Wavelet Transformation, DWT) 结构,提出了一种易于调节的离散时间信号的小波变化,称为可调品质因子小波变换 (Tunable Q-factor Wavelet Transform, TQWT)。这种方法提升了调优小波品质因子的能力,可针对应用信号的部分振荡行为进行选择,提高了设计自定义频分解的灵活性,拥有较为理想的重构能力。唐贵基等^[6]利用 TQWT 分解信号,通过

峭度准则及相关系数准则重构信号,根据重构信号的包络解调谱成功提取出滚动轴承的微弱故障特征。余发军等^[7]则利用谱峭度最大原则,确定 TQWT 的最佳共振因子和尺度带,通过分析重构信号包络谱极值点的频率位置,成功提取轴承故障信息。

目前对旋转机械进行早期的故障状态监测,一般常用的故障特征提取方法易丢失有效信息,且产生的有效故障信号会伴随大量冗余的正常信号,给数据的存储和传输带来了巨大的压力。为了有效缓解冗余数据量和故障早期诊断带来的诊断难题,稀疏分解理论在这种研究热潮下孕育而生。稀疏表示方法的基本思想最早是由 MALLAT^[8]提出的。稀疏表示方法具有低采样率、稳定的重构性能等特点,广泛应用于图像和信号处理等领域。LIU 等^[9-10]研究了稀疏-辅助滤波方法,用于状态监测和机器故障诊断。为了增强稀疏性并克服 L_1 范数正则化带来的不充分性, XU 等^[11-12]提出了非凸稀疏正则化 (Nonconvex Sparse Regularization, NSR) 方法,广泛用于压缩传感和信号处理。然而,大多数 NSR 方法放弃了成本函数的凸性,出现了局部最小化解。SELESNICK 等^[13]研究了损失

收稿日期: 2024-07-08 修回日期: 2024-09-29

基金项目: 国家自然科学基金项目 (11972314)

作者简介: 强煜栋,男,2001年生,陕西渭南人,硕士; 主要研究方向为旋转机械故障诊断、强化学习及路径规划; E-mail: QiangVD@mail.nwpu.edu.cn。

万方义(通信作者),男,1973年生,陕西西安人,博士,副教授; 主要研究方向为飞行器综合测试与健康管理、飞行器结构、机构健康管理及演示试验; E-mail: fwan@nwpu.edu.cn。

引用格式: 强煜栋, 万方义, 张春林, 等. 基于多指标优化 TQWT 稀疏表示的轴承故障诊断[J]. 机械强度, 2026, 48(5): 117-125.

QIANG Yudong, WAN Fangyi, ZHANG Chunlin, et al. Bearing fault diagnosis based on multi-index optimization TQWT sparse representation[J]. Journal of Mechanical Strength, 2026, 48(5): 117-125.



免费获取

函数的凸性,以捕捉任何非正则化和凸优化的优势。WANG 等^[14]提出了一种基于广义极小凹惩罚 (Generalized Minmax-Concave penalty, GMC) 的非凸稀疏正则化轴承故障诊断方法,提供了一个统一的 NSR 和凸优化框架。该方法保持了稀疏正则化最小二乘代价函数的凸性,使用凸优化算法求解全局最小值,并引入了一种稀疏度策略,自适应选择正则化参数。

综上所述,国内外对于滚动轴承故障诊断进行了广泛深入的研究,同时取得了大量成果。然而,稀疏表示的 TQWT 对故障信号进行特征提取的方法在实际应用中,没有解决如何构造 TQWT 的稀疏表达式及自适应选择参数的问题,对于复杂工况下轴承故障特征的提取效果较差。本文针对以上未解决的问题,结合实际工程的需求,提出一种具有非凸正则项系数的稀疏分解方法,并使用多指标特征矩阵进行自适应重构信号的轴承故障诊断方法。构建 TQWT 的稀疏表达式后,通过一个多指标的参数矩阵对小波基函数和稀疏惩罚项参数阈值进行自适应选择,对重构信号进行解调后进行故障诊断,并将该方法分别应用于仿真故障信号和试验信号集中,验证了该方法的可行性。

1 稀疏分解理论模型

1.1 TQWT 基本理论

SELESNICK 等^[15]3560-3575 基于 DWT 结构,提出了一种易于调节的离散时间信号的小波变化,其由品质因子 Q 和冗余因子 r 参数化。该方法提供了改变小波品质因子的能力,针对应用信号的部分振荡行为可进行选择,提高了设计自定义频带分解的灵活性,拥有近似满足完美重构的能力。

TQWT 采用双通道滤波器组对信号进行多层迭代的分解重构,3 层迭代方式分解重构示意图如图 1 所示。

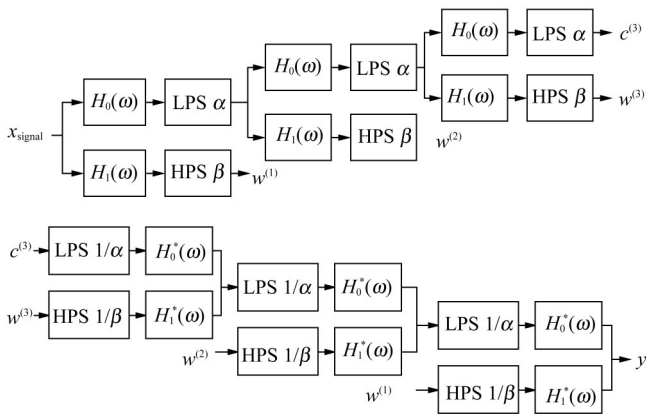


图 1 3 层迭代分解重构示意图

Fig. 1 Schematic diagram of three-layer iterative decomposition and reconstruction

图 1 中, x_{signal} 为输入信号; c^j, w^j 为第 j 次分解得到的尺度系数和小波系数; $H_0(\omega)$ 和 $H_1(\omega)$ 为低通和高通滤波器, $H_0^*(\omega)$ 和 $H_1^*(\omega)$ 为对应的共轭复数; α 和 β 为低通和高通的尺度因子; r 为冗余因子, 对应的定义式分别为

$$\alpha = 1 - \frac{\beta}{r}, \beta = \frac{2}{Q+1} \quad (1)$$

$$Q = \frac{\omega_0}{B_w} \quad (2)$$

式中, B_w 为信号带宽; ω_0 为中心频率。式(1)满足 $0 < \alpha < 1, 0 < \beta \leq 1, \alpha + \beta > 1$ 。

TQWT 的最大分解层数由式(3)决定, 即

$$J_{\max} = \left\lfloor \frac{\lg(\beta/N)}{\lg(1/\alpha)} \right\rfloor = \left\lfloor \frac{\lg[N/4(Q+1)]}{\lg[(Q+1)/(Q+1-2/r)]} \right\rfloor \quad (3)$$

式中, N 为信号长度。

1.2 稀疏分解理论模型

对于早期滚动轴承的故障诊断,测得的故障信号往往淹没在背景噪声中难以提取,而一般的去噪方法往往会去除掉部分有效信息,且在采集过程中伴随着大量的冗余信号,给数据的存储和信号的诊断带来了一系列问题。稀疏表示具有坚实的理论基础,以及较强的鲁棒性,被大量应用于故障诊断领域,能有效缓解数据的冗余问题。信号稀疏的过程是从构建的字典中寻找最少的原子,通过稀疏系数与原子相乘,得到信号的最简洁表达,即对于测得的滚动轴承故障信号,在特定的稀疏完备字典下,通过稀疏系数矩阵表达故障成分的过程。建立故障信号稀疏模型,为

$$y = y_0 + n \quad (4)$$

式中, y_0 为轴承故障信号, $y_0 \in \mathbf{R}^M$; n 为背景噪声。

研究表明,轴承故障信号 y_0 的可逆线性变换的稀疏表示为

$$y_0 = Ax \quad (5)$$

$$x = A^T y_0 \quad (6)$$

式中, $A \in \mathbf{R}^{M \times N'}$ 表示线性变换矩阵; $x \in \mathbf{R}^{N'}$ 表示具有稀疏性的变换系数。

为确保信号稀疏分解能够准确提取轴承故障信息,关键在于构造有效的成本函数和稀疏正则化函数。正则化函数的选择影响轴承瞬态特性的稀疏性。广泛使用的技术是最小化正则化最小二乘的代价函数。

稀疏正则化用于从噪声信号 y 中找到稀疏近似解 x 。 L_1 范数因具有凸性被广泛用于惩罚项,因此 L_1 范数正则化的最小二乘问题可以写为

$$\min_x \left\{ F(x) = \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 \right\} \quad (7)$$

式中, $\|x\|_1$ 为 L_1 范数; λ 为正则化参数,用于平衡损失

项和正则化项之间的权重。

定义软阈值算子为

$$\text{soft}(x; \lambda) = x \frac{\max\{|x| - \lambda, 0\}}{\max\{|x|, \lambda\}} \quad (8)$$

式(7)将稀疏模型转化为基于 L_1 范数正则化的凸优化问题,为求稀疏解提供了 1 个目标函数。

2 多指标优化 TQWT 的稀疏表示方法

2.1 TQWT 增强非凸正则项稀疏表示

线性变换是信号稀疏表示的重要环节。高效的线性变换有利于提高系数的稀疏性。对于轴承故障诊断,本文采用 TQWT 作为线性变换矩阵 A 。

TQWT 可以有效地表示周期或准周期脉冲响应,是一种离散的小波变换,其系数矩阵是一个满秩矩阵,满足可逆的条件。使用 radix-2 基函数进行快速傅里叶变换,可以有效实现适合应用于稀疏正则化反向问题的快速求解。与传统的小波变换相比,TQWT 的品质因子 Q 和冗余因子 r 易于独立指定,可以根据待处理信号的振荡行为对小波基函数进行匹配选择,从而提升稀疏信号表示的稀疏性。

在实际处理轴承故障信号中,使用 L_1 范数作为惩罚项存在一些明显的缺点。 L_1 范数会低估信号的高振幅分量,而在大多数情况下,这些分量往往是故障信号的一部分,因此,提出一种非凸稀疏性的惩罚项,以期增强稀疏性,并保留较多的故障信号特征。

一般地,将非凸正则化的最小二乘问题建模为

$$\min_x \left\{ F(x) = \frac{1}{2} \|y - Ax\|_2^2 + \lambda \psi(x) \right\} \quad (9)$$

式中, $\psi(x)$ 为非凸稀疏性惩罚。

从稀疏性考虑,本文建立一种特殊的非凸稀疏性惩罚,称为 GMC 惩罚,它可以保持函数的凸性,避免函数中存在局部极小值。GMC 惩罚的形式为

$$\phi(x) = \begin{cases} |x| - \frac{\gamma}{2} x^2, & |x| < \frac{1}{\gamma} \\ \frac{1}{2\gamma}, & |x| \geq \frac{1}{\gamma} \end{cases} \quad (10)$$

式中, γ 为稀疏惩罚项阈值。

通过 Moreau 包络将该惩罚项改写为

$$\psi_{MC}(x) = \|x\|_1 - \min_v \{ \|v\|_1 + \frac{\gamma}{2\lambda} \|x - v\|_2^2 \} \quad (11)$$

式中, v 为优化变量,即为寻找最小值的目标。

因此将非凸正则化的最小二乘问题改写为

$$F(x) = \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 - \lambda \min_v \{ \|v\|_1 + \frac{\gamma}{2\lambda} \|x - v\|_2^2 \} \quad (12)$$

GMC 惩罚项没有直接的计算式,因此使用包含

双嵌套循环的迭代算法求出内、外循环的最优解,将式(12)转化为式(13)的鞍点问题,即

$$(x^{\text{opt}}, v^{\text{opt}}) = \arg \min_x \max_v F(x, v) \quad (13)$$

其中,

$$F(x, v) = \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 - \lambda \|v\|_1 - \frac{\gamma}{2} \|A(x - v)\|_2^2 \quad (14)$$

GMC 迭代算法中迭代算法逻辑如下:

- ① Initialize $x^{(0)}, v^{(0)}, 0 < \mu < 2 / \max\{1, \gamma/(1 - \gamma)\} \|A\|_2^2$
- ② for $i=0, 1, 2, \dots$, do
- ③ $\omega^i = x^i - \mu A^T (Ax^i - y) + \mu \gamma A^T A (x^i - v^i)$,
- ④ $u^i = v^i + \mu \gamma A^T A (x^i - v^i)$,
- ⑤ $x^{(i+1)} = \text{soft}(\omega^i; \lambda \mu)$,
- ⑥ $v^{(i+1)} = \text{soft}(u^{(i+1)}; \lambda \mu)$,
- ⑦ end for
- ⑧ return $x^{(i+1)}$

在算法迭代过程中加入了 $\mu \gamma A^T A (x^i - v^i)$ 项,其中, μ 为步长(学习率),避免了 L_1 范数造成的去除高振幅信号信息的问题。

2.2 多指标自适应优化 TQWT 的稀疏表达

在确定基于 TQWT 的稀疏表示故障诊断中选取小波基函数的参数和 GMC 惩罚项的阈值 γ 时,通常以最大限度地提高信号的稀疏表示为目标选择参数,没有考虑故障冲击信号的固有特性,或者仅考虑谱峭度最大的信号分量。通过这种方法选择出的参数会对重构出的信号做进一步的包络谱分析造成较大影响。

考虑到计算量和实际应用^{[15]3560-3575},若 r 过小, $H_0(\omega)$ 和 $H_1(\omega)$ 的频带将相对较窄并且小波的时域响应并不能很好地被局部化,且 r 过小(如 $r=1$)时会导致小波更像 sinc 小波,不利于信号的重建。因此,为了更好地与实际的振动信号相匹配,将冗余因子 r 的值设置为 3。

当品质因子 $Q=1$ 时,会获得 1 个类似于高斯分布的 2 阶导数,如图 2 所示,几乎无法匹配故障特征,而当 Q 值增大,小波的振荡属性也随之增高,时域的波形振荡次数、相邻 2 个子带之间的重叠度也越低,小波基函数会难以匹配共振信号的振荡特性而产生奇异信号。

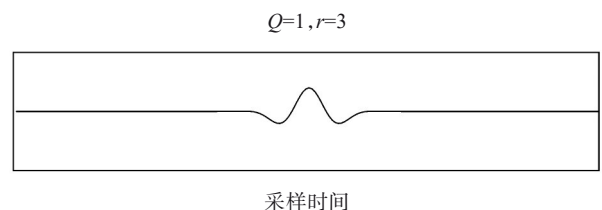


图 2 小波基函数示意图

Fig. 2 Diagram of the signal wavelet basis function

由式(3)可知,最大分解层数 $J_{\max}$ 是由品质因子 Q 和冗余因子 r 决定的,因此分解层数 $J_{\max}$ 的选取范围依赖于品质因子 Q 和冗余因子 r 。为保证信号的稀疏性,将稀疏惩罚项阈值 γ 的范围设置为0.3~1。因此需要找到合理的方法对品质因子 Q 、分解层数 J 和稀疏惩罚项参数阈值 γ 进行自适应选择。

对以上待定参数进行自适应构造时,已有较多用于揭示脉冲故障特征的指标,如峭度、相关峭度和包络谐波噪声比(Envelope Harmonic-to-Noise Ratio, EHNR)等,但这些指标大多缺乏在各种噪声下的鲁棒性和对脉冲故障特征的敏感性。

峭度 K 对故障信号的单一冲击更加敏感,其表达式为

$$K = \frac{\sum [x(n) - u]^4}{\sigma^4 N} \quad (15)$$

式中, u 为信号均值; σ 为序列标准差。

但峭度是1个全局指标,无法对特定信号分量的变化进行表征,对强噪声背景环境下的故障检测敏感度较差。本文对峭度进行改进,引入峭度偏斜度比 K_s 代替峭度,其定义为

$$K_s = \frac{K}{s} \quad (16)$$

$$s = \frac{\sum [x(n) - u]^3}{N} \quad (17)$$

由式(16)、式(17)可知,峭度偏斜度比增强了峭度对冲击信号的敏感性,去除了信号本身对指标值的影响,有效增强了故障特性。

EHNR是描述谐波成分与噪声比率的定量指标,并且在周期性故障特征存在时会增大。此外,EHNR的优点是可以很容易地通过计算自相关函数来估计。

将EHNR记作 E ,并定义为

$$E = \frac{\int_0^{f_c/2} e_s(f) e_s(f + f_m) df}{\int_0^{f_c/2} e_s^2(f_m) df - \int_0^{f_c/2} e_s(f) e_s(f + f_m) df} \quad (18)$$

式中, $e_s(f)$ 为包络谱; f_m 为故障频率; f 为频率。

但是,在工程应用中,间隔为 f 的噪声分量也将被计入谐波项,使得 E 作为指标缺少鲁棒性。

因此本文提出包络频谱的谐波与噪声能量比(Harmonic-to-Noise Energy Ratio of the Envelope Spectrum, ES-HNER)改进EHNR,作为选择小波基函数参数和稀疏惩罚项阈值的指标,即

$$E_H = \frac{\sum_{i=1,2,\dots}^k e_s^2(f_{iM})}{\sum_{i=1,2,\dots}^k e_s^2(f_{iM}) - \sum_{j=1,2,3,\dots}^{N/2} e_s^2(f_j)} \quad (19)$$

式中, $\sum_{i=1,2,\dots}^k e_s^2(f_{iM})$ 为故障特征频率前 k 阶谐波总能量; $\sum_{j=1,2,3,\dots}^{N/2} e_s^2(f_j)$ 为包络谱总能量。

从数学方面考虑,ES-HNER计算故障前 k 阶谐波分量的总能量与包络频谱的其余部分。因此,ES-HNER是一种定性和直接指标,揭示了轴承故障特征的丰富性和强烈性。由式(19)可知,ES-HNER对轴承的故障特征较为敏感,当产生故障时,ES-HNER的值会随之增大,因此可以很好地指示故障信息。

稀疏度 S 可以表示信号的故障稀疏程度,其表达式为

$$S = \frac{\sqrt{\sum x^2(n)}}{\sum |x(n)|} \quad (20)$$

求出参数范围内每1个品质因子 Q 、稀疏惩罚项参数阈值 γ 及所确定的分解层数 J 所决定的子频带稀疏信号的3种指标值序列 $\{K_s[Q(i), \gamma(j), J(k)]\}$ 、 $\{E_H[Q(i), \gamma(j), J(k)]\}$ 、 $\{S[Q(i), \gamma(j), J(k)]\}$,并将该指标序列构成多指标的状态参数矩阵,为

$$M = (x_{ijk})_{Q \times \gamma \times J} = \begin{bmatrix} K_s(1) & K_s(2) & \cdots & K_s(n') \\ E_H(1) & E_H(2) & \cdots & E_H(n') \\ S(1) & S(2) & \cdots & S(n') \end{bmatrix}_{i,j,k} \quad (21)$$

式中, x_{ijk} 为品质因子 Q 取 i 值、第 j 层分解层数、稀疏惩罚项参数阈值 γ 取 k 值所决定的多指标参数状态矩阵。

为确定3种指标对于故障信息的指示能力,引入信息熵作为评判标准。信息熵的数学表达式为

$$H(x) = - \sum_{x \in X} p(x) \lg p(x) \quad (22)$$

式中, $p(x)$ 为发生的概率。

信息熵描述了信息的不确定性,其大小描述了信息中所包含的信息量,值越小代表含有越多的信息,即表示该指标可以指示更多的故障信息。因此,在使用参数状态矩阵进行自适应确定函数时,应赋予信息熵较小的指标和更大的权值。

因此,对式(21)的参数状态矩阵同一类指标求其信息熵,为

$$\begin{cases} p_{ijk} = \frac{x_{ijk}}{\sum_{j=1,k=1}^i x_{ijk}}, \frac{x_{ijk}}{\sum_{i=1,k=1}^j x_{ijk}}, \frac{x_{ijk}}{\sum_{i=1,j=1}^k x_{ijk}} \\ E_i, E_j, E_k = - \sum p_{ijk} \ln(p_{ijk}) \end{cases} \quad (23)$$

计算信号的峭度偏斜度、谐波与噪声能量比和稀疏度时会产生量纲上的差异,因此将3个状态指标进行

归一化后,根据信息熵的大小对其赋予权重,最后再计算出不同指标下稀疏信号子频带的融合特征参数。

$$\left\{ \begin{array}{l} W = 1 - \frac{E_i}{n'} \\ x_{ijk}^* = \frac{x_{ijk} - \min(x_{jk})}{\max(x_{jk}) - \min(x_{jk})}, \frac{x_{ijk} - \min(x_{ik})}{\max(x_{ik}) - \min(x_{ik})}, \\ \frac{x_{ijk} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \\ I = \sum_{i=1}^3 W \cdot x_{ijk}^* \end{array} \right. \quad (24)$$

式中, W 为状态参数权重; I 为融合后的指标参数矩阵。

得到所有的品质因子 Q 、分解层数 J 、稀疏惩罚项参数阈值 γ 决定的融合参数矩阵后,选择最大的融合指标参数所对应的 Q 、 J 、 γ 作为自适应挑选后的小波基函数参数及稀疏惩罚项参数阈值。最后将挑选后的结果应用到 TQWT 的稀疏表达中,对故障信号进行特征提取。

3 故障轴承诊断流程

本文提出的具有非凸正则项系数的稀疏分解、并使用多指标特征矩阵进行自适应重构信号的轴承故障诊断方法如图 3 所示,其具体步骤如下:

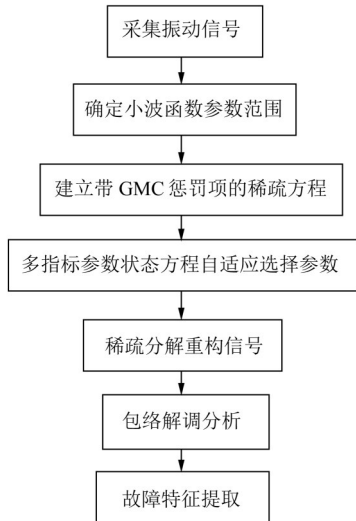


图 3 稀疏分解故障诊断流程图

Fig. 3 Flow chart of sparse decomposition fault diagnosis

1) 使用传感器等设备采集轴承故障信号,对采集的含噪低频振动信号进行分析,根据其共振属性选择品质因子 Q 及稀疏惩罚项参数阈值 γ 的范围。第 2.2 节中的 Q 值需要在一定的范围内,多次试验表明,当 Q 值为 1~10 时,对故障的提取效果最好,因此本文将品质因子的范围设为 1~10。根据确定的 Q 和 γ 计算出分解层数 J 的范围,最后对由其组成的参数矩阵中所有的元素进行可调品质因子小波分解,得到各个子带的

小波系数 $w_i (1 \leq i \leq J)$ 。

2) 针对小波系数 w_i 建立优化的目标函数,设置相应的正则化参数,引入 GMC 惩罚项,用优化算法求出最优解 $\hat{w}$,即低频振荡信号的小波系数最佳稀疏表示。

3) 引入参数状态方程,对每 1 组小波系数参数矩阵进行自适应选择,最终选出最优的品质因子 Q 和分解层数 J 和稀疏惩罚项参数阈值 γ ,重构出含噪成分最少的信号。

4) 对重构信号进行 Hilbert 变换得到其包络谱,观察频域信息后,与理论故障频率对比,得出滚动轴承的故障位置。

4 仿真信号检验

4.1 仿真信号构造及检验

通过构造数字仿真信号,可以验证所提出方法的有效性和可靠性。其中,构造的振动数学模型包含了带有强噪声的滚动轴承元件故障。仿真信号的冲击单元为

$$x(t) = e^{-400t} \cos(1250\pi t) \quad (25)$$

设置仿真信号的采样频率 $f_s = 10\,000$ Hz,故障频率 f_0 为 16 Hz,图 4 所示为仿真的故障信号,为定性进行比较,本文图中的幅值均已作归一化处理。

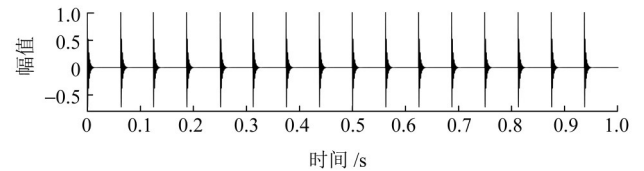


图 4 仿真故障信号

Fig. 4 Simulated fault signal

为验证本文提出方法的优越性,并考虑到滚动轴承振动信号的多分量特性,向冲击信号中加入初始频率为 500 Hz、结束频率为 1 500 Hz 的线性调频信号,起始频率为 1 000 Hz、结束频率为 2 500 Hz 的二次非线性调制信号;并引入调制效应,其载波频率为 1 000 Hz、调制频率为 50 Hz,最后添加信噪比为 -8 dB 的高斯白噪声用于模拟环境噪声,以更接近实际的滚动轴承振动信号。

图 5 所示为添加噪声后的仿真故障信号。

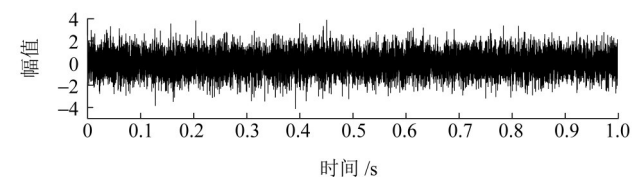


图 5 含噪仿真故障信号

Fig. 5 Simulated fault signal with noise

由图5可知,添加高斯噪声后的信号较为混乱,冲击特征被噪声覆盖,无法得到清楚的故障信息。图6所示为含噪信号的包络谱,无法从中分辨故障成分的有效信息,故需做进一步处理。

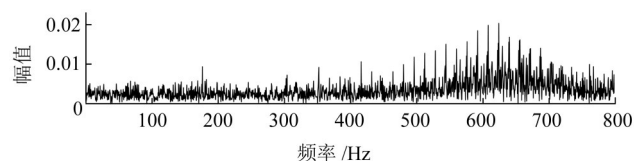


图6 含噪信号频域

Fig. 6 Frequency-domain of noisy signals

使用文中的方法对信号进行处理,根据前文确定出稀疏惩罚项参数阈值 γ 范围为0.3~1,搜索步长为0.1;品质因子 Q 的范围为1~10,搜索步长为1;冗余系数 r 取3;计算品质因子范围内的最大分解层数 J 。

根据多指标参数状态矩阵对所有子频带进行筛选。自适应选择的小波基函数参数为 $Q=2, r=3, J=1$,稀疏惩罚项参数阈值 $\gamma=0.85$ 。图7中蓝色线表示故障特征成分,红色虚线表示小波基函数,由图7可知,自适应构造的小波基函数与仿真的故障成分相似度较高。

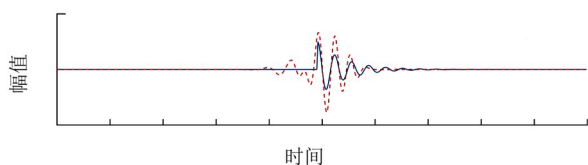


图7 自适应构造的小波基函数

Fig. 7 Adaptive constructed wavelet basis function

应用本文方法对原始信号进行故障提取,得到的信号如图8所示。

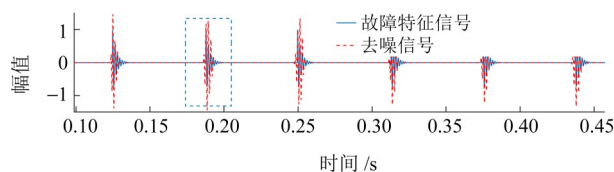


图8 去噪信号时域

Fig. 8 Time-domain of denoised signals

由图8可知,本文的方法可以很清晰地从含噪信号中提取出仿真故障成分。

对去噪信号进行频谱分析,如图9所示。由图9可知,故障特征频率及其谐波特征表现显著,且包络谱线较为清晰,去除了大量的噪声成分。

综上所述,可以证明本文方法可以提取出故障特征,具有可靠性和有效性。

4.2 小波阈值降噪方法对比

小波阈值降噪基于内积匹配原理,可以通过设置阈值的方式去除噪声成分对应的小波系数,是信号去

噪的常用方法之一。本节使用小波阈值降噪方法进行试验。

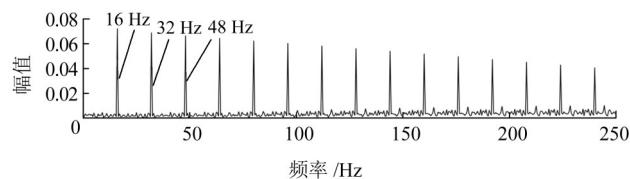


图9 去噪信号频域

Fig. 9 Frequency-domain of denoised signals

对比本文方法的去噪结果,虽然小波阈值降噪能表现出部分冲击成分(图10),但是冲击周期并不明显,信号仍含有较多噪声;图11为小波阈值降噪的频域图。由图11可知,噪声成分仍然较为明显,提取效果较差。

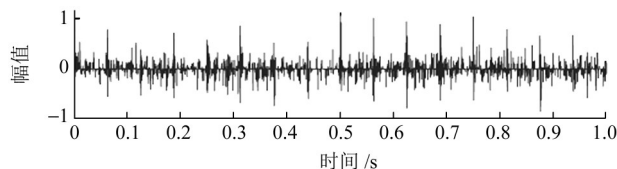


图10 小波阈值降噪信号时域

Fig. 10 Time-domain of threshold-denoised signals

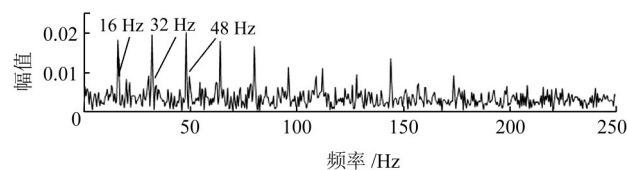


图11 小波阈值降噪信号频域

Fig. 11 Frequency-domain of threshold-denoised signals

在构建仿真信号时,尽管冲击信号的设置较为明显,但采用小波阈值降噪方法仍难以有效提取故障特征。而对于更加复杂的真实信号,小波阈值降噪方法的适用性进一步降低,无法满足故障特征提取的要求。

5 故障轴承信号诊断

5.1 试验说明

在仿真信号验证成功的基础上,通过使用西安交通大学发布的数据库(XJTY-SY)故障数据集^[16]来验证本文方法在实际应用中的可行性。

选择轴承 Bearing1_2 试验数据集作为试验信号,根据滚动轴承理论外圈故障频率计算方法,计算出该轴承在所选工况下的理论外圈故障频率为107.9 Hz,同时计算其倍频成分。

对该试验第60 min时监测信号进行试验,如图12所示。图13所示为去噪后的试验信号,图14所示为去噪信号的频域信息。

图13从时域方面较为清晰地表示出提取的周期性故障信号,而图14从去噪信号的频域方面表示的倍

频成分与理论的故障特征频率吻合程度较高,证明了此方法能够提取出故障信号的特征信息。

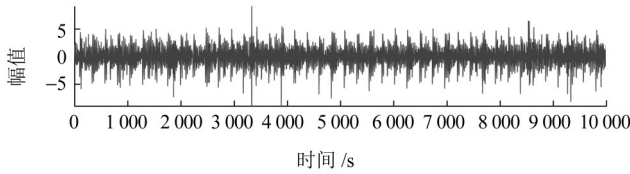


图 12 60 min 试验信号

Fig. 12 60 min test signals

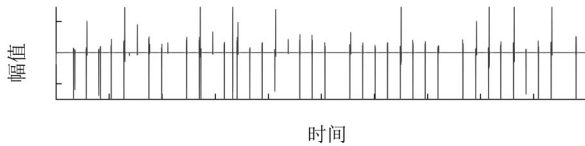


图 13 去噪试验信号

Fig. 13 Denoised test signals

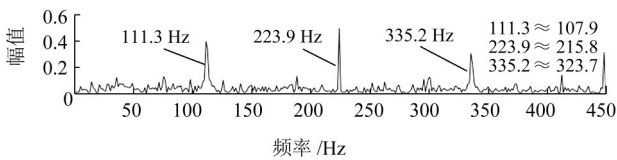


图 14 去噪试验信号频域

Fig. 14 Frequency-domain of denoised test signals

5.2 方法对比

5.2.1 TQWT 方法对比

一般的 TQWT 方法不具有稀疏性,其品质因子固定,无法在试验中自适应确定品质因子 Q 、冗余因子 r 及最大分解层数 $J_{\max}$ 的值;使用这种方法对试验轴承的故障信息提取效果较差,如图 15、图 16 所示分别为使用一般的 TQWT 提取方法得到的故障时域信息和频域信息。

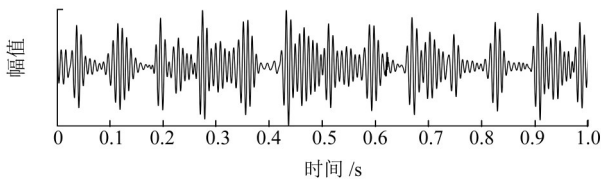


图 15 一般的 TQWT 方法重构信号时域

Fig. 15 Time-domain of reconstructed signals by conventional TQWT method

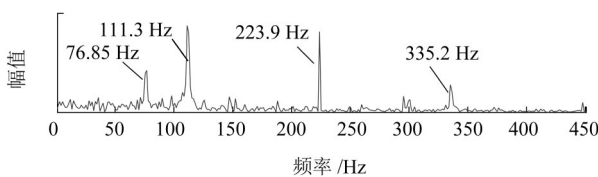


图 16 一般的 TQWT 方法重构信号频域

Fig. 16 Frequency-domain of reconstructed signals by conventional TQWT method

由图 15 可知,未经自适应构造参数的 TQWT 方法未能较好地提取出周期性的故障冲击信号,淹没在较

多的噪声信息中,无法提取出故障信号;而由图 16 可知,存在 76.85 Hz 的错误故障频率倍率,因此使用一般的 TQWT 方法会受噪声的影响,无法准确地提取出故障信息。因此,能够证明本文所提方法可以有效改进参数自适应选择问题,且引入稀疏性后可以有效地去除噪声对试验的影响。

5.2.2 单指标方法对比

仅使用单个指标对试验信号进行故障信号提取,以验证本文使用的参数矩阵的优越性。使用式(19)的 E_H 作为指标进行提取,其试验结果如图 17 所示。

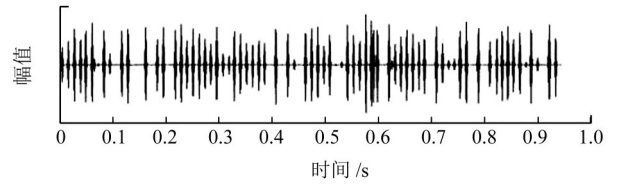


图 17 单指标方法重构信号时域

Fig. 17 Time-domain of reconstructed signals by single index method

由于 ES-HNER 本质上是对信号的能量特性的分析,仅使用该单个指标无法完全表征出信号的其他特性(如仅使用峭度只能表征出信号对单次冲击的敏感度)。由图 18 可知,其提取效果与理论倍频的差距较大,且无法显示更多的倍频信息。因此,本文方法中引入参数矩阵更有优势。

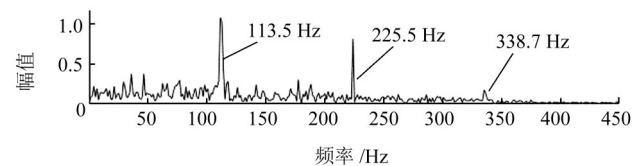
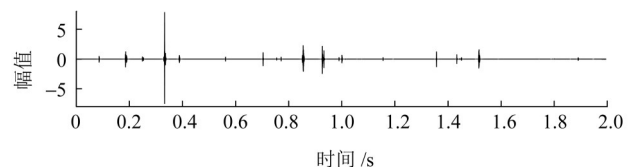


图 18 单指标方法重构信号频域

Fig. 18 Frequency-domain of signals reconstructed by single index method

5.2.3 消融试验对比

为证明本文方法中使用 GMC 作为惩罚项的必要性,设计消融试验进行对比,使用 L_1 范数作为增强稀疏性的惩罚项,其试验结果如图 19、图 20 所示。

图 19 L_1 惩罚项方法重构信号时域Fig. 19 Time-domain of signals reconstructed by L_1 -penalty method

由于使用 L_1 范数作为惩罚项,其特点是容易去除信号的高振幅信息,所以无论是时域还是频域,都无法提取到相应故障特征。因此,本文方法中引入 GMC 作为惩罚项更有优势,可以很好地解决重构信号缺乏

高振幅信息的问题。

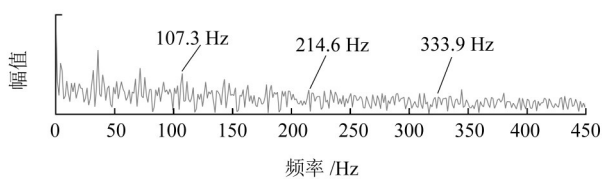


图 20 L_1 惩罚项方法重构信号频域

Fig. 20 Frequency-domain of signals reconstructed by L_1 -penalty method

6 结论

提出了一种结合稀疏分解和自适应小波分析的方法,针对轴承故障诊断中的信号处理与特征提取进行了深入研究,并通过试验验证了其有效性,得出以下主要结论:

1)针对轴承故障特征准确率不高,一般方法易去除部分有效信息、测量信号数据量冗余的问题,本文使用稀疏分解的方法可以很好地解决这些问题。与传统滤波方法相比,非凸正则项系数稀疏分解的方法能更好地保留轴承故障信号,同时降低噪声和其他成分的干扰,显著提升了轴承故障信号的估计精度。

2)针对 TQWT 的稀疏表示方法具有多频带解调特性,使用峭度偏斜度比、谐波与噪声能量比和稀疏值构成多指标参数状态矩阵,对小波基参数及稀疏惩罚项参数阈值进行自适应选择,并使用信息熵对其加权归一化融合。使用该指标对参数进行选择,可以有效避免单一指标的局限性,使频带选择更准确合理,结果更加准确。

3)使用 XJTY-SY 数据集进行试验,并使用多个算法对试验信号进行故障提取,对比结果表明,本文提出的方法能够更好地提取出复杂工况下微弱轴承故障信息。

参考文献 (References)

- [1] 王洪伟. 航空发动机滚动轴承故障诊断与预测关键技术研究 [D]. 南京:南京航空航天大学,2015:1-160.
WANG Hongwei. Research on key technologies of fault diagnosis and prediction of aero-engine rolling bearing [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015: 1-160. (In Chinese)
- [2] 孙鑫威,纪爱敏,杜占涛,等. 动车组齿轮箱滚动轴承变速故障诊断方法[J]. 哈尔滨工业大学学报,2023,55(1):106-115.
SUN Xinwei, JI Aimin, DU Zhantao, et al. Fault diagnosis method for variable speed of rolling bearing in EMU gearbox[J]. Journal of Harbin Institute of Technology, 2023, 55(1): 106-115. (In Chinese)
- [3] TSENG L M, CHUANG C Y, CHUA S K, et al. Identification of coronary culprit lesion in ST elevation myocardial infarction by using deep learning[J]. IEEE Journal of Translation Engineering in Health Medicine, 2022, 11: 70-79.
- [4] KHALIL N, HIPPENSTIEL R. Wavelet transforms of correlation functions of frequency hopped signals[C]//Conference Record of The Thirtieth Asilomar Conference on Signals, Systems and Computers. New York: IEEE, 1996: 1205-1208.
- [5] SELESNICK I W. The double-density dual-tree DWT[J]. IEEE Transactions on Signal Processing, 2004, 52(5): 1304-1314.
- [6] 唐贵基,王晓龙. 可调品质因子小波变换在滚动轴承微弱故障特征提取中的应用[J]. 中国电机工程学报,2016,36(3):746-754.
TANG Guiji, WANG Xiaolong. Application of tunable Q-factor wavelet transform to feature extraction of weak fault for rolling bearing[J]. Proceedings of the CSEE, 2016, 36(3): 746-754. (In Chinese)
- [7] 余发军,周凤星. 基于 EEMD 和自相关函数特性的自适应降噪方法[J]. 计算机应用研究,2015,32(1):206-209.
YU Fajun, ZHOU Fengxing. Adaptive de-noising method based on EEMD and autocorrelation function property [J]. Application Research of Computers, 2015, 32(1): 206-209. (In Chinese)
- [8] MALLAT S G. A theory for multiresolution signal decomposition: the wavelet representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(7): 674-693.
- [9] LIU H N, LIU C L, HUANG Y X. Adaptive feature extraction using sparse coding for machinery fault diagnosis[J]. Mechanical Systems and Signal Processing, 2011, 25(2): 558-574.
- [10] CUI L L, WANG J, LEE S. Matching pursuit of an adaptive impulse dictionary for bearing fault diagnosis [J]. Journal of Sound and Vibration, 2014, 333(10): 2840-2862.
- [11] XU Z B, CHANG X Y, XU F M, et al. $L_{1/2}$ regularization: a thresholding representation theory and a fast solver [J]. IEEE Transactions on Neural Networks and Learning Systems, 2012, 23(7): 1013-1027.
- [12] WANG Y, ZENG J S, PENG Z M, et al. Linear convergence of adaptively iterative thresholding algorithms for compressed sensing [J]. IEEE Transactions on Signal Processing, 2015, 63(11): 2957-2971.
- [13] SELESNICK I, FARSHCHIAN M. Sparse signal approximation via nonseparable regularization[J]. IEEE Transactions on Signal Processing, 2017, 65(10): 2561-2575.
- [14] WANG S B, SELESNICK I, CAI G G, et al. Nonconvex sparse regularization and convex optimization for bearing fault diagnosis [J]. IEEE Transactions on Industrial Electronics, 2018, 65(9): 7332-7342.
- [15] SELESNICK I W. Wavelet transform with tunable Q-factor [J]. IEEE Transactions on Signal Processing, 2011, 59(8): 3560-3575.
- [16] WANG B, LEI Y G, LI N P, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings[J]. IEEE Transactions on Reliability, 2020, 69(1): 401-412.

Bearing fault diagnosis based on multi-index optimization TQWT sparse representation

QIANG Yudong¹ WAN Fangyi¹ ZHANG Chunlin¹ ZHANG Changxing²

(1. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

(2. School of Advanced Materials and Nanotechnology, Xidian University, Xi'an 710126, China)

Abstract: [Objective] Aiming at the problems of low precision in bearing vibration fault feature extraction, fault information easily submerged by strong noise, lack of reasonable evaluation indices for adaptive signal construction, and data redundancy, a study on bearing fault feature enhancement and fault diagnosis was conducted. [Methods] A sparse decomposition method with non-convex regularization coefficients was proposed, and adaptive signal reconstruction is realized based on a multi-index feature matrix. Firstly, the signal is sparsely decomposed via tunable Q-factor wavelet transform (TQWT) to obtain its most concise representation. A non-convex penalty function is constructed to improve the sparsity of the reconstructed signal while retaining key fault feature information. Secondly, a fused index parameter matrix was established by integrating harmonic-to-noise energy ratio, kurtosis-skewness ratio, and sparsity, which was adopted to optimize the parameters of tunable wavelets and the threshold of sparse penalty terms. Finally, signal denoising and reconstruction were completed using the optimized parameters, and demodulation analysis was performed on the reconstructed signal to extract bearing fault feature information. [Results] Simulation and test results show that the proposed method can effectively extract the impact features of early bearing faults and achieve accurate diagnosis of bearing vibration faults.

Key words: Multi-index optimization; Tunable Q-factor wavelet; Sparse decomposition; Fault diagnosis; Signal processing

Corresponding author: WAN Fangyi, E-mail: fwan@nwpu.edu.cn

Fund: National Natural Science Foundation of China (11972314)

Received: 2024-07-08

Revised: 2024-09-29

(编辑:刘欢欢)