

基于非负矩阵分解的单通道故障特征分离方法^{*}

梁霖¹, 栗茂林², 李利邦¹, 刘飞¹, 徐光华^{1,3}

(1. 西安交通大学机械工程学院 西安, 710049) (2. 西安交通大学工程坊 西安, 710049)

(3. 西安交通大学机械制造系统工程国家重点实验室 西安, 710054)

摘要 针对单通道振动信号的多特征分离问题,提出了一种基于正交非负矩阵分解的故障特征提取方法。首先,采用短时傅里叶变换,利用时频分布来描述信号中的局部故障特征,通过核心一致性指标评估子空间维数;然后,在幅值谱矩阵分解的基础上,通过正交性约束实现低维嵌入分量信息的分离,获取局部特征的准确描述;最后,采用相位恢复理论重构出特征波形,对仿真信号和滚动轴承故障数据进行了测试。结果表明,所提出的方法能利用单通道信号有效地分离出微弱的局部故障特征,为机械状态的早期故障诊断识别提供了一种有效手段。

关键词 非负矩阵分解;单通道信号;特征提取;故障诊断

中图分类号 TH165.3

引言

在设备运行状态的监测诊断中,轴承座上的测量传感器能够提供多个部件的信息,包括齿轮振动及啮合、轴承振动,转子振动以及其他干扰噪声等。在某些情况下,振动信息只能利用单传感器采集,考虑到安装成本,往往在每级单侧安装振动传感器。因此,需要利用单通道振动信号的盲分离提取出故障特征分量。

对于包含了多源信息的信号来说,通常可以采用变换域滤波方法实现多源分量的分离。比如:采用自适应 Winner 滤波器实现多信号的分离^[1-2];采用独立主分量分析、盲源分离、基函数法、稀疏分解以及多参数联合估计等方法实现多源信号分解^[3-6]。非负矩阵分解(non-negative matrix factorization, 简称 NMF)作为一种新的矩阵分解算法,克服了传统矩阵分解的很多问题^[7-8],通过对非负矩阵的聚类分析,有效地提取出局部特征^[9-13],同时具有物理意义明确和计算简单等优势,成为单通道振动信号源分离的一种有效手段。

由于设备结构、运行环境等因素,往往导致振动测量信号信噪比较低,多振动分量信息相互干扰,通过 NMF 的局部学习能力,可以有效地将故障振动分量实现分离,从而为运行状态分析提供解决工具。

为此,在单通道信号的非负矩阵分解特征提取中,通过原始观测矩阵的构建和低维嵌入维数的选择,设计基于正交非负矩阵分解的特征分离方法,并通过模拟数据及轴承故障数据进行有效性验证。

1 非负矩阵分解原理

对一个 m 维的随机向量 \mathbf{v} 进行 n 次观测,得到观测矩阵 $\mathbf{V}=[v_1, v_2, \dots, v_n]$,非负矩阵分解的问题可定义为:已知非负矩阵 \mathbf{V} ,找出非负矩阵 \mathbf{W} 和 \mathbf{H} ,使式(1)成立^[14]

$$\mathbf{V} = \mathbf{WH} \quad (1)$$

其中: \mathbf{V} 为 $m \times n$ 维的非负矩阵; \mathbf{W} 为 $m \times k$ 维的基矩阵; \mathbf{H} 为 $k \times n$ 维系数矩阵。

由于 $m \gg k$,从而实现了观测矩阵的约简。

为了寻找合理的 \mathbf{W} 和 \mathbf{H} ,以常用欧几里得距离(Euclidean distance)作为目标函数,则式(1)的逼近可以视为一个优化问题

$$\min D(\mathbf{V} \parallel \mathbf{WH}) = \frac{1}{2} \|\mathbf{V} - \mathbf{WH}\|_F^2 \quad (2)$$

其中: $\|\cdot\|_F$ 为矩阵的 Frobenius 范数。

针对式(2)的优化问题,目前已经提出了多种迭代算法,如乘性迭代算法、梯度下降算法、交替非负最小二乘算法以及改进算法等^[15]。

^{*} 国家自然科学基金资助项目(51575438)

收稿日期:2014-08-06;修回日期:2014-12-31

2 正交非负矩阵分解算法

NMF作为一种基于局部表示整体的策略,在提取过程中,若不同源信号间的相似度越低,所表现出的局部性就越强,分解结果就越好。这一特性表现为基向量之间的正交性。在设备故障诊断中,如数据中源信号彼此间的特征差异不明显,那么分解的向量间就会存在冗余,不利于后续分析。带有正交性约束的非负矩阵分解方法(orthogonal NMF,简称 ONMF)^[16]可以有效地满足这种需求,即通过求解正交约束项的优化,获取的结果带有正交性。针对基矩阵 \mathbf{W} 和权矩阵 \mathbf{H} 施加正交性的目标式可表示为

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} \|\mathbf{V} - \mathbf{WH}\|^2, \quad \mathbf{W}^T \mathbf{W} = \mathbf{I} \quad (3)$$

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} \|\mathbf{V} - \mathbf{WH}\|^2, \quad \mathbf{H}^T \mathbf{H} = \mathbf{I} \quad (4)$$

其中: \mathbf{I} 为单位矩阵。

式(3)和式(4)可以通过引入拉格朗日乘子求得约束问题的近似解。求解的运算量和矩阵维数密切相关,维数较大,计算量就会很大。因此,将正交性因子矩阵的产生包含在目标优化中^[17],即对 \mathbf{W} 和 \mathbf{H} 施加正交性的目标

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} \|\mathbf{V} - \mathbf{WH}\|^2 + \lambda \|\mathbf{W}^T \mathbf{W} - \mathbf{I}\|^2 \quad (5)$$

$$\min_{\mathbf{W}, \mathbf{H} \geq 0} \|\mathbf{V} - \mathbf{WH}\|^2 + \lambda \|\mathbf{H}^T \mathbf{H} - \mathbf{I}\|^2 \quad (6)$$

其中: λ 为正交性系数且 $\lambda \geq 0$ 。

求解式(5)的迭代为

$$\mathbf{W} = \mathbf{W} \otimes \left(\frac{\mathbf{VH}^T + 2\lambda \mathbf{W}}{\mathbf{WHH}^T + 2\lambda \mathbf{WW}^T} \right) \quad (7)$$

$$\mathbf{H} = \mathbf{H} \otimes \left((\mathbf{W}^T \mathbf{V}) / (\mathbf{W}^T \mathbf{WH}) \right) \quad (8)$$

其中: \otimes 表示矩阵元素之间的乘法(除法是矩阵元素之间的除法)。

互换 \mathbf{W} 和 \mathbf{H} 位置可以得到 \mathbf{H} 的正交性惩罚迭代规则

$$\mathbf{W} = \mathbf{W} \otimes \left((\mathbf{VH}^T) / (\mathbf{WHH}^T) \right) \quad (9)$$

$$\mathbf{H} = \mathbf{H} \otimes \left(\frac{\mathbf{W}^T \mathbf{V} + 2\lambda \mathbf{H}}{\mathbf{W}^T \mathbf{WH} + 2\lambda \mathbf{HH}^T} \right) \quad (10)$$

由于正交性约束项是嵌入在目标优化方程中,因此最终的正交性是通过分解逐步得到的,无需额外的约束项,其计算量也有显著下降。

3 基于 ONMF 的单通道特征分离

结合 ONMF 和设备振动信号特点,提出了一种基于单通道机械混合振动信号分离方法,通过构造观测矩阵和选择子空间维数,基于 ONMF 的特

征分离以及波形重构实现特征分离。

3.1 观测矩阵的构造

设备故障往往导致振动信号的非平稳性表现,因此,时频域分析可以有效地提取局部故障信息并构成特征矩阵。目前,时频分析方法有短时傅里叶变换(short-time Fourier transform,简称 STFT)、小波变换和小波包分析等。虽然小波与小波包分解方法极大地改善了信号的时频分析效果,甚至已经逐步成为时频分析的首选方法。但基函数的选取需要和信号中特征类型相匹配,才能较好地提取信号中的局部特征,而设备振动信号中往往包含多源信号,单一基函数势必造成局部故障信息衡量的偏差。

相比而言,STFT 参数较少、计算简便,且对振动信号中的高频突发分量和长周期准平稳分量非常敏感,因此选取 STFT 进行分析

$$\mathbf{S}(t, f) = \int x(t + \tau) w(\tau) \exp(-2j\pi f\tau) d\tau \quad (11)$$

其中: $x(t)$ 为时域信号; $w(t)$ 为窗函数。

为了满足非负矩阵的要求,需要对由 STFT 变换得到的时频分布矩阵取模,保留其幅值信息,实现观测矩阵 \mathbf{V} 的构造,即 $\mathbf{V} = \|\mathbf{S}(t, f)\|$ 。

3.2 低维空间维数的选择

对于高维观测矩阵来说,低维子空间维数 k 是分解中的重要参数,直接影响到分解与提取的效果。 k 值过大会损失降维效果,削弱源信号的信息量, k 值较小会降低分解精度,增大冗余。为了克服传统主分量分析(principal component analysis,简称 PCA)方法中存在的缺陷,采用核心一致性计算方法来选择 k 值^[18],即通过计算数据的平滑系数

$$G(p) = \frac{\text{var}[\{\bar{\lambda}_i\}_{i=p+1}^{I-1}]}{\text{var}[\{\lambda_i\}_{i=p}^{I-1}]} \quad (p=1, 2, \dots, I-2) \quad (12)$$

其中: λ_i 为样本数据协方差的第 i 个特征值, $\bar{\lambda}_i = \lambda_i - \lambda_{i+1}$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_I$; $\text{var}[\{\bar{\lambda}_i\}_{i=p+1}^{I-1}]$ 为方差。

最优分量 k 的选择应满足如下标准

$$k = \arg \min G(p) \quad (p=1, 2, \dots, I-2) \quad (13)$$

当 G 指标最小时,所对应的 p 值即为由该信号获取到的观测矩阵 \mathbf{V} 的最优分解维数 k 。

3.3 基于 ONMF 的单通道信号特征提取方法

基于上述分析,对于包含多特征分量振动信号 $x(t)$,提出了一种单通道信号特征提取方法,其流程如下:

1) 将单通道信号通过 STFT 变换到时频域中,

并保留幅度信息,获得幅值谱 \mathbf{V} ;

2) 利用核心一致性方法估计出 \mathbf{V} 中的低维子空间维数 k 值;

3) 基于 ONMF 算法对 \mathbf{V} 进行 k 维分解,得到基矩阵 \mathbf{W} 和权矩阵 \mathbf{H} ;

4) 在低维子空间中,选择目标信号向量对应的幅值谱 $\mathbf{Y}_s = \mathbf{W}(:,s) * \mathbf{H}(s,:)$;

5) 根据式(14)所示的相位恢复理论方法重构时频谱,并基于短时傅里叶变换的逆变换(inverse short-time Fourier transformation,简称 ISTFT)将时频谱变换到时域中,获得分离的特征分量波形 $x_s(t)$ 。

$$\mathbf{S}_i = \frac{\mathbf{Y}_i}{\sum \mathbf{Y}_i} \cdot \mathbf{F} \quad (14)$$

其中: \mathbf{F} 为混合信号的时频谱; \mathbf{Y}_i 为分离出的第 i 个幅值谱; \mathbf{S}_i 为重构得到的第 i 个时频谱(含相位信息)。

4 仿真分析

为了验证文中特征分离的效果,采用由3个正弦信号构成的仿真信号进行分析。仿真信号表示为

$$x(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + \sin(2\pi f_3 t)$$

令: $A_1 = 2, A_2 = 2.5, f_1 = 200 \text{ Hz}, f_2 = 300 \text{ Hz}, f_3 = 400 \text{ Hz}$ 。同时将3个正弦信号中的部分波形置零,设采样频率为 1 000 Hz,采样点数为 1 000 点。图1和图2分别为3个源信号与混合后的仿真信号波形。

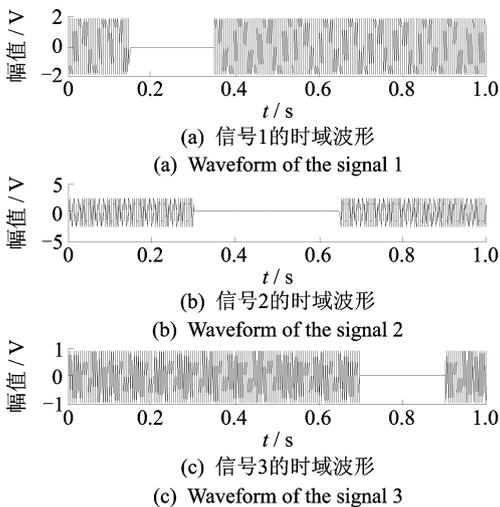


图1 3个源信号的时域波形

Fig. 1 Waveform of three simulation signals

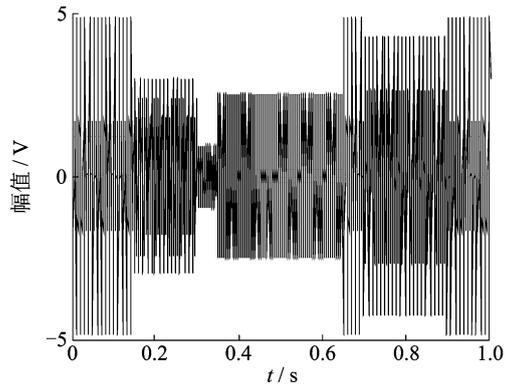


图2 混合信号的时域波形

Fig. 2 Waveform of mixed-signal by three simulation signals

对于混合仿真信号,采用文中的特征提取方法进行分析,首先通过 STFT 获得时频谱矩阵,其中窗函数为汉宁窗。在时频谱矩阵分解中,采用 ONMF 提取基函数 \mathbf{W} 与权分布 \mathbf{H} ,其中获取的 \mathbf{W} 各向量分布如图3所示。作为对比,采用传统 NMF 分解的基矩阵 \mathbf{W} 各向量分布如图4所示。

对比图3和图4的两种基向量分布可知,由于 ONMF 加强了基向量间的独立性,图3所示的3个基向量间互不干扰地分离出源信号。图4中由于缺少约束导致传统 NMF 分解的基向量间存在冗余,无法准确描述3个正弦信号。其次,结合原信号时频谱与所提取的冲击幅度谱,采用 ISTFT 方法可以获得图5所示的时域重构波形。

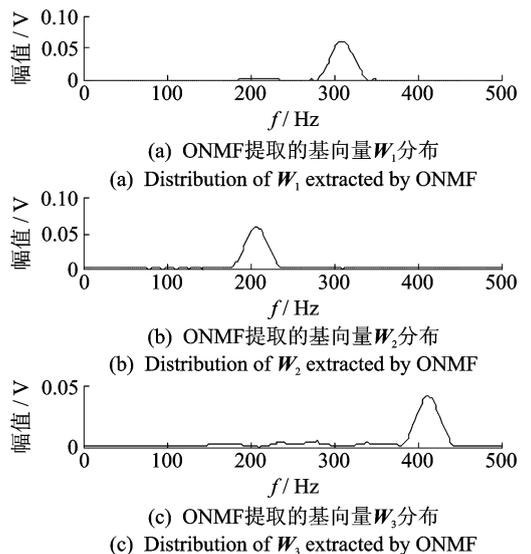


图3 ONMF 提取的基向量 \mathbf{W} 分布

Fig. 3 Distribution of basis vectors \mathbf{W} extracted by ONMF

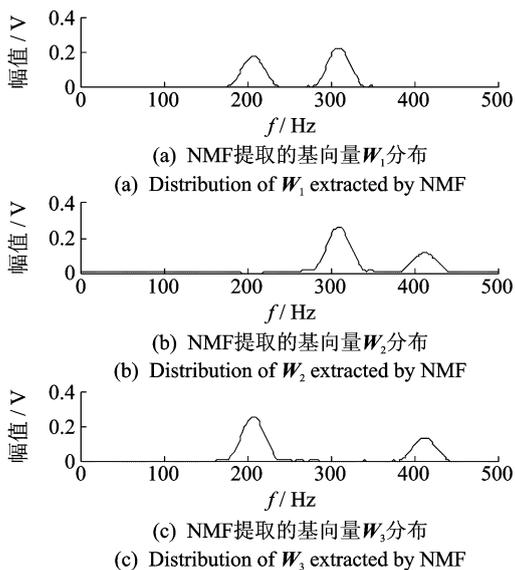
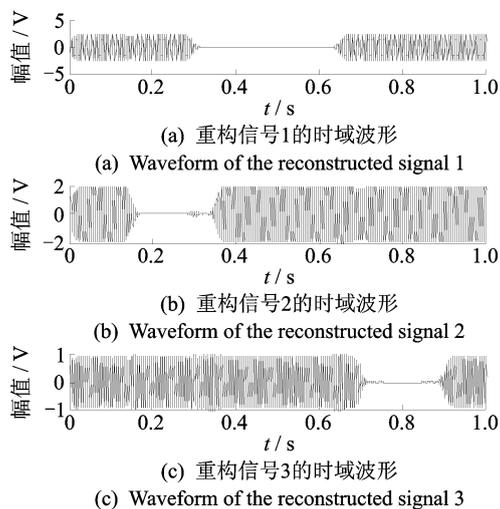
图4 NMF提取的基向量 W 分布Fig. 4 Distribution of basis vectors W extracted by NMF

图5 重构的模拟信号波形

Fig. 5 Waveform of the reconstructed signals

由图3~5对比可见,ONMF在分解冗余量较大的信号时能克服传统NMF的缺点,表明了其在振动信号分离中的优势。

5 应用实例

为验证算法的有效性,采用滚动轴承振动实验台的振动加速度信号进行外环剥落故障特征的提取。其中,轴承型号ZA-2115,转速为2000 r/min,径向负载26695 N,在轴承座上通过PCB 353BB33高灵敏度ICP加速度传感器获取振动信号,采样频率为20 kHz。图6和图7为时域波形和时频分布。

由于噪声干扰,振动信号波形中由故障引起的冲击成分受到显著影响。时频分布中,能量分布集中在5 kHz以下,在部分频带有明显的能量分布,如1.0 kHz,3.0~5.0 kHz,局部冲击对应的特征应分布在这些频率中。

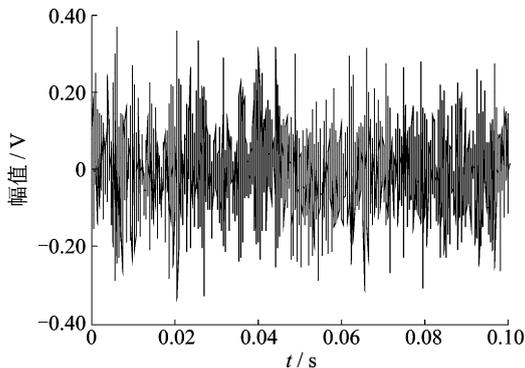


图6 滚动轴承振动信号波形

Fig. 6 Vibration waveform of rolling element bearings

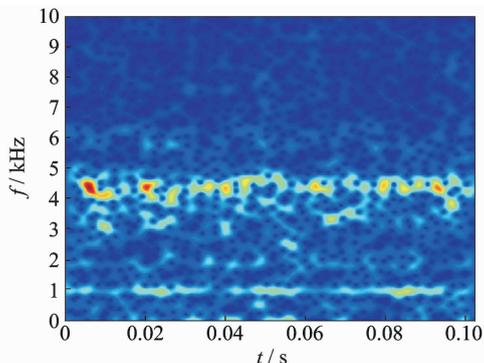


图7 轴承振动信号的时频分布

Fig. 7 Time-frequency distribution of vibration signal

采用ONMF方法分离冲击特征,由图8所示的核心一致性指标的变化曲线确定出最优子空间维数 $k=3$ 。图9~11为分离得到的3个子空间的幅度谱。同图7所示的原始时频分布对比可见,经过ONMF分解,特征被分解在不同的子空间中。其中:图9所示的第1子空间中主要包含3.2和4.2 kHz两个频带;图10所示的第2子空间中主要是1.0和4.6 kHz频带;图11所示的第3子空间具有较好的周期分布特点,对应着外环故障冲击特征表现。因此,将第3子空间时频分布对应的基向量和权向量进行重构,得到图12所示的特征波形。

由图6所示的原始波形信号对比可见,由外环故障引起的冲击特征被有效地提取出来。经过ONMF分离提取之后,信号幅值谱有效地分离出了冲击与噪声,说明了所提方法处理实际数据的有效

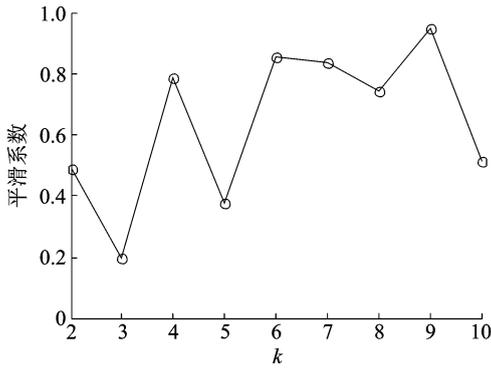


图 8 核心一致性指标的变化曲线

Fig. 8 The curve of consistency index

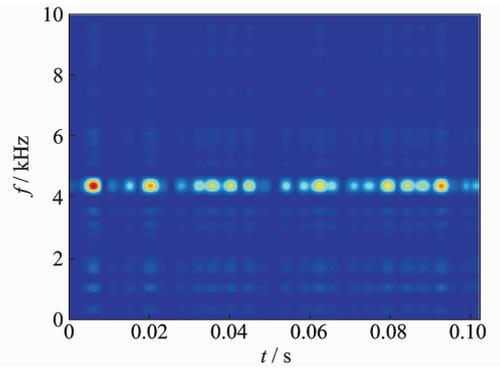


图 11 第 3 子空间时频分布

Fig. 11 Time-frequency distribution of the third subspace

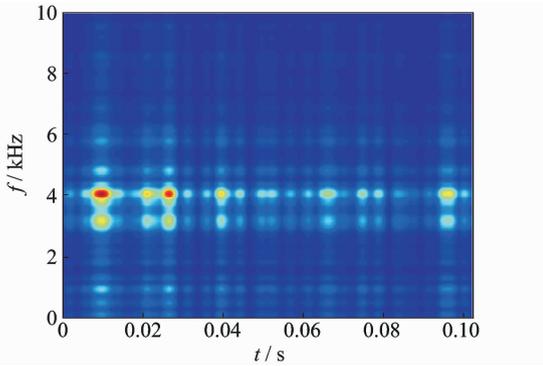


图 9 第 1 子空间时频分布

Fig. 9 Time-frequency distribution of the first subspace

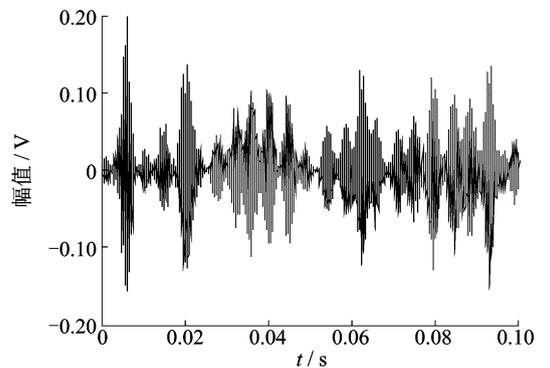


图 12 提取的特征波形

Fig. 12 Waveform of extraction feature

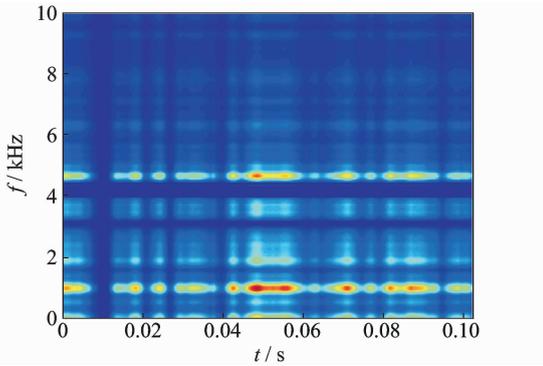


图 10 第 2 子空间时频分布

Fig. 10 Time-frequency distribution of the second subspace

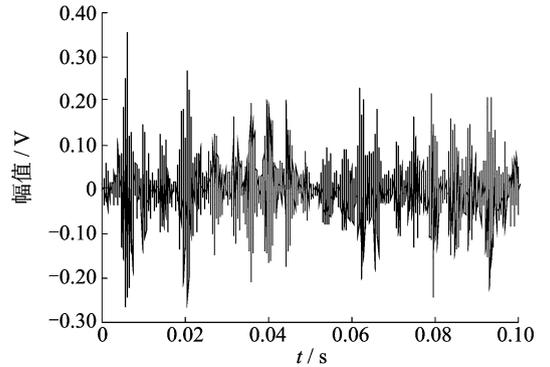


图 13 k=2 时提取的特征波形

Fig. 13 Waveform of extraction feature with k=2

性。为了进一步验证,选择 $k=2$ 时的特征分离做对比,所提取的特征波形如图 13 所示。

同图 12 提取的时域波形对比分析可见,当子空间维数减少时,冲击分量的波形中包含了其他的分量信息,导致提取的效果变差。对其子空间的时频分布进行对比分析,发现除包含主要周期分量外,还含有大量的其他频率分量,从而导致重构波形中含有更多的其他分量。由此可见,利用本算法能够很好地提取出有效的局部故障特征。

6 结 论

1) 带有正交性约束的 ONMF 方法用于单通道的源分离中,大大提高了所提取信息的独立性,能够更准确地描述源信号。

2) 根据振动信号的特点,引入核心一致性方法来评估观测矩阵的隐含子空间维数,可以提高 ONMF 的分解效果,获得较好的子空间时频分布。

3) 通过 STFT 变换得到原始混合信号的幅值谱,在低维子空间中经过选择及 ISTFT 重构,可以

有效提取出局部故障导致冲击特征的波形。因此,该方法对滚动轴承和齿轮箱等设备的早期故障检测具有重要的意义,具有一定的工程应用价值。

参 考 文 献

[1] Sze C J, Liao H Y M, Fan K C. A new image flux conduction model and its application to selective image smoothing[J]. IEEE Transactions on Image Processing, 2001, 10(2): 296-306.

[2] Tuzikov A V, Sheynin S A. Symmetry measure computation for convex polyhedra[J]. Journal of Mathematical Imaging & Vision, 2002, 16(1): 41-56.

[3] Liu Haining, Liu Chengliang, Huang Yixiang. Adaptive feature extraction using sparse coding for machinery fault diagnosis[J]. Mechanical Systems and Signal Processing, 2011, 25(2): 558-574.

[4] Dong Shaojiang, Tang Baoping, Zhang Yan. A repeated single-channel mechanical signal blind separation method based on morphological filtering and singular value decomposition[J]. Measurement, 2012, 45(8): 2052-2063.

[5] 刘婷婷, 任兴民. 独立分量分析在机械振动信号分离中的应用[J]. 振动、测试与诊断, 2009, 29(1): 36-41.
Liu Tingting, Ren Xingmin. Applying independent component analysis (ICA) to obtaining better blind separation of mechanical vibration signals[J]. Journal of Vibration, Measurement & Diagnosis, 2009, 29(1): 36-41. (in Chinese)

[6] 周晓峰, 杨世锡, 甘春标. 一种旋转机械振动信号的盲源分离消噪方法[J]. 振动、测试与诊断, 2012, 32(5): 714-717.
Zhou Xiaofeng, Yang Shixi, Gan Chunbiao. A blind source separation denoising method of rotating machinery vibration signal[J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(5): 714-717. (in Chinese)

[7] 刘维湘, 郑南宁, 游屈波. 非负矩阵分解及其在模式识别中的应用[J]. 科学通报, 2006, 51(3): 241-250.
Liu Weixiang, Zheng Nanning, You Qubo. Non negative matrix factorization and its applications in pattern recognition[J]. Chinese Science Bulletin, 2006, 51(3): 241-250. (in Chinese)

[8] 王科俊, 左春婷. 非负矩阵分解特征提取技术的研究进展[J]. 计算机应用研究, 2014, 31(4): 970-975.
Wang Kejun, Zuo Chunting. Improvements of non-negative matrix factorization for image extraction[J]. Application Research of Computers, 2014, 31(4): 970-975. (in Chinese)

[9] Mehmood A, Damarla T, Sabatier J. Separation of human and animal seismic signatures using non-negative matrix factorization[J]. Pattern Recognition Letters, 2012, 33(16): 2085-2093.

[10] Li Bing, Zhang Peilin, Liu Dongsheng, et al. Feature

extraction for rolling element bearing fault diagnosis utilizing generalized S transform and two-dimensional non-negative matrix factorization[J]. Journal of Sound and Vibration, 2011, 330(10): 2388-2399.

- [11] Wang Qinghua, Zhang Youyun, Cai Lei, et al. Fault diagnosis for diesel valve trains based on non-negative matrix factorization and neural network ensemble[J]. Mechanical Systems and Signal Processing, 2009, 23(5): 1683-1695.
- [12] 李兵, 高敏, 张旭光. 用形态梯度法与非负矩阵分解的齿轮故障诊断[J]. 振动、测试与诊断, 2014, 34(2): 295-300.
Li Bing, Gao Min, Zhang Xuguang. Feature extraction for engine fault diagnosis by utilizing adaptive multi-scale morphological gradient and non-negative matrix factorization[J]. Journal of Vibration, Measurement & Diagnosis, 2014, 34(2): 295-300. (in Chinese)
- [13] 李兵, 米双山, 刘鹏远. 二维非负矩阵分解在齿轮故障诊断中的应用[J]. 振动、测试与诊断, 2012, 32(5): 836-840.
Li Bing, Mi Shuangshan, Liu Pengyuan. Applying two-dimensional non negative matrix factorization to gear fault diagnosis[J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(5): 836-840. (in Chinese)
- [14] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755): 788-791.
- [15] Cichocki A, Zdunek R, Phan A H, et al. Nonnegative matrix and tensor factorizations[M]. West Sussex: John Wiley & Sons Inc, 2009: 131-142.
- [16] Ding Chris, Li Tao, Peng Wei, et al. Orthogonal non-negative matrix t-factorizations for clustering[C] // Proceedings of the 12th Association for Computing Machinery SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, USA: ACM Press, 2006: 126-135.
- [17] Li Zhao, Wu Xindong, Peng Hong. Nonnegative matrix factorization on orthogonal subspace[J]. Pattern Recognition Letters, 2010, 31(9): 905-911.
- [18] Cichocki A, He Zhaoshui. An efficient K-hyperplane clustering algorithm and its application to sparse component analysis[J]. Lecture Notes in Computer Science, 2007, 4492(1): 1032-1041.



第一作者简介: 梁霖, 男, 1973年1月生, 博士、副教授。主要研究方向为机械设备故障诊断技术。曾发表《冲击故障特征提取的非线性流形学习方法》(《西安交通大学学报》2009年第43卷第11期)等论文。

E-mail: lianglin@xjtu.edu.cn.