

DOI:10.3969/j.issn.1001-4551.2015.10.003

基于 EMD 和 PCA 的滚动轴承故障 信号特征提取研究*

张颖¹, 马波^{1*}, 张明¹, 杨鲁伟², 杨俊玲²

(1. 北京化工大学 诊断与自愈工程研究中心, 北京 100029;

2. 中国科学院 低温工程学重点实验室(理化技术研究所), 北京 100190)

摘要: 针对滚动轴承故障信号具有非平稳性的问题, 对滚动轴承非平稳信号特征提取问题进行了研究, 提出了基于经验模态分解 (EMD) 和主成分分析 (PCA) 的滚动轴承故障信号特征提取方法。运用经 EMD 对滚动轴承故障信号进行分解, 得到了多个本征模态分量 (IMF), 计算了每个 IMF 的总能量值, 取能量集中的前 6 层 IMF, 将每层 IMF 频率集中的部分等分成多段, 计算每段能量值, 作为该故障的特征值。运用 PCA 对特征值进行了维度缩减, 将高维度的特征值降低为低维度的特征值, 计算了特征值的累计贡献率, 取累计贡献率达到 80% 的前多个特征值作为每组故障的特征值。研究表明, 该方法可有效提取滚动轴承故障信号特征, 实现高维特征维度缩减。

关键词: 经验模态分解; 主成分分析; 滚动轴承; 特征提取; 非平稳信号

中图分类号: TH133.3; TH165.3

文献标志码: A

文章编号: 1001-4551(2015)10-1284-06

Feature extraction of rolling bearing based on EMD and PCA

ZHANG Ying¹, MA Bo¹, ZHANG Ming¹, YANG Lu-wei², YANG Jun-ling²

(1. Diagnosis and Self-Recovery Engineering Research Center, Beijing University of Chemical Technology, Beijing 100029, China; 2. Key Laboratory of Cryogenics, TIPC, CAS, Beijing 100190, China)

Abstract: Aiming at the problem that fault signal of rolling bearing was non-stationary, aiming at which, how to extract feature of non-stationary signal was studied and the methods based on empirical mode decomposition (EMD) and principal component analysis (PCA) was put forward. Signal was decomposed into several IMFs by means of EMD. The energy of every intrinsic mode function (IMF) was calculated, and some IMFs with greater energy was chosen. Every IMF was divided into several segments according to the range the frequency, and energy of every segment was calculated as the feature value. PCA was used to reduce the dimensions of feature value. Accumulative contribution of feature value was calculated and the first few feature values whose accumulative contribution rate reached 80% was chosen as the final feature value. The results indicate that the method is valid for the feature extraction of rolling bearing, and dimension reduction of high dimensional feature.

Key words: empirical mode decomposition (EMD); principal component analysis (PCA); rolling bearing; feature extraction; non-stationary signal

0 引 言

滚动轴承是离心泵重要部件, 在持续不断的交变

载荷作用下, 极易发生故障, 一旦发生故障, 轻则引起停机影响生产, 重则发生爆炸事故。因此, 实现离心泵滚动轴承故障诊断具有重要意义。

收稿日期: 2015-05-27

基金项目: 国家青年科学基金资助项目(51305020); 国家重点基础研究发展计划(“973”计划)资助项目(2012CB026000); 国家高新技术研究发展计划(“863”计划)资助项目(2014AA041806)

作者简介: 张颖(1992-), 女, 天津人, 主要从事设备故障诊断方面的研究. E-mail: ying.zhang1030@foxmail.com

通信联系人: 马波, 男, 副研究员. E-mail: mabo@mail.buct.edu.cn

滚动轴承发生故障后,经过故障点时产生瞬时冲击,使信号呈现非平稳性^[1]。针对非平稳信号特征提取的问题,国内外进行了大量的研究。文献[2]运用小波包变换提取滚动轴承故障信号特征,并和神经网络方法相结合,实现滚动轴承故障诊断。文献[3]将小波变换和能量算子结合提取气阀故障特征,实验证明该方法能较好地区分气阀不同故障,并表现出抑噪特性。文献[4]将连续小波变换和决策树分类方法结合实现了电动离心泵的故障诊断。文献[5]通过分析对比多种时频分析方法,验证 Hilbert 变换提取非平稳信号上的优势。现有的大多数非平稳信号处理方法,如小波变换、小波包变换等,在处理非平稳信号时,由于受到海森伯测不准原理约束,具有局限性。

1998 年, Hilbert-Huang 提出的一种新的时频分析方法—经验模态分解,从原理上分析该方法不受测不准原理约束,且该方法具有自适应性,使非平稳信号的特征提取不受限于窗函数的选取,已经在很多方面已经得到广泛的应用。文献[6]运用经验模态分解方法(empirical mode decomposition, EMD)和 AR 模型提取滚动轴承故障信号特征,并通过试验验证了该方法的有效性。文献[7]运用 EMD 提取故障信号特征,神经网络对故障进行诊断,试验验证了该方法的可有效实现滚动轴承智能故障诊断。文献[8-10]提出基于 EMD 和峭度的 Hilbert 包络解调方法,用于诊断滚动轴承故障,并验证了该方法的有效性。但是,这些方法在提取故障特征时,得到的特征值维度较高,其中包含很多无效的特征值,为了去除冗余的特征值,降低计算复杂度,提取有效特征,提高诊断效率,可选用主成分分析方法(principal component analysis, PCA)方法解决上述问题^[11-12]。

本研究拟将 EMD 和 PCA 方法相结合,运用 EMD 方法去除信号中的噪声,提取出有用的故障信号特征,PCA 方法去除冗余的特征值,并提取出主成分,以提高诊断效率,降低故障诊断计算复杂度。

1 EMD 方法

经验模式分解方法认为,任何信号由不同的固有模式、简单的振荡信号组成。线性和非平稳模式的零交叉点和极值点的数量都是相同的。EMD 分解的目的就是解决瞬时频率的问题。

任何一个获取的特征信号都可以分解为多个固有模态函数(IMF),同时每个 IMF 也有自己的要求,必须

满足以下几个定义^[13]:

(1)在整个实数域或复数域中,极值与零交叉点的数量顶多相差一个。

(2)在函数上的任一点,信号的上、下包络线所构成的平均包络值必须为零。

上、下包络平均值为 m_1 ,信号 $x(t)$ 和中间的差异 m_1 是第一部分:

$$x(t) - m_1 = h_1 \quad (1)$$

理论上, h_1 是 $x(t)$ 的第一部分,如果 h_1 是固有模态函数。

如果 h_1 不是,那么有:

$$h_1 - m_{11} = h_{11} \quad (2)$$

停止筛选直到 h_k 是本征模量时,有:

$$h_{1(k-1)} - m_{1k} = h_{1k} \quad (3)$$

并可表示为:

$$c_1 = h_{1k} \quad (4)$$

从 $x(t)$ 信号中分离 c_1 , 可得:

$$r_1 = x(t) - c_1 \quad (5)$$

再将 r_1 作为原始信号,重复上述过程,得出第 2 个本征量。依此类推,可得:

$$r_1 - c_2 = r_2; r_2 - c_3 = r_3; \cdots; r_{n-1} - c_n = r_n; \quad (6)$$

这里,假设原始信号为 $x(t)$ 。通过上式综合得出:

$$x(t) = \sum_{j=1}^n c_j + r_n \quad (7)$$

即:

$$x(t) = \sum_{i=1}^n imf_i(t) + r_n(t) \quad (8)$$

式中: $imf_i(t)$ — $x(t)$ 经验模态分解得到的第 i 个固有模态函数(IMF); $r_n(t)$ —去除 n 个固有模态函数(IMF)后所剩的剩余信号。

直到 $r(t)$ 极限值小于两个或者 $|r(t)|$ 很小时可结束分解。

式(7,8)以上过程体现了 EMD 分解的完备性。

2 PCA 方法

主成分分析(PCA)主要用于缩减数据的维度。一般情况下,如果系统有 m 个变量,就可以用 k 个不相关主成分($k < m$)来表示样本之间的差异,然后就可以把原 m 个变量的变化用 k 个主成分来代替。这就将原数据空间分成了两个完全不相关的空间,一个是由 k 个主成分构成的主元子空间,另一个是由其他剩余变量构成的残差子空间,主元子空间用来描述整个系统变化在理想的情况下,而随机噪声包含在残差子空

间^[14]。

分解样本数据矩阵 X :

$$X = U \sum P^T \quad (9)$$

其中: $U \in \mathfrak{R}^{n \times n}$, $P \in \mathfrak{R}^{m \times m}$,

$$\sum = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_m \\ 0 & 0 & \cdots & 0 \end{bmatrix} R^{n \times m} \quad (10)$$

式中: $UU^T = I$, $VV^T = I$; $\sigma_i (i = 1, 2, \dots, m)$ —矩阵 X 的奇异值, 且 $\sigma_1 > \sigma_2 > \cdots > \sigma_m$ 。从而可得:

$$X^T X = (U \sum P^T)^T (U \sum P^T) = (P \sum) I (\sum P^T) = P \Lambda P^T \quad (11)$$

$$\text{式中: } \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_m^2 \end{bmatrix} \in$$

$\mathfrak{R}^{m \times m}$ —特征矩阵; $\lambda_i = \sigma_i^2 (i = 1, 2, \dots, m)$ —特征值; P —所求的特征向量。

本研究进行维度缩减的主要步骤如下:

(1) 从非平稳信号中提取 n 个特征值 $X_i (i = 1, 2, 3, \dots, n)$, 特征值 X_i 维度为 $N \{X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{iN}\} (j = 1, 2, \dots, N)$

(2) 为特征值建立样本矩阵 M :

$$M_{ij} = \frac{X_{ij} - \frac{\sum X_{ij}}{N}}{\sqrt{\sum \left(X_{ij} - \frac{\sum X_{ij}}{N} \right)^2}} \quad (12)$$

(3) 根据样本矩阵 M , 求协方差矩阵;

$$M' = M \times M^T \quad (13)$$

(4) 计算协方差矩阵 M' 的特征值和特征向量。

(5) 计算协方差矩阵 M' 特征值的绝对值并按值得大小进行降序排列, 同时调整相应特征向量的顺序, 得到矩阵 P 。

(6) 计算特征矩阵:

$$K = P \times M \quad (14)$$

(7) 计算特征值得累计贡献率, 根据累计贡献率判定特征矩阵 K 的前 n 行即为对多特征参数处理后最能代表故障征兆的 n 个特征参数。

3 基于 EMD 和 PCA 的特征提取方法

本研究采取 EMD 和 PCA 相结合的方法, 提取滚

动轴承典型的故障特征的流程如图 1 所示。

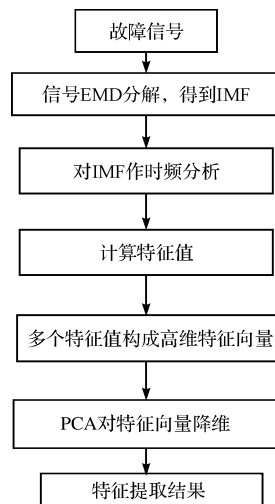


图 1 基于 EMD 和 PCA 的特征提取方法流程

(1) 用传感器采集滚动轴承和轴系故障的加速度信号, 随机选取各状态下的样本信号作为分析信号。

(2) 对分析信号分别进行 EMD 分解, 得到多层 IMF 函数和一层剩余函数 $r(c_1, c_2, c_3, c_4 \dots c_n)$, 不同的 IMF 分量包含着不同的原始信号的部分信息并且频率分布由高到低, 因此, 信号能够在不同的分辨率下准确地显示出来。

(3) 对每一个 IMF 分量进行傅里叶变换, 得到每一层 IMF 函数的频谱图, 频谱图能够反映信号中的不同频率成分的能量。

(4) 计算每一层 IMF 的能量值, 根据计算选择大部分能量所在的前 n 层 IMF 函数, 根据频率成分分布的不同把前 n 层 IMF 函数对应的各层频谱图分成 M 段, 对每一段的能量值进行计算, 得到一个 $N \times M$ 的高维度特征向量。能量值计算, 如下式所示:

$$E = \sum_{i=1}^n A_i^2 \quad (15)$$

式中: E —能量值, n —IMF 的频率值数, A —幅值。

(5) 为了减小计算负担, 找出最能代表故障特征的特征值, 本研究运用 PCA 方法对上述特征向量进行维度缩减, 针对高维度的故障特征的特征值建立样本矩阵, 对样本矩阵的特征值和特征向量分别进行计算, 如果所得到的前 n 个特征的累计贡献率达到 85%, 即可认为该组特征的前 n 个特征具有代表性, 可代表整组的特征。

4 实验及结果分析

本次试验所用的试验台如图 2 所示。其包含了两个支撑座, 双支承转子试验台, 左端为电机。传感器布

置如图 3 所示。

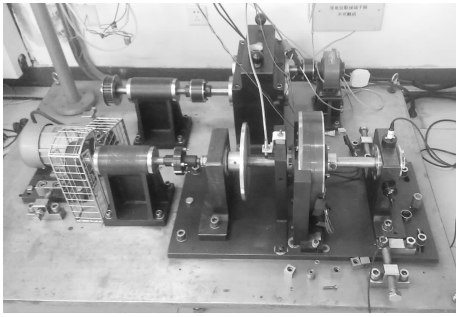


图 2 滚动轴承故障试验台



图 3 传感器布置

该次试验所用轴承型号为 NU205EM,轴承转速为 1 200 r/min,采样频率为 25.6 kHz,每组故障的采样点数为 450 000 个左右,该实验截取其中的 5 000 个点进行分析,转速为 1 200 r/min。该实验所设计的试验故障为滚动轴承外圈故障、滚动轴承内圈故障、滚动轴承滚动体故障。

笔者以滚动轴承为外圈故障为例,说明 EMD 分解和特征值计算过程。运用 EMD 方法将滚动轴承外圈故障信号分解为 11 个 IMF 分量和一个剩余函数如图 4 所示。

本研究对上述 11 层 IMF 函数进行傅里叶变换之后得到频谱如图 5 所示。计算 11 个 IMF 分量的能量值及其占总能量的百分比结果如表 1 所示。

表 1 前 5 层 IMF 能量及占总能量的百分比

IMF 分量	能量值	百分比/(%)
1	1 524	61.49
2	499	20.13
3	256	10.33
4	86	3.5
5	43	1.75

通过表 1 可以看出前 5 层的 IMF 函数能量占总能量的 97.2%,能量主要集中在前 5 层 IMF 函数中,因此

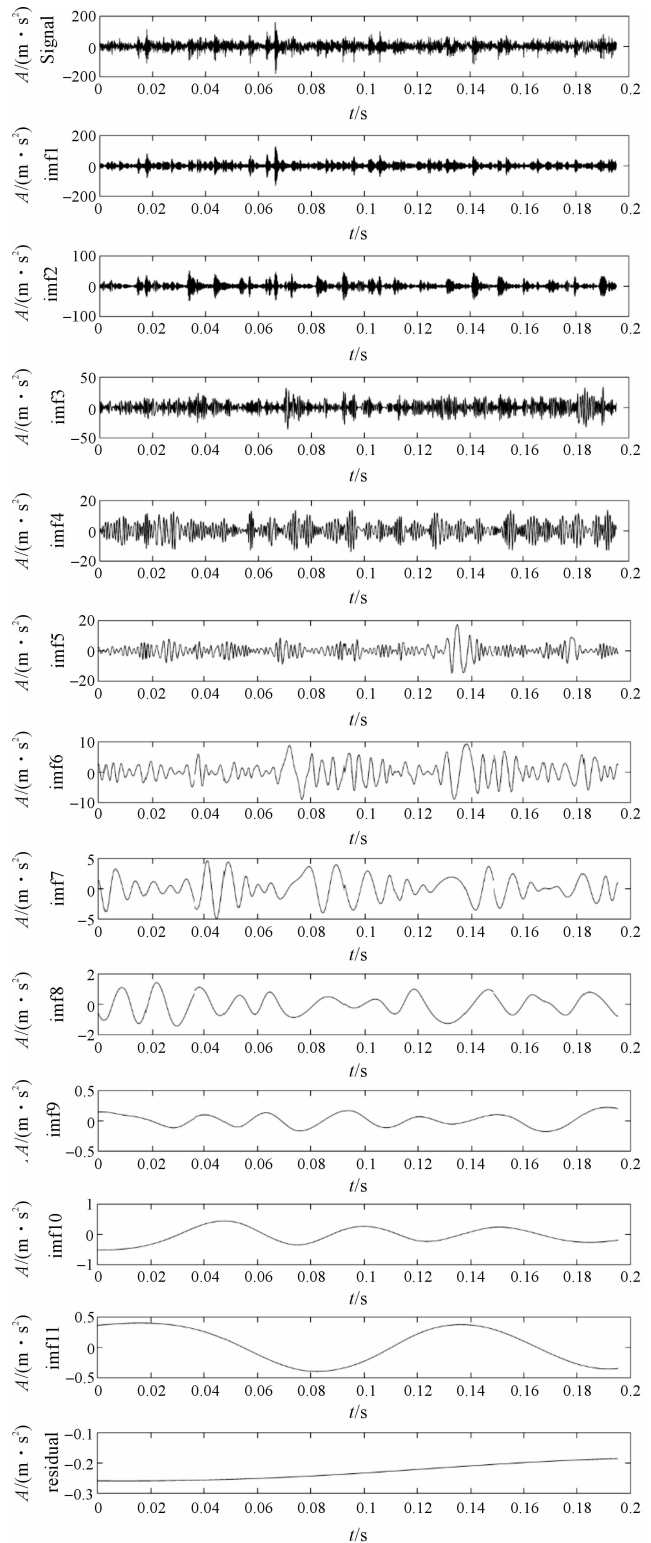


图 4 滚动轴承外圈故障原始波形及 EMD 分解结果

笔者选取经过傅里叶变换的前 5 层 IMF 函数计算特征值。从频谱图中可以看出能量主要分布在 0 ~ 12 000 Hz 频段,将每层 IMF 数据 0 ~ 12 000 Hz 频段均分成 6 段,即 0 ~ 2 000 Hz,2 000 Hz ~ 4 000 Hz,4 000 Hz ~ 6 000 Hz,6 000 Hz ~ 8 000 Hz,8 000 Hz ~ 10 000 Hz,10 000 Hz ~ 12 000 Hz,求每一段的能量值,每层 IMF

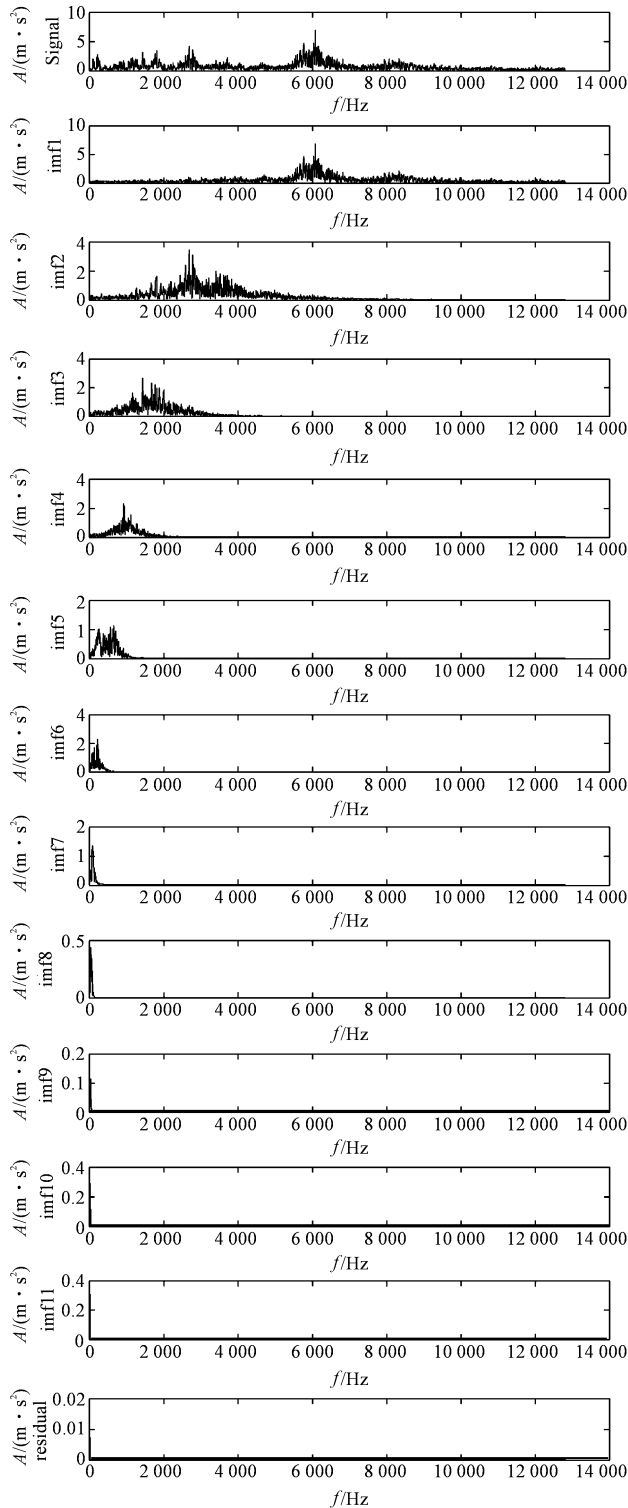


图 5 IMF 分量及剩余函数的傅变频谱

提取 6 个特征值, 每组故障提取 6×5 个特征值, 滚动轴承外圈故障特征值如图 6 所示。

本研究通过计算得到 30 个特征值, 计算每组故障得到的每个特征值的累计贡献率, 根据累计贡献率, 选择最能代表该组故障特征的特征值进行后序计算。滚动轴承外圈故障得到的特征值及其相应的累计贡献率计算结果如表 2 所示。

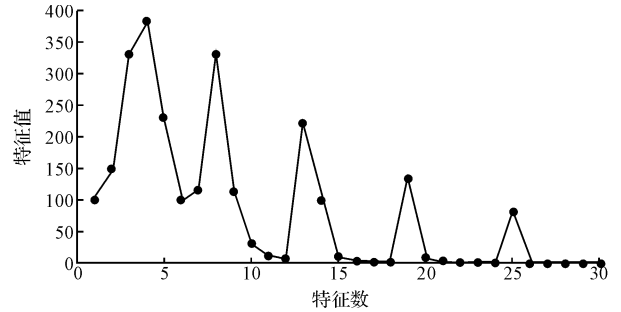


图 6 滚动轴承外圈故障特征值

表 2 滚动轴承外圈故障的特征值及累计贡献率

序号	特征值	累计贡献率/(%)	序号	特征值	累计贡献率/(%)
1	-143.6	4.58	16	34.0	96.86
2	546.3	54.66	17	30.0	96.89
3	-66.2	55.88	18	-44.8	97.56
4	29.3	55.91	19	-22.4	97.83
5	211.3	62.66	20	-72.0	99.23
6	290.3	79.48	21	-9.3	99.35
7	-70.0	80.81	22	-15.3	99.54
9	191.3	91.83	24	7.4	99.64
10	120.0	93.73	25	6.0	99.66
11	141.5	96.51	26	-3.5	99.73
12	35.6	96.58	27	14.2	99.74
13	0.4	96.62	28	-5.7	99.82
14	21.8	96.63	29	-5.5	99.91
15	-14.5	96.80	30	-5.7	100

一般认为, 累计贡献率达到 85% 左右的前 n 个特征值即可作为代表本组故障的特征值, 从表 2 可以看出, 该组特征值的前 8 个特征的累计贡献率达到 85%, 即前 8 个特征值可作为代表滚动轴承外圈故障缩减后的特征值, PCA 方法有效地实现了特征值维度缩减。

本研究尝试将每组故障的前 3 维特征表示在 3 维坐标系中, 结果如图 7 所示。从图 7 中可以看出, 同一种故障的特征值点聚集在一起, 不同故障的前三维特征值被清晰分开。由此可以得出结论, 前三维特征数据即可将不同的故障类型进行区分, PCA 方法能去除冗余特征值, 有效实现故障特征维度缩减。

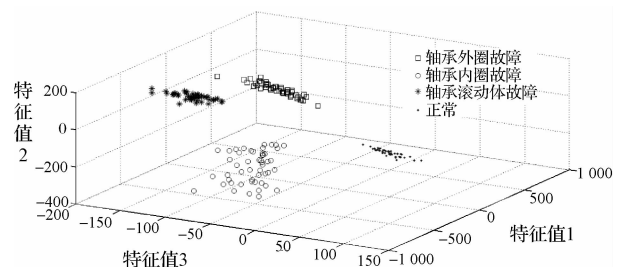


图 7 4 种工况特征值维度缩减后在三维坐标中的表示

5 结束语

本研究针对滚动轴承典型故障信号的非平稳性,运用EMD方法提取信号特征,然后运用PCA方法去除冗余的特征值。主要结论如下:

(1)EMD方法分解非平稳信号,可以得到多个频率由高至低分布的IMF分量,选择能量集中的IMF,可有效去除低频噪声干扰,提取故障信号的高频分量,计算得到信号特征值。

(2)PCA方法可依据累计贡献率,选择少数特征值,来代替高维的特征值,该方法能够保留较完整的故障特征信息,同时PCA对滚动轴承典型故障特征分类有较好的效果。

参考文献(References):

- [1] GELMAN L, MURRAY B, PATEL T H, et al. Vibration diagnostics of rolling bearings by novel nonlinear non-stationary wavelet bicoherence technology[J]. **Engineering Structures**, 2014(80):514-520.
- [2] 刘乐平,林凤涛. 基于小波包特征向量与神经网络的滚动轴承故障诊断[J]. **轴承**, 2008(4):46-48.
- [3] 魏中青,马波,么子云,等. 运用小波包变换与能量算子的气阀故障特征提取[J]. **振动. 测试与诊断**, 2011, 31(1):50-54.
- [4] MURALIDHARAN, SUGUMARAN V. Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump[J]. **Measurement**, 2013, 46(1):353-359.
- [5] 向玲,唐贵基,胡爱军. 旋转机械非平稳振动信号的时频分析比较[J]. **振动与冲击**, 2010(2):42-45.

- [6] JUNSHENG C, DEJIE Y, YU Y. A fault diagnosis approach for roller bearings based on EMD method and AR model[J]. **Mechanical Systems and Signal Processing**, 2006, 20(2):350-362.
- [7] ALI J B, FNAIECH N, SAIDI L, et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals[J]. **Applied Acoustics**, 2015(89):16-27.
- [8] 周浩,贾民平. 基于EMD和峭度的Hilbert包络解调在滚动轴承故障诊断中的应用分析[J]. **机电工程**, 2014, 31(9):1136-1139.
- [9] 赵月刚. 4N6Z凝结水泵轴承故障分析及处理[J]. **流体机械**, 2011, 39(10):55-56.
- [10] 王逍梦,任章. 基于小波变换的二阶传热传感器故障诊断研究[J]. **兵工自动化**, 2013(5):78-82.
- [11] SHAO R, HU W, WANG Y. The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform[J]. **Measurement**, 2014, 54(10):118-132.
- [12] ALVAREZ D G, FUENTE M J, SAINZ G I. Fault detection and isolation in transient states using principal component analysis[J]. **Journal of Process Control**, 2012, 22(3):551-563.
- [13] YU D, CHENG J, YANG Y. Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings[J]. **Mechanical systems and signal processing**, 2005, 19(2):259-270.
- [14] 刘永斌,何清波,孔凡让,等. 基于PCA和SVM的内燃机故障诊断[J]. **振动. 测试与诊断**, 2012, 32(2):250-255.

[编辑:李辉]

本文引用格式:

张颖,马波,张明,等. 基于EMD和PCA的滚动轴承故障信号特征提取研究[J]. **机电工程**, 2015, 32(10):1284-1289.

ZHANG Ying, MA Bo, ZHANG Ming, et al. Feature extraction of rolling bearing based on EMD and PCA[J]. **Journal of Mechanical & Electrical Engineering**, 2015, 32(10):1284-1289.

《机电工程》杂志: <http://www.meem.com.cn>