

文章编号 1004-924X(2006)06-1093-07

# 设备噪声监测中主分量的特征表示

何清波,冯志华,孔凡让

(中国科学技术大学 精密机械与精密仪器系,安徽 合肥 230027)

**摘要:**研究了设备噪声监测中主分量特征表示的提取及应用。在设备噪声时域和频域统计模式特征基础上,通过主分量分析探讨了设备噪声模式的主分量特征表示方法,引入相关度的概念分析了主分量特征表示对设备状态的表征能力,提出了选择有效维数的主分量特征表示进行设备噪声监测的方案。通过在EQ6100型发动机上预先模拟四种连杆轴承磨损故障,测取噪声信号,实例分析显示了低维主分量特征表示可以有效表征设备状态。实验最后对测试集样本进行状态监测得到了100%的准确率,表明了主分量特征表示用于设备噪声监测的有效性。

**关键词:**设备噪声;状态监测;主分量分析;连杆轴承;磨损

**中图分类号:**TH17; TP206; TK418 **文献标识码:**A

## Principle component representations for machine noise monitoring

HE Qing-bo, FENG Zhi-hua, KONG Fan-rang

(*Department of Precision Machinery and Precision Instrumentation,  
University of Science and Technology of China, Hefei 230026, China*)

**Abstract:** The extraction and application of principle component representations were studied for machine noise monitoring. On the base of the time- and frequency-domain statistical features extracted from the machine noise signal, the principle component representations of machine noise pattern were explored, the ability of principle component representations to represent the characteristic of machine condition by introducing the idea of correlation was analyzed, and the machine noise monitoring scheme using the effective low-dimensional principle component representations was proposed. Through previously designing four wearing faults on the connecting rod bearing of EQ6100 model gasoline engine, experimental analysis shows that the low-dimensional principle component representations of engine noise can be conveniently and effectively used for representing the machine conditions. Finally, the excellent result with the accuracy of 100% was obtained by monitoring the conditions of the testing samples, which demonstrates the availability of principle component representations for machine noise monitoring.

**Key words:** machine noise; condition monitoring; principle component analysis; connecting rod bearing; wearing

## 1 引言

机械设备噪声中包含着丰富的设备运行状态信息,设备发生故障时会在噪声变化中得到反映,故适合通过噪声对设备进行状态监测和故障诊断<sup>[1-3]</sup>。同时,声学监测和故障诊断技术具有非接触测量、简便易行、在线测量和不需在设备上增加附加质量影响设备运行等特点<sup>[2-3]</sup>。因此对用振动方法较难诊断的设备内部机件故障,比如发动机连杆轴承磨损故障<sup>[4]</sup>,用噪声进行诊断是一种理想手段。

信号的时域、频域和时频域特征提取技术得到很大发展,人们总是希望提取较多的有效特征参数来进行故障分类和诊断。但是采用特征参数太多会影响模式分类的计算代价甚至分类精度<sup>[5]</sup>。而且它们所反映的状态规律性、敏感性和在模式空间的聚类性、可分性并不相同,在特征分析基础上提取规律性好、敏感性强的特征表示则显得非常需要且极具挑战性<sup>[5-6]</sup>。

主分量分析(Principle component analysis, PCA)是一种通用的多元统计方法,在信息压缩和消除数据间的相关性方面非常有效,被有效用于人脸、字符等特征提取中<sup>[7-8]</sup>。近几年来也有研究者用 PCA 的方法来做状态监测的工作,A. Malhi<sup>[5]</sup>提出一种基于 PCA 的特征抽取方案来保证从多个原始特征中选择最有效的特征来进行轴承状态监测的故障分类,N. Baydar<sup>[9]</sup>把基于 PCA 的多元统计方法用于建立一个正常状态模型进行齿轮箱状态监测。本文致力于研究设备状态监测中低维、有效的新特征提取,利用 PCA 探讨了设备噪声模式的主分量特征表示方法,引入相关度的概念分析了主分量特征对设备状态的表征能力,提出了利用低维主分量特征进行设备噪声监测的方案。通过发动机连杆轴承的实例分析,表明了所提出的状态监测方法非常有效,且计算方便。

## 2 设备噪声统计模式特征

设备噪声信号是一种复杂的随机过程,很难用确定的时间函数表达,进行状态监测和故障诊

断时需要进行特征分析,提取能够反映状态信息的模式特征。为了消除操作环境、传感器因素对数据的影响,使信号分析结果有一个客观、公正的标准,必须对测量信号进行预处理。这里采用均值-方差标准化法,如下面分析,预处理后的信号具有零均值和单位方差的标准。

所测设备噪声信号转换成一维离散数据 $\{y_1, y_2, \dots, y_n\}$ ,  $n$  为样本长度,则数据样本均值为:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (1)$$

样本的方差为:

$$S = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2)$$

则均值-方差标准化后的样本数据为:

$$y_i' = \frac{y_i - \bar{y}}{\sqrt{S}}, \quad (3)$$

经过预处理后,受背景噪声的影响,不同状态的设备噪声信号一般在时域上差别不大,频域上的能量分布则能够在一定程度上反映出状态间的区别。通过分析不同状态信号在频域上的表现,可以发现一些典型的特征频带,反映了不同程度故障的发展<sup>[2]</sup>。本文基于时频分析和频域分析,提取设备噪声信号的时域和频域统计特征指标来全面表达其模式特征:

(1)时域统计特征。首先采用非平稳信号分析工具小波包变换<sup>[2,10]</sup>来对典型特征频带的小波包系数进行重构,以消减背景噪声、得到敏感的状态特征信息。然后提取绝对均值,最大峰值,有效值,方根幅值,方差,斜度,峭度,峰值因子,波形因子,脉冲因子,裕度因子等 11 个时域统计参数,具体计算公式请参见参考文献[1]。

(2)频域统计特征。根据不同状态信号具有不同的频率结构,将预处理后的信号变换到频域,分成 8 个频带分别计算各频带的功率谱能量。在频域提取各频带的功率谱能量以及它们的均值,频率重心,方差,谱峰,峭度等 13 个统计特征<sup>[6]</sup>。

这样总共提取出 24 个特征指标表示设备噪声的模式特征。图 1 展示了某发动机连杆轴承四种磨损状态下噪声样本的 24 维特征分布,其中横坐标是样本序号,从 1 开始每 50 个样本属于一种状态。从图中可见,这些特征对状态的表现并不不同,有的规律而稳定,有的则对状态不敏感、规律性不强。

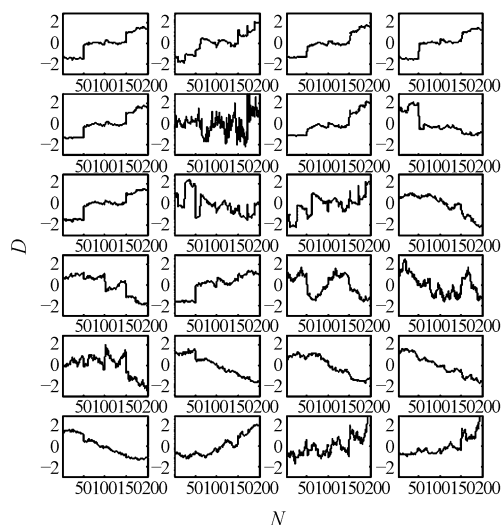


图 1 设备噪声统计模式特征图示

Fig. 1 Statistical pattern features of machine noise

### 3 噪声模式的主分量特征表示分析

前面得到噪声信号的 24 维特征表示,但也看到,它们所反映的状态规律性、敏感性并不相同。而状态监测和诊断需要获得对状态反映规律性好、敏感性强的特征来进行样本分类,所以要在上述特征分析基础上来提取这样的特征表示,这个工作可以靠经验来选择表现好的特征,但构建一个自动的特征选择或提取方案则显得尤其重要<sup>[5]</sup>,其中在已有特征基础上提取新的特征更有挑战性。

PCA 是一种有效的多元统计分析工具。其基本思想是寻找高维数据的一些主分量表示,这些分量具有最大方差,用它们表示原数据具有最小的均方误差。所以 PCA 是一种很好的数据压缩和特征提取方法。PCA 也同样可以用于设备噪声的特征提取中。

设有噪声模式向量矩阵  $\mathbf{X}$ , 其列向量  $x_i$  为某一状态样本对应的  $d$  维模式向量, 则可以得到该矩阵的协方差矩阵为:

$$\mathbf{R}_x = \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x}), \quad (4)$$

其中  $N$  是训练样本的数量,  $\bar{x}$  是各样本模式向量的均值向量:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (5)$$

对  $\mathbf{R}_x$  做特征值分析, 可以求解下式

$$\lambda v = \mathbf{R}_x v, \quad (6)$$

其中  $\lambda$  和  $v$  分别是  $\mathbf{R}_x$  的特征值和特征向量。最多可以得到  $d$  个特征值  $\lambda_i (i=1, 2, \dots, d)$ , 且  $\lambda_1 > \lambda_2 > \dots > \lambda_d$ , 对应的特征向量为  $v_i (i=1, 2, \dots, d)$ 。样本  $x_j$  投影到特征向量  $v_i$  上得到该方向相应的主分量特征:

$$y_{ij} = v_i^T (x_j - \bar{x}), \quad (7)$$

所有特征向量张成一个  $d$  维的正交空间,  $X$  投影到该正交空间得到相应的  $d$  维主分量特征表示。80% 以上的原始数据信息通常保留在前面少数几个主分量中, 所以可以仅用  $m (m < d)$  个特征向量组成一个正交子空间。该子空间相比原始模式空间具有降维的优势, 由它得到的  $m$  维主分量特征表示可以近似表示原始数据结构。

第一个主分量特征表示具有最大的特征值, 即拥有数据变化的最大方差, 所以能够抓住各维原始模式矢量的典型变化特征, 可谓是反映状态特征的综合表示。对设备噪声统计模式特征, 各个特征指标都反映了设备状态的本质统计特征, 只是它们反映了不同侧面的敏感性, 显然综合这些不同侧面敏感性的主分量特征表示可以有效表征设备噪声的敏感模式特征。如图 2 所示, 为对图 1 的噪声模式特征提取出的六个主分量特征表示, 与图 1 对比可见, 第一个主分量表示抓住了各个特征对状态表征的优点, 较好的反映出了不同状态的区别特征, 而其他主分量反映的则是局部变化特征, 反映不同状态特征的能力较弱, 如第六个主分量已经毫无规律。

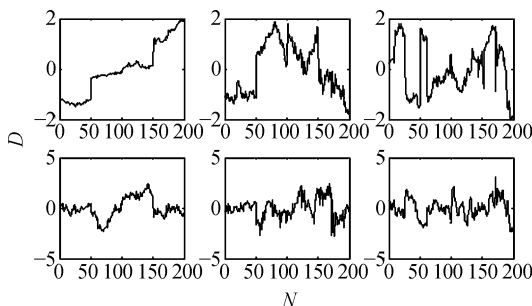


图 2 噪声模式的前六个主分量特征表示

Fig. 2 First six principle component representations of noise pattern

下面利用主分量特征表示与原始数据的相

关度来评估它们对状态特征的代表能力。设  $\mathbf{PC}_i$  ( $i=1, \dots, d$ ) 为第  $i$  个主分量特征表示,  $\mathbf{F}_j$  ( $j=1, \dots, d$ ) 为第  $j$  维原始模式特征向量, 则它们之间的相关度用相关系数的绝对值计算:

$$\rho_{ij} = \left| \frac{E\{[\mathbf{PC}_i - E(\mathbf{PC}_i)][\mathbf{F}_j - E(\mathbf{F}_j)]\}}{\sqrt{D(\mathbf{PC}_i)} \cdot \sqrt{D(\mathbf{F}_j)}} \right|, \quad (8)$$

其中  $E(\cdot)$  表示求期望值,  $D(\cdot)$  表示求方差。把每个主分量特征  $\mathbf{PC}_i$  与各维原始模式向量求相关度得到  $\{\rho_{i1}, \rho_{i2}, \dots, \rho_{id}\}$ , 表示为相关向量  $\boldsymbol{\rho}_i$ , 所有相关向量组成相关度矩阵  $[\rho_1^T, \rho_2^T, \dots, \rho_d^T]$ , 表示为  $\boldsymbol{\rho}$ 。

计算图 2 所示的六个主分量特征与 24 维原始模式特征的相关度矩阵, 如图 3 所示。由图 1、图 2、图 3 可见, 第一个主分量与大部分模式特征都具有很高的相关度, 说明它代表的是状态特征的普遍规律; 第二个主分量的相关度就小了很多, 其中它与第 6、16、17 个模式特征的相关度比较大, 说明它反映的是状态特征的次要规律以及局部变化特征; 第三个主分量则主要反映了第 10 和 11 两个模式特征的规律; 第四和第五两个主分量则分别反映的是第 15 和第 6 个模式特征的规律; 第六个主分量与每个模式特征的相关度都很低, 它反映了模式特征的局部变化信息, 如随机噪声一样, 对状态特征已经没有参考价值。

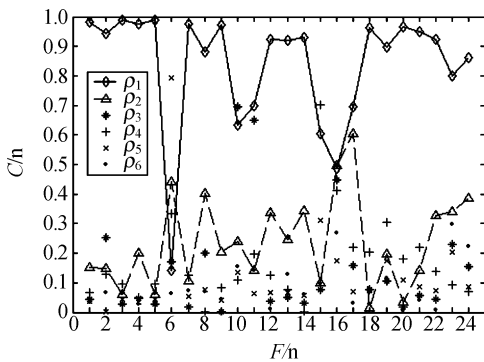


图 3 主分量特征表示与原始模式特征的相关度

Fig. 3 Correlation between the principle component representations and the original pattern features

综合评价主分量特征表示的状态特征表现能力, 可以采用每个主分量特征的相关向量  $\boldsymbol{\rho}_i$  的均值即  $E(\boldsymbol{\rho}_i)$  来测度。把图 2 所示六个主分量特

征的平均相关度与主分量对应的特征值列于表 1, 从表中数据可见, 平均相关度的分布很好的符合了上面的分析, 可以以此来评价主分量特征的代表状态的能力; 同时第一个主分量的平均相关度达到了 0.84, 其他主分量的平均相关度则基本都在 0.30 以下, 印证了第一个主分量特征具有的绝对优越性能。与主分量的特征值对比可知, 平均相关度与特征值的分布基本是一致的, 由此可得出, 在表征设备不同状态的时候, 最大的特征值反映了最大的数据变化信息, 也反映了最典型的状态变化信息。然而需要说明, 不能用特征值去评价主分量特征的状态特征表现能力。

表 1 主分量特征表示的平均相关度与特征值

Tab. 1 Mean correlations and eigenvalues of principle component representations

$\mathbf{PC}_i$	1	2	3	4	5	6
$E(\boldsymbol{\rho}_i)$	0.84	0.24	0.15	0.17	0.12	0.09
$\lambda_i$	17.8	1.93	1.38	1.21	0.92	0.33

## 4 噪声监测方案

首先说明试验情况。加机油口是向内燃机内部加机油的直接通道, 内部机器的运行状况也能通过此通道以声音的方式传出来。所以凭借经验, 可将塑料软管伸入加机油口, 通过声音变化诊断是否存在故障。在 EQ 6 100 型发动机上预先模拟连杆轴承故障, 将轴承与轴颈间隙按正常间隙 (0.04~0.098 mm, 表示为  $\Delta 1$ )、初期磨损 (0.15 mm, 表示为  $\Delta 2$ )、中期磨损 (0.30 mm, 表示为  $\Delta 3$ )、严重磨损 (0.50 mm, 表示为  $\Delta 4$ ) 设置, 分别进行实验。噪声信号的测取是在加机油口处用传声器监听, 放大后送入 B&K 声级计, 对声级计输出的信号进行分析处理。测试是以 A 计权的方式进行的。整个测试过程中, 转速保持在 1 500 r/min 左右, 采样频率为 4 000 Hz。除连杆轴承间隙设置不同外, 其他测试条件保持一致。

分别抽取四种故障间隙工况的噪声信号各 80 组进行分析, 其中 50 组用于训练, 30 组用于测试。样本集与其所属的工况状态的对比关系见表 2 所示。

表 2 样本与所属连杆轴承状态的对应关系

Tab. 2 Corresponding relations between the sample and the connecting rod condition

样本集	正常间隙	初期磨损	严重磨损	严重磨损
训练集 No.	1 - 50	51 - 100	101 - 150	151-200
测试集 No.	1 - 30	31 - 60	61 - 90	91-120

所测噪声信号样本都要进行均值-方差标准化预处理,图 4 所示是经过预处理的四种工况下的典型噪声时域信号,从图中很难辨别它们之间的区别,而从图 5 所示的频域图上来看,除 400 Hz 附近能量随连杆轴承间隙的增大而有所增加外,其他相差无几。把 400 Hz 附近频带作为一个故障状态的特征频带,按照前面所述方法,提取噪声信号的 24 个时频统计特征指标来全面表征设备的模式特征。前面图 1 所示的正是该实验中训练样本集的 24 维统计特征指标。

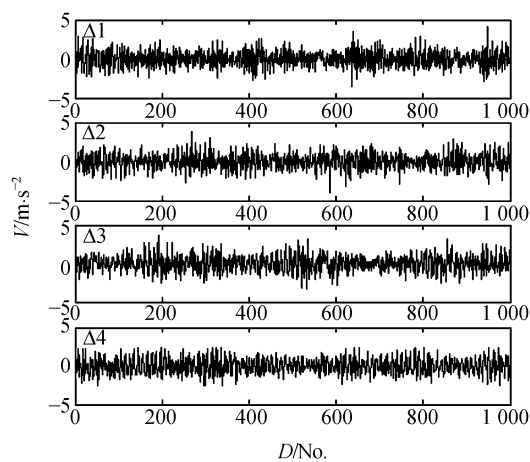


图 4 不同连杆轴承间隙下的噪声时域信号  
Fig. 4 Time domain analysis of signals on different CRB clearances

设备状态监测和故障诊断的实质是模式识别的问题<sup>[6]</sup>,基于上述分析,提出利用低维主分量特征表示的噪声监测方案。在监测样本状态之前,首先利用样本训练集来训练一个 PCA 子空间,然后提取训练样本集的主分量特征表示,通过分析它们对设备状态的表征能力,选择用于状态监测的主分量特征表示维数。这些工作前面已经做了分析,并得出该例中第一个主分量特征对表征状态特征具有非常优越的性能,所以选择第一主分量特征表示来进行状态监测。

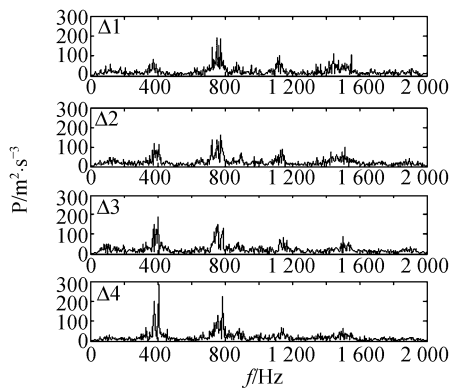


图 5 不同连杆轴承间隙下的噪声频域信号  
Fig. 5 Frequency domain analysis of signals on different CRB clearances

对测试样本也提取相应的一维主分量特征表示,它由测试样本投影到训练样本集模式矩阵  $\mathbf{X}$  的最大特征值对应的特征向量  $\mathbf{v}_1$  上得到。设测试样本的 24 维模式向量为  $\mathbf{x}$ ,则该一维主分量特征表示可以由下式得出:

$$\mathbf{y}_1 = \mathbf{v}_1^T (\mathbf{x} - \bar{\mathbf{x}}), \quad (9)$$

状态监测的关键是对测试样本进行状态分类。计算出训练集一维主分量特征表示在各状态类别的均值  $\text{Mean}_i (i=1, 2, 3, 4)$ ,以此代表各类别,比较测试样本与各类别之间的相似度,用其主分量特征与各类别之间的欧几里得距离来计算,即:

$$\delta_i(\mathbf{x}) = \|\mathbf{y}_1 - \text{Mean}_i\|, \quad (10)$$

其中  $\mathbf{y}_1$  是测试样本  $\mathbf{x}$  的一维主分量特征表示。通过判断与四个状态类别的距离大小,测试样本被分到具有最小距离的类,即与距离最小的状态类具有最大相似度。

现在来监测测试样本的状态,首先测试样本的 24 维模式特征分别用训练样本相应模式特征的均值和方差进行标定,然后用式(9)计算其一维主分量特征表示,如图 6 所示,其中虚线表示的是训练集一维主分量特征的各类别的均值。从图中可见,不同状态测试样本的投影分别分布在相应的四个状态的均值附近,这说明了一维主分量特征表示对设备状态类别表达的有效性。

从图 6 可直观地看出测试样本能够成功地分到所属的状态类别中,用式(10)计算测试样本与四个类均值的相似距离,图 7 是结果的柱状图表示。从中可以明显看出,正常间隙状态样本与正

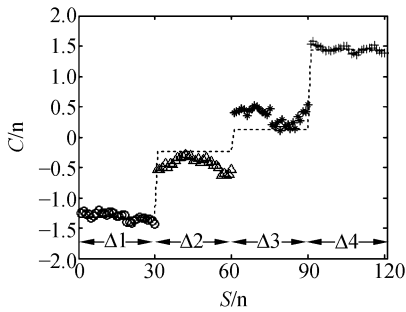


图 6 测试样本集的一维主分量特征表示

Fig. 6 First principle component representation of the testing samples

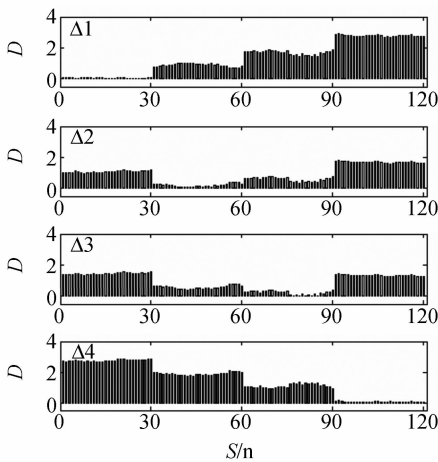


图 7 测试样本与四种状态类别的相似距离

Fig. 7 Similar distances between the testing samples and the four class conditions

常间隙类别的距离最小,初期磨损状态样本与初期磨损类别的距离最小,中期磨损状态样本与中期磨损类别的距离最小,严重磨损状态样本与严重磨损类别的距离最小,其他情况的距离值都比它们大。根据距离越小相似度越大的准则,全部测试样本得到了正确的分类结果,所以低维主分量特征表示可以非常有效地用于设备噪声的状态监测中。

## 5 结 论

设备噪声的时域和频域统计模式特征对设备状态表征的规律性和敏感性并不相同,在此基础上提取的低维主分量特征表示可以综合表征设备状态特征,具有规律性好、敏感性强的优良性能。

通过 PCA 可以自动提取主分量特征表示,可以通过分析主分量特征对设备状态的表征能力选择适合用于表征和监测设备状态的主分量特征维数,本文实例中通过分析选择了一维主分量特征表示。

利用低维主分量特征表示进行设备噪声监测的方案计算操作非常方便,通过在 EQ6100 型发动机上预先模拟四种连杆轴承磨损故障,实例分析表明了所提状态监测方法的有效性和方便性,对试验的测试集样本进行状态监测得到了 100% 的准确率。

## 参考文献:

- [1] 陈克兴,李川奇. 设备状态监测与故障诊断技术[M]. 北京:科学技术出版社,1991.  
CHEN K X, LI CH Q. *Machine condition monitoring and fault diagnosis technology* [M]. Beijing: Science and Technology Press, 1991. (in Chinese)
- [2] 贾继德,孔凡让,刘永斌,等. 发动机连杆轴承故障噪声诊断研究[J]. 农业机械学报,2005,36(6): 87-91.  
JIA J D, KONG F R, LIU Y B, *et al.* Noise diagnosis research based on wavelet packet and fuzzy c-clusters about connecting rod bearing fault [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2005,36(6): 87-91. (in Chinese)
- [3] 常西畅,周艳玲,陈进. 机械设备噪声故障诊断的新进展[C]. 北京:全国振动(诊断、模态、噪声与结构动力学)工程及应用学术会议论文集,2002: 140-143.  
CHANG X CH, ZHOU Y L, CHEN J. The advances of machine noise fault diagnosis [C]. *Proceeding of 2002 National Conference on Vibration Engineering and Application, Beijing*, 2002:140-143. (in Chinese)
- [4] 赵吉文,俞巧云,王建平,等. 往复式机械非平稳信号的混沌与分形[J]. 光学精密工程,2003,11(6):637-642.  
ZHAO J W, YU Q Y, WANG J P, *et al.* Study of fault signals from main bearing of internal-combustion engine using chaos and fractal theories [J]. *Optics and Precision Engineering*, 2003, 11(6): 637-642. (in Chinese)

- [5] MALHI A, GAO R. PCA-based feature selection scheme for machine defect classification [J]. *IEEE Transaction on Instrumentation and Measurement*, 2004, 53: 1517-1525.
- [6] 温熙森,胡芑庆,邱静. 模式识别与状态监控[M]. 长沙:国防科技大学出版社,1997.  
WEN X S, HU N Q, QIU J. *Pattern recognition and condition monitoring* [M]. Changsha: National University of Defense Technology Press, 1997. (in Chinese)
- [7] JOLLIFFE I T. *Principal component analysis*[M]. Springer, New York, 1986.
- [8] TURK M, PENTLAND A. Face recognition using eigenfaces [C]. *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, 1991: 586-591.
- [9] BAYDAR N, CHEN Q, BALL A, *et al.* Detection of incipient tooth defect in helical gears using multivariate statistics [J]. *Mechanical Systems and Signal Processing*, 2001, 15: 303-321.
- [10] FAN X, ZUO M J. Gearbox fault detection using Hilbert and wavelet packet transform [J]. *Mechanical Systems and Signal Processing*, 2006, 20: 966-982.

**作者简介:**何清波(1980—),男,河南台前人,中国科学技术大学精密机械与精密仪器系博士研究生,研究方向为盲信号处理、统计信号处理、设备状态监测;E-mail: heqb@ustc.edu.cn

孔凡让(1951—),男,安徽人,教授,博士生导师,中国科学技术大学工程科学学院副院长,研究方向为智能信息处理、振动分析、测控技术、设备状态监测与故障诊断等。