

文章编号:1671-1637(2009)05-0062-05

# 基于核主成分分析的小波尺度谱图像特征提取

孙丽萍,陈 果,谭真臻

(南京航空航天大学 民航学院,江苏 南京 210016)

**摘 要:**分析了转子不平衡、不对中、碰摩及油膜涡动的尺度谱图像特征,提出利用核主成分分析(KPCA)对故障信号的小波尺度谱进行特征提取的方法。利用 ZT-3 多功能转子试验台获取上述 4 种故障各 32 个样本,对其进行连续小波变换和 KPCA 特征提取,并同时提取了相同样本条件下的尺度谱纹理特征和频谱特征。最后利用参数自适应支持向量机模型对提取的特征进行了分类测试。分析结果表明:KPCA 方法所提取特征的平均识别效果均达到 90% 以上,高于尺度谱纹理特征和频谱特征的分类结果,能够有效提取尺度谱的特征,有利于转子故障的智能诊断。

**关键词:**小波尺度谱;特征提取;故障诊断;转子;核主成分分析

**中图分类号:**V263.6;TP277 **文献标志码:**A

## Image feature extraction from wavelet scalogram based on kernel principle component analysis

SUN Li-ping, CHEN Guo, TAN Zhen-zhen

(School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, Jiangsu, China)

**Abstract:** The scalogram image features of unbalance, misalignment, rub-impact and oil whirl fault were analyzed, and a new feature extraction method from the wavelet scalogram of fault signals was put forward based on kernel principle component analysis(KPCA). By using ZT-3 multi-functional rotor test bed, 32 samples for each type of fault were obtained, continuous wavelet transformation was carried out, and KPCA feature, scalogram texture feature and spectrum feature were extracted. Finally, the extracted features were tested and classified by using parameter self-adaptive support vector machine. Analysis result shows that the average recognition effect of features extracted by KPCA is up to 90%, and is higher than the classification results of scalogram texture feature and spectrum feature, so KPCA can effectively extract the features of scalogram and is helpful for the intelligent diagnosis of rotor faults. 2 tabs, 2 figs, 10 refs.

**Key words:** wavelet scalogram; feature extraction; fault diagnosis; rotor; KPCA

**Author resumes:** SUN Li-ping (1981-), female, graduate student, +86-25-84891850, slpada@163.com; CHEN Guo (1972-), male, professor, +86-25-84891850, cgzyx@263.net.

## 0 引 言

振动信号是旋转机械状态监测与故障诊断的基

本信息来源,当故障发生或发展时会导致动态信号非平稳性的出现。小波分析具有良好的时频局部性、多尺度性和数学显微特性,能较好地解决非平稳

收稿日期:2009-06-01

基金项目:国家自然科学基金项目(50705042);航空科学基金项目(2007ZB52022)

作者简介:孙丽萍(1981-),女,山东莱州人,南京航空航天大学工学硕士研究生,从事图像处理及模式识别研究。

导师简介:陈 果(1972-),男,四川武胜人,南京航空航天大学教授。

信号分析的问题。观察不同故障信号的小波尺度谱,会发现不同故障的尺度谱其纹理分布及灰度变化是有区别的,即故障信号的小波尺度谱能够反映出故障的类别特征和严重程度<sup>[1-2]</sup>。对转子故障信号的小波尺度谱进行有效的特征提取和分类,有利于早期转子故障的识别。

目前对尺度谱进行特征提取的研究中,文献[3]分析了尺度谱纹理的变化,提出将纹理特征作为尺度谱的特征进行故障诊断的方法。文献[4]基于图像处理与识别技术,提出将图像的统计特征参数——图像的一阶矩作为尺度谱的特征。然而上述两种特征提取方法并未对故障的非线性特征进行相应的分析,仅从图像像素的二阶统计特性方面对故障特征进行了描述,未考虑到图像中的高阶统计信息,所提取的尺度谱特征中存在高阶冗余信息,这往往影响分类识别的实际效果。

核主成分分析(KPCA)考虑了包含图像像素间的非线性关系的高阶统计信息,其作为一种有效的非线性特征提取技术,已被广泛应用于人脸识别、故障诊断等多个领域<sup>[5-7]</sup>。由于不同故障类型振动信号的尺度谱在不同时刻、不同频率范围内所表现出的灰度变化有明显差异,并且相同类型信号的尺度谱面貌相似,基于以上特点,本文利用人脸识别领域中广泛应用的核主成分分析,对不同故障类型的尺度谱进行特征提取,利用参数自适应支持向量机模型对所得的特征分类识别,并将识别结果与频谱特征、尺度谱纹理特征的识别结果进行了比较。结果表明,该方法可以更为有效地提取故障信号的尺度谱特征,避免了高阶信息冗余对识别结果的影响。

## 1 基于小波尺度谱的转子故障特征分析

### 1.1 尺度谱及小波函数的选择

小波变换克服了短时傅立叶变换的缺点,通过可变的时频窗口对信号  $x(t)$  的不同频段进行分析,可以实现信号的多分辨率分析。 $x(t)$  的连续小波变换定义为

$$W_x(a, b, \tau) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^*(t - \tau) dt \quad a > 0 \quad (1)$$

窗函数  $\psi_{a,b}(t)$  称为子小波,是由母小波  $\psi(t)$  经伸缩平移产生的,即

$$\psi_{a,b}(t) = a^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

式中: $a$  为尺度参数; $b$  为时间参数; $\psi_{a,b}(t)$  为依赖于参数  $a$  和  $b$  的小波基函数; $\psi_{a,b}^*(t)$  为  $\psi_{a,b}(t)$  共轭。

此时,  $x(t)$  的连续小波反变换为

$$x(t) = \frac{1}{C} \iint_{\mathbb{R}^2} W_x(a, b, \tau) \psi_{a,b}(t) da db \quad (3)$$

$$C = \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt \right|^2 d\omega$$

$$\int_{-\infty}^{\infty} \psi(t) e^{-i\omega t} dt < \infty$$

为对  $\psi(t)$  提出的容许条件。

由连续小波变换公式

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |W_x(a, b, \tau)|^2 da db d\tau$$

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |W_x(a, b, \tau)|^2 da db d\tau$$

可以得知小波变换没有丢失任何信息,并且变换中的能量守恒。定义小波尺度谱为  $|W_x(a, b, \tau)|^2$ ,小波尺度谱被广泛地用于非稳定信号的分析,可看作是连续相关带宽的频谱,可以很快地直观地看到信号的变化情况。

小波基函数  $\psi(t)$  的选择不是唯一的,但其选择应满足小波容许条件,且其定义域应是紧支撑的。该条件使函数的波形具有振荡性,并保证了函数的速降特性以获得时间局域化。实际应用时需要针对不同的信号和不同的处理目的来确定小波基函数的选择。Complex Morlet 小波函数是一个指数衰减正弦曲线信号,而衰减正弦曲线是大多动力学系统的响应,并且 Complex Morlet 小波含有单一的频率,如果一个信号与一个尺度 Complex Morlet 小波高度相关,那么小波的频率就表示了信号被分析的频率。因此,本文选择 Complex Morlet 小波作为母小波,提取故障信号的连续小波尺度谱,Complex Morlet 的形式为

$$\psi(t) = \frac{1}{\sqrt{f_1}} e^{2i f_2 t} e^{-t^2/f_1}$$

式中: $f_1$  为带宽参数; $f_2$  为小波中心频率。本文选择带宽参数  $f_1$  为 1,中心频率  $f_2$  为 1.25。

### 1.2 基于小波尺度谱的转子故障特征分析

本文选取转子不对中、转子不平衡、转静碰摩、油膜涡动等 4 种常见的航空发动机转子故障进行了小波尺度谱特征的分析。从 4 种故障的频率特征上看,转子不对中的频率特征主要为除旋转频率外的 2 倍频或 4 倍频,尚伴有高次偶数倍频;不平衡故障的主要频率特征为旋转频率的单一倍频;转静碰摩比较复杂,由局部摩擦引起的振动频率中,含有 2 倍

频、3 倍频及一些高次谐波和次谐波振动;油膜涡动发生时,故障频率主要是除旋转频率以外的 1/2 倍频。图 1 给出了 4 种典型故障的频谱,图 2 为 4 种故障对应的尺度谱,所用数据均采集自 ZT-3 多功能转子试验台。

由图 1、2 可以看出,4 种故障的频谱及尺度谱均能表现出故障的特征差别,与频谱图相比,尺度谱更能有效地描述频率的变化情况,以及频率变

化的时间范围。不同频率所对应的能量的多少可由尺度谱中光带的明暗程度即图像的灰度变化反映出来,频谱图中频率的幅值越大,尺度谱中相应处的光带越亮,灰度值越大。由上述比较可以知道,故障信号的小波尺度谱能更好地反映出转子故障的时频特性,对转子故障信号的尺度谱进行特征提取并进行有效的识别,有利于转子故障的智能诊断。

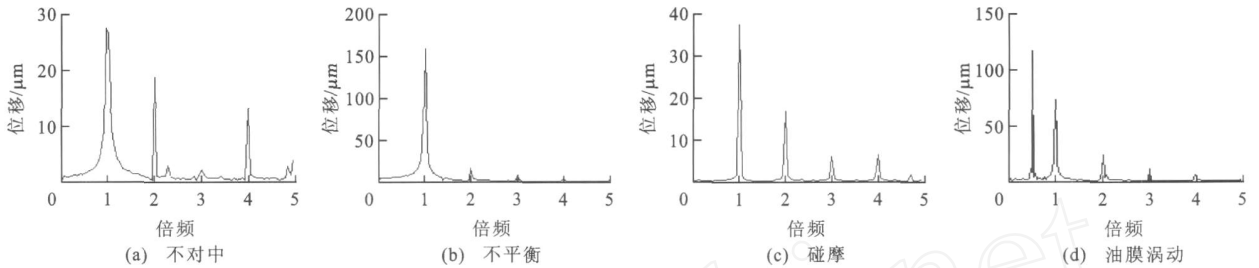


图 1 4 种故障信号的频谱

Fig. 1 Spectrums of four fault signals

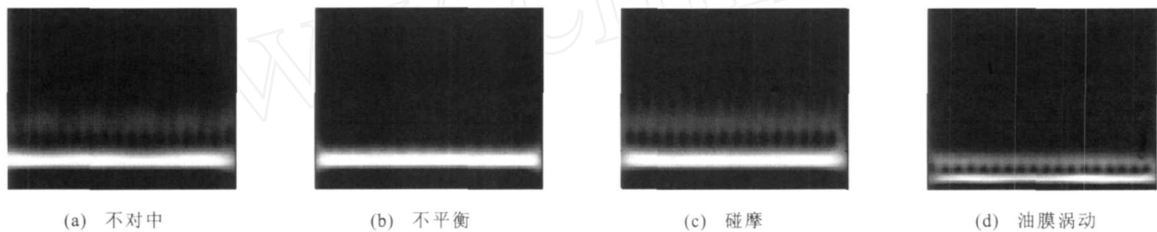


图 2 4 种故障信号的尺度谱

Fig. 2 Scalograms of four fault signals

## 2 基于 KPCA 的尺度谱数字特征提取

尺度谱可以看作一种典型的图像,尺度谱上所表现出的各故障特征差异主要体现在图像灰度明暗的变化,即不同时刻、不同频率范围内所表现出的灰度变化有明显差异。不同的故障类型呈现出的尺度谱面貌特征是有差别的,而相同类型信号的尺度谱面貌相似,正如属于不同人脸的多幅图像。由于尺度谱图像像素间存在着非线性关系的高阶统计信息,可以考虑利用目前人脸识别领域中特征提取的方法——核主成分分析(KPCA),进行尺度谱图像的特征提取,并以所提取的特征进行故障的智能诊断。

### 2.1 核主成分分析的基本原理

核主成分分析最早由 Scholkopf 等借鉴 SVM 的核方法(Kernel Method)思想,将 PCA 拓展到非线性情形而提出的<sup>[5-6]</sup>,是利用核技巧对经典的主成分分析法进行的一种非线性推广。KPCA 通过一个非线性函数将原始向量  $X (X \in \mathbf{R}^N)$  映射到一个高维的特征空间  $F$ ,即

$$F = \{ (X) | X \in \mathbf{R}^N \}$$

在  $F$  上对特征样本进行 PCA 分析<sup>[6-8]</sup>。它可以在输入空间无法线性分类的数据变换到特征空间来实现线性分类。KPCA 具有能有效捕捉数据的非线性特征、对原始空间中数据的分布情况没有要求等优点,被广泛应用于人脸识别等领域。

对于输入空间中的  $M$  个样本向量  $x_k (k = 1, 2, \dots, M)$ ,现引入非线性映射函数,使输入空间中的样本点  $x_1, x_2, \dots, x_M$  变换为特征空间中的样本点  $(x_1), (x_2), \dots, (x_M)$ ,假设映射数据是零均值的,即  $\sum_{k=1}^M (x_k) = \mathbf{0}$ ,则在特征空间  $F$  中,映射函数的协方差矩阵为

$$\bar{C} = \frac{1}{M} \sum_{j=1}^M (x_j) [ (x_j) ]^T \quad (4)$$

对矩阵  $\bar{C}$  做特征矢量分析,设其特征值为  $\lambda$ ,特征矢量为  $v \in F \setminus \{0\}$ ,则有  $\lambda v = \bar{C}v$ 。将每个样本与该式做内积,可得

$$[ (x_k) ] = (x_k) \bar{C}v \quad (5)$$

矩阵  $\bar{C}$  的特征矢量  $v$  可由  $(x_i)$  线性表示为

$$v = \sum_{i=1}^M \lambda_i (x_i) \tag{6}$$

式中:  $\lambda_i$  为相关系数。将式(6)代入式(2)得

$$\sum_{i=1}^M \lambda_i [(x_k) (x_i)] = \frac{1}{M} \sum_{i=1}^M \lambda_i [(x_k) \cdot \sum_{j=1}^M (x_j)] [(x_j) (x_i)] \tag{7}$$

定义  $M \times M$  阶矩阵  $K$  的元素为

$$K_{ij} = (x_i) (x_j) \tag{8}$$

式(4)简化为  $M K = K^2$ , 显然满足

$$M = K \tag{9}$$

通过对式(9)的求解,即可获得特征值和特征矢量。然而对零均值的假设在一般情况下是不成立的,此时  $K$  用  $\tilde{K}$  代替,则有

$$\tilde{K}_{ij} = K_{ij} - \frac{1}{M} \sum_{m=1}^M l_{im} K_{mj} - \frac{1}{M} \sum_{n=1}^M K_{in} l_{nj} + \frac{1}{M^2} \sum_{m,n=1}^M l_{im} K_{mn} l_{nj} \tag{10}$$

式中:对任意  $i, j, l_{ij}$  为 1; 1 为元素均为 1 的矩阵; 为相关系数向量。设所得特征值为  $\lambda_i$ , 其中  $\lambda_1 > \lambda_2 > \dots > \lambda_M$ 。主元  $k$  的贡献率为  $\lambda_k / \sum_{i=1}^n \lambda_i$ , 前  $k$  个主元的累积贡献率为  $\sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i$ 。主元贡献率的大小表示了该主元所携带的原始特征变异信息的大小,贡献率越大,表明它对特征信息的解释能力越强,对原始数据的综合能力也越好<sup>[7-9]</sup>。

2.2 试验结果分析

### 2.2 试验结果分析

为验证利用 KPCA 对尺度谱特征提取的有效性,本文利用 ZT-3 多功能转子故障模拟试验台和 DHDAS 信号测试分析系统分别获取了不同转速下不平衡故障样本、不对中样本、碰摩样本及油膜涡动各 32 个。将所得的 128 个故障信号进行连续小波变换,获得故障信号连续小波尺度谱,并利用核主成分分析对所得尺度谱进行特征提取。本文分别选取了  $p$  阶多项式核函数

$$K(x, x_i) = (x \cdot x_i + 1)^p$$

高斯径向基核函数

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)$$

将尺度谱图像数据变换到特征空间,经过多次试验和试验中核函数参数的调整,最终选取  $p$  为 1.5, 径向基参数  $\sigma^2$  为  $10^5$ 。随机选取 4 种故障样本各 16 幅进行核主成分分析,并计算其主元贡献率。本文进

行了多次试验,选择其中一次,对前 8 个主元贡献率列表,见表 1。

表 1 主元贡献率

Tab. 1 Contribution rates of principle component

特征值序号	$p$ 阶多项式核函数 ( $p=1.5$ )			高斯径向基核函数 ( $\sigma^2=10^5$ )		
	$\lambda$ 值	贡献率/%	累积贡献率/%	$\lambda$ 值	贡献率/%	累积贡献率/%
1	2.075 4	67.28	67.28	4.193	69.75	69.75
2	0.623 1	20.20	87.48	1.052	17.49	87.25
3	0.139 4	4.52	92.00	0.227	3.78	91.03
4	0.058 1	1.88	93.88	0.150	2.49	93.52
5	0.031 5	1.02	94.90	0.077	1.28	94.80
6	0.020 6	0.67	95.57	0.050	0.83	95.62
7	0.013 8	0.45	96.02	0.027	0.45	96.07
8	0.013 5	0.44	96.46	0.024	0.40	96.47

由表 1 可以看出,无论是  $p$  阶多项式核函数或高斯径向基核函数,在特征压缩中均表现出较好的效果,第一主元的贡献率均达到 65% 以上,其中前 3 个特征值的累积贡献率均达到 90% 以上,前 5 个特征值的累积贡献率均接近 95%。为保证特征提取过程中有较好的降维效果并且信息的损失量最少,经试验,选择特征维数为 5。

为验证 KPCA 方法对尺度谱特征提取的有效性,本文对相同样本的 0.20、0.25、0.33、0.43、0.50、0.67、0.75、1.00、2.00、3.00、4.00、5.00、6.00、7.00、8.00、9.00、10.00 倍共 17 个频谱特征及文献[2]中提及的纹理特征进行了提取,并利用参数自适应支持向量机模型<sup>[10]</sup>对获得的尺度谱特征、频谱特征、纹理特征的识别结果做了比较。多次试验中,随机选取 42 个样本作为训练样本,43 个作为测试样本,43 个作为验证样本。选取试验 4 次的结果进行比较,见表 2,其中 KPCA 特征 1 为利用  $p$  阶多项式核函数进行特征变换后所得的特征, KPCA 特征 2 为利用高斯径向基核函数变换后得到的特征。

由表 2 可以看出,利用 KPCA 方法无论采用那种核函数,当核函数的参数选择适当时,提取的小波尺度谱特征比传统的频谱特征和尺度谱纹理特征对故障有着更好的分类效果,最高可达 100%,平均识别效果均为 90.00% 以上,而频谱特征的识别效果平均在 80.00% 左右,尺度谱纹理特征的平均识别效果为 88.75%。显然,本文提出的利用 KPCA 方法对尺度谱提取所得的数字特征比频谱特征和尺度谱纹理特征具有更好的故障识别能力,可以应用于尺度谱的特征提取和故障的智能诊断。

表 2 识别结果比较

Tab. 2 Comparison of recognition results

%

特征	测试 1	验证 1	测试 2	验证 2	测试 3	验证 3	测试 4	验证 4	平均
频谱特征	88.37	82.05	76.74	90.70	79.07	88.37	90.70	81.40	84.68
纹理特征	88.37	90.70	94.29	84.61	90.70	84.62	93.02	83.72	88.75
KPCA 特征 1	95.35	88.37	88.37	94.29	91.43	94.29	97.67	100.00	93.72
KPCA 特征 2	94.29	97.67	97.67	100.00	95.35	90.07	95.35	90.70	95.14

### 3 结 语

针对连续小波变换的尺度谱特征难于自动提取的问题,本文提出利用 KPCA 方法实现小波尺度谱特征的自动提取并以之作为故障识别的特征;利用 ZT-3 多功能转子试验台获取了不平衡、不对中、碰摩及油膜涡动故障的 128 个样本,提取了相同样本下故障信号的尺度谱特征和频谱特征,并利用结构自适应神经网络模型对所得特征进行了分类识别。结果表明,利用 KPCA 方法所提取的尺度谱数字特征具有更强的故障识别能力,通过该方法提取的特征可以有效地实现故障的智能诊断。

#### 参 考 文 献 :

#### References :

- [ 1 ] PENG Z, HE Y, LU Q, et al. Feature extraction of the rub-impact rotor system by means of wavelet analysis[J]. Journal of Sound and Vibration, 2003, 259(4): 1000-1010.
- [ 2 ] 彭志科,何永勇,褚福磊.小波尺度谱在振动信号分析中的应用研究[J].机械工程学报,2002,38(3):122-126.  
PENG Zhi-ke, HE Yong-yong, CHU Fu-lei. Using wavelet scalogram for vibration signals analysis[J]. Chinese Journal of Mechanical Engineering, 2002, 38(3): 122-126. (in Chinese)
- [ 3 ] 陈 果,邓 堰.转子故障的连续小波尺度谱特征提取新方法[J].航空动力学报,2009,24(4):793-798.  
CHEN Guo, DENG Yan. New approach of features extraction for rotor faults from continuous wavelet transform scalogram[J]. Journal of Aerospace Power, 2009, 24(4): 793-798. (in Chinese)
- [ 4 ] 侯敬宏,黄树红,申 强,等.基于小波分析的旋转机械振动信号定量特征研究[J].机械工程学报,2004,40(1):131-135.  
HOU Jing-hong, HUANG Shu-hong, SHEN Tao, et al. Wavelet-based quantitative analysis of vibration signal of rotary machines[J]. Chinese Journal of Mechanical Engineering, 2004, 40(1): 131-135. (in Chinese)
- [ 5 ] SCHOL KOPF B, SMOLA A, MULLER K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation, 1998, 10(5): 1299-1319.
- [ 6 ] KIM K I, JUNG K, KIM H J. Face recognition using kernel principal component analysis [J]. IEEE Signal Processing Letters, 2002, 9(2): 40-42.
- [ 7 ] MAL HI A, GAO R X. PCA-based feature selection scheme for machine defect classification [J]. IEEE Transactions on Instrumentation and Measurement, 2004, 53(6): 1517-1525.
- [ 8 ] HE Qian, LIU Yi-bing, PENGLu. Kernel principal components analysis for early identification of gear tooth crack [C]. IEEE. Proceedings of the 6th World Congress on Intelligent Control and Automation. Dalian: IEEE, 2006: 5748-5751.
- [ 9 ] 李巍华,廖广兰,史铁林.核函数主元分析及其在齿轮故障诊断中的应用[J].机械工程学报,2003,39(8):65-70.  
LI Wei-hua, LIAO Guang-lan, SHI Tie-lin. Kernel principal component analysis and its application in gear fault diagnosis[J]. Chinese Journal of Mechanical Engineering, 2003, 39(8): 65-70. (in Chinese)
- [ 10 ] 陈 果.基于遗传算法的支持向量机分类器模型参数优化[J].机械科学与技术,2007,26(3):347-350.  
CHEN Guo. Optimizing the parameters of support vector machines classifier model based on genetic algorithm [J]. Mechanical Science and Technology, 2007, 26(3): 347-350. (in Chinese)