



孙丽萍

基于 KPCA 的航空发动机滑油滤磨屑图像识别

孙丽萍¹ 陈 果¹ 陈立波² 高绪伟¹

(¹南京航空航天大学 民航学院, 南京 210016; ²北京航空工程技术研究中心, 北京 100076)

摘 要: 针对目前航空发动机滑油滤的检查仍处于目视定性检查水平, 检查结果依赖人员的经验, 既主观又无定量依据的现状, 研究了基于图像识别的航空发动机滑油滤磨屑检测技术。首先构造了油滤图像检测硬件系统; 然后提出了利用核主成分分析(kernel principle component analysis, KP-CA)对滑油滤图像进行特征提取的方法; 最后, 利用实际采集的滑油滤图像进行了实例分析, 并与普通的主成分分析(principle component analysis, PCA)方法进行比较。结果表明, KPCA方法可以更为有效地提取出滑油滤图像的磨损状态特征, 能够有效地提高发动机磨损故障预报的准确率。

关 键 词: 航空发动机; 核主成分分析(KPCA); 滑油滤; 特征提取; 图像识别; 磨损诊断

中图分类号: TP277 **文献标识码:** A **文章编号:** 1003-8728(2010)06-0731-06

Image Recognition of Aero-engine Oil Filter Debris by Kernel Principle Component Analysis

Sun Liping¹, Chen Guo¹, Chen Libo², Gao Xuwei¹

(¹ College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016;

² Beijing Aeronautical Technology Research Center, Beijing 100076)

Abstract: Inspection of the aero-engine oil filters still relies on the experience of the engineers. We present a quantitatively method for aero-engine oil filter inspection. First, we construct the oil filter image detection hardware system. Then, Kernel Principle Component Analysis (KPCA) is used to extract the oil filter image features. Finally, we analyze the actual acquisition oil filter images with KPCA and compare the results with those by the method of traditional Principle Component Analysis (PCA). Comparison result indicates that the method of using KPCA can extract the wear condition features in oil filter images more efficiently. It will improve the prediction accuracy of aero-engine wear faults efficiently.

Key words: aero-engine; KPCA; oil filter; feature extraction; image recognition; wear diagnosis

目前,对于航空发动机滑油润滑零部件的磨损故障,已经形成了较为成熟的监测与诊断技术体系,主要的诊断手段有磁塞检查、颗粒计数、光谱分析及铁谱分析等^[1]。但这些滑油监控技术只能监控滑油中100 μm以下的磨粒。对于滑油系统中100 μm以上的磨粒,由于其极易快速沉降,难以通过定期取

滑油的方式进行监控。100 μm以上磨屑的大量出现,往往是发动机异常磨损的强烈征兆,这些磨屑一般都汇集于滑油滤中。随着新型发动机滑油系统部件和油滤向精细化发展,大量悬浮在滑油中的微小磨粒也将被油滤截获。因此对滑油滤中的磨屑进行定量监控,对于提高发动机磨损故障预报成功率具有重要作用。

通过对某型航空发动机滑油滤故障的统计,发现:

(1) 发动机提前换发率较高,其中磨损和轴承失效故障约占6成以上。发动机滑油系统多发性故

收稿日期:2009-05-18

作者简介:孙丽萍(1981-),硕士研究生,研究方向为图像处理与模式识别,slpada@163.com;陈果(联系人),教授,博士生导师,fgzyx@263.net

障主要集中在主轴轴承磨损、剥落和轴承封严装置偏磨。

(2) 对于滑油滤的检查,外场还处于目视定性检查水平,检查结果依赖人员的经验,既主观又无量依据。

表1为4台因金属屑超量提前返厂的某型航空发动机故障检测情况。其中1台发动机抱轴,2台发动机轴承封严装置严重磨损,1台发动机轴承严重磨损。对上述发动机故障,在部门使用中均发现滑油滤上出现较多金属磨屑,而滑油光谱数据正常。其它返该厂的发动机主轴轴承和封严装置大多有不同程度磨损。

表1 因金属屑故障返厂检修的4台某型发动机情况

| 故障日期 | 返厂原因 | 故障情况 |
|--------------|-------------------|------------|
| 2004. 10. 14 | 振动值超标,滑油滤中发现较多金属屑 | 轴承封严装置严重磨损 |
| 2005. 04. 28 | 振动值超标,滑油滤中发现较多金属屑 | 轴承严重磨损 |
| 2006. 04. 26 | 滑油滤中发现较多金属屑 | 轴承封严装置严重磨损 |
| 2007. 10. 24 | 滑油滤中发现较多金属屑 | 轴承严重磨损 |

以上调研结果可见,航空发动机的轴承和轴承封严装置属易损的故障多发部位,通过滑油光谱和其它现有的监控手段难以有效预测,而滑油滤磨屑监控对于上述部位的磨损或剥落故障较敏感,由此可见,建立标准化、量化的便于外场操作的滑油滤磨屑监控方法,可以有效弥补滑油光谱监控难以检测滑油中较大尺寸磨粒的技术局限性。

由此,笔者研究了一种利用图像识别的航空发动机滑油滤磨损状态识别技术,构建了航空发动机滑油滤监控系统的硬件装置,采集滑油滤磨屑图像;利用目前人工智能领域先进的特征提取方法——核主成分分析(KPCA)方法^[2,3]对采集到的航空发动机滑油滤磨屑图像进行了特征提取和压缩;采用最近邻方法对滑油滤图像的磨损状态进行辨识。最后,利用实际航空发动机滑油滤图像进行分析验证,并与普通的主成分分析(PCA)方法进行比较。

1 航空发动机滑油滤图像获取及特征分析

实际飞机停机坪不具备民用电源,因此采用基于超声波清洗的滑油滤自动清洗方案无法在外场实施,故采用基于气雾喷射原理对滑油滤进行清洗和磨屑收集的方案。滑油滤清洗和磨屑收集系统实物如图1所示。



图1 滑油滤清洗和磨屑收集系统实物

由于发动机滑油系统故障集中在轴承和轴承封严装置,磨损产物以铁磁性的钢、青铜和镍基合金为主,非铁磁性磨屑一般来源于铝合金或镁铝合金等壳体部件偏磨,故障的危害性较小。因此本系统确定以滑油滤上铁磁性磨屑为重点研究对象。

用气雾喷射器原位清洗滑油滤,将清洗下的磨屑通过磨屑过滤装置进行收集。磨屑过滤装置的滤材采用300目滤布,滤布下安装一定尺寸的环形永久磁铁(钕铁硼),使铁磁性磨屑按照磁力线排列成环状,避免过度堆积。采集图像时,取下永磁铁,将过滤组件置于带光源的台架上,采用CCD相机照相的方法采集磨屑图像。图像采集装置如图2所示。



图2 图像采集装置

不同磨损程度的滑油滤图像如图3所示,从图中可以看出,钢、青铜和镍基合金为主的铁磁性磨屑

被磁铁吸引呈环形排列,其他非金属磨屑由于不受磁力作用,在图像中分布较为分散。并且,不同磨损

状态的图像具有不同的磨屑堆积程度,形成了不同的面貌特征。

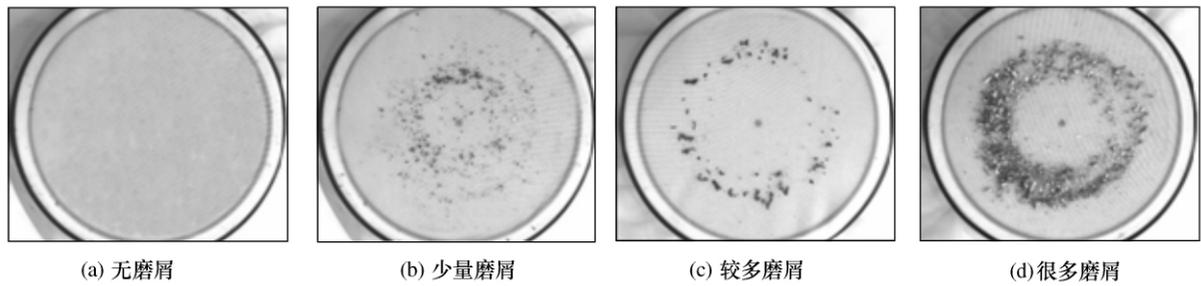


图 3 不同磨损程度的滑油滤图像

如果能够让计算机记住各种磨损状态的油滤图像的面貌,则在图像识别过程中,只需将待识别图像与各典型的不同磨损状态的图像进行比较即可。然而,由于图像数据非常庞大,数据间存在很大的冗余信息,如何有效地对图像数据进行压缩和降维,有效地实现图像特征提取,是有效实现图像识别的关键。受模式识别中的“人脸识别”思想启发^[4-6],如果将不同磨损状态的图像看做不同的人脸,则人脸识别的方法完全可以应用于滑油滤图像的识别中。

目前,主成分分析(PCA)广泛应用于人脸识别^[7],而基于核的主成分分析(KPCA)是目前国际上流行的一种特征提取新方法,它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^[3-5]。与传统主成分分析法相比,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,因此它在包括人脸识别的诸多领域的应用取得很好的效果^[5,6]。有鉴于此,笔者将KPCA方法应用于滑油滤图像的特征提取,实现滑油滤图像的压缩与降维处理,根据所提取的特征,利用最近邻法实现滑油滤图像磨损状态的识别。

2 基于 KPCA 的油滤磨屑图像特征提取

2.1 核主成分分析(KPCA)的原理

最早将核方法思想应用于特征抽取的是 scholkopf 等人^[7],于 1999 年借鉴 SVM 的核方法(Kernel Method)思想,将 PCA 拓展到非线性情形,提出了核主成分分析(KPCA)。核主成分分析(kernel PCA)是 PCA 的改进算法,它采用非线性方法来提取主成分^[2,3],即 Kernel PCA 是通过一个非线性函数 Φ 将原始向量 $X(X \in R^N)$ 映射到一个高维的

特征空间 $F, F = \{\Phi(X) : X \in R^N\}$ 在 F 上进行 PCA 分析。它可以将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。

对于输入空间中的 M 个样本 $x_k(k = 1, 2, \dots, M), x_k \in R^N$,使 $\sum_{k=1}^M x_k = 0$,则其协方差矩阵为

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

对于一般 PCA 方法,即通过求解特征方程

$$\lambda v = C v \quad (2)$$

获得贡献率大的特征值(对应较大的特征值)及与之对应的特征向量。现引入非线性映射函数 Φ ,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 $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_M)$,设

$$\sum_{k=1}^M \Phi(x_k) = 0 \quad (3)$$

则在特征空间 F 中的协方差矩阵为

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

因此,特征空间中的 PCA 是求解方程

$$\lambda v = \bar{C} v$$

中的特征值 λ 和特征向量 $v \in F \setminus \{0\}$,进而有

$$\lambda (\Phi(x_k) \cdot v) = \Phi(x_k) \cdot \bar{C} v (k = 1, 2, \dots, M) \quad (5)$$

注意到上式中 v 可以由 $\Phi(x_i) (i = 1, 2, \dots, M)$ 线性表示出,即

$$v = \sum_{i=1}^M \alpha_i \Phi(x_i) \quad (6)$$

由式(4)~式(6)得

$$\lambda \sum_{i=1}^M \alpha_i (\Phi(x_k) \cdot \Phi(x_i)) =$$

$$\frac{1}{M} \sum_{i=1}^M \alpha_i (\Phi(x_k) \cdot \sum_{j=1}^M \Phi(x_j)) (\Phi(x_j) \cdot \Phi(x_i)) \quad (k=1, 2, \dots, M) \quad (7)$$

定义 $M \times M$ 矩阵 K

$$K_{ij} \equiv \Phi(x_i) \cdot \Phi(x_j) \quad (8)$$

式(7)简化为

$$M\lambda K\alpha = K^2\alpha \quad (9)$$

显然满足

$$M\lambda\alpha = K\alpha \quad (10)$$

则必然满足式(9)。通过对式(10)的求解,即可获得要求的特征值和特征向量。对于测试样本在 F 空间向量 V^k 上的投影为

$$(V^k \cdot \Phi(x)) = \sum_{i=1}^M \alpha_i^k (\Phi(x_i) \cdot \Phi(x)) \quad (11)$$

最后一个问题是假设式(3)在一般情况下是不成立的,此时式(10)中的 K 用 \tilde{K} 代替。

$$\begin{aligned} \tilde{K}_{ij} &= K_{ij} - \frac{1}{M} \sum_{m=1}^M 1_{im} K_{mj} - \\ &\frac{1}{M} \sum_{n=1}^M K_{in} 1_{nj} + \frac{1}{M^2} \sum_{m,n=1}^M 1_{im} K_{mn} 1_{nj} \end{aligned} \quad (12)$$

式中: $1_{ij} = 1$ (对于所有的 i, j)。

2.2 算法的执行步骤

KPCA 算法可以分为以下 3 个步骤:

Step 1 选择适当形式的核函数 $k(x, y)$, 将训练样本的图像数据进行核映射, 实现高维空间上的中心化, 然后根据式(12)计算矩阵 \tilde{K} 。

Step 2 计算 \tilde{K} 的特征值和特征向量, 并进行高维空间上的正规化。

Step 3 对于测试样本, 计算其非线性主分量。

在求出测试样本的非线性主分量后, 可选择合适的分类器进行分类。

2.3 最近邻法介绍

为了对润滑油滤图像的磨损状态进行识别, 在完成 KPCA 特征提取后, 需要设计相应的分类器来实现润滑油滤图像磨损特征的模式分类。笔者选取简单实用的最近邻分类器。最近邻分类器在理论上的错误率不高于 2 倍贝叶斯分类器^[5], 实际应用也具有较好的效果。最近邻分类器是计算待测样本与每一个训练样本的距离, 将“距离”待测样本最近的训练样本所属的类别作为待测样本的类别。最近邻的基本思想为, 设有一组 n 个样本: $X = \{X_1, X_2, \dots, X_n\}$, 每个样本都已标记类别标志。如果这 n 个样本中与待分类样本 X 相距最近的一个样本为 $X'_n \in X$,

则把 X 分到 X'_n 所在的类别中去。

该过程可用判别函数表示。设有 c 类模式样本 $\omega_1, \omega_2, \dots, \omega_c$, 每类样本有 n_i 个 $i=1, 2, \dots, c$, 此时最近邻的判别函数为

$$d_i(X) = -\min_k \|X - X_{i,k}\| \quad (k=1, 2, \dots, n_i) \quad (13)$$

决策法则为: 若有 $d_i(X) > d_j(X)$ $i \neq j$, 则把 X 分到第 i 类中。

3 基于 KPCA 的航空发动机滑油滤图像识别

笔者选取了 128 幅正常滑油滤图像和 128 幅异常滑油滤图像对 KPCA 算法进行验证。其中典型的正常及异常滑油滤图像如图 3 所示, 目前, 正常和异常磨损图像需要由领域专家根据经验确定和认可。实验时, 随机选取 64 幅正常图像和 64 幅异常图像作为训练集, 剩余 128 幅图像作为测试集。

由于光照不均匀等因素对图像造成的干扰, 我们对图像的像素灰度进行了归一化处理, 归一化公式为

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

式中: x 表示归一化之前图像的像素灰度; x' 表示归一化后的图像的像素灰度; x_{\min} 表示 x 中的最小值; x_{\max} 表示 x 中的最大值, 这样像素灰度的范围就变成 $[0, 1]$, 便于后面处理。

由于核函数的选取对 KPCA 的特征提取效果具有很大影响, 笔者选取了多项式核函数及高斯径向基(RBF)核函数对 KPCA 方法的主元累积贡献率进行比较。为了便于表示, 我们将其分别表示为 KPCA1、KPCA2。其中 KPCA1 的核参数 p 取 0.01, KPCA2 的核参数 σ 取 10000。两种核函数的形式可表示为

p 阶多项式核函数(KPCA1)

$$K(x, x_i) = [(x \cdot x_i) + 1]^p \quad (15)$$

高斯径向基函数(RBF)核函数(KPCA2)

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right) \quad (16)$$

在相同参数下, 随机选取 64 幅正常图像和 64 幅异常图像进行计算, 如图 4 为 3 次随机选取训练图像后由 PCA 和 KPCA 算法得到的各主元累积贡献率比较图, 我们对其中图 4(a) 的结果进行列表, 说明 KPCA 方法对特征提取和压缩的优势, 其比较结果如表 2 所示。

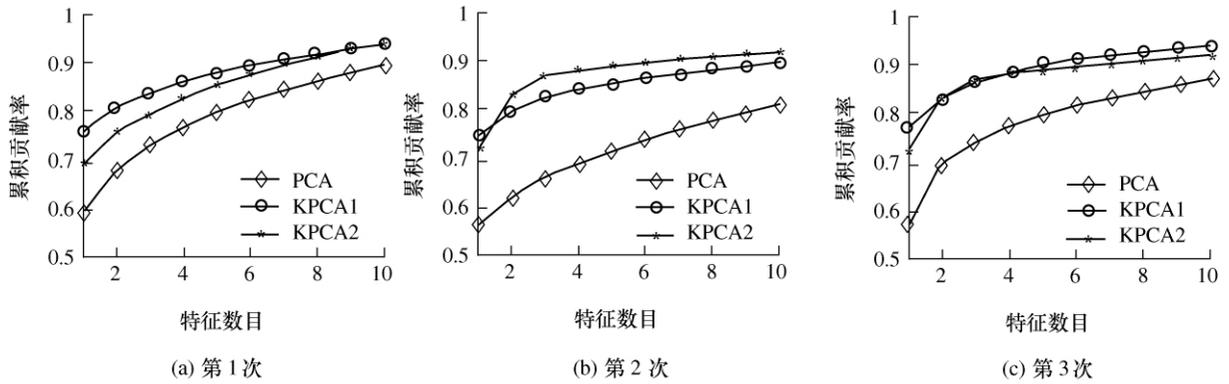


图 4 PCA 与 KPCA 各主元累积贡献率比较图

表 2 PCA 与 KPCA 的效果对比

| 特征值 序号 | PCA | | | KPCA1 | | | KPCA2 | | |
|-----------|---------------|--------|----------|---------------|--------|----------|---------------|--------|----------|
| | λ_i 值 | 贡献率(%) | 累积贡献率(%) | λ_i 值 | 贡献率(%) | 累积贡献率(%) | λ_i 值 | 贡献率(%) | 累积贡献率(%) |
| 1 | 8.801 | 58.71 | 58.71 | 5.179 | 75.56 | 75.56 | 1.299 | 68.96 | 68.96 |
| 2 | 1.345 | 8.97 | 67.68 | 0.363 | 5.3 | 80.86 | 0.12 | 6.39 | 75.35 |
| 3 | 0.764 | 5.09 | 72.77 | 0.193 | 2.82 | 83.68 | 0.069 | 3.71 | 79.06 |
| 4 | 0.566 | 3.78 | 76.55 | 0.151 | 2.21 | 85.89 | 0.063 | 3.36 | 82.42 |
| 5 | 0.463 | 3.09 | 79.64 | 0.116 | 1.7 | 87.59 | 0.052 | 2.77 | 85.19 |
| 6 | 0.348 | 2.32 | 81.96 | 0.096 | 1.4 | 88.99 | 0.042 | 2.22 | 87.41 |
| 7 | 0.301 | 2 | 83.96 | 0.094 | 1.36 | 90.35 | 0.035 | 1.85 | 89.26 |
| 8 | 0.289 | 1.93 | 85.89 | 0.082 | 1.19 | 91.54 | 0.03 | 1.61 | 90.87 |
| 9 | 0.269 | 1.8 | 87.69 | 0.068 | 1.07 | 92.61 | 0.029 | 1.58 | 92.45 |
| 10 | 0.248 | 1.66 | 89.35 | 0.056 | 0.98 | 93.59 | 0.026 | 1.38 | 93.83 |

核函数及其参数的选择是 KPCA 存在的典型问题,目前尚没有统一的选择理论和依据^[3,9]。针对核函数选取的不同而导致识别效果的不同,笔者将在相同的特征维数下,基于 p 阶多项式核函数,高斯径向基核函数的不同参数的 KPCA 算法以及 PCA 算法对滑油滤图像进行特征提取,并利用最近邻法进行识别,得到的识别率如表 3 和表 4 所示。其中表 3、表 4 分别是特征维数为 20 和 10 时,在不同的核函数及参数下 KPCA 所对应的识别结果,从表 3 和表 4 可以看出,不同的核函数和参数对特征提取和结果识别有一定的影响,但是识别率均在

90.63%和 96.88%之间。从表 3 可以看到,当特征维数为 20 时,对于多项式核函数,识别率在 91.41%和 94.35%之间,平均为 92.97%,对于高斯径向基核函数,识别率均在 91.41%和 96.09%之间,平均为 92.84%。从表 4 可以看到,当特征维数为 10 时,对于多项式核函数,识别率均在 90.63%和 94.35%之间,平均为 92.58%;对于高斯径向基核函数,识别率均在 92.19%和 96.88%之间,平均为 94.14%。PCA 在特征维数为 20 时对应的识别率为 91.41%,在特征维数为 10 时对应的识别率为 84.38%。

表 3 特征维数为 20 时不同的核函数及参数下 KPCA 对应的识别结果

| | | | | | | | |
|----------|--------------|--------|--------|--------|--------|--------|--------|
| 多项式核函数 | 核参数 p | 0.01 | 0.05 | 0.1 | 0.5 | 1 | 5 |
| | 识别率 | 94.53% | 93.75% | 93.75% | 92.19% | 91.41% | 92.19% |
| 高斯径向基核函数 | 核参数 σ | 0.1 | 1 | 10 | 100 | 1000 | 10000 |
| | 识别率 | 90.63% | 91.41% | 91.41% | 92.97% | 94.53% | 96.09% |

表 4 特征维数为 10 时不同的核函数及参数下 KPCA 对应的识别结果

| | | | | | | | |
|----------|--------------|--------|--------|--------|--------|--------|--------|
| 多项式核函数 | 核参数 p | 0.01 | 0.05 | 0.1 | 0.5 | 1 | 5 |
| | 识别率 | 94.53% | 93.75% | 92.97% | 91.41% | 92.19% | 90.63% |
| 高斯径向基核函数 | 核参数 σ | 0.1 | 1 | 10 | 100 | 1000 | 10000 |
| | 识别率 | 92.19% | 92.19% | 94.53% | 95.31% | 93.75% | 96.88% |

实验结果说明:在相同的特征维数下,KPCA 的识别率高于普通 PCA 算法,当维数减少时,KPCA 方法更加具有优势,这是因为 KPCA 是一种很好的特征提取方法,它是在 PCA 方法的基础上结合核函数而形成的一种新的特征提取方法;相对 PCA 方法而言,它保留了原始图像的高阶信息,描述了多个像素间的非线性相关性,这样使得提取的特征值包含的信息更为丰富,对后续的训练识别有很大的帮助;另一方面,由于它是将输入数据变换到特征空间,从而提取的特征值是线性可分的,这样又简化了分类器的设计。

4 结论

针对目前航空发动机滑油滤的检查依赖经验,难于定量分析的现状,研究了基于图像识别的航空发动机滑油滤磨屑检测技术。构造了油滤图像检测硬件系统,实现了图像采集;利用核主成分分析(KPCA)对滑油滤图像进行特征提取,并利用实际采集的滑油滤图像进行了实例分析。与普通的 PCA 方法进行比较后,KPCA 方法可以更为有效地提取出滑油滤图像的磨损状态特征,从而可以有效地提高发动机磨损故障预报的准确率。

[参考文献]

- [1] 陈果,陈立波,朱兰琪.一种磨损故障融合诊断新方法及其应用[J].机械科学与技术,2009,28(9):1157~1161
- [2] 边肇祺,张学工.模式识别[M].北京:清华大学出版社,2002
- [3] 肖建华.智能模式识别方法[M].广州:华南理工大学出版社,2005
- [4] 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
- [5] Kirby M, Sirovich L. Application of the K-L procedure for the characterization of human faces[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(1): 103~108
- [6] Malhi 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
- [7] Scholkopf B, Smola A, Muller K-R. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, 10(6): 1299~1319
- [8] Wang L, Zhang Y, Feng J. On the Euclidean distance of images[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1334~1339
- [9] Li J, Li X L, Tao D C. KPCA for semantic object extraction in images[J]. *Pattern Recognition*, 2008, 41(10): 3244~3250