

基于核函数主元分析的航空发动机故障检测方法^{*}

胡金海, 谢寿生, 陈 卫, 侯胜利, 蔡开龙

(空军工程大学 工程学院, 陕西 西安 710038)

摘 要: 航空发动机性能由正常到异常、再由异常发展到完全故障的阶段, 其参数变化具有一定非线性特征。为了有效检测这种具有非线性特征的故障, 提出一种基于核函数主元分析 (KPCA) 的非线性故障检测方法。该方法通过核函数完成非线性变换, 将变量由非线性的输入空间转换到线性的特征空间, 在特征空间中使用线性主元分析 (PCA) 方法计算主元, 构造 T^2 和 SPE 统计量检测故障的发生。通过对某型涡扇发动机进行实例验证分析, 结果表明, KPCA 方法一方面克服了综合参数法由于没有确定的警戒值而无法有效地进行故障检测的不足; 另一方面 KPCA 方法在非线性的故障检测过程中能够提取重要的非线性特征信息, 因而比 PCA 方法能更早地检测到早期潜在故障, 且 KPCA 方法检测错误率更低。因此, KPCA 方法更适用于具有非线性特征的航空发动机故障检测。

关键词: 航空发动机; 性能监控; 故障检测; 核主元分析法; 主元分析法

中图分类号: V263.6 文献标识码: A 文章编号: 1001-4055 (2008) 01-0079-05

An aeroengine fault detection method based on kernel principal component analysis

HU Jin-hai XIE Shou-sheng CHEN Wei HOU Sheng-li CAI Kai-long

(Engineering Inst., Air force Engineering Univ., Xi'an 710038, China)

Abstract The aeroengine parameters possess some nonlinear features when the performance of aeroengine goes from normal to abnormal and further from abnormal to completely faulty. In order to effectively detect the fault owned nonlinear feature, a novel approach of fault detection of aeroengine based on kernel principal component analysis (KPCA) model is presented. KPCA performs nonlinear transformation by kernel function to map the nonlinear input space into linear feature space and computes principal component by performing principal component analysis (PCA) in feature space, and detects faults by utilizing statistics T^2 and SPE . The practical applications in monitoring certain type of turbine-fan engine show that KPCA is superior to PCA in fault detection and is more suitable to fault detection of aeroengine owned nonlinear feature.

Key words Aeroengine; Performance monitoring; Fault detection; Kernel principal component analysis; Principal component analysis

1 引 言

对航空发动机进行连续的性能监控及故障检测是保证航空发动机安全工作、发现及排除早期故障、实现发动机视情维修、延长发动机使用寿命的一种重要技术途径。文献 [1~3] 提出采用综合参数法对航空发动机进行性能监控与故障检测, 该方法通过综合发动机多项性能指标, 得到一个定量反映发动机整体

性能的综合指数, 利用该指数对发动机整体性能进行监控, 发现早期故障。

虽然该方法取得了一定的检测效果, 但从其使用过程来看, 综合参数法仍存在以下不足:

(1) 需要足够多的发动机异常和故障数据作为训练样本来获取各参数的权值; 否则, 得到的权值的推广性较差, 不能有效反映出其它类型的异常和故障, 从而使综合参数法确定的性能综合指数失去意

* 收稿日期: 2006-11-02 修订日期: 2007-08-07。

基金项目: 国家自然科学基金资助项目 (60672179); 军队重点科研基金资助项目 (2003KJ01705)。

作者简介: 胡金海 (1978—), 男, 博士, 讲师, 研究领域为航空发动机状态监控与故障诊断。

Email: hjh19786@163.com

义,而实际发动机异常和故障数据很少且不容易获得;

(2)综合参数法没有给出具体的性能综合指数的警戒值,因此也就无法准确判断到底综合指数下降到什么程度才表示发动机性能严重恶化、存在故障,所以该指数只能定性表征发动机性能的变化趋势,要通过性能综合指数来判断发动机性能正常还是异常(故障),还必须更多依赖使用人员的专业知识和维护经验来辅助判断,这种状况势必会给使用维护人员对发动机性能状况的判断带来一定的困难,并增加了决策的风险性,并且如果存在早期故障,由于没有警戒值,该性能综合指数也并不能完全起到故障检测的作用。

近年来,作为一种基于数据模型的方法,多元统计分析理论为机械设备的状态监测与故障检测的研究提供了一种新的数学工具和解决方案,利用 PCA 模型进行故障检测与诊断技术得到了学术界的广泛重视,其基本思想和理论方法对解决故障检测与诊断问题带来了光明的前景^[4]。

应用 PCA 主元分析模型进行故障检测与诊断的优势在于:只需正常样本数据,不需要故障或异常样本数据;可以构造出反映正常状态及性能的统计量指标,并且可确定统计量的界限值,较容易地实现故障的检测;当检测到故障,可通过贡献率图确定引起故障或异常的变量源。

但 PCA 方法应用在航空发动机故障检测中亦有其不足之处。由于航空发动机结构复杂,工作情况多变,再加以发动机中一些结构的非线性,使得发动机的工作参数具有一定的非线性特征^[5],尤其是从故障初期到故障的发展,直至完全故障,其参数的非线性特征变化更显著。而 PCA 方法是基于线性代数理论的线性空间变换方法,是一种线性映射算法,在处理非线性问题时,往往不能取得好的效果。因此 PCA 方法在检测具有非线性特征的发动机故障时,必定有其不足之处,具体表现为:

(1)线性分解方法压缩和提取信息不充分;

(2)线性方法监测结果不可靠。这是由于较小的主元中可能包含重要的非线性信息,如果舍弃该主元,会导致重要信息的丢失;如果保留该主元,会造成模型复杂。所以有必要采取非线性多元统计分析方法来实施对航空发动机的故障检测。

因此研究人员提出一系列的非线性 PCA 方法,如广义 PCA^[4],神经网络 PCA 方法^[6],主曲线方法^[6]

等。但这些方法都涉及到了复杂的非线性变换及优化问题。本文引入一种新的非线性主元分析的故障检测方法——核主元分析法(KPCA),并将其应用于航空发动机的故障检测中。

2 基于 KPCA 的故障检测原理

2.1 KPCA 基本原理

KPCA 方法是 Scholkopf 等在研究支持向量分类算法提出的^[7]。它的主要思想是通过某种事先选择的非线性映射 Φ 将输入向量 $\mathbf{x}$ 映射到一个高维特征空间 F ,然后在高维特征空间计算主元成分,在高维特征空间中得到的线性主元实质就是原始输入空间的非线性主元。

但是非线性映射 Φ 往往不容易求得,KPCA 是通过巧妙使用核函数来完成从输入空间到特征空间的非线性映射的。

定义核函数 $k(\mathbf{x}, \mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle$,这样特征空间中两向量的内积可以用输入空间中的两变量的核函数来表示。

任一向量 $\mathbf{x}$ 在特征空间 F 中的第 k 主元(得分向量)为^[7]

$$t_k = \langle \mathbf{V}^k, \Phi(\mathbf{x}) \rangle = \sum_{j=1}^n a_j^k \langle \Phi(\mathbf{x}_j), \Phi(\mathbf{x}) \rangle \quad (1)$$

式中 $\mathbf{x}$ 为原始空间的输入向量; $\Phi(\mathbf{x})$ 为特征空间 F 的映射向量; $\mathbf{V}^k$ 为特征空间 F 中第 k 个特征向量。

至于核函数的选择,常见的主要有 3 类^[8]:高斯径向基函数、多项式核函数和神经网络核函数,一般可选为高斯径向基函数 $k(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2 / \sigma)$, σ 为需选定的参数。

2.2 基于 KPCA 的故障检测

KPCA 变换完成后,故障的检测是通过监控两个统计量 T^2 和 SPE 的变化来实现的^[9],其中 SPE 也称为 Q 统计量。前者监控数据在 KPCA 主元空间中的变化,后者监控数据在 KPCA 残差空间中的变化。其具体的表达式见文献[9]。

基于 KPCA 的故障检测过程包括离线正常状态主元模型确定和在线故障检测两部分。

2.2.1 建立正常工作状态的主元模型步骤

(1)将正常状态的数据集进行标准化,变换为均值为 0 方差为 1 的标准数据集;

(2)计算标准数据集 $\mathbf{x}_k \in R^m$ ($k=1, \dots, n$) 的核矩阵 $K \in R^{n \times n}$;

$$K_{ij} = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle \quad (2)$$

(3)在特征空间 F 中,为了使均值为 $\underline{0}$ 需对核矩阵 K 作均值化处理,得到 $\underline{K}$;

$$K_{ij} = (K - l_n K - K l_n + l_n K l_n)_{ij} \quad (3)$$

式中 l_n 是元素为 $1/n$ 的 $n \times n$ 常数方阵。

(4)解决 $n \lambda \alpha = K \alpha$ 的特征值问题,并标准化 α^k ,使 $\langle \alpha^k, \alpha^k \rangle = 1$;

(5)针对正常状态数据向量 $\mathbf{x}$,可通过式 (1)提取非线性主元;

(6)计算正常状态数据集的统计量 T^2 和 SPE ;

(7)确定正常状态统计量 T^2 的控制限 T_a^2 和 SPE 的控制限 SPE_a 。

2.2.2 在线故障检测步骤

(1)从采样中获得新的数据 $\mathbf{x}_i$,并按照正常状态下模型的均值和方差进行标准化;

(2)对标准化后的新数据 $\mathbf{x}_i \in R^m$,计算其内核向量 $\mathbf{k}_i \in R^n$;

(3)对新数据的内核向量进行均值中心化处理 $\underline{\mathbf{k}}_i$;

(4)提取新数据的非线性主元 $\mathbf{t}_i$;

(5)计算新数据 $\mathbf{x}_i$ 的统计量 T^2 和 SPE ;

(6)监控统计量 T^2 和 SPE 是否超过正常状态的控制限 T_a^2 和 SPE_a 。

3 基于 KPCA 的航空发动机故障检测

根据第 2 节的思路和方法步骤,本文采用文献 [3] 的某型涡扇发动机的飞参记录数据对该发动机进行了故障检测。

3.1 KPCA 主元检测模型的建立

选取每次起飞前稳态飞参记录的 9 个参数作为一组表征该次发动机工作状况的参数^[2],9 个参数分别为高压转子换算转速 n_{hcor} 、低压转子换算转速 n_{lcor} 、低压导流叶片角度 a_1 、高压导流叶片角度 a_2 、振动值 B 、滑油压力 p_l 、涡轮后排气温度 T_4 、转差率 S 、尾喷口指示值 le 。

共取 200 组正常工作的飞参记录参数作为发动机正常状态的样本集。根据这 200 组数据建立发动机正常工作状态下的 KPCA 主元模型。采用常用的高斯径向基函数 $k(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2 / \sigma)$ 作为核函数,并依据经验选取 $\sigma = 5m^{[9]}$, m 为原始输入空间的维数。主元个数选取方法采用方差累积贡献率法,方差累积贡献超过 85% 的主元个数为需要保留的主元数。各主元贡献率如表 1 所示 (只列出了一部分),可见前 8 个主元的贡献率累积和百分比为

92.62%,因而可选取前 8 个主元构成主元模型,分别计算该模型的统计量,并得出置信度为 99% 的控制限为 $T_a^2 = 20.812$ $SPE_a = 0.953$ 。

Table 1 Contribution and its cumulative contribution of PC (Principal component) of KPCA

PC	Eigenvalue	Contribution %	Cumulative contribution %
1	0.0060	17.19	17.19
2	0.0054	15.47	32.67
...	...	...	...
8	0.0028	8.01	92.62
...	...	...	...
200	0.0000	0.00	100

为了便于比较,本文还建立发动机正常状态的 PCA 主元模型。主元个数选取方法采用方差累积贡献率法。各主元贡献率如表 2 所示 (只列出了一部分),可见前 3 个主元的贡献率累积和百分比为 95.15%,因而可选取前 3 个主元构成主元模型,此时得出置信度为 99% 的控制限为 $T_a^2 = 11.513$ $SPE_a = 17.236$ 。

Table 2 Contribution and its cumulative contribution of PC (Principal component) of PCA

PC	Eigenvalue	Contribution %	Cumulative contribution %
1	2.9304	40.95	40.95
2	2.3226	32.46	73.41
3	1.5561	21.74	95.15
...	...	...	...
9	0.0023	0.03	100

3.2 应用实例

为了验证所建立的模型的有效性,采用某台工作 109 次,工作时间为 122 h 的发动机飞参记录数据进行了检验。该发动机的使用状况是在装机工作第 36 h 加力喷口调节器故障;而在其故障前的一段时间内,由于调节器的有关工作参数已出现漂移,导致发动机的有关参数也偏离正常值,故该段时间内发动机实际上已处于异常工作状态;在第 36 h 之后由于更换了调节器,发动机性能恢复正常。

为了便于比较,本文列出了文献 [3] 采用综合参数法确定的性能综合指数对该发动机进行性能监控的结果,如图 1 所示。图 1 中两条曲线分别是采用专家调查法和遗传算法确定的性能综合指数,由图 1 可

看出,两种方法得到的性能综合指数值均出现过大幅下降,故该指数可有效反映发动机性能的恶化及故障,并且采用遗传算法确定的性能综合指数能更灵敏、准确地反映发动机性能的变化。但综合参数法并没有给出确定的性能综合指数警戒值,因此性能综合指数无法有效地检测发动机性能的异常及故障。

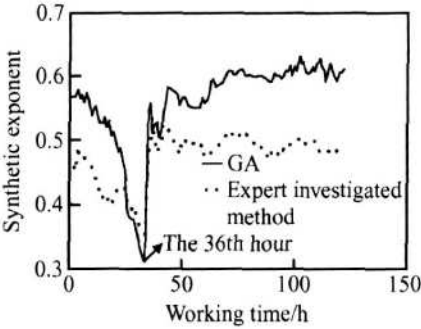


Fig. 1 Monitoring result of performance synthetic exponent

图 2和图 3分别是采用统计量 SPE 和 T^2 监控的结果。图 2和图 3的左半部分为 PCA 模型的检测结果;右半部分为 KPCA 模型的检测结果。表 3为两种模型的监控统计量在检测到故障前后时刻的比较,表 3中的统计量控制限的置信度为 99%。

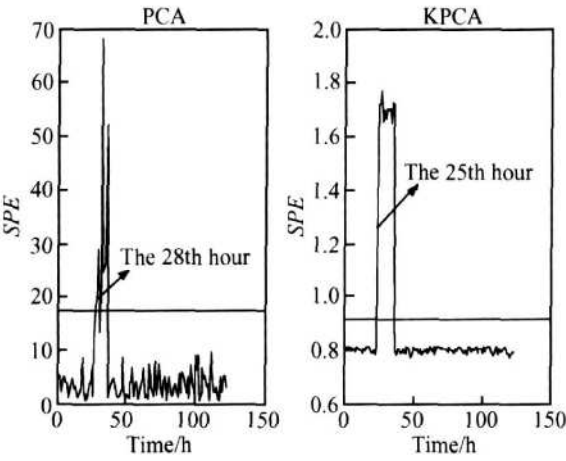


Fig. 2 Monitoring result of statistics SPE

由图 2图 3可知,统计量 SPE 和 T^2 有明确的控制限,因此可很容易地通过判断统计量是否超限来检

测故障。显然,两种模型在发动机故障阶段均检测到故障的发生。但由图 3的比较可知, KPCA 模型比 PCA 模型能更早地检测到故障: KPCA 模型的统计量在第 25 h就已超限,即在该时刻就开始检测到故障

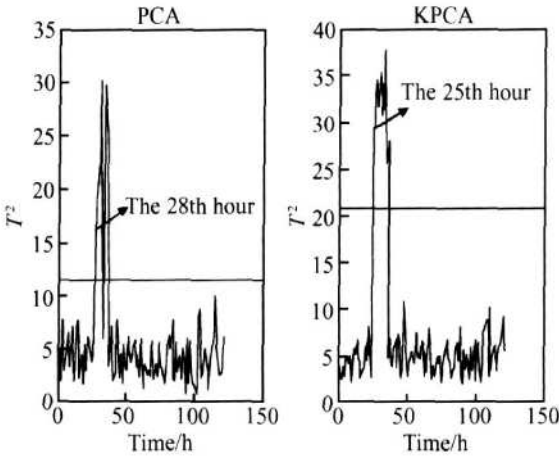


Fig. 3 Monitoring result of statistics T^2

的发生;而 PCA 模型的统计量在此时并没有超限,而是到第 28 h才超限,即到第 28 h才开始检测到故障。另外对于给定的 99% 的检验水平, PCA 模型出现两次错判情况(分别出现在 31.6 h和 34.1 h),而 KPCA 模型则没有出现错判的情况。可见, KPCA 模型的故障检测效果明显优于 PCA 模型。

3.3 结果分析

由表 1,表 2及 2.1节的理论分析可知, KPCA 主元数远多于 PCA 的主元数, KPCA 有 200个主元,而 PCA 只有 9个主元,因此 KPCA 主元携带的信息更为全面、丰富; KPCA 是从 200个主元中选取需保留的主元,而 PCA 模型只是从 9个主元中选取需保留的主元,并且最终 KPCA 选取保留的主元数要多于 PCA 的主元数,这样可避免丢失较小的主元中可能包含的重要非线性特征信息,同时还不会带来更多的噪声信息。这一点已得到证明, Scholkopf等人在 1999年证明 KPCA 可通过合适的主元选取方法来选取主元,达到既可保留更多的主元,同时还不会带来更多的噪声信息^[7]。因此 KPCA 模型选取的主元数较多,一方面

Table 3 Comparison of monitoring statistics of PCA with KPCA

Time/h	PCA				KPCA			
	T^2	T_a^2	SPE	SPE_a	T^2	T_a^2	SPE	SPE_a
24.1	9.124	11.513	12.126	17.236	12.733	20.812	0.892	0.953
25.0	9.657	11.513	14.655	17.236	21.058	20.812	0.967	0.953
26.9	10.234	11.513	15.201	17.236	25.219	20.812	1.133	0.953
28.0	11.935	11.513	17.932	17.236	22.932	20.812	1.501	0.953

可以使选取的主元携带的信息更为充分、全面,提取到重要的非线性特征信息;另一方面还不会引入更多的噪声信息,从而使得其非线性故障检测的效果更好。

PCA 模型虽然前三个主元的方差累积贡献率就已超过 95%,但可能存在重叠信息,夸大了前三个主元对工作状态的反映能力,并且舍弃的较小主元中可能包含一些重要的信息。因此 PCA 模型虽然保留的主元数要少,但其故障检测效果并不佳。

4 结 论

本文对 KPCA 在发动机性能监控与故障诊断中的应用进行了尝试,提出了一种基于 KPCA 的航空发动机故障检测方法。通过对某涡扇发动机进行实例验证分析,得出如下结论:

(1)与综合参数法相比, KPCA 方法不需要故障样本数据,解决了目前发动机故障样本少、较难获取的困难;由于综合参数法没有给出确定的性能综合指数的警戒值,因此无法准确判断是否有故障发生,而 KPCA 采用的监控统计量 T^2 和 SPE 有明确的控制限,故能够容易实现故障检测的目的。

(2)相对于 PCA 方法, KPCA 方法在非线形故障检测过程中能够提取重要的非线性特征信息,因此该方法可提高故障检测的快速性,能更早、更迅速地检测到潜在故障;另一方面可提高故障检测准确率,有助于减少误报警,具有更好的非线性故障检测能力。

(3)KPCA 用于故障检测也存在一些问题,例如核函数选取问题,根据研究对象的非线性程度不同,可选择不同的核函数;由于很难找到由特征空间到原始空间的逆映射函数,且核函数无法提供原测量变量与监控统计量之间的对应关系,故较难实现非线性情形的故障变量识别。这是 KPCA 方法用于故障检测的下一步研究的方向及重点。

参考文献:

- [1] 谢寿生. 某型飞机发动机故障诊断与性能趋势监测研究[D]. 西安: 西北工业大学, 1998
- [2] 胡金海, 谢寿生. 基于遗传算法的航空发动机性能监控与故障诊断[J]. 推进技术, 2003, 24(3): 198~200 (HU Jin-hai XIE Shou-sheng Performance monitoring and fault diagnosis of aeroengine based genetic algorithm[J]. *Journal of Propulsion Technology*, 2003, 24(3): 198~200)
- [3] 侯胜利, 胡金海, 李应红. 基于混沌变量的航空发动机性能监控与故障诊断[J]. 航空动力学报, 2005, 20(2): 314~317
- [4] 赵立杰, 柴天佑, 王 纲. 多元统计性能监视和故障诊断技术研究进展[J]. 信息与控制, 2004, 33(2): 197~201
- [5] 谢寿生. 某型加力涡轮风扇发动机[M]. 西安: 空军工程大学工程学院出版社, 1998
- [6] Dong D, Mcavoy T J. Nonlinear principal component analysis based on principal curves and neural networks[J]. *Computers and Chemical Engineering*, 1996, 20(1): 65~78
- [7] Scholkopf B, Smola A, Muller K R. Kernel principal component analysis[M]. In: Scholkopf B, Burges C, Smola A, eds. *Advances in kernel methods-support vector learning*. Cambridge MA: MIT Press 1999. 327~352
- [8] Baudat G, Anouar F. Kernel-based methods and function approximation[A]. *International Joint Conference on Neural Networks(IJCNN 01)*[C]. 2001, 1244~1249
- [9] LEE JM, YOO C K, CHOI S W, et al. Nonlinear process monitoring using kernel principal component analysis[J]. *Chemical Engineering Science*, 2004, 59(1): 223~234

(编辑: 朱立影)