



葛科宇

# 飞机液压系统磨损故障案例推理诊断方法

葛科宇<sup>1</sup> 陈 果<sup>1</sup> 侯民利<sup>2</sup> 李 爱<sup>1</sup>

(<sup>1</sup>南京航空航天大学 民航学院 南京 210016; <sup>2</sup>成都飞机工业(集团)有限责任公司 成都 610092)

**摘要:** 针对现有案例推理(CBR, case-based reasoning)中案例的检索和匹配对噪声数据和冗余数据较为敏感的问题。提出了一种核主成分分析(KPCA, kernel principal component analysis)与最近邻方法(k-NN, k-Nearest neighbor)相结合的新的案例检索与匹配方法。首先分别对待诊断数据和案例数据进行KPCA压缩,突出数据的主分量;然后,对压缩后的数据进行最近邻匹配;最后,开发出的飞机液压系统磨损综合监控专家系统,并用实际数据进行了验证。

**关键词:** 案例推理; 故障诊断; KPCA; 最近邻方法; 飞机液压系统

中图分类号: V23 文献标识码: A 文章编号: 1003-8728(2011)06-0922-05

## A Fault Diagnosis Method for Aircraft Hydraulic System by Case-based Reasoning

Ge Keyu<sup>1</sup>, Chen Guo<sup>1</sup>, Hou Minli<sup>2</sup>, Li Ai<sup>1</sup>

(<sup>1</sup> College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016;

<sup>2</sup> Chengdu Aircraft Industrial (Group) Co., Ltd, Chengdu 610092)

**Abstract:** Case retrieval and matching in existing case-based reasoning is more sensitive to noise data and redundant data. A new case retrieval model based on KPCA (Kernel Principal Component Analysis) and the k-NN (k-Nearest Neighbor) strategy was proposed. First, KPCA was used to compress diagnostic data and case data respectively, which highlights the main components of data. Then, nearest matching is done on compressed data. Finally, an aircraft hydraulic system condition monitoring expert system (AHMES) was developed. In addition, an actual data validation was carried out, which certified the effectiveness of the method.

**Key words:** case-based reasoning; fault diagnosis; KPCA; k-NN strategy; aircraft hydraulic system

随着航空业的发展,液压流体传动与控制技术在飞机上的应用也越来越广泛,与此同时据相关部门统计,液压系统的故障占整个飞行故障的1/3<sup>[1]</sup>。因此,如何为维修人员的故障诊断提供技术支持,提高飞机完好率和安全系数是飞机综合保障的重要研究内容。

油液监测技术是通过通过对设备中油液理化性能指标和油液中磨损金属颗粒及污染产物的分析,获取设备摩擦副润滑和磨损状态的信息,从而对设备的润滑状

态及磨损故障进行诊断的一种技术。油液监测技术已成为飞机液压系统故障诊断的重要技术手段<sup>[2-4]</sup>。

由于部队在飞机使用过程中记录了大量排故案例,积累了丰富的油样数据,因此,将CBR方法应用于飞机液压系统故障诊断和维护,可以避免规则难以提取、知识获取困难的瓶颈。同时,将历史记录中难以规则化的知识和经验隐含在案例中,以案例形式表达,直观、容易理解,能够反映故障的总体概貌,有利于现场人员参考。

一般地,CBR系统分为案例检索、案例重用、案例修正和案例存储4个步骤,其中案例检索是影响系统性能的关键因素<sup>[5]</sup>。它关系到推理过程的准确性和合理性。原始匹配算法有欧氏距离法、海明距离法、最近邻算法等。但这些匹配策略都有一个

收稿日期: 2010-03-18

基金项目: 成都飞机工业(集团)有限责任公司项目

作者简介: 葛科宇(1985-), 硕士研究生, 研究方向为数据挖掘、智

能诊断与专家系统, kyge1985@126.com; 陈 果(联系

人) 教授, 博士生导师, cgyzx@263.net

固有的缺点,很容易受无关属性和噪声数据的影响。目前,已有一些学者进行了相关的研究,文献[6]利用B-P网络模型来实现相似案例的检索,检索速度快。但匹配精度较低,训练网络的时间较长。文献[7]利用粗糙集理论结合相似案例的检索,通过粗糙集中的属性约简算法对案例冗余特征属性进行了优化,提高了检索效率。

引入了一种新的案例检索匹配方法,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)方法对待分析数据和案例进行特征压缩,然后再进行最近邻匹配,最后利用实际的油样数据进行的验证分析,证明了该方法的有效性。

## 1 基于CBR的飞机液压系统磨损故障诊断关键技术

### 1.1 构建案例库

笔者涉及的飞机液压系统油样案例主要来源于部队使用过程中的排故和维修记录。通过对某型号飞机液压系统故障油样数据进行收集、整理、知识提取,获取了典型飞机故障案例。通过对该型号飞机液压系统故障情况的分析,确定案例应包含对问题整体情况的描述,还应包含对问题的解或解的方法的描述。即

base = {  $C_1, C_2, \dots, C_n$  } 表示案例库,其中  $C_i = C[F_i, S_i]$  表示案例中的第  $i$  个案例;

$F_i = [f_{i1}, f_{i2}, \dots, f_{in}]$  表示案例  $C_i$  的征兆集,  $f_{in}$  表示案例  $C_i$  的第  $n$  个征兆;

$S_i = [S_{i1}, S_{i2}, \dots, S_{in}]$  表示案例  $C_i$  的解决方案集,  $S_{in}$  表示  $C_i$  的第  $n$  个解决方案(或者解决方案的第  $n$  个步骤)。

案例代表的是过去的经验的积累,并不是一条规则,在案例库中每个案例的描述的是一种故障状态,它们是相互独立的。案例库索引策略的合理与否直接影响到案例的检索效率,合理的索引结构可以缩小搜索范围,提高检索效率。根据飞机液压系统故障特点以及专家建议,笔者对飞机液压系统案例库采用基本信息类(如:案例号、飞机型号、液压系统名称)建立索引结构。

### 1.2 基于KPCA案例特征提取意义

设计合理的案例匹配方法是CBR技术的关键之一。然而,由于油样数据(包括:颗粒计数分析数据、光谱分析数据、铁谱分析数据、油品理化性能分析数据)维数较高,数据间存在很大的冗余信息,如何有效地对油样数据进行压缩和降维,有效地实现油样数据特征提取,是提高油样数据匹配的成功率的关键。

目前,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<sup>[8,9]</sup>,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。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。有鉴于此,笔者将KPCA方法应用于飞机液压系统油液的特征提取,实现油样数据的压缩与降维处理,然后根据所提取的特征,利用最近邻法实现飞机液压系统的故障诊断。

### 1.3 核主成分分析(KPCA)原理

最早将核方法思想应用于特征抽取的是 scholkopf 等人<sup>[10]</sup>,于1999年借鉴SVM的核方法思想,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,即Kernel PCA是通过一个非线性函数  $\Phi$  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

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

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  $\Phi$ , 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  $\lambda$  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

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

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,即可获得要求得特征值和特征向量。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。

此时式(10)中的  $K$  用  $\tilde{K}$  代替。

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

式中:  $1_{ij} = 1$  (对于所有的  $i, j$ )。

KPCA 算法可以分为以下三个步骤:

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,可选择最近邻分类器<sup>[10]</sup>进行分类。

## 2 飞机液压系统磨损的 CBR 故障专家系统

### 2.1 诊断流程

笔者所提出的案例检索与匹配方法,已经应用于针对某军用飞机液压系统磨损故障诊所开发的飞机液压系统状态监控专家系统(AHMES),该系统的 CBR 诊断流程如图 1 所示。图 2 为专家系统主界面,图 3 为案例添加界面,图 4 为 CBR 诊断结论界面。目前该系统已成功用于成都飞机工业(集团)有限责任公司,应用在歼 10 飞机液压系统的状态监控。

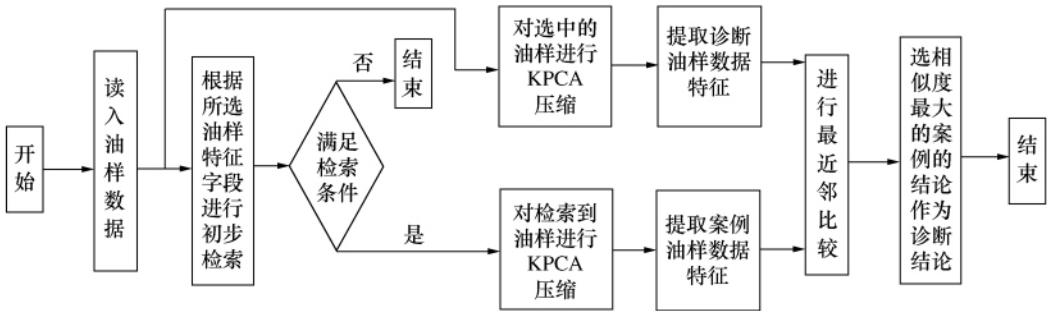


图 1 故障案例检索与匹配的基本流程



图 2 专家系统主界面



图 3 案例添加界面



图 4 诊断结论界面

## 2.2 基于 KPCA 的油样案例诊断

笔者列举一实例,验证基于 CBR 的飞机液压系统磨损故障专家系统(AHMES)故障案例特征提取方法的有效性。成都飞机工业(集团)有限责任公司通过对外场歼 10 飞机液压系统油样分析跟踪,得到国产歼 10 飞机液压系统油样分析数据。挑选 60 条油样案例数据列于表 1。

表 1 歼 10 飞机液压系统油样分析数据

		1	2	3	4	5	6	7	...	60
光谱分析数据 (ppm)	Fe	0.20	0.40	0.60	0.60	0.20	0.30	0.40	...	0.10
	Cr	0.40	0.20	0.20	0.40	0.30	0.30	0.40	...	0.60
	...	...	...	...	...	...	...	...	...	...
	V	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.10
理化分析数	粘度/(mm <sup>2</sup> ·s <sup>-1</sup> )	11.7	11.9	11.7	12.1	11.9	11.7	11.0	...	11.2
	含水量/ppm	82.0	96.9	94.4	87.5	75.7	96.9	51.1	...	44.3
颗粒计数数据	污染度等级(GJB420A)	6	5	5	6	6	6	6	...	7

由表 1 可知油样分析数据有光谱数据、理化分析数据、污染度等级数据组成。该数据集共有 21 维 60 条记录。随着飞机飞行时间的延长和油样分析手段的增加,案例库中案例的条数和维数必将增加。这将给案例的检索和匹配带来很大的困难。因

$$\begin{bmatrix} -0.0003 & 0.0003 & -0.0002 & -0.0004 & 0.0002 & -0.0018 & -0.0020 & \dots & 0.0005 \\ -0.0001 & -0.0001 & -0.0001 & -0.0001 & -0.0001 & -0.0001 & -0.0001 & \dots & 0.0002 \\ 0.0000 & -0.0000 & -0.0000 & -0.0000 & 0.0000 & -0.0001 & 0.0000 & \dots & 0.0007 \end{bmatrix}_{3 \times 60}^T$$

此对油样数据进行 KPCA 压缩和特征提取显得非常必要。经(核参数  $p = 0.001$ ,核函数为  $p$  阶多项式核函数)的 KPCA 压缩后,得到的能量保持率在 95% 的特征矩阵为

CA 方法的识别率进行比较。 $p$  阶多项式核函数为

$$K(x, x_i) = [(x \cdot x_i) + 1]^p \quad (14)$$

从特征矩阵可以看出,KPCA 将原来 21 维的油样案例数据压缩成了 3 维,并且能够将能量保持在 95% 以上。

由于核参数的选取对 KPCA 的特征提取效果具有很大影响,目前尚没有统一的选择理论和依据<sup>[12]</sup>。为了验证经 KPAC 压缩后,不同的核参数对案例油样的识别效果,选取了多项式核函数对 KP-

现将 60 条油样案例叠加 20% 的随机高斯噪声,作为测试样本,进行最近邻识别。识别结果如表 2 所示。从表 2 中可以看出,不同的核参数对特征提取和结果识别有很大的影响,最高识别率可以达到 95.28%,最低仅为 68.34%,平均为 85.81%。

表 2 不同的核参数下 KPCA 对应的识别结果

多项式核函数	核参数 $p$	0.001	0.01	0.05	0.1	0.5	1
	识别率	95.28%	93.33%	96.71%	91.67%	70.00%	68.34%

为了验证 KPCA 对噪声数据的识别效果。现将 60 条油样案例叠加不同强度的随机高斯噪声,作为测试样本,与 PCA 压缩进行比较。图 5 为不同高斯噪声强度下 PCA 与 KPAC 的识别效果比较图。其

中 KPCA 的核参数  $p$  取 0.05。比较结果如表 3 所示。实验表明,在相同的能量保持率和噪声强度下,KPCA 的识别率均大于 PCA。随着高斯噪声强度的增加,KPCA 相对于 PCA 的优势更加明显。

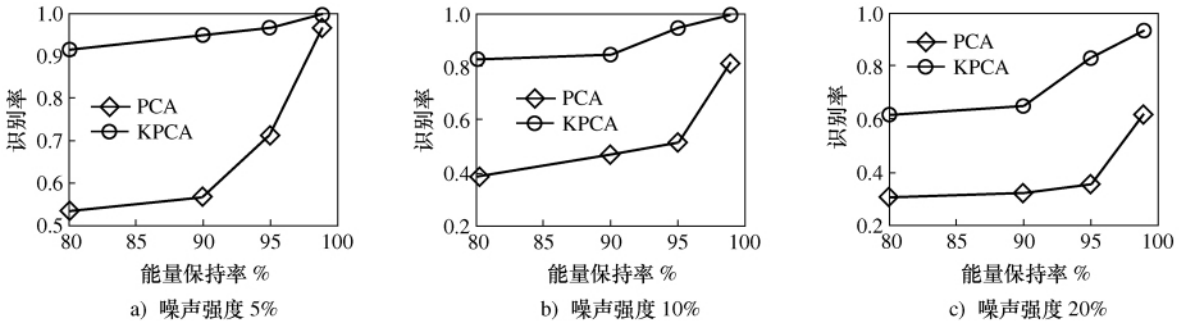


图5 不同强度高斯噪声下的PCA与KPCA识别效果图

表3 不同强度高斯噪声下PCA与KPCA的识别效果对比

高斯噪声强度	5%				10%				20%			
	99	95	90	80	99	95	90	80	99	95	90	80
能量保持率/%	99	95	90	80	99	95	90	80	99	95	90	80
PCA 识别率/%	96.7	71.3	56.7	53.3	81.6	51.6	46.7	38.8	61.6	35	31.6	30
KPCA 识别率/%	100	96.7	95	91.6	100	95	85	83.3	93.7	83.3	65	61.6

为了进一步验证此方法,笔者任意选择一条待诊断条油样数据,如表4所示。在不压缩和经KPCA压缩后的情况下分别与案例库中的案例(表1中的60条数据)进行最近邻匹配。对比结果,如图7所示。从图中可以看出,待诊断油样数据在压缩和不压缩的情况下都是和案例库中的第8条案例距离最近,但经KPCA压缩后,由于突出了油样的主分量,降低了无关冗余属性的影响,与无特征压缩的油样数据相比,案例匹配的最近距离增大,从而具有更

好的抗噪性能。另外,由于特征维数的压缩,也将使得案例存储空间大大减小,随着案例数的增加,计算速度将有明显的改善。

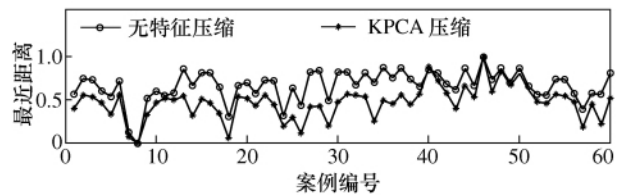


图6 油样数据和案例油样之间最近距离

表4 待诊断油样数据

光谱分析数据(ppm)									理化分析数据		颗粒计数数据
Fe	Cr	Pb	Cu	Sn	Al	Ni	Ag	Si	粘度	含水量	污染度等级
0.11	0.59	0.00	0.40	0.42	0.30	0.02	0.01	16.0	/(mm <sup>2</sup> ·s <sup>-1</sup> )	/ppm	(GJB420A)
B	Na	Mg	Ca	Ba	Zn	Mo	Ti	V			
0.00	0.80	0.01	0.00	0.00	0.30	0.80	0.90	0.10	11.25	44.3	6

3 结束语

针对飞机液压系统故障诊断,规则难以提取、知识获取困难的瓶颈问题,提出一种基于CBR的飞机液压系统诊断方法。运用了关键字段匹配、核主成分分析和最近邻方法相结合的检索模型实现了案例的特征提取、案例检索和匹配,取得了比较满意的匹配效果。利用发了飞机液压系统状态监控专家系统,并用实例验证了该方法应用于飞机液压系统的可行性与有效性。

[参考文献]

[1] 唐有才. 飞机液压系统污染及控制[J]. 机床与液压, 2002, (3): 206 ~ 208  
 [2] 贺石中. 液压设备的润滑磨损故障及监测诊断[J]. 液压气动与密封, 2004 (5): 23 ~ 24  
 [3] 陈果, 陈立波, 宋兰琪. 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

[4] 陈果. 航空发动机磨损故障的智能融合诊断[J]. 中国机械工程, 2005, 16(4)  
 [5] David B L, Andrew K. Case-based CBR capturing and reusing reasoning about case adaptation [J]. *International Journal of Expert Systems*, 1997, 10(2): 197 ~ 213  
 [6] 张荣梅. 智能决策支持系统研究开发及应用[M]. 北京: 冶金工业出版社, 2003  
 [7] 杨善林等. 机器学习与智能决策支持系统[M]. 北京: 科学出版社, 2004  
 [8] 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  
 [9] Scholkopf B, et al. *Kernel PCA Pattern Reconstruction Via Approximate Pre-images* [M]. In ICANN. Springer Verlag, 1998  
 [10] 肖健华. 智能模式识别方法[M]. 广州: 华南理工大学出版社, 2005