

文章编号: 1001-0920(2007)07-0783-04

## 基于模式识别的传感器故障诊断

徐涛, 王祁

(哈尔滨工业大学 自动化测试与控制系, 哈尔滨 150001)

**摘要:** 为满足模式识别故障诊断算法的鲁棒性要求, 在小波包分解提取特征向量的基础上, 提出了有监督模式分类与无监督模式分类相结合的故障诊断方法. 利用小波包分解提取各个频带的能量作为特征向量; 采用 LVQ 神经网络作为有监督的模式分类器进行故障诊断; 运用无监督的减法聚类方法对新型故障模式进行辨识. 最后, 通过动力系统管路流量传感器数据对算法进行检验, 验证了所提出方法的实用性和有效性.

**关键词:** 模式识别; 小波包; LVQ 神经网络; 减法聚类; 传感器故障诊断

中图分类号: TH133; TP183

文献标识码: A

### Sensor fault diagnosis based on pattern recognition

XU Tao, WANG Qi

(Department of Automatic Test and Control, Harbin Institute of Technology, Harbin 150001, China.

Correspondent: XU Tao, E-mail: xtwyh@sohu.com)

**Abstract:** To meet the robustness of the fault diagnosis algorithm, a method is proposed, which combines the supervised classification and unsupervised classification based on the feature extraction with wavelet package decomposition. As the pattern vector, the energy in different frequency with the wavelet package decomposition is calculated. Then, learning vector quantity neural network is employed as the supervised classification for fault diagnosis. As the supervised classification, subtractive clustering is applied to identify the novel fault pattern. Finally, the applicability and effectiveness of the proposed methodology are illustrated by flow sensor data of the dynamical system.

**Key words:** Pattern recognition; Wavelet package; LVQ neural network; Subtractive clustering; Sensor fault diagnosis

### 1 引言

在得不到准确的系统数学模型时, 进行故障检测和诊断的一个有效途径是采用模式识别方法<sup>[1-4]</sup>. 采用模式识别方法进行传感器的故障诊断的前提条件是必须收集并保存好正常操作以及故障状态下的历史数据; 然后将这些历史数据作为训练样本进行处理, 提取出关键信息. 这一过程称为特征提取, 其目标是提取出那些可以明确地识别出的某一类型, 并可以区分其他类型的特征. 接下来的步骤是要设计出一个决策算法, 基于提取出的特征集合, 最优地对采集到的样本进行分类<sup>[5]</sup>. 当新的样本出现时, 首先提取出该样本中的特征信息; 然后便可以基于上述决策算法, 判定是哪一种故障最有可能产生了这一样本.

基于模式识别的故障诊断方法中, 诊断算法应

满足如下鲁棒性要求<sup>[6]</sup>:

1) 故障的精确起始时间可能是未知的, 有可能仅知道大致的时间范围. 诊断算法应对这种不确定性具有鲁棒性.

2) 诊断算法应对故障幅值不敏感. 例如, 传感器偏差这种阶跃型故障有可能具有不同的幅值, 其对应样本的幅值也有所不同, 但其故障原因是相同的.

3) 训练样本中不存在的新故障发生时, 诊断算法应该能够指出这是一种新故障, 与已知的任何故障都不相象.

本文将模式识别的有监督分类与无监督分类相结合, 利用小波包分解提取各个频段的能量作为特征向量, 采用 LVQ (Learning vector quantity) 神经网络作为有监督分类器对传感器进行故障诊断, 采

收稿日期: 2006-03-09; 修回日期: 2006-04-10.

基金项目: 国家自然科学基金项目 (60572010).

作者简介: 徐涛 (1971 →), 男, 辽宁彰武人, 博士生, 从事传感器故障诊断及数据重构的研究; 王祁 (1944 →), 男, 哈尔滨人, 教授, 博士生导师, 从事多传感器信息融合、先进故障诊断理论和人工智能等研究.

用减法聚类作为无监督分类器对测试样本中是否出现新的故障类型进行辨识,从而给出了一种满足上述鲁棒性要求的传感器故障诊断方法。

## 2 基于小波包的特征向量提取

小波包可以对多分辨分析的细尺度进一步分解,从而提高频率分辨率,是一种比多分辨分析更加精细的分解方法,具有更好的时频特性,已广泛应用于特征信息的提取。文献[7]在各个尺度上进行重构,提取时域信号的能量作为特征信息。文献[8]则对小波包分解系数进行削减,使得故障特征更加突出,再经重构后提取时域信号的能量作为特征向量,进行故障分类。由于小波包具有更好的时频局部性特征,这种特征能量的提取方法可以解决鲁棒性要求的第1点:对故障时间具有鲁棒性,但是对于鲁棒性要求的第2点则没有满足。为了使得算法对于故障幅值不敏感,本文在小波包分解之前对传感器信号标准化,从而去除信号幅值的影响,使得算法对故障幅度具有鲁棒性。

具体的特征提取步骤如下:

Step1: 对长度为  $n$  的传感器信号  $X$  进行标准化,即

$$\bar{X} = D^{-1} [X - E(x)]. \quad (1)$$

其中:  $E(x)$  为  $X$  对应的均值,  $D$  为  $X$  的标准差。

Step2: 对  $\bar{X}$  进行三层小波包分解,以  $X_{30}, X_{31}, \dots, X_{37}$  分别表示第3层从低到高的各个节点的分解系数向量。

Step3: 为了更准确地描述信号中能量的变化,对  $X_{30}, X_{31}, \dots, X_{37}$  采取一定的削减算法,即:

1) 计算各个节点分解系数的削减阈值

$$\text{Threshold}_{X_{3i}} = \sqrt[k]{\frac{n^8}{k} X_{3i}}, \quad (2)$$

其中  $k$  表示  $X_{3i}$  的长度。

2) 对应  $X_{3i}$  中的分解系数  $\text{cf}_{S_{3i}}$  进行如下处理:

$$\text{cf}_{S_{3i}} = \begin{cases} \text{cf}_{S_{3i}} / \text{cf}_{S_{3i}} / \text{Threshold}_{X_{3i}}; & \text{cf}_{S_{3i}} / \text{cf}_{S_{3i}} / \text{Threshold}_{X_{3i}}; \\ 0, & \text{cf}_{S_{3i}} / \text{cf}_{S_{3i}} / \text{Threshold}_{X_{3i}} < \text{Threshold}_{X_{3i}}. \end{cases} \quad (3)$$

Step4: 对式(3)中得到的各个节点分解系数分别进行小波包重构,得到各个节点对应的重构数据  $S_{30}, S_{31}, \dots, S_{37}$ 。

Step5: 计算各频带信号的总能量

$$E_{3i} = \int |S_{3i}|^2 dt = \int |S_{3i}|^2. \quad (4)$$

Step6: 因为计算的总能量  $E_{3i}$  值较大,不利于模式分类,所以对  $E_{3i}$  进行归一化处理,即

$$\bar{E}_{3i} = E_{3i} / \sum_{i=0}^7 E_{3i}. \quad (5)$$

Step7: 构造特征向量

$$T = [\bar{E}_{30} \quad \bar{E}_{31} \quad \dots \quad \bar{E}_{37}]^T,$$

作为神经网络的输入。

## 3 基于LVQ神经网络的传感器故障诊断

文献[9,10]分别采用前向BP网络和RBF网络作为有监督分类器对故障模式进行分类。LVQ网络能够对任意输入训练进行分类,不管它们是否线性可分<sup>[11]</sup>,其分类的结果都是明确的。而前馈神经网络则需要进行2次判断来决定分类结果,这一点比前向网络优越。因此,本文选择LVQ网络作为模式分类器对传感器进行故障诊断,判断传感器的状态。

LVQ网络属于基本竞争型人工神经网络,由输入层、竞争层和线性输出层组成。输入有  $N$  个神经元,竞争层有  $M$  个神经元,输出层对应于网络的基本结构如图1所示。其中网络的连接权为  $\{w_{ij}\}$ ,  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ,且约束条件为

$$\sum_{i=1}^N w_{ij} = 1. \quad (6)$$

网络的  $T$  个二值输入学习模式为  $P_k = (p_1^k, p_2^k, \dots, p_N^k)$ ,与其对应的竞争层输出模式为  $A_k = (a_1^k, a_2^k, \dots, a_M^k)$ ,  $k = 1, 2, \dots, T$ 。

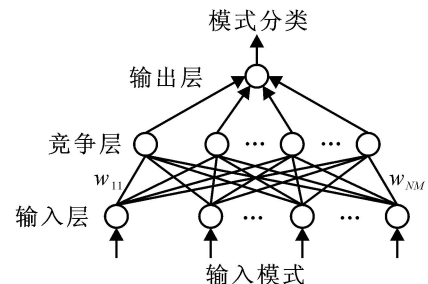


图1 LVQ网络结构

## 4 基于减法聚类的传感器新型故障辨识

减法聚类是一种快速而独立的相似聚类方法<sup>[12,13]</sup>,与其他常用的数据聚类算法不同的是减法聚类算法的聚类中心候选集为样本数据点。因此,采用此方法时的计算量与数据点的数目成线性关系,且与所考虑问题的维数无关。

设样本为  $M$  维空间中的  $n$  个数据点  $(x_1, x_2, \dots, x_n)$ 。不失一般性,假设数据点已经归一化到一个超立方体。因为每个数据点都是聚类中心的候选点,所以数据点  $x_i$  处的密度指标定义为

$$D_i = \prod_{j=1}^n \exp\left(-\frac{x_i - x_j}{(r_a/2)^2}\right), \quad (7)$$

这里  $r_a$  是一个正数。显然,如果一个数据点有多个邻近的数据点,则该数据点具有高密度值。半径  $r_a$  定义了该点的一个领域,半径以外的数据点对该点的密度指标贡献甚微。

在计算每个数据点密度指标后,选择具有高密度指标的数据点为第1个聚类中心。令  $x_{c1}$  为选中的

点,  $D_{cl}$  为其密度指标, 那么每个数据点  $x_i$  的密度指标可用下式修正:

$$D_i = D_i - D_{cl} \exp\left(-\frac{x_i - x_{cl}}{(r_b/2)^2}\right), \quad (8)$$

其中  $r_b$  是一个正数. 显然, 靠近第 1 个聚类中心  $x_{cl}$  的数据点的密度指标显著减少, 从而这些点不太可能成为下一个聚类中心. 常数  $r_b$  定义了一个密度指标函数显著减少的领域, 通常大于  $r_a$ . 为避免出现相距很近的聚类中心, 一般取  $r_b = 1.5 r_a$ .

修正了每个数据点的密度指标后, 选定下一个聚类中心  $x_{c2}$ , 再修正数据点所有密度函数. 该过程不断重复, 直至产生足够多的聚类中心.

首先对训练样本进行减法聚类分析, 确定训练样本的分类数. 进行 LVQ 网络分类之前, 对测试样本进行减法聚类分析, 再与前面的聚类结果进行对比, 确定是否有新的聚类中心出现, 从而实现对新

传感器故障的辨识.

### 5 算法仿真验证

本文以某动力系统管路流量传感器的 4 种典型故障(漂移、偏差、冲击、周期性干扰)为例进行算法研究. 首先提取长度为 400 的传感器信号, 根据第 2 节介绍的步骤提取特征能量. 取 3 层小波包分解提取的特征向量的长度为 8, 所以设计的 LVQ 网络的输入节点数确定为 8. 需要注意的是, 网络的竞争层必须要有足够的神经元, 使得每一类有足够多的竞争神经元. 因此, 本文设计的 LVQ 网络具有 18 个竞争神经元. 网络的输出可分为 5 种模式(1, 2, 3, 4, 5), 直接对应传感器的 5 种状态(正常、漂移故障、偏差故障、冲击故障、周期性干扰).

提取传感器正常及 4 种故障模式下的各 10 组特征向量来训练所建立的 LVQ 神经网络. 表 1 分别给出了每种状态的 1 组训练样本.

表 1 LVQ 训练样本

传感器状态	输 入 样 本	模式
正常	[ 0.176 8 0.248 4 0.319 2 0.190 3 0.364 7 0.558 3 0.472 5 0.317 7 ]	1
漂移	[ 0.426 0 0.149 6 0.202 8 0.367 9 0.518 9 0.285 1 0.229 4 0.465 2 ]	2
偏差	[ 0.999 0 0.007 4 0.009 6 0.018 1 0.025 5 0.014 0 0.011 3 0.022 9 ]	3
冲击	[ 0.700 7 0.681 7 0.088 4 0.068 3 0.081 9 0.092 9 0.094 3 0.087 5 ]	4
周期性干扰	[ 0.348 0 0.235 4 0.215 0 0.390 8 0.514 0 0.309 9 0.226 4 0.461 6 ]	5

表 2 测试结果

测 试 数 据	所属状态	神经网络输出	诊断结果
[ 0.268 1 0.248 4 0.288 8 0.332 5 0.400 5 0.413 1 0.404 2 0.421 8 ]	正常	1	正常
[ 0.369 6 0.149 6 0.208 0 0.378 2 0.533 4 0.293 2 0.235 9 0.478 1 ]	漂移故障	2	漂移故障
[ 0.998 2 0.010 1 0.013 2 0.024 7 0.034 8 0.019 2 0.015 4 0.031 2 ]	偏差故障	3	偏差故障
[ 0.700 3 0.666 7 0.099 8 0.084 4 0.113 8 0.107 9 0.105 4 0.111 0 ]	冲击故障	4	冲击故障
[ 0.330 4 0.226 0 0.214 5 0.392 8 0.518 1 0.312 8 0.235 1 0.466 7 ]	周期性干扰	5	周期性干扰

利用上面得到的 50 组数据对所建立的 LVQ 网络进行训练, 训练后的分类结果如图 2 所示.

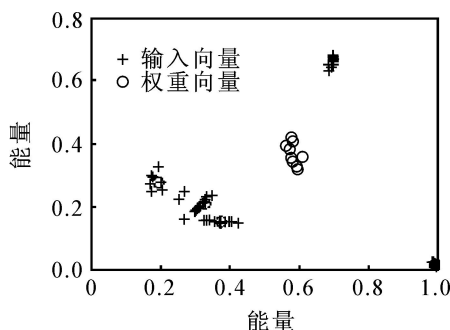


图 2 训练分类结果

另外, 提取传感器正常及各种故障模式的各 20 组测试数据, 此时故障数据的幅值和故障开始时间都与训练数据不同, 从而可检验算法的鲁棒性. 表 2

给出了一组输出结果.

对于这 100 组测试数据, 所建立的 LVQ 网络都识别的结果如表 3 所示, 平均准确率达到 95%. 而采用 RBF 网络的分辨率为 89%.

表 3 诊断结果

测试数据	识别准确率 / %
正常	100
漂移故障	100
偏差故障	100
冲击故障	100
周期性干扰	75

为了验证减法聚类方法对于传感器新型故障的辨识能力, 又仿真出 20 组传感器掉电故障的数据. 首先对原始训练数据及测试数据的集合进行减

法聚类分析. 为显示方便, 本文以特征向量的第 1 维和第 8 维数据为坐标显示二维分类图形. 聚类结果如图 3 所示, 得到了 3 个聚类中心. 然后再对原始训练数据和含有新型故障的数据进行减法聚类分析, 聚类结果如图 4 所示. 可见得到了 4 个聚类中心, 表明出现了新型的故障模式.

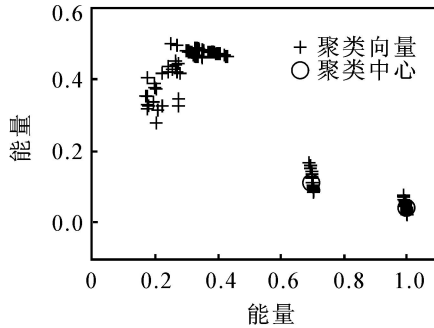


图 3 原始测试数据的减法聚类结果

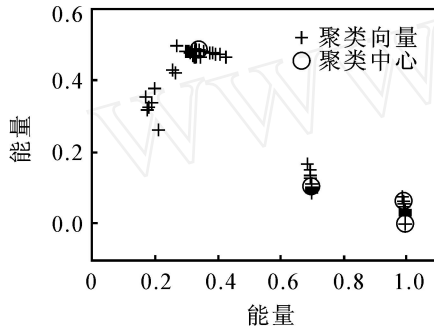


图 4 新测试数据的减法聚类结果

## 6 结 论

本文提出了一种有监督模式分类与无监督模式分类相结合的传感器故障诊断方法, 并结合小波包分解提取特征向量, 满足了诊断算法的鲁棒性要求. 通过流量传感器的故障诊断对所提出的算法进行了验证, 表明算法在完成传感器状态识别的基础上, 能够及时发现新的故障模式.

## 参考文献 (References)

- [1] He Yong, Feng Lei. Diesel fuel injection system faults diagnosis based on fuzzy injection pressure pattern recognition [C]. Proc of the World Congress on Intelligent Control and Automation. Huangzhou, 2004: 1654-1657.
- [2] Diallo Demba Benbouzid, Mohamed El Hachemi, Hamad Denis, et al. Fault detection and diagnosis in an induction machine drive: A pattern recognition approach based on concordia stator mean current vector[J]. IEEE Trans on Energy Conversion, 2005, 20(3): 512-519.
- [3] 代树武, 孙辉先. 基于模型的飞行器电源故障与故障模式识别[J]. 振动与冲击, 2005, 24(3): 55-58.  
(Dai Shu-wu, Sun Hui-xian. Model-based diagnosis and fault mode identification for spacecraft power system [J]. J of Vibration and Shock, 2005, 24(3): 55-58.)
- [4] 刘晓波, 黄其柏. 基于动态核聚类分析的水轮机组故障模式识别[J]. 华中科技大学学报, 2005, 33(9): 47-49, 52.  
(Liu Xiao-bo, Huang Qi-bai. Classification on the modes of hydro-generator unit fault based on dynamic kernel cluster analysis [J]. Huazhong University of Science & Technology, 2005, 33(9): 47-49, 52.)
- [5] Tou J T, Gonzalez R C. Pattern recognition principles [M]. MA: Addison-Wesley, 1974.
- [6] 周东华, 叶银忠. 现代故障诊断与容错控制[M]. 北京: 清华大学出版社, 2000.  
(Zhou Dong-hua, Ye Yin-zhong. Modern fault diagnosis and fault tolerant control [M]. Beijing: Tsinghua University Press, 2000.)
- [7] 孙斌, 周云龙. 基于支持向量机和小波包能量特征的气液两相流流型识别方法[J]. 中国电机工程学报, 2005, 25(17): 93-99.  
(Sun Bin, Zhou Yun-long. Identification method of gas-liquid two-phase flow regime based on support vector machine and wavelet package energy feature [J]. Proceedings of CSEE, 2005, 25(17): 93-99.)
- [8] 李文军, 张洪坤, 程秀生. 基于小波和神经网络的传感器故障诊断[J]. 吉林大学学报(工学版), 2004, 34(3): 491-495.  
(Li Wenjun, Zhang Hong-kun, Cheng Xiusheng. Sensor fault diagnosis based on neural network[J]. J of Jilin University (Engineering and Technology Edition), 2004, 34(3): 491-495.)
- [9] 王国峰, 王子良, 秦旭达, 等. 基于小波包和径向基神经网络轴承故障诊断[J]. 北京科技大学学报, 2004, 26(2): 184-187.  
(Wang Guo-feng, Wang Zi-liang, Qin Xu-da, et al. Accurate diagnosis of rolling bearing based on wavelet package and RBF neural networks[J]. J of University of Science and Technology Beijing, 2004, 26(2): 184-187.)
- [10] Berry, Emily A, Dalby, et al. Reduced bio basis function neural network for identification of protein phosphorylation sites: Comparison with pattern recognition algorithms[J]. Computational Biology and Chemistry, 2004, 28(1): 75-85.
- [11] Kohonen T. Self-organization and associative memory [M]. 3rd ed. Berlin: Springer-Verlag, 1989.
- [12] Chin S L. Fuzzy model identification based on cluster estimation[J]. J of Intelligent and Fuzzy Systems, 1994, 2(3): 267-278.
- [13] Ronald R Y, Dimitar P F. Essentials of fuzzy modeling and control[M]. New York: John Wiley & Sons Inc, 1994.