

文章编号: 1000-5013(2011)01-0013-05

# 利用 AR 模型和支持向量机的 调速阀故障识别

曾小军, 黄宜坚

(华侨大学 机电及自动化学院, 福建 泉州 362021)

**摘要:** 提出一种基于时间序列的自回归(AR)模型和支持向量机故障识别方法. 以液压调速阀的故障识别为例, 利用采集到的调速阀体的振动信号建立 AR 模型; 然后, 将 AR 模型自回归系数和残差方差组成的特征向量输入到支持向量机. 最后, 通过支持向量机完成对调速阀的正常和各种故障工况的分类识别. 实验结果和分析表明, 识别率不仅与核函数的选取有关系, 而且与支持向量机参数的选取也有关系, 以径向基 RBF 为核函数的识别率明显优于以多项式形式为核函数的识别率.

**关键词:** 调速阀; 故障识别; 自回归模型; 支持向量机

**中图分类号:** TH 165<sup>+</sup>.3; TH 137.52<sup>+</sup>2

**文献标识码:** A

调速阀是液压元件中结构较为复杂、容易发生故障的一类典型的液压元件, 因此, 调速阀的故障识别对于液压系统的维护有着重要的意义<sup>[1]</sup>. 典型的故障诊断系统由数据获取、预处理、特征提取和选择, 以及分类决策 4 个部分组成. 信息获取主要针对能够反映机械运行状态的信息, 如振动信号、声音信号等. 预处理的目的是去除噪声, 加强有用信息. 特征选择和提取是为了有效地实现分类识别, 就要对原始数据进行分析变换, 得到最能反映分类本质的特征. 分类决策是将被识别对象归为某一类, 也就是区分正常状态与各种异常状态<sup>[2]</sup>. 本文提出一种基于 AR 模型和支持向量机的液压调速阀故障诊断方法.

## 1 调速阀体的振动信号采集实验

### 1.1 检测系统

调速阀和检测系统的结构简图, 如图 1, 2 所示. 测控系统的硬件有计算机、PS-3030D 型直流电源(北京固纬电子有限公司)、ST-1-03 型非接触式电涡流位移传感器(北京昆仑海岸公司)、数据采集卡 PCI-6014 及接线端子 8LP(美国国家仪器公司). 采用 Lab VIEW 7.1 编写系统检测程序. 通过单通数据连续采集信号, 再对采样信号进行滤波, 滤波后的信号进行单边 Fourier 变换, 求取信号频谱. 全部测试和分析由计算机自动完成.

### 1.2 测试过程

实验中, 人为设置 3 种故障: (1) 节流阀后腔的弹簧变形; (2) 节流阀弹簧里有异物; (3) 节流阀后腔缺圆柱铁芯. 调速阀的故障会在阀体工作时产生的振动信号中表现出来. 分别采集正常和各种故障情况下的振动位移信号进行数据处理分析.

### 1.3 数据预处理

分别选取几组在正常状态和故障状态下采集的数据进行分析. 由于信号的变化趋势会影响故障诊断的全面性和准确性<sup>[3]</sup>, 所以对所取数据先进行预处理. 正常情况和故障 1 在滤波处理前后的电压数据, 如图 3 所示. 由图 3 可知, 位移电压信号是一种平稳时间序列.

**收稿日期:** 2010-05-23

**通信作者:** 黄宜坚(1945-), 男, 教授, 主要从事现代谱分析与故障诊断的研究. E-mail: yjhuang@hqu.edu.cn.

**基金项目:** 福建省科技计划项目(2005H035)

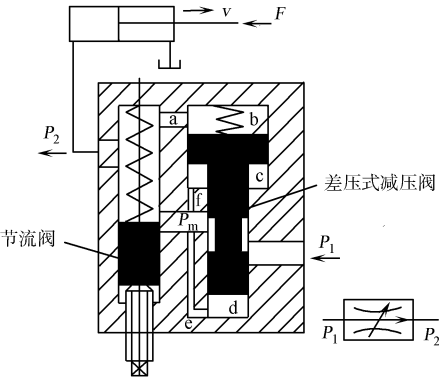


图 1 调速阀结构图

Fig. 1 Configuration of speed control valve

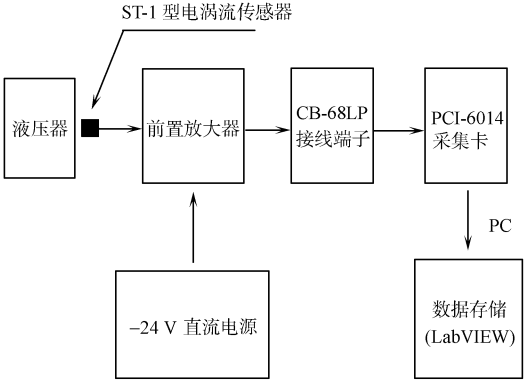
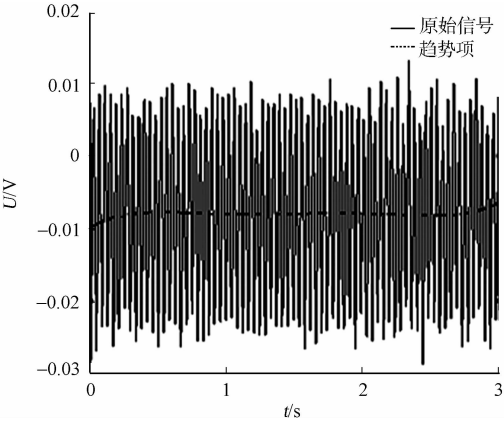
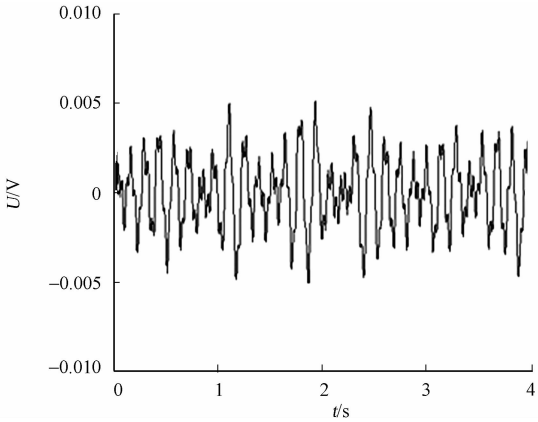


图 2 检测系统结构简图

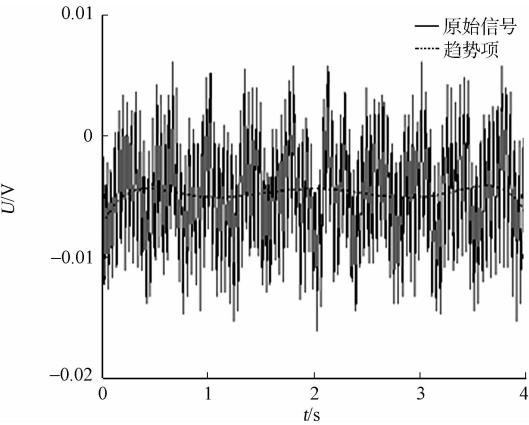
Fig. 2 Schematic of detection system



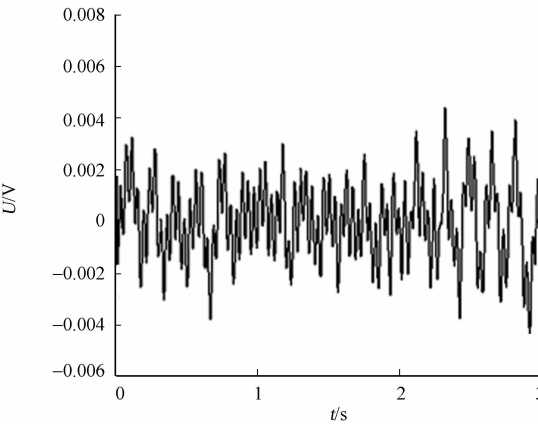
(a) 正常情况滤波处理前



(b) 正常情况滤波处理后



(c) 故障 1 滤波处理前



(d) 故障 1 滤波处理后

图 3 滤波处理前后的电压数据

Fig. 3 Voltage data before and after filtering

## 2 调速阀体振动信号的自回归 AR 模型

### 2.1 特征选择和提取

大量的研究表明,AR 模型的自回归系数对状态变化规律敏感<sup>[4]</sup>. 因此,采用 AR 模型的自回归系数作为特征向量来分析系统的状态变化是十分有效的. 假设采集的调速阀体振动信号是  $x_t$ , 对振动信号建立如下的自回归模型  $AR(p)$ . 即

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + a_t. \quad (1)$$

式(1)中:  $\varphi_i (i=1, 2, \cdots, p)$ ,  $p$  分别是自回归模型  $AR(p)$  的自回归系数和模型阶数;  $a_t$  为模型的残差, 属

于均值为零、方差为  $S^2$  的白噪声系列. 由于自回归系数  $\varphi_i(i=1,2,\cdots,p)$  反映了液压调速阀振动系统的固有特性,模型的残差方差  $S^2$  与液压调速阀振动系统的输出特性密切相关. 因此,可以采用自回归系数  $\varphi_i(i=1,2,\cdots,p)$ 和残差方差  $S^2$  组成的特征向量  $\boldsymbol{\varphi}=[\varphi_1,\varphi_2,\cdots,\varphi_p,S^2]$ 输入到支持向量机来识别液压调速阀的状态.

2.2 自回归 AR 模型的建立

对每种工况下采集的振动信号抽取其中的 1 024 个数据,取 40 组数据,建立 AR 模型. 采用最终预测误差(FPE)准则确定模型的阶数  $P$ ,由最小二乘法估计自回归系数  $\varphi_i(i=1,2,\cdots,p)$ 和模型的残差方差  $S^2$ . 由于采用最终预测误差(FPE)准则<sup>[5]</sup>确定的模型阶数不相等,可以通过对模型阶数补零的方法加以解决.

为了简单起见,取前 6 个 AR 模型自回归系数和残差方差作为特征参数,输入到支持向量机中进行工况的模式识别. 部分样本的特征向量,如表 1 所示.

表 1 部分样本的特征向量  
Tab.1 The feature vectors of part samples

故障类型	$\varphi_1$	$\varphi_2$	$\varphi_3$	$\varphi_4$	$\varphi_5$	$\varphi_6$	$S^2$
正常	0.063 1	−0.238 0	0.090 1	0.168 4	−0.212 3	−0.270 4	770.93
	0.043 5	−0.145 4	0.031 1	0.190 4	−0.274 9	−0.364 6	726.20
故障 1	0.345 0	−0.245 2	0.098 3	0.388 2	0.116 8	−0.212 8	655.17
	0.271 1	−0.242 6	0.092 6	0.332 4	0.078 1	−0.246 8	597.56
故障 2	0.221 5	−0.164 4	0.066 4	0.182 2	0.100 6	−0.221 1	594.86
	0.165 8	−0.111 2	0.062 9	0.095 3	0.083 1	−0.238 3	600.48
故障 3	−0.292 1	−0.003 2	0.069 4	−0.042 7	−0.084 6	−0.150 6	576.93
	−0.221 0	−0.060 3	0.031 6	−0.029 9	−0.116 9	−0.114 0	698.92

3 支持向量机的调速阀故障识别

3.1 故障分类

支持向量机(SVM)是在统计学理论上发展出来的一种新的学习机器<sup>[6-7]</sup>,对于小样本情况的学习机器具有无可比拟的优越性,而对于机械设备系统,尤其是大型机械设备,故障样本的获得往往意味着巨大的经济损失. 因此,对于机械故障类别的识别,支持向量机的优势是明显的.

当数据为两类可分时,设给定训练数据为

$$(x_1,y_1),\cdots,(x_l,y_l), \quad \boldsymbol{x} \in \mathbf{R}^n, \quad y \in \{+1,-1\}.$$

其中: $l$  为训练样本的总数; $n$  为模式空间的维数; $y$  为类别标示. 支持向量机通过某种事先选择的非线性映射,将输入向量  $\boldsymbol{x}$  映射到一个高维特征空间  $Z$ . 在这个空间中构造最优分类超平面,使得训练数据被超平面正确地分开,并且离超平面最近的向量与超平面之间的距离是最大的. 支持向量机通过具有特殊性质的核函数  $K(x_i,x_j)$ 巧妙地避免了直接在高维空间中处理问题.

对于多类问题,采用将多类问题转化为多个两类问题的方法——两两分类法<sup>[8]</sup>. 假设有样本

$$(x_1,y_1),\cdots,(x_l,y_l), \quad \boldsymbol{x} \in \mathbf{R}^n, \quad y \in \{1,\cdots,k\}$$

表示样本的标签,这是一个样本分为  $k$  类的问题.

两两分类的方法将其转化为 $\frac{k(k-1)}{2}$ 个两类问题. 对于第  $i$  类和第  $j$  类的训练数据,需要解决以下两类问题. 即

$$\min_{\boldsymbol{w}^{i,j},b^{i,j},\xi^{i,j}} \frac{1}{2}(\boldsymbol{w}^{i,j})^T \boldsymbol{w}^{i,j} + C(\sum_{t=1}^s \xi^{i,j}), \tag{2}$$

$$[(\boldsymbol{w}^{i,j}) \cdot \boldsymbol{\varphi}(x_i) + b^{i,j}] \geqslant 1 - \xi^{i,j}, \quad y_i = i, \tag{3}$$

$$[(\boldsymbol{w}^{i,j}) \cdot \boldsymbol{\varphi}(x_i) + b^{i,j}] \leqslant -1 - \xi^{i,j}, \quad y_i = j, \quad \xi^{i,j} \geqslant 0. \tag{4}$$

式(2),(3),(4)中: $\boldsymbol{w}^{i,j}$ 为分类平面的法向量; $\boldsymbol{\varphi}(x)$ 为事先选择的非线性映射; $b^{i,j}$ 为超平面的位置; $\xi^{i,j}$ 为

松弛变量; $s$  为第  $i$  类和第  $j$  类个体的总和; $C$  为常数,表示在分类间隔和错误间的折中. 由式(2),(3), (4)可以得到 $\frac{k(k-1)}{2}$ 个决策函数.

当决定样本  $x$  的类别时,采取投票策略,即当决策函数  $f^{i,j}=(\mathbf{w}^{i,j}) \cdot \varphi(x_i)+b^{i,j} \geqslant 0$  时,认为  $x$  属于  $i$  类,则投  $i$  类一票;否则,投  $j$  类一票. 依次类推,经过所有 $\frac{k(k-1)}{2}$ 个决策函数判别后,最终哪一类得票最多,就认为  $x$  属于哪一类. 当得票相同时,简单地将  $x$  归于类标小的那一类.

### 3.2 分类测试

- 根据算法原理,利用 Matlab 编写程序对数据样本进行学习和预测,具体有如下 5 个步骤.
- (1) 准备样本集. 把求出的 40 个自由回归系数  $\varphi_i(i=1,2,\cdots,p)$  和残差  $S^2$  组成特征向量  $\boldsymbol{\varphi}=[\varphi_1, \varphi_2, \cdots, \varphi_p, S^2]$  中的 24(4 种工况各 6 个)个作为支持向量机的训练样本,剩下的 16(4 种工况各 4 个)个作为支持向量机的预测样本. 将正常情况和故障 1、故障 2、故障 3 的训练数据分别标号为 1,2,3,4,即得到所要训练输出的结果.
  - (2) 选择合适的核函数,分别选用径向基函数(RBF)和多项式形式核函数,并对它们在不同 SVM 参数情况下的分类效果作了比较.
  - (3) 采用网格搜索法选择最优参数(惩罚因子  $C$  和核参数  $\sigma^2$ ).
  - (4) 利用上一步网格寻优获得的 SVM 参数,对样本集进行训练以获得支持向量机模型.
  - (5) 利用已建立的支持向量机模型对样本进行预测. 按照前述的故障识别步骤,得到的分类测试结果,如表 2 所示. 表 2 中: $\eta_1$  为分类正确率, $\eta_2$  为识别正确率.

表 2 分类测试结果  
Tab. 2 Test results of classification

核函数类型	$C$	$\sigma^2$	$\eta_1/\%$				$\eta_2/\%$
			正常	故障 1	故障 2	故障 3	
RBF 核	1.7	2.5	100	100	100	100	100.0
RBF 核	2.0	2.5	100	75	100	75	87.5
多项式核 ( $d=1$ )	2.0	—	100	100	75	75	87.5
多项式核 ( $d=2$ )	2.0	—	100	75	100	75	87.5

由表 2 的分类测试结果可知,基于 AR 模型和支持向量机的液压调速阀的故障识别效果不仅与核函数的选取有关,而且还与 SVM 参数的选取有关系. 以径向基 RBF 为核函数的分类效果明显优于以多项式形式为核函数的分类效果.

当选取径向基 RBF 为核函数,SVM 参数  $C=1.7, \sigma^2=2.5$  时,正常和故障 1、故障 2、故障 3 的正确分类率都达到 100%. 由此可见,应用 AR 模型及 SVM 来对液压调速阀进行故障诊断取得了良好的分类效果,表明基于 AR 模型和支持向量机液压调速阀的故障识别方法是可行的,有效的.

## 4 结论

- 以液压调速阀的故障诊断为例,提出一种基于时间序列 AR 模型和支持向量机故障识别方法. 实验取得了良好的分类效果,证明了这种方法的有效性和可行性.
- 该方法的识别效果不仅与核函数的选取有关系,而且与 SVM 参数的选取也有关系. 以径向基 RBF 为核函数的分类效果明显优于以多项式形式为核函数的分类效果. 当选取径向基 RBF 为核函数,SVM 参数  $C=1.7, \sigma^2=2.5$  时,分类正确率达到 100%. 当然,有些情况的分类正确率无法达到 100%,究其原因可能有如下 3 个方面:
- (1) 样本数不够充足. 信息量不够,可能使得支持向量机构建的机器模型无法完全描述振动系统;
  - (2) 参数选择不合理. 所利用的网格搜索方法有其固有的不足,势必导致所选的参数不是最好的参数,这有待于人们研究更优秀的寻优工具;

(3) 核函数的选择. 由表 2 的分类测试结果可知, 基于 AR 模型和支持向量机的液压调速阀的故障识别效果与核函数的选取有关系.

AR 模型是一个“信息的凝聚器”, 可将调速阀体振动系统的特性及工作状态都凝聚于其中, 因而可依据它对液压调速阀体的工作状态进行识别.

实验结果表明, 建立的支持向量机模型能够正确地对调速阀的 4 种工况进行识别, 尽管作为训练数据的样本不多, 仍然具有很好的分类效果. 这说明了 SVM 故障识别的实用性, 验证了 SVM 在处理小样本学习问题上具有独到的优越性.

#### 参考文献:

- [1] 蔡奇志, 黄宜坚. 三谱切片在调速阀故障诊断中的应用[J]. 华侨大学学报: 自然科学版, 2009, 30(1): 16-21.
- [2] 李良敏, 屈良生. 基于遗传编程和支持向量机的故障诊断模型[J]. 西安交通大学学报, 2004, 38(3): 239-242.
- [3] 童进, 吴昭同, 严拱标. 大型旋转机械升降过程故障诊断 HMM-AR 方法研究[J]. 振动与冲击, 1999, 18(2): 79-80.
- [4] 刘天雄, 郑明刚, 陈兆能, 等. AR 模型和分形几何在设备状态监测中的应用研究[J]. 机械强度, 2001, 23(1): 61-65.
- [5] 杨叔子, 吴雅. 时间序列分析的工程应用: 上册[M]. 武汉: 华中理工大学出版社, 1991.
- [6] VAPNIK V. The natural statistical theory[M]. New York: Springer Verlag, 1995.
- [7] NELLO C, JOHN S T. An introduction to support vector machines and other kernel based learning methods [M]. Cambridge: Cambridge University Press, 2000.
- [8] HSU C W, LIN C J. A comparison of methods for multiclass support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(3): 415-425.

## Fault Recognition of Speed Control Valve Based on AR Model and Support Vector Machine

ZENG Xiao-jun, HUANG Yi-jian

(College of Mechanical Engineering and Automation, Huaqiao University, Quanzhou 362021, China)

**Abstract:** A fault recognition method based on time series autoregressive (AR) model and support vector machine has been put forward for fault recognition of hydraulic speed control valve. Firstly, the AR model of vibration signal from speed control valve body is established; then the AR coefficients and error variance are regarded as the feature vectors which are used as an input of support vector machine; lastly, the normal state and all kinds of faults are classified by support vector machine. The result and analysis of the experiment indicate that the recognition rate is affected not only by the selection of the kernel function but also by the selection of parameters of support vector machine, the recognition rate is obviously more better when the kernel function is radial basis function (RBF) than that when the kernel function is polynomial kernel function.

**Keywords:** speed control valve; fault recognition; autoregressive model; support vector machine

(责任编辑: 钱筠      英文审校: 崔长彩)