

基于因子分析的模拟电路故障特征提取技术研究

王月海¹, 卢俊¹, 潘国庆², 冯建呈²

(1. 北方工业大学 信息工程学院, 北京 100144; 2. 北京航天测控技术开发公司, 北京 100037)

摘要: 针对 LS-SVM 算法中小波提取特征存在小波基函数选择和小波分解层次、系数选取的问题, 提出了一种基于因子分析技术的故障特征识别方法; 该方法通过构建采样数据的相关矩阵求出因子载荷和因子得分, 按照累计贡献率自动提取出 1~3 个因子组成特征向量, 从而降低了输入维度, 提高了算法训练诊断效率, 降低了收敛难度; 四运放典型电路的仿真实验结果表明: 文中算法的诊断正确率超过了同类方法, 同时提高了训练时间和诊断效率。

关键词: 因子分析; 故障特征; 因子得分; 特征向量

Research on Feature Extraction for Diagnostics of Analog Circuit Based on Factor Analysis

Wang Yuehai¹, Lu Jun¹, Pan Guoqing², Feng Jiancheng²

(1. College of Information Engineering, North China University of Technology, Beijing 100144, China;
2. Beijing Aerospace Measurement & Control Corp., Beijing 100037, China)

Abstract: This paper presents a fault feature recognition method based on factor analysis techniques for wavelet feature extraction in LS-SVM algorithm existing problem of wavelet bases function selection, wavelet decomposition level and coefficient selection. The method computes factor loadings and factor scores by constructing a correlation matrix of sample data, extract factors 1-3 to compose feature vector automatically according to the cumulative contribution rate, thereby reduce the dimension of the input, improve the efficiency of training and diagnostic algorithm, reduce the convergence difficulty. The simulation results of four op-amp biquad high-pass filter show: The diagnostic accuracy of the algorithm in this paper is beyond similar methods, while increasing the training time and the efficiency of diagnosis.

Keywords: factor analysis; fault feature; factor score; feature vector

0 引言

自 19 世纪 70 年代以来, 模拟电路故障诊断已经成为一个相当活跃的研究领域。而故障特征提取的好坏关系到故障诊断的速度和准确度。尽管目前有许多的特征提取方法已经被研究^[1], 但是由于模拟电路的非线性、连续性、容差等特点使得模拟电路的特征十分复杂, 故障诊断选用的特征包括电压、频率等本原特征, 又包括方差、频谱、能量等高阶统计特征^[2], 还包括小波系数、计算得出的主元等抽象特征。文献 [1] 和 [3] 采用小波变换和主元分析来获取用于训练的最优特征, 这种方式虽然能在一定程度上减少神经网络的输入和提高效率, 但系数选择、小波层次、系数选取没有规律可循, 只能凭经验进行, 诊断效果难以推广^[4]。文献 [3] 对比研究了提取时域特征、频域特征、统计特征和联合特征在 LS-SVM 下作训练诊断的情况, 虽然选取联合特征诊断正确率高, 但是计算比较复杂, 不利于故障检测的自动化。

因子分析可以借助标准化过程自主提取少数的参数, 参数提取方便, 特征量维度进一步缩减, 为模拟电路故障的特征提取提供了一种新的思路。

1 因子分析

因子分析模型是原始变量相关矩阵内部的依赖关系将原始变量经过一系列标准化过程归结为少数综合因子, 得出因子表达式, 求出因子得分值, 然后根据因子得分在因子所构成的空间中把变量进行分类。应用因子分析来提取特征的步骤如下:

1) 获取故障数据并标准化:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{S_j}}, i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (1)$$

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2)$$

$$S_j = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \quad (3)$$

其中: x_{ij} 表示原始故障样本集, \bar{x}_j 表示原始故障样本均值, S_j 表示原始故障样本方差, n 表示样本个数, p 表示每个样本的观测指标数目。

2) 求相关矩阵, 分析样本之间的相关性。

$$\text{相关矩阵: } \mathbf{R} = \frac{1}{n-1} \mathbf{Z}'\mathbf{Z} \quad (4)$$

$$\text{其中 } \mathbf{Z} = [z_1, z_2, \dots, z_n]' \quad (5)$$

3) 求出初始公共因子, 采用主成分法确定因子载荷^[6]。

设有 n 个变量, 由相关阵出发求解主成分, 找出 n 个主成分。将所得的 n 个主成分按由大到小的顺序排列, 记为 Y_1, Y_2, \dots, Y_n , 令 $F_i = Y_i / \sqrt{\lambda_i}$, $a_{ij} = \sqrt{\lambda_i} \gamma_{ji}$, 则有如下关系式:

收稿日期: 2014-04-04; 修回日期: 2014-05-12。

作者简介: 王月海(1975-), 男, 山东莒南人, 博士, 副教授, 硕士研究生导师, 计算机学会会员, 主要从事计算智能和智能机器人方向的研究。

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + \varepsilon_1 \\ X_2 = a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + \varepsilon_2 \\ \vdots \\ X_n = a_{n1}F_1 + a_{n2}F_2 + \dots + a_{nm}F_m + \varepsilon_n \end{cases} \quad (6)$$

这样就得到载荷矩阵 A 和一组初始公因子 F_i 。

$$A = (\sqrt{\lambda_1}\gamma_1, \sqrt{\lambda_2}\gamma_2, \dots, \sqrt{\lambda_m}\gamma_m) \quad (7)$$

式 (7) 中, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ 为样本相关阵 R 的特征根, $\gamma_1, \gamma_2, \dots, \gamma_p$ 为对应的标准正交化特征向量, 且 $m < p$ 。

4) 因子旋转:

对得到的初始公共因子进行线性组合, 找到实际意义更明显的公共因子。

$$\begin{cases} F'_1 = d_{11}F_1 + d_{12}F_2 + \dots + d_{1m}F_m \\ F'_2 = d_{21}F_1 + d_{22}F_2 + \dots + d_{2m}F_m \\ \vdots \\ F'_m = d_{m1}F_1 + d_{m2}F_2 + \dots + d_{m2}F_m \end{cases} \quad (8)$$

式中, F_1, F_2, \dots, F_m 是初始公共因子, F'_1, F'_2, \dots, F'_m 是新的公共因子。

5) 因子得分:

采用采用回归的思想求出线性组合系数的估计值, 建立如下以公共因子为因变量、原始变量为自变量的回归方程:

$$F_j = \beta_{j1}X_1 + \beta_{j2}X_2 + \dots + \beta_{jp}X_p \quad j = 1, 2, \dots, m \quad (9)$$

在最小二乘意义下, 可以得到 F 的估计值:

$$F = A'R^{-1}X \quad (10)$$

式 (10) 中, A 为因子载荷矩阵; R 为原始变量的相关阵; X 为原始变量向量。这样在得到一组样本值后, 就可以代入式 (10) 求出公共因子的估计得分, 从而用少数公共因子去描述原始变量, 当因子数 m 较少时, 还可以用因子法直观地描述样本的分布情况。

2 多分类支持向量机

标准支持向量机最初是用来解决二分类问题的, 多分类问题常常被转化为二分类问题^[7]。本文模拟电路故障诊断, 采用一对多算法^[8], 将一个 C 分类问题转化为 C 个二分类问题, 其中第 i 个分类器是通过将属于第 i 类的样本点视为正类, 其余所有的样本点视为负类训练而成的^[9]。设第 i 类和剩余类之间的最优超平面为:

$$D_i(x) = \omega_i^T \Phi(x) + b_i = 0 \quad (11)$$

当采用离散判别函数时, 其决策规则为: 把样本 x 归为 $D_i(x) > 0$ 的第 i 类。当有多个 $D_i(x) > 0$ 时, 会存在不可分区域。

当采用连续判别函数时, 其决策规则为: 把样本 x 归为 $\arg \max_{i=1, \dots, C} D_i(x)$ 的第 i 类。当有多个 $D_i(x)$ 相等时, 会存在不可分区域。

为解决离散判别的可能存在的不可分区域问题^[10], 提出了基于模糊决策函数的支持向量机。对于第 i 类, 在与最优超平面 $D_j(x) = 0$ 垂直的方向上定义模糊隶属度函数 $m_{ij}(x)$ 如下:

$$\text{当 } i = j \text{ 时 } m_{ii}(x) = \begin{cases} 1, & D_i(x) \geq 1 \\ D_i(x), & \text{其他} \end{cases} \quad (12)$$

$$\text{当 } i \neq j \text{ 时 } m_{ij}(x) = \begin{cases} 1, & D_j(x) \leq -1 \\ -D_j(x), & \text{其他} \end{cases} \quad (13)$$

样本 x 属于第 i 类的隶属度函数定义为

$$m_i(x) = \min_{j=1, \dots, C} m_{ij}(x) \quad (14)$$

令:

$$k = \arg \max_{i=1, \dots, C} m_i(x),$$

则样本 x 被分为第 k 类。

该方法结构简单、全局最优、泛化能力强, 等价于采用连续决策函数的一对多方法。

利用因子分析提取特征使用多分类 LS-SVM 进行模拟电路测试的仿真流程如图 1 所示。

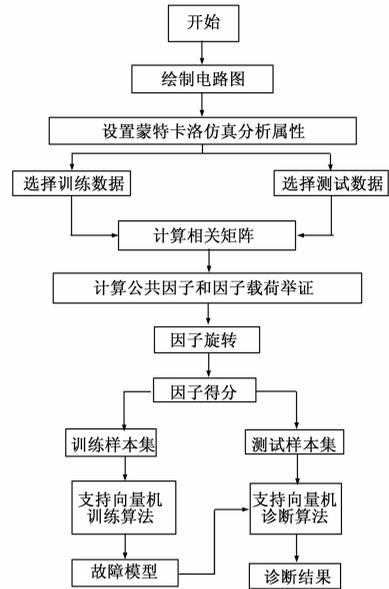


图 1 基于因子分析的 LS-SVM 的训练诊断过程

3 仿真

3.1 仿真电路和故障设置

为方便比较仿真电路选用图 3 所示四运放高通滤波器, 电路的故障类别、元器件的正常值和故障值如表 1 所示^[2]。13 种故障包括 NF, $C1 \uparrow$, $C1 \downarrow$, $R4 \uparrow$, $R4 \downarrow$, $C2 \uparrow$, $C2 \downarrow$, $R3 \uparrow$, $R3 \downarrow$, $R2 \uparrow$, $R2 \downarrow$, $R1 \uparrow$, $R1 \downarrow$, 其中 \uparrow 和 \downarrow 分别表示高于和低于正常值。为了获取不同故障类别的训练数据和测试数据, 需要人为的对电路中的元器件注入故障, 同时让其电阻和电容的容差分别保持在 5% 和 10% 的范围内。

表 1 电路故障类别及其元器件的正常值与故障值

Fault ID	Fault Class	Nominal	Faulty-value	Value
F0	NF			
F1	$C1 \uparrow$	5 nf		10 nf
F2	$C1 \downarrow$	5 nf		2.5 nf
F3	$R4 \uparrow$	1.6 kΩ		2.5 kΩ
F4	$R4 \downarrow$	1.6 kΩ		0.5 kΩ
F5	$C2 \uparrow$	5 nf		15 nf
F6	$C2 \downarrow$	5 nf		1.5 nf
F7	$R3 \uparrow$	6.2 kΩ		12 kΩ
F8	$R3 \downarrow$	6.2 kΩ		3.3 kΩ
F9	$R2 \uparrow$	6.2 kΩ		12 kΩ
F10	$R2 \downarrow$	6.2 kΩ		2 kΩ
F11	$R1 \uparrow$	6.2 kΩ		15 kΩ
F12	$R1 \downarrow$	6.2 kΩ		3 kΩ

3.2 特征提取

假定当电路的某个元器件高于或低于先前描述的正常值

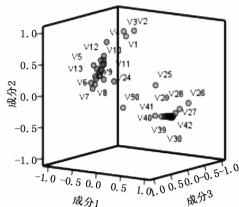


图 2 故障 F1 的旋转空间中公共因子 3 维图

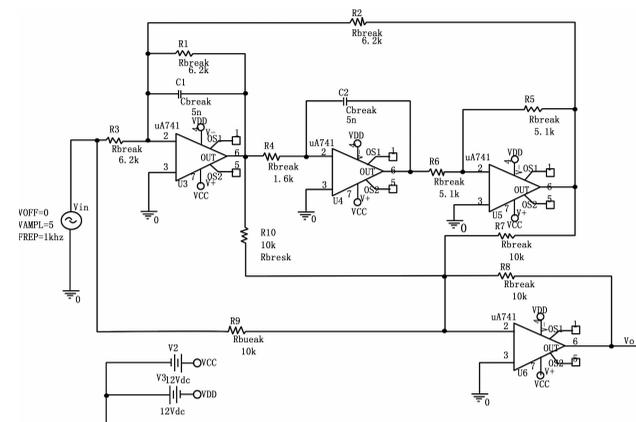


图 3 四运放高通滤波器

(其它元器件保持在各自的容差范围内) 时, 获取一条故障时域曲线。对仿真电路的每种故障类别使用 p-spice 仿真软件进行 Monte Carlo 分析 60 次; 其中 20 次做训练样本, 40 次做测试样本。总共有 260 个训练样本, 520 个测试样本。对应每次 M-C 分析获得输出点在采样时间内的电压值序列。然后将所获得的原始数据送入 SPSS 软件中进行处理, 求出公共因子和相应的因子得分。图 2 展示了故障 F1 做因子旋转后新的公共因子 3 维图, 其中成分 1、成分 2、成分 3 分别表示新的第一、第二、第三个公共因子。最后将公共因子构造成 SVM 训练样本集和测试样本集后, 使用 LS-SVM 进行故障的训练和诊断。

3.3 仿真结果

采用先前介绍的特征提取方法构造训练样本集和测试样本集。由于有 13 种故障, 所以总共有 $13 \times 40 = 520$ 个测试样本。对于仿真测试电路的诊断结果如表 2 所示: 故障 C1 ↑ 有 3 个

表 2 13 种故障的诊断结果

	NF	C1 ↑	C1 ↓	R4 ↑	R4 ↓	C2 ↑	C2 ↓	R3 ↑	R3 ↓	R2 ↑	R2 ↓	R1 ↑	R1 ↓
NF	40												
C1 ↑		37										3	
C1 ↓			40										
R4 ↑				40									
R4 ↓					40								
C2 ↑						38		2					
C2 ↓							40						
R3 ↑								40					
R3 ↓									40				
R2 ↑										40			
R2 ↓											40		
R1 ↑												40	
R1 ↓													40

测试样本被误诊为故障 R1 ↑, 故障 C2 ↑ 有 2 个测试样本被误诊为故障 R3 ↑, 故障 R2 ↓ 有 3 个测试样本被误诊为故障 R1

↓, 其余测试样本均诊断正确。从表 2 中可以得出结论: 对于仿真电路采用因子分析提取特征公共因子来构造训练样本集, 对于具体某一种故障最佳的诊断正确率高达 100%, 最坏的情况也有 90% 的正确率。

3.4 结果分析

表 3 是本文方法与文献 [2] 中的实验结果进行比较的结果, 可以看出: 1) 因子分析提取故障特征时比先小波预处理再用主元提取特征的诊断正确率要高; 2) 因子分析方法明显要比先小波预处理再用主元提取特征的步骤少, 处理过程更简洁。

表 3 因子分析与小波+主元的实验结果比较

方法	训练样本个数	测试样本个数	诊断正确率
因子分析	260	520	98.46%
小波+主元	260	520	94.8%

4 结束语

提出了一种基于因子分析和 LS-SVM 相结合的模拟电路故障诊断方法, 首先用 p-spice 仿真软件采集待测电路的不同类别的故障信号, 然后用 SPSS 软件对信号进行处理, 抽象出代表每种故障信号的特征, 构造训练样本集和测试样本集, 再送入 LS-SVM 训练与诊断程序进行分类识别, 判断出故障。仿真实验结果表明: 因子分析技术能够有效的提取故障样本的重要特征, 相较 [2] 在准确率保持相同数量级, 但显著降低特征向量的维度, 从而提高诊断算法训练收敛性和训练时间, 进而证实了所提出方法的正确性和有效性。

参考文献:

[1] Aminian M, Aminian F. Neural-network based analog-circuit fault diagnosis using neural transform as preprocessor [J]. IEEE Trans. Circuits Syst. II, 2000, 47 (2): 151-156.

[2] Long B, Tian S, Wang H J. Diagnostics of filtered analog circuits with tolerance based on LS-SVM using frequency features [J]. J Electron Test, 2012, 28 (3): 291-300.

[3] Aminian F, Aminian M, Collins H W. Analog fault diagnosis of actual circuits using neural networks [J]. IEEE Trans. Instrum. Meas., 2002, 51 (3): 544-550.

[4] 王月海, 程 冉, 蒋爱民, 等. 模拟电路故障特征提取的小波基选取方法研究 [J]. 计算机测量与控制, 2011, 19 (6): 1329-1330.

[5] Long B, Tian S L, Miao Q, et al. Research on features for diagnostics of filtered analog circuits based on LS-SVM [J]. IEEE Autotestcon. Baltimore, MD, 2011, 360-366.

[6] 张洪波, 何怡刚, 等. 主成分分析法与概率神经网络在模拟电路故障诊断中的应用 [J]. 计算机测量与控制, 2008, 16 (12): 1789-1791.

[7] 杨晓伟, 郝志峰. 支持向量机的算法设计与分析 [M]. 北京: 科学出版社, 2013.

[8] Cui J, Wang Y R. A novel approach of analog circuit fault diagnosis using support vector machines classifier [J]. Measurement, 2011, 44: 281-289.

[9] Nello Cristianini, John Shawe-Taylor. 支持向量机导论 [M]. 北京: 机械工业出版社, 2005.

[10] Mangasarian O, Musicant D. Lagrangian support vector machines [J]. Journal of Machine Learning Research, 2001, 1: 161-177.