

文章编号:1001-9081(2006)08-1977-03

## 基于聚类预处理和支持向量机的模拟电路故障诊断技术

崔 江<sup>1</sup>,王友仁<sup>1</sup>

(南京航空航天大学 自动化学院, 江苏 南京 210016)

(cuijiang@nuaa.edu.cn)

**摘 要:**支持向量机(SVM)最初源于两种分类问题,用于存在较多故障模式的模拟电路诊断问题,易造成识别重叠区域。为此提出了利用动态聚类算法作为 SVM 预分类器的故障诊断方法,首先采用模糊 C-均值(FCM)算法对训练样本进行聚类,然后分别对两大类进行内部的子聚类,每一次的聚类都产生两种模式并对各个模式内的故障模式样本训练产生对应的 SVM 网络,最后采用二叉树形式把所有的模式分开。实验结果表明,采用该方法对测试样本的诊断正确率可以达到 99% 以上。

**关键词:**支持向量机;模糊 C 均值;模拟电路;故障诊断

**中图分类号:**TP391.5 **文献标识码:**A

## Analog circuit faults diagnosis based on clustering preprocess and SVM

CUI Jiang, WANG You-ren

(Nanjing University of Aeronautics and Astronautics, College of Automation and Engineering, Jiangsu Nanjing 210016, China)

**Abstract:** Support Vector Machines (SVM) is invented from binary classification problem and it is difficult to be used directly in analog circuit faults diagnosis because of multi-fault modes and overlapped recognition areas. A diagnosis method was proposed by using dynamic clustering method as a preprocessor of SVM. Firstly, Fuzzy C-Means (FCM) algorithm was used to generate father and mother clustering gather, then their children were generated by binary tree structure, and every clustering would train a SVM network to be stored. The simulation experimental results show that accuracy of this diagnosis can be higher than 99%.

**Key words:** Support Vector Machines(SVM); Fuzzy C-Means(FCM); analog circuits; fault diagnosis

### 0 引言

模拟电路的故障诊断方法大体可分为两大类:基于模型的方法和基于知识的方法。例如:参数辨识方法<sup>[1]</sup>、网络撕裂方法、故障字典方法<sup>[2]</sup>、神经网络方法<sup>[3]</sup>、基于事件推理的方法(Case Based Reasoning, CBR)等。其中,基于模型的方法一般用于线性电路的故障诊断和分析,难以处理非线性容差电路的故障问题;故障字典法最具有工程应用价值,但该方法的基本原理大部分基于屏蔽支路原理(迭代算法耗时长且难以处理容差问题)或者诊断的是硬故障模式;人工神经网络(Artificial Neural Network, ANN)识别方法也是故障诊断领域研究的热点之一,但神经网络具有统计特性,必须在一定量的样本训练下才能获取期望的网络输出和泛化能力,且当故障模式较多、故障特征样本的数据维数较高时不易收敛;CBR 方法不需建立复杂模型,随着诊断过程的不断积累其有效性越来越高,但是该方法需要在线更新且需要考虑容差问题,适合长期使用设备的故障诊断。

影响模拟电路故障诊断的因素较多,主要有:模拟元器件的容差、电路呈现出的非线性特性、电路模型难以建立等。为了区别较多的故障模式并达到较高的故障诊断率,采用特征提取方法得到的故障特征往往具有较高的数据维,采用神经网络等常规方法进行学习(训练)显得较为困难。作为一种机器学习方法和模式识别方法,SVM 在样本的非线性、稀疏性和

高维模式识别方面具有独特优势,并已经在故障诊断领域得到实际应用<sup>[4-6]</sup>。模拟容差电路的故障特征样本往往具有数量少、数据维数高等特点,应用 SVM 进行样本的训练可以很好地解决神经网络训练速度慢、容易陷入局部极值等缺点。

SVM 算法最初源于二类问题,但模拟电路故障诊断往往需要诊断多个故障模式,因此需要 SVM 多分类器。而 SVM 多分类器一般有两种类型<sup>[5]</sup>:组合多个二值 SVM 形成多类 SVM 分类器;改造二值 SVM 内部结构和算法以形成 SVM 多分类器。本文采用第一种方法构造二叉树 SVM(或层次 SVM)以实现模拟电路故障的诊断。

如果随机选取故障模式中的两种进行组合,那么产生的组合数目很大且不同组合对诊断的效果影响很大。根据 SVM 的原理,不同类别模式的分类间隔最大就可以使得真实风险最小,同时使得支持向量个数较少,那么得到的 SVM 网络泛化(推理)能力较强。基于该思想,本文利用 FCM 算法预分类(分为两类)全部训练样本,然后再设计对应的二叉树多类 SVM 分类器。由于聚类算法可以粗略、较好地预分类(分为两类)全局数据样本,从而优化了二叉树 SVM 分类器的各层结构,且使得各层两类数据之间的分类间隔最大,因此得到的 SVM 网络具有支持向量个数少、泛化能力强等特点。实验证明,对模拟电路软故障的诊断能力可达 99% 以上,优于同一条件下小波神经网络、聚类算法和故障字典方法的诊断精度,并和 Prony 分析方法的诊断精度较为接近。

收稿日期:2006-02-;修订日期:2006-04- **基金项目:**国家自然科学基金资助项目(60374008,60501022);航空科学基金资助项目(04152068);南京航空航天大学青年基金资助项目(Y0521033)

**作者简介:**崔江(1977-),男,山东潍坊人,讲师,硕士,主要研究方向:模拟/混合电路故障诊断、计算机和嵌入式系统测控;王友仁(1963-),男,江苏南京人,教授,博士,主要研究方向:电路测试与故障诊断、仿生硬件理论与技术、机载测控系统自主修复。

## 1 基于 FCM 预处理的二叉树 SVM 多分类器

基于 FCM 预处理的二叉树 SVM 设计方法和流程如图 1 所示。

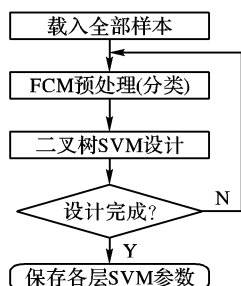


图1 基于 FCM 预分类的二叉树 SVM 设计

首先把全部训练样本采用 FCM 进行第一层聚类(聚为两类 A 和 B),然后设计二值 SVM 进行两类样本的训练和分类;接着对 A 类或 B 类分别应用相同的方法进行深层次的聚类,直到把全部的模拟电路故障模式分开为止,这样就得到一个类似倒立二叉树结构的支持向量机分类器。整个分类器结构设计完成之后,需要存储对应各层的 SVM 训练参数。

### 1.1 FCM 预处理

FCM 算法是一种常用的动态聚类算法,其目的是把  $n$  个样本动态划分至  $C$  个类别中,使得各个样本与其所在类之间的误差平方和最小<sup>[4]</sup>。

首先选定初始聚类中心和聚类数目,然后分别迭代计算隶属度和对应的新聚类中心,直到两个参量稳定为止。采用标准 FCM 算法进行预分类的过程和编程步骤如下:

- 1) 设定  $C = 2$ , 初始化  $m_j$  为随机变量;设一共有  $K$  种故障模式,且每种故障模式的训练样本为  $T$  个,那么总样本个数  $n = K \times T$ ;
- 2) 对训练数据按照迭代公式完成两类聚类,分别称为父类(以“+1”标签表示)和母类(以“-1”标签表示);
- 3) 计算各故障模式下样本对父类和母类的隶属度及聚类中心距离,统计隶属度较大或者距离较远的样本数目(可以设定一定的隶属度阈值进行判断,例如隶属度大于等于 0.97 的样本即可以认为满足判断条件),如果该数目大部分属于父类(或者为母类),那么设定该故障模式的归属为父类(或者母类);如果不满足判断条件,回至步骤 2) 重新聚类或者重新改变判断阈值;直到所有的故障模式都得到明确归属;
- 4) 对 3) 步骤中得到的两类数据进行 SVM 训练,得到对应的 SVM 训练参数并保存;
- 5) 对 3) 步骤中得到的两类数据重新进行子类的划分,分别跳至步骤 2) 进行聚类,直到完成步骤 2) ~ 5)。聚类过程中,设定的判断阈值可以随着子类分类的深入而逐步调整;
- 6) 完成所有模式的分类和 SVM 的训练;最终得到  $K - 1$  个 SVM。

### 1.2 二值 SVM 分类器

SVM 的基本思想可以概述为:通过一定的非线性核函数映射把低维空间内非线性可分的问题映射到高维空间内,使其线性可分,经过变换以后的分类函数依靠支持向量形成分类<sup>[6]</sup>。

假设待分类样本为集合  $\{(x_i, y_i)\} (i = 1, 2, \dots, n, \text{且 } x_i \in R^d, y_i \in \{+1, -1\})$ , 在  $d$  维线性空间内的分类面方程为:

$$W \cdot X + b = 0 \quad (1)$$

其中,  $W$  为超平面的法向量,  $\cdot$  是向量的点积运算,  $b$  为偏

差。该方程对两类样本的分类间隔为  $\frac{2}{\|W\|}$ , 上式为最优非线性可分分类面的条件是满足下列约束不等式:

$$y_i[(W \cdot x_i) + b] - 1 + \xi_i \geq 0 \quad (2)$$

其中,  $\xi_i \geq 0$  为松弛项。在此约束条件下,最优分类面问题等价于求取下列二次规划函数的最小值:

$$\Phi(W, \xi) = \frac{(W \cdot W)}{2} + C \sum_{i=1}^n \xi_i \quad (3)$$

其中,  $C$  为惩罚因子。构造下列 Lagrange 函数(4), 通过对  $W$  和  $b$  求偏导令其结果为 0 即可得到较为简单的对偶问题。

$$L(W, b, \alpha) = \Phi(W, \xi) - \sum_{i=1}^n \alpha_i (y_i (W \cdot X + b) - 1 + \xi_i) \quad (4)$$

其中,  $0 \leq \alpha_i \leq C$  为拉式变换系数。根据最优化理论中的 KKT (Karush-Kuhn-Tucker) 条件最后得到的最优分类函数为:

$$f(x) = W^* \cdot X + b^* = \sum_{i=1}^n \alpha_i^* y_i (x_i \cdot X) + b^* \quad (5)$$

通过把点积运算换成符合 Mercer 条件<sup>[4]</sup> 的核函数运算, 而其他条件不变, 就得到了 SVM 的判别泛函(6) 和决策函数公式(7)。

$$w(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (6)$$

$$f(x) = W^* \cdot X + b^* = \sum_{i=1}^n \alpha_i^* y_i K(x_i, X) + b^* \quad (7)$$

## 2 电路故障诊断和分析

### 2.1 故障模式设定和注入

仿真电路为一个滤波器电路,如图 2 所示<sup>[7]</sup>。为方便分析,假设元器件产生单一软故障(见表 1), 这样一共有故障类别 17 种(包括电路无故障模式 F0), 其中,“↓”表示故障值波动低于正常容差范围,“↑”表示故障值波动高于标称值范围。其中电阻元件的正常容差为 5%, 电容元件的正常容差为 10%。单一软故障模式可以分为两种,即对应元件参数值变大的故障和参数值变小的故障,两种模式对应的故障容差区间分别设置为[正常容差, 50%]和[-50%, 正常容差]。采用正态故障注入方式设定电路故障模式,即在整个故障容差范围内计算其故障中间值和左右偏离容差,然后以中间值作为产生的故障标称值,左右偏离容差作为仿真的故障容差。

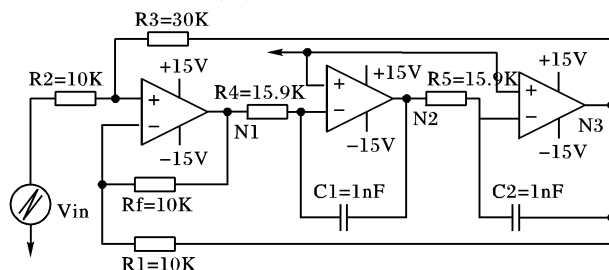


图2 滤波器电路

### 2.2 故障特征集提取

本文选取测试信号  $V_{in}$  (如图 2 所示) 为宽度 1us、幅度 +5V 的单脉冲电压信号。该激励信号含有丰富的频域信息, 在不同故障模式下, 电路的频域和时域输出都有所差别, 因此利用 Haar 小波包作为特征提取器采集电路故障特征信息。

本文选取三个运算放大器的输出节点 N1、N2 和 N3 进行电压信号的采集和小波包分解变换, 每个节点提取 5 个特征

变换系数,一共得到 15 个特征系数作为故障模式特征集,形成各个故障模式的特征训练集(每种故障模式 50 个样本)和测试集合(每种故障模式 100 个样本)。

表 1 故障模式设定表

故障元件	故障代码	故障类别	故障元件	故障代码	故障类别
R1	F1	↓	R5	F9	↓
R1	F2	↑	R5	F10	↑
R2	F3	↓	Rf	F11	↓
R2	F4	↑	Rf	F12	↑
R3	F5	↓	C1	F13	↓
R3	F6	↑	C1	F14	↑
R4	F7	↓	C2	F15	↓
R4	F8	↑	C2	F16	↑

### 2.3 故障诊断结果和分析

按照 FCM 的预分类步骤,最后得到全部数据训练集合的一个二叉树结构,树的每个节点都采用一个 SVM 进行训练,这样一共得到的 SVM 个数为 16 个。

SVM 的核函数选择较为困难,一般通过实验的方法选取和确定。本文分别选取多项式函数、径向基函数 RBF 和 Sigmoid 函数作为核函数进行验证,结果发现 RBF 作为核函数的分类精度较高,因此最终确定核函数为 RBF 函数,其中核宽为 8。训练结束以后分别存储对应的参数(例如:支持向量位置,拉式系数和对应偏差)。

利用训练好的 SVM 对 17 种故障模式下的  $17 \times 100$  个测试样本进行测试,并和故障字典法(存储对应的小波系数平均特征)、小波神经网络方法(采用 BP ANN 对小波系数进行训练,由于网络很难收敛,故设计了多个子网络分别诊断)、模糊 C 均值方法(取聚类中心矢量和对应的隶属度函数,采用 M 范数作为诊断距离)以及随机选取的一组二叉树 SVM 进行了比较,其结果如表 2 所示。结果表明,采用本文方法具有较高的诊断正确率,并与参考文献[7]中的 Prony 分析方法较为接近,但采用 Prony 作为特征提取器涉及到了许多复杂运算,例如:Z 域变换求取指数系数、病态 Toeplitz 和 Vardermonte 矩阵的求解等,计算量较大,且对噪声较为敏感,最后的模式识别采用 BP 神经网络实现,在某些情况下会产生较大失真;而小波包分解是一种常用的特征提取方法,虽然

在运算的时候也涉及到了矩阵运算,但是运算量和 Prony 方法相比差别较大,且小波包分解技术具有一定的消噪功能,最终采用 SVM 作为模式识别方法不存在局部极小值问题。

表 2 各种诊断方法的比较

诊断方式	诊断正确率(%)
故障字典方法	81.50
改进的 BP ANN	91.00
动态聚类(FCM)	93.88
随机二叉树 SVM	97.70
本文方法	99.35
Prony 方法平均诊断正确率	99.43

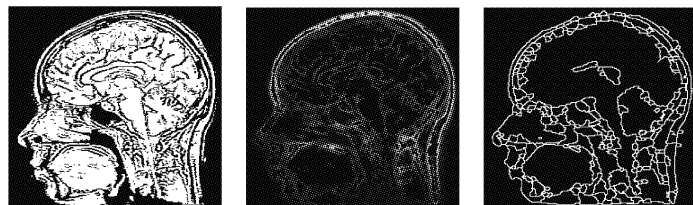
### 3 结语

支持向量机来源于两类算法,当用于模拟电路故障诊断时,必须设计多个故障类别的分类算法。采用 FCM 算法预分类训练数据,可以较好的控制数据的整体分类,然后采用二叉树支持向量机训练聚类故障特征集数据,最后利用本文算法和故障字典法、小波神经网络法、聚类算法等的性能在同一条件下进行了分别测试,实验数据表明,采用本文算法可以较好的诊断和分类模拟电路单一故障。

#### 参考文献:

- [1] SAVIR J, GUO Z. On the Detectability of Parametric Faults in Analog Circuits[ A]. Proceedings of the 2002 IEEE International Conference on Computer Design[ C]. 2002.
- [2] 陈圣俭, 洪炳容, 王月芳, 等. 可诊断容差模拟电路软故障的新故障字典法[ J]. 电子学报, 2002, 28(2).
- [3] LIU Q, YU X, FENG Q. Fault Diagnosis Using Wavelet Neural Networks[ A]. Neural Processing[ C]. 2003. 115 - 123.
- [4] 边肇祺, 张学工. 模式识别[ M]. 北京: 清华大学出版社, 2000.
- [5] 黄席樾, 张著洪, 何传江, 等. 现代智能算法两类及应用[ M]. 北京: 科学出版社, 2005.
- [6] VAPNIK V. The Nature of Statistical Learning Theory[ M]. N. Y.: Springer, 1995.
- [7] EL-YAZEED MFA, MOHSEN AAK. A Preprocessor for Analog Circuit Fault Diagnosis Based on Prony's Method[ J]. AEU international journal of electronics and communications, 2003, 57(1): 16 - 22.

(上接第 1976 页)



(a) 基于边缘的二值图像 (b) 改进多尺度梯度图像 (c) 分水岭处理后图像

图 4 改进多尺度分水岭算法实验结果

### 5 结语

本文提出的基于边缘滤波的多尺度分水岭算法,结合了多尺度滤波和边缘滤波的优点,在强化边缘点的同时消除了局部极小值效应,对于灰度不均匀的磁共振(MR)医学图像有较好的效果。如何适应到其他环境中还需做进一步研究。

#### 参考文献:

- [1] VINCENT L, SOILLE P. Watersheds in digital space: An efficient algorithms based on immersions imulation[ J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(6): 583 - 598.
- [2] 王蜀, 李永宁, 陈楷民, 等. 基于数学形态学的医学图像分割[ J]. 计算机应用, 2005, 25(10): 2381 - 2382.
- [3] 刘喜英, 吴淑权, 徐向明. 基于改进分水岭算法的医学图像分割的研究[ J]. 微电子技术, 2003, 31(4): 39 - 42.
- [4] 罗玲, 解梅, 陈杉. 基于多尺度形态滤波的分水岭图像分割方法[ J]. 计算机辅助设计与图形学学报, 2004, 16(2): 168 - 173.
- [5] 赵晨光, 宋利伟, 庄天戈. 基于分水岭区域差异性的多级图像分割[ J]. 上海交通大学学报, 2004, 38(9): 1489 - 1491.
- [6] 卢官明. 一种计算图像形态梯度的多尺度算法[ J]. 中国图象图形学报, 2001, 6(3): 214 - 218.