

基于核方法的模拟电路故障诊断

陈文华

(浙江师范大学 数理与信息工程学院, 浙江 金华 321004)

摘要: 核方法通过非线性映射将原始数据嵌入到高维特征空间, 然后进行线性分析和处理, 为基于知识的数据分析带来新的方法和模式; 传统方法无法解决故障特征数据维数高、在故障样本交叠严重时多分类性能较差的问题, 因此在电路故障特征数据预处理阶段, 提出了分步骤分别在时域对电路输出电压波形进行小波包分析和在频域测量电路幅频特性的方法来提取电路故障特征; 预处理后的故障特征向量只是 8 维向量, 减少了 SVM 的训练时间; 将该方法应用于国际标准电路中的 CTSV 滤波器电路的故障诊断, 结果表明: 该方法能突出不同故障的特性, 故障诊断正确率达到 98.57% (414/420)。

关键词: 核函数; 小波包; 故障诊断; 支持向量机

Analog Circuit Fault Diagnosis Based on Kernel Methods

Chen Wenhua

(College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, China)

Abstract: New methods and models for knowledge-based data analysis are brought by kernel method. It uses nonlinear mapping, the original data is embedded into a high dimensional feature space, then conduct the linear analysis and processing. The traditional methods cannot solve the problem of poor classification performance in the case of high dimensionality of fault feature data and serious overlap of fault samples, therefore in the circuit fault feature data pre-processing stage, we propose a method of extracting circuit fault feature. The method step by step, respectively, conduct wavelet packet analysis of Circuit Output Voltage Waveform in the time domain and measure amplitude-frequency characteristics of the circuit in the frequency domain; Preprocessed fault feature vectors is just eight-dimensional vector, thus reducing training time of SVM. When the method is applied in the fault diagnosis of the CTSV filter circuit in international standard circuit, the results have shown that this method can highlight the different characteristics of the fault, and the fault diagnostic correct rate of 98.57% is achieved.

Key words: kernel function; wavelet packet; fault diagnosis; support vector machine

0 引言

模拟电路故障诊断经过多年的研究, 已初步奠定了故障诊断的基础理论, 但已有的故障诊断方法大多集中在讨论单故障、硬故障的诊断, 具有容差的模拟电路多故障、软故障诊断的研究明显不足; 利用智能计算技术进行模拟电路故障诊断时, 大多数研究者都集中于神经网络的应用, 但这种基于经验风险最小化原则的学习机器, 如何从有限的故障样本中得到泛化能力较强的决策函数还是个难题, 因其渐进性的前提条件得不到满足, 导致在故障诊断这种小样本问题中难以取得令人满意的结果^[1]。

目前模拟电路故障智能诊断面临的主要难题是典型故障样本的严重不足以及诊断知识的发现问题, 基于统计学习理论的支持向量机器学习算法为两类核方法之一, 很好地执行了结构风险最小化原则, 其应用于故障诊断最大的优势在于它适合于小样本决策, 能在有限特征信息情况下, 最大限度地发掘数据中隐含的分类知识^[2]。

本文分别用对被诊断电路输出波形进行小波包分解和测量

电路频率响应的方法提取电路故障特征, 然后分步用支持向量机对随机选择的故障样本进行训练和预测, 取得了很高的故障诊断率。

1 核方法简介

核方法的基本思想是将低维向量空间中的随机向量利用非线性函数映射到高维特征空间, 然后在高维特征空间中设计线性的学习算法。如果特征空间各坐标分量间的相互作用仅限于内积, 则无需知道映射函数的具体形式, 只要利用满足 Mercer 定理的核函数替换线性算法中的内积, 就可实现原空间中对应的非线性算法。

2 小波包分析方法简介

小波包分析方法将信号频带进行多层次划分, 对小波分析没有进一步分解的高频分量作进一步分解, 并能根据被分析信号的特征, 自适应地选择频带, 使之与信号频谱相匹配, 从而提高时-频分辨率。

3 待诊断电路故障特征提取方法

提取故障特征是模拟电路故障诊断的关键, 也是构造样本集的基础, 同时是进行电路故障诊断和测试的难点。单一的方法往往难以达到高的故障正确诊断率, 因此, 本文先后采用两种故障特征提取方法。

方法一: 小波包分析提取故障特征的方法。运用小波包分

收稿日期: 2014-01-06; 修回日期: 2014-02-28。

基金项目: 江苏省教育厅 2012 年度一般项目(Y201226067)。

作者简介: 陈文华(1966-), 男, 湖北人, 理学硕士, 主要从事电路故障智能诊断与测试技术及嵌入式系统的研究。

解提取故障特征值是基于“能量—故障”的诊断方法^[3]，在小波包分解系数单支重构的基础上，以各频段信号能量为基础来构造故障特征向量。该方法直接利用各频率成份能量的变化来诊断故障，无需知道系统的模型结构，原因在于各频率成份信号能量中包含丰富的故障信息，某种或某几种频率成份能量的改变即代表对应的故障模式。

具体步骤为：首先，记录一段输出电压波形，持续时间只需要为 ms 级，然后用 Matlab 中的 wpdec 函数对波形的采样点数据进行 3 层小波包分析得到有 8 个叶结点的完整的小波包二叉树，最后用 wenergy 函数对此小波包二叉树进行分析得到与 8 个叶结点对应的 8 个频段的能量百分比。对此百分比进行归一化处理即得 8 维的故障特征向量。

方法二：基于频率响应提取故障特征的方法^[4]。电路的频率响应反映了电路对不同频率输入信号传输能力的差异，在电路拓扑结构确定的情况下完全由构成电路的元器件参数决定。元器件参数的变化必然导致频率响应的变化。测量出各个模式下的输出电压的频率响应曲线，找出差异最大频率范围，在该范围内均匀选取 8 个有效采样点，各采样点对应的电压幅度值组合在一起就构成了各个模式的 8 维故障特征向量。

4 仿真研究

本文选择 ITC' 97 国际标准电路中的 CTSV 滤波器为例进行仿真和分析。CTSV 滤波器电路原理图如图 1 所示。电路中电阻和电容等无源器件的容差均设为 10%。无源器件的标称值为 $R_1 = R_2 = R_3 = R_4 = R_5 = 10\text{ k}\Omega$, $R_6 = 3\text{ k}\Omega$, $R_7 = 7\text{ k}\Omega$, $C_1 = C_2 = 20\text{ nF}$ 。

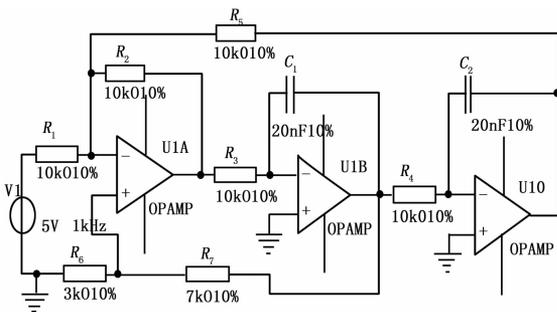


图 1 ITC' 97 的 CTSV 滤波器原理图

经分析发现 R_3 和 R_4 灵敏度极低，可不考虑其软故障。需要诊断的待测量电路的软故障包括 8 种单一软故障，分别为 $F_1 (R_1 + 50\%)$, $F_2 (R_1 - 50\%)$, $F_3 (C_1 + 50\%)$, $F_4 (C_1 - 50\%)$, $F_5 (C_2 + 50\%)$, $F_6 (C_2 - 50\%)$, $F_7 (R_5 + 50\%)$, $F_8 (R_5 - 50\%)$ ；3 种多软故障组合，分别为 $F_9 (R_1 + 50\% \& C_1 + 50\%)$, $F_{10} (R_1 - 50\% \& C_1 - 50\%)$, $F_{11} (R_5 + 50\% \& C_2 + 50\%)$ ，其中符号“+”和“-”分别表示故障值上偏和下偏离元件的标称值；加上电路正常工作状态 NF，则共有 12 种状态模式。

首先用上文所述方法一提取故障特征向量。用 NI-Multisim10 电路仿真软件对 CTSV 电路进行仿真，所有的故障模式均采用正态蒙特卡罗分析，每种故障模式仿真 35 次。信号源 V_1 向电路输入频率为 1 kHz、幅度为 5 V 正弦电压作为电路的激励信号，在 OPAMP U1C 输出端采集各种故障状态下

电路的响应电压信号。对电路每次响应信号的 $1 \sim 1001\ \mu\text{s}$ 信号进行采样可得 1 000 个数据点。首先对数据点进行小波包分析以提取故障特征。经过分析和比较，当基小波采用 ‘db1’ 小波，分解层次选为 3 层时可达较高的正确诊断率；具体为提取第 3 层从低频到高频共 8 个频段的能量，以各频段能量百分比构造 8 维特征向量，分别得到各种故障模式下的 35 个样本集，这样共可得 420 个 8 维样本集。随机选择其中的 320 个样本用以训练 SVM，其余 100 个样本用来测试 SVM 的预测准确性。

参数的选择对 SVM 的预测准确性有很大的影响。关于参数的优化选取，国际上并没有公认的最好方法，目前常用网格搜索法和交叉验证 (CV, Cross Validation) 法来确定参数。CV 是用来验证分类器性能的一种统计分析方法，基本思想是在某种意义下将原始数据进行分组，一部分作为训练集 (train set)，另一部分作为验证集 (validation set)。首先用训练集对分类器进行训练，再利用验证集来测试训练得到的模型，以得到的分类准确率作为评价分类器的性能指标。常见的 K 重交叉验证 ($K-CV$) 方法是把原始数据划分成类别分布相似、大小相同的 K 个样本子集，然后进行 K 次交叉验证过程，每次取其中的 $K-1$ 个子集组合在一起作为训练集，剩下的 1 个子集作为测试集，这样会得到 K 个模型，用这 K 个模型最终的验证集的分类准确率的平均数作为此 $K-CV$ 下分类器的性能指标。 $K-CV$ 可以有效地避免过学习和欠学习的发生。通过比较发现核函数采用径向基函数 (RBF) 时预测准确率最高。网格搜索是确立参数对 (c, g) ，其中 c 为惩罚系数， g 为径向基函数唯一的参数。目标是选择最好的参数对 (c, g) ，得到预测准确率最高的交叉验证结果，具体流程如图 2 所示。

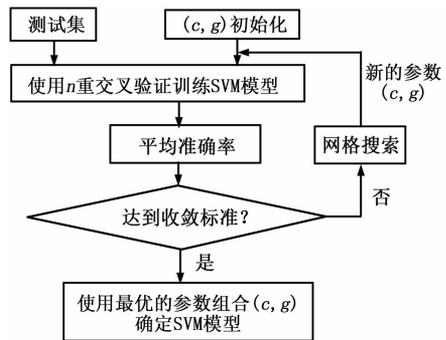


图 2 参数 (c, g) 确定流程图

利用网格搜索法确定参数 c 和 g 时，需预先设定初值，可先让参数呈指数增长的规律来快速初步确定其取值范围。例如取 $c = (2^{-5}, 2^{-3}, \dots, 2^{15})$, $g = (2^{-15}, 2^{-13}, \dots, 2^3)$ ；参数 c 和 g 在二维坐标系上构成一个二维网格。在该网格上依次选取每一组参数对 (c, g) 的值对支持向量机进行训练并将训练后的支持向量机在测试集上进行测试，计算测试准确率，根据准确率的测试结果确定最佳的参数对 (c, g) 。若测试准确率达不到要求，可在现有稀疏网格搜索的基础上选定一个准确率较高而数值范围较小的搜索区域，减小搜索步长进行进一步的精细搜索以求得更高的精度。假如通过粗略搜索得到 (c, g) 的值为 $(2^3, 2^{-5})$ ，以该参数对为中心值，在其附近区域

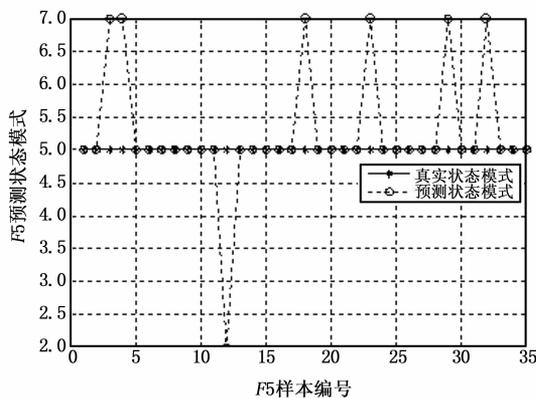
减小搜索步长，重新划分 $c = (2^1, 2^{1.25}, \dots, 2^3)$, $g = (2^{-7}, 2^{-6.75}, \dots, 2^{-3})$, 按上述方法继续寻找更为精确的参数值。经过粗略搜索和精细搜索两个步骤可以得到比较精确的参数对 (c, g) 。网格搜索的交叉验证法简单易行，并且可以同时搜索两个参数值，最终获得的参数对 (c, g) 能使判决函数分类准确率达到较优；该方法不容易遗漏好的参数，针对具体问题试算的方法往往能得到比较小的泛化误差。该方法的缺点是需要反复进行多次训练和检验，计算量大，比较耗时。

总体结果如下：训练样本预测准确率 = 96.562 5% (309/320)；测试样本预测准确率 = 96% (96/100)。各状态模式诊断结果如表 1 所示。

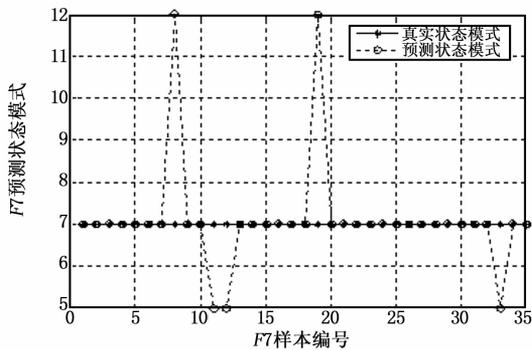
表 1 小波包分析提取故障特征方法的测试样本预测准确率

故障模式	F1	F2	F3	F4	F5	F6
样本总数	35	35	35	35	35	35
正确诊断数	35	35	35	35	28	35
诊断正确率 (%)	100	100	100	100	80	100
故障模式	F7	F8	F9	F10	F11	NF
样本总数	35	35	35	35	35	35
正确诊断数	30	34	34	35	34	35
诊断正确率 (%)	85.71	97.14	97.14	100	97.14	100

由表可知只有 F5 和 F7 的诊断准确率稍低，详细结果如图 3 所示。



(a) F5 诊断结果



(b) F7 诊断结果

图 3 F5 和 F7 的诊断结果

仔细分析误诊结果发现：F5 的 7 个误诊样本中有 6 个误诊为 F7，F7 的 5 个误诊样本中有 3 个误诊为 F5，基本上是 F5 与 F7 之间交错。为了提高 F5 与 F7 的诊断准确率可采取

如下所述的方法：只要诊断结果为 F5 和 F7 则认为其不可信，尽管诊断准确率也分别达到了 80% 和 85.7%。遇到这种情况，需要进入下一步：用基于频率响应提取故障特征的方法来提取电路的故障特征。测量电路的频率响应，经分析和比较发现当 F5 和 F7 发生时，在低于大约 2 kHz 的频率范围内，电路的幅频特性曲线有显著的差别。在频率范围 1~2 001 Hz 内均匀选取 8 个有效采样点，各采样点对应的电压幅值组合在一起就分别构成了 2 种故障模式的 8 维故障特征向量。用正态蒙特卡罗分析对这 2 种故障模式分别仿真 35 次共可得到 70 个样本，用其训练 SVM。再次用正态蒙特卡罗分析对这 2 种故障模式分别仿真 35 次得到 70 个样本，用训练好的 SVM 对新样本进行预测，2 种故障模式的诊断准确率都为 100%。综合使用两种故障特征提取方法，故障诊断正确率高达 98.57% (414/420)。

表 2 给出不同方法下的诊断正确率^[5]。这里 K&SVM 表示利用传统的、未经优化的 K 均值聚类方法和支持向量机进行组合诊断；K&FSVM 表示利用传统的、未经优化的 K 均值聚类方法和模糊支持向量机进行组合诊断；MCBKM 表示利用优化的类互均衡核 K-均值法 (Mutual cluster balanced K-means, 简称 MCBKM) 技术和模糊支持向量机进行组合诊断。Ours 表示本文采用的方法进行组合诊断。

表 2 不同方法对应的诊断正确率

诊断方法	K&SVM (%)	K&FSVM (%)
单一软故障	90.98	92.74
多软故障	82.47	90.98
诊断方法	MCBKM (%)	Ours (%)
单一软故障	94.96	98.57 (276/280)
多软故障	94.49	98.10 (103/105)

5 结论

本文提出的故障特征提取方法和分步故障诊断方法是一种简单、物理概念清晰的方法。从诊断性能来看，本文所提出的诊断方法较传统的 K&SVM 和 K&FSVM 以及 MCBKM 更适用于容差模拟电路的故障特征数据。并且对软多故障的诊断，本文所提出的方法性能最优，在保持精度的同时有效降低了故障特征数据的维数，具有很高的故障识别率。对该容差模拟电路的多故障诊断正确率最低达到 97.14%，具有很好的诊断性能和广阔的工程应用前景。

参考文献：

[1] 毛先柏. 基于支持向量机的模拟电路故障诊断研究 [D]. 武汉: 华中科技大学, 2009.
 [2] 谢保川, 刘福太. 支持向量机在模拟电路故障诊断中的应用 [J]. 计算机仿真, 2006, 23 (10): 167-171.
 [3] 王 雪. 测试智能信息处理 [M]. 北京: 清华大学出版社, 2008.
 [4] 毛先柏, 李昌禧. PCA-SVM 在模拟电路故障诊断中的应用 [J]. 计算机测量与控制, 2009, 17 (7): 1250-1252.
 [5] 唐 静, 秦 亮, 史贤俊, 等. 基于核理论均衡聚类和支持向量机的模拟电路诊断方法 [J]. 中南大学学报 (自然科学版), 2011, 42 (9): 108-112.