

一种基于 EMD 的模拟电路故障特征提取方法

侯青剑, 王宏力

(第二炮兵工程学院控制科学与工程系, 陕西 西安 710025)

摘要: 为克服小波分析中混频现象对模拟电路故障特征提取造成的不准确, 提出了一种基于经验模式分解的模拟电路故障特征提取方法。该方法通过对模拟电路输出信号进行 EMD 得到若干个内禀模态函数, 以各 IMF 的能量作为故障判别的特征。并针对 EMD 中的端点效应问题, 提出了基于遗传优化的最小二乘支持向量机预测方法。仿真实验结果证明了该方法的有效性。

关键词: 模拟电路; 故障特征; 经验模式分解; 端点效应; 最小二乘支持向量机; 遗传算法

中图分类号: TP 206+ .3 **文献标志码:** A

Method of fault feature extraction for analog circuits based on EMD

HOU Qing-jian, WANG Hong-li

(Dept. of Control Science and Engineering, The Second Artillery Engineering Inst., Xi'an 710025, China)

Abstract: To overcome the imprecision of fault feature extraction for analog circuits because of the mixing phenomenon of wavelet analysis, a method of fault feature extraction for analog circuits based on empirical mode decomposition (EMD) is put forward. The method gets several intrinsic mode functions (IMF) through the EMD of analog circuit output signals, and the energy of each IMF is regard as the feature to distinguish faults. Aiming at the problem of end effect on EMD, a predictive method of least square support vector machines (LSSVM) based on genetic optimization is put forward. The results of simulation experiment prove the validity of this method.

Keywords: analog circuit; fault feature; empirical mode decomposition; end effect; least square support vector machine; genetic algorithm

0 引言

模拟电路广泛应用于军工、通讯、自动控制、测量仪表、家用电器等诸多方面, 因此, 模拟电路的故障诊断技术也受到了极大的关注。在过去的几十年中, 形成了一系列的模拟电路故障诊断理论与方法。但由于模拟电路故障诊断的复杂性, 尤其是非线性模拟电路软故障诊断的复杂性, 目前还需要进一步深入研究。

故障特征的提取是故障诊断的关键步骤之一。当电路出现故障时, 其传递函数的幅频特性将会改变, 对于相同的输入, 其输出将会发生改变, 各频率成分的能量将会发生改变。因此, 具有多分辨率分析特性的小波变换被用于故障特征的提取^[1-2]。经过小波包分解, 提取各频带能量作为故障特征被证实可以用于模拟电路的故障诊断。但小波包分解存在着严重的混频现象, 对故障特征的提取带来了不利的影响。

经验模式分解(empirical mode decomposition, EMD)是近年来发展起来的一种新的信号分析方法, 它不仅吸取了小波变换多分辨率分析的优点, 同时其分解过程又具有自适应性, 对信号各频率成分的描述更为精细^[3]。EMD 方

法已被广泛应用于海洋波动分析、地震信号分析、旋转机械故障诊断等方面^[4]。但是, 在应用 EMD 时存在一个比较棘手的端点效应问题。在运用 EMD 方法对非平稳信号进行分解时, 在数据的两端会产生发散现象, 并且这种发散的结果会逐渐向内“污染”整个数据序列而使得结果严重失真。因此, 在运用 EMD 对信号进行分析时, 必须对信号两端进行延拓以抑制端点效应。

本文以非线性模拟电路的软故障为对象, 引入基于最小二乘支持向量机(least square support vector machine, LSSVM)预测的 EMD 端点效应抑制方法, 提出了一种基于 EMD 的故障特征提取方法。通过 LSSVM 预测对信号两端进行延拓, 并采用遗传算法对 LSSVM 参数进行自动优化, 然后运用 EMD 分解得到若干个内禀模态函数(intrinsic mode function, IMF)分量, 提取各 IMF 分量的能量为故障特征以实现故障诊断。

1 EMD 及特征提取方法

EMD 方法的基本思想是: 假如一个原始数据序列 $X(t)$ 的极值点数目比零点数目多两个(或两个以上), 或者由极大值点决定的上包络与由极小值点决定的下包络的均值不

为零, 该数据序列就需要进行平稳化处理。具体处理方法是: 找出 $X(t)$ 的所有极大值点, 用三次样条函数对极大值点序列插值构成 $X(t)$ 的上包络线; 同理利用极小值点得到 $X(t)$ 的下包络线; 上下包络线的均值为平均包络线 $m_1(t)$; 将原数据序列 $X(t)$ 减去该平均包络 $m_1(t)$ 后得到一个新数据序列 $h_1(t)$, 即

$$X(t) - m_1(t) = h_1(t) \tag{1}$$

一般来讲, $h_1(t)$ 仍然不是一个平稳数据序列, 为此需对它重复上述处理过程, 即

$$h_1(t) - m_{11}(t) = h_{11}(t) \tag{2}$$

重复进行上述处理过程 k 次, 直到所得到的 $h_{1k}(t)$ 是一个 IMF 为止。这样就得到了第一个 IMF 分量 $C_1(t)$, 即

$$C_1(t) = h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \tag{3}$$

第一个 IMF 分量代表原始数据序列中最高频的组成成分。将原始数据序列 $X(t)$ 减去第一个分量 $C_1(t)$, 可以得到一个去掉高频组成成分的差值数据序列 $r_1(t)$ 。由于 $r_1(t)$ 仍然包含具有较长周期组成成分的信息, 所以 $r_1(t)$ 把当作新的数据对其进行上述平稳化处理过程, 得到第二个 IMF 分量 $C_2(t)$ 。如此重复下去直到最后一个差值序列 $r_n(t)$ 不可再被分解为止, 此时 $r_n(t)$ 代表原始数据序列的均值或趋势

$$\begin{aligned} X(t) - C_1(t) &= r_1(t) \\ &\vdots \\ r_{n-1}(t) - C_n(t) &= r_n(t) \end{aligned}$$

原始数据序列可由这些 IMF 分量和一个均值或趋势项表示

$$X(t) = \sum_{i=1}^n C_n(t) + r_n(t) \tag{4}$$

各 IMF 分量的能量 E_i 定义为

$$E_i = \sum_{j=1}^{M_i} |C_i^j|^2 \tag{5}$$

式中, $i=1, \dots, n$, n 为 IMF 分量的个数; M_i 为第 i 个 IMF 分量的数据长度; C_i^j 为第 i 个 IMF 分量的第 j 个数据。

2 基于 LSSVM 的端点效应处理

针对 EMD 过程中的端点效应问题, Huang 等提出在数据两端增加两组特征波的方法进行数据延拓, 但并未公开确定合适的特征波的具体方法, 并已将该方法在美国申请了专利。目前, 人们已提出了一些抑制端点效应的方法, 包括神经网络数据延拓法、ARMA 模型数据延拓法、多项式数据延拓法等, 这些方法对抑制端点效应都有一定的效果^[5]。本文提出了基于 LSSVM 预测的数据延拓方法以克服端点效应, 并通过遗传算法对 LSSVM 参数进行自动优化。

2.1 LSSVM 回归算法^[6]

首先介绍线性回归算法, 设输入样本为 n 维向量, 某区域的 l 个样本及其值表示为

$$(x_1, y_1), \dots, (x_l, y_l) \in R^n \times R$$

线性函数设为

$$f(x) = \mathbf{w}^T x + b \tag{6}$$

优化问题为

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^l \xi_i \tag{7}$$

约束为

$$y_i = \mathbf{w}^T x_i + b + \xi_i, \quad i = 1, \dots, l \tag{8}$$

定义拉格朗日函数为

$$L = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (\mathbf{w}^T x_i + b + \xi_i - y_i) \tag{9}$$

对各参数求偏导, 并令其为零, 表示成矩阵形式, 得

$$\begin{bmatrix} \mathbf{I} & 0 & 0 & -\mathbf{x} \\ 0 & 0 & 0 & -\mathbf{1}^T \\ 0 & 0 & \gamma & -\mathbf{I} \\ \mathbf{x}^T & \mathbf{1} & \mathbf{I} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ b \\ \xi \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbf{y} \end{bmatrix} \tag{10}$$

式中, \mathbf{I} 为单位矩阵

$$\mathbf{x} = [x_1, \dots, x_l], \quad \mathbf{y} = [y_1, \dots, y_l]^T,$$

$$\mathbf{1} = [1, \dots, 1]^T, \quad \xi = [\xi_1, \dots, \xi_l]^T, \quad \alpha = [\alpha_1, \dots, \alpha_l]^T$$

最后的解由下面的矩阵确定

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \mathbf{x}^T \mathbf{x} + \gamma^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{11}$$

对于非线性回归, 使用一个非线性映射 ϕ 把数据映射到一个高维特征空间, 选取合适的核函数 $K(x, y)$, 再在高维特征空间进行线性回归, 最后非线性回归函数可以表示为

$$f(x) = \sum_{i=1}^l \alpha_i K(x_i, x) + b \tag{12}$$

2.2 基于遗传优化的 LSSVM 回归算法

LSSVM 的正规范化参数 γ 和核函数参数对 LSSVM 的性能有较大影响, 但又没有成熟的参数选取方法^[7]。本文采用遗传算法对 LSSVM 的正规范化参数 γ 和核函数参数进行二进制编码, 以十重交叉验证的均方差 (MSE) 为指标实现参数优化。算法实现的具体步骤为:

步骤 1 对待优化参数进行编码, 并建立初始种群。

以径向基核函数 (RBF) $K(x_i, x_j) = \exp\left[-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right]$

为例, 对正规范化参数 γ 和核函数参数 σ^2 采用二进制进行编码, 根据精度要求设定编码长度及种群数量;

步骤 2 计算种群中个体的适应度。根据步骤 1 中编码的定义, 对各种群体进行译码, 以译码后得到的值为 LSSVM 参数, 利用学习样本对 LSSVM 进行训练, 以学习样本的十重交叉验证均方差的倒数为个体的适应度;

步骤 3 选择。从当前种群中选择适应度高的优良个体直接进入下一代中去;

步骤 4 交叉。根据交叉概率, 并按照某种规律选择两个个体, 将他们的部分二进制编码进行交换, 产生新的个体并添加到新的种群中;

步骤 5 变异。随机地改变一些个体的编码, 产生新的个体并添加到新的种群中;

步骤 6 以种群最大适应度收敛或以指定进化代数为结束, 反复执行步骤 2-5。最佳个体译码后得到的结果即为最优的 LSSVM 参数。

2.3 基于 LSSVM 的数据序列延拓方法

假设数据序列的采样时间间隔是 τ 得到的数据序列为

$$x(0), x(\tau), \dots, x(i\tau), \dots, x((n-1)\tau) \quad (13)$$

该数据序列也是对未来值进行预测时的输入序列。由 n 时刻的前 m 个值预测第 n 个值的问题可表示为寻找如下的对应关系 f 的问题

$$x_n = f(x_{n-1}, x_{n-2}, \dots, x_{n-m}) \quad (14)$$

式中, x_i 是 $x(i\tau)$ 的缩写形式, $i=1, 2, \dots$

构造如下的训练样本对: 输入 (x_1, x_2, \dots, x_m) 对应某一时刻的输出为 x_{m+1} , 输入 $(x_2, x_3, \dots, x_{m+1})$ 对应某一时刻的输出为 x_{m+2} , 并依此类推, 由 l 个训练样本就可构建 $l-m$ 个训练样本对。采用构造的训练样本训练 LSSVM 模型, 并通过遗传算法对 LSSVM 的参数进行优化。

当模型训练完成后, 对未来值第 1 步预测的形式为

$$x_{m+1}^* = f(x_n, x_{n-1}, \dots, x_{n-m+1}) \quad (15)$$

第 2 步预测为

$$x_{m+2}^* = f(x_{m+1}^*, x_n, x_{n-1}, \dots, x_{n-m+2}) \quad (16)$$

后续各步的预测依此类推, 并由此形成多步预测。其中, x_i 为第 i 点的真实值, x_i^* 为第 i 点的预测值, $i=1, 2, \dots$

3 仿真研究

本文以仿真信号为例进行了仿真研究。该仿真信号为

$$X(t) = 0.6\cos\left|\frac{2\pi}{10}t\right| + \cos\left|\frac{2\pi}{20}t\right| + 0.5\sin\left|\frac{2\pi}{80}t\right| \quad (17)$$

式中, $t \in [-57, 57]$ 。

采样频率为 1 Hz。其波形如图 1 所示。

3.1 小波包能量特征提取方法

对仿真信号进行 4 层 db1 小波包分解, 在尺度 4 上形成 16 个频带, 以文献[8]的方法计算各频带信号的能量, 图

2 为各频带能量直方图。

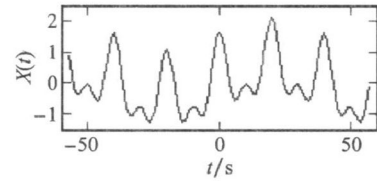


图 1 仿真信号的时域波形

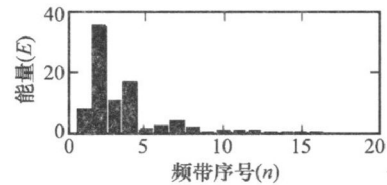


图 2 频带能量直方图

由仿真信号可知, 第 1、2 和 4 频带的能量应当较大, 但图 2 中第 3 频带和其它几个频带的能量也较大, 因此可以看出, 小波包分解出现了较为严重的混频现象, 为故障特征的提取带来了不利影响。

3.2 基于 EMD 的特征提取方法

对仿真信号直接进行 EMD, 得到的各 IMF 分量如图 3 所示。以径向基函数为核函数, 对参数 ν 和 σ^2 进行二进制编码, 长度均为 20, 搜索范围分别为 100-100 000 和 0.1-100, 种群个数为 20, 代沟为 0.9, 最大遗传代数为 30, 选择方法为随机遍历抽样, 单点交叉概率为 0.7, 变异概率为 0.017 5, 采用本文提出的方法对仿真信号进行延拓后经 EMD 得到的各 IMF 分量如图 4 所示。图中实线为真实信号中的各分量, 虚线为 EMD 分解得到的 IMF 分量。

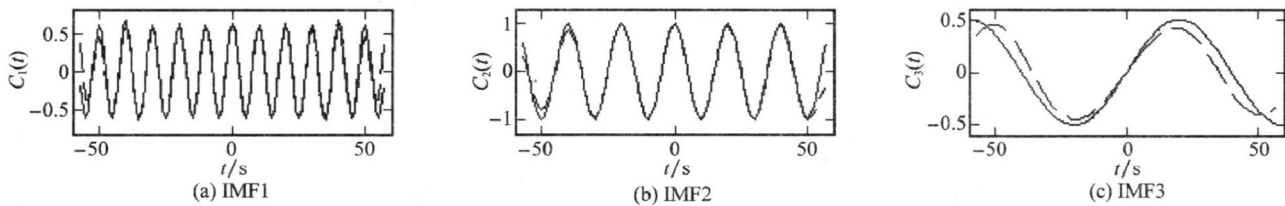


图 3 直接进行 EMD 的 IMF 分量

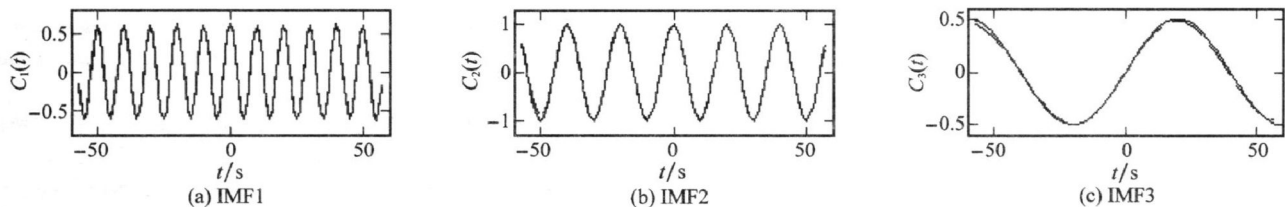


图 4 延拓后 EMD 的 IMF 分量

从图 3 和图 4 可以看出, 信号的直接 EMD 产生了严重的端点效应, 本文提出的信号延拓方法有效地抑制了 EMD 中的端点效应问题, 得到了准确的 IMF 分量。

计算图 4 中的各 IMF 分量能量并绘出能量直方图, 如图 5 所示。

由图 5 可见, 经 EMD 提取的能量特征更为准确, 对信号的描述也更加精细, 能够较好地表征信号的特征。

4 实例应用

带通滤波器电路如图 6 所示, 各参数值及输入信号见

图。其中, V_3 、 V_4 的相位设为 90° 。在不考虑电路容差的情况下, 假设 $R_1=3\text{ k}\Omega$ 为故障 1, $C_1=8\text{ nF}$ 为故障 2。通过 Pspice 仿真, 对运算放大器输出电压进行采样, 采样起始时间相同, 采样间隔为 $1\ \mu\text{s}$, 共采样 189 个点, 以本文提出的方法对电路在正常、故障 1 和故障 2 情况下的采样数据进行特征提取, 遗传操作的设置与 3.2 节相同, 其前 3 个 IMF 分量的能量特征直方图如图 7 所示。

由图 7 可以看出, 电路在 3 种不同状态下的能量特征有显著不同, 该特征可以表征电路的不同状态, 能够有效地

对故障进行判别。

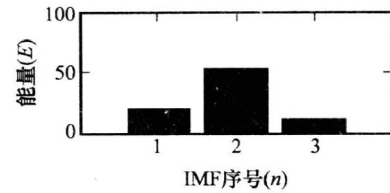


图 5 IMF 能量直方图

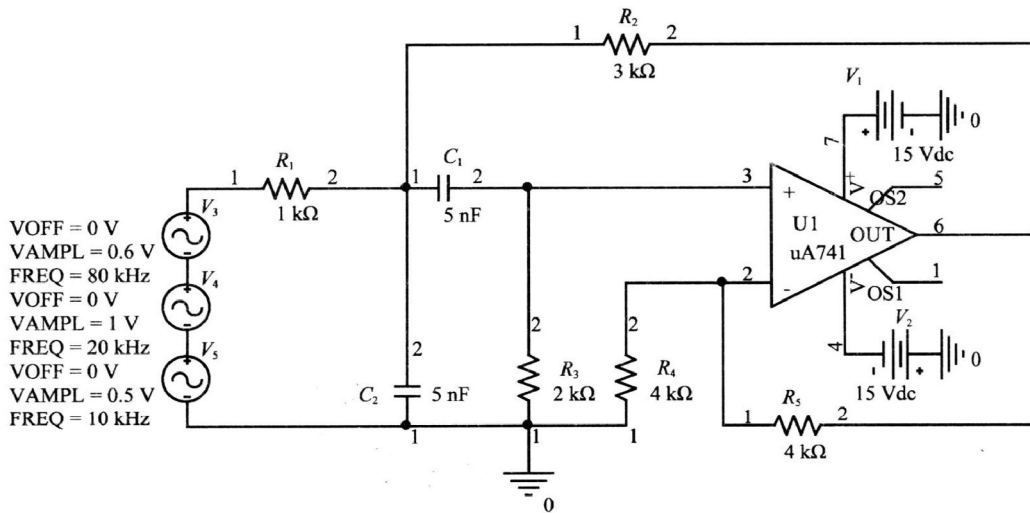


图 6 带通滤波器电路

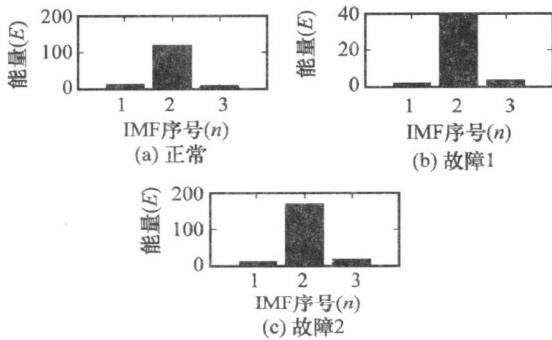


图 7 3 种状态下的能量特征直方图

5 结 论

本文针对 EMD 中严重的端点效应问题, 提出了基于 LSSVM 预测的数据延拓方法, 并通过遗传算法优化 LSSVM 的性能, 仿真结果表明, 该方法有效地抑制了 EMD 中的端点效应。在此基础上, 提出了一种新的模拟电路故障特征提取方法, 该方法通过 EMD 得到各 IMF, 并以各 IMF 的能量为特征进行故障判别。实际模拟电路的特征提取实验结果表明, 该方法能够有效地表征电路的不同状态, 可以作为判别电路故障的特征。

本文中仅考虑了电路的两种故障, 但在实际应用中, 可以使用同样的方法得到电路在其他故障时的特征, 利用这

些特征可以对电路的不同故障进行诊断。

另外, 由于遗传算法的引入, 信号 EMD 的时间稍长, 但对于离线故障诊断来说影响不大。为满足某些实际应用中对于实时性的需求, 对本文提出的方法进行深入研究, 缩短特征提取时间, 是下一步的主要工作。

参考文献:

- [1] 谭阳红, 叶佳卓. 模拟电路故障诊断的小波方法[J]. 电子与信息学报, 2006, 28(9): 1748-1751.
- [2] 王奉涛, 马孝江, 邹岩崑, 等. 基于小波包分解的频带局部能量特征提取方法[J]. 农业机械学报, 2004, 35(5): 177-180.
- [3] Huang N E, Zheng S, Steven R L, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[C] // *Proc. of R. Soc. London. A*, 1998, 454: 903-995.
- [4] 王奉涛, 马孝江, 邹岩崑, 等. 基于局域波法的时频局部能量特征提取方法[J]. 中国机械工程, 2003, 14(23): 2012-2014.
- [5] 贾嵘, 王小宇, 蔡振华, 等. 基于最小二乘支持向量机回归的 HHT 在水轮发电机组故障诊断中的应用[J]. 中国电机工程学报, 2006, 26(22): 128-133.
- [6] Vapnik V N. The nature of statistical learning theory[M]. *New York: Springer-Verlag*, 1995.
- [7] 江江, 陆颂元. 汽轮发电机组故障诊断 GA-SVM 模型方法的研究[J]. 汽轮机技术, 2005, 47(1): 1-16.
- [8] 吴桂峰, 翟玉庆, 陈虹, 等. 基于小波-神经网络的电机振动故障诊断[J]. 控制工程, 2004, 11(2): 152-154.