

基于非线性频谱数据驱动的动态系统故障诊断方法

张家良¹, 曹建福¹, 高峰¹, 韩海涛²

(1. 西安交通大学 机械制造系统工程国家重点实验室,
西安 710049; 2. 第二炮兵工程学院 101教研室, 西安 710025)

摘要: 基于非线性频谱数据驱动方法, 研究了动态系统的故障诊断问题. 利用一维非线性输出频率响应函数提出一种非线性频谱特征提取方法, 为了提高实时性, 采用变步长自适应辨识算法进行求解; 根据估计偏差实时地改变步长, 兼顾了收敛速度与稳态误差; 获取了非线性频谱特征之后, 利用最小二乘支持向量机分类器进行故障识别. 通过对提升设备的故障诊断问题进行实验研究, 所得结果表明, 所提出的算法识别率高, 能满足在线诊断要求.

关键词: 故障诊断; 非线性频谱; 自适应辨识; 支持向量机

中图分类号: TP277

文献标志码: A

Fault diagnosis approach of dynamic system based on data driven of nonlinear spectrum

ZHANG Jia-liang¹, CAO Jian-fu¹, GAO Feng¹, HAN Hai-tao²

(1. State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. Staff Office 101, The Second Artillery Engineering University, Xi'an 710025, China. Correspondent: CAO Jian-fu, E-mail: cjf@mail.xjtu.edu.cn)

Abstract: The problem of fault diagnosis for the dynamic system is studied based on the data driven method of nonlinear spectrum. An extraction method of nonlinear frequency spectrum feature is proposed by using one dimensional nonlinear output frequency response function. In order to improve timeliness, the variable step size adaptive identification algorithm is used to solve the nonlinear output frequency response function. The step size is changed according to estimating error so that convergence rate and steady state error are both considered. After obtained nonlinear frequency spectrum feature, the least square support vector machine classifier is used to fault identification. The fault diagnosis of hoisting equipment is researched, and experiments show that the proposed algorithm has the good high recognition rate that can fulfill the demand of online diagnosis.

Key words: fault diagnosis; nonlinear spectrum; adaptive identification; support vector machine

0 引言

非线性动态系统的故障诊断是近年来国内外研究的一个十分活跃的领域, 同时由于风电系统、数控机床、飞行器等诸多装备都是非线性系统, 对这个问题研究越来越重要. 目前, 非线性系统的故障诊断方法主要包括: 基于数学模型、信号处理和人工智能的方法, 但由于非线性系统本身的复杂性, 直到现在还没有从根本上解决其故障诊断问题.

基于数据驱动的故障诊断方法已成为故障诊断研究的一个热点, 在多元统计分析、机器学习及信息融合等方面已取得了一些成果^[1-5]. 现有的大多数数

据驱动诊断方法都未充分利用系统数据中的非线性特性, 故对非线性动态系统故障诊断问题的研究是一个重要方向. 而基于非线性频谱的方法根据非线性谱的变化来进行故障诊断, 它充分利用了系统数据中的非线性特性. 过去基于非线性频谱的诊断方法主要采用广义频率响应函数(GFRF)来获取频谱特征^[6-7], 但GFRF辨识模型的结构复杂且计算量大. 为了简化计算, 英国学者Lang等^[8]提出了一种非线性输出频率响应函数(NOFRF). NOFRF是一维频域函数, 其结构简单、计算数据量小, 可作为动态系统诊断模型, 但目前对它的实时求解算法还缺乏研究. 获得非线性频谱数

收稿日期: 2012-09-03; 修回日期: 2012-12-21.

基金项目: 陕西省科技项目(2010K08-13).

作者简介: 张家良(1982-), 男, 博士生, 从事工业系统故障诊断的研究; 曹建福(1963-), 男, 教授, 从事先进机器人控制、工业系统故障诊断与非线性系统等研究.

据后, 另一个需要解决的关键问题是故障检测与识别. 非线性频谱图包括二维图形、三维图形等, 图形的形状复杂, 直接根据图形差异或简单的阈值判断难以进行准确的故障检测与识别.

本文利用NOFRF提取非线性频谱特征, 研究非线性动态系统的故障诊断问题. 为了提高非线性频谱数据获取的实时性, 克服收敛速度与稳态误差之间的矛盾, 提出一种NOFRF的变步长频域最小均方(VSSFRLMS)自适应辨识算法. 获取了非线性频谱数据后, 通过最小二乘向量机(LSSVM)设计故障分类器以进行故障识别. 利用所提出的方法对提升设备的故障诊断问题进行研究, 实验结果表明了该算法的有效性.

1 基于数据驱动的非线性频谱特征提取

对动态系统获取其非线性频谱特征包括两种方法, 即基于解析的方法和基于数据的方法. 基于数据的方法是通过采集系统输入输出数据, 利用Volterra级数描述通过系统辨识方法来求解. 结合NOFRF描述及自适应辨识算法, 本文给出一种基于数据驱动的非线性频谱特征提取方法.

设输入为 $u(t)$, 非线性系统 n 阶NOFRF定义为^[8]

$$G_n(j\omega) = \frac{\int_{\omega_1+\dots+\omega_n=\omega} H_n(j\omega_1, \dots, j\omega_n) \prod_{i=1}^n U(j\omega_i) d\sigma_{n\omega}}{\int_{\omega_1+\dots+\omega_n=\omega} \prod_{i=1}^n U(j\omega_i) d\sigma_{n\omega}}. \quad (1)$$

其中: $H_n(j\omega_1, \dots, j\omega_n)$ 为 n 阶广义频率响应函数, $\int_{\omega_1+\dots+\omega_n=\omega} H_n(j\omega_1, \dots, j\omega_n) \prod_{i=1}^n U(j\omega_i) d\sigma_{n\omega}$ 表示 n 维超平面 $\omega_1 + \omega_2 + \dots + \omega_n = \omega$ 上的积分, $U(j\omega_i)$ 为输入频谱, $\int_{\omega_1+\dots+\omega_n=\omega} \prod_{i=1}^n U(j\omega_i) d\sigma_{n\omega} \neq 0$.

任意非线性系统可以用有限阶次NOFRF表示, 即有

$$Y(j\omega) = \sum_{n=1}^N Y_n(j\omega) = \sum_{n=1}^N G_n(j\omega) U_n(j\omega). \quad (2)$$

其中: $U_n(j\omega)$ 为 $u^n(t)$ 的傅立叶变换, $Y_n(j\omega)$ 为系统的 n 阶输出频率响应, N 为模型的最高阶次.

令 $G_n^R, G_n^I, U_n^R, U_n^I$ 分别表示 $G_n(j\omega)$ 和 $U_n(j\omega)$ 的实部和虚部, $\text{Re}Y(j\omega)$ 和 $\text{Im}Y(j\omega)$ 分别表示 $Y(j\omega)$ 的实部和虚部, 则由式(2)可得

$$\begin{bmatrix} \text{Re}Y(j\omega) \\ \text{Im}Y(j\omega) \end{bmatrix} = \begin{bmatrix} U_1^R, \dots, U_N^R, -U_1^I, \dots, -U_N^I \\ U_1^I, \dots, U_N^I, U_1^R, \dots, U_N^R \end{bmatrix} \begin{bmatrix} G^R \\ G^I \end{bmatrix}. \quad (3)$$

其中: $G^R = [G_1^R, \dots, G_N^R]^T$, $G^I = [G_1^I, \dots, G_N^I]^T$.

对非线性系统进行 $\bar{N} (\bar{N} \geq N)$ 次激励, 输入为 $u_i(t) = a_i \bar{u}(t) (i = 1, \dots, \bar{N}, a_{\bar{N}} > \dots > a_1 > 0)$, 输出为 $y_1(t), \dots, y_{\bar{N}}(t)$, 输入输出频谱分别为 $U_i(j\omega) = a_i \bar{U}(j\omega)$ 和 $Y_i(j\omega)$, 则由式(3)可得

$$Y = A_U \begin{bmatrix} G^R \\ G^I \end{bmatrix}. \quad (4)$$

其中

$$Y = [\text{Re}Y_1(j\omega), \dots, \text{Im}Y_{\bar{N}}(j\omega)]^T, \\ A_U = \begin{bmatrix} a_1 \bar{U}_1^R, \dots, a_1^N \bar{U}_N^R, -a_1 \bar{U}_1^I, \dots, -a_1^N \bar{U}_N^I \\ a_1 \bar{U}_1^I, \dots, a_1^N \bar{U}_N^I, a_1 \bar{U}_1^R, \dots, a_1^N \bar{U}_N^R \\ \vdots \\ a_{\bar{N}} \bar{U}_1^R, \dots, a_{\bar{N}}^N \bar{U}_N^R, -a_{\bar{N}} \bar{U}_1^I, \dots, -a_{\bar{N}}^N \bar{U}_N^I \\ a_{\bar{N}} \bar{U}_1^I, \dots, a_{\bar{N}}^N \bar{U}_N^I, a_{\bar{N}} \bar{U}_1^R, \dots, a_{\bar{N}}^N \bar{U}_N^R \end{bmatrix}.$$

采用最小二乘算法求解式(4), 可得

$$\begin{bmatrix} G^R \\ G^I \end{bmatrix} = (A_U^T A_U)^{-1} A_U^T Y. \quad (5)$$

上述计算过程不需要被诊断系统的结构模型信息, 完全通过输入输出数据来获取非线性频谱特征. 根据式(5)通过批量最小二乘算法来辨识NOFRF, 需涉及到矩阵的求逆运算, 计算量较大. 为了减小计算量, 提高实时性, 可以采用自适应辨识算法. 文献[9]采用分块LMS自适应算法辨识NOFRF, 根据非线性频谱数据提取故障特征, 研究了非线性模拟电路的故障诊断问题. 标准LMS自适应算法的步长较大时, 收敛速度快, 但稳态误差大; 步长较小时, 稳态误差小, 但收敛速度慢. 为了解决收敛速度与稳态误差的矛盾, 本文采用变步长最小均方(VSSFRLMS)自适应算法辨识NOFRF, 该算法在收敛过程中根据输出偏差实时地改变步长.

图1为NOFRF的VSSFRLMS自适应辨识算法结构图. 其中: $u(t)$ 为输入信号, $\tilde{Y}(j\omega)$ 为实际输出频谱, $Y(j\omega)$ 为估计输出频谱, $e(j\omega)$ 为输出频谱偏差, $U = [U_1(j\omega), \dots, U_N(j\omega)]^T$, $G = [G_1(j\omega), \dots, G_N(j\omega)]^T$.

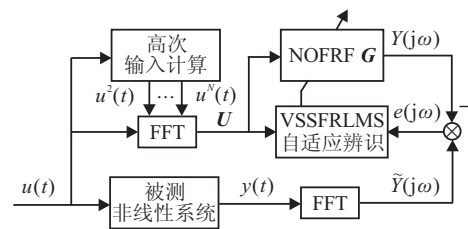


图1 VSSFRLMS自适应辨识算法结构

NOFRF的VSSFRLMS自适应辨识算法为

$$G_{k+1} = G_k + 2\mu_k e_k(j\omega) \frac{(U_k)^*}{\gamma + U_k^T(U_k)^*}, \quad (6)$$

其中“*”表示共轭复数, 为步长, 是一个较小的正数.

步长 μ_k 由下式确定:

$$\mu_k = \begin{cases} \mu_{\max}, & \mu_k' > \mu_{\max}; \\ \mu_k', & \mu_{\min} \leq \mu_k' \leq \mu_{\max}; \\ \mu_{\min}, & \mu_k' < \mu_{\min}. \end{cases} \quad (7)$$

其中: $\mu_k' = \alpha\mu_{k-1} + \beta|e_{k-1}(j\omega)|^2$, $0 < \alpha < 1$, $\beta > 0$; μ_{\max} 为步长上限, μ_{\min} 为步长下限.

由式(7)可看出步长 μ_k 的大小由估计误差决定: 当误差较大时 μ_k 较大, 使得算法收敛速度较快; 当误差较小时 μ_k 较小, 使得算法稳态误差较小. 因此提出的自适应辨识算法能兼顾收敛速度与估计误差.

选择采样频率为 f_s , 采样长度为 L , 令 $\mathbf{G}_1 = [0, \dots, 0]^T$, $\mu_1 = \mu_{\max}$, NOFRF 的 VSSFRLMS 自适应算法的步骤如下所示.

Step 1: 采集第 k ($k \geq 1$) 组输入、输出数据并分别进行傅立叶变换, 得到输入频谱向量 $\mathbf{U}_k = [U_{(1)}^{(1)}(j\omega), \dots, U_{(L(N,m))}^{(N)}(j\omega)]^T$ 与输出频谱 $\tilde{Y}_i(j\omega)$;

Step 2: 计算估计输出频谱 $Y_k(j\omega) = \mathbf{G}_k^T \mathbf{U}_k$;

Step 3: 计算偏差 $e_k(j\omega) = \tilde{Y}_i(j\omega) - Y_k(j\omega)$, 若 $|e_k(j\omega)| \leq \varepsilon$, 则结束辨识, 否则转 Step 4;

Step 4: 计算 NOFRF 向量 \mathbf{G}_{k+1} ;

Step 5: 计算步长 μ_{k+1} , 令 $k = k + 1$, 转 Step 1.

2 基于 SVM 的故障分类器设计

SVM 是一种小样本学习方法, 对于大型机电设备而言, 更多故障样本的获得往往意味着巨大的经济损失, 故本文选择 SVM 构建故障分类器. Suykens^[10] 通过改变标准 SVM 的约束条件和风险函数提出了 LSSVM. LSSVM 的训练只需求解一个线性方程组, 这种算法易于实现, 训练效率高.

采用 NOFRF 的频谱幅值作为特征变量, 设由频谱特征变量构成的训练样本为 $S_1: \{\tilde{\mathbf{G}}_i, y_i\}, i = 1, 2, \dots, M$. 其中, $\tilde{\mathbf{G}}_i \in R^p$ 为模式向量, $y_i \in \{-1, 1\}$ 为类别标号, M 为样本容量. 根据约束优化理论及 KKT 条件, LSSVM 的约束优化问题转换为求解线性方程组问题, 即

$$\begin{bmatrix} 0 & -\mathbf{Y}^T \\ \mathbf{Y} & \Omega + c^{-1} \cdot \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}. \quad (8)$$

其中: $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_M]$, $\alpha_i \in R$ 为拉格朗日乘子; $\mathbf{Y} = [y_1, y_2, \dots, y_M]^T$; \mathbf{I} 为 M 维单位矩阵; $\mathbf{1}$ 为 M 维单位列向量; b 为分类阈值; $c > 0$ 为惩罚因子; Ω 为 M 维对称方阵, $\Omega_{i,j} = y_i y_j K(\tilde{\mathbf{G}}_i, \tilde{\mathbf{G}}_j)$, $K(\cdot, \cdot)$ 为核函数.

求解式(8)可以得到 $\boldsymbol{\alpha}$ 和 b , 根据 LSSVM 构建的非线性系统故障分类器结构如图 2 所示. 其中: $\tilde{\mathbf{G}} = [\tilde{\mathbf{G}}_1, \dots, \tilde{\mathbf{G}}_l]^T$ 为非线性频谱特征向量; $\tilde{\mathbf{G}}_1, \dots, \tilde{\mathbf{G}}_M$ 为支持向量; $\alpha_1, \dots, \alpha_M$ 与 y_1, \dots, y_M 分别表示对应支持向量的拉格朗日乘子与类别标号; \hat{y} 为故障分类结果.

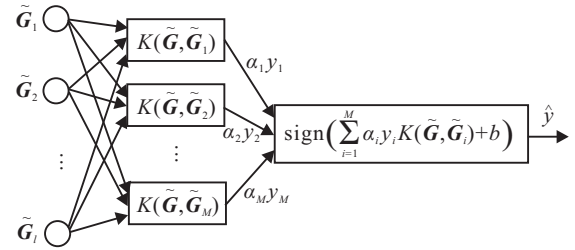


图 2 LSSVM 故障分类器结构

结合 NOFRF 与 LSSVM 的非线性系统故障诊断过程如图 3 所示. 首先实时采集输入输出数据 $u_i(k)_{i=1, \dots, L}$ 和 $y_i(k)_{i=1, \dots, L}$, 对时域数据进行 FFT 变换得到输入输出频谱数据; 然后利用 VSSFRLMS 自适应辨识算法估计 NOFRF 并提取非线性频谱特征变量; 最后将提取的频谱特征输送到 LSSVM 分类器进行分类, 得到故障识别结果.

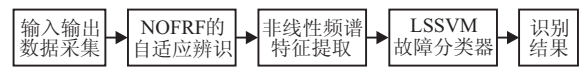


图 3 非线性系统故障诊断结构

3 故障诊断实验

提升机盘式液压制动器的动态模型可以表示为

$$m\ddot{x}(t) + B\dot{x}(t) + k_1x(t) + k_2x^2(t) - u(t) = 0. \quad (9)$$

其中: B 为摩擦阻尼系数, m 为弹簧、连杆与制动盘的总质量, k_1 为线性弹性系数, k_2 为非线性塑性系数, $u(t)$ 为碟型弹簧承受的压力, $x(t)$ 为连杆位移.

在长时间的使用过程中, 碟型弹簧会由于疲劳而损坏, 表现为 k_1 减小, k_2 增大. 某升降机制动器正常状态下参数为 $m = 1$, $B = 15$, $k_1 = 200$, $k_2 = 20$, 选择激励信号为 $u(t) = 50 \cos(0.4\pi t) + 75 \cos(1.6\pi t) + 100 \cos(4\pi t)$, 采用 VSSFRLMS 算法辨识前 3 阶 NOFRF, 采样频率为 12.8 Hz, 每次采样长度为 64, $\alpha = 0.8$, $\beta = 0.05$, $\gamma = 0.001$, $\mu_{\max} = 0.1$, $\mu_{\min} = 0.02$. 正常状态下系统实际输出频谱与由 NOFRF 得到的估计输出频谱如图 4 所示. 由图 4 可以看出, 估计输出频谱与实际输出频谱偏差非常小, 说明 VSSFRLMS 自适应算法的辨识精度较高, 并验证了采用前 3 阶 NOFRF 描述输出频域响应的正确性.

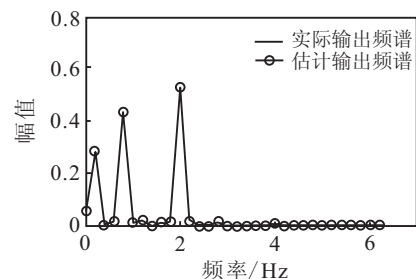


图 4 实际输出频谱与估计输出频谱

输出信噪比为 40 dB 时, 采用标准 LMS 算法与本文算法得到的输出频谱均方误差曲线见图 5, 其中

LMS算法的步长分别为0.02与0.1。由图5可知:步长为0.02时,曲线收敛速度较慢;步长为0.1时,稳态误差较大;本文算法的稳态误差最小且收敛速度最快。

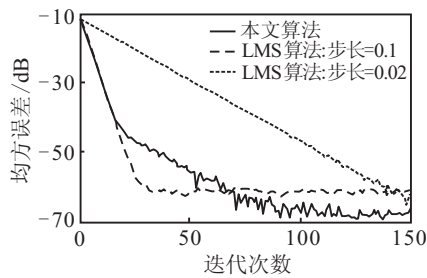


图5 输出频谱均方误差曲线

设故障状态下提升机制动器参数为 $m = 1$, $B = 15$, $k_1 = 190$, $k_2 = 30$, 提升机正常及故障状态下的1阶与2阶NOFRF幅值如表1所示,3阶NOFRF幅值较小,可以忽略。其中, $G1(\cdot)$ 表示1阶NOFRF, $G2(\cdot)$ 表示2阶NOFRF。由表1可以看出,提升机发生故障时,NOFRF幅值发生改变,其中2阶NOFRF幅值显著增大,因此可以通过非线性频谱进行诊断。

表1 非线性频谱数据表

NOFRF	幅 值	
	正常	故障
G1(0.2Hz)	5.05×10^{-3}	5.39×10^{-3}
G1(0.8Hz)	5.22×10^{-3}	5.54×10^{-3}
G1(2Hz)	4.76×10^{-3}	4.79×10^{-3}
G2(0Hz)	2.48×10^{-6}	4.20×10^{-6}
G2(0.4Hz)	1.86×10^{-6}	2.67×10^{-6}
G2(0.6Hz)	2.75×10^{-6}	4.95×10^{-6}
G2(1.2Hz)	2.67×10^{-6}	4.50×10^{-6}
G2(1.6Hz)	3.07×10^{-6}	5.19×10^{-6}
G2(1.8Hz)	2.45×10^{-6}	3.91×10^{-6}
G2(4Hz)	0.78×10^{-6}	1.14×10^{-6}

假定制动器正常时 k_1 服从 [198, 200] 的均匀分布, k_2 服从 [20, 25] 的均匀分布; 故障时 k_1 服从 [185, 195] 的均匀分布, k_2 服从 [30, 35] 的均匀分布。采用蒙特卡洛方法进行故障诊断仿真实验, 计算机CPU为AMD速龙X2-245处理器, 内存为2G。LSSVM采用多项式核函数, 多项式次数 $q = 3$, 惩罚因子 $r = 0.4$, 利用100组正常与故障数据进行训练。训练完成后, 对100组正常与故障数据分别采用本文算法与固定步长LMS自适应算法获取非线性频谱数据进行故障诊断。固定步长为0.02时, 本文算法平均诊断时间为0.22s, 识别率为98%; 固定步长算法的平均诊断时间为0.35s, 识别率为98%。由诊断结果可以看出, 本文算法减少了诊断时间, 识别率高, 能够满足在线故障诊断的要求。

4 结 论

为了解决非线性动态系统的故障诊断问题, 本文对基于非线性频谱数据驱动的故障方法进行了研究。在简化的非线性频谱模型NOFRF基础上提出了一种

自适应辨识算法, 降低了非线性频谱数据求解的计算量。通过LSSVM分类器利用提取的频谱特征进行故障识别。对提升机进行了故障诊断实验, 结果表明所提出的诊断算法所需时间少、识别率高。可见, 基于非线性频谱数据驱动的故障方法可有效地对非线性动态系统进行故障诊断, 具有重要的实用价值。

参考文献(References)

- [1] Majid N A A, Taylor M P, Chen J J J, et al. Aluminium process fault detection by multiway principal component analysis[J]. Control Engineering Practice, 2011, 19(4): 367-379.
- [2] Nguyen V H, Golival J C. Fault detection based on kernel principal component analysis[J]. Engineering Structures, 2010, 32(11): 3683-3691.
- [3] Cui P L, Li J H, Wang J Z. Improved kernel principal component analysis for fault detection[J]. Expert Systems with Applications, 2008, 34(2): 1210-1219.
- [4] Hsu C C, Chen M C, Chen L S. Intelligent ICA-SVM fault detector for non-Gaussian multivariate process monitoring[J]. Expert Systems with Applications, 2010, 37(4): 3264-3273.
- [5] Liu X F, Ma L, Mathew J. Machinery fault diagnosis based on fuzzy measure and fuzzy integral data fusion techniques[J]. Mechanical Systems and Signal Processing, 2009, 23(3): 690-700.
- [6] Tang H, Liao Y H, Cao J Y, et al. Fault diagnosis approach based on Volterra models[J]. Mechanical Systems and Signal Processing, 2010, 24(4): 1099-1113.
- [7] 刘本德, 胡昌华. 基于Volterra频域核辨识的非线性模拟电路故障诊断[J]. 控制与决策, 2009, 24(8): 1167-1171. (Liu B D, Hu C H. Fault diagnosis of nonlinear analog circuits based on Volterra frequency domain kernel identification[J]. Control and Decision, 2009, 24(8): 1167-1171.)
- [8] Lang Z Q, Billings S A. Energy transfer properties of nonlinear systems in the frequency domain[J]. Int J of Control, 2005, 78(5): 345-362.
- [9] 韩海涛, 曹建福, 马红光, 等. 非线性输出频域响应函数的自适应辨识算法及应用[J]. 西安交通大学学报, 2011, 45(10): 77-81. (Han H T, Cao J F, Ma H G, et al. An adaptive identification algorithm of nonlinear output frequency response functions and its application[J]. J of Xi'an Jiaotong University, 2011, 45(10): 77-81.)
- [10] Suykens J A K, Vandewalle. Least square support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.

(责任编辑: 李君玲)