

# 基于 EMD 和 SVDD 的铸钢支座故障诊断

杨永超, 汪同庆

(重庆大学光电技术及系统教育部重点实验室, 重庆 400030)

**摘要:**针对重庆轻轨铸钢支座系统故障诊断中缺乏故障样本的问题,提出一种基于经验模态分解(EMD)和支持向量数据描述(SVDD)的故障诊断方法。对采集到的振动脉冲响应信号进行 EMD 分解,提取第一、第二模态的能量和平均值作为特征输入到 SVDD 分类器进行训练和分类。实验结果表明,采用 EMD 分解后提取的特征能有效地浓缩故障信息,使 SVDD 分类器具有分类效果好、计算效率高等优点。

**关键词:**支持向量数据描述;铸钢支座系统;经验模态分解;故障诊断

## Malfuction Diagnosis of Cast Steel Pedestal Based on EMD and SVDD

YANG Yong-chao, WANG Tong-qing

(Key Laboratory of Optoelectronic Technology & Systems, Ministry of Education, Chongqing University, Chongqing 400030)

**【Abstract】** A method of malfunction diagnosis based on Empirical Mode Decomposition(EMD) and Support Vector Data Description(SVDD) is proposed to solve the problem of lacking malfunction samples in Chongqing light-rail's cast steel pedestal system diagnosis. The vibration impulse response signal is captured and decomposed by EMD, and the first mode's and the second mode's energy and average are extracted as features to input the SVDD classifier for training and classifying. Experimental result indicates that, extracting the features after EMD decomposition can concentrate the malfunction information effectively, which makes the SVDD classifier have well representation and high efficiency.

**【Key words】** Support Vector Data Description(SVDD); cast steel pedestal system; empirical mode decomposition; malfunction diagnosis

### 1 概述

在重庆市轻轨交通系统中,支撑轨道梁的是固定在墩台上的铸钢支座,其服役状态直接影响着轨道梁安全。如果支座的结构或强度受损,如锚固螺杆螺母的松动、支座本身出现裂纹等可能会使轨道梁与墩台的连接失效,引起重大的交通事故。然而,在轻轨正常运行期间,铸钢支座系统出现故障的概率很小,且体积较大、安装位置特殊,其故障样本很难从现场获得,在很大程度上限制了铸钢支座系统的故障诊断方法的研究,目前国内尚属空白。

支持向量数据描述(Support Vector Data Description, SVDD)的方法是一种单值分类算法,由文献[1-2]提出发展起来,其理论基础源于文献[3]提出的支持向量机(Support Vector Machine, SVM)。SVDD 是一种特殊的分类方法,所要解决的问题是将某一目标类数据与其他不属于目标类的数据区分开来<sup>[4]</sup>。应用该方法,仅仅依靠正常运行状态下的数据样本,而不需要故障样本,就可以建立起单值分类器,对机械设备的运行状态进行诊断<sup>[5]</sup>。支持向量数据描述在说话人识别方面已经得到了成功应用。

经验模态分解(Empirical Mode Decomposition, EMD)是文献[4]在 1998 年提出的。该方法是一种自适应的信号处理方法,被认为是近年来对以傅里叶变换为基础的线性、稳态频谱分析的一个重大突破,且具有自适应的信号分解和降噪能力,目前已成功应用于图像处理 and 旋转机械故障诊断<sup>[6]</sup>等领域。

### 2 EMD 方法及振动信号研究

EMD 的本质是将一个时间序列的信号分解成不同尺度

的本征模函数(Intrinsic Mode Function, IMF)。所谓本征模函数,必须满足 2 个条件:(1)对于一系列数据,极值点和过零点数目必须相等或至多相差一点;(2)在任意点,分别由局部极大点和局部极小点构成的上下包络线的平均值为零。每一个 IMF 可以用 Hilbert 变换来求取瞬时频率。根据 EMD 原理,它实现的过程就是一个逐层分解的过程,形象地称之为“筛”的过程,将高频成分逐步“筛”去,直到只剩下序列的趋势或均值。由于 EMD 方法在“筛”的过程中构成上下包络线的 3 次样条拟合函数在数据序列的两端的拟合点是不确定的,因此会出现发散现象,并随着“筛”的过程的逐渐向内“污染”整个数据序列,使结果失真,也就是出现通常所说的端点效应<sup>[7]</sup>。

根据 EMD 算法,以铸钢支座系统 1 号锚固螺杆在 500 N·m 的紧固力矩下的脉冲响应信号对其进行 4 阶经验模态分解,结果如图 1 所示。由图可知,第一模态包含了原信号中的主频分量,第二模态包含了原信号幅频中的共振频率分量,IMF3、IMF4 等模态分量因为包含的信息量相对较少可以忽略不计。这样可以简化所要处理的信号,在分解的过程中有效地滤除噪声和干扰,克服在处理过程中遇到的 FFT 混叠、泄漏、频谱分辨率不高等可能出现的问题,达到了信号分析消除信号中多余成分、去除噪声和干扰、将信号转换成

**基金项目:**国家科技支撑计划基金资助项目(2007BAG06B06);重庆大学自然科学基金资助项目

**作者简介:**杨永超(1981-),男,硕士研究生,主研方向:计算机测控技术,信号处理;汪同庆,教授、博士生导师

**收稿日期:**2009-04-30 **E-mail:** yangyongchao20@163.com

符合要求的形式的目的。

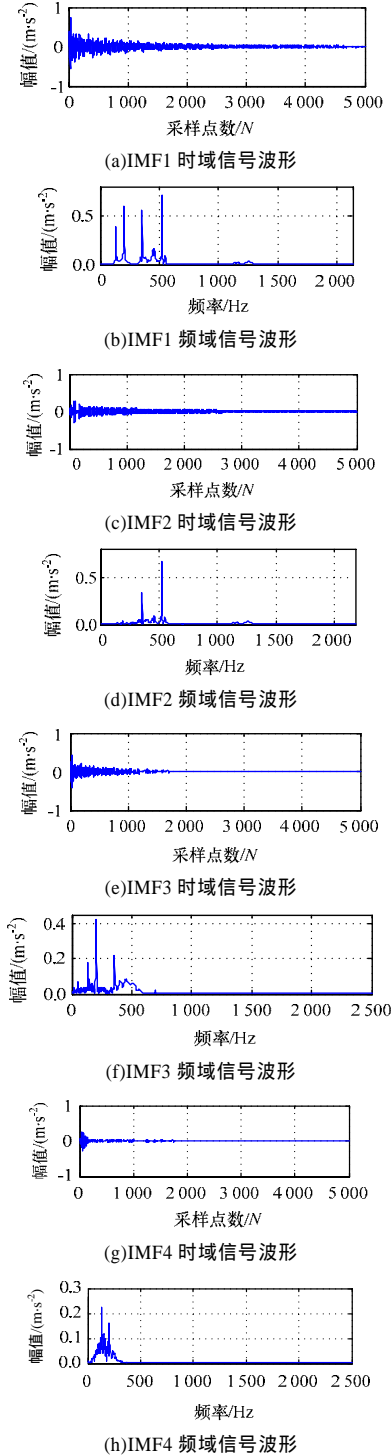


图1 脉冲响应信号的 EMD 分解

### 3 支持向量数据描述算法

SVDD 的基本思想在于建立一个封闭紧凑的超球体，将目标类对象全部或尽可能多地包含在该超球体内，而非目标类对象没有或尽可能少地包含在该超球体内，其目标是要构建单值分类器，形成由中心为  $a$ 、半径为  $R$  的超球体。为了增强其分类的鲁棒性，引入松弛因子  $\xi_i$ ，则超球体应满足

$$\min f(R, a, \xi) = R^2 + C \sum_{i=1}^N \xi_i, \quad i = 1, 2, \dots, N$$

$$\text{约束条件: } \|x_i - a\|^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0 \quad (1)$$

其中， $C$  为常数，起控制对错样本惩罚程度的作用，以实

现在错样本的比例和算法复杂程度之间的折中。

将式(1)转化为 Lagrange 极值问题：

$$f(R, a, \xi, \alpha, \lambda) = \sum_{i=1}^N \alpha_i (x_i \cdot x_i) - \sum_{i=1, j=1}^N \alpha_i \alpha_j (x_i \cdot x_j) \quad (2)$$

其中， $0 \leq \alpha_i \leq C$ ， $\lambda_i \geq 0$  为 Lagrange 系数，且  $\sum_{i=1}^N \alpha_i = 1$ 。

引入高斯径向基核函数：

$$K(x, y) = \exp\left[-\frac{\|x - y\|^2}{2\sigma^2}\right] \quad (3)$$

其中，给定不同的  $\sigma$  对于同一样本集可能产生不同的超球体边界，为达到满意的分类效果，还需要调整  $\sigma$  到合适的值。

引入核函数  $K(x_i, x_j)$  后，可以使低维空间的非线性问题转化为高维空间的线性问题，那么，式(3)变为如下形式：

$$f(R, a, \xi, \alpha, \lambda) = \sum_{i=1}^N \alpha_i K(x_i, x_i) - \sum_{i=1, j=1}^N \alpha_i \alpha_j K(x_i, x_j) \quad (4)$$

在实际计算中，多数的  $\alpha_i$  将为 0，在计算中将被忽略。其不为 0 的  $\alpha_i$  所对应的  $x_i$  称之为支持向量，这部分支持向量决定了  $a$  和  $R$  的值。其中，

$$R^2 = (x_k \cdot x_k) - 2 \sum_{i=1}^N \alpha_i (x_i \cdot x_k) + \sum_{i=1, j=1}^N \alpha_i \alpha_j (x_i \cdot x_j) \quad (5)$$

判断一个新样本  $z$  是否属于目标样本集，如果决策函数

$$R_z = \|z - a\|^2 = K(z, z) - 2 \sum_{i=1}^N \alpha_i K(z, x_i) + \sum_{i=1, j=1}^N \alpha_i \alpha_j K(x_i, x_j) \leq R^2 \quad (6)$$

成立，则  $z$  为目标样本；否则， $z$  为非目标样本。

从以上可以看出，SVDD 算法解决了单值分类的问题且不使用类标，惩罚参数  $C$  作为对引入异常点的惩罚，使得在训练集包含噪声的情况下 SVDD 算法仍能工作。

## 4 实验研究

### 4.1 特征提取

铸钢支座系统的下端由 4 根锚固螺杆固定在轨道墩台上面，其上端与轨道梁连接在一起并支撑着轨道梁，在正常情况下，每根锚固螺杆的紧固力矩为 800 N·m。

通过大量实验对不同紧固力矩下 IMF1 及 IMF2 等模态的分析验证发现，随着紧固力矩的增加，支座系统的接触刚度增大，系统的固有频率会逐渐升高，同时引起 IMF1 和 IMF2 的模平均值和能量逐渐增大。也就是说 EMD 分解后 IMF1 及 IMF2 等模态的模态能量与模平均值和紧固力矩之间存在较为线性的变化规律，根据这种变化规律可以进行铸钢支座系统中锚固螺杆松动故障的检测。图 2 为经验模态分解后的 IMF1, IMF2 的模态能量与紧固力矩之间的关系。

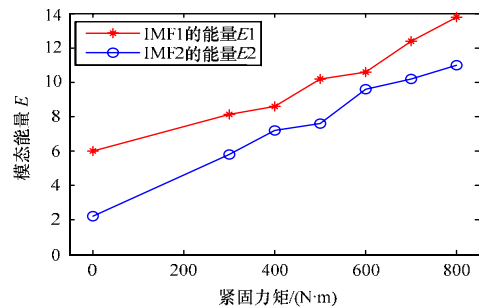


图2 IMF1, IMF2 的模态能量与紧固力矩之间的关系

在实验中，对铸钢支座系统的脉冲响应信号进行 EMD 分解后，选择 IMF1 和 IMF2 的模态能量、IMF1 和 IMF2 的模态平均值等 4 个特征进行提取。以这 4 个特征组合而成的特征集为铸钢支座系统故障诊断的依据，并作为 SVDD 分类

器的训练样本。

模态能量  $E_m$  和平均值  $T_m$  的计算如下：

$$E_m = \sum IMF(m)_{ij}^2, \quad m=1,2, i, j=0,1, \dots, N-1$$

$$T_m = \frac{1}{N} \sum |IMF(m)_{ij}|, \quad m=1,2, i, j=0,1, \dots, N-1 \quad (7)$$

其中,  $IMF(m)_{ij}$  表示 IMF 模态矩阵值;  $N$  为数据长度。

#### 4.2 故障诊断

现改变 1 号锚固螺栓的紧固力矩来模拟铸钢支座的故障情况。从试验台上采集铸钢支座系统在正常和故障 2 种工况下的振动脉冲响应信号, 数据长度为 5 000。

故障工况: 1 号锚固螺栓的紧固力矩分别为 0 N·m, 300 N·m, 500 N·m, 共采集 15 组。

正常工况: 4 个锚固螺栓的紧固力矩均为 800 N·m, 共采集 60 组, 其中 30 组用于训练分类器, 另外 30 组用于测试分类器效果。

把用于测试分类器效果的正常工况的 30 组和故障工况下的 15 组共计 45 组测试样本进行编号, 其中, 前 30 组正常样本为 1~30, 后 15 组故障样本为 31~45。

首先不进行特征提取, 将正常工况下的 30 组数据直接输入到 SVDD 分类器进行训练, 建立单值分类器, 并计算出对应的  $\alpha_i, a$  和  $R$ 。然后把 45 组待测样本输入到分类器进行验证, 根据式(6)分别计算各个测试样本  $z$  到超球体中心  $a$  的距离。由于特征数据为多维, 不能画出其特征分布图, 因此引入参数  $\varepsilon_k = R_k - R$ ,  $R_k$  表示第  $k$  个样本  $x_k$  到超球体中心的距离,  $R$  表示分类器所形成的超球体的半径, 以距离分布图来观察其分类效果。如果  $\varepsilon_k > 0$ , 表示  $x_k$  属于非目标样本, 即为故障样本; 反之  $\varepsilon_k \leq 0$ , 则表示  $x_k$  属于目标样本, 即为非故障样本。对上述的 45 个测试样本进行分类以后, 根据  $\varepsilon_k$  的计算结果画出的距离分布图如图 3 所示。

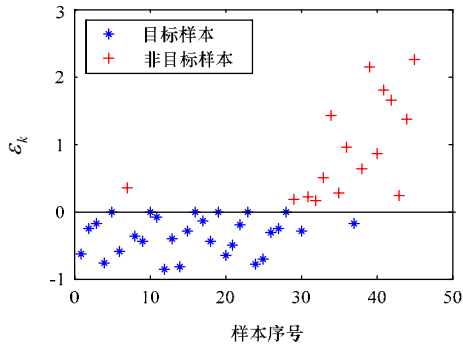


图 3 未进行特征提取的距离分布图

然后对训练样本和测试样本进行 EMD 分解并进行特征提取后, 输入到 SVDD 分类器中, 其分类效果如图 4 所示。

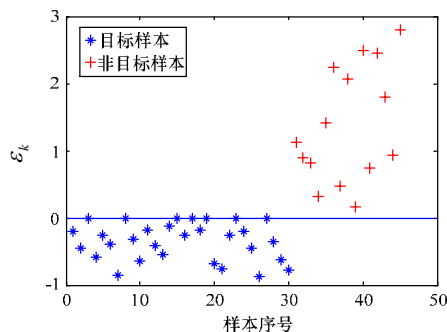


图 4 进行特征提取后的距离分布图

从图 4 可以看出, 如果不对数据样本进行特征提取而直接输入到分类器时, 由于数据信号中含有较大的噪声或干扰, 分类效果很差。一般来说, 分类误差有 2 种, 一种是目标样本被拒绝, 另一种是非目标样本被接受。通常, 第 2 种分类误差带来的风险更大。把第 7 号和第 29 号目标样本误判为非目标样本, 这对实际情况尚可接受, 而把第 37 号非目标样本误判为目标样本, 这在实际检测中是不允许发生的, 可能会引起重大交通事故。从图 4 可以看出, 经过特征提取以后, 分类器可以很好地将目标样本和非目标样本分开, 实现正常样本和故障样本的准确识别。这说明对数据样本进行 EMD 分解并提取特征后, 可以去掉信号中所含的噪声和干扰, 能够有效地浓缩样本的故障信息, 使得分类器可以进一步提高分类效果。

从上述可以看出, 在训练分类器时, 只需要输入正常状态时的数据样本, 而不需要任何故障样本。在实际的机械设备检测中, 只需要采集设备完好或者故障前的数据训练分类器, 然后检测时将待测设备的数据输入分类器, 就可以根据分类情况来判定是否存在故障。

#### 4.3 与其他分类方法的比较

常用于机械故障诊断的单值分类方法还有支持向量机 (SVM) 法、BP 神经网络法等。将上述相同的 30 组训练样本和 45 组测试样本进行 EMD 特征提取后, 分别输入到 SVM 和 BP 神经网络, 得到其分类结果和计算时间如表 1 所示。

算法	诊断结果		计算时间/s
	正常	故障	
SVDD	30	15	3.728
SVM	27	18	3.334
BP 神经网络	36	9	26.582

由表 1 可知, SVDD 的分类结果最好, SVM 次之, BP 神经网络最差, 且有漏检情况, 这在实际检测时是不允许存在的。在计算时间上面, SVDD 和 SVM 均不超过 4 s, 虽然 SVDD 比 SVM 时间稍长, 但不会影响检测效率, 而 BP 神经网络太耗时间, 在待测样本比较多时, 效率是相当低的。因此, SVDD 算法在轻轨铸钢支座系统故障诊断中的总体效果最优。

#### 5 结束语

在机械设备系统的故障诊断中, 正常样本很容易获取, 但故障样本一般很难、或者要花费很大代价才能够获得, 如果把支持向量数据描述的方法应用于其中, 仅依靠系统正常运行时的数据样本就能够对设备的运行状态进行监测和故障诊断。本文将 SVDD 和 EMD 方法相结合, 应用于重庆轻轨轨道梁铸钢支座系统的故障诊断, 有效地解决了因故障样本难以获取继而难以进行诊断的问题。与其他方法相比, 该方法具有计算效率高、分类效果好等优点, 如果对该方法进一步研究, 缩短计算时间, 完全可以满足在线诊断的要求。因此, 该方法在机械故障诊断领域具有非常重要的推广价值和广泛的应用前景。

#### 参考文献

- [1] Tax D M J, Duin R P W. Support Vector Domain Description[J]. Pattern Recognition Letters, 1999, 20(11-13): 1191-1199.
- [2] Tax D M J, Duin R P W. Support Vector Data Description[J]. Machine Learning, 2004, 54(1): 45-66.

(下转第 250 页)