

文章编号: 1001-0920(2007)07-0778-05

基于最小二乘支持向量机的汽轮机故障诊断

杨奎河^{1,2}, 单甘霖¹, 赵玲玲²

(1. 军械工程学院 光学与电子工程系, 石家庄 050003; 2. 河北科技大学 信息科学与工程学院, 石家庄 050054)

摘要: 提出一种小波包分析与最小二乘支持向量机相结合的汽轮机故障诊断模型. 对故障信号功率谱进行小波分解, 简化了故障特征向量的提取. 用二次损失函数取代支持向量机中的不敏感损失函数, 将不等式约束条件变为等式约束, 从而将二次规划问题转变为线性方程组的求解. 选用 RBF 函数作为核函数, 并提出对核函数的参数进行动态选取, 提高了诊断的准确率. 仿真结果表明该模型具有较强的非线性处理和抗干扰能力.

关键词: 小波包分析; 故障诊断; 最小二乘支持向量机; 概率神经网络; 汽轮机

中图分类号: TP391

文献标识码: A

Steam turbine fault diagnosis based on least squares support vector machine

YANG Kui-he^{1,2}, SHAN Gan-lin¹, ZHAO Ling-ling²

(1. Department of Optics and Electron Engineering, Ordnance Engineering College, Shijiazhuang 050003, China; 2. College of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050054, China. Correspondent: YANG Kui-he, E-mail: ykh@hebust.edu.cn)

Abstract: A steam turbine fault diagnosis model, in which wavelet packet analysis and least squares support vector machine (LSSVM) are combined effectively, is proposed. The power spectrums of fault signals are decomposed by using wavelet analysis, which simplifies choosing method of fault eigenvectors. The non-sensitive loss function is replaced by quadratic loss function and the inequality constraints are replaced by equality constraints. Consequently, quadratic programming problem is simplified as the problem of solving linear equation groups. RBF function is used as kernel function, and it is proposed to choose the parameter of kernel function dynamically, which enhances the preciseness rate of diagnosis. The simulation results show that the model has strong non-linear solution and anti-jamming ability.

Key words: Wavelet packet analysis; Fault diagnosis; Least squares support vector machine; Probabilistic neural networks; Steam turbine

1 引言

近年来,神经网络在汽轮机故障诊断中的应用研究取得了许多成果,但它在网络结构选择、网络训练以及提高网络的泛化能力等方面尚存在许多问题,限制了它在工程实际中的广泛应用.支持向量机(Support Vector Machine, SVM)是由 Vapnik 等人在统计学习理论的基础上建立起来的一种机器学习方法^[1].SVM 通过结构风险最小化原理来提高泛化能力,较好地解决了小样本、非线性、高维数、局部极小等实际问题,已在模式分类、回归预测、概率估计以及控制理论等领域得到应用^[2-4].最小二乘支持

向量机(Least Squares Support Vector Machine, LSSVM)是基于 SVM 的一种改进算法^[5-7],它是 SVM 在二次损失函数下的一种形式,用二次损失函数取代 SVM 中的不敏感损失函数,通过构造损失函数将原 SVM 中算法的二次寻优变为求解线性方程,简化了计算的复杂性.

本文将小波包分析与 LSSVM 进行了有机结合,应用小波变换频带分析技术进行设备故障的特征向量提取,并采用 LSSVM 进行汽轮机常见故障的诊断,取得了较好的诊断效果.

2 小波频带分析和特征向量提取

收稿日期: 2006-04-29; 修回日期: 2006-08-07.

基金项目: 国家自然科学基金项目(60572062); 中国博士后科学基金项目(2005038515).

作者简介: 杨奎河(1966—),男,河北邯郸人,教授,博士后,从事智能控制技术和故障诊断的研究;单甘霖(1962—),男,石家庄人,教授,博士生导师,从事装备故障诊断的研究.

表 1 故障原因与征兆表

故障样本	0	1	2	3	4	5	6	7	8
频段	0.01 ~ 0.39f	0.40 ~ 0.49f	0.50f	0.51 ~ 0.99f	f	2f	3 ~ 5f	奇数倍 f	高频 > 5f
不平衡	0.00	0.00	0.00	0.00	0.90	0.05	0.05	0.00	0.00
气动力偶	0.00	0.30	0.10	0.60	0.00	0.00	0.00	0.00	0.00
不对中	0.00	0.00	0.00	0.00	0.40	0.50	0.10	0.00	0.00
油膜涡动	0.10	0.80	0.00	0.10	0.00	0.00	0.00	0.00	0.00
径向碰摩	0.10	0.10	0.10	0.10	0.20	0.10	0.10	0.10	0.10
共生松动	0.00	0.00	0.00	0.00	0.20	0.15	0.40	0.00	0.25

小波分析的多分辨率分析是将信号投影到一组互相正交的小波函数构成的子空间上,形成信号在不同尺度上的展开,从而提取信号在不同频带的特征,同时保留信号在各尺度上的原有特征.但多分辨率分析只对信号的低频部分进行分解,高频部分保留不动.小波包分析是对小波变换的一种改进,它能够为信号提供一种更加精细的分析方法.它可以将小波变换中没有细分的高频部分作进一步分解,从而提高信号通频带的频率分辨率.因此,小波包分解可对信号在全频带内进行正交分解,同时可以在低频和低频部分进行分解,自适应地确定信号在不同频段上的分辨率.用小波包分析技术,可以将信号分解到任意精细的频带上,在这些频带上进行能量统计,形成特征向量,更趋合理性.

应用小波包分析技术可以将信号分解到不同的频带之内,对这些频带的信号进行分析,称为频带分析技术.通常可根据振动信号的频率范围,将信号在一定尺度上进行分解,从而提取相应频带内的信息.此外,还可以对各频带内的信号进行统计分析,形成反映信号特征的特征向量.若分析各频带内的信号能量,则称之为频带的能量分析.

与 Fourier 频谱分析技术相同,小波频带技术的理论依据也是 Parseval 能量积分等式.

因在 $f(x)$ 时域上的能量为

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx, \quad (1)$$

$f(x)$ 的小波变换为

$$C_{j,k} = W(2^j, 2^j k) = \int_{-\infty}^{+\infty} 2^{-j/2} \overline{w(2^{-j}x - k)} f(x) dx, \quad (2)$$

二者由 Parseval 恒等式关联得

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \sum_{k=1}^n |C_{j,k}|^2, \quad (3)$$

其中 j, k 和 n 分别为小波包的尺度指标、位置指标和频率指标.由上式可知,小波变换系数 $C_{j,k}$ 具有能量的量纲,因此可用于能量分析.

电力系统设备故障诊断技术中对汽轮机等旋转

机械故障诊断的研究是很深入和完善的.将多分辨率分析应用到功率谱的特征提取上,可以方便而有效地提取出特征向量.运用小波包分析技术,可对旋转机械常见的故障进行能量分析.经过大量的实验,可以建立起汽轮机故障原因与征兆对应表^[8].用汽轮机常见的不平衡、不对中、油膜涡动等 6 种常见故障作为诊断模型的输出,利用振动信号频谱的 9 个频段上不同频率谱的谱峰能量值作为特征量,形成汽轮机振动故障的训练样本 (f 为工频),如表 1 所示.

3 LSSVM 诊断模型

3.1 SVM 原理

SVM 是从线性可分情况下的最优分类面发展而来的.设有 N 个样本 x_i 及其所属类别 y_i ,表示为 $\{(x_i, y_i)\}$.其中: $x \in R^d; y_i \in \{1, -1\}; i = 1, 2, \dots, N; d$ 为输入空间的维数.对于标准的 SVM,其分类间隔为 $2/\|w\|$,使分类间隔最大相当于 $\|w\|^2$ 最小.因此,使分类间隔最大的优化问题可表示为如下的二次规划问题:

$$\min_{w,b} J(w, b) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i; \quad (4)$$

$$\text{s.t. } y_i [w^T(x_i) + b] - 1 - \xi_i, \quad (5)$$

$$\xi_i \geq 0, i = 1, 2, \dots, N. \quad (6)$$

其中: $\xi_i \geq 0 (i = 1, 2, \dots, N)$ 为松弛因子,用来保证在线性不可分情况下分类的正确性;惩罚因子 C 为某个指定的常数,用来控制对错样本的惩罚程度,实现在错样本的比例与算法复杂程度之间的折衷; σ 为非线性变换函数,可通过 σ 将非线性问题转化为某个高维空间中的线性问题,在变换空间求最优分类面.

设有非线性映射 $\sigma: R^d \rightarrow H$ 将输入空间的样本映射到高维的特征空间 H 中.当在特征空间 H 中构造最优超平面时,训练算法仅使用空间中的内积,即 $(x_i) \cdot (x_j)$,而没有单独的 (x_i) 出现.因此,如果能够找到一个函数 K 使得

$$K(x_i, x_j) = (x_i) \cdot (x_j),$$

则高维空间的内积运算就可以用原空间中的函数实现,甚至可不必知道变换的形式.统计学习理论指出,根据 Hilbert-Schmidt 原理,只要一个函数 $K(x_i, x_j)$ 满足 Mercer 条件,那么它就对应某一变换空间中的内积.

SVM 算法就是通过求解上述二次规划问题来实现对样本的正确分类.

3.2 LSSVM 分类算法

自从 Vapnik 等人提出 SVM 算法以来,已经出现了 SVMlight, Smooth SVM 和 LSSVM 等多种改进算法.其中 LSSVM 中的优化指标采用了平方项,只有等式约束,而没有标准 SVM 的不等式约束,从而推出了不同的一系列等式约束,而不是二次规划问题.这样,将二次规划问题转变成线性方程组的求解,简化了计算复杂性.对于 LSSVM,其寻优目标函数式应变为

$$\min_{w,b} J_{LS}(w, b) = \frac{1}{2} \|w\|^2 + \frac{1}{2} \sum_{i=1}^N \xi_i^2; \quad (7)$$

$$\text{s.t. } y_i [w^T(x_i) + b] = 1 - \xi_i, \quad \xi_i = 1, 2, \dots, N. \quad (8)$$

其中参数 ξ_i 是类似于 SVM 中的参数 C ,用于对 $J_{LS}(w, b)$ 进行控制.为求解这个问题的优化问题,引入如下拉格朗日函数:

$$L_{LS}(w, b, \xi) = \frac{1}{2} \|w\|^2 + \frac{1}{2} \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \lambda_i \{y_i [w^T(x_i) + b] - 1 + \xi_i\}, \quad (9)$$

其中 λ_i 为 Lagrangian 乘子,可正可负.在鞍点(极值)处将上式分别对 $w, b,$ 和 λ_i 求导,并令它们等于零,得

$$\frac{\partial L_{LS}}{\partial w} = 0, \quad w = \sum_{i=1}^N \lambda_i y_i (x_i); \quad (10)$$

$$\frac{\partial L_{LS}}{\partial b} = 0, \quad \sum_{i=1}^N \lambda_i y_i = 0; \quad (11)$$

$$\frac{\partial L_{LS}}{\partial \lambda_i} = 0, \quad \lambda_i = \xi_i; \quad (12)$$

$$\frac{\partial L_{LS}}{\partial \xi_i} = 0, \quad y_i [w^T(x_i) + b] - 1 + \xi_i = 0. \quad (13)$$

上述条件可写成一个线性系统,即

$$\begin{bmatrix} I & 0 & 0 & -Z^T \\ 0 & 0 & 0 & -Y^T \\ 0 & 0 & I & -I \\ Z & Y & I & 0 \end{bmatrix} \begin{bmatrix} w \\ b \\ \xi \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (14)$$

其中

$$Z = [(x_1)^T y_1, \dots, (x_N)^T y_N]^T,$$

$$Y = [y_1, \dots, y_N]^T, \quad 1 = [1, \dots, 1]^T,$$

$\lambda = [\lambda_1, \dots, \lambda_N]^T, \quad \xi = [\xi_1, \dots, \xi_N]^T,$
 $I \in R^{(N \times N)}$ 为单位矩阵.消除 w 和 λ 后,式(14)简化为

$$\begin{bmatrix} 0 & Y^T \\ Y & ZZ^T + I \end{bmatrix} \begin{bmatrix} b \\ \xi \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (15)$$

定义 $Q = ZZ^T \in [q_{ij}]_{N \times N}$,并对 Q 矩阵应用 Mercer 条件,则该矩阵的元素可表示为如下形式:

$$q_{i,j} = y_i y_j (x_i)^T (x_j) = y_i y_j K(x_i, x_j), \quad (16)$$

其中 $K(x_i, x_j)$ 为核函数.常用的核函数有:

1) 多项式核函数

$$K(x_i, x_j) = ((x_i \cdot x_j) + c)^d, \quad d = 1, 2, \dots;$$

2) RBF(Radial Basis Function) 核函数

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right);$$

3) Sigmoid 核函数

$$K(x_i, x_j) = \tanh((x_i \cdot x_j) + c).$$

用最小二乘法即可求解式(15).在 LSSVM 算法中,将二次规划问题转变为线性方程组的求解.求解上述问题后得到的最优分类函数为

$$y(x) = \text{sign}\left[\sum_{i=1}^N \lambda_i y_i (x, x_i) + b\right]. \quad (17)$$

3.3 LSSVM 中参数选取算法的改进

RBF 核函数可以将样本非线性地映射到更高维的空间,与多项式核函数相比只有较少的参数参与运算,减少了模型选择的复杂性;而 Sigmoid 核函数在取某些参数值时则可能无效.因此,本文采用 RBF 函数作为核函数.在 LSSVM 算法中,规则化参数 σ 和 RBF 核函数的标准化参数 c 一般都根据经验选取一个固定的值,但针对不同的样本集,最优的参数值是变化的,因此在一定程度上影响了故障诊断的正确率.经用 UCI 标准分类数据库数据实验发现,规则化参数 σ 的取值对计算的结果影响不明显,而随着 c 的取值不同,计算的结果波动较大,因此在计算时对 c 进行动态选取.

识别出错率 $y = f(c)$ 受参数 c 变化影响情况如图 1 所示.实验结果表明,在参数 σ 固定的情况下,识

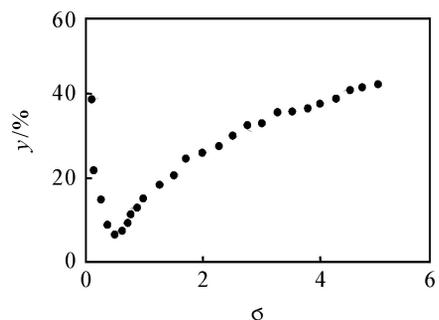
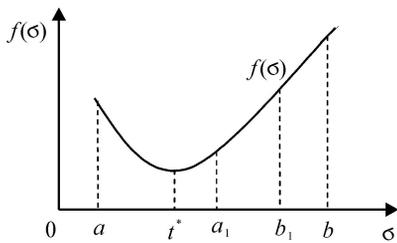


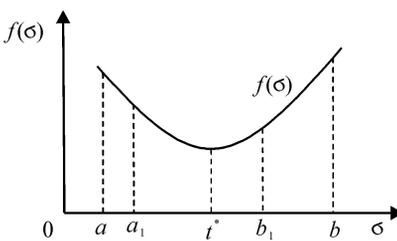
图 1 参数变化影响实验结果

别出错率 y 每次得到的结果都有一定的随机性,并不能用一个简单的函数式表示.

实验结果也显示出,虽然参数 σ 在某一点得到的识别出错率有一定的随机性,但振幅度很小.因此,在一个相对较大区间范围内,识别出错率 $y = f(\sigma)$ 可近似看作在较大区间 $[a, b]$ 上的下单峰函数,在此区间上 t^* 处及附近较小邻域范围内函数 y 可取得极小值.若在区间 $[a, b]$ 内取任两点 a_1 和 b_1 , $a_1 < b_1$,并计算函数值 $f(a_1)$ 和 $f(b_1)$,可能出现以下两种情况(如图 2 所示): $f(a_1) < f(b_1)$,这时极小点 t^* 必在区间 $[a, b_1]$ 内; $f(a_1) > f(b_1)$,这时极小点 t^* 必在区间 $[a_1, b]$ 内.因此,只要在区间 $[a, b]$ 内取 2 个不同点,并算出它们的函数值加以比较,就可以将搜索区间 $[a, b]$ 缩小成 $[a, b_1]$ 或 $[a_1, b]$,缩小后的区间仍包含极小点.如果要继续缩小搜索区间 $[a, b_1]$ 或 $[a_1, b]$,就需要在上述区间内再取出 1 点算出其函数值,并与 $f(a_1)$ 或 $f(b_1)$ 加以比较即可.只要缩小后的区间包含点 t^* 及附近较小邻域,则区间缩小得越小, y 的取值就越接近函数的极小点取值区域.但是,当 y 的取值进入 t^* 处附近较小邻域范围时,出错率 y 会出现振荡.因此,可以在 y 第 1 次振荡时停止区间搜索,并取振荡前一次的 y 的取值作为 σ 的值.为使算法计算简单,本文在具体计算时取 $a = 0.1, b = 5, a_1$ 和 b_1 取在 $[a, b]$ 的三等分处.



(a) $f(a_1) < f(b_1)$ 时的极小点位置



(b) $f(a_1) > f(b_1)$ 时的极小点位置

图 2 搜索区间

4 仿真结果

为了验证 LSSVM 诊断模型对汽轮机常见故障的诊断准确率,分别用两种诊断模型进行仿真计算.模型 1 采用概率神经网络(Probabilistic Neural Networks, PNN)模型^[9,10].PNN 是由径向基函数网络发展而来的一种前馈型神经网络,其理论依据

是贝叶斯决策理论中的贝叶斯最小风险准则,可用来解决分类和故障诊断问题.PNN 由输入层、径向基层和竞争层组成.输入层设 9 个输入节点对应训练样本中振动信号频谱的 9 个频段,网络的隐层神经元个数与输入样本矢量的个数相同,该网络的竞争层是输出层,其神经元个数等于训练样本数据的种类个数,在本模型中取 6 来对应 6 种故障类型.当训练样本足够多时,PNN 收敛于一个贝叶斯分类器.

模型 2 采用 LSSVM 故障诊断模型,用 One-Against-One 方法实现多类分类^[11].由于规则化参数 γ 取值变化对计算结果影响不明显,设参数 γ 为定值,取 $\gamma = 10$.为了比较 RBF 核函数的标准化参数 σ 动态选取的效果,在诊断过程中对参数 σ 分别采用动态选取和取不同定值的方法进行诊断.

采用小波包分析技术对汽轮机振动信号进行能量分析,提取出故障特征向量作为样本集.设 D_1 为加入噪声前的故障样本数据矩阵, D_2 为加入噪声后的故障样本数据矩阵,仿真需要的含有噪声的样本数据矩阵元素通过下式获得:

$$D_2(i, j) = D_1(i, j) \times (1 + \gamma \times \text{rands}(1)). \quad (18)$$

其中: γ 为噪声控制系数,分别取 $\gamma = 0, 0.2, 0.4, 0.6$; $\text{rands}(1)$ 为随机函数,可生成一个 -1 到 1 之间的随机函数.利用上式对每种故障情况分别产生 160 组带噪声的测量参数,共 960 组样本,其中 600 组作为训练集,360 组作为测试集.用 PNN 及 LSSVM 进行诊断,诊断统计结果如表 2 所示.

表 2 故障诊断准确率 %

噪声控制系数	PNN 模型	LSSVM 模型			
		$\sigma = 0.1$	$\sigma = 1$	$\sigma = 2$	动态
0.0	100	100	100	99.7	100
0.2	98.1	98.1	97.2	95.3	98.3
0.4	92.2	94.4	93.3	91.9	95.6
0.6	87.8	92.2	89.7	86.4	93.6

从表 2 可以看出,诊断准确率受噪声控制系数的影响,随着样本噪声的增加,两种模型的诊断准确率都出现下降,但 LSSVM 模型的下降速度明显低于 PNN 模型,显示出较强的鲁棒性;在 LSSVM 模型中,动态选取参数 σ 时具有较高的诊断准确率.

5 结 论

本文通过对汽轮机故障信号功率谱进行小波包分解,简化了故障特征向量的提取.仿真结果表明,LSSVM 模型的故障诊断能力优于 PNN 模型,显示了 LSSVM 较强的诊断鲁棒性.与 PNN 相比较,

SVM 模型是专门针对小样本情况的,其目标是得到现有信息下的最优解而不仅仅是样本数趋于无穷大时的最优值,得到的是全局最优点,解决了神经网络无法避免的局部极值问题.LSSVM 则对 SVM 进行改进,用二次损失函数取代 SVM 中的不敏感损失函数,将不等式约束条件变为等式约束,将二次规划问题转变为线性方程组的求解,用最小二乘法实现了 SVM 算法,简化了计算的复杂性,核函数参数的动态选取进一步提高了诊断的正确率。

参考文献(References)

- [1] Vapnik V, Golowich S, Smola A. Support vector method for function approximation, regression estimation and signal processing[J]. Advances in Neural Information Processing Systems, 1996, 9(2): 281-287.
- [2] Chen S, Samangan A K, Hanzo L. Support vector machine multiuser receiver for DS-CDMA signals in multipath channels [J]. IEEE Trans on Neural Networks, 2001, 12(3): 604-611.
- [3] Sebald D J, Bucklew J A. Support vector machine techniques for nonlinear equalization[J]. IEEE Trans on Signal Processing, 2000, 48(11): 3217-3226.
- [4] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Trans on Geoscience Remote Sensing, 2004, 42(8): 1778-1790.
- [5] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers [J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [6] Lukas L, Suykens J A K, Vandewalle J. Least squares support vector machines classifiers: A multi two-spiral benchmark problem[C]. Proc of the Indonesian Student Scientific Meeting. Manchester, 2001: 289-292.
- [7] Suykens J A K, Lukas L, Vandewalle J. Sparse least squares support vector machine classifiers[C]. Proc of the European Symposium on Artificial Neural Networks. Bruges, 2000: 37-42.
- [8] 虞和济, 陈长征, 张省, 等. 基于神经网络的智能诊断[M]. 北京: 冶金工业出版社, 2002.
(Yu He-ji, Chen Chang-zheng, Zhang Xing, et al. Intelligent diagnosis based on neural networks [M]. Beijing: Metallurgy Industry Press, 2002.)
- [9] 叶志锋, 孙健国. 基于概率神经网络的发动机故障诊断[J]. 航空学报, 2002, 23(2): 155-157.
(Ye Zhi-feng, Sun Jian-guo. Probabilistic neural networks based engine fault diagnosis [J]. Aviation Transaction, 2002, 23(2): 155-157.)
- [10] 李冬辉, 刘浩. 基于概率神经网络的故障诊断及应用[J]. 系统工程与电子技术, 2004, 26(7): 997-999.
(Li Dong-hui, Liu Hao. Method and application of fault diagnosis based on probabilistic neural network [J]. Systems Engineering and Electronics, 2004, 26(7): 997-999.)
- [11] Hsu C W, Lin C J. A comparison of methods for multiclass support vector support vector machines[J]. IEEE Trans on Neural Networks, 2002, 13(2): 415-425.

(上接第 777 页)

参考文献(References)

- [1] Shozo M, Nishihara H, Furuta K. Control of unstable mechanical system: Control of pendulum [J]. J of Control, 1976, 23(5): 673-692.
- [2] Awtar S, King N, Allen T. Inverted pendulum systems: Rotary and arm-driven — A mechatronic system design case study[J]. Mechatronics, 2002, 12(3): 357-370.
- [3] 王永, 孙颖娥, 周烽. 多级旋转式倒立摆的建模与分析[J]. 电机与控制学报, 2004, 8(3): 258-263.
(Wang Yong, Sun Ying-e, Zhou Feng. Modeling and analysis of the multiple rotational inverted pendulum [J]. Electric Machines and Control, 2004, 8(3): 258-263.)
- [4] Passino M, Yurkovich S. Fuzzy control[M]. Beijing: Tsinghua University, 2001.
- [5] Young K D, Utkin V I, Ozguner U. A control engineer's guide to sliding mode control [J]. IEEE Trans on Control System Technology, 1999, 7(3): 328-342.
- [6] Huang S J, Lin W C. Adaptive fuzzy controller with sliding surface for vehicle suspension control[J]. IEEE Trans on Fuzzy Systems, 2003, 11(4): 550-559.
- [7] Liu D T, Yi J Q, Zhao D B, et al. Adaptive sliding mode fuzzy control for a two-dimensional overhead crane [J]. Mechatronics, 2005, 15(5): 505-522.
- [8] Lin Chih-min, Mon Yr-jen. Decoupling control by hierarchical fuzzy sliding-mode controller [J]. IEEE Trans on Control Systems Technology, 2005, 13(4): 593-598.