

doi: 10.11835/j.issn.1000-582X.2013.07.005

二叉树支持向量机的旋转机械故障诊断

朱新才¹, 邓 星¹, 周 雄², 胡腾飞¹, 郭 蕾¹

(1. 重庆理工大学汽车学院, 重庆 400054; 2. 重庆科技学院机械学院, 重庆 401331)

摘 要:针对 SVM 二叉树多类分类优先级的确定问题, 通过旋转机械故障实验平台和数据采集系统, 采集旋转机械故障实验台转子正常、转子不平衡、转子不对中、转子轴承内圈裂缝、转子轴承外圈裂缝 5 种工况下的振动信号, 进行零均值化处理; 选择信号的主要频段进行信号重组, 提取其时域无量纲特征值, 利用并联式 SVM 的正检率大小确定 SVM 二叉树多类分类的优先级, 进行故障类型的识别。通过实验, 实现了训练样本的完全可分, 说明此种方法的有效性。

关键词: 旋转机械; 振动; 信号重组; 支持向量机; 故障诊断

中图分类号: TH11 **文献标志码:** A **文章编号:** 1000-582X(2013)07-021-06

Fault diagnosis of rotating machinery based on SVM 2PTMC

ZHU Xincal¹, DENG Xing¹, ZHOU Xiong², HU Tengfei¹, GUO Lei¹

(1. Automobile College, Chongqing University of Technology, Chongqing 400054, China;

2. College of Mechanical Engineering, Chongqing University of Science and Technology, Chongqing 401331, China)

Abstract: A method is proposed to determine the priority class of binary tree SVM. Firstly, the vibration signals of rotating machinery are collected through a rotating machinery fault experimental platform and data acquisition system. The signals are from 5 different conditions, i. e. rotor normal, rotor unbalance, rotor misalignment, rotor bearing inner ring cracks and rotor bearing outer ring cracks. Then the signals are disposed by zero-mean and the main frequency band of the vibration signals are reconstructed to, extract the dimensionless time domain as characteristic value. Finally, the priority class of SVM 2PTMC can be determined by the correct inspection rate of parallel SVM. Training samples can be completely divided in experiments, which verifies the effectiveness of this method.

Key words: rotating machinery; vibration; signal reconstruction; support vector machine (SVM); fault diagnosis

旋转机械在许多工业部门是关键核心设备, 为此需要对旋转机械进行监测。一般情况下采集旋转机械的振动信号作为原始依据, 这种振动信号一般是一种高频振荡信号且其统计量也不为常数, 可以看成是一个以时间为自变量的非平稳信号^[1]。对这些非平稳信号进行数字处理, 从中获取旋转机械故

障的征兆信息。在旋转机械出现故障征兆时能及时的发出预警信号及识别故障类型, 从而来保证旋转机械运行在正常工作条件下, 同时也可以对设备维修提供依据。

近年来人工智能方法如神经网络、模糊理论和遗传算法等在故障诊断中的应用, 能够处理传统故

收稿日期: 2013-02-24

基金项目: 国家自然科学基金资助项目 (51078375)

作者简介: 朱新才(1954-), 男, 重庆理工大学教授, 主要从事流体力学及其控制与应用的研究, (E-mail) cqgzzxc@163.com。

障诊断方法无法解决的问题,使得故障诊断的人工智能方法在近几年得到广泛的认可和应用^[2-5]。基于结构风险最小化原理的统计学习理论是近年来计算学习理论的重要研究成果,为研究有限训练样本情况下的统计模式识别,并为更广泛的机器学习问题建立一个较好的理论框架。同时,该理论也发展了一种高效的分类器学习算法——支持向量机(SVM)^[6-8]。再结合现代信号处理的手段,如小波变换、希尔伯特-黄变换、局部特征尺度分解等来提取特征值,都存在着参数选择困难的问题。笔者从统计学的角度看待信号,能够很好避免这一问题,并利用 MATLAB 中支持向量机工具箱,对旋转机械的 5 种基本状况进行诊断。

1 旋转机械故障特征及检测分析方法

旋转机械的零部件、整机随着机械的运行都会有不同程度的振动。在理想的情况下,不同的故障类型对应着不同的激励,由此旋转机械的振动信号也会有所不同。而实际上存在着许多不确定的、非线性的因素,即使是相同的故障类型的旋转机械振动信号也不能实现完美重合。这就需要对含有丰富运行状态信息的振动信号进行合理的处理,从而提高旋转机械的故障诊断的准确性。

为了实现对旋转机械的故障诊断,本试验采用旋转机械的振动信号来进行旋转机械的故障诊断。首先测出转子正常状态下的振动信号,然后测出在其他条件不变的情况下,人为设置各种故障的振动信号。正常状态下的振动信号作为标准信号,将其其他故障信号与之相应的标准信号进行对比,通过相同的数据处理来评判故障类型与已知的故障类型的匹配情况。本实验主要由旋转机械故障实验平台对转子正常、转子不平衡、转子不对中、轴承内圈裂缝、轴承外圈裂缝 5 种工作状态进行测试,利用振动加速度传感器和数据采集系统测得旋转机械运行时的振动信号。图 1 为旋转机械故障实验平台。

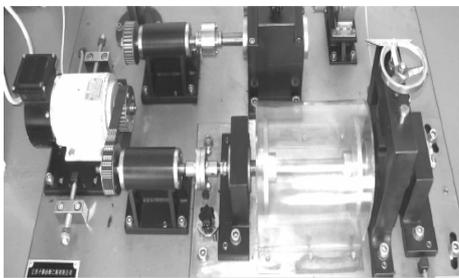


图 1 为旋转机械故障实验平台

2 旋转机械振动信号的特征值选取

2.1 零均值化

对于振动信号,首先对其进行零均值化处理,去掉傅里叶变换后的直流分量。

设振动信号 $x_i(n) = [x_{i1}, x_{i2}, \dots, x_{in}]$;
那么零均值化处理后的振动信号 x'_i

$$x'_{ik} = x_{ik} - \frac{1}{n} \sum_{j=1}^n x_{ij}, k = 1, 2, \dots, n。$$

2.2 信号重构

信号重构是为去掉或减弱信号中不必要的频率成分,保留主要的频率成分。目前也有许多算法,如小波分解后通过设定各个细节的阈值,再进行逆变换来重构信号;希尔伯特-黄变换中使用的经验模式分解后减去残余量,将得到的固有模式序列相加来重构信号等。笔者采用的总体思路是:选取零均值化处理后的信号经过傅里叶变换的频域信号相对较大的幅值——取前信号长度的 25%,令其他频率的幅值为 0,利用傅里叶逆变换得到一个复数序列,其实部就为重构信号。其实质就是一个滤波过程,将幅值影响较小的频率成分滤波。

设零均值化处理后的信号 $x'_i(n) = [x'_{i1}, x'_{i2}, \dots, x'_{in}]$,那么重构信号如图 2。

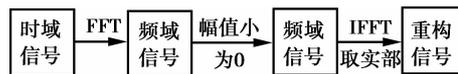


图 2 重构信号原理方框图

2.3 提取特征值

特征值的提取是对于重构信号而言的。笔者选取变异系数、偏度、峰度、波高率、极大值变动率、极小值变动率、方根均值、自乘平方值和对数平均值 9 个无量纲作为特征值如下。

变异系数: $cv = E(|x|)/D(|x|)$; 偏度: $g_1 = E\{[(x - E(x))/\sqrt{D(x)}]^3\}$; 峰度: $g_2 = E\{[(x - E(x))/\sqrt{D(x)}]^4\}$; 波高率: $B_{bg} = |x|/D(|x|)$; 极大值变动率: $P_{max} = \frac{x_{max}}{E_{mau}}$; 极小值变动率: $P_{min} = \frac{x_{min}}{E_{miu}}$; 方根均值: $E_{sqrt} = E(\sqrt{|x|}/\sqrt{D(|x|)})$; 自乘平方值: $E_{zc} = E(x \cdot x)/D(|x|)^2$; 对数平均值: $E_{log} = \sum_{i=1}^n [\lg(|x_i|) + 1]/\lg[D(|x|)]$ 。其中 E_{mau} 为极大值的均值; E_{miu} 为极小值的均值。

3 支持向量机(SVM)故障类型识别原理

旋转机械故障诊断的实质就是一个模式识别问题,判断旋转机械的运行状态是否正常、故障类型是什么和故障发生的具体部位等都可以归结为模式识别的问题^[1]。

支持向量机具有坚实的数学和理论基础,专门针对有限样本情况,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,在小样本学习问题上有其独特的优势;在解决高维问题中,支持向量机通过引入一些具有特殊性质的核函数,将实际问题通过非线性变换转换到高维的特征空间,在高维空间中构造线性决策函数来实现原空间中的非线性决策函数,使得求解支持向量机的过程只与训练样本的数目有关,与样本的维数无关,从而有效避免了“维数灾难”;同时其算法最终转化成为一个凸二次规划的优化问题,从理论上说得到的将是全局最优解^[9-10]。

支持向量机真正具有高价值的地方就是用来解决非线性问题。它是通过一个非线性映射 φ ,把样本空间映射到一个高维乃至无穷维的特征空间,使在特征空间中可以应用线性支持向量机的方法解决样本空间中的高度非线性分类问题。那么,特征空间中的分类约束条件转换为

$$y_i[\mathbf{w} \cdot \varphi(x_i) + b] - 1 \geq 0, i = 1, 2, \dots, n; \quad (1)$$

且 \mathbf{w} 具有最小的二范数

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 = \min_{\mathbf{w}, b} \frac{1}{2} \mathbf{w}^T \mathbf{w}, \quad (2)$$

式中:

$$y_i = \begin{cases} 1, & \mathbf{w} \cdot \varphi(x_i) + b \geq 1; \\ -1, & \mathbf{w} \cdot \varphi(x_i) + b \leq -1. \end{cases} \quad (3)$$

这是一个凸二次规划优化问题,其解可通过求解下面的拉格朗日函数获得,即

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^n \alpha_i \{y_i [\mathbf{w} \cdot \varphi(x_i) + b] - 1\}, \quad (4)$$

式中 $\alpha_i \geq 0$ 为拉格朗日乘子。

分别对 \mathbf{w} 和 b 求偏导数,并令它们等于 0,有

$$\begin{cases} \frac{\partial L(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w}' = \sum_{i=1}^n \alpha_i y_i \varphi(x_i), \\ \frac{\partial L(\mathbf{w}, b, \alpha)}{\partial b} = 0 \Rightarrow \sum_{i=1}^n \alpha_i y_i = 0. \end{cases} \quad (5)$$

将式(5)代入拉格朗日函数式(4),有

$$L(\mathbf{w}, b, \alpha) = \sum_{i=1}^n \alpha_i' - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i' \alpha_j' y_i y_j \langle \varphi(x_i) \cdot \varphi(x_j) \rangle. \quad (6)$$

根据 Mercer 条件,可以用核函数 $k(x_i, x_j) = \langle \varphi(x_i) \cdot \varphi(x_j) \rangle$, 因此,非线性支持向量机就可以转化为目标函数,即

$$\max(L) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j k(x_i, x_j). \quad (7)$$

其约束条件为

$$\begin{cases} \sum_{i=1}^n \alpha_i y_i = 0, \\ \alpha_i \geq 0, \end{cases} \quad (i = 1, 2, \dots, n) \quad (8)$$

的一个二次规划问题。

4 旋转机械故障诊断实验

通过旋转机械故障实验平台,测得 5 种故障的数据文件,为了保证实验结果的准确性,笔者将总数据文件的 70% 作为训练数据文件,将 30% 的数据文件作为测试文件^[11]。将这些数据文件通过数据处理后得到的特征值作为训练样本和测试样本。总的文件数、训练样本数和测试样本数如表 1 所示。

表 1 数据准备

机器状态	文件总数	训练文件数	测试文件数	训练样本数	测试样本数
正常	40	30	10	30	10
不平衡	24	16	8	16	8
不对中	30	22	8	22	8
内圈裂缝	40	30	10	30	10
外圈裂缝	24	16	8	16	8

图 3 是 5 种工况下振动信号的时域波形。由于本实验研究的是 5 种故障类型,因此利用支持向量机进行分类时,可采用一对多组合分类方法,根据一对多组合分类思想,为每一种故障类型构造一个 SVM。故障诊断的 SVM 模型如图 4 所示。

在测试点,1 表示分类到第 1 类,-1 表示分类到其他类。SVM1 用来识别正常和其他情况; SVM2 用来识别转子不平衡和其他情况; SVM3 用来识别转子不对中和其他情况; SVM4 用来识别轴承内圈裂缝和其他情况; SVM5 用来识别轴承外圈裂缝和其他情况。

借助 MATLAB 自带的有关支持向量机的函

数,选取支持向量机的核函数为多项式核函数,编写训练程序和测试程序。程序执行后,得到支持向量机的训练结果表 2 和测试结果表 3。

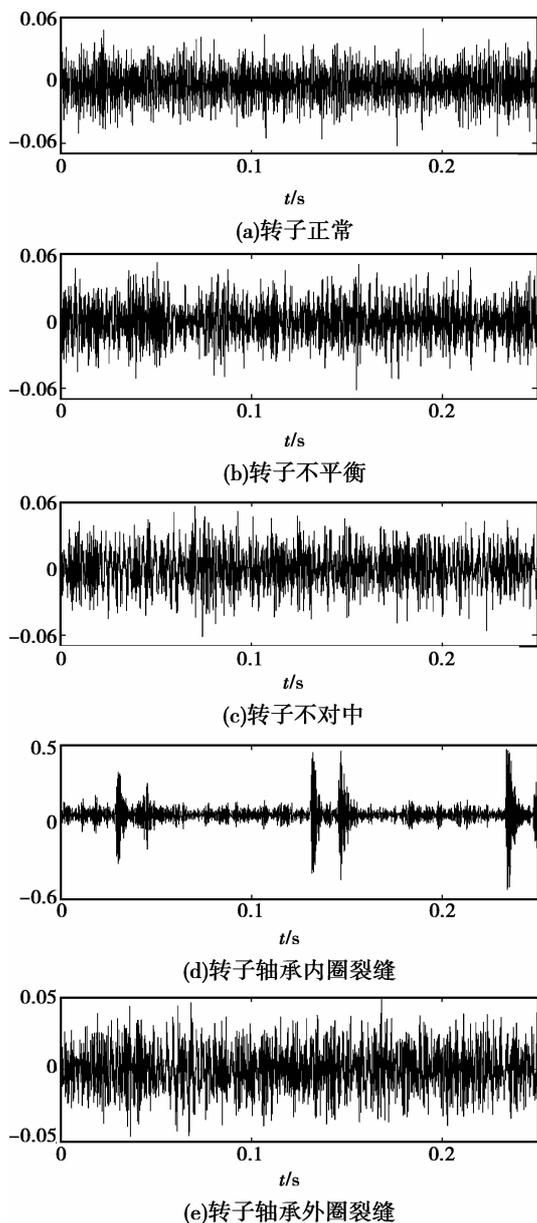


图 3 5 种工况振动信号的时域波形

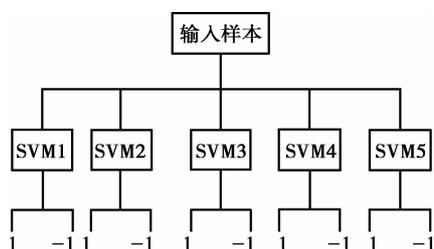


图 4 故障诊断 SVM 模型

表 2 支持向量机的训练结果

机器状态	正常	不平衡	不对中	内圈裂缝	外圈裂缝
正常	30	0	0	0	0
不平衡	0	12	0	0	2
不对中	0	2	19	0	2
内圈裂缝	0	0	0	30	0
外圈裂缝	0	2	3	0	12
正检率%	100	75	86	100	75

表 3 支持向量机的测试结果

机器状态	正常	不平衡	不对中	内圈裂缝	外圈裂缝
正常	10	0	0	0	0
不平衡	0	5	0	0	1
不对中	0	2	7	0	1
内圈裂缝	0	0	0	10	0
外圈裂缝	0	1	1	0	6
正检率%	100	63	88	100	75

表中列对应着输入样本的类别,行对应着分类器输出样本的类别。

通过对旋转机械故障实验台中的转子的振动信号的分析 and 特征提取,从表 3 和表 4 中可以很明显的看出,正常信号和转子轴承内圈裂缝的识别准确率为 100%。其次为轴承外圈裂缝、转子不对中和转子不平衡。其使用的模型为并行结构的一对多支持向量机分类模型。在支持向量机与统计学习理论的已有研究中,重点在如何避免判别平面的法向量为零的情况的发生,如采用非线性判别函数或者级联结构等^[12-15]。而基于 SVM 的二叉树多类分类算法正是级联结构,该方法简单、直观,重复训练样本少^[16]。

依据表 3 和表 4 中正检率的大小来确定二叉树多类分类算法的优先级,其与文献^[16]根据各种状态发生频率的高低排序确定不同。优先级依次为正常信号、转子轴承内圈裂缝、转子不对中、转子轴承外圈裂缝和转子不平衡。本文支持向量机的二叉树多类分类算法分类模型如图 5。

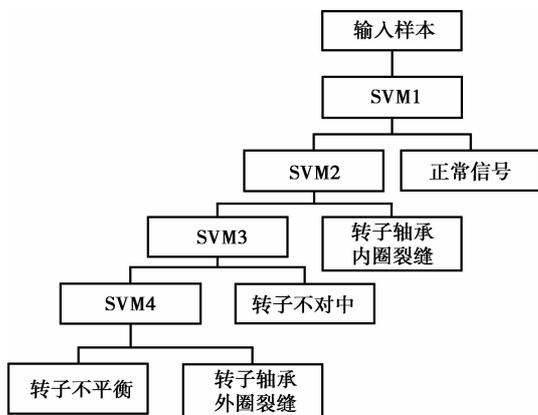


图 5 故障诊断 SVM 二叉树多类分类模型

图 5 中, SVM1 用来识别正常和异常; 如果异常, 再用 SVM2 来识别轴承内圈裂缝和其他异常; 如果是其他异常, SVM3 用来识别转子不对中和其他两类故障, SVM4 用来识别转子不平衡和转子轴承外圈裂缝。支持向量机二叉树多类分类算法的训练结果如表 4、测试结果如表 5 所示。

表 4 支持向量机的训练结果

机器状态	正常	不平衡	不对中	内圈裂缝	外圈裂缝
正常	30	—	—	—	—
不平衡	0	16	0	0	0
不对中	0	0	22	0	0
内圈裂缝	0	—	—	30	—
外圈裂缝	0	0	0	0	16
正检率%	100	100	100	100	100

表 5 支持向量机的测试结果

机器状态	正常	不平衡	不对中	内圈裂缝	外圈裂缝
正常	10	—	—	—	—
不平衡	0	7	0	0	2
不对中	0	1	7	0	0
内圈裂缝	0	—	—	10	—
外圈裂缝	0	0	1	0	6
正检率%	100	88	88	100	75

很明显训练结果转子正常、转子不平衡、转子不对中、转子轴承内圈裂缝和转子轴承外圈裂缝的正检率都达到 100%。

5 结 语

笔者提出利用主要频段进行信号重构, 利用并联式 SVM 模型的正检率大小来确定二叉树多类分类算法的优先级。通过检验, 训练结果的正检率都达到了 100%, 实现了训练样本的完全可分, 测试结果也较改进前有所提高, 取得了较好的诊断结果, 表明该方法的有效性和实用性。笔者所使用的信号重构方法的参数为 0.25, 经过对比可以进一步优化以提高诊断效果。

参考文献:

- [1] 陈晓平, 和卫星, 马东玲, 等. 基于符号熵与支持向量机的滚动轴承故障诊断[J]. 中国机械工程, 2010, 21(17): 2079-2082.
CHEN Xiaoping, HE Weixing, MA Dongling, et al. Symbol entropy and SVM Based rolling bearing fault diagnosis[J]. China Mechanical Engineering, 2010, 21(17): 2079-2082.
- [2] 杨超, 王志伟. 基于小波分析和模糊神经网络的齿轮故障诊断研究[J]. 噪声与振动控制, 2010, 30(4): 64-67, 88.
YANG Chao, WANG Zhiwei. Fault diagnosis of gears based on wavelet analysis and fuzzy neural network[J]. Noise and Vibration Control, 2010, 30(4): 64-67, 88.
- [3] Rao K S R, Mahammad A Y. Neural networks applied for fault diagnosis of AC motors[C]// Proceedings of the 2008 International Symposium on Information Technology, August 26-28, 2008, Kuala Lumpur, Malaysia. Piscataway: IEEE Press, 2008: 1-6.
- [4] Ekici S, Yildirim S, Poyraz M. A transmission line fault locator based on Elman recurrent networks[J]. Applied Soft Computing, 2009, 9(1): 341-347.
- [5] Wang D W, Fung R Y K, Ip W H. An immune-genetic algorithm for introduction planning of new products[J]. Computers and Industrial Engineering, 2009, 56(3): 902-917.
- [6] Saini L M, Aggarwal S K, Kumar A. Parameter optimisation using genetic algorithm for support vector machine-based price-forecasting model in National electricity market[J]. IET Generation, Transmission & Distribution, 2010, 4(1): 36-49.
- [7] He Q B, Yan R Q, Kong F R, et al. Machine condition monitoring using principal component representations[J]. Mechanical Systems and Signal Processing, 2009, 23(2):

446-466.

- [8] Cortes C, Vapnik V. Support-vector networks [J]. *Machine Learning*, 1995, 20(3): 273-297.
- [9] 胡良谋, 曹克强, 徐浩军, 等. 支持向量机故障诊断及控制技术[M]. 北京: 国防工业出版社, 2011.
- [10] 姚富光, 钟先信. 一类支持向量机中引导式量子粒子群优化机理及实现[J]. *重庆大学学报*, 2011, 34(2): 108-112, 119.
- YAO Fuguang, ZHONG Xianxin. Mechanism and implementation of directional quantum-behaved particle swarm optimization in OC-SVM [J]. *Journal of Chongqing University*, 2011, 34(2): 108-112, 119.
- [11] 屈梁生, 张西宁, 沈玉娣. 机械故障诊断理论与方法[M]. 西安: 西安交通大学出版社, 2009.
- [12] Bennett K P, Blue J A. A support vector machine approach to decision trees[C] // *Proceedings of the 1998 IEEE World Congress on Computational Intelligence/The 1998 IEEE International Joint Conference on Neural Networks*, May 4-9, 1998, Anchorage, AK. Piscataway: IEEE Press, 1998, 3: 2396-2401.
- [13] Fröhlich H, Chapelle O, Schölkopf B. Feature selection for support vector machines using genetic algorithm[J]. *International Journal on Artificial Intelligence Tools*, 2004, 13(4): 791-800.
- [14] Wan V, Renals S. Speaker verification using sequence discriminant support vector machines [J]. *IEEE Transactions on Speech and Audio Processing*, 2005, 13(2): 203-210.
- [15] Howley T, Madden M G. The genetic kernel support vector machine: description and evaluation[J]. *Artificial Intelligence Review*, 2005, 24(3/4): 379-395.
- [16] 黄扬帆, 张慧敏, 徐子航, 等. 基超球体支持向量机的不完全二叉树多类分类算法[J]. *重庆大学学报*, 2012, 35(6): 125-128, 140.
- HUANG Yangfan, ZHANG Huimin, XU Zihang, et al. An incomplete binary tree SVM multi-class classification algorithm based on hypersphere[J]. *Journal of Chongqing University*, 2012, 35(6): 125-128, 140.

(编辑 张 苹)

(上接第 15 页)

- LI Minqiang, KOU Jisong. Coordinate multi-population genetic algorithms for multi-modal function optimization[J]. *Acta Automatica Sinica*, 2002, 28(4): 497-504.
- [10] Castro L N D, Zuben F J V. Learning and optimization using the clonal selection principle [J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(3): 239-251.
- [11] Castro L N D, Zuben F J V. The clonal selection algorithm with engineering applications[EB/OL]. [2011-11-17]. <http://wenku.baidu.com/view/ed1f630e76c66137ee0619e8.html>.
- [12] 杨志才. 化工生产中的间歇过程[M]. 北京: 化学工业出版社, 2001.
- [13] 田志锋, 尚宏利, 姚威. 自动化集成电镀生产线的生产调度问题[J]. *重庆理工大学学报: 自然科学*, 2011, 25(6): 38-44.
- TIAN Zhifeng, SANG Hongli, YAO Wei. Schedule problem for automatic integrated electroplated product lines[J]. *Journal of Chongqing University Institute of Technology: Natural Science*, 2011, 25(6): 38-44.
- [14] 李富明, 朱云龙, 尹朝万, 等. 可变机器约束的模糊作业车间调度问题研究[J]. *计算机集成制造系统*, 2006, 12(2): 169-173, 197.
- LI Fuming, ZHU Yunlong, YIN Chaowan, et al. Research on fuzzy job shop scheduling with alternative machines [J]. *Computer Integrated Manufacturing Systems*, 2006, 12(2): 169-173, 197.
- [15] 刘爱军, 杨育, 邢青松, 等. 多目标模糊柔性车间调度中的多种群遗传算法[J]. *计算机集成制造系统*, 2011, 17(9): 1954-1961.
- LIU Aijun, YANG Yu, XING Qingsong, et al. Multi-population genetic algorithm in multiobjective fuzzy and flexible job shop scheduling [J]. *Computer Integrated Manufacturing Systems*, 2011, 17(9): 1954-1961.

(编辑 张 苹)