

# 基于粒子群优化支持向量机的 LIBS 钢液 Mn 元素定量分析

杨友良 王鹏\* 马翠红

河北联合大学电气工程学院, 河北 唐山 063000

**摘要** 为了更好地监测钢液成分,利用激光诱导击穿光谱(LIBS)技术,建立了基于径向基函数(RBF)核函数的支持向量机模型,采用粒子群算法优化支持向量机的参数,通过计算元素特征谱线的积分强度和Fe元素内标归一化来降低仪器和环境带来的干扰。将实验数据进行主成分降维后,对钢液中Mn元素的浓度进行定量分析,得到均方根误差(MSE)为0.599%,相对标准偏差(RSD)为8.26%,相关系数为0.997。结果显示,粒子群优化支持向量机回归定量分析方法可以用于LIBS钢液成分分析,其分析性能较传统的定标方法有一定提高。

**关键词** 光谱学;激光诱导击穿光谱;支持向量机;定量分析

中图分类号 TN247

文献标识码 A

doi: 10.3788/LOP52.073004

## Quantitative Analysis of Mn Element in Liquid Steel by LIBS Based on Particle Swarm Optimized Support Vector Machine

Yang Youliang Wang Peng Ma Cuihong

College of Electrical Engineering, Hebei United University, Tangshan, Hebei 063000, China

**Abstract** In order to make better use of laser induced breakdown spectroscopy (LIBS) in the liquid steel composition monitoring, a model of support vector machine based on radial basis function (RBF) kernel function is established by using particle swarm optimized support vector machine. In order to reduce the interference of instrument and environment, integral intensity of spectral line and Fe normalization is used. The experimental data is subjected to principal component analysis to carry out quantitative analysis of the concentration of Mn element in molten steel, it is obtained that the mean square error (MSE) is 0.599%, the relative standard deviation (RSD) is 8.26%, the correlation coefficient is 0.997. The results show that the particle swarm optimized support vector machine regression method can be used to analyze LIBS of liquid steel composition, its analytical performance is improved compared with traditional calibration methods.

**Key words** spectroscopy; laser-induced breakdown spectroscopy; support vector machine; quantitative analysis

**OCIS codes** 300.6365; 140.3440; 140.3460

### 1 引言

随着钢铁生产日趋大型化、高速化、连续化和测控一体化,利用光谱技术分析物质组成已经成为这个行业重要的发展方向<sup>[1-2]</sup>。激光诱导击穿光谱(LIBS)技术是一种利用激光激发等离子体,然后通过等离子体发射光谱分析物质化学组成的技术,具有快速、实时、多元素同时测量等优点<sup>[3-6]</sup>,已被作为一项极具潜力的物质成分分析技术尝试应用于冶金行业。

支持向量机(SVM)由Vapnik于1995年提出,是一种以统计学习理论为基础,建立在VC(Vapnik-Chervonenkis)维、结构风险最小和核函数原理基础上的新型机器学习方法。支持向量机回归算法适用于小样本,可以融合先验知识,已被广泛运用于各种预测领域。支持向量机的参数选取会直接影响到预测结果的准确度,采用粒子群算法对支持向量机的两个参数进行寻优,可以得到最佳的预测模型<sup>[7-8]</sup>。本文从

收稿日期: 2015-01-30; 收到修改稿日期: 2015-03-04; 网络出版日期: 2015-06-19

基金项目: 国家自然科学基金(61271402)

作者简介: 杨友良(1961—),男,教授,主要从事复杂工业系统的建模与控制等方面的研究。E-mail: 68211246@qq.com

\*通信联系人。E-mail: wp01201213@163.com

LIBS技术的原理出发,建立了基于径向基函数(RBF)核函数的粒子群优化(PSO)寻参支持向量机模型,并以钢液中Mn元素的浓度进行了验证性实验,测试效果良好。

## 2 理论背景

### 2.1 定量分析模型的建立

薄等离子体处于局部热平衡状态,忽略自吸收效应,激光诱导击穿光谱谱线强度表示为

$$I = FC_s \frac{A_{ki} g_k}{U_s(T)} \exp(-E_k/k_B T), \quad (1)$$

式中  $C_s$  为钢液中待测元素  $s$  的浓度,  $I$  为实验中测得的特征谱线强度,  $U_s(T)$  为发射元素  $s$  配分函数,  $k, i$  分别为特征谱线的上下能级,  $g_k$  为上能级简并度,  $A_{ki}$  为特征谱线的跃迁几率,  $E_k$  为跃迁能级的上能级能量,这些参数的取值可参考美国国家标准与技术研究所(NIST)的原子光谱数据库。 $T$  为等离子体温度,  $F$  为与实验条件相关的参数,它们在同一次实验测量中为定值,可通过自由定标法和玻尔兹曼曲线计算得到,  $k_B$  为玻尔兹曼常数。

令  $A = \frac{U_s(T)}{FA_{ki} g_k \exp(-E_k/k_B T)}$  等价变换计算出待测元素浓度的表达式为

$$C_s = AI. \quad (2)$$

理论上,通过(2)式即可求得待测元素浓度,但由于参数  $A$  受到实验环境和样品特性的影响较大,实验中难以精确定量。当待测元素浓度很大时,需要考虑到谱线自吸收效应的影响,待测元素浓度  $C_s$  和特征谱线强度  $I$  不再呈线性关系,而是呈指数关系<sup>[9]</sup>,即

$$C_s = K_b I^b, \quad (3)$$

式中  $K_b$  为比例系数,  $b$  为自吸收系数。

将光谱强度  $I$  作为变量代入支持向量机回归目标函数得

$$C_s = \sum_{i \in S} a_i k_{\text{ibs}}(I_i, I) + b, \quad (4)$$

式中  $S$  为支持向量集,  $a_i$  和  $I_i$  分别为相应支持向量的拉格朗日乘子与谱线强度。  $k_{\text{ibs}}$  为核函数,  $b$  为常数。综合考虑(2)、(3)式中钢液元素浓度和光谱强度的关系,定义  $k_{\text{ibs}}$  为RBF核函数,具体表示为

$$k_{\text{ibs}} = \exp\left(\frac{-\|I - I_i\|^2}{2\sigma^2}\right). \quad (5)$$

### 2.2 粒子群优化寻参原理

合理选择惩罚参数  $c$  和RBF核参数  $g$  是建立支持向量机模型的关键。这两个参数通常通过试凑法和遍历寻优法获得。试凑法除了要求参数选取者的算法分析能力,还要求其有一定的选取经验<sup>[10]</sup>;遍历寻优法的求解过程受所选择的步长影响严重,过大的步长会导致跳过全局最优解,过小的步长又十分耗时<sup>[11]</sup>。为避免上述方法的限制,在参数选取时可以采用粒子群优化算法。

粒子群算法是一种基于迭代模式的优化算法,源于对鸟群捕食行为的模拟,即通过对群体中粒子之间的信息传递来达到寻优目的。下述内容为其基本算法<sup>[12]</sup>。

设在一个  $D$  维搜索空间中,由  $n$  个粒子组成种群  $M = (M_1, M_2, \dots, M_n)$ , 其中第  $i$  个粒子的位置为  $X_i = (X_{i1}, X_{i2}, \dots, X_{id})^T$ , 速度为  $V_i = (V_{i1}, V_{i2}, \dots, V_{id})^T$ , 个体极值为  $P_i = (P_{i1}, P_{i2}, \dots, P_{id})^T$ , 种群的全局极值为  $P_g = (P_{g1}, P_{g2}, \dots, P_{gd})^T$ 。粒子迭代寻优过程中速度和位置的更新公式为

$$V_{id}(k+1) = \omega V_{id}(k) + c_1 r_1 [P_{id}(k) - X_{id}(k)] + c_2 r_2 [P_{gd}(k) - X_{id}(k)], \quad (6)$$

$$X_{id}(k+1) = X_{id}(k) + V_{id}(k+1), \quad (7)$$

式中  $\omega$  为惯性因子,主要用于平衡全局搜索和局部搜索能力。  $d = 1, 2, \dots, n$ ;  $i = 1, 2, \dots, D$ ;  $k$  为进化代数;  $c_1$  和  $c_2$  为加速因子,为非负值;  $r_1$  和  $r_2$  为随机数,取值空间为[0,1]。

每个粒子都需要在历次迭代过程中基于目标函数来决定自适应值的取值(实验中目标函数为均方根误差),根据自适应值确定当前粒子的最优位置  $P_{id}(k)$  和群体最优位置  $P_{gd}(k)$ ,通过(6)、(7)式调整各个粒子的速

度及位置。当迭代次数达到设定值或搜索到的最佳位置满足设定的最小自适应值时,结束寻优。

采用粒子群算法对支持向量机两个参数进行寻优,过程如图 1 所示。

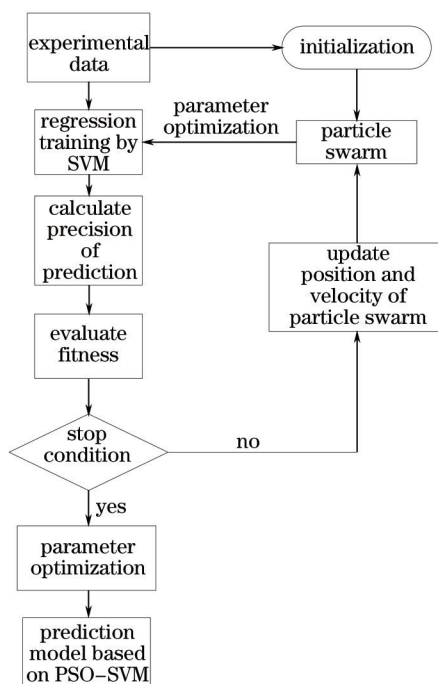


图 1 粒子群优化支持向量机寻参流程图

Fig.1 Particle swarm optimized support vector machine flow chart

### 3 实验部分

#### 3.1 实验装置

图 2 为 LIBS 实验平台结构图。激光器为德国 INNOLAS 公司的 Nd:YAG 激光器,其输出波长为 1064 nm,最大输出能量为 400 mJ,频率为 10 Hz,脉冲激光器通过焦距为 50 mm 的凸透镜会聚在样品表面产生等离子体。光谱仪为海洋光学 LIBS2500-7 高分辨率光谱仪,使用 7 个线性 CCD 阵列探测器用于 200~980 nm 带宽的分析,光学分辨率为 0.1 nm,最小积分时间为 1 ms,实验时激光能量选取 150 mJ,光谱探测延时为 2  $\mu$ s。

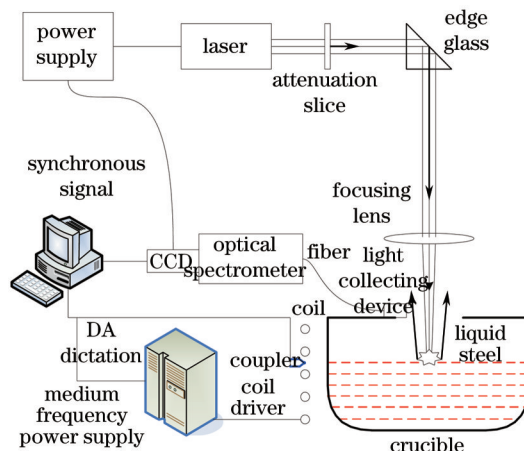


图 2 LIBS 实验平台示意图

Fig.2 Schematic of LIBS experimental setup

#### 3.2 实验过程

实验中为了熔合金钢样本,采用中频炉对合金钢进行加热。中频炉的加热原理是通过电子管振荡电路产生高频电磁场,由高频线圈对钢铁样品进行感应,生成涡电流,进而产生热量使钢铁样品迅速熔化。

实验过程如下:

- 1) 向坩埚中放入需要熔化的合金钢样品,检查所有设备是否完全关闭(包括激光器、光谱仪);
- 2) 将中频炉的输出功率控制在 5 kW 以下,对中频炉进行预热。预热结束后,缓慢增加输出功率,保持直流电流在 70 A 左右,时刻注意控制柜面板上的电流表读数,随着中频炉工作的进行,电流值可能下降,此时要缓慢调节输出功率,使电流值保持在合适的位置。熔化过程中根据实际情况增加电流,但不能超过 85 A;
- 3) 待钢水表面趋于平稳之后,进行测量。脉冲激光经过 45°放置的反射镜后,经过聚焦透镜与钢水直接作用,产生等离子体;
- 4) 通过收光器采集等离子体发射光谱,利用光纤传送至光谱仪,光谱仪将所采集的光谱信号进行分光。光谱仪出口处的探测器对光谱信号进行探测,并将光谱信号转化成数字信号传输到计算机进行分析。图 3 为实验得到的 400~600 nm 范围典型的 LIBS 光谱数据。

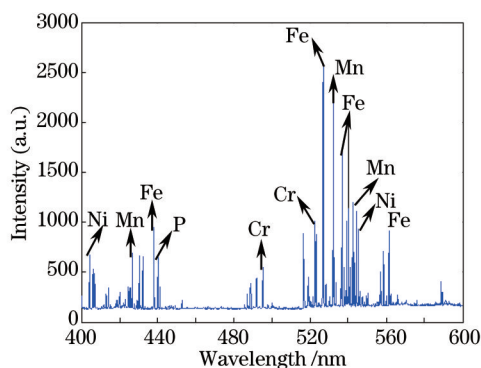


图 3 400~600 nm 部分的典型 LIBS 光谱

Fig.3 Typical LIBS of sample in the range of 400~600 nm

### 3.3 实验样品与特征提取

实验样品是由国家标准物质网提供的钢铁光谱分析标准样品。样品编号和元素组成如表 1 所示。

表 1 实验样品的主要成分

Table 1 Main composition of experimental samples (%)

Sample number	Fe	Cr	Ni	Mn	C	Si	Cu	Co	P	S
ERM279-2	81.786	15.642	1.603	0.258	0.088	0.516	0.107	0.000	0.000	0.000
ERM281-1	70.329	18.170	9.370	0.786	0.048	0.929	0.076	0.023	0.012	0.016
ERM282-1	67.150	16.720	10.860	1.640	0.086	0.734	0.109	0.000	0.019	0.004
ERM284-2D	67.496	16.811	10.720	1.745	0.020	0.537	0.183	0.053	0.026	0.024
ERM287-1D	67.439	18.610	10.350	1.480	0.016	0.569	0.203	0.148	0.027	0.001
ERM289-1D	55.238	14.630	24.680	1.106	0.049	0.531	0.000	0.065	0.012	0.003
ERM292-1D	68.927	18.000	10.090	1.744	0.037	0.402	0.039	0.026	0.018	0.006
ERM295-1D	48.360	19.510	24.400	1.758	0.017	0.418	1.481	0.045	0.017	0.001
ERM296-1D	82.020	11.820	2.790	0.676	0.117	0.242	0.1498	0.0218	0.018	0.003
ERM298-1D	63.380	24.720	7.056	0.398	0.0146	0.262	0.201	0.055	0.020	0.001
ERM276-2C	91.36	4.975	0.203	0.365	0.399	1.034	0.183	0.000	0.001	0.020
ERM278-1	78.795	18.110	0.236	0.405	0.903	0.336	0.077	0.000	0.015	0.005

对 12 个熔化后的实验样品分别进行 30 组测试,由于仪器自身特点和测量过程中不可避免的干扰,谱线的特征峰会发生变形,因此计算元素 Mn 的 20 条特征谱线的积分强度并以基体元素 Fe I 426.87 nm 处的积分强度作为内标(回归分析时 Mn 元素的浓度也用其自身浓度和参考元素 Fe 的浓度比值代替),得到 Mn 的 20 条特征谱线相对强度。利用 Matlab 对 12 种样品的 360 组光谱数据进行主成分分析<sup>[13]</sup>,根据主成分累计解释的贡献率确定最佳主成分数为 5,得到的前 5 个主成分的得分矩阵为自变量矩阵,大小为 360×5,30 次测量中不同样品对应的 Mn 元素浓度和参考元素 Fe 的浓度比为因变量矩阵,大小为 360×1,二者共同构成了支持向量机回归输入数据集。

## 4 回归训练与结果分析

### 4.1 回归训练

在经过处理的 360 组输入数据中,随机抽取 240 组作为训练集,剩余 120 组作为预测集。采用 2.1 节中的



定量分析模型和2.2节中的粒子群优化寻参方法,基于台湾大学林智仁等开发设计的 libsvm3.1 软件包对钢液中的 Mn 元素进行粒子群优化支持向量机回归训练。

在运用粒子群算法优化支持向量机参数的过程中,粒子维度  $D=2$ ,最大优化代数 100,粒子群数目  $n=50$ ,加速因子  $c_1=1.5$ ,  $c_2=1.7$ ,惯性因子  $\omega=1$ ,惩罚参数  $c$  的搜索范围为  $[0,500]$ ,核参数  $g$  的搜索范围为  $[0,100]$ 。粒子群适应度(均方根误差)变化如图 4 所示,最优惩罚参数  $c$  为 25.4639,核参数  $g$  为 1.6156。图 5 为钢液中 Mn 元素的训练集预测结果,图 6 为钢液中 Mn 元素的测试集预测结果。

利用传统的一元线性回归方法对钢液中的 Mn 元素含量进行定量分析,选取 Mn 元素的特征谱线 Mn I 403.076 nm 和 Mn 元素的浓度进行线性拟合定标,结果如图 7 所示。

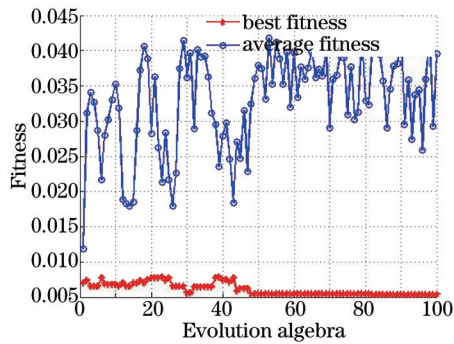


图 4 适应度变化曲线

Fig.4 Fitness curves

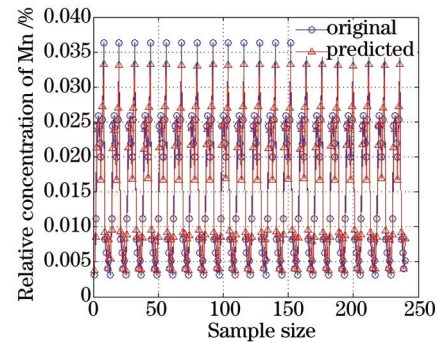


图 5 支持向量机训练集预测结果

Fig.5 Prediction results of the training set in SVR

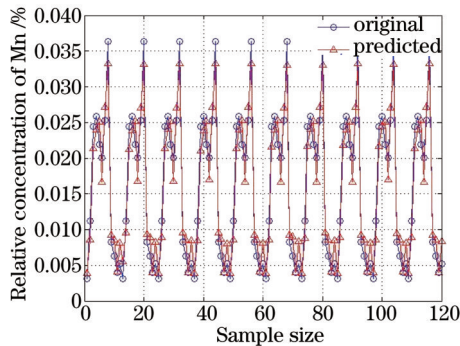


图 6 支持向量机测试集预测结果

Fig.6 Prediction results of the test set in SVR

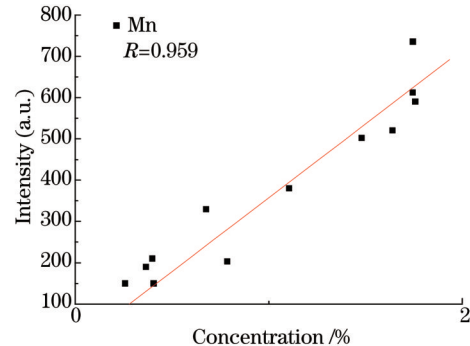


图 7 Mn 元素一元非线性回归定标曲线

Fig.7 Unitary nonlinear regression calibration curve of Mn

## 4.2 结果分析

相对标准偏差(RSD)用于衡量回归预测的重复性,即独立预测结果之间的一致程度。计算公式为

$$f_{\text{RSD}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} / \bar{x} \times 100\% , \quad (8)$$

式中  $f_{\text{RSD}}$  为相对标准偏差,  $x_i$  为预测值,  $\bar{x}$  为预测平均值,  $n$  为预测次数。

均方根误差(MSE)用来评价回归预测的准确性,即预测结果与真实值之间的一致程度。计算公式为

$$f_{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i)^2} , \quad (9)$$

式中  $f_{\text{MSE}}$  为均方根误差,  $\hat{x}_i$  为钢液中 Mn 元素浓度的真实值,  $x_i$  为 Mn 元素浓度的预测值。

由支持向量机回归测试集预测结果可知,钢液中 Mn 元素浓度的 RSD 为 8.26%,MSE 为 0.599%,相关系数  $R=0.997$ ,而传统的一元线性定标法测得 Mn 元素浓度的 RSD 为 15.7%,MSE 为 5.31%, $R=0.959$ 。说明粒子群优化支持向量机对钢液中 Mn 元素浓度定量分析较传统的定量分析方法在重复性和准确性上均有一定改善。

## 5 结 论

结合激光诱导击穿光谱定量分析原理设计了RBF支持向量机回归分析模型,并利用粒子群算法对参数进行优化。为了降低实验环境和仪器设备带来的干扰,光谱数据的输入值选用Mn元素特征谱线的积分强度并用基体元素Fe的特征谱线的积分强度作为内标。采用主成分分析技术对输入的光谱数据进行降维处理,利用粒子群优化支持向量机对Mn元素的浓度进行预测。结果表明,粒子群优化支持向量机回归可以用于LIBS钢液成分的定量分析,学习和泛化能力良好,对元素Mn的浓度预测中,相对标准偏差RSD为8.26%,均方根误差MSE为0.599%,相关系数 $R=0.997$ 。

## 参 考 文 献

- 1 Li Junxiang, Yang Youliang, Meng Fanwei, *et al.*. Matrix correction method used for liquid steel online quantitative analysis by LIBS[J]. *Laser & Optoelectronics Progress*, 2013, 50(3): 031406.  
李俊香, 杨友良, 孟凡伟, 等. 用于LIBS钢液在线定量分析的基体校正方法[J]. *激光与光电子学进展*, 2013, 50(3): 031406.
- 2 Sun Lanxiang, Yu Haibin, Xin Yong, *et al.*. Quantitative analysis of Mn and Si of alloy steels by laser-induced breakdown spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2010, 30(12): 3186-3190.  
孙兰香, 于海斌, 辛 勇, 等. 采用激光诱导击穿光谱技术测定合金钢中锰和硅的含量[J]. *光谱学与光谱分析*, 2010, 30(12): 3186-3190.
- 3 Li Min, Zhu Xinyong, Xu Yuan, *et al.*. Quantitative determination of Cu in lake water by laser induced breakdown spectroscopy[J]. *Laser & Optoelectronics Progress*, 2013, 50(1): 013001.  
李 敏, 朱心勇, 徐 媛, 等. 应用LIBS技术定量检测湖水样品中的铜[J]. *激光与光电子学进展*, 2013, 50(1): 013001.
- 4 Lin Yongzeng, Yao Mingyin, Chen Tianbing, *et al.*. Laser induced breakdown spectroscopy of Cu and Cr detection of Gannan navel orange planting soil[J]. *Laser & Optoelectronics Progress*, 2013, 50(5): 053002.  
林永增, 姚明印, 陈添兵, 等. 激光诱导击穿光谱检测赣南脐橙种植土壤的Cu和Cr[J]. *激光与光电子学进展*, 2013, 50(5): 053002.
- 5 Feng Weilei, Wang Fujuan, Zeng Wanqi, *et al.*. CCD spectrum measurement system for laser induced breakdown spectroscopy[J]. *Laser & Optoelectronics Progress*, 2013, 50(1): 013002.  
冯为蕾, 王福娟, 曾万祺, 等. 应用于LIBS的CCD光谱测量系统[J]. *激光与光电子学进展*, 2013, 50(1): 013002.
- 6 L St-Onge, E Kwong, M Sabsabi, *et al.*. Quantitative analysis of pharmaceutical products by laser-induced breakdown spectroscopy[J]. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2002, 57(7): 1131-1140.
- 7 Li Yujun, Tang Xiaojun, Liu Junhua. Application of least square support vector machine based on particle swarm optimization in quantitative analysis of gas mixture[J]. *Spectroscopy and Spectral Analysis*, 2010, 30(3): 774-778.  
李玉军, 汤晓君, 刘君华. 基于粒子群优化的最小二乘支持向量机在混合气体定量分析中的应用[J]. *光谱学与光谱分析*, 2010, 30(3): 774-778.
- 8 Ye Meiyang, Wang Xiaodong. The support vector machine method of chaotic optical system identification[J]. *Acta Optica Sinica*, 2004, 24(7): 953-956.  
叶美盈, 汪晓东. 混沌光学系统辨识的支持向量机方法[J]. *光学学报*, 2004, 24(7): 953-956.
- 9 Wang Chunlong, Liu Jianguo, Zhao Nanjing, *et al.*. Quantitative analysis of laser-induced breakdown spectroscopy of heavy metals in water based on support-vector-machine regression[J]. *Acta Optica Sinica*, 2013, 33(3): 0330002.  
王春龙, 刘建国, 赵南京, 等. 基于支持向量机回归的水体重金属激光诱导击穿光谱定量分析研究[J]. *光学学报*, 2013, 33(3): 0330002.
- 10 Wang Rui. Method analyze about support vector machine's parameter[J]. *Journal of Chongqing Normal University (Natural Science Edition)*, 2007, 24(2): 36-38.  
王 睿. 关于支持向量机参数选择方法分析[J]. *重庆师范大学学报(自然科学版)*, 2007, 24(2): 36-38.
- 11 Gong Yonggang, Tang Shiping. A novel parameters optimization of SVM for large data sets[J]. *Computer Simulation*, 2010, 27(9): 204-207.  
龚永罡, 汤世平. 面向大数据的SVM参数寻优方法[J]. *计算机仿真*, 2010, 27(9): 204-207.
- 12 Yu Shixing. Support Vector Machine (SVM) Based on Intelligent Algorithm Combined With Wood NIR Application Research[D]. Beijing: Beijing Forestry University, 2014.  
于仕兴. 基于智能算法的支持向量机结合木材近红外光谱应用研究[D]. 北京: 北京林业大学, 2014.
- 13 Wang Qianqian, Huang Zhiwen, Liu Kai, *et al.*. Classification of plastics with laser-induced breakdown spectroscopy based on principal component analysis and artificial neural network model[J]. *Spectroscopy and Spectral Analysis*, 2012, 32(12): 3179-3182.  
王茜茜, 黄志文, 刘 凯, 等. 基于主成分分析和人工神经网络的激光诱导击穿光谱塑料分类识别方法研究[J]. *光谱学与光谱分析*, 2012, 32(12): 3179-3182.