

文章编号: 1006-4710(2013)03-0338-05

基于最小二乘支持向量机的色彩空间转换模型

陈梅^{1,2}, 刘昕¹

(1. 西安理工大学 印刷包装工程学院, 陕西 西安 710048;

2. 杭州电子科技大学 数字媒体与艺术设计学院, 浙江 杭州 310018)

摘要: 以 CMYK 到 CIE $L^*a^*b^*$ 色彩空间的转换为研究对象, 采用最小二乘支持向量机建立了相应的转换模型。实验选取了 IT8.7/3 色靶中的 710 个色块作为训练样本, 36 个色块作为测试样本对所提模型进行了验证, 并与已有的神经网络模型进行了比较。仿真结果表明, 基于最小二乘支持向量机的色彩空间转换模型具有较高的转换精度, 所有测试样本转换后的色差值均小于 1。

关键词: 色彩空间; 支持向量机; 非线性; 色彩管理

中图分类号: TS8

文献标志码: A

Color Space Conversion Model Based on the Least Square Vector Machine

CHEN Mei^{1,2}, LIU Xin¹

(1. Faculty of Printing and Packaging Engineering, Xi'an University of Technology, Xi'an 710048, China;

2. School of Media and Design, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: With CMYK to CIE $L^*a^*b^*$ color space conversion as the research object, the least squares support vector machine is adopted to establish the corresponding conversion model. The 710 color blocks selected from IT8.7/3 color targets are served as the training specimen, and 36 color blocks are served as the testing specimen to test the proposed model and also to compare with the existing neural network model. The simulation results indicate that color space conversion model based on the least squares support vector machine is of higher conversion accuracy so that the color difference value of all the testing specimen after having been converted is less than 1.

Key words: color space; support vector machine; non-linear; color management

凡涉及颜色表现的领域都有色彩管理, 色彩管理在解决信息传播、数字图像输出、计算机视觉、印刷复制等各种设备的色彩空间转换中具有重要作用。色彩空间转换作为色彩管理的关键技术, 由与设备无关的色彩空间 CIE $L^*a^*b^*$ 作为“枢纽中心”, 通过它们就可以在各种设备之间传递所有的颜色。而研究不同色彩空间之间转换模型, 并提高转换精度一直是国内外相关学者研究的重要内容^[1-4]。

目前, 人们对于颜色转换空间的研究总体上可分为理论模型法和经验模型法。前者需要充分的理论依据和数学公式来支持, 而后者不需要充分的理论依据, 而是利用测量大量的色谱数据, 通过各种数学算法拟合两色空间的转换关系。

最小二乘支持向量机 (Least squares support vector machine, LS-SVM)^[5] 是在统计学习理论基础上发展起来的一种简化的支持向量机学习方法, 并被成功应用于非线性建模等问题中^[6-8]。CMYK 作为印刷复制领域颜色表征的主要形式, 实现 CMYK 色彩空间到与设备无关色彩空间 CIE $L^*a^*b^*$ 的转换具有重要意义。本文结合 LS-SVM 良好的非线性建模及泛化性能, 提出基于 LS-SVM 的色彩空间转换模型, 并用于实现 CMYK 到 CIE $L^*a^*b^*$ 的色彩空间转换。

1 最小二乘支持向量机算法

最小二乘支持向量机是一种遵循结构风险最小

收稿日期: 2013-05-17

基金项目: 浙江省教育厅科研资助项目 (Y201326949)。

作者简介: 陈梅, 女, 博士生, 副教授, 研究方向为色彩管理。E-mail: cmei@hdu.edu.cn。

刘昕, 男, 博士, 教授, 研究方向为图像处理与印刷适性、印刷复制技术。E-mail: liuxin@xaut.edu.cn。

化原则的核函数学习机器,是 SVM 的一种扩展。相比较于标准型 SVM,LS-SVM 把凸二次规划变成了线性方程组求解问题,求解所需的资源较小,这使得参数选择更为简单,计算速度也更快,但不能保证全局最优解。这种方法较好地解决了非线性、小样本等实际问题。

将低维输入空间的样本映射到高维特征空间(Hilbert 空间),通过非线性映射 $\phi(\bullet)$,就可以将低维非线性回归问题转化为高维特征空间中的线性回归的问题,这是非线性支持向量机的核心。

给定 N 个样本点 $\{x_i, y_i\}$,其中 $x_i \in \mathbf{R}^n, y_i \in \mathbf{R}, i \in [1, N]$ 。构造最优决策函数:

$f(x) = w^T \phi(x) + b, (w \in \mathbf{R}^h, b \in \mathbf{R}) \quad (1)$

在上式中, b 为偏置量, h 为高维特征空间维数。根据结构风险最小化原则和问题求解目标,式(1)需满足如下条件:

$$\begin{cases} \min J = \frac{1}{2} w^T w \\ \text{s. t. } |y_i - w^T \phi(x) + b| \leq \varepsilon, i = 1, \dots, N \end{cases} \quad (2)$$

根据 LS-SVM 算法,定义误差损失函数为误差的二次项,上述问题可转为:

$$\begin{cases} \min_{w,b,e} J(w,b,e) = \frac{1}{2} (w^T w + \gamma \sum_{i=1}^N e_i^2) \\ \text{s. t. } y_i = w^T \phi(x_i) + b + e_i, i = 1, \dots, N \end{cases} \quad (3)$$

式中, γ 为惩罚系数,其作用是当样本超出误差 e 时控制惩罚程度。为了解决最优化问题,用拉格朗日乘子法求解这个具有等式约束的二次规划问题,定义 Lagrange 函数:

$$L(w,b,e,\alpha,\gamma) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N e_i^2 - \sum_{i=1}^N \alpha_i (w^T \phi(x_i) + b + e_i - y_i) \quad (4)$$

式中 α_i 为 Lagrange 乘子。

通过对(4)式中 b, w, e_i 和 α_i 分别求偏导,得(5)式所示的矩阵方程:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (5)$$

其中 $y = [y_1, y_2, \dots, y_N]^T, \mathbf{1} = [1, 1, \dots, 1]^T, \alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T, \Omega_{ij} = K(x_i, x_j)$ 为核矩阵, K 为定义的核函数,对于非线性回归问题,核函数可以实现到高维空间的映射。

由 Mercer 条件可得:

$\Omega_{ij} = \phi(x_i, x_j) = K(x_i, x_j), i, j = 1, \dots, N \quad (6)$

由(5)式可以求得 α_i 和 b ,结合(1)式可得最小

二乘支持向量机函数模型为:

$$y(x) = \sum_{i=1}^N \alpha_i K(x, x_i) + b \quad (7)$$

LS-SVM 中所有的训练样本都认为是支持向量,与 SVM 相比,LS-SVM 用等式约束代替不等式约束,求解过程变为求解一组等式方程(如(5)式所示),而不是求解二次规划问题,所以求解速度更快。

2 基于 LS-SVM 的 CMYK 到 L*a*b* 的色彩空间转换模型

2.1 模型的建立与分析

模型示意图如图 1 所示。模型以颜色的 CMYK 值作为输入,以 CIE L*a*b* 值作为模型的相应输出。

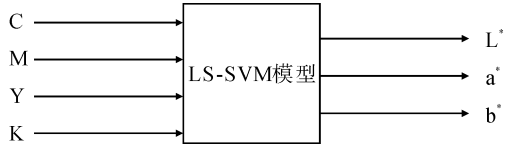


图 1 基于 LS-SVM 的色彩空间转换模型示意图
Fig.1 Model of color space conversion based on LS-SVM

对于 CMYK 四色印刷方式,设 $X = [C, M, Y, K]^T$ 为 4×1 的 CMYK 响应值向量, $Y = [L^*, a^*, b^*]^T$ 是 $n \times 1$ 的 CIE L*a*b* 色彩空间下的响应向量,从 CMYK 到 L*a*b* 的色彩空间转换问题就是把样本的网点覆盖率响应值作为输入变量,把相应颜色的 CIE L*a*b* 色彩空间响应值作为模型输出变量,模型的映射关系是从四维到三维模式,然后建立 X 与 Y 之间的映射关系:

$$Y = f(X) \quad (8)$$

若有训练样本集:

$$S = \{ (x_i, y_i), x_i \in \mathbf{R}^4, y_i \in \mathbf{R}^3, i = 1, 2, \dots, m \}$$

其中 $x_i = (C_i, M_i, Y_i, K_i)^T$ 是 CMYK 响应值, $y_i = [L_i^*, a_i^*, b_i^*]^T$ 是 CIE L*a*b* 色彩空间下的响应向量。根据上述的颜色转换模型,样本集 S 设置为:

$$S = \{ [(C_i, M_i, Y_i, K_i)^T, (L_i^*, a_i^*, b_i^*)^T] \mid i = 1, 2, \dots, m \} \quad (9)$$

根据式(9)确定的色彩空间转换模型,响应值 (C_0, M_0, Y_0, K_0) 对应的 CIE L*a*b* 值为 $f(C_0, M_0, Y_0, K_0)$ 。

2.2 仿真实验

为了验证上述所建立的 CMYK 到 CIE L*a*b* 色彩空间的 LS-SVM 转换模型的有效性,本文采用 IT8.7/3 样张作为实验样本进行测试,该色靶采用 Epson Stylus Pro 9910 数字喷墨印刷机印制。

IT8.7/3 样张如图 2 所示。

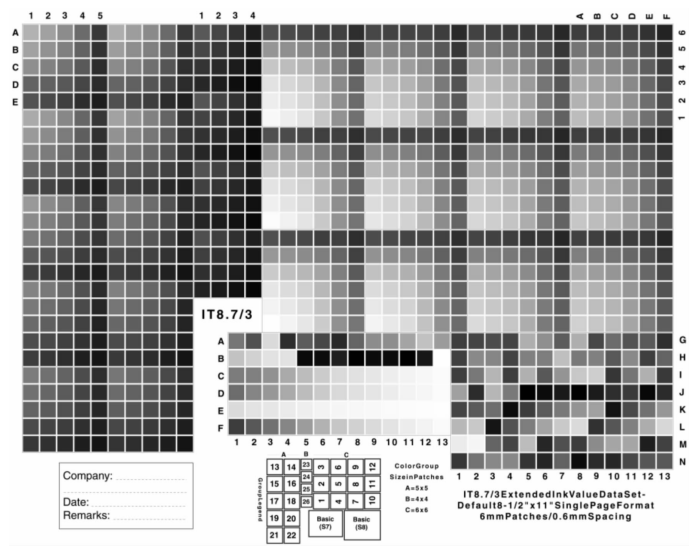


图2 IT8.7/3 色靶
Fig.2 IT8.7/3 Color target

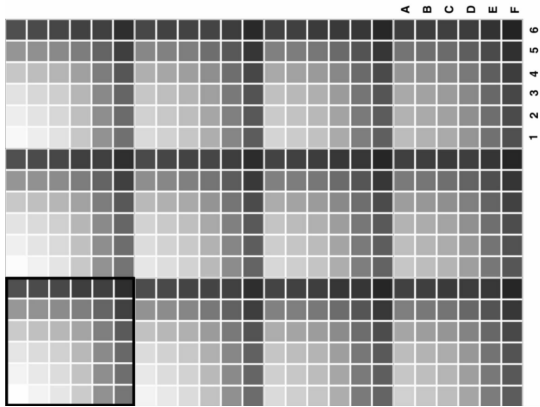
IT8.7/3 是专为 Printer 打印机或是色彩输出机所准备的工业标准。样本总共包括 928 个色块,其中每个色块都有规定的网点面积百分率。为了准确反映印刷工艺的特征,需要大量的颜色。本文在选取用于建模的色块时,考虑到数据在整个色空间的分布,选取了除基本色块外的 746 个扩充色块作为实验对象。它们的网点面积率组合见表 1。

表 1 IT8.7/3 标准色靶推荐用于建立特性文件的数据
Tab.1 The profile data of IT8.7/3 standard color target

系列	K/%	C、M、Y/%	色块数量
1	0	0,10,20,40,70,100	6×6×6=216
2	20	0,10,20,40,70,100	6×6×6=216
3	40	0,20,40,70,100	5×5×5=125
4	60	0,20,40,70,100	5×5×5=125
5	80	0,40,70,100	4×4×4=64
总色块			746

模型的输入模式为 CMYK 色彩空间,输出模式为 CIE L*a*b* 色彩空间,模型的映射关系是从四维到三维模式,因此建模样本点选用的是色靶标准色块的 C、M、Y、K 网点面积覆盖率以及其所对应的目标 CIE L*、a*、b* 测量值组成的数据对,其中测量值均是在 D50 光源和 2° 视场条件下采用 X-rite530分光光度仪测量得到。全部色块样本中随机取 746 个色块作为建模样本,并随机分成建模集和预测集,建模集有 710 个样本,预测集有 36 个样本。

仿真实验的实验样本本取 IT8.7/3 样张 C 区黑框中的色块。样本区域及所取样本如图 3(a)所示,C 区是由 12 组色块形成,每组包含 36(6×6)色块,A~F:蓝增强,1~6:红增强。为了方便识别,本文对所选取的 36 个色块分别采用 S₁~S₃₆进行标记,如图 3(b)所示。



(a) IT8.7/3 色 - C 区

S ₆	S ₁₂	S ₁₈	S ₂₄	S ₃₀	S ₃₆
S ₅	S ₁₁	S ₁₇	S ₂₃	S ₂₉	S ₃₅
S ₄	S ₁₀	S ₁₆	S ₂₂	S ₂₈	S ₃₄
S ₃	S ₉	S ₁₅	S ₂₁	S ₂₇	S ₃₃
S ₂	S ₈	S ₁₄	S ₂₀	S ₂₆	S ₃₂
S ₁	S ₇	S ₁₃	S ₁₉	S ₂₅	S ₃₁

(b) 取样色块

图3 IT8.7/3 色 - C 区及取样色块
Fig.3 IT8.7/3 color target-C and color sample

由于高斯径向基核函数^[9](如式(10)所示)具有良好的可分性和局部性,将其作为 LS-SVM 模型的核函数,且模型中参数 $\delta^2=10,\gamma^{-1}=0.02$ 。

$$K(x,\mathbf{x}_i)=\exp\left(-\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{2\delta^2}\right)\tag{10}$$

表2 测试样本 CMYK 值
Tab.2 The CMYK value of test samples

编号	C	M	Y	K	编号	C	M	Y	K	编号	C	M	Y	K	编号	C	M	Y	K
S ₁	10	40	0	0	S ₁₀	10	40	100	0	S ₁₉	10	40	70	20	S ₂₈	40	0	0	60
S ₂	40	0	10	0	S ₁₁	40	0	0	20	S ₂₀	40	0	100	20	S ₂₉	40	100	0	60
S ₃	100	40	10	0	S ₁₂	100	40	0	20	S ₂₁	100	40	100	20	S ₃₀	40	70	20	60
S ₄	10	10	20	0	S ₁₃	10	10	10	20	S ₂₂	40	20	0	40	S ₃₁	40	40	40	60
S ₅	40	70	20	0	S ₁₄	40	70	10	20	S ₂₃	40	0	20	40	S ₃₂	40	20	70	60
S ₆	100	10	40	0	S ₁₅	100	10	20	20	S ₂₄	40	100	20	40	S ₃₃	40	0	100	60
S ₇	10	100	40	0	S ₁₆	10	100	20	20	S ₂₅	40	70	40	40	S ₃₄	40	100	100	60
S ₈	40	20	70	0	S ₁₇	40	20	40	20	S ₂₆	40	40	70	40	S ₃₅	40	0	40	80
S ₉	100	100	70	0	S ₁₈	100	100	40	20	S ₂₇	40	20	100	40	S ₃₆	40	40	70	80

表3 模型输出仿真值与色块测量值比较
Tab.3 Comparison of the output values with the simulated and measured

编号	L* a* b* 测量值			L* a* b* 仿真值			ΔE	编号	L* a* b* 测量值			L* a* b* 仿真值			ΔE
	L*	a*	b*	L*	a*	b*			L*	a*	b*	L*	a*	b*	
S ₁	69.95	19.92	-10.31	69.46	19.62	-9.98	0.67	S ₁₉	54.87	9.14	37.73	54.67	9.05	37.12	0.65
S ₂	76.26	-11.27	-8.19	75.64	-10.96	-7.90	0.75	S ₂₀	58.56	-22.07	55.26	58.29	-21.35	55.54	0.82
S ₃	41.50	-13.10	-35.46	41.56	-12.75	-34.64	0.89	S ₂₁	30.67	-30.50	13.71	30.95	-29.81	13.57	0.75
S ₄	80.58	-0.14	9.56	79.88	-0.05	9.50	0.71	S ₂₂	51.80	1.13	-15.90	51.66	1.20	-15.46	0.47
S ₅	48.56	25.05	-8.53	48.48	24.65	-8.24	0.50	S ₂₃	56.53	-11.08	0.17	56.30	-10.77	0.29	0.40
S ₆	47.12	-40.59	-14.26	47.07	-39.70	-13.85	0.98	S ₂₄	24.88	36.70	-10.54	25.27	36.07	-10.21	0.81
S ₇	41.74	59.83	8.34	41.80	59.04	8.30	0.79	S ₂₅	35.24	15.05	3.96	35.43	14.84	4.01	0.28
S ₈	62.54	-11.45	35.21	62.19	-11.14	34.65	0.73	S ₂₆	41.97	-1.45	22.06	42.02	-1.33	21.75	0.33
S ₉	20.28	13.63	-11.31	20.76	13.45	-10.96	0.62	S ₂₇	48.09	-11.88	45.68	48.02	-11.56	44.91	0.84
S ₁₀	63.56	11.08	64.33	63.41	11.07	64.20	0.20	S ₂₈	46.99	-5.83	-10.50	46.95	-5.63	-10.17	0.39
S ₁₁	65.39	-7.55	-15.28	64.98	-7.31	-14.85	0.63	S ₂₉	20.45	33.19	-15.89	20.93	32.63	-15.45	0.86
S ₁₂	35.14	-6.16	-36.88	35.33	-5.95	-36.03	0.90	S ₃₀	28.19	15.33	-4.33	28.51	15.12	-4.12	0.44
S ₁₃	69.62	0.95	0.85	69.13	1.02	0.96	0.51	S ₃₁	35.08	1.62	5.16	35.27	1.68	5.19	0.20
S ₁₄	40.72	21.72	-12.35	40.80	21.38	-11.98	0.51	S ₃₂	36.41	-7.84	21.65	36.57	-7.60	21.35	0.42
S ₁₅	40.61	-28.05	-23.69	40.69	-27.41	-23.10	0.87	S ₃₃	41.28	-16.48	40.30	41.35	-16.07	39.64	0.78
S ₁₆	33.70	53.05	-1.70	33.92	52.10	-1.54	0.99	S ₃₄	20.00	12.24	14.32	20.48	12.09	14.17	0.53
S ₁₇	55.67	-6.65	10.49	55.46	-6.43	10.41	0.32	S ₃₅	32.39	-9.02	6.28	32.63	-8.75	6.28	0.36
S ₁₈	17.95	16.35	-20.11	18.47	16.12	-19.59	0.77	S ₃₆	22.71	-0.39	11.90	23.14	-0.29	11.79	0.46

从表3可以看出,基于LS-SVM的CMYK到L*a*b*的色彩空间转换模型具有较高的计算精度,大多数测试样品的色差值都能控制在1以内,符合实际应用过程中对于色差精度的要求。在建模点数量足够充分、取值均匀的情况下,基于LS-SVM进行建模则完全可以实现CMYK到L*a*b*色彩空间的转换。

测试样本的具体数据如表2所示。将测试样本数据输入到LS-SVM模型当中,仿真得到如表3所示的结果,表中包含了与测试色块对应的测量值和仿真值的比较,并给出了相应的色差 ΔE 值。

2.3 与神经网络的比较
为了验证本文所提色彩空间转换模型的有效性,本文将所提的LS-SVM算法与神经网络(NN)方法^[2]进行比较。其中,综合考虑拟合精度和收敛速度,NN模型采用三层神经网络结构,隐层神经元个数为8^[10]。比较结果如表4所示。
从表4可以看出,对于所有样本而言,基于

LS-SVM 的色彩空间转换模型所得出色差均小于 1, 而 NN 模型所得到的大部分色差均大于 1, 因此可以

看出, 基于 LS-SVM 模型的色彩空间转换方法能得到更为精确的转换结果。

表 4 LS-SVM 模型与神经网络模型结果比较

Tab. 4 Comparison between LS-SVM model and Neural Network model

编号	ΔE		编号	ΔE		编号	ΔE		编号	ΔE	
	LS-SVM	NN		LS-SVM	NN		LS-SVM	NN		LS-SVM	NN
S ₁	0.67	0.78	S ₁₀	0.20	1.23	S ₁₉	0.65	1.47	S ₂₈	0.39	1.73
S ₂	0.75	1.35	S ₁₁	0.63	1.64	S ₂₀	0.82	1.91	S ₂₉	0.86	1.97
S ₃	0.89	1.58	S ₁₂	0.90	1.86	S ₂₁	0.75	1.61	S ₃₀	0.44	1.48
S ₄	0.71	0.87	S ₁₃	0.51	1.48	S ₂₂	0.47	1.96	S ₃₁	0.20	1.42
S ₅	0.50	1.69	S ₁₄	0.51	3.27	S ₂₃	0.40	1.60	S ₃₂	0.42	2.11
S ₆	0.98	1.90	S ₁₅	0.87	1.69	S ₂₄	0.81	1.82	S ₃₃	0.78	1.24
S ₇	0.79	1.15	S ₁₆	0.99	1.70	S ₂₅	0.28	2.02	S ₃₄	0.53	1.62
S ₈	0.73	0.90	S ₁₇	0.32	0.61	S ₂₆	0.33	0.60	S ₃₅	0.36	1.28
S ₉	0.62	1.68	S ₁₈	0.77	1.63	S ₂₇	0.84	3.83	S ₃₆	0.46	2.49

3 结 语

介绍了基于 LS-SVM 的 CMYK 到 CIE L* a* b* 色彩空间的转换。将采集到的 CMYK 的网点面积覆盖率通过 LS-SVM 控制器转换到 CIE L* a* b*, 并通过比较测量值 CIE L* a* b* 与仿真值的 CIE L* a* b* 的色差值, 实验测试验证了 LS-SVM 具有较高的逼近精度, 利用基于该控制器的系统检测能较好地实现 CMYK 到 CIE L* a* b* 色彩空间的转换。

参考文献:

[1] 刘一方, 张云峰, 迟静, 等. 基于 SSLUT 的颜色空间转换的快速算法[J]. 山东大学学报:工学版, 2013, 43(1): 41-47.

Liu Yifang, Zhang Yunfeng, Chi Jing, et al. A fast algorithm for color space transformation based on SSLUT[J]. Journal of Shandong University(Engineering Science), 2013, 43(1): 41-47.

[2] 陈路, 李小东. 基于 BP 神经网络的 CMY 到 XYZ 颜色空间转换算法研究[J]. 包装工程, 2007, 28(7): 63-64.

Chen Lu, Li Xiaodong. Research of algorithm transforming CMY to XYZ color space based on BP Neural Network[J]. Packaging Engineering, 2007, 28(7): 63-64.

[3] 智川, 周世生, 石毅. 颜色空间转换的模糊神经网络辨识算法[J]. 北京理工大学学报, 2011, 31(6): 722-726.

Zhi Chuan, Zhou Shisheng, Shi Yi. Color space transform based on fuzzy and neural network identification algorithm [J]. Transactions of Beijing Institute of Technology, 2011, 31(6): 722-726.

[4] 曹从军. 色彩管理关键技术 CIE L* a* b* 与 CMYK 变换算法的研究[D]. 西安: 西北大学, 2008.

Cao Congjun. Research on the key technology of color management algorithm transforming CIE L* a* b* and CMYK color space [D]. Xi'an: Northwest University, 2008.

[5] Suykens J A, Gestel T V, Brabanter J D. Least square support vector machines [M]. Singapore: World Scientific, 2002: 22-25.

[6] 孙丽萍, 范宇, 张冬妍, 等. 基于 LSSVM 的木材干燥在线建模研究[J]. 仪器仪表学报, 2009, 30(9): 1991-1995.

Sun Liping, Fan Yu, Zhang Dongyan, et al. Online modeling for wood drying based on least squares support vector machine[J]. Chinese Journal of Scientific Instrument, 2009, 30(9): 1991-1995.

[7] 吴建斌, 田茂. 基于支持向量机的地雷识别研究[J]. 仪器仪表学报, 2009, 30(7): 1487-1491.

Wu Jianbin, Tian Mao. Landmine recongnition research base on SVW[J]. Chinese Journal of Scientific Instrument, 2009, 30(7): 1487-1491.

[8] 王焱, 郭伟. 基于最小二乘 SVM 的混沌跳频通信码预测方法[J]. 电子测量与仪器学报, 2007, 21(5): 64-68.

Wang Yi, Guo Wei. Chaotic-FH code prediction method based on LS-SVM[J]. Journal of Electronic Measurement and Instrument, 2007, 21(5): 64-68.

[9] 赵莹. 支持向量机中高斯核函数的研究[D]. 上海: 华东师范大学, 2007.

Zhao Ying. The study on gauss kernel function in support vector machine[D]. Shanghai: East China Normal University, 2007.

[10] 褚辉, 赖惠成. 一种改进的 BP 神经网络算法及应用[J]. 计算机仿真, 2007, 24(4): 75-77.

Chu Hui, Lai Huicheng. An improved back-propagation NN algorithm and its application [J]. Computer Simulation, 2007, 24(4): 75-77.