

文章编号: 1001 - 9014 (2010) 01 - 0045 - 04

# 利用地物波谱学习的遥感影像波段模拟方法

程 熙<sup>1,3</sup>, 沈占锋<sup>1</sup>, 骆剑承<sup>1</sup>, 沈金祥<sup>2,3</sup>, 胡晓东<sup>1,3</sup>, 朱长明<sup>1,3</sup>

(1. 中国科学院遥感应用研究所, 北京 100101; 2 中国科学院 新疆生态与地理研究所, 新疆 乌鲁木齐 830011;  
3 中国科学院研究生院, 北京 100049)

**摘要:**针对已有遥感影像模拟方法难以在影像光谱维上扩展的不足,提出了一种基于地物波谱学习的遥感影像波段模拟方法.以地物波谱库作为先验知识,通过支持向量机拟合地物在不同观测波段范围内反射率之间的复杂非线性关系,进而在多光谱遥感影像已有波段的基础上模拟一个新的波段影像.通过模拟 TM 红波段影像的实验,证明本方法能较为准确地模拟出真实的光谱影像,其模拟结果可靠.进一步将该方法应用于模拟 RS 真彩色影像,验证了本方法的实用性.本方法能够有效地解决多光谱影像波段缺损的问题,并在一定程度上可解决较高空间分辨率遥感影像光谱维的不足,为建立地物波谱与遥感像元波谱的定量联系提出了新的思路.

**关键词:**波段模拟;地物波谱;支持向量机

**中图分类号:** TP701 **文献标识码:** A

## METHOD ON SMULATING REMOTE SENSING MAGE BAND BY USING GROUND-OBJECT SPECTRAL FEATURES STUDY

CHENG Xi<sup>1,3</sup>, SHEN Zhan-Feng<sup>1</sup>, LUO Jian-Cheng<sup>1</sup>, SHEN Jin-Xiang<sup>2,3</sup>,  
HU Xiao-Dong<sup>1,3</sup>, ZHU Chang-Ming<sup>1,3</sup>

(1. Institute of Remote Sense Application Institute, Chinese Academy of Sciences, Beijing 100101, China;  
2. Xinjiang Ecology and Geography Institute, Chinese Academy of Sciences, Urumq 830011, China;  
3. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** A method for remote sensing image band simulation based on ground-object spectral features study by support vector machine (SVM) was proposed, since the existing simulation methods had difficulties in the spectrum bands extension. By taking spectral features into account as a priori knowledge and according to the study of the complicated and nonlinear relationships among ground-object spectra in different bands by SVM, a new band image was simulated on the basis of the existing sensor bands of multi-spectral remote sensing images. Based on TM red band simulation experiments, it is confirmed that the method can precisely simulate the real spectral images and the simulation results are reliable. Furthermore, the method was used to simulating RS true-color experiments. The results verify that it is a practical method to reconstruct the missing bands of multi-spectral image. The method effectively resolves the band defection problem of multi-spectral images, and it also resolves the spectral bands deficiency of high spatial resolution remote sensing image to some extent. The method puts forward a new idea for establishing links between ground-object spectrum and remote sensing pixel spectrum.

**Key words:** band simulation; ground-object spectrum; support vector machine (SVM)

### 引言

遥感影像模拟是在遥感先验知识及已有遥感影像的基础上,通过遥感模型计算,获取特定条件下的模拟影像的技术.通过遥感影像模拟获取地物波谱在不同条件下的表现形式有利于深化遥感成像机理的研究,在对不同类型的传感器的性能进行差异分

析、对新设计的传感器性能进行论证等方面具有很大的应用价值<sup>[1,2]</sup>.此外,通过遥感影像模拟技术可解决多光谱影像波段缺损等问题,对于遥感数据预处理工作也具有重要的意义.

目前遥感影像模拟方法主要有以下两种方式.一种是“图图”模拟,即由已知的遥感图像模拟特定条件下的遥感图像,如叶泽田<sup>[2]</sup>等人利用机载高光

收稿日期: 2009 - 09 - 05, 修回日期: 2009 - 12 - 02

Received date: 2009 - 09 - 05, revised date: 2009 - 12 - 02

基金项目: 国家科技支撑计划重大项目课题 (2006BAJ02A01, 2006BAJ14B08); 国家“863”计划项目课题 (2009AA12Z123, 2007AA12Z141)

作者简介: 程 熙 (1982-), 男, 四川乐山人, 博士生, 主要从事遥感信息提取与应用等方面的研究, E-mail: chengx2@tm.com.

谱 M V I S 数据定量模拟了 SPOT-HRV 的地面反射率图像。图 1 模拟技术已广泛应用于利用已有的传感器影像模拟尚未发射的卫星影像 (如 HJ1 星波段模拟等), 为传感器设计方案论证提供依据。该方法实际上是光谱能量重新分配的过程, 其缺点是需要依赖于原有图像的光谱信息, 难以实现影像光谱维的扩展。另一种是“谱图”模拟, 即以地物波谱数据为基础, 通过遥感物理模型模拟特定条件下的遥感图像, 如 Verhoef<sup>[3]</sup> 等人提出通过计算叶片的光谱反射率、单叶片透射率、地表状况等参数并通过 MODTRAN 模型模拟传感器接收到的图像; 苏理宏<sup>[4]</sup> 等在波谱知识库支持下预测任意视场、光照和观测条件下的地物波谱反射率并实现材料波谱到像元波谱的转换等。“谱图”模拟可以模拟地物波谱测量范围内任意波段的像元光谱值, 但计算复杂度很高, 且只能在理论上计算一定数量模拟数据, 难以生成可靠的可视模拟图像<sup>[5]</sup>。此外, 陈方<sup>[5]</sup> 等人提出依据光谱数据库等先验知识的支持, 在宽光谱光学遥感数据提供的图像纹理信息和地类信息的基础上, 获取细分光谱波段的模拟光学遥感图像。这种方法可看作是“图图”、“谱图”结合的模拟方法, 但这种线性模拟方程只能在已有“宽光谱”的波段范围内进行计算且不适合影像地物复杂的情况。

通过地面实测获取的地物波谱数据是进行遥感理论研究与应用分析的基础<sup>[6]</sup>, 它以较窄的光谱采样间隔记录了大量地物从紫外到远红外波谱范围的反射率或辐射值, 并由此形成了大量地物波谱信息的集合, 即地物波谱库。地物波谱库为我们深入研究地物在不同波段内反射率之间的关系和规律并将其用于遥感影像模拟提供了有力的数据支持。但地物波谱特征由自身复杂的物理化学特性所决定, 难以确定其特征函数的形式, 加之地物类型繁多且受到波谱时间与空间效应的影响<sup>[6]</sup>, 采用数理统计方法难以表达其复杂的非线性关系。本文提出了一种基于地物波谱学习的遥感影像模拟方法, 以地物波谱作为先验知识, 通过机器学习方法来拟合各类地物在不同观测波段范围内反射率之间的复杂非线性关系, 进而在遥感影像已有波段的基础上推算感兴趣波段的模拟影像。

## 1 研究方法

支持向量机 (support vector machine, SVM) 是基于统计学习模式中的结构风险最小化原则的一种机器学习方法, 它较好地解决了小样本、非线性、高维

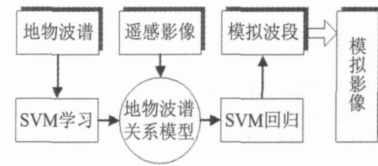


图 1 基于地物波谱学习的遥感波段模拟方法思路  
Fig 1 The sketch of band simulation based on ground-object spectrum study

数和局部极小点等实际问题, 在高维空间具有良好的学习性能和泛化能力, 已被引入遥感影像分析领域并得到成功应用<sup>[7-9]</sup>。本文将 SVM 用于学习典型地物光谱之间的高阶性质, 提出基于地物波谱学习的遥感波段模拟方法思路, 如图 1 所示。

首先将地物波谱库数据作为 SVM 的学习样本, 以拟合地物在不同观测波段范围内反射率之间的隐含关系, 构建地物波谱在对应传感器已知波段与待模拟波段范围内反射率之间的非线性关系模型, 进而将遥感影像像元在已有波段的反射率值作为参数输入该模型以获得该像元在模拟波段的反射率值并合成模拟遥感影像。整个算法包括以下几个主要步骤。

### 1.1 波谱库数据预处理

地物波谱库以连续光谱曲线的形式表示各类地物在不同波长值对应的反射率值, 或是以 ASCII 文件的形式提供给使用者。首先根据不同的波谱宽度 (参照传感器波段设置和待模拟波段) 对波谱库窄光谱数据重采样, 获得各类典型地物对应于传感器成像波段和模拟波段的地表反射率, 得到波谱库样本集  $T = \{ (x_i, y_i) \}_{i=1, 2, \dots, l}$ , 其中样本值为  $x_i \in R^N$ ,  $x_i$  为第  $i$  类地物对应于各传感器波段的反射率组成的  $N$  维向量 ( $N$  为传感器波段数), 目标值为  $y_i \in R$ ,  $y_i$  为第  $i$  类地物对应于模拟波段的反射率值,  $l$  为地物样本个数。由于地物波谱采集与遥感影像成像存在机理与环境差异, 其间还存在空间尺度的问题; 另一方面, 由于地形坡度、坡向、阴影或者太阳高度与强度季节变换, 遥感影像上地表同样物质或目标物的亮度值也会不一样<sup>[6]</sup>。波段间比值计算可以在一定程度上减少这样外部因素的影响, 但若地物在被除波段内反射率值极小, 会产生极大的样本值而影响 SVM 的泛化性能, 很可能导致无效的数学结果。因此, 我们考虑选择一个波段作为标准波段, 计算其它各波段与其反射率之间的归一化波段比值 (NBR)

$$N(x_i) = x_i / \sqrt{x_i^2 + s_i^2}, \quad (1)$$

作为 SVM 学习样本, 此时 SVM 样本集  $T_N = \{ (N(x_i), N(y_i)) \}_{i=1, 2, \dots, l}$ ,  $x_i \in R^{N-1}$ ,  $y_i \in R$  其中  $s_i$

为第  $i$  类地物对应于标准波段的反射率值. 归一化波段比值将波段间的关系限定在  $[0, 1]$  的范围内, 使得 SVM 的学习算法更为稳定和有效, 能够有效地避免无效的模拟值.

## 1.2 地物波谱 SVM 学习

本文利用 SVM 进行地物波谱学习, 并对影像指定波段的反射率进行模拟预测, 这实际上是 SVM 用于解决回归问题的推广形式, 可称之为支持向量回归. 其核心思想是将学习样本从低维空间映射到高维特征空间, 使得在特征空间中构造的线性特性对应于原始空间的复杂非线性特性, 其原理可简单阐述如下文.

若以地物波谱库数据集  $T_n$  作为 SVM 的学习样本, 则对应支持向量回归的决策函数形式为:

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) K(x_i / \sqrt{x_i^2 + s_i^2}, x) + b, \quad (2)$$

其中,  $a_i, a_i^*$  为拉格朗日乘子,  $b$  为阈值;  $x_i / s_i$  为作为支持向量的样本因子向量, 即  $a_i, a_i^*$  不同时为零时所对应的样本向量;  $K(x_i, x)$  为核函数, 其作用是将样本集映射到高维空间实现非线性函数拟合. 关于 SVM 学习算法、支持向量求解等原理的详细论述, 可参见相关文献 [9~11].

通过样本集学习得到式 (2), 即构建了各类地物在不同波谱宽度上反射率的非线性映射模型. 这里需要注意的是, SVM 需要人为给定参数 (不敏感系数、正则化参数  $C$  和相应的核参数等), 参数的选取直接影响 SVM 的性能, 对预测数值结果有很大的影响. 须通过交叉验证或遗传算法等方法进行 SVM 模型参数优化<sup>[12]</sup>, 选取合适的参数进行学习, 以得到最佳的波段模拟精度.

## 1.3 遥感影像波段模拟

对于已有遥感影像, 需要在几何纠正的基础上进行辐射定标和大气校正, 目的是消除传感器损耗和大气影响, 将遥感数据 DN 值转换为直接反应地物特征的地表反射率值. 在此基础上按地物波谱库训练样本集的格式由式 (1) 计算遥感影像各像元在已有各波段的反射率值与标准波段反射率值间的归一化比值  $(N(x_{1i}), N(x_{2i}), \dots, N(x_{ni}))$ , 其中  $x_{ni}$  为影像第  $i$  个像元在第  $n$  个波段上的反射率,  $s_i$  为该像元在标准波段的反射率值) 作为输入因子向量. 对应于任意的影像输入, 由式 (2) 确定的输出即为各像元在模拟波段与标准波段内反射率的归一化比值  $N(y_i)$ . 由式 (1) 反解

$$y_i = N(y_i) / \sqrt{1 - N(y_i)^2} \times s_i, \quad (3)$$

得到该像元在模拟波段的反射率值.

## 2 实验与结果分析

### 2.1 模拟 TM 红波段实验

实验思路是通过地物波谱学习方法模拟 TM 的红波段影像, 并与 TM 原有的红波段影像相比较以验证此方法的模拟精度和可靠性.

#### (1) 实验数据

实验影像数据选取新疆某地的一小块 TM 影像 (获取时间为 1999 年 10 月 15 日), 大小为  $501 \times 501$  像素; 影像区包含农田、裸地、人工地物、道路、岩石和水体等, 地物类型丰富, 具有很好的代表性. 考虑到 TM 数据热红外波段 6 空间分辨率过低、全色波段 8 在波谱角度上并不是一个独立的波段, 所以实验中采用是 TM 波段 1~5、7 共 6 个波段的数据. 数据经大气校正后去除大气环境对地物反射率的影响.

实验用波谱库数据是 USGS 发布的第六版地物波谱数据 (sp1ib06a) 和 RSI 公司 ENV I 软件内置的波谱数据, 共 1365 条. 根据 TM 的波段设置将其重采样为与 TM 的 B、G、R、N R、SW R (波段 5、7) 波段相对应的反射率值.

#### (2) 模拟 TM 红波段

实验中考考虑到地物在短波红外波段 SW R5 的光谱反射率较强, 且对应的 TM 波段范围较大, 因此对窄光谱进行重采样的反射率值相对较大, 将其作为标准波段可得到相对小的比值, 有助于减少实验的误差. 由式 (1) 计算其它各波段向量与其的归一化比值, 这样 SVM 波谱库样本集的输入值为  $(N(B_i), N(G_i), N(NR_i), N(SWR7_i))$ , 目标值为  $N(R_i)$ .

SVM 选用 epsilon-SVR 作为回归方法, 核函数为高斯核  $K(x_i, x_j) = \exp(-x_i - x_j)^2)$ ,  $> 0$ , 由交叉验证法取得最佳正则化参数  $C$  值为 1024, 高斯核宽度为 0.015, 不敏感系数为 0.12 经对波谱库数据集学习后, 获得支持向量 537 个. 以训练样本集作为测试集, 测试回归结果的均方根误差为 0.032.

由式 (1) 计算实验影像各像元在 B、G、N R、SW R7 波段反射率值与标准波段 SW R5 波段反射率值的归一化比值, SVM 回归结果即为该像元在模拟红波段与标准波段内反射率的归一化比值  $N(R_i)$ , 由式 (3) 即可得到模拟红波段影像 R, 如图 2(a) 所示.

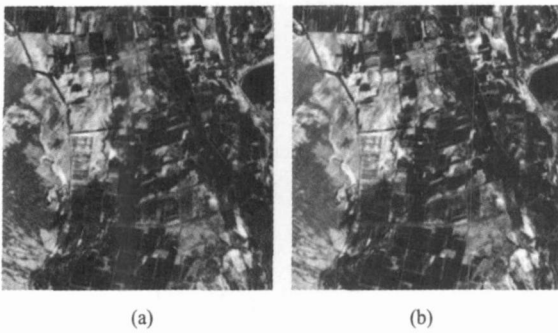


图 2 模拟 TM 红波段与原始波段对比图 (a)模拟红波段 (b)原始红波段

Fig 2 The comparison between simulated-band and original band (a) simulated R (b) original R

表 1 模拟波段与原始波段基本统计值对比

Table 1 The statistical comparison between simulated-band and original band

波段	统计值	中值 Med	均值 Mean	标准差 Std Dev	信息熵 Entropy
模拟波段 R		79	93.272	47.864	7.29
原始波段 R		87	99.229	57.105	7.57

### (3) 实验结果分析

为了验证所得模拟红波段的可靠性,将模拟波段 R 与影像原始波段 R (图 2(b))进行对比分析,如图 2 所示。

从目视效果看,整体上模拟波段影像与原始波段几乎一致,光谱还原准确;细节上,模拟波段对比度较原始波段略低,表现为影像更为平滑,表明本方法的 SVM 预测模型具有较高的鲁棒性,模拟结果弱化了噪声的影响。从定量统计分析上,计算模拟波段和原始波段的基本统计信息,见表 1。

由于受波谱库中典型地物波谱样本数量的限制,实验得到的模拟波段影像的亮度值处于一个相对较小的范围内,加上 SVM 的模拟结果弱化了斑点噪声的影响,因此模拟波段的均值、中值、标准差和信息熵等统计值均较原始波段低(表 1)。进一步计算得到模拟波段对应于原始波段各像元亮度值平均绝对误差为 5.957,为了计算误差在空间上的分布,将两波段影像平均划分为 9 个子区,分别计算各子区每一像元的亮度值平均绝对误差,如图 3 所示。可见模拟波段的误差在影像上的分布比较平均,说明在保证地物波谱库样本数量的前提下,本方法模拟受遥感影像上地物类型影响不大,具有较好的稳定性。

最后计算两波段间相关系数为 0.978(相关系数反映了波段间信息的重复性,相关性越大),说明两波

9.275	3.743	3.081
9.425	4.105	3.179
8.293	4.276	8.004

图 3 模拟误差空间分布

Fig. 3 Spatial distribution of simulation error

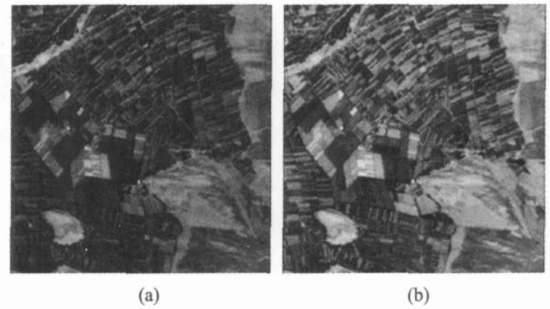


图 4 地物波谱学习法模拟 IRS 真彩色影像结果 (a) 模拟 IRS 蓝波段 (b) 模拟 IRS 真彩色影像

Fig. 4 IRS true-color simulation based on ground-object spectrum study (a) simulated IRS blue band (b) simulated IRS true-color image

段信息的重复量越多,故可表明两个波段信息几乎一致。实验证明,采用本方法模拟的 TM 红波段能够较准确地还原原始波段光谱信息,模拟结果可信。

## 2.2 模拟 IRS 真彩色影像实验

由于蓝波段受天空光影响较大,SPOT 和 RS 均未设置此波段,这导致 SPOT 和 RS P6 等数据不能直接合成真彩色影像,在实际应用中受到了一定的限制。实验利用本文提出的地物波谱学习方法模拟蓝波段,与 RS P6 数据原有的绿、红波段合成真彩色影像。

实验数据选取新疆某地的一块 RS P6 多光谱数据,分辨率为 23m,大小为 500 × 500 像素,并已进行大气校正;实验方法和波谱库数据与上一实验完全相同,计算得到模拟的 RS 蓝波段(参照 TM 蓝波段:0.45 ~ 0.52 μm)如图 4(a)所示,合成 RS 真彩色图像如图 4(b)所示。

从图中可以看出,用本方法模拟得到的蓝波段与 RS 原有的绿、红波段合成的真彩色影像色调自然,光谱信息丰富,目视效果令人满意。实验表明,利用本方法能较好地解决因传感器波段缺损引起的应用问题。

## 3 结论

本文提出的基于地物波谱学习遥感影像波段模拟方法利用机器学习来表达复杂的地物波谱规律,

(下转 62 页)

- 1949, **39** (5): 327—343.
- [5] Clark Jones R. Factors of merit for radiation detectors[J]. *Journal of the Optical of Society American*, 1949, **39** (5): 344—356
- [6] Owen M W. Imaging infrared fundamentals: an alternative perspective[C]. *Proceedings of SPIE on Infrared Imaging Systems: Design, Analysis, Modeling, and Testing III*, 1992
- [7] Sillion F X, Peuch C. *Radiosity and Global Illumination* [M]. San Francisco: Morgan Kaufman Publishers Inc, 1994.
- [8] Donald Heam, Pauline Baker m. *Computer Graphics with OpenGL Third Edition* [M]. America: Pearson Education Inc, 2004.
- [9] Henrik W J, Per C. High quality rendering using ray tracing and photon mapping[C]. *ACM SIGGRAPH 2007 courses*, Aug 2007.
- [10] Ingo W, Carsten B, Philipp S. Distributed Interactive Ray Tracing of Dynamic Scenes [C]. In *Proceedings of the IEEE Symposium on Parallel and Large-Data Visualization and Graphics*, 2003.

(上接 48页)

避免了采用简单统计回归方法推广性弱的缺点,解决了以往方法不能在影像光谱维上进行扩展的问题,以较小的计算代价获取了较精确的模拟影像,模拟方法可靠.本方法有效地解决了多光谱影像波段缺损的问题,并在一定程度上可解决较高空间分辨率遥感影像光谱维的不足,更重要的是为建立地物波谱与遥感像元波谱间的定量联系提出了新的思路.

本方法波段模拟的精度,主要依赖于以下因素.一是地物波谱样本的完整性:由于地物波谱存在时间效应与空间效应<sup>[6]</sup>,故地物波谱的先验知识越丰富,模拟精度越高.理想的情况是加入影像区实地采集的地物波谱数据作为支持向量机的学习样本.二是遥感影像的空间分辨率:由于混合像元普遍存在<sup>[8]</sup>,将遥感影像像元作为纯像元来进行计算降低了影像波段模拟的精度,因此认为遥感影像的空间分辨率越高,波段模拟精度越高.在此基础上考虑混合像元分解、成像系统影响等问题进行遥感影像波段高精度模拟是我们进一步的研究方向.

## REFERENCES

- [1] CHEN Fang, NU Zheng, LIAO Chujiang. Analysis on simulation of remote sensing image and its application[J]. *Geoinformation Science* (陈方,牛铮,廖楚江. 遥感图像模拟技术方法与应用分析. 地球信息科学), 2006, **8** (3): 114—118.
- [2] YE Ze tian, GU Xing fa. Simulation of remote sensing images based on MIVIS data [J]. *Acta Geodaetica et Cartographica Sinica* (叶泽田,顾行发. 利用 MIVIS 数据进行遥感图像模拟的研究. 测绘学报), 2000, **29** (3): 235—239.
- [3] Verhoef W, Bach H. Simulation of hyperspectral and direction radiance images using coupled biophysical and atmospheric radiative transfer models[J]. *Remote Sensing of Environment*, 2003, **87**: 23—41.
- [4] SU Lihong, LI Xiaowen, LIANG Shunlin, et al. Data frame and spectral simulation for remote sensing spectral data base [J]. *Geoinformation Science* (苏理宏,李小文,梁顺林,等. 典型地物波谱库的数据体系与波谱模拟. 地球信息科学), 2006, **4** (7): 7—14.
- [5] CHEN Fang, NU Zheng, QN Yu-chu, et al. Simulation of an image with a subsection of spectral band using an image with a wider spectral band[J]. *Opto-Electronic Engineering* (陈方,牛铮,覃驭楚,等. 基于宽光谱光学遥感图像的细分光谱光学遥感图像的模拟. 光电工程), 2007, **34** (5): 89—96.
- [6] Zhao Yingshi. *Remote Sensing Application And Analysis: Principles and Methods* [M]. Beijing: Science Press (赵英时. 遥感应应用分析原理与方法. 北京: 科学出版社), 2003: 54—57.
- [7] LUO Jian-cheng, ZHOU Cheng-hu, LEUNG Yee, et al. Support vector machine for spatial feature extraction and classification of remotely sensed imagery[J]. *Journal of remote sensing* (骆剑承,周成虎,梁怡,等. 支撑向量机及其遥感影像空间特征提取和分类的应用研究. 遥感学报), 2002, **2** (1): 50—55.
- [8] TAN Kun, Du Pei-Jun. Hyperspectral remote sensing image classification based on support vector machines[J]. *J. Infrared Millim. Waves* (谭琨,杜培军. 基于支撑向量机的高光谱遥感图像分类. 红外与毫米波学报), 2008, **27** (2): 123—128.
- [9] JWU Bo, ZHANG Liang-pei, LI Ping-xiang. Unmixing hyperspectral imagery based on support vector nonlinear approximating regression [J]. *Journal of remote sensing* (吴波,张良培,李平湘. 基于支撑向量回归的高光谱混合像元非线性分解. 遥感学报), 2008, **27** (2): 123—128.
- [10] Huang C, Davis L S. An assessment of support vector machines for land cover classification[J]. *International Journal of remote sensing*, 2002, **23**: 725—749.
- [11] Vapnik V N. An Overview of Statistical Learning Theory [J]. *IEEE Transactions on Neural Networks*, 1999, **10** (5): 988—999.
- [12] GUO Lei, XIAO Huai-Tie, FU Qiang. SVM Model Optimal Multi-parameter selection method for imbalanced data target recognition [J]. *J. Infrared Millim. Waves* (郭雷,肖怀铁,付强. 非均衡数据目标识别中 SVM 模型多参数优化选择方法. 红外与毫米波学报), 2009, **28** (2): 141—145.