

文章编号: 1007-4619 (2002) 01-0050-07

# 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

骆剑承<sup>1</sup>, 周成虎<sup>1</sup>, 梁 怡<sup>2</sup>, 马江洪<sup>3</sup>

(1 中国科学院 地理科学与资源研究所信息室, 北京 100101; 2 香港中文大学地理系, 香港; 3 西安交通大学, 西安 710049)

**摘要:** 提出了基于支撑向量机(SVM)的遥感影像空间特征提取的新方法, 并以 SPOT 全色波段影像上城市特征信息的提取为应用实例, 并与人工神经网络(ANN)等特征提取方法进行综合比较, 认为 SVM 方法不但能够获得比较高的分类精度, 而且在学习速度、自适应能力、特征空间高维不限制、可表达性等方面具有优势。

**关键词:** 支撑向量机, 遥感影像, 空间特征

**中图分类号:** TP751.1 **文献标识码:** A

## 1 引言

遥感空间特征信息分类和提取主要方法包括在领域知识和地理辅助信息支持下的数理统计、神经网络、基于符号逻辑推理方法、分形等<sup>1-4</sup>。一般过程是: 首先通过预处理获得影像的基本空间特征分布, 然后用提取模型对其中的结构单元进行学习记忆, 最后提取影像结构信息, 还可进一步加入一定的辅助知识获得其有关的属性, 比如通过地形因子进行综合判别, 进行推断、判别、决策。其中, 提取模型是关键。人工神经网络方法(ANN)主要包括反向传播神经网络(BPNN)和自组织映射神经网络(ARTMAP)。但对于高维、复杂映射的特征提取问题, BPNN 存在学习速度慢、难以收敛、网络无反馈记忆等问题; ARTMAP 具有自组织反馈、增量式学习、高度复杂映射等特点, 与 BPNN 相比较, 更接近于模拟人的感知记忆系统, 因此适合于应用在高维空间的映射和分类中<sup>5-6</sup>。但由于 ARTMAP 自组织增量式的学习方式, 遇到高度复杂的数据集会使网络急剧膨胀, 影响特征提取的效率, 另外 ARTMAP 同样有神经网络难以对结果作解释的缺陷。

支撑向量机(Support Vector Machine, SVM)是由 Vapnik 针对两类线性可分数据的分类问题提出的,

后来被推广到可处理线性不可分数据<sup>7</sup>。SVM 可用于各种学习和表示, 如神经网络、径向基函数(RBF)、样条及多项式估计等, SVM 提供了函数变化的新形式, 还可用于预报学习问题之外的场合, 如 SVM 参数化可用于解大型线性算子方程组、CT、信号/图像压缩等问题<sup>8,9</sup>。

SVM 方法不同于常规统计和神经网络方法, 它不是通过特征个数变少来控制模型的复杂性。它提供了一个与问题维数无关的函数复杂性的有意义刻画。使用高维特征空间, 使得在高维特征空间中构造的线性决策边界可对应于输入空间的非线性决策边界, 概念上, 通过使用具有很多个基函数的线性估计量, 使在高维空间控制逼近函数的复杂性提供了很好的推广能力; 计算上, 在高维空间上利用线性函数的对偶核, 解决了数值优化的二次规划求解问题。SVM 结合了几个不同概念: (1) 归纳原理的新实施, SVM 是选择有限数据集提供最小期望风险的最优模型; (2) 用事先定义的非线性基函数集把输入样本映射到高维的空间; (3) 用具有复杂性约束的线性函数来逼近或判别高维空间的输入样本; (4) 用优化对偶理论使高维特征空间中的模型参数估计易于计算。

本文引进 SVM 方法, 提出基于 SVM 的遥感影像空间特征提取的具体模型和方法, 目的是解决传

收稿日期: 2000-10-12; 修订日期: 2000-12-26

基金项目: 中国科学院创新项目(KZCX1-Y-02), 国家自然科学基金(40101021)。

作者简介: 骆剑承(1970—), 男, 汉族。1991年浙江大学遥感专业毕业。1999年获中科院地理科学与资源研究所博士学位。现于资源与环境信息系统国家重点实验室从事基础研究, 并为香港中文大学博士后研究人员, 研究方向是空间数据挖掘、遥感图像处理、空间信息认识等。发表学术论文二十余篇。

统空间特征提取模型处理高维数据存在的难以收敛、计算复杂和结果难以解释等问题<sup>[10, 11]</sup>。首先简单介绍 SVM 的基本原理, 然后通过模拟数据的分类实验提出 SVM 遥感影像空间特征提取的具体框架, 并以香港城区 SPOT 影像城市空间特征划分为应用实例, 对其精度、效率、表达等几个方面进行综合分析和比较, 最后是 SVM 空间特征提取和分类方法的结论和展望。

## 2 支撑向量机(SVM)基本原理

### 2.1 两类问题

SVM 的原理是用分离超平面作为分离训练数据的线性函数。SVM 允许直接用训练数据来描述分离超平面, 可以直接解决分类问题, 无需把密度估计作为中间步骤。设训练数据由  $n$  个样本  $(x_1, y_1), \dots, (x_n, y_n)$  构成,  $x \in R^d, y \in \{+1, -1\}$ , 由超平面决策函数来分离:

$$D(x) = (w \circ x) + w_0$$

其中  $w$  和  $w_0$  为适当的系数。定义数据样本可分性的约束为

$$(w \circ x_i) + w_0 \geq +1, \quad \text{若 } y_i = +1$$

$$(w \circ x_i) + w_0 \leq -1, \quad \text{若 } y_i = -1, \quad i = 1, \dots, n$$

或

$$y_i [(w \circ x_i) + w_0] \geq 1, \quad i = 1, \dots, n$$

对给定的训练数据集, 分离超平面可表达为上述形式。从分离超平面到最近数据点的最小距离, 被称为空隙, 用  $\tau$  表示。空隙直接与分离超平面的推广能力有关, 空隙越大, 类间的可分性越大, 因此选取分离超平面的条件是使空隙达到极大。支撑向量是在空隙边沿上的数据点, 或等价地使  $y_i [(w \circ x_i) + w_0] = 1$  的数据点, 也是最接近于决策曲面的数据点, 它们最难被分类, 可决定决策曲面位置, 最优超平面的决策曲面可用支撑向量集来描述。

设样本  $x'$  到分离超平面的距离为  $|D(x')| / \|w\|$ 。假定空隙  $\tau$  存在, 则所有的训练模式服从不等式:

$$\frac{y_k D(x_k)}{\|w\|} \geq \tau, \quad y_k \in \{-1, 1\}, k = 1, \dots, n$$

找最优超平面就是极大化空隙  $\tau$  的  $w$ 。为限制解, 固定  $\tau$  和  $w$  范数之积的刻度, 令  $\tau \|w\| = 1$ 。这样, 极大化空隙  $\tau$  等价于极小化  $w$  的范数。最优分离超平面满足  $y_i [(w \circ x_i) + w_0] \geq 1 (i = 1, \dots, n)$  并关于  $w$  和  $w_0$  极小化  $\eta(w) = \|w\|^2$ 。可用二次

规划(QP)解这个优化问题:

$$\begin{cases} \min & h(w) = \frac{1}{2} \|w\|^2 \\ \text{s.t.} & y_i [(w \circ x_i) + w_0] \geq 1, \quad i = 1, \dots, n \end{cases}$$

首先, 用拉氏乘子构造无约束优化问题:

$$Q(w, w_0, \alpha) = \frac{1}{2} (w \circ w) - \sum_{i=1}^n \alpha_i \{y_i [(w \circ x_i) + w_0] - 1\}$$

其中  $\alpha_i$  是拉氏乘子。函数的鞍点提供了优化问题的解。

然后, 由 Kuhn-Tucker 条件, 只根据参数  $\alpha_i$  来表示上面函数中的参数  $w, w_0$ 。函数  $Q$  就变成只需关于拉氏乘子  $\alpha_i$  极大化的对偶问题。根据 Kuhn-Tucker 定理<sup>[12]</sup>, 通过解偏导数, 可求得  $Q$  的解  $w^*, w_0^*, \alpha^*$ , 具有下列性质:

(1) 系数  $\alpha_i^*, (i = 1, \dots, n)$  满足:

$$\sum_{i=1}^n \alpha_i^* y_i = 0, \quad \alpha_i^* \geq 0, \quad i = 1, \dots, n$$

(2) 向量  $w^*$  是训练集中向量的线性组合:

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i, \quad \alpha_i^* \geq 0, \quad i = 1, \dots, n$$

(3) 只有当数据样本  $(x_i, y_i)$  满足约束

$$y_i [(w \circ x_i) + w_0] = 1$$

时, 相应的参数  $\alpha_i^*$  才不为 0。即:

$$\alpha_i^* [y_i (w^* \circ x_i + w_0^*) - 1] = 0, \quad i = 1, \dots, n$$

其中, 使  $y_i (w \circ x_i + w_0) - 1 = 0$  (或等价地  $\alpha_i^*$  非零) 的数据样本是支撑向量。

二次优化问题的对偶可定制如下:

给定训练数据  $(x_i, y_i), i = 1, \dots, n$  和正则化参数  $C$ , 在约束

$$\sum_{i=1}^n y_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq \frac{C}{n}, \quad i = 1, \dots, n$$

之下, 找极大化函数:

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \circ x_j)$$

的参数  $\alpha_i, i = 1, \dots, n$ 。超平面决策函数为:

$$D(x) = \sum_{i=1}^n \alpha_i^* y_i (x \circ x_i) + w_0^*$$

其中系数  $\alpha_i^*, i = 1, \dots, n$  是对偶问题的解。注意  $\alpha_i^*$  非零的数据样本是支撑向量, 而且这个问题可只根据输入数据向量之间的内积  $(x \circ x')$  来表示。

### 2.2 非线性高维映射

最优超平面是理想状态下的逼近函数, 复杂性

与维数无关,对高维也能提供很好的推广能力。SVM可进一步通过构造高维基函数映射来进行非线性扩展,相应的高维空间称为特征空间。上述优化问题需计算向量之间的内积,这也是需要训练数据  $x$  的唯一运算。若用基函数大集合(即  $g_j(x), j=1, \dots, m$ ),那么解优化问题就需要确定基函数定义的特征空间中的内积。令  $g_j(x), j=1, \dots, m$  表示事先定义的非线性变换函数集,这些函数把向量  $x$  映射到  $m$  维特征空间中,并产生超平面,然后再把特征空间的线性决策边界映射到输入空间的非线性决策边界。用非线性变换函数  $g_j(x)$  来产生特征,决策函数就变为:

$$D(x) = \sum_{j=1}^m w_j g_j(x).$$

其中求和的项数依赖于特征空间的维数。这里去掉了零阶“阈值”项  $w_0$ ,因为它通过在特征空间加入一个常数基函数(即,  $g(x)=1$ )来表示。在对偶形式中,决策函数为:

$$D(x) = \sum_{i=1}^n \alpha_j H(x_i, x).$$

内积核( $H$ )是基函数  $g_j(x)$  的一个表示。对给定的基函数集  $g_j(x)$ ,内积核  $H$  由  $H(x, x') = \sum_{j=1}^m g_j(x) g_j(x')$  来确定,其中  $m$  可取高维。高维特征空间中,特征向量之间内积的计算可间接由支撑向量和输入空间的向量之间的核( $H$ )计算来完成。内积在对偶表示中的展开式允许在输入空间中非线性决策函数的构造,使非常高维特征空间的产生在计算上成为可能。

基函数的选择对应于构造特征所用的函数类的选择,根据 Hilbert-Schmidt 理论,基函数  $H(x, x')$  需要满足 Mercer 条件的对称函数<sup>[8]</sup>。用于学习机的常用基函数类可对应于计算内积的核函数。几个常见的多元逼近函数类和它们的内积核包括多项式、径向基函数、神经网络 SIGMOID 函数等,如:

(1)  $q$  阶多项式内积核:

$$H(x, x') = [(x \cdot x') + 1]^q$$

(2) 径向基函数:

$$f(x) = \text{sign} \left[ \sum_{i=1}^n \alpha_i \exp \left\{ - \frac{|x - x_i|^2}{\sigma^2} \right\} \right]$$

其中  $\sigma$  定义宽度,内积核为:

$$H(x, x') = \exp \left\{ - \frac{|x - x'|^2}{\sigma^2} \right\}$$

### 2.3 多类问题

SVM 的最初提出是针对两类问题,但是可以很

方便地扩展到多类问题的划分中去<sup>[9,13]</sup>。一般有两种分类方式。其中简单的扩展方法是把多类问题分解为两类问题,然后用 SVM 进行训练。也就是每次将其中一个类别的训练数据作为一个类别,其他不属于该类别的训练数据作为另外一个类别。因此对于  $K(K>2)$  类别划分问题,可用  $K$  组支撑向量集表达的决策函数来实现输入空间的划分。另外还有一种方法是建立  $K(K-1)/2$  个 SVM,即对每两类之间训练一个 SVM 将这两类分开。其中前一种方法计算简单,计算量小;后一种方法更能准确地对多类问题进行划分,但是在类别比较多的情况下,计算量也相对复杂。我们选用的是第一种分类方式。图 1 是我们用 SVM 对一组二维输入空间的模拟数据的划分结果(其中图 1(a)为两类问题;图 1(b)为多类问题;实心圆点是支撑向量)。可以看出支撑向量点是进行类别划分的关键点。

## 3 基于 SVM 遥感影像空间特征信息提取

### 3.1 实验区及遥感数据

从卫星上俯视香港,可以看出由于自然条件与历史发展等因素而造成的香港独特的城市空间结构及形态模式(如图 2)。香港城市主体位于维多利亚海湾两岸的港岛和南九龙半岛。香港岛位于香港地区的南端,历史上最早就是从港岛北部的维多利亚港附近的平地 and 低丘区开埠;密集的建筑群在港岛北部狭长的海岸带随地势变化基本上呈东西向延伸。九龙半岛也是香港地区的商业、贸易、交通、文化中心,地势相对平坦,街道呈棋盘状密集分布,原老机场——启德国际机场就位于九龙市区中东部的观塘海滨;九龙市区西面的滨海码头被不断填海扩建。本实验选用 SPOT 全色波段影像,空间分辨率为 10m,影像获取时间是 1999-02-03,影像大小为 600 行、700 列。

### 3.2 提取过程

基于 SVM 方法,利用 SPOT 全色波段数据,我们进行了香港城市空间特征信息的提取和分类实验。应用 SVM 进行遥感影像空间特征提取的基本过程包括如下:

(1) 预处理:对影像数据作基本预处理,包括图像的拉伸、噪声消除等,使得空间结构信息突出,并不受过多干扰。

(2) 样本选取:根据空间特征单元的大小,选取

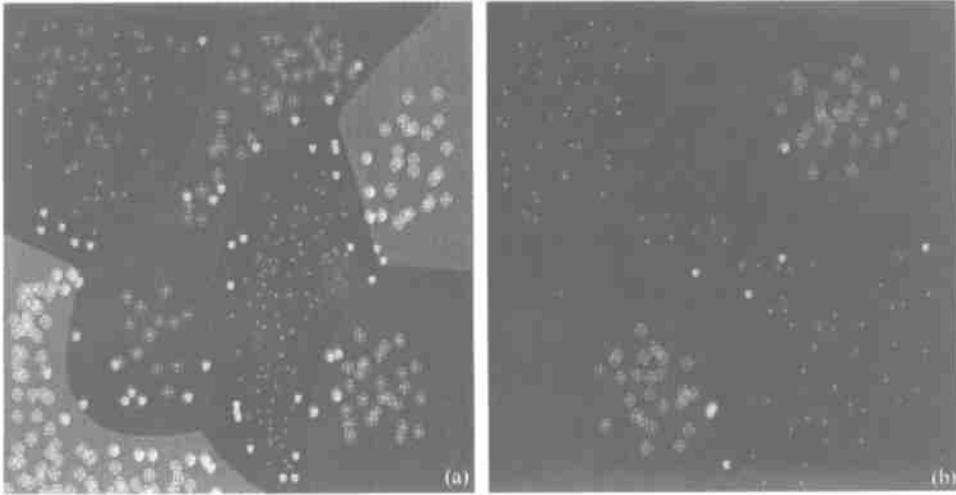


图 1 SVM 分类试例(模拟数据)

Fig. 1 Classifying map of SVM with by simulation data



图 2 香港维多利亚两岸主要城区 SPOT 全色影像

Fig. 2 SPOT Panchromatic image covered on Study area

空间特征单元样本。由于遥感影像中的空间特征信息单元不仅是其光谱属性的反映,更主要是空间域上多个像元的光谱特性的组合,因此空间特征信息的提取和分类一般是通过目标像元为中心的一定大小的窗口区域来表示空间特征单元。

(3) 高维映射: 根据内积核函数,把空间特征样本映射到高维特征空间中去。我们采用的是径向基函数作为核函数。关键是宽度  $\sigma$  的选取,对输入样本进行统一刻度的归一化处理,这样就可以通过定义相对稳定的宽度值来保证高维映射的有效性。

(4) 构造判别函数: 用 SVM 在特征空间中找出各类别特征样本与其他特征样本的分离超平面,得

到代表各空间特征的支撑向量集及其相应的乘子参数  $\alpha_i$ , 形成判断各特征类别的线性判别函数。

(5) 提取和分类决策: 原始图像中未知空间单元,通过内积核映射到特征空间中,作为各类别特征判别函数的输入,然后输出属于各类别的结果,最后通过竞争判断该单元属于相应的空间特征类别。

本次实验在选取样本之前,通过实地调查和分析,试验区一共包含了 4 个主要的空间特征: C1 为水体, C2 为山体, C3 为裸露地, C4 为城市建成区。然后用一定大小的窗口模板去选取空间特征的样本,其中样本选取应该遵循以下几个原则:

(1) 由于空间特征信息在遥感影像上的表现差异较大,包括建筑成份、纹理、宽度、方向、周围单元等都存在差异,因此所选取的样本应该广泛。

(2) 样本选取的模板窗口大小须合适,基本原则是应该稍大于空间特征单元的大小。如果窗口选取太小,则影响信息提取的精度;而窗口太大,影响训练效率而且并不一定能提高精度;不同大小的模板,能提取不同粗细的特征信息。

(3) 样本的选取应该考虑特征结构具有的方向性。可以通过旋转样本窗口的方法来选取不同方向的结构样本。我们对于某一个位置的窗口,分别用旋转  $0^\circ, 90^\circ, 180^\circ, 270^\circ$  等角度及其镜向共 8 个来选取样本。

### 3.3 结果分析

我们分别用  $2 \times 2, 3 \times 3, 4 \times 4, 5 \times 5$  等窗口大小

进行了空间特征信息的取提和分类,分别选取了  $240 \times 8$  个样本数据(240 个点,8 个方向,共 1920 个样本)和 200 个测试样本( $50 \times 4$ )分别用 BPNN、ARTMAP 神经网络方法和 SVM 进行了训练和测试。测试样本是对照香港城市分布地图以及作者长期实地考察的经验选取的,因此测度精度可以代表最后的分类精度。SVM 方法中,我们选用了径向基函数作为内积核进行非线性映射,我们对输入向量作了归一化处理,可以选用统一的参数, $\sigma$ (宽度)为 0.5。

表 1 结构信息提取参数列表

Table 1 Parameters of Spatial Information Extraction

窗口 参数		2×2	3×3	4×4	5×5
		A 向量维数	4	9	16
SVM	训练时间	8s	9s	12s	15s
	分类时间	105s	175s	240s	30s
	测试精度	73.00%	78.00%	82.00%	84.50%
BPNN	训练时间	600s	800s	不收敛	不收敛
	分类时间	8s	12s	无	无
	测试精度	59.00%	63.50%	无	无
ARTMAP	训练时间	55s	78s	105s	160s
	分类时间	56s	61s	125s	180s
	测试精度	73.00%	74.00%	80.50%	84.50%

综合比较结果和主要测试参数如表 1 所示。我们通过用 SVM、ARTMAP 和 BPNN 等提取模型进行遥感影像空间特征信息提取的应用,从过程到结果进行了综合分析,可以得出以下结论:

(1) SVM 与 BPNN 神经网络方法相比较: SVM 的学习效率得到提高,从而使训练时间大大减小,而且 SVM 能够保证收敛;对于 BPNN,当输入向量的维数很高或训练样本数据量较大时,对应的多层前向网络结构也会更加复杂,导致 BPNN 训练速度缓慢,甚至会因为振荡剧烈,使学习难以收敛。

(2) SVM 与 ARTMAP 相比较:都能保持高度复杂映射,而 SVM 训练和提取过程的计算复杂性下降,对划分过程能够用参数形式进行表达。

(3) 选取合适窗口:选择训练样本时窗口模板太小,影响信息提取的精度;选择的窗口太大,对应的计算量就越多,而且由于窗口内包含的信息太复杂,使精度反而下降。因此窗口大小应该与所提取的结构信息相吻合。本次实验工作的最后选取窗口为  $4 \times 4$ (如表 2 所示),空间特征提取和分类结果如图 3 所示。

(4) 在训练过程中,会出现过映射现象,也即由

表 2  $4 \times 4$  窗口模板测试误差矩阵(84.5%,  $\kappa = 0.793$ )Table 2 The Error Matrix of  $4 \times 4$  window

分类 实际	C1	C2	C3	C4	总共
C1	50	10	0	0	60
C2	0	33	0	1	34
C3	0	0	48	2	50
C4	0	7	2	9	47
总共	50	50	50	50	200



图 3 香港城市信息分类结果

Fig. 3 The result of urban information classification

于对训练数据超过一定的过饱和的映射状态后,反而对其他数据的分类或识别能力下降,因此在训练过程需要加入一定的人工交互,以获得最佳的学习状态。

## 4 结论与展望

与传统的神经网络方法相比较, *SIM* 更适用于进行复杂、高维的数据分析处理,因此我们提出了基于 *SIM* 的遥感影像空间特征提取的方法,初步应用于城市空间特征信息的提取和分类,结果表明 *SIM* 在学习、分类效率、可表达性等方面要优于神经网络方法。今后,基于 *SIM* 的遥感空间特征信息提取及分析在几个方面需要进一步深入:(1) *SIM* 学习完毕之后,简化类别判别公式,降低计算量,提高判别效率;(2) 基于 *SIM* 多源空间数据的集成;(3) 如何创建基于 *SIM* 的空间特征单元的影像知识库,形成对特征形状信息、阴影、交通网络、相互关系等主要

结构单元进行 SVM 学习、记忆、识别等的系列化处理过程; (4) 用 SVM 方法进行对城市纹理结构信息的初步提取和识别, 主要包括城市建筑结构、道路网络等所构成的有规则的空间结构单元; (5) 在以上结构信息的提取和识别基础上, 进行形态分析、动态扩散监测及趋势分析。

## 参 考 文 献 (References)

- [1] A-M. Flygare. A Comparison of contextual classification methods using Landsat TM [J]. *INT. J. Remote Sensing*, 1997, **18**(18): 3835—3842.
- [2] U. BHATTACHARYA and S. K. PARUL, An improved back-propagation neural network for detection of roadlike features in satellite imagery [J]. *INT. J. Remote Sensing*, 1997, **18**(6): 3379—3394.
- [3] Francesca Medda, Peter Nijkamp, and Piet Rietveld, Recognition and Classification of Urban Shapes [J]. *Geographical Analysis*, 1998, **30**(3): 304—314.
- [4] F. J. CORTIJO and N. PEREZ DE LA BLANCA, Improving Classical Contextual Classification [J]. *INT. J. Remote Sensing*, 1998, **19**(8): 1591—1613.
- [5] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multi-dimensional maps [J]. *IEEE Transactions on Neural Networks*, 1992, **3**: 698—713.
- [6] Zhou Cheng-hu, Luo Jian-cheng, etc. Remote Sensing Image Geo-Understanding and analysis [M]. Beijing: Science Press, 1999. [周成虎, 骆剑承等, 遥感影像地学理解与分析, 科学出版社, 1999.]
- [7] Vapnik, V., The Nature of Statistical Learning Theory [M]. New York: Springer Verlag, 1995.
- [8] Burges, C. J. C., A tutorial on support vector machines for pattern recognition [J]. *Data Mining and Knowledge Discovery*, 1998, **2**(1): 121—167.
- [9] Bernhard Scholkopf, Christopher J. C. Burges, Alexander J. Smola, Advances in Kernel Methods: Support Vector Learning [M]. The MIT Press, 1999.
- [10] Martin Brown, Hugh G. Lewis, and Steve R. Gunn, Support Vector Machines For Spectral Unmixing [C]. *IGRASS' 99*, 1999, **2**: 1363—1365.
- [11] Lothar Hemes, Dieter Friauff, Jan Puzicha and Joachim M. Buhmann, Support Vector Machines for Land Usage Classification in Landsat TM Imagery [C], in: Proc. of the IEEE International Geoscience and Remote Sensing Symposium, 1999, **1**: 348—350.
- [12] Strang, G., Introduction to Applied Mathematics [M]. Wellesley, MA: Wellesley-Cambridge Press, 1986.
- [13] Cecilio Angulo and Andreu Catala, K-SVCR. A Multi-class support vector machine [C]. R. Lopez de mantaras, E. Plaza (Eds.): ECML 2000, LNAI 1810, 2000, 31—38.

# Support Vector Machine for Spatial Feature Extraction and Classification of Remotely Sensed Imagery

LUO Jian-cheng<sup>1</sup>, ZHOU Cheng-hu<sup>1</sup>, LEUNG Yee<sup>2</sup>, MA Jiang-hong<sup>3</sup>

(1. LREIS, Institute of Geographical Science and Natural Resources Research, CAS, Beijing 100101, China;

2. Department of Geography, the Chinese University of Hong Kong, Hong Kong, China;

3. Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract:** Extracting and classifying spatial features from high-resolution satellite sensor imagery, especially from the image covering urban areas, is a very significant but challenging task. However, it is very difficulty to be implemented and the main obstacle comes from high-dimensional and complicated properties of spatial features. In this paper, the Support vector machine (SVM) is introduced as a new technique for solving a variety of learning and function estimation problems, and it has shown great potential in pattern recognition and computer vision problems. The optimal decision surface of a SVM is constructed from finite support vectors, which are conventionally determined by solving a quadratic programming (QP) problem and independent of the dimension of the features. Based on the idea of SVM, a new approach for spatial feature extraction and classification on high-resolution satellite sensor has been developed by the experimental case of spatial feature classification in central urban area of Hong Kong with SPOT Panchromatic image. The algorithms are synthetically compared with other established algorithms which are represented with conventional neural network algorithms including BPNN and ARTMAP. Finally, the conclusions can be drawn that SVM holds some particular advantages on its faster learning speed, self-adaptive and representative capability, no limitation by high-dimension property in feature space, etc.

**Key words:** Support Vector Machines; remote sensing image; spatial features