

# ICA 结合纹理特征的 SVM 盐渍化信息提取研究

张 飞<sup>1,2,3</sup>,塔西甫拉提·特依拜<sup>1,2</sup>,丁建丽<sup>1,2,3</sup>,买买提·沙吾提<sup>1,2</sup>,田 源<sup>1,2</sup>

ZHANG Fei<sup>1,2,3</sup>,TASHPOLAT·Tiyip<sup>1,2</sup>,DING Jian-li<sup>1,2,3</sup>,ILYAS·Nurmuhammat<sup>1,2</sup>,TIAN Yuan<sup>1,2</sup>

1.新疆大学 资源与环境科学学院,乌鲁木齐 830046

2.新疆大学 绿洲生态教育部重点实验室,乌鲁木齐 830046

3.新疆大学 研究生院,乌鲁木齐 830046

1.College of Resources and Environment Science,Xinjiang University,Urumqi 830046,China

2.Key Laboratory of Oasis Ecology,Xinjiang University,Urumqi 830046,China

3.Graduate School of Xinjiang University,Urumqi 830046,China

E-mail:zhangfei3s@yahoo.com.cn

**ZHANG Fei,TASHPOLAT·Tiyip,DING Jian-li,et al.Study on soil salinization information extraction with SVM's classification based on ICA and texture features.Computer Engineering and Applications,2010,46(9):193-197.**

**Abstract:** This paper takes the Delta Oasis of Weigan and Kuqa rivers for example,using ETM+data,discusses the extraction method of soil salinization.This paper reports the SVM's classification method based on Independent Component Analysis(ICA) and Texture features(ICA-T-SVM method).Meanwhile,the paper introduces the fundamental theory of SVM algorithm and ICA, then incorporates ICA and texture features.The classification result is compared with ICA-SVM classification,single data source SVM classification,Maximum Likelihood Classification(MLC) and neural network classification qualitatively and quantitatively.The research result shows that this method can effectively solve the problem of low accuracy and fracture classification result in single data source classification.It has high spread ability toward higher array input.The overall accuracy is 93.4183%,which increases by 3.4123% compared with SVM based on ICA and increases by 3.4237% compared with single data source SVM classification,and increases by 4.9796% compared with maximum likelihood classification,even increases by 7.7144% compared with neural network classification,and thus acquires good effectiveness.The classification results are easier interpreted when compared with the conventional classification method.Therefore,the classification method based on SVM(Support Vector Machine) and incorporating the ICA and texture features can be adapted to RS image classification and monitoring of soil salinization, furthermore,provides effective way for the things remote sensing information extraction.

**Key words:** Independent Component Analysis(ICA);texture features;Support Vector Machines(SVM);soil salinization;classification

**摘 要:**以渭干河-库车河三角洲绿洲为例,利用 ETM+数据,探讨了该绿洲盐渍化土地覆盖信息的提取方法。提出了基于 ICA 与纹理特征的 SVM 复合的分类方法(简称 ICA-T-SVM 法),通过此方法对该绿洲进行分类研究,并将分类结果与基于 ICA 的 SVM 法(简称 ICA-SVM)、单源数据(光谱)SVM 法、最大似然法(MLC)、神经网络法(Neural Network)分类结果进行定性和定量比较分析。研究表明:该方法能够有效地解决单数据源分类效果破碎、分类精度不高等问题,并对高维输入向量具有较高的推广能力。总精度达到 93.4183%,比基于 ICA 的 SVM 法提高了 3.4123%,比单源信息的 SVM 分类法提高了 3.4237%,比最大似然法提高了 4.9796%,比神经网络法提高了 7.7144%,取得了良好的效果。与传统的分类方法的比较表明,文中所提出的分类方法具有明显的优越性和良好的前景,因此该方法更适合于遥感图像分类和盐渍化信息提取,是地物遥感信息提取的有效途径。

**关键词:**独立成分分析;纹理特征;支持向量机;盐渍地;分类

**DOI:**10.3778/j.issn.1002-8331.2010.09.055 **文章编号:**1002-8331(2010)09-0193-05 **文献标识码:**A **中图分类号:**TP79

## 1 引言

支持向量机(Support Vector Machine,SVM)是由 Corte 和

Vapnik 于 1995 年首先提出的<sup>[1]</sup>,作为一种最新的也是最有效的统计学习方法,近年来成为模式识别与机器学习领域一个新

**基金项目:**新疆教育厅创新研究群体基金项目(No.XJEDU2004G04);中国科学院西部之光项目(No.XJEDU2005I07);新疆绿洲生态重点实验室开放课题(No.XJDX0201-2007-01, No.XJDX0201-2008-03);新疆大学青年教师科研启动基金(No.QN070122)。

**作者简介:**张飞(1980-),男,在读博士,主研方向为遥感与地理信息系统在生态学和资源环境方面的应用;塔西甫拉提·特依拜(1958-),男,博士生导师,教授,主研方向为干旱区遥感应用;丁建丽(1974-),男,硕士生导师,教授,主研方向为资源遥感与 3S 技术应用;买买提·沙吾提(1976-),男,在读博士,主研方向为国土整治与 3S 技术的应用;田源(1980-),男,在读硕士,主研方向为遥感技术的理论和应用研究。

**收稿日期:**2008-09-09 **修回日期:**2008-11-24

的研究热点。SVM 基于结构风险最小化原则,与传统的人工神经网络方法相比,SVM 不仅结构简单,而且各种技术性能尤其是泛化能力明显提高,这已被大量实验证明<sup>[2-4]</sup>,另外,SVM 具有小样本学习、抗噪声性能、学习效率高与推广性好的优点,能够用于解决空间信息处理分析领域的遥感影像处理。遥感图像分析与处理是 SVM 应用一个热门的研究方向<sup>[5]</sup>,一些主要应用如土地利用分类<sup>[6]</sup>、混合像元分解<sup>[7]</sup>、遥感影像融合<sup>[8]</sup>、多光谱/高光谱遥感分类等<sup>[9]</sup>。此外,SVM 算法在模式识别(字符识别、文本自动分类、人脸检测、头的姿态识别)、回归估计、概率密度函数估计、函数逼近、数据挖掘、三维物体识别等方面都有很好的应用<sup>[10-12]</sup>。目前针对多光谱数据进行分类还是仅仅局限于传统的分类方法,不但运算速度慢,分类精度低,而且出现了严重的 huges 现象,而在多光谱遥感分类中 SVM 具有明显的优势。

独立成分分析(Independent Component Analysis,ICA)<sup>[13]</sup>是近年来发展起来的一种新的多维数字信号处理技术。最早是作为一种新的盲源分离技术出现,目前,独立成分分析被一些学者应用到特征提取领域<sup>[14]</sup>。由于独立成分分析着眼于数据间的高阶统计特性,用独立成分分析进行特征提取可使提取得到的特征分量之间不仅互不相关,而且还尽可能统计独立,更有效地利用了数据在统计关系上的本质特征。该方法的目的是将观察到的数据进行某种线性分解,使其分解成线性独立的成分。Y.Rai<sup>[15]</sup>提出了一种简化的 ICA 方法;A.Kocsor 和 J.Csirik<sup>[16]</sup>对 Fast ICA 进行了相关研究和应用,并给出了代码;F.Theis<sup>[17]</sup>提出了一种基于几何方法的 ICA 学习算法;Z.Shi、H.Tang、Y.Tang 等<sup>[18]</sup>也对 Fast ICA 方法进行了相关研究。ICA 方法的兴起为将具有相关性的数据指标转换为互相独立的数据指标提供了强有力的基础。

遥感影像不但包含光谱信息,还包含空间信息。纹理信息是一种常用的空间信息,被越来越多地运用到信息提取中。颜梅春等<sup>[19]</sup>利用 IKONOS 影像提取竹子信息;孙小芳等<sup>[20]</sup>利用纹理信息提取城市绿地;赵萍<sup>[21]</sup>、张锦水<sup>[22]</sup>等将光谱信息与多源信息结合研究遥感影像分类。这些研究表明,纹理信息有助于提高影像分类精度。该文在 ICA 与纹理特征分析的基础上提出了一种改进的支持向量机分类方法,称之为 ICA-T-SVM 分类方法。结果表明分类精度有所提高,此方法能够有效地监测土壤盐渍化信息。

## 2 研究方法

### 2.1 独立成分分析

独立成分分析的主要思想为:通过线性变换,从一组观察信号矩阵  $X=[x_1, x_2, \dots, x_m]$  分离出一组统计独立的信号源矩阵  $S=[s_1, s_2, \dots, s_n]$  的估计矩阵  $Y=[y_1, y_2, \dots, y_n]$ ,  $Y$  也是统计独立的,且假定源信号  $S$  为非高斯分布<sup>[23]</sup>,则  $X$ 、 $Y$  和  $S$  之间的关系为:

$$Y=WX=WAS=\hat{s}$$

其中,  $A$  为  $m \times n$  的未知满秩矩阵,称为混叠矩阵;  $W=A^{-1}$  为分离矩阵;  $\hat{s}$  为  $S$  的估计。独立成分分析算法的任务是找出分离矩阵  $W$ ,使得  $Y$  尽量相互独立,且逼近  $S$ 。由中心极限定理可知,非高斯性可以作为随机变量相互依赖的度量,因此,估计信号  $Y$  是否独立的问题演变为计算分量之间的非高斯性是否达到最大的问题。目前独立成分分析的实现方法很多,应用较为广泛的是 Bell 和 Sejnowski 的信息最大化(infomax)算法以及 Hyvarinen 的 Fast ICA 算法。

### 2.2 纹理特征分析

图像的纹理一般理解为图像灰度在空间上的变化和重复,或图像中反复出现的局部模式(纹理单元)和它们的排列规则。纹理是遥感影像上重要的信息,是进行目视判读和计算机自动解译的重要基础。遥感影像专题信息提取中,在原始影像光谱信息的基础上加上纹理信息可以使分析的准确性和精度提高<sup>[24]</sup>。

纹理分析指的是通过一定的图像处理技术抽取出纹理特征,从而获得纹理的定量或定性描述的处理过程<sup>[25]</sup>。包括两方面的内容:检测出纹理基元和获得有关纹理基元排列分布方式的信息。其基本方法可分为统计方法、结构方法和谱方法三大类。统计方法是指在不知道纹理基元或尚未检测出基元的情况下进行纹理分析,主要描述纹理基元或局部模式随机和空间统计特征,如灰度共生矩阵法、随机场模型法等。结构方法主要在已知基元的情况下进行,描述纹理基元及其周期性排列的空间几何特征和排列规则,如形态学、图论、拓扑等方法。谱方法是建立在多尺度分析与时、频分析基础之上的纹理分析方法,如 Gabor 变换、小波变换、分形学等<sup>[26]</sup>。灰度共生矩阵(Gray Level Co-occurrence Matrices, GLCM)可以描述影像各像元灰度的空间分布和结构特征,在利用影像的纹理特征改善影像的地学目标分类效果方面具有优势<sup>[27]</sup>。

### 2.3 支持向量机分类法

#### 2.3.1 支持向量机

支持向量机的核心思想就是(图 1)把低维空间中的难于线性划分的样本数据通过非线性映射  $\varphi$  映射到高维(可能是无限维)特征空间,在高维特征空间中构造具有低 VC 维的最优分类超平面<sup>[28]</sup>,使分类风险上界最小,从而实现了样本空间的非线性分类,在高维空间中构造最优分类器<sup>[29]</sup>。图 2 是 SVM 求解非线性分类问题过程的示例。高维空间的非线性分类问题还包括二类划分和多类划分,二类划分还可以分线性可分类问题和线性不可分类问题。

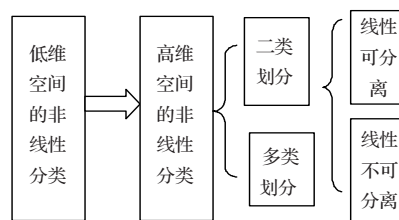


图1 支持向量机(SVM)的基本思路

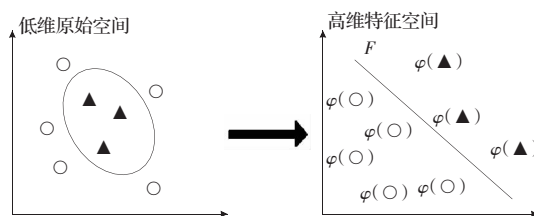


图2 非线性映射图

#### 2.3.2 两类线性可分的支持向量机

SVM 是从线性可分情况下的最优分类超平面(Optimal Hyperplane, OHP)发展来的,基本思想可用两类线性可分情况来说明。如图 3 所示,实心三角点(▲)和空心圆点(○)代表两类的训练样本。假如这两类样本(训练集)是线性可分的,则机器学习的结果是一个超平面(二维情况下是直线)或称为判别

函数,该超平面可以将训练样本分为正负两类。

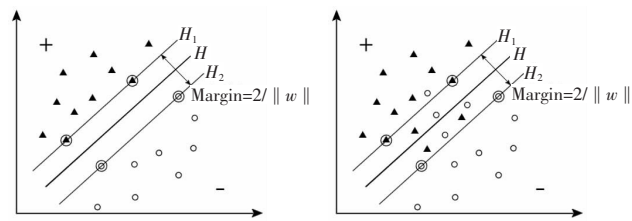


图3 线性可分情况下的最优分类超平面示意图

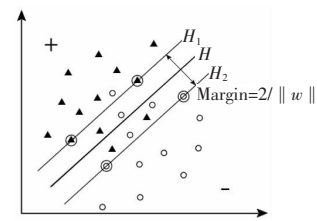


图4 线性不可分情况下最优分类超平面示意图

图中  $H$  为分类线,  $H_1, H_2$  分别为过各类中离分类线最近的样本且平行于分类线的直线, 它们之间的距离叫做分类空隙或分类间隔 (Margin)。标记为“ $\blacktriangle$ ”和“ $\odot$ ”的样本点就是过两类样本中离分类面最近的点且平行于最优分类面的超平面上的训练样本, 被称为支持矢量。各种数学公式见参考文献[30-32]。

### 2.3.3 线性不可分的支持向量机

对于线性不可分的情况 (如图 4 所示) 一般引入一个松弛项  $\zeta_i \geq 0$ , 折衷考虑最少错分样本和最大分类间隔。式(1)变为式(2):

$$\min(\varphi(w)) = \min\left[\frac{1}{2} \|w\|^2\right] = \min\left[\frac{1}{2} w g w\right] \quad (1)$$

$$\min((\varphi, \zeta)) = \frac{1}{2} (w g w) + c \sum_{i=1}^n \zeta_i \quad (2)$$

满足条件  $y_i(w g x_i + b) \geq 1 - \zeta_i$ , 其中  $\zeta_i \geq 0, i=1, 2, \dots, n, \sum_{i=1}^n \zeta_i$

是测试错误数目的一个上界; 正则化参数  $c$  由实验决定, 它可以控制机器的复杂性和不可分离点数之间的平衡, 即控制对错分样本惩罚的程度, 加大  $c$  的值可以减少错分样本的数目。若  $\zeta_i$  都为 0, 则变为线性可分情况下的式(1)。相应的对偶形式为:

$$\max(Q(a)) = \max\left[\sum_{i=1}^n a_i - \frac{1}{2} \sum_{j=1}^n a_i a_j y_i y_j (x_i g x_j)\right] \quad (3)$$

其满足条件:

$$0 \leq a_i \leq c, \sum_{i=1}^n a_i y_i = 0, i=1, 2, \dots, n$$

### 2.3.4 非线性支持向量机

非线性支持向量机问题的基本思想是通过非线性变换转化为某个高维空间中的线性问题, 在变换空间求最优分类面。各种数学公式见参考文献[30-32]。

实际上, 正是核函数的引入使得支持向量机得以实用化, 因为它避免了显示高维空间中向量内积而造成的大量运算。目前常用的核函数包括以下 4 种, 在实际应用中, 通常需要根据问题的具体情况选择合适的核函数以及对应的参数进行处理, 常用的核函数见表 1。

表 1 常用的一些核函数

核函数(Kernel Function)	定义(Definition)	参数 (Parameters)
线性函数(Linear)	$K(x_i, x_j) = x_i^T x_j$	
多项式函数(Polynomial)	$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$	$\gamma, r, d$
径向基函数(Radial Basis Function)	$K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2), \gamma > 0$	$\gamma$
S 形函数(Sigmoid)	$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$	$\gamma, r$

在这 4 种核函数中, 应用最广泛的是 RBF 核, 无论是低

维、高维、小样本、大样本等情况, RBF 核函数均适用, 具有较宽的收敛域, 是较为理想的分类依据函数<sup>[33]</sup>。

### 2.3.5 SVM 的多类问题

SVM 方法最初提出是针对二分类问题, 然而二分类的实际应用是很有局限性的。但是可以把它很方便地推广到多类问题的划分问题中去。SVM 的多类划分问题有着更广泛的实际应用。在遥感领域上土地覆盖的分类问题也是个多类分类问题。

解决多类分类主要是两个方向: 一是将多个分类面的参数求解合并到一个最优化问题中, 通过求解该最优化问题一次性实现多类分类<sup>[34]</sup>, 但这种方法实现难度大、计算复杂度高, 在现实中使用很少; 另一是通过某种方式构造一系列的二类分类器并将它们组合在一起来实现多分类<sup>[34]</sup>。这种方法形式简单、组成灵活, 在现实中得到广泛应用, 二类别的分类器组合有一对一(one versus one, 1-1)分类器, 一对多(one versus rest, 1-v-r)分类器和多对多三种组合方法。一对一方法一般在  $n$  类分类中, 需构造  $n(n-1)/2$  个分类器, 计算量比较大, 本研究使用的就是这种方法; 一对多方法是用一类和剩下其他所有类别判别分类, 一个分类器分出其中某一类。在  $n$  类分类中, 要构造  $n$  个分类器, 这种方法相对计算量较小, 缺点就是存在混分和漏分样本问题; 多对多分类有决策树算法, 先把多类划为二大类, 再在二类中继续往下划分, 直到最后的一对一类判别。

## 2.4 ICA 结合纹理特征的 SVM 分类方法的构建

基于上述思想可构建其分类方法: (1)对影像进行预处理包括几何校正、辐射校正和坐标转换等技术。在此基础上, 选取提取盐渍化信息的最佳光谱波段参与后续的 ICA 分析以及纹理特征提取。这里选取 TM1、TM3、TM5 光谱波段<sup>[35]</sup>; (2)对所选择的光谱波段进行 ICA 分析; (3)提取样本的纹理特征。Haralick<sup>[36]</sup>曾经提出 14 种由灰度共生矩阵计算出来的统计量, 但是 Barald<sup>[37]</sup>认为对于遥感图像来说 Angular second moment (角二阶矩)、Homogeneity (均一性)、Dissimilarity (相异性)、Entropy (熵) 4 种统计量效果最好。然后采用 3x3 窗口计算 3 个光谱波段各自的 4 个纹理特征值, 共有 12 个纹理特征值, 加上 3 个光谱波段的光谱值, 再加上 3 个光谱波段 ICA 变化后的特征值, 共有 18 个特征向量。通过实地考察和结合地形图目视判读, 将该区地面覆盖类型粗分为水体、农田、绿洲、重度盐渍土、中度盐渍土、轻度盐渍土和其他 (其包括岩石、泥土、泥沙等) 等 7 个大类; (4)SVM 样本训练。选择径向基函数(RBF)核函数将训练样本映射到高维特征空间。在选择了最佳核函数与参数后, 即可训练整个训练样本集, 得到分类器模型。最后应用该模型对整个研究区进行分类; (5)输出分类结果图; (6)精度分析。

## 3 渭干河-库车河三角洲绿洲盐渍化信息提取

### 3.1 研究区概况

渭干河-库车河三角洲绿洲位于天山南麓, 塔里木盆地北部, 地理坐标介于 41°06'~41°38'N, 81°26'~83°17'E, 是一个典型而完整的山前冲积扇平原, 辖阿克苏地区的库车、沙雅和新和三个县。渭干河出口后呈辐射状分布, 形成东西长约 64 km, 南北宽约 160 km 的大扇裙绿洲, 地势北高南低, 平均海拔 920~1 100 m, 总面积 8 346.5 km<sup>2</sup>。属大陆性暖温带极端干旱气候, 光、热资源丰富, 多年年均降水量 51.6 mm, 水面蒸发量 1 992.0~2 863.4 mm ( $\varphi$ 20 cm 蒸发皿观测值), 温度 10.7℃, 日照 2 888.7 h, 全年无霜期 209 d。土壤以轻壤和沙壤为主, 土壤



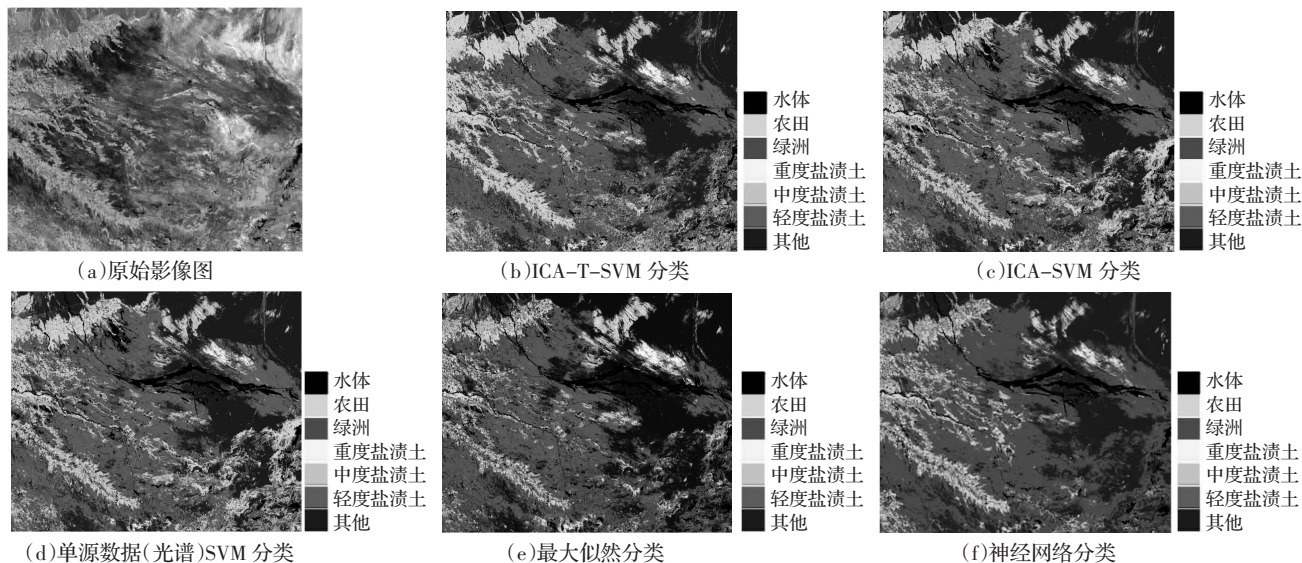


图5 原始影像图及其各种分类方法

pH 值为 7.9~8.0,土壤耕层含盐量一般为 2%左右。盐分多以氯化物、硫酸盐为主,灌区近 50%为非盐化土,主要分布在灌区上游、河道及大型灌区或有完善灌溉系统的土地上,重盐化土面积为 26%,主要分布在非农区、荒地。本区经济为绿洲经济,无灌溉则无农业。渭干河-库车河三角洲绿洲水资源利用的特点是水资源组成单一,渭干河是其唯一的灌溉水源。年径流量多年平均  $2.21 \times 10^9 \text{ m}^3$ ,变差系数 0.119,水资源利用程度高,利用效率低,灌排失调,土壤盐渍化严重。绿洲耗水组成多,生态耗水占有重要位置<sup>[38]</sup>。

### 3.2 SVM 模型选择

利用 SVM 模型对遥感图像分类首先需要考虑设置 SVM 模型参数,它包括核函数类型的选择和核函数参数的选择。Vapnik 等人的研究表明,SVM 的性能与所选用的核函数的类型关系不大,而核函数的参数和误差惩罚因子  $C$  是影响 SVM 性能的主要因素<sup>[39]</sup>。统计学习理论目前对支持向量机在实际应用中参数选择的问题给出了一些建议和解释,但还没有给出实际可行的方案。目前也只有通过实验方法来确定最佳参数。因此,在使用支持向量机进行分类和预测时,如何选择适当的参数就成为非常重要的问题。首先选择核函数,笔者选择径向基(RBF)函数作为核函数,主要原因是:首先,RBF 函数可以将样本非线性地规划到更高维的空间中,从而解决类标签和属性间非线性的关系问题,这是线性核函数无法解决的。另外,多次实验和 ROLIF 等的研究表明 RBF 核的分类精度一般高于多项式核和 Sigmoid 核,而线性核精度最低<sup>[39]</sup>。在此,多项式和 Sigmoid 内积核函数所要确定的参数比 RBF 核函数多,选取模型的工作量大。使用 RBF 核函数需要选择的参数有两个,即为核参数  $\gamma$  和惩罚因子  $C$ ,采用交叉验证算法(cross-validation)来确定这两个参数<sup>[40]</sup>。经验证明, $C$  和  $\gamma$  以指数增长方式( $C=2^{-5}, 2^{-3}, \dots, 2^{15}, \gamma=2^{-5}, 2^{-3}, \dots, 2^{15}$ )搜索是快速确定  $C$  和  $\gamma$  参数的比较好的办法。最佳 SVM 模型参数为  $\gamma=0.125, C=128$ 。这一结论与依力亚斯江·努尔麦麦提<sup>[41]</sup>所得出的结果一致。

此次研究选取库车河三角洲绿洲区域,首先通过目视解译,切割一块有代表性的且盐渍土壤面积较大的区域为实验区进行分类,去除绿洲外围的部分山地及戈壁区,最终所选区域图像位于北纬  $41^{\circ}8'14''$  至  $41^{\circ}46'39''$ ,东经  $82^{\circ}55'36''$  至

$83^{\circ}54'11''$  之间。图像大小为  $2\ 924 \times 2\ 416$  像元。

### 3.3 各种分类方法的分类结果

为了验证基于 ICA 和纹理特征的 SVM 分类方法,用同一训练样本分别采用 ICA-T-SVM 法、ICA-SVM 法、单源数据(光谱)SVM 法、最大似然法(MLC)以及神经网络法对其进行分类,原始影像图及分类结果如图 5 所示,并将分类结果与基于 ICA 和纹理特征的 SVM 分类方法相比较。各种方法分类结果精度见图 6。

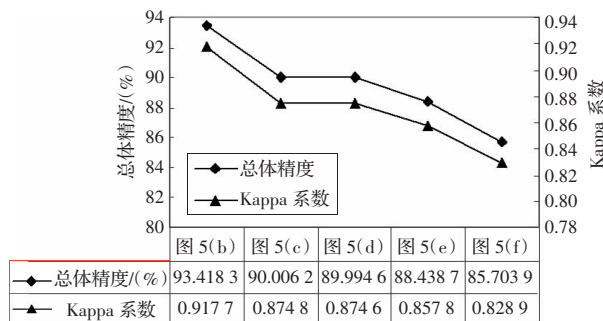


图6 各种方法分类结果比较图

### 3.4 实验结果及分析

(1)由图 5(b)可以得出:总体上盐渍地主要分布在渭干河和库车河的下游,渭-库绿洲的东和东南部地区。盐渍地的分布在绿洲内部呈条形状分布,而在绿洲外部呈片状分布,且绿洲外部重度盐渍地交错分布在中轻度盐渍地中。利用 ICA 变化特征并加入纹理特征后使得光谱信息比较接近的 3 类地物(重度、中度、轻度盐渍地)的区分性增大,从而精度有所提高。

(2)由图 5(a)~(f)可以得出:基于 ICA 和纹理特征的 SVM 和基于 ICA 的 SVM 分类结果比较理想,MLC 和神经网络分类结果相对较差。结合 ICA 和纹理特征的 SVM 分类结果图多为面状分布,单源数据(光谱)SVM 法、最大似然法、神经网络法分类地物分布细碎,导致“Pepper and Salt”现象较多。

(3)由图 6 可以得出:结合 ICA 和纹理特征的 SVM 分类法进行遥感影像分类,较基于 ICA 的 SVM 分类方法、基于单源数据(光谱)SVM 法、MLC 法、神经网络法在分类精度上有了较大提高,分类总体精度提高了 7.714 4%,Kappa 系数提高了

0.088 8。农田与绿洲,重度盐渍地与中度盐渍地、轻度盐渍地间存在一定的错分现象。主要原因是该区农田、绿洲和重、中、轻度盐渍地类型复杂,有些地区农田与绿洲,重、中、轻度盐渍地间杂分布,而使图像难以区分。不过通过该文实验可以看出:ICA 结合纹理特征的 SVM 分类方法对这些类型的分类精度较其他方法高,有效地减少了错分现象。一般来说,采用传统的分类方法区分出农田与绿洲,重度盐渍地与中度盐渍地难度较大,因为两者光谱特征相差不大,线性不可分。SVM 通过引入核函数进行非线性变换映射至高维空间,提取它们的非线性特征,增强不同类型之间的可分性,有效地提高了盐渍化信息的提取精度。

#### 4 结论与讨论

(1)阐述了基于 ICA 和纹理特征的 SVM 遥感土壤盐渍化信息提取方法,实验结果也证明了该方法的有效性。文中利用 ICA 变化特征、光谱信息、纹理信息,基于 SVM 进行图像分类,具有较高的分类精度。精度优于 ICA-SVM 分类方法、单源数据(光谱)SVM 分类方法、MLC 分类方法、神经网络分类法。分类结果总精度分别从 85.703 9%、88.438 7%、89.994 6%和 90.006 2% 提高到 93.418 3%,取得了比较好的分类结果。

(2)从图像定性分析和精度结果的定量分析来看,结合 ICA 和纹理特征的 SVM 分类法对复杂、高维输入向量具有更高的推广能力,传统的最小距离法和最大似然法由于信息量的增多,干扰信息对分类的负面影响较大,该文方法要优于传统的最大似然法和神经网络法。

(3)支持向量机基于结构风险最小化原理,它结构简单,而且各种技术性能尤其是泛化能力较好。但 SVM 算法在训练过程中并没有考虑数据之间可能存在的相关性,并且统计学习理论是针对小样本情况进行的,因而 SVM 有时对数据的分类并不是十分准确,在大样本情形下还会出现训练速度过慢的缺点。为此,在 ICA 与纹理特征的基础上提出了一种改进的支持向量机分类模型——ICA-T-SVM 模型。实验结果表明,对于样本比较多的数据集,与标准的 SVM 算法相比,ICA-T-SVM 算法不仅能够节约训练时间,同时还能够提高分类的正确率,从而为数据挖掘提供了一种有效的数据分类技术。

#### 参考文献:

[1] Cotes C, Vapnik V. Support vector networks[J]. *Machine Learning*, 1995, 20: 273-295.

[2] Bartlett P L, Taylor J S. Generalization performance on support vector machines and other pattern classifiers[M]. Cambridge, MA: MIT Press, 1999.

[3] Sholkopf B, Sung K, Burges C J C, et al. Comparing support vector machine with Gaussian kernels to radial basis function classifiers[J]. *IEEE Trans Signal Processing*, 1997, 45: 2758-2765.

[4] Vapnik V N. *Statistical learning theory*[M]. [S.l.]: Publishing House of Electronics Industry, 2004.

[5] 谭琨, 杜培军. 基于支持向量机的高光谱遥感图像分类[J]. *红外与毫米波学报*, 2008, 27(2): 123-128.

[6] Huang C, Davis L S, Townshend J R G. An assessment of support vector machines for land covers classification[J]. *International Journal of Remote Sensing*, 2002, 23: 725-749.

[7] Brown M, Lewis H G, Gunn S R. Linear spectral mixture models

and support vector machines for remote sensing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2000, 38(5): 2346-2360.

[8] 董广军, 张永生, 范永弘. 基于多特征多分辨率融合的高光谱图像分类[J]. *红外与毫米波学报*, 2006, 25(2): 123-126.

[9] Zhang J, Zhang Y, Zhou T. Classification of hyper-spectral data using support vector machine[C]// *IEEE International Conference on Image Processing*, 2001: 882-885.

[10] Sundararaghavan V, Zabarar N. Classification and reconstruction of three-dimensional microstructures using support vector machines[J]. *Computational Materials Science*, 2005, 32: 223-239.

[11] Yao Y, Marcialis G. Combining flat and structured representations for fingerprint classification with recursive neural networks and support vector machines[J]. *Pattern Recognition*, 2003, 36: 397-406.

[12] Zhang Y, Shen D. Design efficient support vector machine for fast classification[J]. *Pattern Recognition*, 2005, 38: 157-161.

[13] 宦若虹, 杨汝良. 基于 ICA 和 SVM 的 SAR 图像特征与目标识别[J]. *计算机工程*, 2008, 34(13): 24-26.

[14] 徐正光, 武楠, 穆志纯. 基于独立分量分析的人耳识别方法[J]. *计算机工程*, 2006, 32(19): 178-180.

[15] Rai Y A. Simplified approach to independent component analysis[J]. *Neural Compute & Application*, 2003, 12: 173-177.

[16] Kocsor A, Csirik J. Fast independent component analysis in kernel feature spaces[C]// *LNCS*, 2001, 2234: 271-281.

[17] Theis F. Overcomplete ICA with a geometric algorithm[C]// *LNCS*, 2002, 2415: 1049-1054.

[18] Shi Z, Tang H, Tahng Y. A fast fixed-point algorithm for complexity pursuit[J]. *Neurocomputing*, 2005, 64: 529-536.

[19] 颜梅春, 张友静, 鲍艳松. 基于灰度共生矩阵法的 IKONOS 影像中竹林信息提取[J]. *遥感信息*, 2004(2): 30-34.

[20] 孙小芳, 卢健, 孙小丹. 城市地区高分辨率遥感影像绿地提取研究[J]. *遥感技术与应用*, 2006, 21(2): 159-162.

[21] 赵萍, 傅云飞, 郑刘根, 等. 基于分类回归树分析的遥感影像土地利用/覆被分类研究[J]. *遥感学报*, 2005, 9(6): 708-715.

[22] 张锦水, 何春阳, 潘耀忠, 等. 基于 SVM 的多源信息复合的高空间分辨率遥感数据分类研究[J]. *遥感学报*, 2006, 10(1): 49-57.

[23] Comon P. Independent component analysis: A new concept[J]. *Signal Processing*, 1994, 36(3): 287-314.

[24] 周廷刚, 郭达志, 盛业华. 灰度矢量多波段遥感影像纹理特征及其描述[J]. *西安科技学院学报*, 2000, 20(4): 336-338.

[25] 杨淑莹, 胡军, 曹作良. 基于图像纹理分析的目标物体识别方法[J]. *天津理工学院学报*, 2001, 17(4): 31-33.

[26] 薛重生, 王霞. 基于分形几何的遥感图像纹理分析方法及应用[J]. *地质科技情报*, 1997, 16: 99-105.

[27] Haralick R M, Shanmugan K, Dinsrein I. Textural features for image classification[J]. *IEEE Trans on Syst Man Cybern*, 1973, 6: 610-621.

[28] Webb A R. *Statistical pattern recognition*[M]. 2nd ed. [S.l.]: Publishing House of Electronics Industry, 2004: 5-6, 106-111.

[29] Chapelle O, Haffner P, Vapnik V. SVMs for histogram based image classification[J]. *IEEE Transactions on Neural Network*, 1999, 21(6): 78282.

[30] 张锦水, 何春阳, 潘耀忠, 等. 基于 SVM 的多源信息复合的高空间分辨率遥感数据分类研究[J]. *遥感学报*, 2006, 10(1): 49-57.

[31] 张友静, 高云霄, 黄浩, 等. 基于 SVM 决策支持树的植被类型遥感分类研究[J]. *遥感学报*, 2006, 10(2): 191-196.

[32] 陈波, 张友静, 陈亮. 结合纹理的 SVM 遥感影像分类研究[J]. *测绘工程*, 2007, 16(5): 23-27.