

基于知识的遥感影像的 BAYES 分类方法研究 ——以福州市琅岐岛土地覆盖/土地利用类型为例

李春华,沙晋明

(福建师范大学地理科学学院,福州 350007)

摘要:提高计算机遥感影像的分类精度,是遥感应用中研究的主要问题之一。以福州市琅岐岛土地覆盖/土地利用类型为例,以遥感影像解译知识为基础,使用 TM、Aster 的融合影像和 NDVI 生成的植被覆盖度影像,并结合 DEM、土地利用等地理辅助数据,将 DEM 和 NDVI 因子作为待分类影像的波段加入其中,构成新的待分类影像,运用 Bayes 分类方法,通过循环迭代的方法消除先验概率对分类精度的影响,实例证明比运用单一的分类方法精度明显提高。

关键词:土地覆盖/土地利用;Bayes 分类方法;NDVI;地理辅助数据

中图分类号:TP79;F301.24

文献标识码:A

文章编号:1005-3409(2006)06-0126-03

Knowledge Based Bayes Image Classification Approach ——Land Cover/Land Use of Fuzhuo Lanqi Island as a Study Case

LI Chun-hua, SHA Jin-ming

(Institute of Geography, Fujian Normal University, Fuzhou 350007, China)

Abstract: To improve remote sensing image classification accuracy is one of the main topics in the field of the remote sensing application research. The research is based on the land cover/land use classification of Fuzhou Lanqi island, on the basic of image interpretation knowledge, using merged image of TM and Aster and the NDVI image, integrating DEM and land use geographic data, applying Bayes classification method. The method was approved to be effective to improve classification accuracy.

Key words: land cover/land use; Bayes classification approach; NDVI; geographic assistant data

近年来,基于知识的光谱特征遥感信息模型得到了快速发展,该方法是将知识作为分类的辅助信息,参与分类过程计算,可有效的改善和提高分类精度^[1]。这里所指的知识范围很广,但归纳起来可分为地学知识和专家知识。本研究中将地形参数和植被指数作为分类的辅助信息,分别将 DEM 和 NDVI 因子作为待分类影像的波段加入其中,构成新的待分类影像,对遥感影像的特征空间参数定义进行扩展,运用 Bayes 监督分类方法, Bayes 判别分类是建立在 Bayes 决策规则的基础上的模式识别,是一种典型的和应用广泛的监督分类方法,分类错误小而精度高^[2],但是实际操作中,使用者往往忽略了先验概率对分类精度的影响,本文阐述了如何在 ERDAS IMAGINE 8.7^[3]支持下,运用循环迭代法消除先验概率对分类精度的影响,并对分类精度进行评价。

1 研究区概况

琅岐岛现在是福建省福州市经济开发区,地处东海之滨,闽江入海口,位于东经 119°35',北纬 26°33'。属冲积小平原海岛,三面环江,一面靠海。琅岐岛为亚热带海洋性气候,四季温和,最高气温 28℃,最低气温 -2℃,年均降水 1 600 mm。全岛总面积 92 km²,其中陆地面积 59 km²,是福建省的第五大岛。地势中部高四周低。琅岐岛生态环境保

护较好,绿化覆盖 15.7%。岛内和周边地区均无大气、水源污染,这里的空气、土壤、水源经农业部监测认定达到生产绿 A 食品的要求,是福建省绿 A 蔬菜副食品生产基地之一。盛产缢蛏、文蛤、鸭蛋、西瓜、甜瓜、葡萄,是著名的鱼米瓜果之乡。自从 2000 年 7 月 19 日成立琅岐经济开发区以来,琅岐岛发挥其生态优势,大力开发生态观光农业和海滨生态旅游,卓有成效。

2 数据基础

本文研究的影像资料为福州幅 2002-03-07 的 TM 影像(分辨率为 30 m×30 m)和福州南半部分 2002-02-05 的 ASTER(分辨率为 15 m×15 m)影像,1 6000 琅岐岛航空相片。背景资料为 1:5 万琅岐岛土地利用现状图,琅岐岛 5 m 等高线,还有实地考察的资料。

3 数据处理

利用 ENVI 软件,对数据进行处理:

3.1 研究区相关提取

由于本文研究区只是整景影像数据的一部分,故须在 ENVI 的支持下,借助琅岐岛县界图,将之叠在配准好的影像图上作为掩膜,提取琅岐岛行政区范围的影像数据。

* 收稿日期:2005-09-13

基金项目:福建省国际合作重点项目“基于 GIS 的植被覆盖变化及其水热状况的定量遥感”

作者简介:李春华(1975-),女,福建师范大学地理科学学院硕士研究生,从事遥感与地理信息系统研究。

3.2 影像融合

利用 ASTER 数据的第 2 波段(可见光 - 近红外波段,分辨率 15 m)和 TM1、2、3、4、5、7 波段(分辨率 30 m)运用主成分分析方法进行融合,图像融合后在分辨率上提高了一倍,达到了 15 m,光谱性质也发生了变化,不同的地物表现出来的显示色彩有很大的差别,更有利于目视判读(如图 1)。这样既可以利用 ASTER 高分辨率的特点,还可以利用 TM 数据多光谱的特点,从而实现空间分辨率的提高以及波段信息的提取。根据美国查维茨 Chavez 等(1984)提出的最佳指数(OIF)的概念,可以评价图像质量好坏,融合后信息多出近 20%,融合后的光谱波形如图 2 所示。

由光谱曲线图可以看出利用主成分变换融合后的 6 个波段的数据能够较明显的区分开来,并且亮度适中,表现力强,可用融合后的影像作为待分类图像。

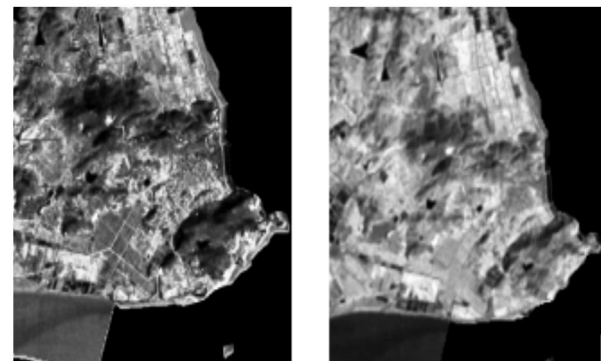


图 1 融合后的影像 融合前 TM 影像

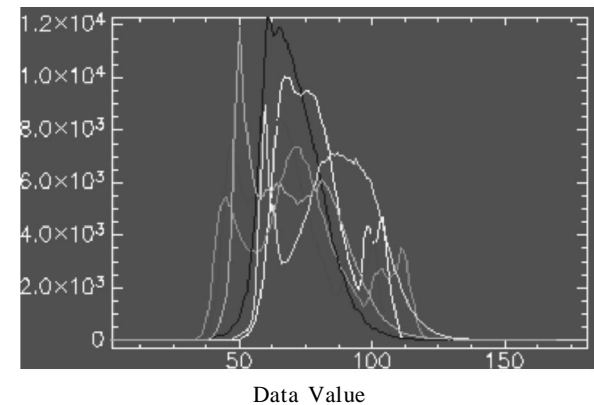


图 2 ASTER(Band2)与 TM 融合图像直方图

3.3 将 DEM 和 NDVI 因子作为待分类影像的一个波段加入融合后的影像,构成新的待分类影像

琅岐岛为亚热带海洋性气候,四季温和,植被覆盖率很高,NDVI可用于植被识别,如果值小于 0,表明地面覆盖为非植被(建筑物、水体、裸地等),另外,由 NDVI 所反映的不同的植被覆盖度^[4]并辅以 DEM 高程因子,实验证明可有效地区别林地、园地、耕地^[5]。

(1)将 DEM 高程因子作为待分类影像的一个波段加入融合后的影像,步骤如下: 把在 MapInfo 里数字化好的琅岐 5 m 等高线 tab 文件转成 e00 文件,再转成 Shapefile 文件,建立等高线之间的拓扑关系。在 ENVI 中使用 CONVERT CONTOUR TO DEM 命令,选取高程值作为生成影像的灰度值,将 DEM 转换为 15 m 格子大小的栅格图像,然后对 DEM 进行投影坐标转换,使其与遥感影像的投影坐标系统一致 使用 Layerstack 将以上 DEM 作为待分类影像的一个波段加其中。图 3 是将 DEM,BAND5,BAND4 作为红绿蓝波段合成后的影像图。

(2)将 NDVI 因子作为待分类影像的一个波段加入融合后的影像,步骤如下: 利用 ENVI 软件算出 NDVI 值为 float 型。使用 BAND MATH 命令将 NDVI 值归一化到 0 ~ 255 之间^[6],其算法如下: $NDVI_n = 255 \cdot (NDVI - NDVI_{min}) / (NDVI_{max} - NDVI_{min})$ 其中, $NDVI_n$ 表示标准化的 NDVI 值, $NDVI_{max}$ 、 $NDVI_{min}$ 分别表示 NDVI 中的最大值和最小值。使用 Layerstack 将以上 NDVI 作为待分类影像的一个波段加其中。图 4 是将 NDVI,BAND5,BAND4 作为红绿蓝波段合成后的影像图。

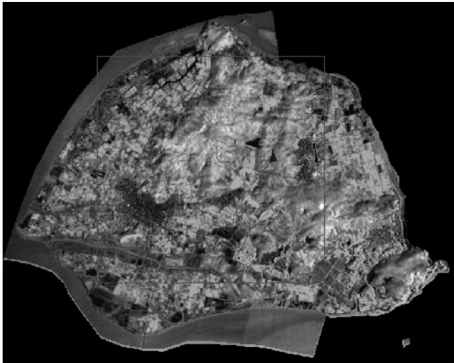


图 3 DEM,BAND5,BAND4 合成后的影像图

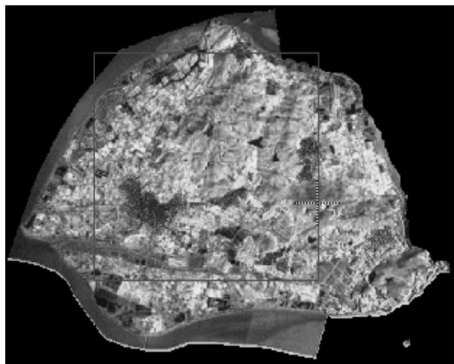


图 4 NDVI,BAND5,BAND4 合成后的影像图

4 分类方法

将叠加了 DEM 和 NDVI 因子的影像作为待分类影像。参考《欧盟土地分类》标准,辅以过度期间《全国土地分类》数据,将土地覆盖/土地利用类型划分为人工地面(城镇、交通用地、工矿用地)、农用地(耕地、果园)、林地(森林、草地、绿化用地)、水体(内陆水体、海洋水体)、湿地(盐田、滩涂)五个一级类别。以外出考察的样本类型为依据,并在影像上标出考察的区域,辅以航空相片和土地利用现状图,确保选取训练区样本类型的准确性以及精度的需要。

以下着重阐明如何运用 Bayes 分类方法对影像进行分类,通过循环迭代法减少先验概率对分类精度的影响。

4.1 Bayes 判别原理^[7]

Bayes 判别分类是建立在 Bayes 准则基础上,通过计算各个像元属于各类的概率,将该像素归属于概率最大的类。设待分类遥感图像 X 由 m 个多谱图像组成,则遥感图像 x 的像元 x 的值可由 m 维向量表示:

$$X = [x_1 \quad x_2 \quad \dots \quad x_m] \tag{1}$$

若研究地区可分为 G 类,则任意像元必来自其中的某一类。当各类总体为多元正态总体 $N(u_g, \sigma_g)$ 时,像元特征向量 X 在 g 类的概率密度为:

$$f_g(X) = \frac{1}{(2\pi)^{m/2} |\sigma_g|^{1/2}} \exp \left[-\frac{1}{2} (X - u_g)^T \sigma_g^{-1} (X - u_g) \right]$$

式中, u_g 与 σ_g 第 g 类总体的均值向量和协方差矩阵, $g = 1, 2, \dots, G$ 根据 Bayes 公式, 在 X 出现的条件下, 其归属第 g 类的归属概率为:

$$P(g/X) = [P_g f_g(X)] / \sum_{k=1}^G [P_k f_k(X)] \quad (2)$$

式中: $P_i (i = 1, 2, \dots, G)$ 为第 i 类出现的先验概率, 对于已知试验区而言, 先验概率 p_i 可以根据各类估计面积 A_i 作近似计算 ($P_i = A_i / \sum_{k=1}^G A_k$)。

显然, $P(g/X)$ 越大, 像元来自 g 类的概率就越大, 所以 $P(g/X)$ 表示归属于 g 类的概率, 并称为像元 X 的归属概率。

在多类判别时, Bayes 判别式为:

当 $P(g^*/X) = \max_{1 \leq g \leq G} P(g/X)$ 则 $X \in g^*$ 也就是说, 根据像元 X 求出 $P(g/X) (g = 1, 2, \dots, G)$ 后, 若 $g = g^*$ 时 $P(g^*/X)$ 极大, 则 X 归属于 g^* 类。

4.2 先验概率对分类精度的影响

Bayes 判别公式 (2) 中所需要的参数如均值、协方差矩阵可以通过样本数据估计, 但先验概率必须在分类前确定, 而在分类前研究者无法确切知道各地类占总面积的比例。本文提出的循环迭代法过程如下: 用相同的训练区进行多次分类, 第一次分类时各类别先验概率赋值为 1, 即忽略先验概率, 以后每一次分类都用上一次分类所得到的面积比率对先验概率赋值, 这样每一次先验概率对分类精度的影响应该都比上一次小, 迭代直到相邻二次分类各类面积的变化率接近 0 或分类精度稳定时, 即认为近似消除了先验概率对分类结果的影响。

各类别先验概率赋值为 1, 等同于假设各类别分布面积相等。循环迭代法能够进行的关键在于第一次分类后的结果比这一假设更符合实际的类别分布。kappa 指数可解释为: 若某一分类结果 kappa 指数为 0.7, 表示这种分类方法优于随机赋予各点一类别的方法的 70%。随机赋予各像元点一类别, 分类结果必然是各类别的面积相等。因此, 若第一次的分类结果 kappa 指数大于 0, 则表示该分类结果与各类别分布面积相等的假设更符合实际的类别分布, 就可以运用循环迭代法来消除先验概率对 Bayes 分类结果的影响。

并且, 笔者对几种常用的图象处理软件作了研究, 发现 ERDAS 软件中完全提供了先验概率的设置以及多次循环迭代的实现: 在默认的情况下, ERDAS 分类模版 (signature editor) 设置各类别的先验概率都为 1, 我们可以通过选择 use probabilities 选项, 重新设置分类模版的 Prob. 值。Prob. 值可通过分类影像的 Attribute 属性表求得。

4.3 以琅岐岛为例进行分类精度评价

4.3.1 检验区采用实地调查的采样数据

在 Erdas 软件的支持下, 利用 Bayes 最大似然判别法, 首先在各类别先验概率赋值为 1 的前提下, 检验 DEM 高程因子和 NDVI 因子对分类精度的影响, 表 1 和表 2 分别为选择 1, 2, 3, 4, 5, 7 波段和 1, 2, 3, 4, 5, 7 波段、DEM、NDVI 因子对影像进行分类的混淆矩阵、总体精度和 Kappa 系数。从表 1、表 2 的分类结果显示, 将 DEM 和 NDVI 因子作为待分类影像的波段加入其中构成新的待分类影像, 对遥感影像的特征空间参数定义进行扩展, 分类总体精度由 91.015% 提高到 94.5817%, Kappa 系数由 0.8761 提高到 0.9250, 分类精度明显提高。

4.3.2 对迭代后的分类精度进行检验

利用表 2 的分类结果的后验概率作为下一次迭代的先验概率, 重复迭代 4 次后, 发现总体精度和 Kappa 系数趋于稳

定。表 3 为五次分类结果的总体精度和 Kappa 指数, 表 4 为最后一次迭代的分类结果的精度和混淆矩阵。比较分类的结果, 总体精度提高了 0.888%, Kappa 系数提高了 0.009。

表 1 1, 2, 3, 4, 5, 7 波段参与分类的结果

混淆矩阵					
总体精度 = (18576/20080) 91.0159 %					
Kappa 系数 = 0.8761					
地面实况 (像元素)					
分类	水体	湿地	人工地面	林地	农用地
水 体	2750	66	0	0	0
湿 地	98	2063	80	56	0
人工地面	7	24	3017	197	88
林 地	15	0	122	8334	251
农用地	0	0	339	459	2112
合 计	2870	2155	3558	9046	2451

表 2 1, 2, 3, 4, 5, 7 波段、DEM、NDVI 因子参与分类的结果

混淆矩阵					
总体精度 = (18992/20080) 94.5817 %					
Kappa 系数 = 0.9250					
地面实况 (像元素)					
分类	水体	湿地	人工地面	林地	农用地
水 体	2750	15	0	0	0
湿 地	93	2107	82	34	1
人工地面	1	32	3182	252	215
林 地	26	0	11	8754	36
农用地	0	1	283	6	2199
合 计	2870	2155	3558	9046	2451

表 3 五次分类结果的总体精度和 Kappa 系数

第 1 次分类	1	2	3	4	5
总体精度	94.5817	95.0132	95.3475	95.4682	95.4697
Kappa 系数	0.925	0.9294	0.9327	0.934	0.934

表 4 分类的最后结果及精度检验

混淆矩阵					
总体精度 = (19171/20080) 95.4697 %					
Kappa 系数 = 0.9340					
地面实况 (像元素)					
分类	水体	湿地	人工地面	林地	农用地
水 体	2761	22	0	0	0
湿 地	105	2124	42	74	1
人工地面	4	34	3292	252	215
林 地	16	0	17	8757	26
农用地	0	0	280	4	2236
合 计	2870	2155	3558	9046	2451

5 结 语

(1) 遥感信息融合技术, 可以充分利用不同数据源的空间和波谱特性, 是一种非常有效的信息增强方法。本研究中使用 TM 和 ASTER 影像的融合, 增强了光谱信息, 空间分辨率提高了一倍。

(2) 利用辅助信息, 建立基于知识的遥感影像分类方法, 可以有有效的改善和提高土地利用信息的分类精度。从分类的精度验证表明, 在琅岐岛土地利用信息提取中, DEM 高程因子和 NDVI 植被指数作为待分类的影像的一个波段, 有效的扩展了遥感影像的特征空间, 提高了分类精度。本研究中仅提取 DEM 的高程因子, 可以满足土地利用一级分类精度的要求。后续的工作中将进一步研究 DEM 的坡度、坡向因

(下转第 132 页)

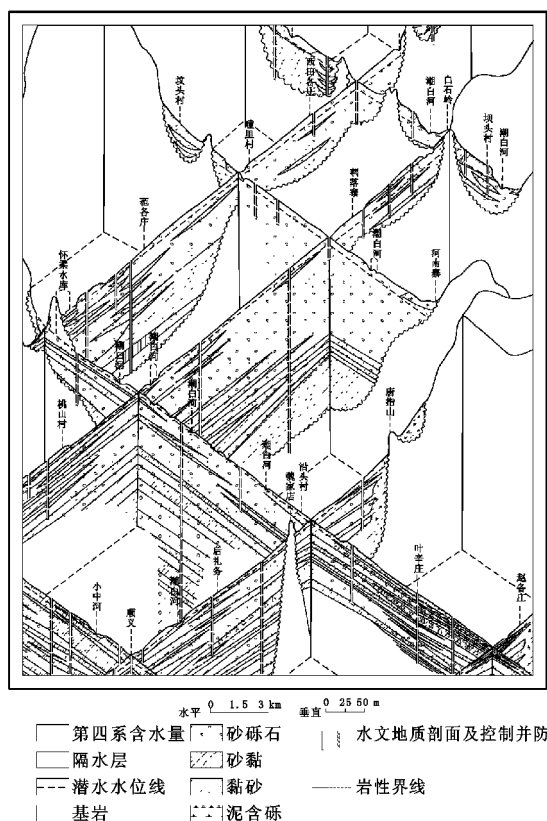


图 3 潮白河流域调蓄水文地质联合剖面图

潮白河 1981 年断流引起了地下水补、径条件的变化,因此将 1981 年 12 月与 2000 年 12 月地下水平均水位对比,计算出调蓄库容量为 $5.997 \times 10^8 \text{ m}^3$ 。

参考文献:

- [1] Hauser, V L, Lotspeich, F B. Artificial ground water recharge through wells[J]. Soil Water Conserve, 1967, 22(1): 11 - 55.
- [2] 北京市用水调研课题组. 北京市用水调研与需求预测研究报告[R]. 2002.
- [3] Pyne, R D G, Garcia - Bengochea, J I. Aquifer storage Recovery: A New Water Supply alternative[A]. Artificial Recharge of Ground Water - Proceedings of the International Symposium[M]. New York: the American Society of Civil Engineers, 1989. 107 - 121.
- [4] 崔亚莉, 谢振华, 邵景力, 等. 北京市平原区地下水合理开发利用数值模型研究[R]. 北京: 中国地质大学, 2003.
- [5] 刘予, 贾三满, 孙颖, 等. 南水北调(北京段)环境地质问题调查评价报告[R]. 北京市水文地质工程地质大队, 2004.
- [6] 惠士博, 谢森传, 张思聪. 南水北调来水后北京的水资源优化配置[J]. 中国水利, 2001, 1: 14 - 15.
- [7] 武晓峰, 唐杰. 地下水人工回灌与再利用[J]. 工程勘察, 1998, (4): 37 - 39.

(上接第 128 页)

子对二级分类精度的影响,以及植被指数与坡度、坡向之间内在联系。

(3) 运用循环迭代法对影像进行基于 Bayes 判别的最大似然法分类,保证了错分损失最小,可消除先验概率对分类结果的影响,提高分类总体精度和 Kappa 系数。

参考文献:

- [1] 杨存建,周成虎. 基于知识的遥感图像分类方法的探讨[J]. 地理学与国土研究, 2001, 17(1): 72 - 77.
- [2] 章孝灿,黄智才,赵元洪. 遥感数字图像处理[M]. 杭州:浙江大学出版社, 1997. 224 - 228.
- [3] 党安荣,等. ERDAS IMAGINE 遥感图像处理[M]. 北京:清华大学出版社, 2003.
- [4] 朱云燕,朱翔,李卓卿. 归一化差值植被指数在土地覆盖遥感动态调查中的应用[J]. 云南环境科学, 2003, (4): 9 - 10.
- [5] 术洪磊. 基于知识的遥感影像分类与制图综合方法研究[J]. 北京:北京大学, 1995.
- [6] 江振蓝,沙晋明. 福州市植被生态的遥感分析[J]. 福建地理, 2003, 18(3): 3 - 6.
- [7] 梅长林,周家良. 实用统计方法[M]. 北京:科学出版社, 2002.

潮白河调蓄区已具备的入渗场地包括河道、砂石坑、渠道和水库,所有入渗场地都可通过不同途径接受水库放水。根据入渗场地的地理位置不同和入渗能力大小,本次研究中将该地区的联调工程划分为三部分:潮白河调蓄工程、应急备用水源地调蓄工程和潮河总干渠调蓄工程。

永定河、潮白河冲洪积扇地下水调蓄模型预测结果表明,地下水人工调蓄效果明显,调蓄工程可行,对地下水具有积极的恢复作用,具有重大的社会和环境效益。

并且,提出联合调蓄方案建议:在南水北调水源进京之前,以应急开采地下水,来弥补地表水库供水量的不足;南水北调进京后,应建立外来水与本地水、地表水与地下水的联网调蓄 - 供水工程。

为此,应加强对南水北调工程实施后地表水与地下水联合调蓄工程与调蓄方案的实施研究,并征求有关部门的意见,修改完善后列入北京市水利工程规划,力争分阶段逐步实施。

4 结 语

目前,我国已有 300 多个大中型城市面临供水不足的问题。水资源的危机将成为个世纪制约我国经济发展和人民生活水平提高的重要因素^[7]。水是一种可再生的自然资源,但是当水的自然循环过程受到破坏,它也会一去不返的离开我们。因而,应该尊重自然,保护自然,改掉过去的攫取式的资源开采方式,树立利用与养蓄并重的新的发展观。水资源的养蓄,即是对过去水资源开发利用实践的总结,也是一个全新的概念,它需要我们从水的循环过程、资源的可持续利用、城市生态建设需求、规划与管理、人文和社会经济建设前景分析等多方面出发,考虑水资源在新的历史条件下的开发利用方式。笔者只是就自己在水资源研究工作中获得的一些浅见,不当之处,还望各界专家指正。

(4) 本文将遥感影像的光谱特征和地理辅助信息相结合进行研究,在后续的工作中将进一步研究用分形和神经网络方法来描述图像的纹理结构特征,采用图像的纹理结构特征和影像的光谱特征相结合的方法进行分类研究。