

基于 TM 影像的水体信息提取^{*}

徐金鸿, 陈淑娟

(重庆交通大学 土木建筑学院测绘系, 重庆 400074)

摘 要:在分析各种地物在 TM 遥感影像中特性的基础上, 采用决策树方法对影像进行分类。通过分析水体的影像特征, 找出水体与其他地物之间的灰度差异, 进而用关系表达式来实现水体信息的提取。

关键词:水体; 遥感; 决策树分类

中图分类号: TP79; P331

文献标识码: A

文章编号: 1005-3409(2008)06-0161-03

Extraction of Information About Water Body Based on TM Image

XU Jin-hong, CHEN Shu-juan

(School of Civil Engineering & Architecture, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: In this research, the decision tree classification was used to TM image. The water body information was extracted, the image characteristic of the water body is analyzed and the digital number difference is discovered between the water body and the other thing.

Key words: water body; remote sensing; the decision tree classification

水无论作为一个独立的环境因子, 还是作为一种资源, 都被人们格外的重视。水体的面积监测是调查水资源的一个重要方面, 也是洪水灾害检测的重要内容^[1]。近 20 年来随着科学技术的飞跃, 水体信息获得的方法也得到了很大的发展。早期的遥感水体提取方法主要是利用目视解译^[2], 人工目视解译可充分利用判读人员的知识, 擅长提取空间相关信息, 但花费较多时间, 并存在个人差异。目前基于计算机技术的各种水体识别方法相继得到了应用。刘建国利用比值法提取 MSS 影像中的水体信息^[3]。陆家驹等分别利用阈值法、色度判别法、比率测算法从 TM 影像中识别水体^[4]。陆家驹采用山地阴影间接判读方法对 TM 影像和航片进行了流域水系的目视判读, 发现 TM 影像对流域水系分析非常有效^[5]。刘建波等利用密度分割法从 TM 影像中提取水体的分布范围, 但是对于山区的水体来说, 会将阴影也提取出来^[6]。因此利用遥感影像的波谱特征和水体在影像上的特性, 研究自动化提取水体的方法, 获得它的范围甚至其他的特征将是十分有意义的工作。本文采用广州市 Landsat 5 TM 卫星遥感数据作为基础信息源, 时相为 1998 年 10 月 3 日。利用决策树分类方法来进行地物分类, 并提取水体信息。

1 决策树算法原理

决策树算法可以像分类过程一样被定义, 依据规则把遥感数据集一级级往下细分以定义决策树的各个分支。决策树由一个根结点 (Root nodes)、一系列内部结点 (Internal

nodes) (分支) 及终极结点 (Terminal nodes) (叶) 组成, 每一结点只有一个父结点和二个或多个子结点。如果把自然界中的地物看成一个原级 T (根结点), 开始考虑地物的分类时, 首先可以考虑组分, 将原级 T 分为 T1 (植被) 和 T2 (土壤) 两大类, 称为“一级”分类, 进而每大类中又可再进一步分类, 如 T1 只用于进一步地有效细分类。这就是决策树分类器特征选择的基本思想。事实上, 决策树分类器的特征选择过程不是由“原级”到“终极”的顺序过程, 而是由“终极”到“原级”的逆过程, 即在预先已知“终极”类别样本数据的情况下, 根据各类别的相似程度逐级往上聚类, 每一级聚类形成一个树结点, 在该结点处选择对其往下细分的有效特征。依此往上发展到“原级”, 完成对各级各类组的特征选择。在此基础上, 再根据已选出的特征, 从“原级”到“终极”对整个影像实行全面的逐级往下分类。对于每一级处的特征选择, 可以依据散布矩阵准则来进行, 但为了使一个可分离性准则既能用于特征选择, 又适用于聚类, 可使用 Bhattacharyya 距离准则 (简称 B - 距离准则) 来表征两个类别 W_i 和 W_j 之间的可分性, 其表达式为

$$D_{ij} = \frac{1}{8} (M_i - M_j)^T \left[\frac{i+j}{2} \right]^{-1} (M_i - M_j) + \frac{1}{2} \ln \frac{0.5(M_i + M_j)}{|M_i|^{\frac{1}{2}} \cdot |M_j|^{\frac{1}{2}}}$$

式中: M_i, M_j —— W_i 和 W_j 类集群的均值向量; i, j —— W_i 和 W_j 的协方差矩阵; D_{ij} —— W_i 和 W_j 之间的 B - 距离。

* 收稿日期: 2007-12-24

基金项目: 重庆市教委科研项目 (KJ080418); 重庆交通大学校内科研基金资助

作者简介: 徐金鸿 (1976 -), 男, 江西新干人, 副教授, 主要从事 3S 技术研究。E-mail: xujh814@163.com

据上述分类思想,以样本数据为对象逐级找到分类树的结点,并在每个结点上记录所选出的特征影像编号及相应判别函数的参数,从而有可能反过来顺着从“原级”到“终级”的过程,按:若 $D_{ij} > 0$,则 x 属于 w_i ;否则, x 属于 w_j 的判别规则,逐级地在每个结点上对样本数据以外的待分类数据进行分类,这便是决策树分类法的原理。从中可以看到,判别函数的确定经常是与特征选择密切相关的。一旦分类结束,不仅各类之间得到区分,同时还确定了各类的类别属性^[7]。

2 水体在 TM 遥感影像中的特性

要想从影像中提取水体,了解水体在 TM 影像中的特性是很有必要的,这可以为我们提取水体提供有利的参考依据。从宏观的角度看,陆地上水体主要表现为湖泊、河流、沟渠、水库、池塘等,可分为面状水体和带状、线状水体。在卫星遥感影像上,湖泊、水库、池塘大致表现为面状区域,河流、沟渠表现为线状结构。在 TM 图像上,对于水体来说,水体几乎全部吸收了近红外和短波(中)红外波段内的全部入射能量,所以水体在近红外和短波(中)红外波段的反射能量很少,而植物、土壤在这两个波段内的吸收能量较少,而且又具有较高的反射特性,这就使得水体在这两个波段上与植被、土壤有明显的区别。水体在这两个波段上呈现出暗色调,而土壤、植被则呈现出相对较亮的色调。但是,在山区,由于山体阴影的影响,使得近红外、短波(中)红外波段在阴坡面的反射能量特别低,从而同样造成它们在图像上呈现出明显的暗色调。水体与阴影的混淆使得难以在这两个波段上利用阈值法来提取水体。在可见光波段内,图像上记载的反射信息主要来自水面、水中悬浮物质和水体底部物质的反射。关于水体的吸收和辐射特性,不仅与其本身的性质有关,而且还与其所含的物质类型和大小有关。一般地,进入清澈水体的大部分太阳光在水下 2 m 内被吸收。吸收的程度取决于波长。近红外波段在水体 1/10 m 深处就被吸收,使得在近红外影像上,即使是浅水也是暗的影像色调。可见光波段的吸收会因为研究水体的特征变化而变化。从透过水体拍摄水底的详细程度来看,透射力最好的是在 $0.48 \sim 0.60 \mu\text{m}$ 之间的波段。在清澈、平静的水体中,此波段范围可透射 20 m。虽然蓝色波段透射力好,但它们散射厉害且有“水下模糊”的现象,而红色波段只能透射几米。

3 基于 TM 的水体信息提取

3.1 水体及背景地物的灰度值分析

TM 中的第四波段(近红外),植物在该波段有强烈的反射,水体有强烈的吸收,容易区分植被、水体、土壤。TM 中第 5 和第 7 波段属于中红外,但是第 5 波段提供的光谱信息最丰富,完全可以代替第 7 波段。TM 影像中水体信息的提取主要是依据水体在 7 个波段上光谱的不同特征以及其他地物与水体的区别,通过分析水体及背景地物的灰度值,可以只用单个波段或多个波段组合来提取。在遥感影像上就测定每一种地物其各波段的灰度值,利用文中的遥感影像并从中抽取一些典型而具有代表性的灰度值数据列于表 1。从表 1 可见:在第 2 和第 3 波段上,水体与植被的灰度值相接近,与其他地物

的灰度值有区别;在第 4 波段上,水体与其他地物的灰度值有区别;在第 5 波段上,水体与其他地物的灰度值有明显的区别。此外水体在第 2 和第 3 波段上,水体的灰度值大于阴影。将这两个波段相加可以增大这种差异。在第 4 波段和第 5 波段上,阴影的值一般都大于水体,将这两个波段相加,可以增大这种差异。将波段 2 与波段 3 相加,波段 4 与波段 5 相加,并做出改进后的地物波谱,可以看出,只有水体具有波段 2 加波段 3 大于波段 4 加波段 5 的特征。

表 1 1998 年的 TM 影像典型地物灰度均值

波段	水体	植被	沙	阴影	建筑物
TM1	75	72	111	67	88
TM2	30	31	60	26	39
TM3	29	28	83	23	46
TM4	16	64	77	50	39
TM5	10	58	137	43	64
TM6	135	138	150	134	155
TM7	5	20	85	14	40

表 2 TM 影像典型地物灰度均值

波段	水体	林地	建筑物	草地	道路
TM1	78	65	87	70	103
TM2	33	29	39	29	57
TM3	32	25	46	26	89
TM4	17	124	40	52	83
TM5	10	89	69	48	154
TM6	134	137	157	138	152
TM7	5	24	44	17	97

3.2 水体信息的提取

根据分析,判别条件为当灰度值 $(\text{TM2} + \text{TM3}) > (\text{TM4} + \text{TM5})$ 时,利用对水陆界线反映较好的 TM 的第 5 波段,即中红外波段,通过反复实验,最后确定其阈值为 17,水体信息基本准确提取出来了(见图 1)。用 TM4,3,5 作假彩色合成影像,用目视判读的方法,检验阈值法提取水体的效果。所提取的珠江和湖泊的轮廓与目视判读的一致,而较窄的河流没有被提取出来,但是这些河流由目视判读是很难判读出来的。

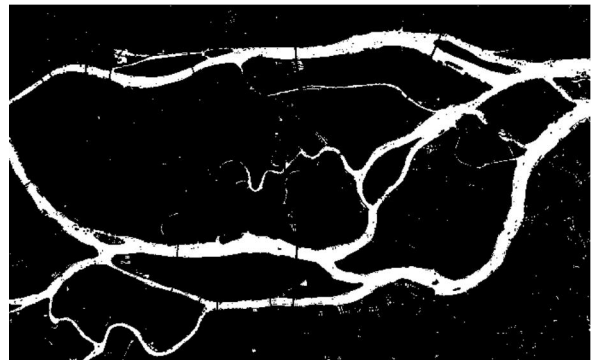


图 1 水体信息提取(图中白色为水体)

3.3 决策树分类

根据“中国资源环境数据库”土地利用遥感分类体系将土地利用类型分为耕地、林地、草地、水域、建设用地、未利用地 6 类,并充分考虑到实验区实际情况,利用影像丰富的波谱信息,对影像中 5 种不同的地物:水域、林地、建筑物、草

地、道路的灰度值进行提取列于表 2 中,据此建立决策树分类模型。

水体的提取可以直接采用灰度值判别式 $(TM2 + TM3) > (TM4 + TM5)$ 作为判断条件提取水体,而道路的波谱特征在波段 2、波段 3 和波段 5 上的波谱特征与其他地物区别明显,通过试验确定了 $(b2 > 60)$ and $(b5 > 90)$ 为判断条件,提取出道路。同样的方法通过分析林地、草地、建筑物的波谱值,然后反复试验确定判断条件,把林地、草地、建筑物都提取出。

4 结论

针对广州市的 TM 遥感影像,分析该区域水体与背景地物的在遥感影像上的灰度值特征,采用灰度值判别式 $(TM2 + TM3) > (TM4 + TM5)$ 进行水体提取,而利用决策树对遥感影像进行了分类。通过目视解译,分类效果较好。

(上接第 160 页)

5.2.3 非法采砂

由于经济利益的驱动,近年来北京河道内外的非法采砂现象愈演愈烈。据北京矿业协会不完全统计,目前北京市和周边地区取得采矿证的各种砂石生产企业共 100 多家,年产砂石约 7000 万 t,但非法采砂点的数量远远高于合法采砂企业,仅北京地区就有非法采砂点 200 处以上,年产砂石量达 2000 多万 t。潮白河河床已被挖得千疮百孔,沙坑星罗棋布,毁坏了地表植被,破坏了土壤结构,直接造成沙化土地面积的扩大。

6 结论

基于 RS 和 GIS 技术,针对北京潮白河流域沙化的具体情况,利用 SPOT5 卫星遥感图像,结合植被覆盖度、土壤类型、土地利用类型这 3 种辅助数据,对研究区的沙化强度进行了遥感图像解译,并对其沙化现状及成因进行了初步分析和探讨。结果显示,植被覆盖度为 30% ~ 60% 的沙化土地面积较大;绝大部分沙化土地为林地、草地和未利用地这 3 种土地利用类型;沙化是自然和人为双重作用下的结果。研究成果为北京潮白河流域沙化防治和生态环境改善提供了依据。

参考文献:

- [1] 国家林业局. 全国荒漠化监测主要技术规定[Z]. 1998.
- [2] 贺然,徐伟,王棒,等. 北京市典型沙地景观动态变化分析:以康庄为例[J]. 水土保持研究, 2007, 14(5): 145-148.
- [3] 程复,尚润阳,丁国栋,等. 北京土地沙化现状分析和综合治理[J]. 水土保持研究, 2006, 13(6): 89-91.

参考文献:

- [1] 周成虎,骆剑承,杨晓梅,等. 遥感影像地学理解与分析[M]. 北京:科学出版社, 2001.
- [2] 杨晓平. 基于 TM 遥感图像的流域地貌研究[J]. 科技通报, 2003, 19(2): 151-153.
- [3] 刘建国. 陆地卫星 MSS 图像地表水域信息的机助识别提取[J]. 环境遥感, 1989, 4(1): 19-26.
- [4] 陆家驹,李士鸿. TM 资料水体识别技术的改进[J]. 环境遥感, 1992, 7(1): 17-23.
- [5] 陆家驹. 多种遥感资料河网水体的有效解译[J]. 水利学报, 1993(1), 41-45.
- [6] 刘建波,戴昌达. TM 图像在大型水库库情监测管理中的应用[J]. 环境遥感, 1996, 11(1): 53-58.
- [7] 李爽,张二勋. 基于决策树的遥感影像分类方法研究[J]. 地域研究与开发, 2003, 22(1): 18-21.
- [4] 李海龙. 风沙灾害危险性评价及防沙治沙安全格局构建研究[D]. 北京:北京林业大学, 2006.
- [5] 曾永年,冯兆东. 沙质荒漠化遥感监测与环境影响研究进展[J]. 山地学报, 2005, 23(2): 218-227.
- [6] 北京市潮白河管理处. 潮白河水旱灾害[M]. 北京:中国水利水电出版社, 2004: 3-32.
- [7] 陈浩,周金星,陆中臣,等. 荒漠化地区生态安全评价:以首都圈怀柔县为例[J]. 水土保持学报, 2003, 17(1): 58-62.
- [8] 齐善忠,罗芳,肖洪浪. 干旱内卢河流域土地利用变化及其对环境的影响:以西北地区黑河流域为例[J]. 生态环境, 2006, 15(4): 757-760.
- [9] 陈建平,王功文,丁火平,等. 北京荒漠化演化模拟与评价[M]. 北京:地质出版社, 2004: 12-51.
- [10] Zhang Kebin, Yang Xiaohui. Desertification Assessment Indicator System in China[J]. Forestry Studies in China, 2002, 4(1): 44-48.
- [11] 北京市林业局. 北京市第三次土地荒漠化和沙化监测报告[R]. 2005.
- [12] 陈广庭. 北京平原土壤机械组成和抗风蚀能力的分析[J]. 干旱区资源与环境, 1991, 5(1): 103-113.
- [13] 北京市水务局. 2005 年北京市水资源公报[N]. 2005.
- [14] 王建,李文君,宋冬梅,等. 近 30 年来民勤土地荒漠化变化遥感分析[J]. 遥感学报, 2004, 8(3): 282-288.
- [15] 张国祯. 北京市沙化土地现状评价及其防治策略研究[D]. 北京:北京林业大学, 2007.