

扎龙湿地沼泽的提取模型研究

韩 敏¹, 程 磊¹, 刘 权²

(1. 大连理工大学电子与信息工程学院自动化系, 大连 116023; 2. 水利部松辽水利委员会水文局, 长春 130021)

摘要: 利用扎龙湿地 TM 图像, 通过分析各类地物的光谱特征、谱间结构特征以及采样剖面线, 发现了扎龙湿地沼泽的光谱特征, 并由此建立了一个基于知识发现的沼泽提取模型。提取结果表明, 建立的沼泽提取模型能够实现自动提取扎龙湿地的芦苇沼泽, 其精度可以达到 98%。

关键词: 沼泽提取; 光谱模型; 知识发现; 扎龙湿地; TM 图像

中图分类号: TP 75 文献标识码: A 文章编号: 1001-070X(2004)01-0043-04

0 引言

湿地具有涵养水源, 蓄洪防灾, 补充地下水等多种重要的生态功能。传统上, 湿地资源要素指标的获取是通过野外实地调查方法, 这些工作费时、费力, 且往往受人为因素的影响。遥感技术的发展为获取湿地资源要素指标提供了新的途径。利用遥感技术进行湿地资源调查具有客观、准确、经济和高效等特点, 大大提高了工作效率和调查的精度, 促进了湿地的研究与进展。

湿地研究的关键在于定量化获取和分析湿地信息。多光谱分类技术为专题信息提取提供了先进的手段^[1,2]。与目视判读相比, 多光谱分类技术因把识别问题转化为光谱特征空间的定量求解而具有客观和高效的优点。但这种方法建立在数理统计之上, 而没有建立在遥感知识的挖掘和发现上, 在分类中所获取到的知识通常不可传递, 也难以解释, 而训练样区的重复选取也限制了遥感信息提取的自动化。目前, 信息提取的方法已有一定的发展^[3], 知识已被用于信息提取^[4,5]过程中, 但是基于知识信息提取方法的发展仍不成熟。

本文采用基于知识发现的信息提取方法, 通过分析扎龙湿地 TM 图像各类地物的光谱特征, 提取沼泽的光谱特征知识, 建立扎龙湿地沼泽提取模型。

1 扎龙湿地及其数据

扎龙自然保护区位于黑龙江省西部松嫩平原, 总面积 2 100 km², 是一个以鹤类等大型水禽为主的珍稀鸟类和湿地生态类型的国家级自然保护区, 在世界鹤类保护及湿地生态系统保护事业中占有重要位置, 1992 年被列入国际重要湿地名录。扎龙湿地不仅为鹤类等珍稀鸟类提供了栖息地, 而且对调节黑龙江省西部地区干旱气候、平衡降雨、蓄水、分洪等方面也发挥了重要作用。

本文采用覆盖扎龙湿地的 1/2 景 TM 图像研究扎龙湿地的土地覆被分类。图像时相为 2001 年 10 月 5 日, 图像大小为 4 871 × 2 871 像素。图像中心点的地理坐标为 46°57'N, 124°15'E。地面分辨率为 30 m。插页彩片 14 为 TM4 TM3 TM2 假彩色合成图像。通过目视判读, 结合土地利用现状图, 确定扎龙湿地 TM 图像区域内主要包括 8 大地类, 分别是沼泽、居民地、林地、旱地、水域、草地、河滩地和盐碱地。在假彩色合成图像上, 许多像元都表现为纯净像元。如沼泽和湖泊像元。但由于用地形状不规则, 仍有许多像元表现出混合像元的特征, 这给判读带来一些困难。插页彩片 14 的图像获取时间为 10 月份, 此时图像上沼泽呈现墨绿色, 紫红色及黑色, 与 8、9 月份呈现红色的沼泽相比, 10 月份的沼泽较易与林地和草地等地物区分。

2 沼泽的自动提取

2.1 沼泽光谱知识的发现

通过对图像各类地物进行光谱特征分析,可以发现沼泽光谱知识,从而将沼泽与其它地物区分开来。基于多光谱特征分类器的总体分类原则在于:

(1)对于多变量图像,不能孤立地根据个别变量的数值进行分类,而是要从像元点在高维特征空间中的位置及聚类情况出发进行分类。

(2)分类的实质是把多维特征空间划分为若干区域,每个区域相当于一类。分类或划分区域范围标准差的标准为:由每类的统计特征出发,研究它所归属的区域。

本文以每类均值向量作为中心,把在标准范围内的点归入一类。首先对图像上的不同地物进行灰度值采样,然后对所有采样灰度值数据进行均值统计,结果如表 1 所示。

表 1 各地物采样波谱灰度值的均值

地物类型	TM1	TM2	TM3	TM4	TM5	TM7	样本数/个
沼泽	61	46	54	54	67	47	1 000
居民地	79	67	78	64	75	68	500
林地	62	53	56	84	68	43	500
旱地	75	66	83	86	111	84	800
水域	71	57	56	33	14	12	800
草地	70	56	66	66	93	72	600
河滩地	73	61	72	69	109	94	800
盐碱地	119	116	142	117	150	140	200

利用这些数据可以绘出各地物的光谱曲线图(图 1)。

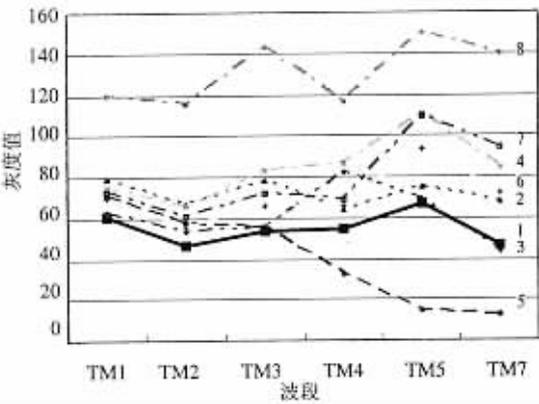


图 1 各地物的光谱曲线

1—沼泽; 2—居民地; 3—林地; 4—旱地;
5—水域; 6—草地; 7—河滩地; 8—盐碱地

2.1.1 光谱灰度均值特征分析

根据表 1 和图 1,在各波段上根据灰度值对各地物进行排序,以分析各类地物的灰度值在每个波段上的差异。

物进行排序,以分析各类地物的灰度值在每个波段上的差异。

(1)在 TM1 波段上,各地物灰度值依次为盐碱地>居民地>旱地>河滩地>水域>草地>林地>沼泽,沼泽易与林地混淆。

(2)在 TM2 波段上,各地物灰度值依次为盐碱地>居民地>旱地>河滩地>水域>草地>林地>沼泽,沼泽易与林地混淆。

(3)在 TM3 波段上,各地物灰度值依次为盐碱地>旱地>居民地>河滩地>草地>林地>水域>沼泽,沼泽易与林地和水域混淆。

(4)在 TM4 波段上,各地物灰度值依次为盐碱地>旱地>林地>河滩地>草地>居民地>沼泽>水域,沼泽易与居民地和水域混淆。

(5)在 TM5 波段上,各地物灰度值依次为盐碱地>旱地>河滩地>草地>居民地>林地>沼泽>水域,沼泽易与林地和水域混淆。

(6)在 TM7 波段上,各地物灰度值依次为盐碱地>河滩地>旱地>草地>居民地>沼泽>林地>水域,沼泽易与居民地和林地混淆。

在 TM1~TM3 波段上,沼泽具有最低光谱值。在 TM1 和 TM3 波段上,沼泽和林地及水体的光谱值接近,而在 TM2 波段沼泽和其它地物有一定的差距,所以沼泽具有 TM2 波段的光谱值低于某个值的特征。

2.1.2 光谱谱间结构特征分析

根据表 1 和图 1,分别对每个地物在各波段上的灰度值进行排序,其结果如下:

- 沼泽 TM5>TM1>TM4>TM3>TM7>TM2
- 水域 TM1>TM2>TM3>TM4>TM5>TM7
- 居民地 TM1>TM3>TM5>TM7>TM2>TM4
- 林地 TM4>TM5>TM1>TM3>TM2>TM7
- 旱地 TM5>TM4>TM7>TM3>TM1>TM2
- 草地 TM5>TM7>TM1>TM3>TM4>TM2
- 盐碱地 TM5>TM3>TM7>TM1>TM4>TM2
- 河滩地 TM5>TM7>TM1>TM3>TM4>TM2

从图 1 和以上分析可以看出,最容易与沼泽混淆的是林地和水域。但观察光谱曲线发现:

(1)林地与沼泽光谱差异最大的波段是 TM4,其次是 TM2 波段。

(2)水域与沼泽光谱差异最大的波段是 TM5 和 TM7 波段。

(3)水域和林地的光谱都具有 TM4 波段>TM5 波段的特征,而沼泽并不具备。

(4)沼泽的光谱具有 TM3 波段与 TM4 波段很

接近的特征,林地和水域却并不具备。

因此利用这些差异可将林地,水域和沼泽区分开来。

2.1.3 采样剖面线分析

除了光谱灰度均值特征分析和光谱谱间结构特征分析方法之外,还可以采用采样剖面线分析的方法分析光谱特征知识。采样剖面线分析方法可以获得更直观的光谱特征知识,尤其是沼泽内部以及沼泽与其周围地物的光谱知识。图 2 为沼泽及其周围地物的光谱采样剖面线。

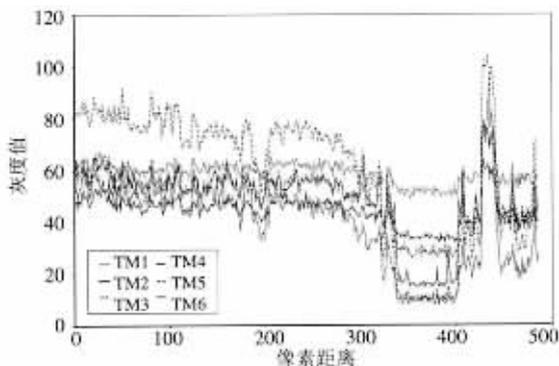


图 2 沼泽的采样光谱剖面线

从图 2 中可以看出,沼泽内部光谱特征变化不大,但沼泽与周围旱地的光谱值有比较明显的差异,它和周围水体的光谱值在某些波段上有明显差异。水体在 TM_4 波段、 TM_5 波段和 TM_7 波段上灰度值都比较低,并且具有 $TM_4 > TM_5$ 的特征,而沼泽具有 TM_3 波段与 TM_4 波段很接近的特征。这个结论验证了前面光谱谱间结构特征分析得出的结论。

2.2 沼泽提取模型

根据以上的光谱分析,本文利用阈值法提出基于光谱知识的沼泽提取模型。沼泽提取模型为

$IF(TM_4 - TM_3 < K_1) \text{ and } (TM_2 < K_2) \text{ and } (TM_4 \leq TM_5)$

THEN 为沼泽

式中, K_1, K_2 为选取的阈值。第一个表达式 $TM_4 - TM_3 < K_1$ 是根据以上分析 TM_3 与 TM_4 很接近的特征得到的,该表达式的含义是,如果 $TM_4 - TM_3$ 的值小于某个设定值 K_1 ,那么该像元就为沼泽;第二个表达式 $TM_2 < K_2$ 是根据在 TM_2 上沼泽具有最低的光谱值这个特征得到的,该表达式的含义是,如果 TM_2 的光谱值小于某个设定值 K_2 ,那么该像元就为沼泽;第三个表达式 $TM_4 \leq TM_5$ 是用于剔除水体和林地的,该表达式表示沼泽一定满足该条件 $TM_4 \leq TM_5$ 的关系,而水体和林地不满足该条件,从而将其剔除。利用上述模型,可以达到将沼泽从图像上提取出来的目的。

3 沼泽提取结果分析

3.1 歪取结果

模型中需要设定一些阈值,它是通过不断地试验来确定的。考虑到地物光谱特征差异的相对性,阈值一般也是相对量。

首先,利用表达式 $TM_4 - TM_3 < K_1$ 和 $TM_2 < K_2$ 对沼泽进行提取。经过不断试验,发现当 K_1 为 8, K_2 为 52 时可以将沼泽比较好地提取出来。当 K_1 取 7、6 或者 K_2 取 51、50 时,提取结果变化不大,当 K_1, K_2 取更小值时,就会有漏提的现象出现;当 K_1 取 9、10 或 K_2 取 53、54 时,提取结果变化仍然不大,而当 K_1, K_2 取更大值时,就会有多提的现象出现。在利用这两个表达式提取出来的结果中,有一些误提的水体和林地。利用第 3 个表达式 $TM_4 \leq TM_5$ 可以将水体和林地剔除。最后,在 GIS 的支持下,进一步剔除误提的地物。利用沼泽提取模型得到的沼泽提取结果如图 3 所示。



图 3 提取的沼泽

3.2 提取精度分析

利用目视判读方法对提取结果进行精度分析。精度分析包括漏提误差和误提误差的分析。漏提误差可以通过在假彩色合成图像上随机选取样本,将其目视判读结果与沼泽提取结果图上进行比较而得出;误提误差则可以通过在沼泽提取结果图上随机选择样本,再与假彩色合成图像上相应像素的目视判读结果比较得出。所以,首先,在假彩色合成图像

上随机选择 1 000 个样本点,检验发现有 39 个没有提取,漏提误差为 3.9%;然后在沼泽提取结果图上随机选择 1 000 个样本点,检验发现有 7 个点误提为沼泽,误提率为 0.7%。

综上所述,这种基于知识的提取方法是可行的。利用本文所提出的沼泽提取模型对扎龙湿地的芦苇沼泽进行提取是有效的,其精度可以达到 98%。

4 结 论

本文提出的沼泽提取模型能够有效地提取扎龙湿地的芦苇沼泽。通过分析扎龙湿地的 TM 图像灰度均值特征、谱间结构特征及采样剖面线,发现了扎龙湿地沼泽的光谱特征知识,即沼泽满足第 4 波段和第 3 波段的光谱值非常接近,在第 2 波段上具有最低的光谱值以及第 4 波段光谱值不大于第 5 波段光谱值的特征,并由此建立了一个基于知识发现的沼泽提取模型。提取结果表明,建立的沼泽提取模型能够实现自动提取扎龙湿地的芦苇沼泽,并且能

够很好的剔除与沼泽特征相近的林地和水体,提取精度可以达到 98%。

参 考 文 献

- [1] Ronald E McRoberts, Daniel G Wendt, Mark D Nelson, et al. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates[J]. Remote Sensing of Environment, 2002, 81(1): 36-44.
- [2] 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.
- [3] Joo-Hyung Ryu, Joong-Sun Won, Kyung Duck Min. Water-line extraction from Landsat TM data in a tidal flat: A case study in Gomsu Bay, Korea[J]. Remote Sensing of Environment, 2002, 83(3): 442-456.
- [4] 杨存建, 周成虎. 基于知识发现的 TM 图像居民地自动提取研究[J]. 遥感技术与应用, 2001, 16(1): 1-6.
- [5] 周成虎, 骆剑承, 杨晓梅, 等. 遥感图像地学理解与分析[M]. 北京: 科学出版社, 2001.
- [6] 方余兴, 孙振华. 上海市湿地资源调查中的遥感技术应用[J]. 国土资源遥感, 2001, (3): 25-27.

MARSH INFORMATION EXTRACTION BASED ON KNOWLEDGE DISCOVERY

HAN Min¹, CHENG Lei¹, LIU Quan²

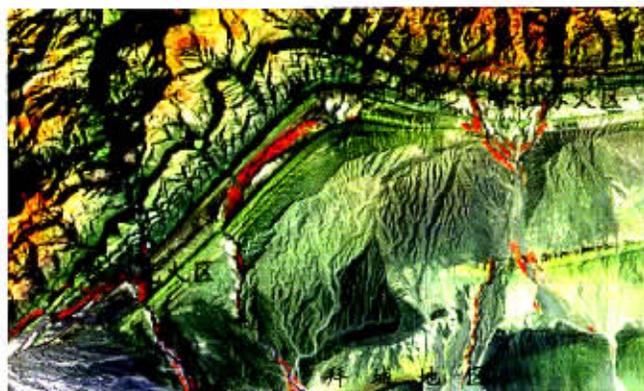
(1. School of Electronic and Information Engineering, Dalian University of Technology, Dalian, 116023, China; 2. Hydrology Bureau of Songliao Water Resources Committee, Changchun, 130021, China)

Abstract: This paper deals with the way of extracting reed marsh information in Zhalong wetland based on the knowledge of spectrum characteristics. Spectral analyses of TM image covering Zhalong wetland were conducted, which included analysis of spectrum average value characteristics, analysis of characteristics of different spectrum structures and analysis of the sampling profile. This led to the discovery of the spectrum character knowledge of the marsh in Zhalong wetland. With the results obtained, a model of marsh extraction based on knowledge discovery was established. The marsh information extraction shows that the model proposed in this paper can extract reed marsh in Zhalong wetland automatically, and the extraction accuracy can reach 98%.

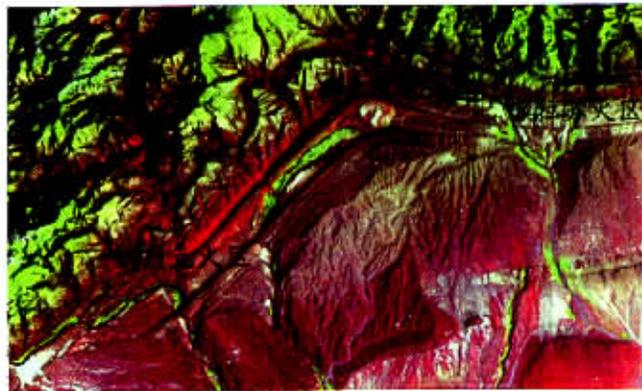
Key words: Marsh extraction; Spectrum model; Knowledge discovery; Zhalong wetland; TM image

第一作者简介: 韩敏(1959—),女,工学博士。大连理工大学电子与信息工程学院副教授。1982年毕业于大连理工大学电子系。1992年至1999年留学日本,获日本国立九州大学工学博士学位。目前承担国家自然科学基金重点项目(50139020)3S信息系统部分。参加国家重点基础研究发展计划(973)项目研究专题,及教育部科学技术重点项目等5个课题。研究方向为:神经网络,3S系统及混沌序列分析等。共发表论文60余篇。其中被SCI、EI、ISTP检索论文23篇次。

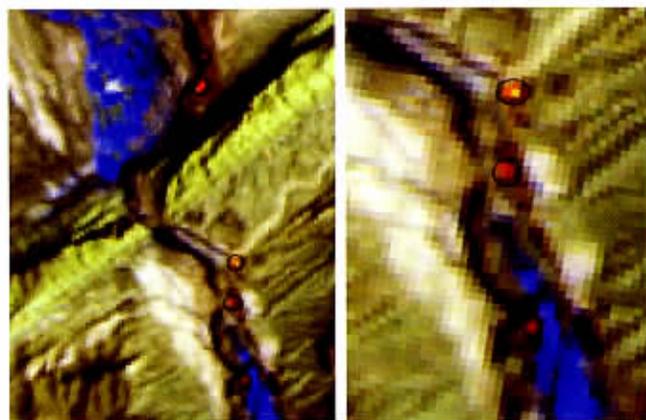
(责任编辑:肖继春)



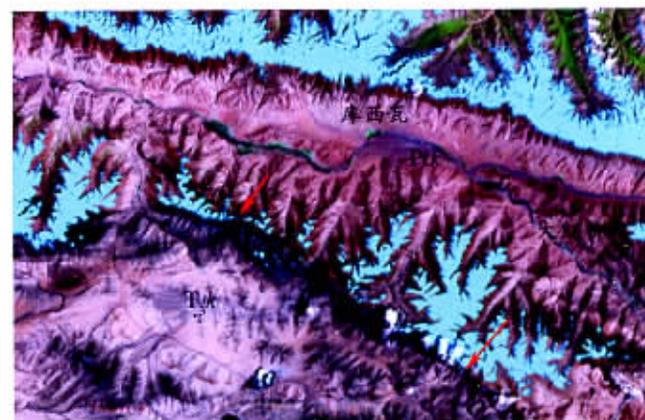
彩片9 拜城地区火区 TM4TM5TM2 假彩色合成图像



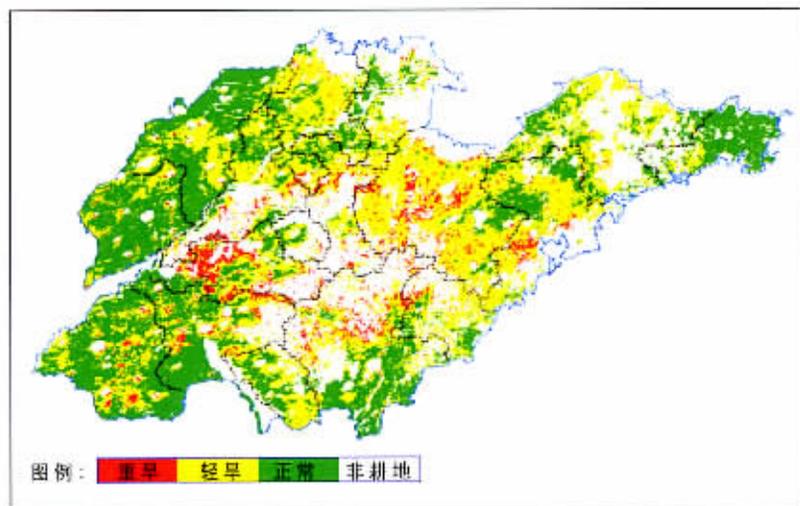
彩片10 拜城地区火区 TM6TM4TM3 假彩色合成图像



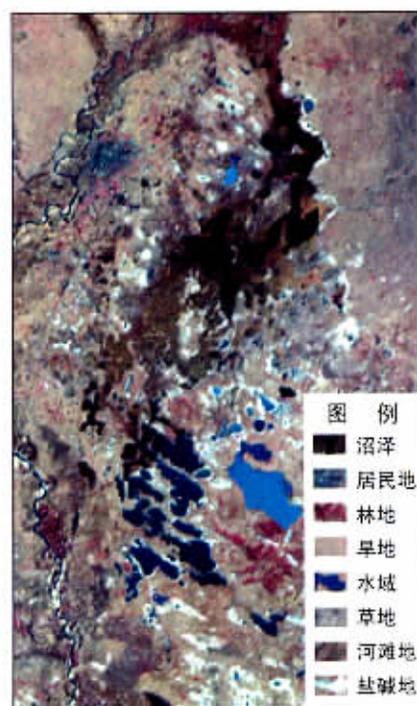
彩片11 TM7TM5TM4 假彩色合成图像
(左: 铁列克火区局部放大图; 右: 进一步放大图)



彩片12 原克勒青河群(T₂k)中的深变质岩(P₂k)影像
(箭头所指为三十里营房-郭扎错断裂)



彩片13 气象卫星遥感监测山东旱情分布图
(山东省气象中心)



彩片14 TM4TM3TM2 假彩色合成图像

(彩片12见刘刚等一文; 彩片13见赵玉金等一文; 彩片14见韩敏等一文)