

应用遥感及地理信息系统进行植被制图*

吴炳方 黄 绚 田志刚

(中国科学院地理研究所资源与环境信息系统国家重点实验室 北京 100101)

摘 要 利用卫星遥感数据制作复杂地形环境的植被图面临的最主要问题是精度, 单纯对遥感数据(TM 或 SPOI) 进行监督或非监督分类的精度低于 50%。本文选择美国亚利桑那州 Santa Catalina 山脉的 Pusch Ridge 作为研究区, 分析地理信息系统模型在改善植被分类精度中的作用。结果表明, 通过结合辅助数据和应用地理信息系统模型, 其精度可以从 37.41% 提高到 71.67% (SPOT 数据, 非监督分类), 或从 50.07% 提高到 61.50% (TM 数据, 监督分类)。同时表明用 SPOT 数据进行山区植被制图的效果好于 TM 数据。

关键词 植被制图, GIS 模型, 辅助数据

1 概 述

复杂地形环境下的卫星影像分类精度可通过利用辅助数据和由影像提取空间要素来改善 (Richard, 1982)。分类过程可通过地理信息系统集成辅助数据, 并用空间分析方法提高分类精度。

本研究是试验 GIS 空间及光谱分析模型在提高卫星数据分类精度中的作用。另外比较二种卫星数据在多山地区植被制图中的效果 (SPOT 和陆地卫星)。

2 研究区

我们选择美国亚利桑那东南部的 Santa Catalina 山脉的 Pusch Ridge (亦为科罗拉多国家森林区) 作为研究区, 进行植被制图, 并试验 GIS 在辅助遥感数据分类中的作用。Pusch Ridge 面积 230.65km², 从东到西长 25km, 向北伸展约 15km。高程从 Lemmon 山顶的 2791m 到西南部图森市附近的 853—975m。山峰急剧隆起, 并被深谷切割。

3 数据采集及预处理

对 1988 年 5 月 7 日的 SPOT HRV 卫星 CCT 磁带数据作了辐射校正和大气校正,

* 这项研究是世界实验室资助培训计划的一部分。这项培训计划是瑞士世界实验室 ICSC 资助黄河下游实时洪水监测和洪水管理系统项目中的一部分内容。培训计划于 1992 年元月 15 日至 1992 年 5 月 15 日由美国亚利桑那大学、干旱地区研究办公室和亚利桑那遥感中心执行。作者感谢 Stuart E. Marsh 博士和 Jiang Li 博士(亚利桑那遥感中心)给予的巨大支持和指导。感谢 Charles F. Hutchinson 教授和 Ellen Schuster 女士给予的帮助。感谢 Robert A. Clark 教授(水文和水资源系, 亚利桑那大学)提供机会, 为我们安排参加这项非常有益的学术培训计划。

收稿日期: 1994 年 6 月 21 日; 收到修改稿日期: 1994 年 8 月 15 日

它的有关数据从 CCT 磁带头文件中查得 (CNEC & SPOT Image, 1988)^[2] XS1、XS2、XS3 3 个波段的绝对校正参数为 0.81801、0.86965 和 0.97075。改进的黑体技术用于大气散射校正 (Chavez, 1988)^[11]。数据获得时的大气条件非常好, 所以其相关散射模型为(波长)⁻⁴。

1989 年 5 月 19 日的陆地卫星 TM CCT 磁带数据覆盖了大部分研究区。影像云量极少, 但未进行关于大气影响及光谱方面的校正。使用两种影像的目的是比较 TM 及 SPOT 影像在多山地区植物制图中的应用效果。

SPOT 和 TM 影像的几何纠正是从各自的 33 和 22 个地面控制点组中选出的 15 个和 16 个点来完成, 误差分别为 0.83 和 0.81RMS。由于 SPOT 多光谱影像的地面空间分辨率为 20 米, 因而用了双向线性插值的方法进行重采样, 其输出像元的大小为 30 米 × 30 米。对 TM 影像则采用了最近邻点方法。SPOT 影像的非监督分类是在几何纠正前完成的, 并对其分类结果图使用相同的 GCPS, 用最近邻点法纠正到 UTM 坐标系中。

辅助数据包括地形图、行政边界图、植被图(由航片上解释得到), 道路/公路图及水系图, 所有的图件的比例尺均为 1:24000 并用 Auto CAD 数字化输入计算机中。地形高程模型 (DEM) 是通过用 AutoCAD 数字化 6 幅 1:24000 地形图然后通过 ARC/INFO 编辑归并转换至 IDRISI 格式, 并在 IDRISI 中进行 30 米 × 30 米空间分辨率的插值得到 (图版 1 图 1)。10° 差的坡度及坡向数据从 DEM 中求得。其坡向用 0 至 360° 顺时针方向标准方位角命名。DEM、坡度和坡向数据按给定的间距重新进行划分, DEM 按 61m 的间距划分成 34 级, 坡度按 10° 间距划分为 9 级; 坡向划分为 16 级。边界图则被转换成位图以便通过多次复合操作提取 Pusch Ridge 研究区的像元。还计算了离河流距离的数据层 (图版 1 图 2)。

收集了植被及其种类的野外数据及其地形条件等。Pusch Ridge 地区被划分成 6 种土被类型: 河岸植物、混合针叶林、美国红松、栎树林、Sonoran 沙漠草场和沙漠灌木 (见表 1) (Whittaker, 1965)^[7,10]。从航片上读得的植物图用于以下两方面: (1) 做为最大似然分类训练区的参考资料; (2) 做为试验点确定分类的精度。由于土被类别来自大量不同的植被和地貌特征, 绝对的制图精度难以估算。

表 1 植被序列的高程分布 (m)

Tabel 1 Elevation range (in meters) of plant communities

植物种群	北 坡	南 坡
沙漠灌木	792—1158	792—1402
沙漠草地	1036—1524	1158—1829
栎树林	>1097	>1219
美国红松	>1524	>1768
混合针叶林	>1585	
岩石		
河岸植物	792—1829	

有两种分类方法被用来进行卫星影像制图。第一种是直接应用 ERDAS 系统中的非监督最大似然分类, 分类后的 GIS 文件被转换为 IDRISI 格式, 用 GIS 模型程序提

高分类精度。第二种方法是利用 ERDAS 系统中的监督最大似然分类器。所有的精度评价均使用 Kappa 一致性指数进行整体精度评价,而条件 Kappa 一致性指数被定为特定类别的精度指标 (Rosenfiel, 1986)^[4]。

4 非监督分类

利用 ERDAS 系统进行的非监督分类分为两步:非监督训练和最大似然法分类。训练用来揭示数据原始模式的参数。

ERDAS^[4] 系统提供了四种非监督训练方法——序列法、统计法、ISODATA 法和 RGB 聚类法。本项研究采用了统计法,这种方法考虑相邻像元的相似特点。需要输入的参数是 L 和 U ——标准差的上下限。 V 变差系数, N 最大集群数, S 最小光谱距离, W 某个集群允许的最小窗口数。 L 和 U 分别定为 0.1 和 1.2, V 必须是 2.0,因为 SPOT 数据三个波段的最小值为 42—80, N 和 S 的确定需要一些试错分析(见表 2)。评价标准是所有 SPOT 三个波段特征空间的 TM 距离和总的分类精度。由表 2 可见,如果 S 等于 2.0,16 个集群具有最大的离散度,当 S 减小时集群的数目增加,16、23 和 29 个集群的光谱特性参数被用来进行最大似然法分类,分类后的像元归并成 6 种植被类型。整体精度计算是通过分类图像与 IDRISI 中的数字植被图进行的正交制表。23 个集群的分类图像得到了最高的精度——37.41% Kappa 指数,但是仍很不够。因为植被类型的分布与高程、坡向、坡度等相关,地理信息系统模型程序被用来提高分类精度,这些程序用 23 个集群分类的图像作为基本图像。

表 2 集群数的试错分析

Table 2 Trial-and-error analysis of cluster number

N	30	30	20	30	40	50	60	90	90
S	3	2.4	2.0	2.0	2.0	2.0	2.0	1.0	0.5
集群数	6	14	10	17	16	17	21	23	29
平均 JM	1341	1358	1376	1370	1393	1385	1387	1389	1412
最小 JM	1006	914	1099	966	1083	962	807	788	989
总的 Kappa(%)	33.16							37.41	37.03

程序 1 这个程序将非监督分类的 23 个集群合并为 4 种植被类型: Sonoran 沙漠灌木,沙漠草场,林地和混合针叶林。对于 23 个集群中的 i 集群,第一步首先我们利用 IDRISI 中的 ASSIGN 将 i 集群从分类图中分离出一个位图。例如我们可以对集群 2 使用 ASSIGN 的值文件(Eastman, 1992)^[3] 来建立位图。同样北面或南面覆盖的位图可以通过坡向图来建立。将这两个位图通过乘法运算叠加从而得到属于集群 i 并且是南坡或北坡的覆盖。第二步:北坡覆盖的位图与数字高程模型和植被指数 SAVI 相乘,得到集群 i 像元在北坡高程和 SAVI 值。第三步:利用 CROSSTAB 来建立高程和 SAVI 值在北坡的正交相关图像,所形成的新的类别说明了集群在北坡高程和植被指数的集合。第四步:显示正交相关图像并将新类别与地面真实情况进行比较,划出各类别与植被类型

关系的规律。第五步: 再次利用 ASSIGN, 根据有关规律将正交相关图像的分类再归类到植被类型。第六步: 重复第二到第五步, 处理集群 i 在南坡出露的象元。第七步: 以上六个步骤对于所有 23 个集群重复进行。最后, 利用加法运算将所有再分类的图像合成临时植被图像。

包括八个步骤的上述程序对于提高 Sonoran 沙漠灌木, 沙漠草场, 林地和混合针叶林的分类精度是有效的。精度从 37.41% 提高到 66.57%。表 3 给出了个别类型的条件 Kappa 值。其中第 4 列中的美国红松包含了混合针叶林的像元。河岸植物没有从其它植被类型中分出来。

表 3 SPOT 数据非监督分类和 GIS 模型的精度分析(%)

Table 3 Overall and conditional kappa agreement index of unsupervised classification and GIS modelling programs with SPOT data (Percent)

植被类型	非监督	程序 1	程序 2/3	小区
沙漠灌木	45.61	74.53	74.25	74.35
沙漠草地	17.84	60.48	65.02	65.12
栎树林	51.77	72.55	75.39	79.57
美国红松	37.00	68.93	67.93	70.12
混合针叶林	0.0	8.47	39.42	33.33
河岸植物	0.0	0.01	65.27	57.16
总的 Kappa	37.41	66.57	70.35	71.67

程序 2 这个程序用于把混合针叶林从红松中分离出来。再次利用 ASSIGN 建立红松位图。红松位图与坡向和坡度图像相乘以记录红松的坡度和坡向属性。通过建立正交相关图像得到所有存在红松像元的坡度和坡向的类型组合, 通过显示来确定再分类的标准。

程序 3 这个程序用以从其它植被类型中提取河岸植被像元。由于在 Pusch Ridge 地区所有植被类型中都夹有河岸植被, 只是大部分布在主要峡谷河流附近。水系数据被用于计算每一像元到最近的水系的距离。生成的距离图像被再分类从而产生新的位图。例如, 距离值小于 3 个像元的像元赋值 1, 而大于 3 个像元的赋值 0。因为沿水系生长的河岸植被的宽度小于 200 米。将位图与 SAVI 数据相乘即得到沿水系所有像元的 SAVI 值。SAVI 值大于 0.1 的像元被取为河岸植被像元。

使用以上 3 个程序之后, 植被图的总体精度增加到 70.35%。从植被图中, 我们发现区域的东南发生了严重误分类, 因此在这一子区再次使用程序 1 从而使整体精度增至 71.67%, 与仅仅使用 SPOT HRV 非监督分类方法得到的 37.41% 的精度比较这一结果是令人满意的(图版 1 图 3)。

5 监督分类

监督分类的目的是比较 SPOT 和 TM 在 Pusch Ridge 地区对山区和沙漠景观植被的分类能力。植被图被作为地表真实状态, 选择可靠的代表不同植被类型的样区, 用

ERDAS 系统中的 DIGSCRN 显示, 屏幕上确定了 6 种植被类型的样区位置。描述这些区域的多边形作为一个独立的文件存贮, 从而确定位于多边形内像元的 SPOT 和 TM 数据的特征。

我们采用了最大似然法, 各类的先验概率由地面真实情况取得。3 个波段 SPOT 数据、4 和 6 个波段 TM 数据的分类精度分别是 36.87%, 49.33% 和 50.07% (表 4), 6 个波段 TM 数据的分类误差最小, 但仍不令人满意。

表 4 SPOT 和 TM 的监督分类
Tabel 4 Supervised classification of SPOT and TM data

	数据	波 段	平均 JM	最小 JM	总的 Kappa
1	SPOT	XS1, XS2, XS3	1256	861	36.87%
2	TM	TM1, TM2, TM3, TM4	1284	627	49.33%
3	TM	TM1, TM2, TM3, TM4, TM5, TM7	1306	714	50.07%

对现有的 SPOT 和 TM 图像增加波段可以提高监督分类的精度。增加的波段包括滤波后的图像, 植被指数, 朝向和水系的距离, 所有辅助波段的数据成为每一像元的附加特性。利用 10 个, 13 个和 16 个 SPOT, TM 和 TM/SPOT 波段数据分别得到了三个新图像(表 5), 虽然 ERDAS 允许最多对 16 个波段的数据进行同时分类, 但是减少数据项有利于节省计算机的运行时间并提高精度, 因为多项数据层并不意味着高的精度。用 JM 距离得到最佳分类的波段子集。表 5 是 8, 10 和 12 波段数据子集。它们的统计距离在表 6 列出。表 6 还列出了总体 Kappa 指数。SPOT 数据比 TM 数据分类精度高。但差别不大。坡度数据对提高分类精度的作用很小。但它对个别类型的分类精度有所提高。例

表 5 TM, SPOT 和 TM/SPOT 的波段组合
Tabel 5 TM, SPOT and TM/SPOT image bands

NO.	SPOT	TM	TM/SPOT
波段 1	SPOT XS1*	TM1!	TM1#
波段 2	SPOT XS2*	TM2*	TM2
波段 3	SPOT XS3	TM3*!	TM3*#
波段 4	SPOT XS3 High Pass	TM4*!	TM4*#
波段 5	SPOT XS3 Low Pass*	TM5*!	TM5*#
波段 6	SPOT XS3 Texture*	TM7	TM4 Low Pass
波段 7	SPOT SAVI*	TM4 High Pass!	TM4 Texture#
波段 8	Elevation Model*	TM4 Low Pass*!	TM SAVI*#
波段 9	Aspect*	TM4 Texture!	SPOT XS2*#
波段 10	Distance*	TMSAVI*	SPOT XS3 High Pass
波段 11		Elevation*!	SPOT XS3 Low Pass*#
波段 12		Aspect!	SPOT XS3 Texture#
波段 13		Distance*!	SPOT SAVI
波段 14			Elevation*#
波段 15			Aspect#
波段 16			Distance*#

注: * 8 个波段组合用; ! 10 个波段组合用; # 12 个波段组合用

如,混合针叶林和红松植被的条件 Kappa 指数在用坡度图代替 XS3 的低通滤波图像后精度分别增加了 1% 和 0.3%。

表 6 波段组合分析

Table 6 Divergence and overall Kappa agreement index of datasets

	数据	波段数	平均 JM	最小 JM	总的 Kappa%
1	SPOT	8	1393	1207	63.10
2	SPOT	10	1395	1215	63.40
3	TM	8	1403	1335	60.12
4	TM	10	1405	1336	61.50
5	TM/SPOT	8	1406	1358	61.75
6	TM/SPOT	12	1409	1365	61.51
7	SPOT + Slope	10	1398	1260	63.82

6 分类结果

所有分类都使用了最大似然法分类方法,不同点在于采用不同的训练样区和波段组合。Kappa 指数的总精度在 37.41% (最佳非监督分类)和 63.4% (监督分类采用 10 个波段 SPOT 和辅助数据)之间。表 7 说明了以条件 Kappa 指数为指标的各类精度。

表 7 分类精度(%)

Table 7 Conditional KIA of individual class (Percent)

植被类型	SPOT 非 监督+模型	SPOT 10 个 波段监督	SPOT + 坡度 10 个波段	TM 监督 10 个波段	TM/SPOT 监督波段	SPOT 监 督+模型
沙漠灌木	74.35	81.12	76.81	69.25	69.94	83.89
沙漠草地	65.12	33.72	38.68	41.26	40.80	52.30
栎树林	79.57	88.57	87.29	85.84	87.78	81.94
美国红松	70.12	54.93	56.17	50.34	51.40	55.12
混合针叶林	33.33	25.29	25.35	24.70	17.34	26.30
河岸植物	57.16	33.34	27.51	11.24	16.05	34.22
总的 Kappa	71.67	63.40	63.82	61.50	61.75	68.72

应该注意到,无论采用什么方法,沙漠草场、混合针叶林和河岸植被的分类精度都低于其它类。因为沙漠草场介于沙漠灌木和林地之间,草场易与位于高程低的灌木和高程高一些的林地混淆。混合针叶林和红松分布在高山地区,它们都具有类似的植被特征和分布模式,因此很难从红松植被系列中将混合针叶林分离出来。

利用 3 个 SPOT 波段进行监督最大似然法分类得到了 36.87% 的总精度。增加辅助数据——高程、坡向和水系距离使精度提高到 63.40%。TM 数据总精度由 50.07% 增加到 61.5%,用辅助数据之后,SPOT 的精度比 TM 要高,原因不详,可能与他们的空间和统计特性直接相关。例如 TM 没有覆盖整个 Pusch Ridge,高程模型数据与多变量常规分布不一致,TM 数据没有经过辐射和大气纠正。另一方面,在 SPOT 数据的几何纠正过程中采用了双线性插值重采样方法。表 7 中最后一项显示了通过使用地理信息模型程

序对 10 个波段的 SPOT 进行监督分类的结果,精度为 68.72%(图版 1 图 4),与原来的分类精度比较,增加了 5.32%。这说明了一个重要事实:GIS 模型程序对提高分类精度的效率直接与集群数相关,间接与光谱特性相关。

3 个波段 SPOT 数据的最大似然法非监督分类得到了 37.41% 的总精度。应用了 GIS 模型程序后,总的 Kappa 指数增加到 71.67% (见表 3)。

7 结 论

我们开发了三个 GIS 模型程序来处理山区和沙漠景观地区遥感数据,这些程序同时被用来在分析中有效地综合辅助数据,这种方法基于自然环境和植被种属分布模式之间的紧密联系。通过使用这些程序,总精度增加了 34%。

将辅助数据结合到监督分类中之后,精度由 36.87% 显著增至 63.40% (SPOT),对于 TM 则从 50.07% 增至 61.50%。但 SPOT, TM 和辅助数据的结合没有取得预期的结果。

最终图像的总精度达到 71.6%,这需要非监督分类方法与 GIS 模型程序的综合应用。SPOT 数据的监督分类与辅助数据和 GIS 模型综合得到的精度为 68.72%, TM 数据为 61.50%。

GIS 模型程序能够显著地提高多山地区的卫星数据分类精度。但 GIS 模型程序的开发是一种耗时的工作并需要实际野外工作经验。

参 考 文 献

- [1] Chavez, Pat S. An Improved Dark-Object Substration Technique for Atmospheric Scattering Correction of Multispectral, Remote Sens. Environ. 1988, 24: 459—479.
- [2] CNEC & SPOT Image, SPOT User's handbook, Vol. 1 Reference Manual; Vol. 2 SPOT Handbook 1988.
- [3] Eastman, J. Ronald, IDRISI User's Guide, Version 4.0, Clark University, Ma. USA, 1992.
- [4] ERDAS Inc., ERDAS Field GUIDE: Version 7.5, Second Edition 1991.
- [5] Hendricks, David M., Arizona Soils, Library of Congress Cataloging in Publication Data 1985.
- [6] Joria, Peter E. and Sean C. Ahearn, A Comparision of the SPOT and Landsat Thematic Mapper Satellite Systems for Detecting Gypsy Moth Defoliation in Michigan, Photogrammetric Engineering & Remote Sensing, 1991 Vol. 57 No. 12, pp.1605—1612.
- [7] Lowe, Charles H. and David E. Brown, Biotic Communities of the American Southwest-United States and Mexico, Desert Plants. 1982, Vol. 4 No. 1—4.
- [8] Richard, J. A. et al., A Means for Utilizing Ancillary Information in Multispectral Classification, Remote Sensing of Environment, 1982, 12: 463—477.
- [9] Rosenfiel, George H. and Katherine Fitzpatrick-Lins, A Coefficient of Agreement as a Measure of Thematic Classification Accuracy, Photogrammetric Engineer. & Remote Sensing, 1986, Vol. 52 No. 2 pp.223—227.
- [10] Whittaker, R. H. and W. A. Nicing, Vegetation of the Santa Catalina Mountains, Arizona: A gradient Analysis of the South Slope, Ecology, 1965, Vol. 46 No. 2, Summer.

Vegetation Mapping Using Remote Sensing and Geographic Information System Technique at Pusch Ridge, Tucson, State Arizona, USA

Wu Bingfang, Huang Xuan, Tian Zhigang

(LREIS, Institute of Geography, Chinese Academy of Sciences Beijing 100101)

Abstract The classification accuracy of vegetation mapping by satellite imagery in a complex terrain environment can be improved by using ancillary data and imagery spatial features extracted from the images (Richard, 1992). This study is to test the role of GIS spatial and spectral analysis model in aiding the classification of satellite data. Three GIS programs is developed, which improve the accuracy of unsupervised classification for SPOT data from 37.41% to 71.67%.

Second objective is to test the ability of two satellite system, SPOT and Landsat Thematic Mapper (TM), in mapping of vegetation in mountain region. Both data are processed with supervised classification in-corporating with ancillary data. The accuracy with SPOT data is higher than with TM.

Key words Vegetation Mapping, GIS Model, Ancillary data