

黄土高原防护林区土地利用 遥感分类方法研究*

张彦忠 张玉贵

摘要 利用陆地卫星 5 号 TM 数据,对位于黄土高原防护林区的山西省隰县进行了大面积(1 400 km²)的土地利用的计算机自动分类研究。提出了在地形破碎的黄土高原防护林区做遥感图像自动分类的方法和步骤,即把无监分类和有监分类有机结合的混合分类方法,提高了分类精度。本方法为在实践中有效地利用遥感数据提供了一条新途径。

关键词 遥感图像分类、土地利用分类、黄土高原、防护林

土地利用自动分类是遥感技术应用的一个主要研究方向。自从 1972 年对地观测陆地卫星发射升空,就开展了大量的用多光谱数据进行土地利用制图的研究^[1]。研究中所使用的分类技术主要是无监或有监方法,这方面已开发了很多自动分类方法^[2~4]。但都是建立在小面积图像(一般是 512×512 像元,最大为 1 024×1 024 像元)试验研究的基础上。真正对大面积地区进行系统的土地利用自动分类却少有报道。另一方面,利用遥感技术进行土地利用调查,特别是在林业方面的应用,对处于地形不规则或覆盖类型的光谱特性不均匀的地区,遇到了更大的困难。因此在地形地貌破碎,沟壑纵横切割的黄土高原地区应用遥感技术已明显地成为主要问题。

据以上两方面,即研究区面积大,地形破碎,冲蚀沟谷交错,本研究提出了用非监分类和有监分类相结合的方法,对研究区的遥感数据做了土地利用自动分类,并对分类结果的精度进行了初步评价,整体效果是令人满意的。更进一步的研究工作正在进行之中。

1 研究区和数据资料

研究区选在山西省隰县,位于黄土高原,属“三北”防护林区,处在昕水河流域上游。面积约 1 400 km²。基本属于侵蚀黄土塬,少部分属侵蚀黄土丘陵。地形由东向西倾斜,地表大部分为新生代第四纪马兰黄土所覆盖。其地貌特征主要是:土层深厚,地面破碎,沟壑密布,坡陡沟深,水土流失严重。该县东及东北部有一定面积的天然次生林和大片疏林灌草丛的土石山区,自然植被较茂盛,针阔叶林片状混交并有成片的落叶松灌丛和温性草原灌丛。以上这些特点,在该县 TM 假彩色合成图上可明显看出。

考虑到以上特征,首先确定了该县的适合于遥感的主要土地利用类型。防护林区土地利用遥感分类系统,应以作为基本信息源的航天影像为依归,从土地的用途、经营特点、利用方式、覆盖特征入手建立。本研究所用的信息源是陆地卫星 TM 图像,其地面分辨率为 28.5 m,考虑到这些,同时结合当地实际情况,特制定如下分类系统。

1994—12—21 收稿。

张彦忠助理研究员,张玉贵(中国林业科学研究院资源信息研究所 北京 100091)。

* 本研究系国家“八五”攻关课题《“三北”防护林体系和植被动态监测及信息管理系统研究》的一部分。

(1)耕地:种植作物的土地,包括水田和旱地。该县以旱地为主,多分布于塬面上,在沟谷和缓坡上也有分布,由于种植作物不同,在 TM 图像上有多种表现形式。

(2)园地:该县以果园为主,由于分布零星,多属果粮间作地,故在遥感图像上难于区分,在自动分类中未做单独处理。

(3)草地:以生长草本植物为主,用于畜牧的土地,包括天然草地,改良草地和人工草地。多分布于较陡的山坡上,较深的沟谷内。

(4)灌丛:覆盖度大于 40% 的灌木林和较高的草丛。多分布于海拔较高的山区。

(5)针叶林:该县以油松(*Pinus tablaeformis* Carr.) 为主。与其它植被和分布范围有明显不同,影像能够区分。

(6)阔叶林:主要有刺槐(*Robinia pseudoacacia* L.)、杨树(*Populus*)和其它天然次生林,绝大多数分布于土石山区。

(7)混交林:针叶树和阔叶树混交,另外还有零星灌木林分布于其中。

(8)城镇居民点用地:较大城镇能在影像上呈现。因多住窑洞,小居民点不能显示。

(9)水域:该县水域分布较小,与沟谷耕地相混杂。

(10)其它用地:难以确定名称定义的用地和难利用地。

由于是计算机自动分类,最后确定的类型为:耕地,草地,灌丛,以油松为主的针叶林,以刺、槐杨树为主的阔叶林,其它阔叶林及混交林,城镇居民点用地共 7 类。

所用遥感数据取自 Landsat-5 卫星 1992-06-08 记录的 TM-CCT 数据,轨道号为 126/34。由于属夏季影像,区内主要作物和其它天然植被都处于生长日期,有利于植被的分类。其它数据资料还有研究区 1:5 万和 1:10 万地形图,早期土地利用图,黑白航空像片(部分地区)。

2 方法

2.1 研究区 TM 数据的一般特征分析和几何精纠正

2.1.1 数据特征分析 TM 图像中包含着十分丰富的地表信息,其信息载体便是 7 个波段的数据。很明显,只有在对这信息载体的基本特征有了充分的认识之后,才有可能通过数字图象处理技术,把隐含其中的有用信息充分地提取出来。因此,在作一地区具体的 TM 图

表 1 研究区 TM 数据统计特征参数

| 项目 | TM ₁ | TM ₂ | TM ₃ | TM ₄ | TM ₅ | TM ₇ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 最小值 | 56 | 22 | 20 | 25 | 19 | 8 |
| 最大值 | 141 | 82 | 112 | 147 | 184 | 121 |
| 均值 | 91.506 | 47.700 | 59.615 | 77.305 | 112.916 | 66.165 |
| 标准差 | 14.381 | 10.398 | 17.881 | 12.916 | 26.619 | 22.728 |
| TM ₁ | 1.000 | | | | | |
| TM ₂ | 0.984 | 1.000 | | | | |
| TM ₃ | 0.983 | 0.991 | 1.000 | | | |
| TM ₄ | 0.098 | 0.186 | 0.131 | 1.000 | | |
| TM ₅ | 0.919 | 0.942 | 0.941 | 0.323 | 1.000 | |
| TM ₇ | 0.963 | 0.973 | 0.980 | 0.159 | 0.974 | 1.000 |

像分类之前,对其数据特征进行认真的分析是取得成功的必要前提,它包括亮度值分布分析,各波段统计特征分析,相关性分析等。表1是研究区内TM 6个波段(除 TM_6 外)的统计特征参数。从表中可以看出, TM_5 亮度值覆盖范围最宽达166个灰级,其次是 TM_4 为122个灰级,表明这两个波段包含的信息量较丰富,波段1,2和3之间相关系数大,波段5和波段7的相关系数也达97.4%。

据各波段的标准差和相关系数,本研究在分类过程中只选用了3,4,5和7四个波段。因可见光波段较少(未用波段1和2),这样将减少大气对分类结果的影响,同时减少机器运行时间。

2.1.2 几何精纠正 由于来自卫星、地球及传感器等因素的影响,图像会发生畸变,必须选择畸变函数,利用地面控制点,采用最小二乘拟合法,来建立像元行列号与地理坐标系之间的关系,从而使TM图像与地形图配准。控制点的选取对于纠正精度是至关重要的。本文选择了均匀分布于整个研究区的13个控制点,建立了二元二次多项式,用于对分类结果图像进行纠正。纠正精度在X和Y方向上都小于半个像元。特别指出本研究用于分类的图像是未经纠正的。因为对图像进行纠正后,不可避免要损失信息,影响分类结果精度,故只用该多项式对分类结果进行了纠正,这样既节约时间,又相对提高了分类精度。

2.2 分类方法及步骤

图像分类,实质上就是据一定的原则把图像中的每一个像元归为某一类的过程。遥感数据包括的像元通常是由多个光谱通道记录的光谱特征。一个像元可以定义为N维特征(光谱)空间内的一个点。专题信息可在多光谱图像分类中提取。因此,一个图像分类系统的输出即是一幅专题图。在该专题图中,原始图像内的每个像元都归属于多个光谱类中的某一类。光谱类是遥感数据固有的,由分析者定义和标记的。多光谱图像分类分为监督和非监督分类或两种方法的混合。本研究采用了混合分类方法,非监督训练确定了光谱类型,监督分类指定了每一个未知像元所属的光谱类别。

2.2.1 非监督训练 虽然进行了实地调查,并做了训练样地,由于研究区面积较大,不可能对区内地物分布有详细的了解,故采用非监督训练方法,生成初步的光谱分类图。然后利用所收集到的反映真实地表信息的其它资料,对初步分类图进行分析。参照分类系统,使光谱类与信息类一一对应。在这里有的光谱类被合并。本文用ISODATA(Iterative Self-Organizing Data Analysis Technique)训练方法,即迭代自组合数据分析技术,处理过程如下:

(1)计算每像元与初始类别中心的距离,将该像元分配到最近的类别中。

(2)计算并改正重新组合的类别中心,重新进行(1)的过程,如果重新组合的像元数在某一门限值以下,或当类别数在一定的范围内,而类别中心间的距离在门限值以上,且类内,方差的最大值为门限值以下时,分析終了。

(3)输出类别统计量和初始分类图(光谱类别图)。

经过ISODATA分析过程,得到训练结果,本研究中,得到多类光谱类别中心,利用散点图和类别间的距离等参数对其进行分析,对距离较近,或散点图上基本重叠的类进行了合并,形成新的中心,并把每一中心与信息类建立联系,做为监督分类的输入。

2.2.2 监督分类 对研究区的遥感数据建立了训练样本后,即可输入分类器,进行监督分类。本研究应用高效的极大似然分类法,其判别函数为:

$$D = \ln(a_j) - [0.5 \ln(|C_j|)] - [0.5(X - M_j)^T(C_j^{-1})(X - M_j)]$$

这里 D 为权重距离; j 为指定的类别, a_j 是未知像元属于 j 类的先验概率, C_j 为 j 类样本像元的协方差矩阵, $|C_j|$ 为矩阵 C_j 的行列式值, C_j^{-1} 为 C_j 的逆矩阵, T 为转置函数, X 为像元数据矢量, M_j 为 j 类样本的均值矢量。由于无法确定每一类别的权重, 故在分类时, 对每一类来说, 权重都相等, 即为 1。当 D 最小时, 则该像元归属于 j 类。

2.2.3 类后处理和精度分析 由于采用的分类方法, 是以像元为最小处理单位的, 受同物异谱、异物同谱及其它不可知因素的影响, 分类结果必然存在着很多离散的细小图斑。为此, 采用了邻域处理技术, 对分布过于分散的小图斑进行了处理, 对分类结果图像起到了整饰的作用。文中采用了 3×3 窗口滤波处理, 使该窗口内的中心像元占大多数。

类后处理还应该包括分类精度的评价和分析, 不同的评价方法, 得出的精度有所不同, 将另文讨论。在本文中, 只利用该县原有的土地利用图和部分航片及原始彩色合成遥感图像, 对分类结果的精度进行了评价, 表明分类结果是令人满意的。尤其对林地的分类, 可以分出针叶林、稀疏阔叶林、高郁闭度的阔叶林、针阔混交林和部分灌林草丛, 其图斑整齐明显。表 2 列出了几种主要类型的分类精度。

表 2 几种主要类型的分类精度

| 土地利用/覆盖类型 | 分类精度(%) |
|-----------|---------|
| 耕地 | 80 |
| 针叶林 | 85 |
| 高郁闭度的阔叶林 | 83 |
| 稀疏阔叶林 | 70 |
| 针阔混交林 | 76 |
| 灌林草丛 | 70 |
| 城镇居民点 | 90 |

由于当时正值初夏季节, 农田上所种作物不同, 故在训练样本中农田有多类, 这可称之为覆盖类型, 但最终要得出土地利用类型, 故在做类后处理和精度验证时, 把这几种合并为一类, 称之为耕地。

3 结语

本文结合了遥感图像大面积覆盖和计算机快速处理的双重特点, 对研究区 1400 km^2 的面积进行了客观的土地利用自动分类研究, 建立了一套具体的大面积的土地利用自动分类的方法和步骤, 使遥感技术在实际应用中得以真正的实现。通过本研究, 得到如下几点体会:

(1) 利用计算机对地形地貌破碎的黄土高原防护林区进行遥感数据的土地利用自动分类研究是可行的。对林地和非林地的识别非常容易。能够快速获得林地面积数据。

(2) 运用监督分类和非监督分类两种方法的有机结合, 可以克服前者的盲目性, 后者的不确定性, 获得事半功倍的好效果, 这一方法能够充分利用光谱信息。

(3) 在对非监督分类结果进行分析时, 综合了现地调查及人工目视解译的知识, 这要求操作者要有较强的遥感和地理等基础知识。这也是提高监督分类精度的关键所在。

(4) 本文仅从如何全面地利用地物光谱信息的角度出发, 对研究区进行了土地利用/覆盖分类研究。研究区位于黄土高原, 有其特殊性, 为了进一步提高分类精度, 还可以对原始图像进行增强处理, 充分利用遥感图像的光谱、纹理形状等信息。另外, 还可以引进 GIS 数据, 这方面的研究正在进行。

参 考 文 献

- 1 Campbell J B. Introduction to remote sensing. New York: Guilford Press, 1987.
- 2 张彦忠, 张福祥. TM 数据在桔果经济林区土地利用分类和专题信息提取中的应用. 遥感技术与应用, 1992, 7(3): 15~21.
- 3 杨凯, 孙家柄, 卢健, 等. 遥感图像处理和方法. 北京: 测绘出版社, 1988. 392~448.
- 4 K. ARAI. A supervised thematic mapper classification with a purification of training samples. Int. J. Remote sensing, 1992, (13): 2039~2049.

Land Use Classification Using Remote Sensing Data in Protection Forest Region of Loess Plateau

Zhang Yanzhong Zhang Yugui

Abstract A study on the use of landsat-5 TM data to conduct the land use auto-classification in a large area was carried out in Xixian County, Shanxi Province. A method of integrated supervised classification with unsupervised classification is proposed. The precision of classification is improved. The results are very satisfactory. The study provided a new efficient approach to apply remote sensing data in practice.

Key words remote sensing, land use classification, loess plateau, protection forest

Zhang Yanzhong, Assistant Professor, Zhang Yugui (The Research Institute of Forest Resource Information Techniques, CAF Beijing 100091).