

基于遥感和陆表信息集成的广东省土地覆盖分类方法研究

陈劲松 梁守真 余晓敏 王贺

(中国科学院深圳先进技术研究院 深圳 518055)

摘要 土地覆盖信息是估算地-气间的生物物理过程和能量交换的关键参数,也是区域和全球尺度气候和生态系统过程模型所需要的重要参量。如何高效地利用遥感数据提取土地覆盖信息是当前研究迫切需要解决的问题。面向对象的分类方法不但充分利用了遥感数据的光谱信息,同时也利用了影像的纹理结构信息和更多的地物分布信息关系,在遥感分类中具有较大的潜力。研究基于2010年多时相的环境卫星数据、TM数据以及DEM数据,并结合地表采集的4000多个样点数据,采用面向对象的分类方法对广东省土地覆盖进行分类。经采样验证,广东省土地覆盖平均精度为85%,分类结果精度远高于常规的分类算法,说明结合陆表信息的面向对象分类方法比常规的分类算法更具有优势,可以实现高精度的土地覆盖分类。

关键词 土地覆盖分类;面向对象;遥感;广东

Study on Land Cover/use Classification of Guang Dong Province by Integral Use of Remote Sensing Data and Field Survey

CHEN Jin-song LIANG Shou-zhen YU Xiao-min WANG He

(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract Land cover/use information is a key parameter for estimating biophysical process and energy exchange between land and atmosphere, and meanwhile it is also an important parameter for regional and global climate models and ecosystem process models. Remote Sensing is an important and effective tool for mapping land cover /use on a large scale. But the quality of remote sensing data and weather condition have great impacts on the accuracy of land cover /use information from remote sensing data. This research proposes a new methodology for mapping land cover /use of Guang Dong province using object-oriented classification with remote sensing data and field survey information as inputs. Object-oriented classification has been increasingly employed for mapping land cover/use because of its advantages of using multi information of remote sensed objects such as spatial distribution, shape, size, spectral and texture over some traditional classification methods which often only use spectral information of objects. The remote sensing data used in this research include multitemporal and multi spectral China HJ-1A/B and TM data. Firstly multi temporal remote sensing data were processed and geo-rectified. The classification rules were then established using object-oriented classification method with processed remote sensing data, DEM, history land cover/use data and field observation data as inputs. Finally a new classification method is developed by integrating multi information from remote sensing data and field observations into classification rules. In situ observation data is used to validate classification result. The result shows that the average accuracy of mapping land cover /use of Guangdong province is 85%, much higher than that of traditional classification method. The accuracy can be further improved by using more effective classification rules and more field survey in future.

Keywords land-cover classification; object-oriented; remote sensing; guangdong

基金项目: 中国科学院战略性先导科技专项XDA05050107-3; 广东海南固碳参量遥感监测, 2011-2014。

作者简介: 陈劲松, 研究员, 博士生导师, 主要研究方向为定量遥感及遥感环境监测应用, E-mail: js.chen@siat.ac.cn; 梁守真, 助理研究员, 研究方向为环境和生态遥感; 余晓敏, 研究工程师, 研究方向为遥感数据处理和生态监测应用; 王贺, 研究实习员, 研究方向为微波遥感数据处理和应用。

1 引 言

土地覆盖 (Land Cover) 是随着遥感技术的发展而出现的一个新概念, 它是指地球表面当前所具有的自然和人为影响所形成的覆盖物, 是自然过程和人类活动共同作用的结果, 具有特定的时间和空间属性, 其形态和状态可在多种时空尺度上变化^[1,2]。土地覆盖的变化在很大程度上影响着经济、生态、资源、环境、人口等因素的可持续发展, 是全球变化研究的重要组成部分, 也是资源与环境研究的核心内容之一。随着全球经济的快速发展, 地表的土地覆盖类型发生了重大的变化, 及时掌握土地覆盖信息对政府各部门也具有十分重要的意义。近年来, 土地覆盖监测受到了越来越多的重视, 国内外的专家学者也对土地覆盖分类方法进行了大量的研究。然而, 目前大部分方法主要是基于象元、基于特征或两者相结合的方法, 这些方法都没有充分利用遥感影像的其他信息, 如地物的几何信息、结构信息以及地物之间的拓扑关系等, 不能充分发挥遥感的优势^[3,4]。与以上传统的分类方法相比较, 面向对象的分类方法是以具有某种相似特征的多个临近像元为基本的分类单元, 更多地考虑了影像对象间的语义信息、纹理信息和拓扑关系等信息, 这种分类方法更加便于影像的解译与信息提取。另外, 遥感影像会受到天气、传感器性能、影像分辨率等多种因素制约。因此, 在实际应用中, 需要结合多种陆表信息, 以提高土地覆盖分类的精度。

在中等尺度的土地覆盖遥感分类研究中, TM 遥感影像被广泛采用, 但是 TM 时间分辨率较低, 我国南方地区云雨较频繁, 影像受云覆盖的影响较大, 要获得高质量的多时相 TM 遥感影像比较困难。相较于 TM 影像, 我国发射的环境卫星具有更高的时间分辨率, 两颗卫星平均 2-3 天就可以将同一地区覆盖一遍, 且与 TM 有相同的空间分辨率, 可用于区域土地覆盖分类。因此, 将多时相的 TM 以及环境卫星影像相结合, 可以为土地覆盖分类提供更加充足的数据。此外, 在分类过程中, 如果将一部分野外采集数据集集成到分类过程中, 一方面可使样点中所包含的信息能够充分被利用, 另一方面可提高分类的精度。因此, 本文利用面向对象的理论知识, 结合 TM、环境卫星数据、DEM 数据以及地面采集数据, 将广东省作为研究区, 对其土地覆盖分类进行研究。

2 研究区概况与数据

广东省位于中国南岭以南, 地势复杂, 主要呈北高南低的趋势, 面积约为 17.98 万平方公里, 包括山地 (31.7%)、丘陵 (28.5%)、台地 (16.1%) 和平原 (23.7%)。气候以亚热带季风气候为主, 从北向南逐渐由亚热带气候向热带气候过渡, 南部雷州半岛属季风热带的北缘, 全省年均降水量为 1366 mm, 年平均气温 22℃。该地区植被类型丰富, 森林类型具有亚热带向热带过渡的特征。植被四季常绿, 种类分布的区域性与季节性规律并不明显, 尤其在广东南部, 多数以混杂为主, 在遥感影像上表现较为破碎, 不同类型植被光谱差异不明显。广东省生态、矿产资源丰富, 农林业生产发展条件优越, 是中国热带经济作物重要生产基地之一。近年来, 随着广东经济的进一步发展, 省内逐步形成了地少人多的局面。为了指导土地资源的合理开发利用, 建立本地区真实、可靠的土地覆盖数据显得尤为重要。

本研究主要采用 2010 年的多时相环境卫星 HJ-1A/B 影像以及美国 Landsat TM 影像, 结合 DEM 以及地表测量数据进行土地覆盖分类研究。HJ-1A 星和 HJ-1B 星上搭载的两台 CCD 相机可获取 4 个波段的影像数据, 地面像元分辨率为 30 米, 各重访周期为 4 天, 两颗运行后, 平均 2~3 天同一地区可以覆盖一遍。与 Landsat TM 卫星数据相比, 其具有明显的多时相数据采集能力, 可基本反映植被的时间变化规律, 但其波段数要少于 TM 影像。地表数据采集工作基于沿道路两旁 2 公里视觉范围内的地物采样, 整个研究区内共采集样点 4000 余个, 样点分布图如图 1 所示。DEM 数据为 ASTER GDEM, 地面分辨率为 30 米。

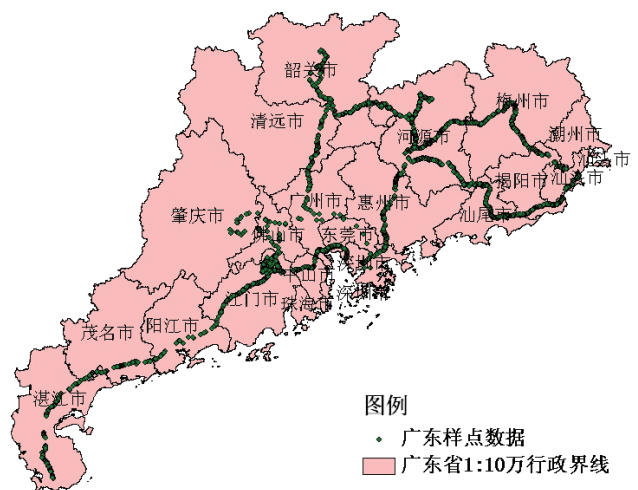


图1 广东省地表采集数据分布图

3 研究方法

本研究具体技术流程如图2所示，主要包括三个部分：数据预处理、土地覆盖分类、精度评价。

3.1 数据预处理

在进行分类前，首先需要对TM数据、环境卫星数据以及辅助数据进行预处理。TM数据的预处理主要包括去条带、辐射校正、地形校正和大气校正。环境卫星数据的预处理主要包括辐射校正、几何校正、地形校正和大气校正。辅助数据的预处理主要包括坐标投影转换。最后所有的空间数据被转换到统一的坐标系中(UTM投影、WGS 84椭球体和基准面)。此外，基于处理后的TM数据和环境卫星数据初步建立遥感解译标志库。

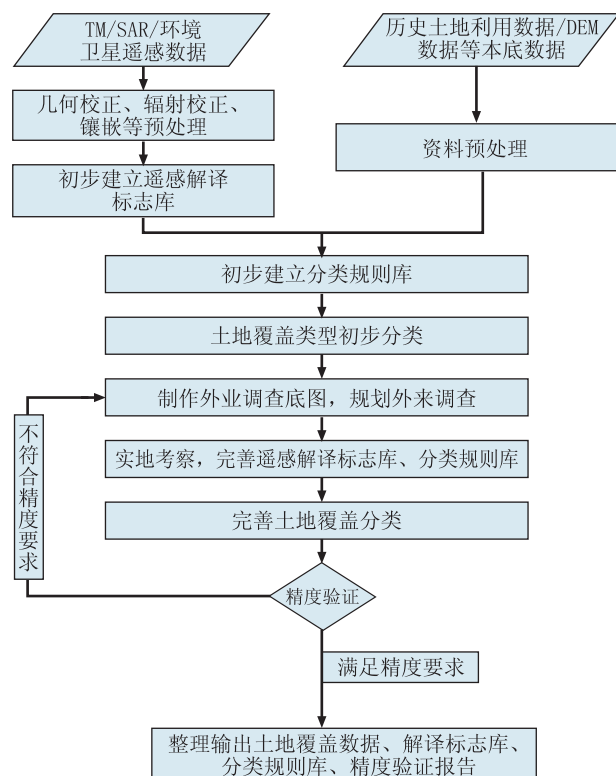


图2 总体技术路线图

3.2 土地覆盖分类

自动分类是计算机图像处理的初期便涉及的问题。传统的遥感自动分类主要依赖地物的光谱特性，采用数理统计的方法，基于单个像元来进行，如监督分类和非监督分类方法。本研究采用面向对象的分类方法对图像像元进行归类。与其他传统的基于像元的影像分类方法相比，面向对象分类的基本处理单元是有意义的影像对象和它们的相互关系，而不是单个的像元。该方法不但顾及了遥感数据的光谱信息，同时

也利用了影像的纹理结构信息和更多的上下文语义关系。其基本思路是：首先根据地表景观的多尺度效应，基于多尺度分割技术构建具有不同尺度大小的影像对象，形成具有不同分辨率的影像目标层次^[5]。然后在各类地物所对应的最优观测尺度下，根据分割对象的光谱特征、空间特征、纹理特征等，按照一定的规则进行模糊分类。整个分类的技术流程如图3所示。

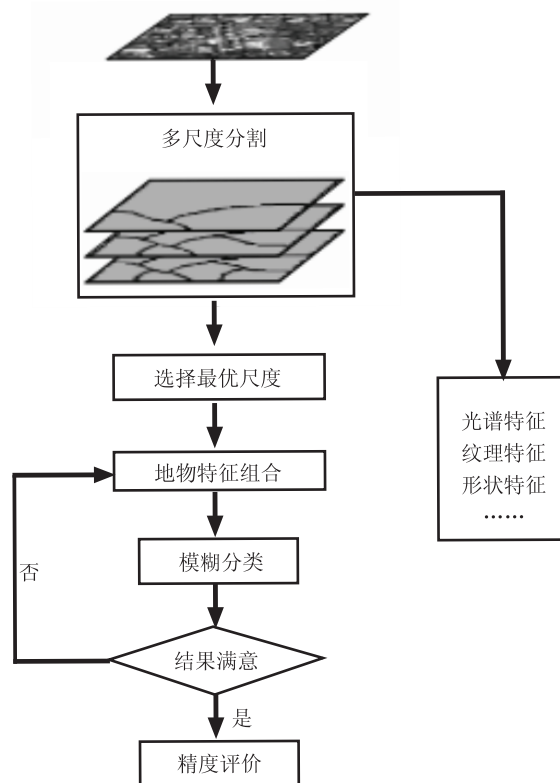


图3 分类技术流程图

(1) 影像分割

分割对象的好坏直接影响到整个分类的精度，并且不同的分类对象对分割需求也各不相同，如何有效准确地使用各种信息，得到满足各个分类需求的最佳分割方案首先要考虑需要解决的技术难点。影像分割的标准为：①影像对象的平均异质性应该被减少到最小；②像素的平均异质性应该最小化，像素所归属的影像对象的异质性应该被分配到每一个像素中^[6]。影像分割主要涉及到四个参数：图层权重、分割尺度、形状因子(颜色因子)、紧致度(平滑度)。

分割图层及图层权重的选择直接影响到分割对象，因此在分割时，要保证所使用到的信息尽可能全。通过数据变化方法能够生成信息丰富的图像。LBV变换利用多光谱图像多个波段的光谱信息，经过数据变换可得到反映地物总辐射水平的L图像、反映可见光-近红外辐射平衡的B图像以及反映波段辐射

变化矢量（速度和方向）的V图像。将L、B、V图像分别赋予红、蓝、绿色进行彩色合成后，能得到色彩鲜艳、层次分明、地物类别易于区分的图像。该变换方法能突出植被、水体和裸露地表等典型地物，为信息识别和提取提供了方便。同时LBV彩色合成图像上颜色与地物类型存在逻辑对应关系，据此可准确地提取地物信息。可以根据不同的图层对于分割结果的重要性而取不同的权重值。

影像信息分割时尺度的选择是很重要的，它直接决定以后分类结果精度高低。分割尺度要求既能突出整体地物的连续性，又能保证一些细节地物的特征。对于一种确定的地物类型，最优分割尺度应使得生成的影像对象能清楚地表示该地物类型的边界，不能过于破碎，也不能出现边界模糊，而且要求清楚提取类别信息的情况下使用的数据量最小。要想确定影像上对各类地物都能较好地进行分割的分割尺度，就必须确定一个综合各类地物的最优分割尺度。

由于分割过程中相邻的相似像元会被合并成一个多边形对象，因此在空间域和光谱域中对象不断增长的异质性都必须最小才能保证整幅影像在给定尺度的情况下所有影像对象的平均异质性最小。这不仅是指光谱上的异质性，而且也包括形状上的异质性最小。由异质性函数(式1)^[7]可知，任何一个影像对象的异质性值是由光谱信息权重、形状信息权重、光谱异质性值、形状异质性值四个变量计算而得到的，如何确定最优参数对于对象构建的准确性有着很大的影响。目前应用过程中多是通过小区域实验或反复实验确定最优参数。

$$f = w * h_{color} + (1 - w) * h_{shape} \quad (1)$$

其中， f 表示异质性， h_{color} 表示光谱异质性， h_{shape} 表示形状异质性， w 表示权重。

通过大量实验，本研究发现：使用不受云污染的HJ波段及其L、B、V图像参与分割，各波段的分割权重设为1，分割尺度设为30，形状因子设为0.1，紧致度设为0.5，能够得到较好的分割结果。

(2) 对象特征提取和筛选

提取稳定、有效的对象特征是土地覆盖信息提取的关键。影像对象特征空间可分为光谱特征、几何特征、纹理特征、空间关系特征、语义特征、区域地学特征和干扰特征，此外还可基于实际应用的需要建立更适用且更有意义的自定义特征。本研究主要采用的光谱特征有NDVI、RVI、DVI、NDWI、SBI、BAI等，形状特征有面积、长宽比等。

提取对象各种特征信息的最终目的是为了从尽可能多的不同角度综合地描述某一类型对象区别于其他类别的本质特征。一般来说少量特征很难全面地表述出一个类别的本质特征，但试验证明，并不是特征越多越好。由于可能存在相关性，一些对象特征信息对图像分类的作用有限，甚至会使分类器无法运行，因此需要对其进行筛选。SEaTH算法^[8]是目前较流行的特征筛选方法，该算法采用基于训练目标的统计方法来确定有效特征。这些训练目标代表了从全部图像对象中分离出来的一个小子集，是每个对象类的典型对象。采用对象类中相互间两两可分性的统计度量（J-M距离，式5）确定每个对象类的有效特征，随后计算所选择特征容许的最大可分离阈值。

$$J = 2(1 - e^{-B}) \quad (2)$$

$$\text{其中 } B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2} \ln \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2} \quad (3)$$

式中 m_i 和 σ_i ，分别代表两个类别样本对象某个特征分布的均值和方差。

当两个类别样本特征分布的均值和方差相等时，式(2)和式(3)均等于0，此时 C_1 和 C_2 类完全重叠。

(3) 基于规则集的土地覆盖分类

根据野外采集的样点数据，分析不同类别影像对象的各种特征，找到合适的分离阈值，然后将分割对象的各种特征函数组合起来，形成类别描述，最后基于模糊数学评价的方法，依照这些分类规则识别各影像对象。如果对得到的分类结果不满意，可以重新调节分类规则集，形成新的类别描述，重新分类，直到分类结果满意为止。总的来说，该分类过程是透明的、可调节的，可以通过不断调节类别描述，来改善分类结果。

最后再根据分类的实际情况，制作外业调查底图，进行野外调查，从而完善遥感解译标志库和分类规则库，完善土地覆盖分类。

3.3 精度评价。

精度评价也就是对分类结果进行评价，确定分类的精度和可靠性。精度评价对于遥感分类很重要，因为一幅分类图像的精度直接影响着分类数据的有用性和利用其进行科学研究的合理性。最常用的方法是建立混淆矩阵对分类结果进行检验，混淆矩阵既可计算出总精度、Kappa系数等表示分类总的精度情况的指标，又可计算出制图精度、用户精度等表示单个类别分类情况的指标^[9]。本研究中便采用此法来评价广东省土地覆盖分类结果。建立混淆矩阵所需要的样本来

自野外采样点和高分辨率图像目视解译点。在选择样本时为保证每一类地物都有对应的检验样本，研究中采用分层采样的方法（即分别对每个类别进行采样），每种地类选择一定的样本，依据选择样本建立混淆矩阵，计算分类精度。若满足精度要求，则整理输出土地覆盖数据、解译标志库、分类规则库和精度验证报告。若不满足精度要求，则对分类精度差的区域增加野外调查点，进一步完善解译标志库，从而提高分类精度。

4 结果与讨论

基于面向对象的分类方法，得到了广东省2010年的土地覆盖分类结果，如图4所示。广东省的土地覆盖共分为5种一级类型和27种二级类型。基于采样点建立混淆矩阵，计算得到广东省土地覆盖分类总体平均精度为85%。

通过对分类结果分析，发现影响分类精度的主要原因是林地之间以及水田和旱地之间的混淆导致。林地在TM图像上光谱相似，利用一两个时相的数据很难区分不同类型的林地，需要多时相的遥感数据，但是在广东地区云雨较多，很难获取高质量的多时相影像。在广东地区，水田和旱地年际甚至年内转换频繁，而验证点只是某一时间的数据，因此可能会导致精度偏低。此外北部山区土地覆盖破碎，像元混合严

重，也影响了土地覆盖的分类结果。因此要监测土地覆盖破碎度高的区域的土地覆盖变化，还需要更多的辅助数据，同时，混合像元分解也是不可回避的问题。

总的来看，结合地表数据，采用面向对象的分类方法能够高精度地提取土地覆盖信息，相比于传统的分类算法，具有一定的优势。但是研究也发现，尽管该方法的分类精度较高，但是仍然存在着一些遥感分类普遍难以解决的问题。比如混合像元、更细致的林地类之间的区分，加入更高分辨率或者受天气条件影响较小的雷达数据或许有助于这些问题的解决，这有待于以后更加深入的研究。

参 考 文 献

- [1] 史培军,宫鹏,李晓兵,等.土地利用/覆盖变化研究的方法与实践[M].北京:科学出版社,2000:1-2.
- [2] 韩涛.遥感监测土地覆盖变化的方法及研究进展[J].干旱气象,2004,22(2):76-81.
- [3] 钱巧静,谢瑞,张磊,等.面向对象的土地覆盖信息提取方法研究[J].遥感技术与应用,2005,20(3):338-342.
- [4] 杜凤兰,田庆久,夏学齐,等.面向对象的地物分类法分析与评价[J].遥感技术与应用,2004,15(1):20-23.
- [5] 高伟.基于特征知识库的遥感信息提取技术研究[D].中国地质大学博士学位论文,2010.
- [6] Volker W. Object-based classification of remote sensing data for change detection[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2004,58:225-238.
- [7] Definiens Image gmbh. ecognition user guide[EB/OL]. <http://www.definiens-imaging.com/document>.
- [8] Bruzzone L, Serpico S B. A technique for feature selection in multiclass problems[J]. International Journal of Remote Sensing, 2000, 21(3): 549-563.
- [9] Hay A M. The derivation of global estimation from a confusion matrix[J]. International Journal of Remote Sensing, 1988,9(8): 1395-1398.

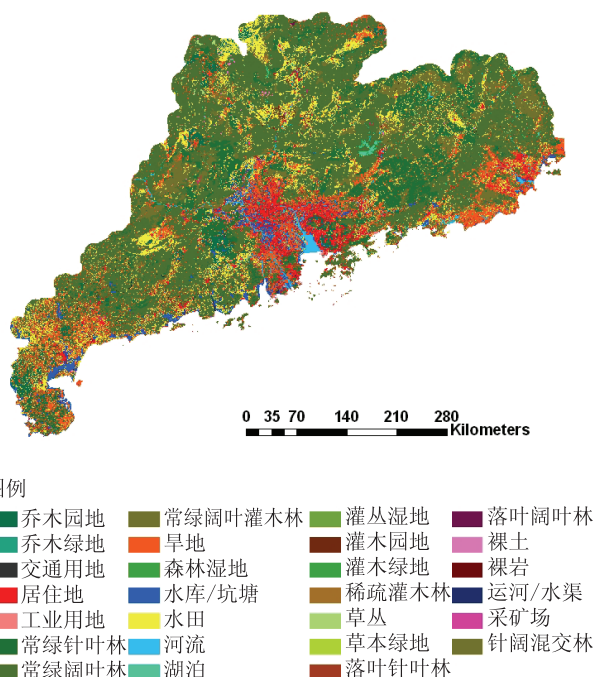


图4 广东省2010年土地覆盖分类图