

博士文丛(第三辑)

JIYU ZHISHI DE JIANGNAN DIANXINGQU
TUDI LIYONG/FUBEI FENLEI YANJIU

基于知识的江南典型区 土地利用/覆被分类研究

赵萍 著

安徽人民出版社

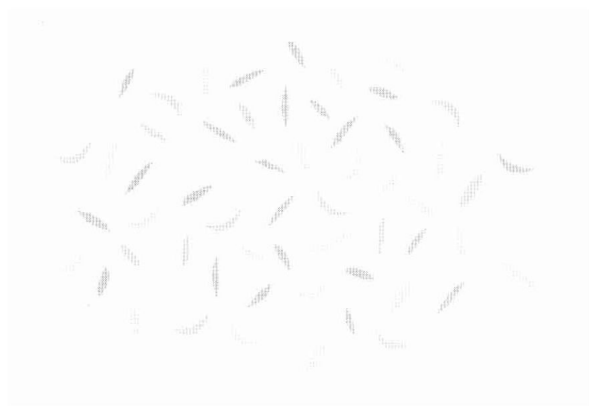
博士文丛(第三辑)

JIYU ZHISHI DE JIANGNAN DIANXINGQU
TUDI LIYONG/FUBEI FENLEI YANJIU

基于知识的江南典型区 土地利用/覆被分类研究



赵萍 著



安徽人民出版社

图书在版编目 (CIP) 数据

基于知识的江南典型区土地利用 / 覆被分类研究 / 赵萍著. —合肥: 安徽人民出版社, 2008.6

ISBN 978-7-212-03284-5

I. 基… II. 赵… III. 遥感地面调查—应用—土地利用—研究—华东地区 IV. F321.1

中国版本图书馆 CIP 数据核字 (2008) 第 080818 号

基于知识的江南典型区土地利用 / 覆被分类研究

Knowledge-based Landuse/cover Classification

in the Typical Areas of the Lower Reaches of Yangtze River

赵 萍 著

出版发行: 安徽人民出版社

地 址: 合肥市政务文化新区圣泉路 1118 号出版传媒广场 8 楼

发 行 部: 0551-3533258 3533268 3533292 (传真) 邮编: 230071

组 编: 安徽师范大学编辑部 电话: 0553-3883578 3883579

经 销: 新华书店

印 刷: 芜湖新欣传媒有限公司

开 本: 965 × 1270 1/32 印张: 4 字数: 105 千

版 次: 2008 年 6 月第 1 版 2008 年 6 月第 1 次印刷

标准书号: ISBN 978-7-212-03284-5

定 价: 12.00 元

本版图书凡印刷、装订错误可及时向承印厂调换

《博士文丛》编委会

主 任：蒋玉珉

副主任：王世华 王 伦

委 员（以姓氏笔画为序）：

王 伦	王世华	王先俊
王绍武	朱 平	朱士群
朱家存	孙国正	孙胜忠
巫 俊	李琳琦	陆 林
张 毅	张传开	房列曙
周志俊	赵广超	姚本先
聂刘旺	黄时中	龚基云
阎蒙钢	蒋玉珉	谢昭新

摘要

本研究以提高江南典型区土地利用 / 覆被自动分类的精度为目标,围绕基于知识的遥感影像分类方法的过程展开。首先从目视解译的原理出发,分析了研究区土地利用 / 覆被类型在遥感影像上的表征,并对相关特征加以分析与表达;然后从知识的概念、知识的获取、知识库的构建以及知识的推理几方面深入探讨了基于知识的遥感影像分类方法,其中重点研究了知识获取的决策树方法,并在此基础上对绍兴和江宁两个试验区进行了基于知识的遥感影像土地利用 / 覆被分类以及传统的监督分类和逻辑通道法分类试验;最后从不同分类方法的精度评价结果的比较和分类规则的分析两方面,对基于知识的遥感影像土地利用 / 覆被分类方法的可行性进行了评价。

当前,土地利用 / 覆被变化(LUCC)研究是全球变化研究的核心内容,而土地利用 / 覆被数据的获取是其主要的基础工作。近年来,飞速发展的遥感技术已成为土地利用 / 覆被调查的一种有效的技术手段。然而,遥感影像分类技术的提高却远远落后于遥感传感器技术的发展,不能满足实际应用的需要。因而,实现遥感影像解译的自动化与高精度定量化,以满足人们从海量遥感数据中快速、准确地获取专题信息的需要,不仅是当前遥感自身发展的前沿,也是遥感应用的迫切要求。

在我国江南地区,土地利用 / 覆被情况相对比较复杂,受遥感

影像分辨率的限制以及成像过程中诸多因素的影响,影像中“同物异谱,异物同谱”现象比较突出,用传统的基于单个像元光谱特性的监督或非监督自动分类方法很难获得令人满意的分类效果。而基于知识的遥感影像分类方法,不仅可以综合影像的光谱特征和结构特征,而且可以综合关于地物分布的地学规律和专家经验,这就为集成多源数据、解决复杂地表的土地利用 / 覆被分类提供了一条有效的途径。但是,知识的获取一直是其得以广泛应用的“瓶颈”,通常还是采用人工手段,由知识工程师和解译专家协同完成。首先由解译专家通过对地面特征在影像上表征的分析,以及对大量具有代表性样本的统计,总结出反映地面特征的光谱属性、空间分布及时相变化的内在规律;然后由知识工程师通过与解译专家的交谈抽取这些知识,并用适当的知识表示方法转换为计算机可存储的内部形式存入知识库。这种间接的、机械式的知识获取手段不仅需要耗费大量的时间和精力,而且不同的专家间也会存在知识的冲突,很大程度上影响了遥感影像分类的智能化水平和客观性,阻碍了它的广泛应用。20 世纪 90 年代,新兴的空间数据挖掘和知识发现 (Spatial Data Mining and Knowledge Discovery, SDMKD) 技术为遥感影像分类知识的获取开辟了一条新的途径。它可以发现空间数据中普遍的几何知识、空间分布规律、空间关联规则、空间分类规则、空间特征规则、空间区分规则和空间演变规则等。

基于上述背景,本研究选取浙江绍兴地区和江苏南京江宁区这两个典型的江南低山丘陵地区为试验区,分别对中尺度遥感影像 TM 和 SPOT - XI 进行了基于知识的遥感影像土地利用 / 覆被分类方法的研究。主要研究内容和结论包括:

1. 通过实地调查和影像判读,分别对上述两个试验区内各种地物类型在影像上的表征进行定性分析,并在此基础上对图像的光

谱响应特征、纹理特征、位置布局特征和空间几何特征加以分析和表达。通过各类地物典型样本的灰度统计分析试验区内各类地物的光谱响应特征和谱间关系;采用基于灰度共生矩阵的描述方法对影像的纹理特征进行提取,并通过对各类地物样本纹理特征的统计分析,确定最佳纹理统计参数;分析地物的空间几何特征,选用面积、周长及基于面积和周长的形状描述指标用于水体类型的细分;针对试验区内地物分布的地域特征,选取地理坐标、DEM、坡度和坡向作为空间布局特征的描述指标。分类试验表明,对于我国江南低山丘陵区来说,辅助纹理特征和位置布局特征可以较好地改善土地利用/覆被分类的精度。纹理特征一方面突出了居民地、道路等自身空间几何结构比较明显的人工构筑物的信息;另一方面增强了山体起伏较大地区的阴影与其它地物的区分。地形因子主要是辅助了水田和林地、园地等具有明显空间分布规律的土地利用/覆被类型的识别。空间几何特征则可以有效地对水体类型进行划分。由此也说明,基于知识的分类方法不仅可以提高分类的精度,而且可以提高分类的等级。

2. 基于知识的遥感影像分类方法的研究及该方法在两个试验区的土地利用/覆被分类的试验。这部分主要研究了知识获取的决策树方法、基于产生式规则的知识库构建方法和基于知识库的推理策略。其中重点论述了如何利用分类回归树(Classification And Regression Tree, CART)算法构建决策树获取分类规则的方法。CART算法具有以下的特点:对输入数据没有任何统计分布的假设要求,测试变量之间既可以独立也可以相关,预测变量既可以是连续的数字变量也可以是离散的类别变量;可以处理特征集与目标类别间的非线性关系;可以有效地处理高维数据;能包容数据的缺失和错误;能明确地判定变量对于分类的重要性;结构容易理解,可以方便地用规则来表示等等。本研究通过分类回

归树算法对样本的学习构建决策树获取了大量的分类规则,在两个试验区的土地利用 / 覆被分类试验中均获得了满意的应用效果。由此说明,利用分类回归树算法构建决策树,获取分类规则是可行的,是解决基于知识的遥感影像分类方法在土地利用 / 覆被分类中推广应用的“瓶颈”问题的一种有效方法。

3.从不同分类方法的精度评价结果的比较和分类规则的分析两方面,对基于知识的遥感影像土地利用 / 覆被分类方法的可行性进行分析与评价。为了便于比较,对试验区进行逻辑通道法和监督分类试验,并利用基于混淆矩阵的方法分别对基于知识的遥感影像分类方法、逻辑通道法、监督分类法进行精度评价与比较。结果表明,尽管试验区土地利用 / 覆被类型复杂,而且缺乏同一时期的实际验证资料,分类精度的评价存在一定的偏差,但总体说来,由于综合了地物的光谱响应特征、纹理特征和位置布局特征,基于知识的分类方法的分类精度比传统的监督分类和逻辑通道法都有了一定程度的提高。在地表状况更为复杂和地块细碎的江宁试验区,精度的改善更加明显。通过对分类规则的分析可知,所建立的分类规则也反映了实际规律,这说明了利用分类回归树算法构建决策树,获取分类规则是合理有效的。

4.基于知识的分类方法可以有效地综合其它各种辅助数据用于改善分类的精度。但是,辅助数据本身的质量以及辅助数据与影像之间的配准精度,在一定程度上也会影响分类的精度。因此,在实际应用中必须注意辅助数据的选择和影像与各种辅助数据之间的精确配准。

综上所述,本研究主要是提出了对于地块破碎,土地利用 / 覆被情况复杂的江南典型区,利用多源数据进行基于知识的遥感影像自动分类的新方法,也为其它地区中尺度遥感影像的土地利用 / 覆被分类提供了有益的指导。

ABSTRACT

This research concerns with the knowledge-based remote sensing image classification in order to improve the accuracy of landuse/cover classification in the typical areas of the lower reaches of Yangtze River. At first, the research analyzes and describes the detailed characteristics of landuse/cover types in the testareas based on the principles of interpretation. Then it elaborates in depth on the knowledge-based remote sensing image classification from several aspects, including the concept of knowledge, the attainment of knowledge, the establishment of the knowledge base and the deduction based on the knowledge base. The emphasis is made on how to use the decision tree to obtain the rules of classification. After that, three methods are investigated to perform landuse/cover classification in two testareas. These methods are knowledge-based classification, logic channel classification and supervised classification. At last, results of accuracy assessments on the three classification methods are compared, and the rules of knowledge-based classification are analyzed. Based on these analyses, the feasibility of application of the knowledge-based classification method is recognized.

Currently the landuse/cover change (LUCC) research is the core of global change research, and the attainment of landuse/cover data

is the primary foundation work. In recent years, rapid development of remote sensing technology has become an effective way of landuse/cover investigation. However, the classification methods of remote sensing images are far behind the development of the sensors. In order to satisfy the requirements of obtaining thematic information rapidly and correctly from the magnitude remote sensing data, the realization of automatization and high precision quantification of image interpretation is an essential requirement for the development and application of remote sensing.

Due to the complexity of landuse/cover in the typical areas of the lower reaches of Yangtze River, the limitation of image resolution and the disturbance of other factors in the imaging process, the phenomena of 'same objects with different spectrum and different objects with similar spectrum' is serious. It's very difficult to obtain the satisfactory effect by supervised or unsupervised automatic classification methods based on spectral characteristics. While the knowledge-based classification is not only helpful in synthesizing the spectrum and structure characteristics of images, but also in integrating the geographical principles and expert's experience of the distribution of the objects on the earth, which provides an effective solution to integrate multi-resource data to accomplish the complex landuse/cover classification. However, knowledge attainment is a bottleneck of the wide application of knowledge-based image classification method. Currently the way to attain the knowledge is mostly finished manually by knowledge engineers and interpreters. The interpreters generalize the principles of the spectral characteristics, the distribution and the change of the

typical samples and their experiences and professional knowledge. The knowledge engineers extract the knowledge by talking with the interpreters, and transform them to the specific forms that the computer can deal with. This mechanical approach is not only of low efficiency, time consuming, but also much more subjective. It decreases the intelligence and objectivity of knowledge-based image classification in some extent and impedes its wide application. In the 1990s, the forthcoming technique, Spatial Data Mining and Knowledge Discovery (SDMKD) breaks a new path for attaining the knowledge. It can discover the spatial geometry characteristics, the principle of the spatial distribution, spatial associated rules, distinct rules and spatial evolvement rules from the spatial data.

Based on this background, this research takes two typical low hill and upland areas of the lower reaches of Yangze River as test areas: Shaoxing city in Zhejiang province and Jiangning county which is near Nanjing city in Jiangsu province. The knowledge-based landuse/cover classification by middle-scale images, TM and SPOT - XI is discussed. The main research contents and conclusions include the following:

1. The features of all types of objects on the images are analyzed through field investigation and image interpretation. And these associated characteristics of all types of objects are analyzed and expressed. The spectral response characteristics of image are analyzed by the statistics of typical samples of all types of objects. The textural characteristics are analyzed based on the grey co-occurrence matrix. The proper parameters are selected by statistics of typical samples of all types of objects. The spatial geometry characteristics are analyzed,

choosing four measures like area, perimeter, the shape index and compactness to express the water's geometry characteristics. The principles of distribution of all types of objects in the test areas are discussed, choosing geographical coordinate, DEM, slope and aspect as measures. The latter investigations suggest that assisted by the textural characteristics and the principles of the distribution of all types of objects can improve the accuracy of land use/cover classification in the low hill and upland area of lower reaches of Yangtze River. On the one hand, the textural characteristics can enhance the information of man-made constructions such as residential areas and roads that have obvious spatial geometry structures. On the other hand, they can enhance the difference between shadow and other objects at mountain undulate area. The topographical factors mainly increase the difference of those land use/cover types with obvious spatial distribution principles, such as paddy, woodland and garden, etc. The spatial geometry characteristics can be used to subclassify the water effectively, which suggests that the knowledge-based classification method can not only improve the accuracy, but also increase the level.

2. To describe the knowledge-based remote sensing image classification method and apply it to the land use/cover classification in two test areas. The definition of the concept of knowledge, the knowledge attainment by decision tree, the establishment of the knowledge base and the deduction based on it are discussed. Emphasis is made on the method of using decision tree grew by Classification And Regression Tree (CART) to obtain the rules of classification. The CART has following advantages: It has no

assumption requirements of statistical distribution of input data, they can either be independence or correlation, either be continuous or discrete. It can deal with the non-linear relationship between eigenvector and target variable. It can dispose multi-dimension data effectively. It contains the missing data and mistakes. It can clearly judge the importance of variables for classification. The structure is clear and easy to be understood which can be transformed to rules. This research sets up many rules from decision tree grew by CART, and gets the satisfactory results of landuse/cover classification in two testareas. So it is feasible to attain the classification rules by this method, which will be an effect tool to solve the bottleneck problem of wide application of knowledge-based landuse/cover classification.

3.To analyze and evaluate the knowledge-based classification from the accuracy assessments and comparisons of classification by three methods and the analysis of the classification rules. The logic channel classification and supervised classification are also performed on the landuse/cover classification in two testareas. The accuracy assessments and comparisons of the three classification methods are analyzed separately based on the confusion matrix. The results suggest that although the situation of landuse/cover is complex in two testareas and shortage of concrete validates at the same time, there are some errors of the accuracy assessment, the accuracy of the knowledge-based classification method is higher than the supervised classification and logic channel classification because it synthesizes the characteristics of spectral response, textural characteristics and the principles of spatial distribution. And as for the Jiangning testarea, it is more complex with the finer and more fragmental

parcels, so the accuracy is improved obviously. Furthermore, it can be learnt from the analysis of the rules that the classification rules are reasonable, which react the fact of the testareas. This also suggests that it is effective to set up the rules base on the decision tree grew by CART.

4.The knowledge-based classification method can effectively synthesize all kinds of assistant information to improve the accuracy. But the quality of assistant data and the match precision between the data also may affect the accuracy. So we must pay attention to the assistant data selection and the rigorous match between the different data in the practical application.

As the above mentioned, this research mainly puts forward a new method of knowledge-based automatic classification based on the multi-resource data in the typical testareas of the lower reaches of Yangtze River where the parcels are fragmental, and the types of landuse/cover are complex. It can provide a good guidance for landuse/cover classification by middle-scale remote sensing images in other areas.

目 录

第 1 章 绪 论	1
1.1 研究背景	1
1.2 研究思路与技术路线	11
1.3 研究内容与目的	14
1.4 本书的框架结构	14
第 2 章 研究区域概况及数据资料	16
2.1 研究区域概况	16
2.2 数据资料与数据预处理	19
2.3 研究区土地利用 / 覆被类型的影像表征	23
第 3 章 特征的分析与表达	26
3.1 图像光谱响应特征的分析与表达	26
3.2 纹理特征的分析与表达	32
3.3 空间几何特征的分析与表达	44
3.4 位置布局特征的分析与表达	47
第 4 章 知识的获取与知识库的构建	49
4.1 知识的定义	49
4.2 知识获取的决策树方法	51
4.3 知识库的构建	59
4.4 试验区分类知识的获取与知识库的构建	60

第 5 章 基于知识库的推理分类与评价	76
5.1 基于知识库的推理分类	76
5.2 分类精度的评价	81
5.3 分类规则的分析	88
5.4 分类方法的可行性评价	89
第 6 章 结 语	90
6.1 结论与讨论	90
6.2 创新点	92
6.3 展 望	92
主要参考文献	94
附 录	108
中文图表索引	108
英文图表索引	110
后 记	113

CONTENTS

Chapter 1 Introduction	1
1.1 Background	1
1.2 Research Thoughts and Techniques	11
1.3 Research Contents and Purposes	14
1.4 Structure of the Book	14
Chapter 2 Survey of Study Area and Data	16
2.1 Survey of Study Area	16
2.2 Data and Data Preprocessing	19
2.3 Image Characteristics of Landuse/Cover in the Study Area	23
Chapter 3 Analysis and Expression of Characteristics	26
3.1 Analysis and Expression of Image Spectrum Characteristics	26
3.2 Analysis and Expression of Image Texture Features	32
3.3 Analysis and Expression of Spatial Geometric Features	44
3.4 Analysis and Expression of Local Distribution Characteristics	47
Chapter 4 Knowledge Acquisition and Knowledge	
Base Construction	49
4.1 Definition of the Knowledge	49
4.2 Decision Tree Method	51
4.3 Knowledge Base Construction	59
4.4 Classification Knowledge Acquisition and Knowledge Base	
Congstruction of Study Area	60