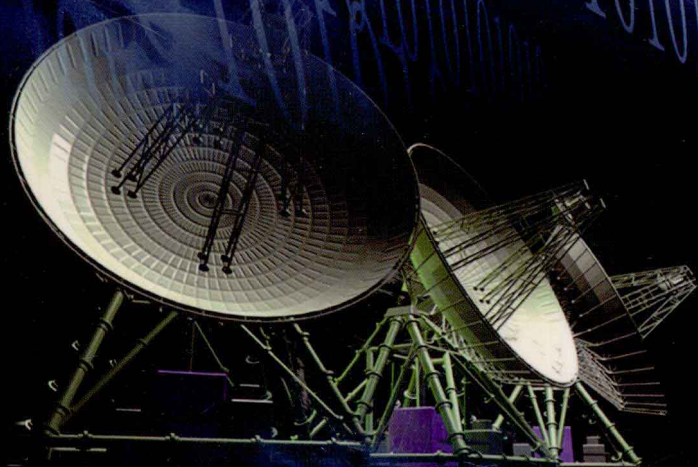


遥感信息的不确定性研究 分类与尺度效应模型

柏延臣 王劲峰 著

FENLEI YU CHIDUXIAOYING MOXING



地质出版社

遥感信息的不确定性研究

分类与尺度效应模型

柏延臣 王劲峰 著

地质出版社

· 北 京 ·

内 容 提 要

遥感数据专题分类不确定性问题及与其相关的尺度问题是当前遥感研究和应用中的重要问题。本书对当前遥感专题分类不确定性评价方法和尺度问题进行了详尽的评述;发展了基于“扩展概率矢量”的像元尺度上评价遥感分类不确定性的一般模型;发展了多分类器结合的遥感专题分类方法并验证了该方法在降低分类不确定性方面的有效性;发展了以统计可分性为基础的遥感分类不确定性的尺度效应评价模型;提出了利用模拟点扩散函数进行遥感信息尺度转换的方法;发展了简单快速的利用变异函数进行遥感信息合适分辨率选择的方法。

本书可供遥感、地理信息系统、测绘及其相关学科的科研人员和高等院校师生参考。

图书在版编目 (CIP) 数据

遥感信息的不确定性研究/柏延臣, 葛咏, 王劲峰著. -北京: 地质出版社, 2003.4
ISBN 7-116-03815-9

I. 遥… II. ①柏…②葛…③王… III. 遥感-信息技术-不确定系统-研究
IV. TP7

中国版本图书馆 CIP 数据核字 (2002) 第 026697 号

YAOGAN XINXI DE BUQUEDINGXING YANJIU

责任编辑: 祁向雷

责任校对: 李 玫

出版发行: 地质出版社

社址邮编: 北京海淀区学院路 31 号, 100083

电 话: (010) 82324508 (邮购部); (010) 82324577 (编辑室)

网 址: <http://www.gph.com.cn>

电子邮箱: zbs@gph.com.cn

传 真: (010) 82310759

印 刷: 北京印刷学院实习工厂

开 本: 787mm × 960mm^{1/16}

印 张: 7.25

字 数: 304 千字

印 数: 1—800 册

版 次: 2003 年 4 月北京第一版·第一次印刷

定 价: 30.00 元 (共两册)

ISBN 7-116-03815-9/T·109

(凡购买地质出版社的图书, 如有缺页、倒页、脱页者, 本社发行处负责调换)

目 录

第一章 引言	(1)
第一节 研究背景	(1)
第二节 本书内容及组织结构	(3)
第二章 遥感分类信息的不确定性研究: 进展与问题	(6)
第一节 基本概念	(6)
第二节 遥感信息中的不确定性来源	(7)
第三节 遥感数据分类不确定性评价方法	(10)
第三章 遥感信息提取中的尺度问题	(17)
第一节 自然和社会科学中的尺度问题	(17)
第二节 遥感信息尺度问题研究	(20)
第三节 遥感信息提取中不确定性和尺度的关系	(27)
第四章 像元尺度的遥感数据专题分类不确定性表达与评价	(29)
第一节 引言	(29)
第二节 像元尺度上遥感数据专题分类的不确定性度量	(31)
第三节 对概率矢量所表达的不确定性的理解	(43)
第五章 降低分类不确定性——多分类器结合方法	(46)
第一节 各种遥感数据分类方法比较	(46)
第二节 多分类器结合的分类方法	(53)
第三节 标准的多分类器结合方法	(55)
第四节 不同分类器结合方法进行遥感数据分类的实验	(57)
第六章 遥感数据分类不确定性的尺度效应	(68)
第一节 引言	(68)
第二节 遥感数据分类不确定性的尺度效应研究	(69)
第三节 统计可分性度量	(71)
第四节 遥感数据分类不确定性的尺度效应分析	(73)
第七章 遥感信息尺度转换及合适分辨率的选择方法	(78)
第一节 引言	(78)
第二节 基于模拟点扩散函数的尺度转换方法	(79)
第三节 遥感信息尺度效应的空间统计学分析	(81)
第四节 实例分析	(85)

第五节 讨论	(92)
第八章 总结与展望	(94)
第一节 内容总结	(94)
第二节 研究展望	(95)
参考文献	(98)

第一章 引言

第一节 研究背景

一门学科在不同的发展阶段总是具有相应的研究特点。地球科学作为一门古老的科学，经过了从综合研究到部门研究再到综合研究的长期发展过程。这种发展轨迹既是现实应用的需要，也是不同阶段学科发展水平的必然。资源、环境、人口是事关全人类发展的问题，要求进行综合的、从区域到全球的多尺度综合研究，这对现代地球科学的发展提出了新的要求。空间科学、信息科学以及计算机技术的发展为地球科学的发展提供了新的机遇，从而产生了一门新兴的交叉学科：地球信息科学。其中地理信息系统作为一种采集、存储、管理、分析和显示空间定位信息的一种技术，是地球信息科学的核心技术之一。陈述彭院士指出：地理信息系统的出现是近 40 年来地理学发展中的一个革命性事件。同时，现代地球科学大范围、多尺度和动态研究的特点，使遥感成为最有价值的地理信息系统的信息源。因此，如何从遥感数据中提取高精度的信息，是遥感研究和应用中的核心问题之一。如何表达、评价及如何降低所提取信息的不确定性是重要的研究方面之一。

遥感信息提取中的不确定性是当前遥感研究的一个热点。人们总是希望从遥感数据中提取的信息完全客观准确地反映实际情况，但由于自然环境的复杂性，以及自然环境与遥感波谱相互作用的复杂性，从传感器记录的光谱信号中提取的关于地表的信息中，总是存在不确定性，因此，在使用从遥感数据得到的专题图或某一地表参数的分布信息时，需要了解这些信息的不确定性。

空间数据的不确定性是评价空间数据质量的最主要方面。地理空间数据的质量对数据生产者和用户来说都是一个非常重要的考虑因子，它可以使数据生产者正确描述他们的数据符合生产规范的程度，也是用户决定数据是否符合他们应用目的的依据。同时，当这些信息作为地理信息系统数据源，在地理信息系统中经过各种操作或进行各种模型运算，并最终为决策支持服务时，数据源中的不确定性会在每一步操作和模型运行中传播，使模型运行的结果具有不确定性，从而使决策具有不确定性。要评价模型输出以及决策中的不确定性，必须首先评价数据源的不确定性。因此，研究从遥感数据中提取的信息的不确定性具有重要和现实的意义。

地理空间数据的不确定性在世界各地得到广泛研究。例如,在地球观测系统(EOS)的科学计划中,具有专门的“EOS 验证计划(Validation Program for EOS)”。我国已经在这一领域开展了研究。2001年,在中国科学院遥感应用研究所成立了遥感误差分析实验室。两年一度的“自然资源与环境科学中的空间精度评价国际会议”已经成为研究空间数据精度的科学家们的盛会。该会议已经成功召开五届。第一届于1994年在美国 Virginia 的 Williamsburg 召开,第二届在美国 Colorado 召开(1996),第三届在加拿大 Quebec City 召开(1998),第四届在荷兰的 Amsterdam 召开(2000),第五届在澳大利亚的 Melbourne 召开(2002)。在历届会议文集中,有许多论文是讨论遥感信息提取中的不确定性问题。大量的关于遥感信息不确定性的研究论文发表在《Photogrammetric Engineering and Remote Sensing》,《IEEE Transactions on Geoscience and Remote Sensing》,《Remote Sensing of Environment》,《International Journal of Remote Sensing》以及《International Journal of Geographical Information Science》等刊物上。

多尺度是遥感信息的一个基本特征。现有的遥感数据的空间分辨率包括从0.61 m到数十公里,可以从多个空间尺度进行对地遥感观测。不同空间分辨率的遥感数据所包含的地表信息的详尽程度不同。同时,地表景观本身具有多尺度的层次组织结构。从遥感数据中提取不同尺度层次的景观特征的信息需要不同空间尺度(空间分辨率)的遥感数据。根据所研究的目标,选择合适的空间分辨率的遥感数据可以提高信息提取的精度,降低分析结果的不确定性。一般高空间分辨率的遥感数据含有更多的信息,但并不是空间分辨率越高,信息提取的精度越高,不确定性越低。就遥感数据土地覆被分类来说,分类精度和地表覆被类型单元的大小有关,过高的空间分辨率反而会降低分类精度。从这个意义上,研究尺度因子在遥感信息提取中的作用,尺度与不确定性的定性和定量关系,以便选择合适的空间分辨率的遥感数据,不但可以降低信息提取的不确定性,而且还可以排除选择数据时的盲目性,并降低数据成本。

另一方面,现有的地球观测系统中的遥感器只能提供若干个离散的尺度(空间分辨率)的遥感数据,而不是连续变化尺度上的数据。但在实际应用中,常常需要特定空间分辨率的数据,而现有的遥感传感器并不能提供这样空间分辨率的数据。这种情况下,就需要对现有空间分辨率的数据进行尺度转换,以满足不同应用的需求。因此,如何对遥感数据(或者,更广义地,空间数据)进行尺度转换,就成为重要的研究问题。

遥感信息的尺度问题研究是当前国内外的研究热点。国外至少在20世纪70年代末期就开展了相关领域的研究。我国已在这一领域开展了许多研究工作。由李小文院士主持的我国重点基础研究项目(973项目)“地球表面时空多变要素的定量遥感理论及应用”中,“在像元尺度上,以研究尺度效应和尺度转换理论为核心,检验与反演目标参数有关的基本物理概念、定理、定律的适

用性”被列为总体目标和要解决的关键科学问题之一。此外，一系列关于遥感中尺度问题的国际会议也相继召开并出版了相关文集。例如，1993年在法国召开的热红外遥感尺度问题国际会议，尺度问题被认为是空间对地观测中的主要挑战；1997年召开了遥感和GIS中的尺度问题国际会议；1996年召开了关于利用遥感信息进行水文学中尺度转换的国际会议。2001年7月在中国黄山召开的第一届国际热红外遥感会议上，尺度问题被列为最主要的讨论主题之一。

遥感信息的不确定性和尺度问题绝不是独立存在的问题。遥感数据的空间尺度（空间分辨率）是影响遥感信息提取不确定性的的重要因素，要提取高精度的信息，需要选择合适空间分辨率的遥感数据。同时，在遥感信息不确定性的评价过程中，需要考虑从遥感数据中提取的信息与参考数据空间尺度的差异。例如，当以点位观测数据验证遥感信息提取的精度时，常常需要对点位观测信息进行尺度扩展。

第二节 本书内容及组织结构

遥感信息的不确定性和尺度问题是一个涉及广泛的多学科问题。传统的遥感分类不确定性的评价方法以基于误差矩阵（又称混淆矩阵）的方法为代表。这类方法可以在类别尺度上表达和评价遥感数据分类结果的不确定性，但无法反映分类不确定性的空间分布结构，使评价结果既不便于可视化表达与分析，也不便于不确定性在空间模型中的传播方面的研究。同时，这类方法还受到采样方法的限制和对地面真实状况的了解等因素的局限。因此，如何在像元尺度上评价遥感数据分类不确定性是当前面临的主要问题之一。在不确定性评价基础上，降低分类不确定性一直是遥感研究中的主要问题。对于特定的研究区域，遥感分类不确定性除了与分类方法有关外，还和所用的遥感数据的空间尺度，即空间分辨率有关。对于遥感数据的空间分辨率如何影响数据分类的不确定性，目前尚未得到普遍重视，更缺乏系统的研究。同时，对于特定的应用，如何选取合适空间分辨率的数据和进行数据的尺度转换评价，也是当前迫切需要研究的问题。针对这些问题，本书以遥感信息提取中的不确定性和尺度效应为核心，着重研究遥感专题分类的不确定性表达和评价，从分类器的角度降低遥感分类不确定性的方法，遥感信息专题分类不确定性的尺度效应，以及遥感信息的尺度转换方法及合适分辨率的选择问题。本书共分八章：

第一章为引言，提出遥感信息不确定性和尺度效应研究的背景、意义，介绍全书的研究内容及组织结构。

第二章讨论不确定性的概念，遥感专题分类不确定性评价的历史、方法、进展，以及存在的问题。

第三章讨论遥感信息尺度问题的概念、研究问题和研究方法，研究进展以及存在的问题，并进一步讨论遥感信息不确定性和尺度之间的关系。

第四章研究像元尺度上遥感数据专题分类不确定性的表达和评价问题。分析了基于概率矢量的不确定性表达方法及从概率矢量中计算的各种不确定性度量指标，指出概率残差和信息熵是像元尺度上评价分类不确定性的两个较好的指标。针对概率矢量只能从贝叶斯分类过程中得到的特点，提出了“扩展的概率矢量”的概念，使基于概率矢量的分类不确定性评价方法也可以用于传统的各种距离分类器和模糊分类器等。此外，还对基于概率矢量的不确定性指标所表达的不确定性的意义进行了探讨。

第五章研究从分类器的角度降低分类不确定性的途径。将近年来模式识别领域中发展的多分类器系统的思想引入到遥感数据的专题分类，特别验证了标准的多分类器结合方法在遥感专题分类中的有效性，对抽象级和测量级的多分

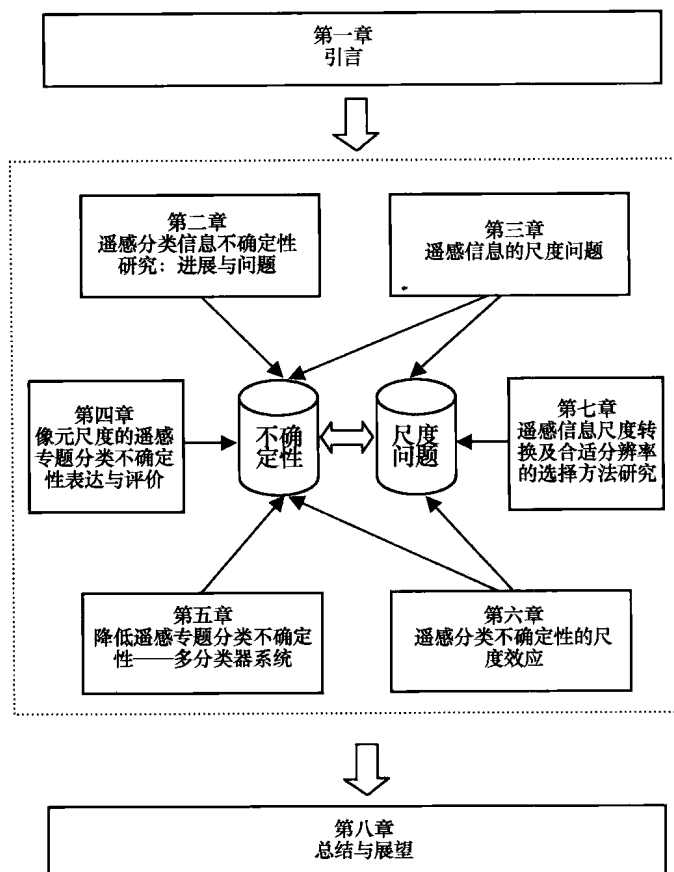


图 1-1 本书主要研究内容及框架

Fig.1-1 Main content and framework of the whole book

类器结合方法的特点进行了评价。

第六章研究了遥感专题分类不确定性的尺度效应问题。遥感专题分类不确定性的来源有多种。传统的直接将不同空间分辨率数据的分类不确定性直接比较的方法无法区分由尺度变化导致的不确定性的变化和由分类方法本身导致的不确定性，而且计算强度大。针对这种情况，提出了以遥感专题分类中类别特征之间的统计可分性为度量研究分类不确定性尺度效应的方法，指出了随着遥感数据空间分辨率变化，不同类对之间分类不确定性变化结构的差异性及其与类别空间分布结构之间的关系。

第七章研究了遥感信息的尺度扩展方法及合适尺度（空间分辨率）的选择问题。遥感信息尺度扩展方面，通过以模拟的不同空间分辨率遥感传感器的点扩散函数与高空间分辨率的遥感数据进行卷积运算，将高分辨率数据尺度扩展到不同的低分辨率，并以遥感数据反演的地表辐射温度为例，与传统的图像尺度扩展方法进行了比较；合适空间尺度的选择方法方面，在分析和评价了已有方法的基础上，提出了通过分析以一个像元大小为步长的图像变异函数的结构，简单快速的选择合适分辨率的方法，并与传统方法的结果进行对比，验证了该方法的可靠性。

第八章对全书进行了总结，并在此基础上对进一步的研究做了展望。

本书的主要结构组织如图 1-1 所示。

第二章 遥感分类信息的不确定性研究：进展与问题

从遥感数据中获取地面目标分布的专题信息（如土地覆被/土地利用等），是从遥感数据中获取的主要信息之一。通过对遥感数据进行目视解译或自动分类，可以获得区域和全球尺度的土地覆被/土地利用变化信息。受数据源特性、数据处理方式，分类方法等多种因素的影响，遥感数据的专题分类结果不可避免地具有不确定性。对分类结果的不确定性进行分析，不但可以使用户了解数据的质量和适用性，还可以帮助检验分类方法的有效性，以便改进或建立新的分类方法。

第一节 基本概念

- 误差 (Error): Heuvelink (1993) 将误差定义为：“现实与对现实的表达之间的差异。它不但包括错误 (mistakes or faults)，而且包括统计意义上的‘变化 (variation)’ …”。从这个意义上说，误差是一个贬义词。正如 Chrisman (1991) 所指出的，“…Error is a bad thing…”。在大部分遥感分类精度或误差评价的文献中，分类误差指某一像元被赋予的类型与此像元所代表的真实类型的差别。

- 精度 (Accuracy): 精度被定义为“观测、计算或估计值与真实值之间的接近程度 (closeness)” (AGI, 1991)。在统计意义上，精度也可以理解为观测、计算或估计的均值与真实均值之差 (Burrough, 1986)。在这个意义上，为了评价数据的精度，就必须有一个更高精度的数据作为“实况”。但是，正如 Drummond (1995) 所指出的那样，“真实状况也许永远无法得到…”。Aronoff (1989) 将遥感分类精度定义为：“…给地图上某一位置赋予的类别为该位置真实类别的概率”。这个定义强调统计意义上的精度。Story and Congalton (1986) 将分类精度分为总体精度 (overall accuracy)、生产者精度 (producer's accuracy) 和用户精度 (user's accuracy)，其中“用户精度”对应于 Aronoff (1989) 所定义的精度。

- 精确度 (Precision): 精确度被定义为“某一值被表达的精确程度 (exactness)，而不管是对是错” (AGI, 1991)。Unwin (1995) 将精确度更通俗地定义为“测量值的小数点位数”。这意味着精确度是和精度无关的量。高的精确

度并不意味着高的精度，相反，可能虽然数据的表达非常精确，但是完全错误的。Goodchild (1993) 和 Campbell (1987) 将属性数据的精确度定义为“详尽程度 (detail)”或遥感分类过程中定义的类别数目，它代表分类所得专题图的综合水平。

- 质量 (quality): 数据质量定义为数据的“适用性 (fitness for use)”。既然数据质量反映数据的适用性，那么它和特定的数据应用目标相关。因此质量是一个相对的概念。

- 不确定性 (uncertainty): 不确定性的概念被越来越多的引入空间数据的处理和空间数据质量评价中来。Heuvelink (1993) 将不确定性作为误差 (error) 的同义词。Goodchild (1995) 则将不确定性看作是一个误差的更一般的 (generic) 的测量，且不隐含关于误差源以及误差能否纠正的任何信息。Van der Wel (2000) 将不确定性理解为“...a useful concept to express the inability to be confident of, and knowledgeable about the truth value of a particular data characteristic”。本文中的不确定性的意义采用 Van der Wel 的理解，并将不确定性作为比误差更一般的度量。而误差和精度则是不确定性的一种特定的表达方式。

第二节 遥感信息中的不确定性来源

在遥感数据的生命周期中，从数据的获取、处理、分析、数据转换等各种操作中，都会引入不同类型和不同程度的不确定性，并在随后的各种处理过程中传播。最终的总的不确定性则是各种不确定性不断积累的结果。不同的操作也可能引入不同的不确定性。图 2-1 展示了典型遥感信息处理过程中误差的累积。

在遥感信息处理过程中，每一步都可能引入不同类型的不确定性。Bedard (1987) 从地学应用的角度将遥感信息中的误差分为位置误差 (locational error)

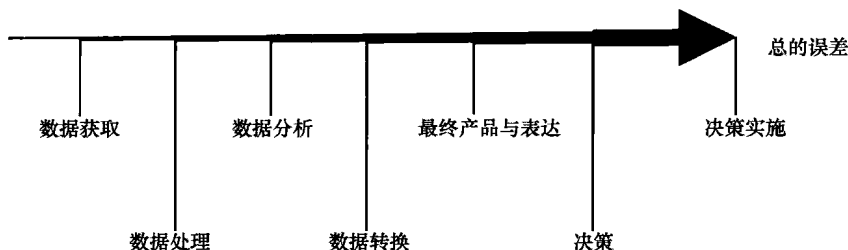


图 2-1 典型遥感信息处理过程中误差的累计

(据 Lunetta et al. 和 Ppul Smits, 1998)

Fig. 2-1 The accumulation of error in a 'typical' remote sensing information processing

(by Lunetta et al. 1991 and Paul Smits, 1998)

和描述性误差 (descriptive error); Chrisman (1991) 则将其称为位置误差 (positional error) 和属性误差 (attribute error); Veregin (1989) 则将其称为图形误差 (cartographic error) 和专题误差 (thematic error)。Goodchild and Wang (1988) 从数据处理不同阶段出发, 将遥感信息中的误差分为数据源误差 (source error), 处理误差 (processing error) 和产品误差 (product error); 无独有偶, Beard (1989) 将其分为源误差, 处理误差和应用误差 (use error)。

在遥感信息处理流程的各个阶段, 引入不确定性的基本因素如图 2-2 所示。

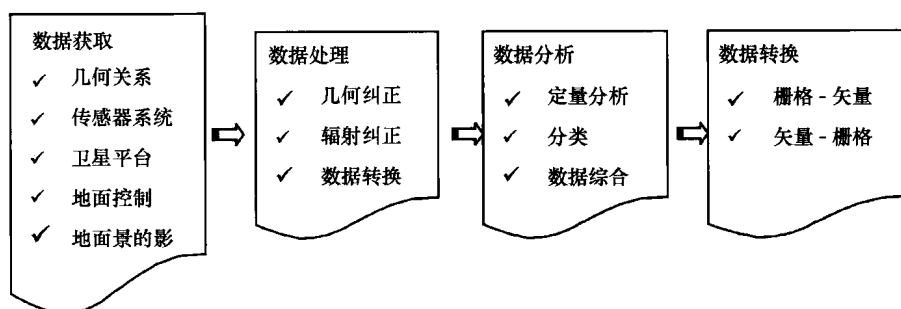


图 2-2 遥感信息处理流程中引入不确定性的因素

(据 Lunetta, 1991)

Fig.2-2 The factors introducing uncertainty in the processing flow of remote sensing data

(by Lunetta, 1991)

(1) 在数据获取阶段, 传感器与地面景的几何关系直接影响图像的质量。理想状态下, 在一景图像中的辐照几何关系应该是一个常数。当在瞬时视场角 (IFOV) 较大时, 这种几何关系在一个范围内变化, 导致图像内的辐射畸变。同时, 这种大视场角也会导致图像内的几何畸变。

(2) 传感器系统的特性对图像质量以及信息提取的不确定性有显著影响。传感器的物理参数不但决定获取遥感数据的空间、时间和辐射分辨率, 还决定数据的信噪比。同时, 传感器工作的电磁波范围决定了云对数据获取的影响以及大气吸收、辐射和散射等造成的遥感数据的辐射畸变。由传感器参数而决定的空间分辨率直接影响遥感信息提取 (如分类) 的精度。卫星平台的轨道高度, 飞行速度和传感器的瞬时视场角一起决定图像的几何特性。卫星平台的稳定性严重影响遥感系统的几何精度。遥感图像所覆盖的地面景的复杂性也对遥感数据中的不确定性有影响。地形起伏不但会导致图像几何畸变, 而且会导致图像内迎光坡和背光坡均匀区域内光谱响应的巨大差异, 导致信息提取结果中大的不确定性。对于特定物理参数的传感器所获取的遥感图像, 地表景观分布的复杂性和地表单元的大小共同直接影响遥感数据的分类不确定性。这些在数

据获取时引入的不确定性有些可以纠正,有些则无法纠正或处理,但它们终将影响最后遥感信息中的不确定性。例如,葛咏(2001)从机理的角度研究 SAR 的不确定性传递。

(3) 遥感数据的处理过程也会引入不确定性。根据不同的应用目标,需要对遥感数据进行不同的处理。一般来说,对图像进行几何纠正、辐射纠正和数据转换是最基本的处理步骤。遥感图像的几何纠正一般通过选取地面控制点和图像上同名的点,在地面控制点坐标和同名点的图像坐标之间建立多项式,而在两个坐标系之间建立联系,然后将遥感图像从图像坐标系转换到地面控制点的坐标系。通常,利用均方根误差(root-mean-square error, RMSE)衡量参考像元的位置精度(Janssen and Van der Wel, 1994)。地面控制点的选取对图像几何纠正精度起决定性作用。一般地面控制点从大比例尺的地形图选取。由于地形图的不同地图综合程度,控制点的位置总是存在一定偏差。即使地面控制点位置通过 GPS 精确得到,对于不同空间分辨率的遥感数据而言,在图像是选取相应于地面控制点的同名点的位置也会存在一定偏差。几何纠正过程中的重采样过程也可能引起图像整体或局部亮度值的变化,从而影响数据分析结果的精度。Smith and kovalick (1985)比较了不同重采样方法对图像分类精度的影响。辐射纠正的目的是消除大气效应、地形效应以及传感器引入的遥感数据中的辐射畸变。不同的辐射纠正方法以及辐射纠正模型中参数的精度常常影响辐射纠正效果,并可能引入新的不确定性。

上述不确定性会在遥感数据的进一步分析中传播并导致分析结果的不确定性,数据分析过程本身也会引入新的不确定性。就遥感数据分类而言,在分类过程的各个阶段都可能引入不确定性。在分类系统设计过程中,类型定义的不完备或模糊性,没有考虑混合像元,不同分类体系集成中的不兼容等(Lunetta et al., 1991),都会在分类结果中引入不确定性;在分类过程中,训练样本数量、训练样本的代表性直接影响分类结果的精度。分类算法本身也是一个重要的不确定源。不同的分类算法所产生的分类结果,其不确定性会有很大差别。

由于从遥感数据的获取、处理、分析和转换等一系列步骤中都有不同类型和程度的不确定性引入,并在进一步分析中传播。在遥感信息提取过程中,不但要设法纠正数据获取过程中引入的不确定性,而且要选择合适的、对误差敏感的分析和处理方法,使最后提取的信息包含最小的不确定性。由于本文的重点是研究遥感数据分类结果的不确定性,我们将重点探讨遥感数据分类方法以及由此导致的结果的不确定性。但我们必须明确,最终分类结果的不确定性是数据获取、处理和分析等过程中不同不确定性传播、积累的结果。只有在数据源和各处理步骤完全相同的前提下,才有可能通过评价分类结果的不确定性来评价分类方法的有效性。

第三节 遥感数据分类不确定性评价方法

自 20 世纪 70 年代以来, 遥感数据就被广泛用于土地利用/土地覆被制图、资源调查、环境和自然灾害监测等领域, 但直到 80 年代, 人们才开始研究遥感数据分类的不确定性问题。遥感数据分类的不确定性评价方法也经历了一个逐步细化和严格的过程。Congalton (1994) 将分类不确定性评价的发展分为四个阶段: 第一阶段的精度评价方法以目视判断为主, 这种精度评价方法是一种定性的评价方法, 而且具有很大的主观性。第二阶段精度评价方法由定性发展到定量方法。这一阶段中, 精度评价主要通过比较分类所得的专题图中各类别的面积范围 (或面积百分比) 与地面或其他参考数据中相应类别的面积范围 (或面积百分比)。与第一阶段的方法比较, 这种评价方法具有定量和客观的优点。但这种方法的最大局限在于其非定位 (non-site specific) 本质。因为分类专题图中的某些类别面积即使占有正确的百分比, 但它可能在错误的位置。因此, 这种评价方法可能掩盖分类结果的真实精度。第三个阶段以定位 (site specific) 类别比较和精度测量 (accuracy metrics) 为特征。在这一阶段, 精度评价通过比较特定位置的分类结果中的类别和地面实况或其他参考数据中相应点的类别为基础, 并在此比较基础上发展了各种精度测量 (如总体精度等)。第四阶段的评价方法是在第三阶段方法基础上的细化和发展。其核心是误差矩阵方法, 特点是在充分利用误差矩阵信息的基础上, 发展各种精度测量 (如 kappa 系数), 且统计上更为严格。

误差矩阵精度评价方法依然是当前遥感分类精度评价的核心方法, 而且有人建议将其作为分类精度评价的标准方法 (Smits et al., 1999)。但以误差矩阵为基础的精度评价方法存在诸多局限性。随着对遥感数据分类问题认识的深入, 以及不同精度评价目标的需求, 科学家们发展了许多新的误差评价方法和指标。同时, 基于误差矩阵的精度评价方法也在不断发展和完善。

表 2-1 列举了一些主要的遥感分类精度评价方法及其精度测量指标。从评价方法的角度, 可以将它们归类为基于误差矩阵的方法、基于模糊分析的方法和其他各种方法三大类。

一、基于误差矩阵的分类精度评价方法

1. 误差矩阵及其精度测量

误差矩阵 (Error matrix) 又称混淆矩阵 (Confusion matrix), 是一个用于表示分为某一类别的像元个数与地面检验为该类别数的比较阵列 (Congalton, 1991)。通常, 阵列中的列代表参考数据, 而行代表由遥感数据分类所得的类别数据。一个典型的误差矩阵如图 2-3 所示。从误差矩阵中可以直观地得到每

表 2-1 各种分类精度评价方法及特点

Table 2-1 An overview of the classification accuracy assessment methods

作 者	关 键 字	总体精度	每一类精度	分类器比较
Rosenfield (1981)	方差技术	✓	—	—
Aronoff (1982, 1985)	生产者风险和用户风险	✓	✓	—
Congalton (1991, 1999)	误差矩阵, Kappa 系数	✓	✓	✓
Czaplewski (1992), Zhu (1996)	多变量合成估计	✓	—	—
Gopal and Woodcock (1994), Woodcock and Gopal (2000)	模糊分析	✓	✓	✓
Ma and Redmond (1995)	Tau 系数	✓	✓	✓
Richards (1996)	贝页斯估计	✓	✓	✓
Paul Smits (1999)	基于成本的误差评价	✓	✓	✓
Zhuang et al. (1995)	Tukey 分析	—	—	✓

(据 Paul C. Smits et al., 1999, 经修改)

		参考数据				行总数
		A	B	C	D	n_{i+}
分类数据	A		N_{12}	...	N_{1k}	N_{1+}
	B	N_{21}		...	N_{2k}	N_{2+}
	C	...	...		...	...
	D	N_{k1}	N_{k2}	...		N_{k+}
列总数		n_{+1}	N_{+2}	...	N_{+k}	N

图 2-3 误差矩阵示意图

Fig. 2-3 The illustration of error matrix

一类别的包含误差 (commission error) 和丢失误差 (omission error)。包含误差指不该属于某类别的像元被分为该类别的误差, 它由该类别所在行的非对角线元素之和除以该行总和而得; 丢失误差指该属于某一类别的像元未被分为该类别的误差, 它由该类别所在列的非对角线元素之和除以该列的总和而得。

误差矩阵除了清楚地显示各类别的包含误差和丢失误差外, 还可以从误差矩阵中计算出各种精度测量指标, 如总体精度 (Overall accuracy), 生产者的精度 (Producer's accuracy) 和用户的精度 (User's accuracy) (Story and Congalton, 1986)。总体精度是误差矩阵内主对角线元素之和 (正确分类的个数) 除以总的采样个数。生产者精度和用户精度可以表示某一单个类别的精度。生产者精

度为某类别正确分类个数除以该类的总采样个数（该类的列总和）；而用户精度定义为正确分类的该类的个数除以分为该类的采样个数（该类的行总和）。总体精度、生产者精度和用户精度的计算公式见 R.G. Congalton and K.Green (1998)。

除了以上各种描述性的精度测量，在误差矩阵基础上利用各种统计分析技术，可以用于比较不同的分类方法，其中最常用的是 Kappa 分析技术。

Kappa 分析技术 (Cohen, 1960; Stehman, 1996; Congalton and Mead, 1983) 是一种多变量统计分析技术，它在统计意义上反映分类结果在多大程度上优于随机分类结果，并可以用于比较两个分类器的误差矩阵是否具有显著差别 (Congalton, 1999; Smits, 1999)。Kappa 分析的结果是 KHAT 统计。从误差矩阵中可以计算总体分类的 KHAT 统计值和各类别的条件 Kappa 系数 (Conditional Kappa Coefficient)。其计算公式见 R.G. Congalton and K.Green (1998)。

一般的 Kappa 系数及其方差的估计方法是在假设采样模型为多项式模型的基础上发展的，而只有简单随机采样方法满足这个假设 (Congalton, et al., 1999)。Stehman (1996) 探讨了分层采样条件下的 Kappa 系数及其方差的估计方法。

除了 Kappa 分析技术，可以通过“边际拟合 (Margfit)”技术将误差矩阵归一化（或标准化）以便于不同误差矩阵之间的比较。通过对误差矩阵做归一化处理，可以消除误差矩阵生成过程中因样本数的不同而造成的差异，使不同误差矩阵中任意相对应的元素之间具有可比性。由于在归一化过程中考虑了非对角线元素的信息，因此，与误差矩阵中的总体精度相比，归一化的精度更能代表分类的真实精度 (Congalton et al., 1998)。

2. 基于误差矩阵精度评价方法的问题

(1) 精度测量指标。虽然从误差矩阵可以得到诸如总体精度、生产者精度、用户精度，以及 Kappa 系数等多个精度度量指标，并且已经成为遥感数据分类精度评价的核心方法 (Foody, 2001; Smits, 1999)。但在实际应用中，仍然存在许多问题 (Foody, 1992; Pontius, 2000)。

Foody (1992) 认为，由于在 Kappa 系数计算过程中实际高估了偶然一致性 (chance agree)，使总体分类精度被低估。Ma and Redmond (1995) 同样认识到这个问题，并建议用 Tau 系数代替 Kappa 系数作为误差矩阵的精度指标。有些科学家甚至认为，作为一个非基于概率的测量，Kappa 系数不适合作为一个精度测量指标 (Stehman and Czaplewski, 1998)。也有一些科学家认为，应该根据不同的目标使用不同的精度测量 (Lark, 1995; Stehman, 1997, 1999; Koukoulas and Blackburn, 2001)，且在精度评价报告中提供原始的误差矩阵和多个精度测量以全面描述分类精度 (Arbia et al., 1998; Muller et al., 1998; Stehman, 1997)。