

# WebGIS 中地理空间 Metadata 管理系统研究与设计

李琦 陈爱军 夏曙东  
(北京大学数字地球工作室, 北京 100871)

**摘要** 为实现 WebGIS 中地理空间信息的有效管理和网络共享, 结合国际地理空间信息共享的标准规范, 对 WebGIS 中地理空间 Metadata 的定义、作用、内容、技术实现及其应用进行了详细的研究和设计, 提出了与 WebGIS 集成的地理空间 Metadata 管理系统解决方案, 并给出了体现具体技术实现的原型系统, 为国家空间信息基础设施和数字地球的建设提供了理论基础和技术原型, 在当今信息化与数字地球时代, 对地理空间信息的网络发布和共享具有积极的推动作用和深远的社会意义。

**关键词** WebGIS 地理空间 Metadata 地理空间数据

中图法分类号: P208 文献标识码: A 文章编号: 1006-8961(2000)10-0811-07

## Research and Design of Geospatial Metadata Management System on WebGIS

LI Qi, CHEN Ai-jun, XIA Shu-dong  
(The CyberGIS Studio, Peking University, Beijing 100871)

**Abstract** The massive information and world-covered network makes the information management and sharing more urgent than ever, in order to attain invalid management and sharing of geospatial information on WebGIS, the detailed research and design on definition, role, content, implementation, application of geospatial metadata has been discussed in this paper. Combined with the frontier of international geospatial metadata standard, the geospatial metadata standard draft has been advanced here. Also taking advantages of the newest Internet technology and integrated with WebGIS, geospatial metadata management system, which had been running on Internet, has been designed and achieved. And some of demonstrations of various functions of this system have been presented finally. In the current times of information and Digital Earth, it will promote the sharing of geospatial information and propel the progress of Chinese standard, at the same time, this system provided the bases for constructing National Spatial Information Infrastructure (NSII) and Digital Earth.

**Keywords** WebGIS, Geospatial metadata, Geospatial data

## 0 引言

电子技术和通信技术的广泛使用, 使得全球范围内对数字地理信息和地理信息系统应用的需求也越来越大, 并且越来越多的地理科学和信息技术科

学之外的个人、组织、机构也涉入到这一领域, 开始生产、处理和修改数字地理信息, 使得人们对地理空间及其与事物间空间联系的重要性有了一个再认识的过程. 它们从各自的角度出发, 已形成了许多空间数据, 它们不仅满足各自的需要, 同时这些数据也可以满足其它应用需要, 但问题是, 目前我们不知道存

在什么样的数据、已有数据的质量如何、以及怎样访问和使用这些数据成果。因此,人们迫切需要采取一定的办法来避免数据的重复性建设,同时需协调不同数据部门之间的资源共享,这样,随着地理空间数据集的数量、复杂性和多样性的增加,一个适应数据集共享的标准化规范和管理系统也就应运而生,这就是 Metadata 网络管理系统,它也是 WebGIS 系统(即网络地理信息系统),必不可少的一部分,通过它可以了解 WebGIS 系统所提供地理空间数据的情况,如:有什么数据?数据质量如何?数据有那些格式?以什么方式在哪儿可以得到数据?等等和数据有关的信息。通过这些信息可以实现不同部门、不同专业领域的地理空间数据的网络共享,从而可避免因地理空间数据的重复收集、录入和处理而导致的大量时间、资金、人力和物力的浪费。

## 1 Metadata 的定义及其作用

对于空间元数据标准内容的研究,目前国际上主要有欧洲标准化委员会(CEN/TC 287)、美国联邦地理数据委员会(FGDC)和国际标准化组织地理信息/地球信息技术委员会(ISO/TC 211)三个组织在进行。

关于空间元数据的定义,欧洲标准化委员会 CEN/TC 287 认为空间元数据是“描述地理数据集内容、表示、空间参考、质量以及管理的数据”,而美国联邦地理数据委员会 FGDC 和国际标准化组织地理信息/地球信息委员会 ISO/TC 211 则认为:空间元数据是“关于数据的内容、质量、条件以及其它特征的数据”<sup>[1,2]</sup>。

总的来说,空间元数据是“关于数据的数据”,它在地理信息中用于描述地理数据集的内容、质量、表示方式、空间参考、管理方式以及数据集的其它特征,它是实现地理空间信息共享的核心标准之一。

地理空间元数据与数据字典的主要区别在于:元数据是对关于整个数据集本身及其内容的全面分层次规范化的描述,且任何数据集的元数据描述格式和内容都是相同的,因而可以用相同的管理系统对所有数据集的元数据进行管理和维护;而数据字典只是描述数据集中的部分内容,且没有统一的规范和标准,不同数据集生产者只是根据不同需求对数据集内容作出描述或说明,因此不可用相同的管理系统进行统一的管理和维护。

空间元数据主要有下列几个方面的作用<sup>[2-4]</sup>

(1) 用来组织和管理空间信息,并挖掘空间信息资源,这正是数字地球的特点和优点所在。通过它可以在 Intranet 或 Internet 上准确地识别、定位和访问空间信息。

(2) 帮助数据使用者查询所需空间信息。比如,它可以按照不同的地理区间、指定的语言以及具体的时间段来查找空间信息资源。

(3) 维护和延续一个机构对数据的投资。空间元数据可以确保一个机构对数据投资的安全。空间数据集建立后,随着机构中人员的变换以及时间的推移,后期接替该工作的人员会对先前的数据了解甚少或一无所知,这样便对先前数据的可靠性产生质疑,而通过空间元数据内容,则可以充分描述数据集的详细情况。同样,当用户使用数据引起矛盾时,数据提供单位也可以利用空间元数据维护其利益。

(4) 用来建立空间信息的数据目录和数据交换中心。数字地球中包含着各行各业的不同内容的信息,诸如农业、林业、交通、水利等许许多多专题信息,然而很少有一个组织可以提供所有方面的数据内容。因此,通常由一个组织产生的数据可能对其它组织也有用,而通过数据目录、数据代理机、数据交换中心等提供的空间元数据内容,用户便可以方便地得到符合自己要求的地理空间数据,并很容易地使用它们,从而它已成为实现地理空间信息的跨部门、跨行业和跨区域共享的有效解决途径之一。当然,要想真正实现空间信息的全球共享,还涉及到处理空间信息的软件系统、空间信息的软件模型和数据格式以及国家的政策法规等。

(5) 提供数据转换方面的信息。在未来的空间信息中,均应当包含空间元数据信息,以便使用户在获取包含空间信息数据集的同时就可以得到空间元数据信息。通过空间元数据,人们便可以接受并理解数据集,且可以与自己的空间信息集成在一起,进行不同方面的分析决策,使地理空间信息实现真正意义上的共享,发挥其最大的应用潜力。

目前对于地理空间元数据的应用需求主要集中在目录、历史记录、地理空间数据集内部以及可读性等4个方面。其具体的实施应用,关键是在获取和整理地理空间数据的同时,要严格按照地理空间元数据的标准规范,建立该数据集的元数据和建设相应的元数据管理系统,并向上一级数据交换中心提供该数据集的元数据标准规范(实现所需规范和技术

可参看本文其它部分内容).

## 2 Metadata 的分类

按照 Metadata 所描述的数据内容, Metadata 可分为数据集系列 Metadata、数据集 Metadata、要素类型和要素实例 Metadata、属性类型和属性实例 Metadata(如图 1)<sup>[5]</sup>.

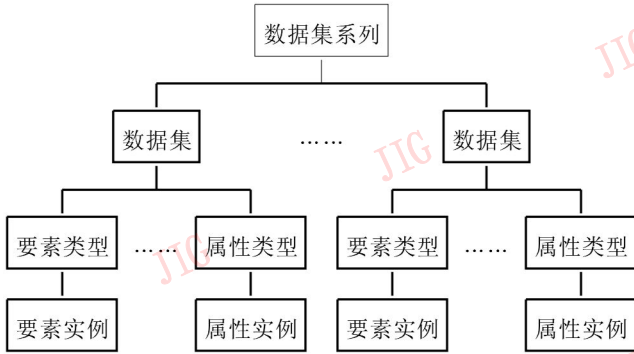


图 1 Metadata 类别结构图

### (1) 数据集系列 Metadata

数据集系列 Metadata 是指一系列拥有共同主题、日期、分辨率以及方法等特征的空间数据系列或集合, 它也是用户用于概括性查询数据集的主要内容. 通常, 数据集系列的定义由数据集生产者具体定义, 比如航空摄影时, 飞机在一条航带上用同一摄影机和相同参数拍摄的一系列航片, 或按照行政区划组成的国家资源环境数据库中的某一区域库等内容, 都是组成数据集系列数据的数据集.

在软件实现上, 如果拥有数据集系列 Metadata 模块, 则既可以使数据集生产者方便地描述宏观数据集, 而且也可以使用户很容易地查询到数据集的相关内容, 以实现空间信息资源的共享. 当然, 要获取数据集的详细信息, 还需通过数据集 Metadata 来实现.

### (2) 数据集 Metadata

数据集 Metadata 模块是整个 Metadata 标准软件的核心, 它既可以作为数据集系列 Metadata 的组成部分, 也可以作为后面数据集属性以及要素等内容的父代 Metadata 数据集系列. 在 Metadata 软件标准设计的初级阶段, 通过该模块便可以全面反映数据集的内容. 当然随着数据集的变化, 为避免重复记录 Metadata 元素内容以及保持 Metadata 元素的实时性, 它便可通过继承关系仅仅只需更新变化了的信息, 这时 Metadata 软件系统的层次性便显得异

常重要.

### (3) 要素类型和要素实例 Metadata

要素类型在数据集内容中相对比较容易理解, 它是指由一系列几何对象组成的具有相似特征的集合, 比如数据集中的道路层、植被层等, 便是具体的要素类型.

要素实例或具体要素是具体的要素实体, 它用于描述数据集中的典型要素, 而且通过它可以直接获取有关具体地理对象的信息. 该模块是 Metadata 体系中详细描述现实世界的重要组成部分, 也是未来数字地球中走向多级分辨率查询的依据, 比如武汉长江大桥便是一个具体要素. 因此, 我们通过数据集系列、数据集、要素类型等层次步骤, 便可以逐级对地理世界进行描述, 用户也可以按照这一步骤, 沿网络获取详细的数据集内容信息.

### (4) 属性类型和属性实例 Metadata

属性类型是用于描述空间要素某一相似特征的参数, 比如桥梁的跨度便是一个属性类型; 属性实例则是要素实例的属性, 比如某一桥梁穿越某一道路的跨度. 属性类型和属性实例是与要素类型和要素实例对应的模块, 它们是地理数据集软件层次结构或继承关系的组成部分, 也是 Metadata 软件系统的高级阶段内容.

## 3 Metadata 的内容

描述空间信息的空间元数据内容按照部分、复合元素和数据元素来组织(如图 2 所示).

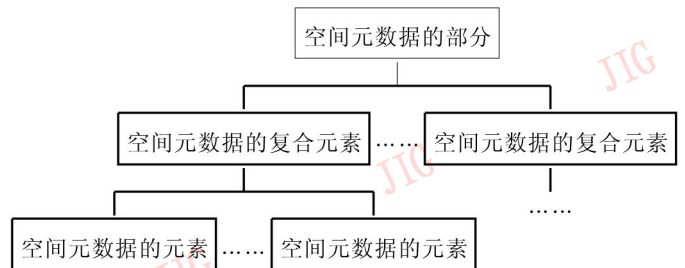


图 2 空间元数据内容组织示意图

空间元数据标准体系的内容具体分为 8 个基本内容部分和 4 个引用部分, 共由 12 个部分组成, 具体的标准化内容以及它们之间的相互关系如图 3 所示.

本文介绍的空间元数据标准由两层组成, 其中第一层是目录层, 它所提供的空间元数据复合元素和数据元素是数字地球中查询空间信息的目录信息, 它相对概括了第二层中的一些必选项信息, 是空

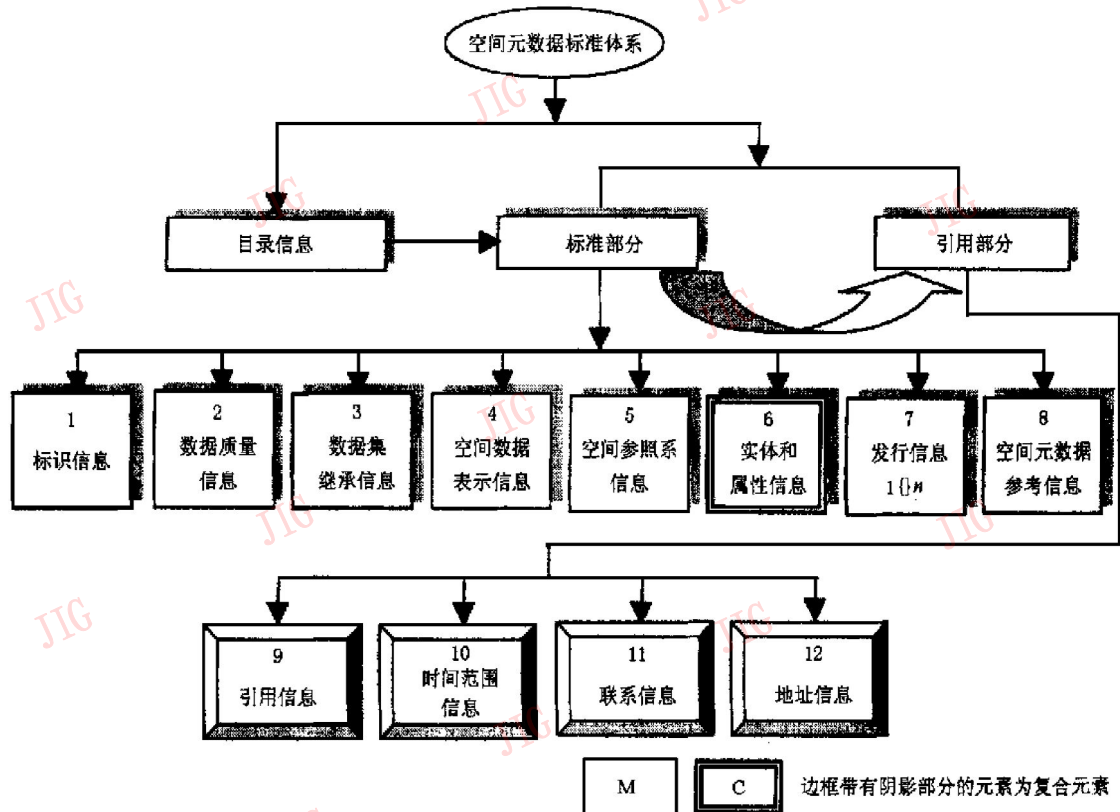


图3 空间元数据内容标准的组织框架

图中, M—该类型方框内的内容为必选项; C—该类型方框内的内容为条件可选项

元数据体系内容中比较宏观的信息; 第二层是空间元数据标准的主体, 它由 8 个标准部分(图 3(1~8)) 和 4 个引用部分(图 3(9~12)) 组成, 包括了全面描述地理空间信息的必选项、条件可选项、以及可选项的内容<sup>[1, 2, 6, 7]</sup>.

下面对元数据本身及其组成地理空间元数据的 12 个部分做较为详细的说明。

(1) 空间元数据 它是关于数据集内容、质量、表示方式、空间参考、管理方式以及数据集的其它特征的数据, 它位于整个标准体系的最上段, 属于复合元素(compound), 由两个层次组成。

在构成空间元数据标准内容的两个层次中, 第一层目录信息主要用于对数据集信息进行宏观描述, 它适合在数字地球的国家级空间信息交换中心或区域以及全球范围内管理和查询空间信息时使用。第二层则作为详细或全面描述地理空间信息的空间元数据标准内容, 是数据集生产者在提供数据集时必须提供的信息。

(2) 标识信息 是关于地理空间数据集的基本信息。通过标识信息, 数据集生产者可以对有关数据集的基本信息进行详细的描述, 诸如描述数据集的名称、作者信息、所采用的语言、数据集环境、专题分

类、访问限制等, 同时用户也可以根据这些内容对数据集有一个总体的了解。

(3) 数据质量信息 是对空间数据集质量进行总体评价的信息。通过这部分内容, 用户可以获得有关数据集的几何精度和属性精度等方面的信息, 也可以知道数据集在逻辑上是否一致, 以及它的完备性如何, 这是用户对数据集进行判断, 以及决定数据集是否满足他们需求的主要判断依据。数据集生产者也可以通过这部分内容对数据集的质量评价方法和过程进行详细的描述。

(4) 数据集继承信息 是建立该数据集时所涉及的有关事件、参数、数据源等的信息, 以及负责这些数据集的组织机构信息。通过这部分信息便可以对建立数据集的中间过程有一个详细的描述, 比如当一幅数字专题地图的建立经过了航片判读、清绘、扫描、数字地图编辑以及验收等过程时, 应对每一过程有一个简要描述, 使用户对数据集的建立过程比较了解, 也使数据集生成的每一过程的责任比较清楚。

(5) 空间数据表示信息 是数据集中表示空间信息的方式。它由空间表示类型、矢量空间表示信息、栅格空间表示信息、影像空间表示信息以及传感器波段信息等内容组成, 它是决定数据转换以及数

据能否在用户计算机平台上运行的必须信息. 利用空间数据表示信息, 用户便可以在获取该数据集后对它进行各种处理或分析了.

(6) 空间参照系信息 是有关数据集中坐标的参考框架以及编码方式的描述, 它是反映现实世界与地理数字世界之间关系的通道, 诸如地理标识码参照系统、水平坐标系统、垂直坐标系统以及大地模型等. 通过空间参照系中的各元素, 可以知道地理实体转换成数字对象的过程以及各相关的计算参数, 使数字信息成为可以度量和决策的依据. 当然, 它的逆过程也是成立的, 即可以由数字信息反映出现实世界的特征.

(7) 实体和属性信息 是关于数据集信息内容的信息, 包括实体类型、实体属性、属性值、域值等方面的信息. 通过该部分内容, 数据集生产者可以详细地描述数据集中各实体的名称、标识码以及含义等内容, 也可以使用户知道各地理要素属性码的名称、含义以及权威来源等.

在实体和属性信息中, 数据集生产者可以根据自己数据的特点, 在详细描述和概括描述之间选择其一, 以描述数据集的属性等特征.

(8) 发行信息 是关于数据集发行及其获取方法的信息, 包括发行部门、数据资源描述、发行部门责任、订购程序、用户订购过程以及使用数据集的技术要求等内容. 通过发行信息, 用户可以了解到数据集在何处, 怎样获取、获取介质以及获取费用等信息.

(9) 空间元数据参考信息 是有关空间元数据当前现状及其负责部门的信息, 包括空间元数据日期信息、联系地址、标准信息、限制条件、安全信息、以及空间元数据扩展信息等内容, 它是当前数据集进行空间元数据描述的依据. 通过该空间元数据描述, 用户便可以了解到所使用的描述方法的实时性等信息, 从而加深了对数据集内容的理解.

以下 4 部分内容作为地理空间元数据的引用部分, 自己不单独使用, 而是被图 3(1~ 8) 部分所引用. 这 4 部分内容在整个元数据标准规范中多次重复出现, 为了减少本标准规范的冗余度, 增强组成规范的内容的层次性和独立性, 所以对这 4 部分内容单独处理. 在具体实现某一数据集的元数据时, 该 4 部分内容会多次出现在图 3(1~ 8) 部分内容中.

(10) 引用信息 是引用或参考该数据集时所需的简要信息, 它自己从不单独使用, 而是被标准内容部分有关元素引用. 它主要由标题、作者信息、参

考时间、版本等信息组成.

(11) 时间范围信息 是关于有关事件的日期和时间的信息, 该部分是引用标准内容部分有关元素时要用到的信息, 它自己不单独使用.

(12) 联系信息 是与数据集有关的个人和组织联系时所需的信息, 包括联系人的姓名、性别、所属单位等信息. 该部分是引用标准内容部分有关元素时要用到的信息, 它自己不单独使用.

(13) 地址信息 是同组织或个人通讯的地址信息, 包括邮政地址、电子邮件地址、电话等信息. 该部分是描述有关地址元素的引用信息, 它自己不单独使用.

## 4 Metadata 网络管理系统设计与实现

WebGIS 中用于实现空间信息共享的空间元数据网络管理系统主要由权限验证功能(服务器端验证)、输入和合法性校验功能(客户端校验)、查询功能(服务器端查询)与返回和显示功能(服务器端格式化查询结果并返回, 客户端显示)等组成. 整个空间元数据网络管理系统逻辑图及其各功能模块间的关系如图 4 所示. 此系统可以作为数字地球中的子系统, 许多个这样的子系统的组合, 便可以基本上实现数字地球中的空间信息共享. 下面分别介绍各功能模块的功能<sup>[5, 8]</sup>.

### (1) 权限功能

提供用户登录界面, 当用户输入用户名和保密字后, 该信息会被传到服务器端, 并与保存在服务器端的合法用户名和保密字进行匹配, 只有二者完全匹配, 才是合法的用户, 才可以查询空间元数据内容. 同时, 系统还设置了高级用户, 高级用户具有录入和修改空间元数据内容的权限.

### (2) 输入功能

输入功能由输入结构、输入空间元数据、以及输入空间数据信息三部分组成. 其中, 输入结构是描述空间元数据在数据库中组织方式的信息, 它是客户端解码的依据; 输入空间元数据是空间元数据元素信息的核心部分, 通过该输入界面, 数据集生产者便可以充分地描述数据集内容; 输入空间数据信息是用于链接空间数据库的依据, 这部分内容是可选界面, 在数据库处于在线状态时, 具有非常重要的意义. 当然这些模块应当遵循前面所讲的层次关系, 使得在输入空间元数据元素时, 避免多余的重复.

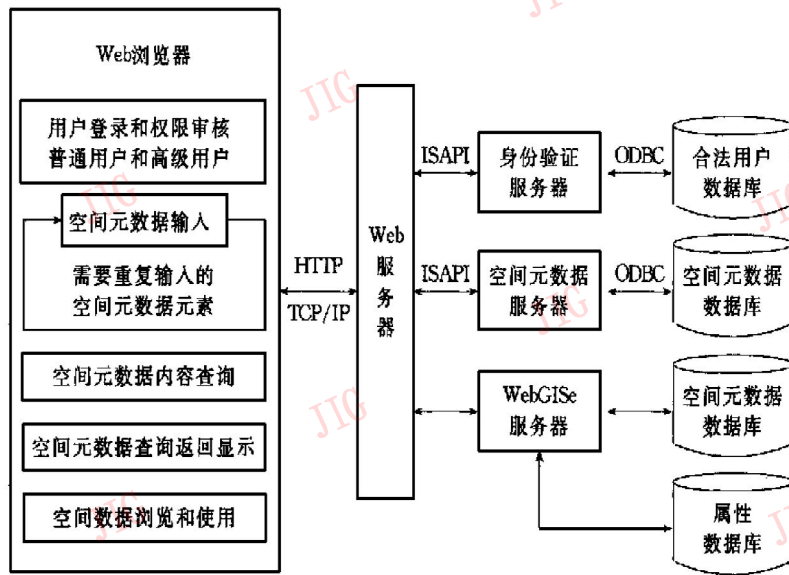


图4 空间元数据网络管理系统逻辑图

### (3) 查询功能

查询功能由初始化查询条件和返回结果两部分组成. 当用户在客户端提交查询条件和算法时, 服务器即将对该规则进行解译, 用来操作空间元数据数据库, 并将查询结果信息返回给客户端, 形成返回结果清单, 由此清单进一步获取用户所需的空间信息.

在软件设计时, 返回的清单项目至少应包括空间元数据标识码、中文标题, 以及数据集的范围信息(比如东西南北4个边界坐标点)等内容, 其中空间元数据标识码主要用于空间元数据内容之间的链接, 一般不用于显示.

### (4) 返回功能

返回功能由请求空间元数据、返回空间元数据、返回结构、请求空间数据、返回空间数据等内容组成.

当用户发出一个查询请求时, 它由客户端通过TCP/IP和HTTP协议传输给服务器, 这时服务器便对请求作出响应, 从空间元数据集合中查找合适的信息, 并对找到的信息进行HTML标准格式化, 然后返回给客户端, 返回结果中包含空间元数据标识码, 它用于进一步查询空间元数据的详细内容. 这样, 按照上述方式, 并利用空间元数据标准软件系统, 便可以通过网络实现地理空间信息的输入、查询、返回等操作, 从而实现数字地球中的空间信息共享. 以上所有操作都是基于TCP/IP协议和HTTP协议的.

实现后的Metadata管理系统基于Internet, 具有统一的用户界面, 操作简单, 使用方便, 具有权限的任何上网用户都可以访问该系统, 并可利用该系统进行地理空间信息的查询和获取.

## 5 Metadata 网络管理系统的应用

任何具有权限的用户都可以通过Internet访问该元数据管理系统, 可以查询Metadata, 并根据查询结果浏览地理空间数据集或获取所需要的数据集, 高级用户可以增加、修改、删除和维护Metadata内容, 由此就提供了一个实现WebGIS中地理空间信息共享的有效解决方案. 同时, 随着信息高速公路覆盖全球, 知识经济蓬勃发展, 网络进入千家万户, 数字地球正在迈向其日新月异的未来, 元数据标准体系在数字地球中的重要作用日益突出. 由具有空间信息共享功能的WebGIS站点作为数字地球中的一个信息节点, 由点连线, 由线及面, 覆盖全球, 从而真正实现信息化地球. 通过信息化地球上的高级元数据标准体系来对现实世界进行详细描述, 未来数字地球中还可以实现信息的多级分辨率查询和精确定位, 由此可见, Metadata网络管理系统的应用将为国家空间信息基础设施和数字地球的建设奠定坚实的理论基础和技术基础.

### 参考文献

- 1 ISO/Tc211 Metadata Working Document: Geographic Information, 1997, <http://www.fgdc.gov/>.
- 2 Metadata, Clearinghouse, Standards, NSDI, 1998, <http://www.fgdc.gov/>.
- 3 Harlan J Onsrud, Gerard Rushton Eds. Metadata Issues In Sharing Geographic Data. In: Sharing Geographic Information. New Brunswick: Rutgers, The State University of New Jersey, 1995: 499.
- 4 中国资源环境数据(1: 400万). 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, 1996.

- 5 赵永平. 基于国家空间信息基础设施的 Metadata 研究及其共享示范体系的建立[博士论文]. 北京: 北京大学, 1998, 7.
- 6 OpenGIS Topic11. OpenGIS Abstract Specification-Metadata (version 3), <http://www.opengis.org/technology/space.htm>.
- 7 赵永平, 承继成. 地理信息数据描述中元数据标准化的研究. 中国标准化, 1998, (3): 6~ 9.
- 8 李浩川. 基于网络环境下的地理空间数据组织、管理及其在可持续发展决策支持系统中的应用[硕士论文]. 北京: 北京大学, 1998, 7.



**李 琦** 教授、博士生导师, 中国图象图形学会技术委员会主任委员、全国地理信息标准化技术委员会委员、中国 GIS 协会资源与环境专业委员会副主任. 已在国际、国内刊物上发表论文 50 多篇. 研究领域为空间信息科学与技术、数字地球.



**陈爱军** 1972 年生, 北京大学数字地球工作室博士生. 已在国际国内刊物上发表 10 多篇论文. 研究领域为空间信息的网络发布与共享、空间信息科学、数字地球.



**夏曙东** 1972 年生, 北京大学数字地球工作室博士生. 已在国际国内刊物上发表数篇论文. 研究领域为空间信息科学、数字地球.

## 图象工程(下册)——图象理解与计算机视觉

章毓晋 编著

北京: 清华大学出版社, 2000

### 内 容 提 要

本书主要介绍图象理解和计算机视觉的基本概念、研究内容和实用技术, 以及近年来国际上有关的最新研究成果和应用实例.

本书包括五部分. 第一部分(第 1、2 章)概述了一些有关图象和视觉的基础知识; 第二部分(第 3、4 章)是对一些图象采集、处理和分析技术的推广; 第三部分(第 5、6 章)讨论有关立体匹配和场景恢复的技术; 第四部分(7、8、9 章)介绍有关对场景进行解释所涉及的知识表达、匹配理解等内容; 第五部分(附录 A 和附录 B)介绍了一些相关技术, 包括统计和结构模式识别、人工神经网络、数学形态学等, 以及有关的参考文献目录. 书中还提供了大量例题与习题.

本书可作为信息和信号处理、通信与电子系统、模式识别、机器人视觉、生物医学工程等学科专业课教材, 也可供上述学科及信息工程、电子工程、计算机科学与技术、机器人自动化、遥感和军事侦察等领域的科技工作者和高等院校的师生参考.

开本: 787× 1092 1/16 印张: 17.5 字数: 425 千字; 版次: 2000 年 8 月第 1 版 2000 年 8 月第 1 次印刷; 书号: ISBN 7-302-03980-1/TN·112; 定价: 20.00 元