

# 地理信息系统教程 复习

2020年1月12日 15:38

wangjw 整理

## 1、 GIS基本概念

数据(data)是人类在认识世界和改造世界过程中, 定性或定量对事物和环境描述的直接或间接原始记录, 是一种未经加工的原始资料,是客观对象的表示

信息( information)是用文字、数字、符号、语言、图像等介质来表示事件、事物现象等的内容、数量或特征,从而向人们(或系统)提供关于现实世界新的事实和知识, 作为生产、建设、经营、管理、分析和决策的依据。  
信息具有客观性、适用性、可传输性和共享性等特征。

地理信息

信息系统是具有采集、管理、分析和表达数据能力的系统。在计算机时代, 信息系统都部分地或全部地由计算机系统支持, 并由计算机硬件、软件、数据和用户四大要素组成。另外, 智能化的信息系统还包括知识。

地理信息系统( geographical information system 或geo-information system, 简称GIS), 有时又称为“地学信息系统”或“资源与环境信息系统”。它是一种特定的十分重要的空间信息系统, 是在计算机软硬件系统支持下, 对整个或部分地球表层(包括大气层)的有关地理分布数据进行采集、储存、管理运算、分析、显示和描述的技术系统。

GIS的基本内涵如下:

(1) GIS的物理外壳是计算机化的技术系统, (2) GIS 的操作对象是空间数据, 即点、线、面、体这类有三维要素的地理实体和地理现象。(3) GIS的技术优势在于它的数据综合、模拟与分析评价能力,可以得到常规方法或普通信息系统难以得到的重要信息, 实现地理空间过程演化的模拟和预测。(4) GIS与测绘学和地理学有着密切的关系。

## 2、 地理信息基本特征

具有信息的客观性、适用性、信息的可传输性和信息的共享性, 但从其本身而言, 地理信息还具有一些独特的特性, 包括:

(1)空间相关性:任何地理事物都是相关的, 并且在空间上相距越近则相关性越大, 空间距离越远则相关性越小, 同时地理信息的相关性具有区域性特点。

(2)空间区域性:区域性是地理信息的天然特性, 不仅体现在数据上的分区组织, 而且在应用方面也是面向区域的, 即一个部门或专题必然也是面向所管理或服务的区域的。

(3)空间多样性:在不同的地方或区域, 地理数据的变化趋势是不同的, 地理信息的多样性意味着地理信息的分析结果需要依赖其位置, 才能得出合乎逻辑的解释。地理信息的多样性也体现在不同区域对地理信息的需求不一样, 特别是对于地理信息服务, 信息的生产、信息的存储和信息的使用安排需要考虑不同地方对信息的需求。

(4) 空间层次性:地理信息的层次性首先体现在同一区域上的地理对象具有多重属性, 例如, 某区的土壤侵蚀研究, 相关因素包括该地区的降雨、植被覆盖、土壤类型等;其次是空间尺度上的层次性, 不同空间尺度数据具有不同的空

间信息特征。

### 3、 GIS基本功能

#### 数据采集功能

数据是GIS的血液，贯穿于GIS的各个过程。数据采集是GIS的第一步，即通过各种数据采集设备如数字化仪、全站仪等来获取现实世界的描述数据,并输入GIS(图1.2)。GIS 应该尽可能提供与各种数据采集设备的通信接口。

#### 数据处理和编辑

通过数据采集获取的数据称之为原始数据，原始数据不可避免地含有误差。

为保证数据在内容、逻辑、数值上的一致性和完整性，需要对数据进行编辑、格式转换、拼接等一系列的处理工作。也就是说,GIS应该提供强大的、交互式的编辑功能,包括图形编辑、数据变换、数据重构、拓扑建立、数据压缩、图形数据与属

#### 数据储存、组织和管理功能

计算机的数据必须按照一定的结构进行组织和管理,才能高效地再现真实 环境和进行各种分析。由于空间数据本身的特点，一般信息系统中的数据结构和数据库管理系统并不适合管理空间数据,GIS必须发展自己特有的数据存储、组织和管理的功能。目前常用的GIS数据结构主要有矢量数据结构和栅格数据 结构两种，而数据的组织和管理则有文件-关系数据库混合管理模拟模式、全关 系型数据管理模式、面向对象数据管理模式等。

#### 空间查询和空间分析功能

虽然数据库管理系统一般提供了数据库查询语言，如SQL语言,但对于GIS而言,需要对通用数据库的查询语言进行补充或重新设计,使之支持空间查询。

#### 数据输出功能

通过图形、表格和统计图表显示空间数据及分析结果是GIS项目的必需。

作为可视化工具,不论是强调空间数据的位置还是分布模式乃至分析结果的表 达，图形是传递空间数据信息最有效的工具。GIS脱胎于计算机制图,因而GIS 的一个主要功能就是计算机地图制图，包括地图符号的设计、配置与符号化、地 图注记、图幅整饰统计图表制作、图例与布局等项内容。此外对属性数据也要 设计报表输出，并且这些输出结果需要在显示器、打印机、绘图仪或数据文件输 出(图1.6)。GIS软件亦应具有驱动这些设备的能力。

### 4、 GIS组成

GIS运行环境包括计算机硬件系统、软件系统、网络、空间数据和管理与应用人员。其核心部分是计算机软硬件系统，空间数据反映了GIS 的地理内容,而管理人员和用户则决定系统的工作方式和信息表示方式

### 5、 GIS与CAD的关系

计算机辅助设计(CAD)是计算机技术用于机械、建筑、工程和产品设计的 系统，它主要用于范围广泛的各种产品和工程的图形，大至飞机小到微芯片等。

CAD主要用来代替或辅助工程师们进行各种设计工作，也可以与计算机辅助制 造(CAM)系统共同用于产品加工中作实时控制。GIS与CAD系统的共同特点是二者都有坐标参考系统，都能描述和处理图形数据及其空间关系，也都能处 理非图形属性数据。它们的主要区别是,CAD处理的多为规则几何图形及其 组合，图形功能极强,属性功能相对较弱。而GIS处理的多为地理空间的自然目标和人工目标，图形关系复杂，需要有丰富的符号库和属性库,GIS需要有

较强的空间分析功能,图形与属性的相互操作十分频繁,且多具有专业化的特征。此外,CAD一般仅在单幅图上操作,海量数据的图库管理的能力比GIS要弱。

## 6、 地理格网的含义,格网的划分体系

### 1.地理格网的含义

地理格网是指按一定的数学规则对地球表面进行划分而形成的格网。有人认为地理格网是地图分幅的代名词,并赋予“国土控制格网”的名称;有人认为地理格网是地球特定区域某种属性的统计单位,而常称之为“地理定位格网”但不论是“国土控制格网”或者是“地理定位格网”的说法,都曾造成某些误解使有的部门牵强附会地引用这个标准去表示现状分布的地理要素。为避免误解,标准最后定名为“地理格网”。

### 2.格网划分体系

地理格网可以按经纬度坐标系统划分(称之为地理坐标格网),也可以按直角坐标系统划分(称之为直角坐标格网)。两者各有其用处,也各具缺点。

地理坐标格网体系着眼于全球范围宏观研究的需要其优点是便于进行大区域乃至全球性的拼接,它不随投影系统的选择而改变格网的位置,但这种格网所对应的实地大小不均匀,高纬度地区较小,低纬度地区较大。我国领土所覆盖的面积较大,这种差别尤为明显。

直角坐标格网体系着眼于现实世界大量系统和数据生产单位实际采用直角坐标系的客观需求。具有实地格网大小均匀的优点,它在局部的小区域是可行的。但直角坐标格网所对应的实地位置随着选用地图投影的不同而改变。若采用高斯投影的6°带进行分割,则在分带的边缘会产生许多不完整的网格,无法进行全国性的整体拼接。然而这两种划分体系都可以互相转换

## 7、 空间数据模型的定义

空间数据模型是以计算机能够接受和处理的数据形式,为了反映空间实体的某些结构特性和行为功能,按一定的方案建立起来的数据逻辑组织方式,是对现实世界的抽象表达。分为概念模型、逻辑模型、物理模型。

## 8、 空间关系主要有哪些 - 空间拓扑关系

空间关系是指地理空间实体之间相互作用的关系。空间关系主要有:

(1)拓扑空间关系:用来描述实体间的相邻、连通、包含和相交等关系。

(2)顺序空间关系:用于描述实体在地理空间上的排列顺序,如实体之间前后、上下、左右和东、南、西、北等方位关系。

(3)度量空间关系:用于描述空间实体之间的距离远近等关系。

对空间关系的描述是多种多样的,有定量的,也有定性的;有精确的,也有模糊的。各种空间关系的描述也非绝对独立,而是具有一定联系。对空间关系的描述和表达,是GIS能够进行复杂空间分析的重要原因。

## 四

### 9、 什么叫空间数据结构, 实体数据结构、栅格数据结构

空间数据结构是指对空间逻辑数据模型描述的数据组织关系和编排方式, 对地理信息系统中数据存储、查询检索和应用分析等操作处理的效率有着至关重要的影响。

同一空间逻辑数据模型往往采用多种空间数据结构, 例如游程长度编码结构、四叉树结构都是栅格数据模型的具体实现。空间数据结构是地理信息系统沟通信息的桥梁, 只有充分理解地理信息系统所采用的特定数据结构, 才能正确有效地使用系统。在地理信息系统中, 较常用的有栅格数据结构和矢量数据结构, 除此之外还有混合数据结构、镶嵌数据结构和超图数据结构等。空间数据结构的选择取决于数据的类型、性质和使用的方式, 应根据不同的任务目标, 选择最有效和最合适的数据结构。

实体数据结构也称Spaghetti数据结构, 是指构成多边形边界的各个线段, 以多边形为单元进行组织。按照这种数据结构, 边界坐标数据和多边形单元实体一一对应, 各个多边形边界点都单独编码并记录坐标。例如对图4.1所示的多边形A、B、C、D, 可以采用两种结构分别组织。

以规则栅格阵列表示空间对象的数据结构称为栅格数据结构。阵列中每个栅格单元上的数值表示空间对象的属性特征。即栅格阵列中每个单元的行列号确定位置, 属性值表示空间对象的类型、等级等特征。每个栅格单元只能存在一个值。

### 10、 栅格单元确定的几种方法

#### 1. 栅格数据的参数

亦一个完整的栅格数据通常由以下几个参数决定:

(1) 栅格形状: 栅格单元通常为矩形或正方形。特殊的情况下也可以按经纬网划分栅格单元

(2) 栅格单元大小: 也就是栅格单元的尺寸, 即分辨率。栅格单元的合理尺寸应能有效地逼近空间对象的分布特征, 以保证空间数据的精度。但是用栅格来逼近空间实体, 不论采用多细小的栅格, 与原实体比都会有误差。通常以保证最小图斑不丢失为原则来确定合理的栅格尺寸。设研究区域某要素的最小图斑面积为 $s$ , 栅格单元的边长 $L$ 用如下公式计算:

$$L=0.5 * \sqrt{s}$$

(3) 栅格原点: 栅格系统的起始坐标应当和国家基本比例尺地形图公里网的交点相一致, 或者和已有的栅格系统数据相一致。并同时使用公里网的纵横坐标轴作为栅格系统的坐标轴。这样在使用栅格数据时, 就容易和矢量数据或已有的栅格数据配准。

(4) 栅格的倾角: 通常情况下, 栅格的坐标系统与国家坐标系统平行。但有时候, 根据应用的需要, 可以将栅格系统倾斜某一个角度, 以方便应用。

#### 2. 栅格单元值的选取

栅格单元取值是惟一的, 但由于受到栅格大小的限制, 栅格单元中可能会出现多个地物, 那么在决定栅格单元值时应尽量保持其真实性, 对于图4.7所示的栅格单元, 要确定该单元的属性取值, 可根据需要选用如下方法。

(1) 中心点法: 用位于栅格中心处的地物类型决定其取值。由于中心点位于代码为 $c$ 的地物范围内, 故其取值为 $C$ 。这种方法常用于有连续分布特性的地理现象



(2)面积占优法:以占矩形区域面积最大的地物类型作为栅格单元的代码。

从图上看,B类地物所占面积最大,故相应栅格单元代码为B。

(3)重要性法:根据栅格内不同地物的重要性,选取最重要的地物类型作为相应的栅格单元C代码。设图中A类地物为最重要的地物类型,则栅格代码为A。这种方法常用于有特殊意义而面积较小的地理要素,特别是点状和线状地理要素,如城镇、交通线、水系等。在栅格代码中应尽量表示这些重要地物。

(4)百分比法:根据矩形区域内各地理要素所占面积的百分比数确定栅格单元的取值,如可记面积最大的两类B和A,也可根据B类和A类所占面积百分比数在代码中加入数字。

由于采用的取值方法不同,得到的结果也不尽相同。

逼近原始精度的第二种方法是缩小单个栅格单元的面积,即增加栅格单元的总数,行列数也相应地增加。这样,每个栅格单元可代表更为精细的地面矩形单元,混合单元减少。混合类别和混合的面积都大减小,可以大大提高量算的精度;接近真实的形态,表现更细小的地物类型。然而增加栅格个数,提高数据精度的同时也带来了一个严重的问题,那就是数据量的大幅度增加,数据冗余严重。

## 11、什么是链码结构

链码数据结构首先采用弗里曼(Freeman)码对栅格中的线或多边形边界进行编码,然后再组织为链码结构的文件。链式编码将线状地物或区域边界表示为:由某一起始点和在某些基本方向上的单位矢量链组成。单位矢量的长度为一个栅格单元,每个后续点可能位于其前继点的8个基本方向之一(图4.15)。

图4.16所示的线实体和面实体可编码为表4.12所示的方式。具体编码过程是:起始点的寻找一般遵从从上到下、从左到右的原则。当发现没有记录过的点,而且数值也不为零时,就是一条线或边界线的起点。记下该地物的特征码及起点的行列数然后按顺时针方向寻迹,找到相邻的等值点,并按8个方向编码。

如遇不能闭合的线段结束后可以返回到起始点再开始寻找下一个线段。已经记录过的栅格单元,可将属性代码置零,以免重复编码。

## 五

## 12、什么叫空间数据

空间数据又称几何数据,它用来表示物体的位置、形态、大小分布等各方面的信息,是对现实世界中存在的具有定位意义的事物和现象的定量描述。根据在计算机系统中对地图是对现实世界的存储组织、处理方法的不同,以及空间数据本身的几何特征,空间数据又可分为图形数据和图像数据

地理空间数据是指以地球表面空间位置为参照的自然、社会和人文景观数据,可以是图形、图像、文字、表格和数字等(图1.12),由系统的建立者通过数字化仪扫描仪、键盘、磁带机或其他通信系统输入GIS,是系统程序作用的对象,是GIS所表达的现实世界经过模型抽象的实质性内容。不同用途的GIS其地理空间数据的种类、精度都是不同的,但基本上都包括3种互相联系的数据类型:

1.某个已知坐标系中的位置 即几何坐标、标识地理实体和地理现象在某个已知坐标系(如大地坐标系、直角坐标系、极坐标系、自定义坐标系)中的空间位置,可以是经纬度、平面直角坐标、极坐标,也可以是矩阵的行、列数等。

2.实体间的空间相关性 即拓扑关系,表示点、线、面实体之间的空间联系,如网络节点与网络线之间的枢纽关系,边界线与面实体间的构成关系,面实体与岛或内部点的包含关系等。空间拓扑关系对于地理空间数据的编码、录入

格式转换、存贮管理、查询检索和模型分析都有重要意义,是地理信息系统的特色之一。

3.与几何位置无关的属性 即常说的非几何属性或简称属性(attribute),是与地理实体和地理现象相联系的地理变量或地理意义。属性分为定性和定量两种,前者包括名称、类型、特性等,后者包括数量和等级。定性描述的属性如岩石类型、土壤种类、土地利用类型行政区划等;定量的属性如面积、长度、土地等级、人口数量、降雨量、河流长度、水土流失量等。非几何属性一般是经过抽象的概念,通过分类、命名、量算、统计得到。任何地理实体和地理现象至少有一个属性,而地理信息的分析、检索和表示主要是通过属性的操作运算实现的。因此,属性的分类系统、量算指标对系统的功能有较大的影响。

地理信息系统特殊的空间数据模型决定了地理信息系统特殊的空间数据结构和特殊的数据编码,也决定了地理信息系统具有特色的空间数据管理方法和系统的空间数据分析功能,成为地理学研究和资源管理的重要工具。

### 13、 空间数据库与一般数据库比较有什么特点

空间数据库与一般数据库相比,具有以下特点:

- ①数据量特别大。地理信息系统是一个复杂的综合体,要用数据来描述各种地理要素,尤其是要素的空间位置和空间关系等,其数据量往往很大
- ②不仅有地理要素的属性数据(与一般数据库中的数据性质相似),还有大量的空间数据,即描述地理要素空间分布位置的数据,并且这两种数据之间具有不可分割的联系。
- ③数据应用广泛。例如地理研究、环境保护、土地利用和规划、资源开发、生态环境、市政管理、道路建设等。

空间数据库的组成,从类型上分有栅格数据库和矢量数据库两类,其中栅格数据包括航空遥感影像数据和DEM数据;矢量数据则包括各种空间实体数据(图形和属性数据)。如图5.2所示。

### 14、 空间数据的基本特征

#### 1.空间特征

每个空间对象都具有空间坐标,即空间对象隐含了空间分布特征。这意味着在空间数据组织方面,要考虑它的空间分布特征。除了通用性数据库管理系统或者文件系统关键字的索引和辅关键字索引外,一般都需要建立空间索引。

#### 2.非结构化特征

在当前关系数据库管理系统中,数据记录中每条记录都是定长的(结构化),数据项不能再分,不允许嵌套记录。而空间数据不能满足这种定长(结构化)要求。若用一条记录表达一个空间对象,其数据项可能是变长的,例如一条弧段的坐标,其长度是不可限定的,可能是两对坐标,也可能是成百上千对坐标;另一方面,一个对象可能包含另外的一个或多个对象,例如一个多边形,可能含有多条弧段。若一条记录表示一条弧段,则该多边形的记录就可能嵌套多条弧段的记录,故它不满足关系数据模型的结构化要求,从而使得空间图形数据难以直接采用通用的关系数据管理系统。

#### 3.空间关系特征

空间数据除了空间坐标隐含了空间分布关系外,还通过拓扑数据结构表达了多种空间关系。这种拓扑数据结构一方面虽然方便了空间数据查询和空间分析,但另一方面也给空间数据的一致性和完整性维护增加了复杂度。特别是有些几何对象,没有直接记录空间坐标的信息,如拓扑的面状实体仅记录组成它的弧段标识,因而进行查找、显示和分析操作时都需要操纵和检索多个数据文件。

#### 4.多尺度与多态性

一不同观察比例尺具有不同的尺度和精度,同一地物在不同情况下也会有形态差异。例如,城市在空间上占据一定的范围,在较大比例尺中可以作为面状空间实体对象,而在较小比例尺中,则是作为点状空间对象来处理的。

## 5.分类编码特征

一般情况下，每个空间对象都有一个分类编码。这种分类编码往往是按照国家标准，或者行业标准、地区标准来应用的，每一种地物类型在某个GIS中的属性项个数是相同的。因而在许多情况下，一种地物类型对应一个属性数据表文件。

。当然，如果几种地物类型的属性项相同，也可以多种地物类型共用一个属性数据表文件。

## 6.海量数据特征

GIS中数据量非常庞大，远大于一般的通用数据库。可称之为海量数据。一个城市地理信息系统数据量可达几十GB。如果考虑影像数据的存储，可能达到几百个GB。这样的数据量在城市管理的其他数据库中是很少见的。由此，需要在二维空间上划分块或图幅，在垂直方向上划分层来进行组织。

## 15、 矢量数据的三种数据管理方法

对于矢量数据，其位置数据和属性数据通常是分开组织的。这一特点使得在管理时需要同时顾及空间位置数据和属性数据，其中属性数据很适合用关系数据库来管理，空间位置数据则不太适合用关系数据库管理。空间数据管理方式与数据库发展是密不可分的，按照发展的过程，对矢量数据的管理有文件-关系数据库混合管理、全关系管理、对象-关系数据库管理等方式。

## 16、 什么叫空间索引，常用的空间索引方法

经对研究区空间数据输入并建立空间数据库以后，得到了一个庞大的数据库，如何从该数据库中快速检索提取所需的空間数据来满足空间分析、模拟与决策的需要是一个重要的问题。空间索引就是指依据空间对象的位置和形状或空间对象之间的某种空间关系按一定的顺序排列的一种数据结构，其中包含空间对象的概要信息，如对象的标识、外接矩形及指向空间对象实体的指针。作为辅助性的空间数据结构，空间索引介于空间操作算法和空间对象之间，它通过筛选作用，大量与特定空间操作无关的空间对象被排除，从而提高空间操作的速度和效率。空间索引的性能的优劣直接影响空间数据库和地理信息系统的整体性能，它是空间数据库和地理信息系统的一项关键技术。常见的空间索引一般是自顶向下、逐级划分空间的各种数据结构，比较有代表性的包括BSP树、K-D-B树、R树、R+树和CELL树等。此外结构较为简单的格网空间索引有着广泛的应用。

对象范围索引

格网索引

四叉树空间索引

R树、R+树索引

## 六

## 17、 数字化绘图出现的问题与解决办法

## 18、 属性数据编辑的内容

属性数据校核包括两部分：

①属性数据与空间数据是否正确关联，标识码是否惟一，不含空值；

②属性数据是否准确，属性数据的值是否超过其取值范围等。

对属性数据进行校核很难，因为不准确性可能归结于许多因素，如观察错误、数据过时和数据输入错误等。属性数

据错误检查可通过以下方法完成:

①首先可以利用逻辑检查,检查属性数据的值是否超过其取值范围,属性数据之间或属性数据与地理实体之间是否有荒谬的组合。在许多数字化软件中,这种检查通常使用程序来自动完成。例如,有些软件可以自动进行多边形节点的自动平差,属性编码的自动查错等。

②把属性数据打印出来进行人工校对,这和用校核图来检查空间数据准确性相似。

对属性数据的输入与编辑,一般在属性数据处理模块中进行。但为了建立属性描述数据与几何图形的联系,通常需要在图形编辑系统中设计属性数据的编辑功能,主要是将一个实体的属性数据连接到相应的几何目标上,亦可在数字化及建立图形拓扑关系的同时或之后,对照一个几何目标直接输入属性数据。一个功能强的图形编辑系统可提供删除、修改、拷贝属性等功能。

## 19、属性数据的内容

属性数据即空间实体的特征数据,一般包括名称、等级、数量、代码等多种形式。属性数据的内容有时直接记录在栅格或矢量数据文件中,有时则单独输入数据库存储为属性文件,通过关键码与图形数据相联系。

属性数据一般采用键盘输入。输入的方式有两种:一种是对照图形直接输入;另一种是预先建立属性表输入属性,或从其他统计数据库中导入属性,然后根据关键字与图形数据自动连接。

## 20、空间数据质量评价的重点

误差

准确度

偏差

精密度

不确定性

## 21、空间数据质量的控制-常用的几种方法

空间数据质量控制常见的方法有以下几种:

(1)传统的手工方法:质量控制的手工方法主要是将数字化数据与数据源进行比较,图形部分的检查包括目视方法、绘制到透明图上与原图叠加比较,属性部分的检查采用与原属性逐个对比或其他比较方法。

(2)元数据方法:数据集的元数据中包含了大量的有关数据质量的信息,通过它可以检查数据质量,同时元数据也记录了数据处理过程中质量的变化,通过跟踪元数据可以了解数据质量的状况和变化。

(3)地理相关法:用空间数据的地理特征要素自身的相关性来分析数据的质量。例如,从地表自然特征的空间分布着手分析,山区河流应位于微地形的最低点,因此,叠加河流和等高线两层数据时,若河流的位置不在等高线的汇水线上且不垂直相交,则说明两层数据中必有层数据有质量问题,如不能确定哪层数据有问题时,可以通过将它们分别与其他质量可靠的数据层叠加来进一步分析。因此,可以建立一个有关地理特征要素相关关系的知识库,以备各空间数据层之间地理特征要素的相关分析之用。

## 22、数据生产过程中质量控制

2.空间数据生产过程中的质量控制 数据质量控制应体现在数据生产和处理的各个环节。下面仍以地图数字化生成空间数据过程为例,介绍数据质量控制的措施。

(1)数据源的选择:由于数据处理和使用过程的每个步骤都会保留甚至加大原有误差,同时可能引入新的数据误差,



因此,数据源的误差范围至少不能大于系统对数据误差的要求范围。

所以对于大比例尺地图的数字化,原图应尽量采用最新的二底图,即使用变形较小的薄膜片基制作的分版图,以保证资料的现势性和减少材料变形对数据质量的影响。

(2)数字化过程的数据质量控制:主要从数据预处理、数字化设备的选用、对点精度、数字化限差和数据精度检查等环节出发。

1)数据预处理:主要包括对原始地图、表格等的整理、脊清或清绘。对于质量不高的数据源,如散乱的文档和图面不清晰的地图,通过预处理工作不但可减少数字化误差,还可提高数字化工作的效率。对于扫描数字化的原始图形或图像,还可采用分版扫描的方法来减小矢量化误差。

2)数字化设备的选用:主要按手扶数字化仪、扫描仪等设备的分辨率和精度等有关参数进行挑选。这些参数应不低于设计的数据精度要求。般要求数字化仪的分辨率达到0.025 mm,精度达到0.2 mm;对扫描仪的分辨率则不低于300DPI(dots per inch)。

3)数字化对点精度(准确性):数字化对点精度是指数字化时数据采集点与原始点重合的程度。般要求数字化对点误差小于0.1 mm。

4)数字化限差:数字化时各种最大限差规定为,曲线采点密度2 mm、图幅接边误差0.2 mm、线画接合距离0.2 mm、线画悬挂距离0.7 mm。对于接边误差的控制,通常当相邻图幅对应要素间距离小于0.3 mm时,可移动其中一个要素以使两者接合;当这一距离在0.3 mm与0.6 mm之间时,两要素各自移动一半距离;若距离大于0.6 mm,则按一般制图原则接边,并作记录。

5)数据的精度检查:主要检查输出图与原始图之间的点位误差。般要求,对直线地物和独立地物,这一误差应小于0.2 mm;对曲线地物和水系,这一误差应小于0.3 mm;对边界模糊的要素应小于0.5 mm。

## 23、 数据入库流程

入库流程一般在数据库建库设计阶段就基本确定,不同数据源,不同的空间数据库库体,它们在具体的人库过程中,需要完成的工作各不相同,但通常包括图6.43所示的主要工作。首先,对待入库数据进行全面质量检查,包括资料完整性检查、数据完整性检查数据正确性检查,并完成检查报告。如果质量不合格,则将数据返回生产单位进行修改,修改后重新进行质量检查直至满足入库要求方可进入下一步。其次,对检查合格的数据进行整理,包括以下工作:按数据组织规则建立数据文件存储目录;按数据命名规则对成果数据统一命名;文件资料数字化;根据入库内容对数据字典及元数据进行相应更新;将成果数据存入指定目录。最后,将数据入库,完成全部入库工作。

## 24、 什么叫元数据,元数据的作用

### 1.元数据与元数据的作用

元数据的英文名称是“metadata”,它是关于数据的数据。在地理空间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他特征,它是实现地理空间信息共享的核心标准之一。

元数据的主要作用是:

(1)帮助用户了解和分析数据:元数据提供丰富的引导信息,以及由纯数据得到的分析、综述和索引等。根据元数据提供的信息,用户可对空间数据库进行浏览、检索、研究、分析,了解数据的基本情况和数据的可用性,获取方法等内容。

(2)空间数据质量控制:不论是统计数据还是空间数据都存在数据精度问题,影响空间数据精度的原因主要有两个方面:一是源数据的精度,是数据加工处理工程中精度质量的控制情况。空间数据质量控制内容包括:①有准确定义的数据字典,以说明数据的组成,各部分的名称、表征的内容等;②保证数据逻辑科学地集成,如植被数据库中不同亚类的区域组合成大类区。这要求数据按定逻辑关系有效的组合;③有足够的说明数据来源、数据的加工处理工程、数据释译的信息。通过按一定的组织结构集成到数据库中构成数据库的元数据信息系统可以实现上述功能。

(3)在数据集成中的应用:元数据记录了数据格式、空间坐标体系数据的 表达形式、数据类型数据使用软硬件环境数据使用规范数据标准等信息。这些 信息在数据集成的一系列处理中,如数据空间匹配、属性致化处理 数据交 换等方面是必需的。

(4)数据存贮和功能实现:元数据系统用于数据库的管理可以实现数据库设计和系统资源利用方面开支的合理分配,避免数据的重复存贮,并可以高效查 询检索分布式数据库中任何物理存贮的数据,减少用户查询数据库及获取数据的时间。

七

## 25、 缓冲区概念

缓冲区是根据数据库中的点、线、面地理实体,自动建立其周围一定宽度范 围的多边形,来表征特定地理实体对邻域的影响范围。缓冲区查询是在不破坏 原有空间目标的关系的情况下,只检索缓冲区范围内涉及的空间目标。

根据用户给定的点缓冲、线缓冲或面缓冲的距离,形成一个缓冲区的多边 形,再根据多边形检索的原理,从该缓冲区内检索出所要的空间对象。

## 26、 GIS空间分析概念

空间分析是从空间数据中获取有关地理对象的空间位置、分布、形态、形 成和演变等信息的分析技术,是地理信息系统的核心功能之一,它特有的对 地理信息的提取、表达和传输的功能,是地理信息系统区别于一般管理信息 系统的主要功能特征。

在空间分析的研究和实践中,很多在应用领域具有普 遍意义的、涉及空间位置的分析手段和方法被总结、提炼出来,形成了在GIS 软件中均包含的一些固有的空间分析功能模块。这些功能具有一定的通用 性质,故而称之为GIS基本空 间分析,具体的有叠置分析、缓冲区分析、窗口 分析和网络分析。了解GIS基本空间分析对于进一步掌握复杂空间分 析方 法,具有一定的指导意义。

八

## 27、 什么是叠置分析

叠置分析是地理信息系统中常用的提取空间隐含信息的方法之一,叠置分 析是 将有关主题层组成的各个数据层面进 行叠置产生的一个新数据层面,其结 果综合了原来两个或多个层面要素所具有的属性,同时叠置分析不仅生成了新的空间关系,而且还将输入的多个数据层的属性联系起来产生新的属性关系。

其中,被叠加的要素层面必须是基于相同坐标系统的、基准面相同的、同一区域 的数据。

按照GIS中最常用的两种数据结构将叠置分析分成矢量数据叠置分析和栅 格数据叠置分析,如图8.1所示。

根据操作形式的不同,叠置分析可以分为图层擦除、交集操作、图层合并 等;根据操作要素的不同,可以将矢量数据 叠置分析分成点与多边形叠加、线 与多边形叠加、多边形与多边形叠加,栅格数据叠置分析分为单层与多层栅格 数据叠置分析。要注意的是这里也要对属性进行一定的操作、所指的属性是 较为简单的属性值,但注解属性、尺 度属性、网络属性等均不能作为输入的属 性值。

## 28、 缓冲区分析的内容

缓冲区分析是地理信息系统中常用的一种空间分析 方法,是对空间特征进 行度量的一种重要手段。缓冲区分析是 研究根据数据库的点、线、面实体,自动 建立其周围一定宽度范围内的缓冲区 多边形实体,从而实现空间数据在水

平方 向得以扩展的信息分析方法。它是地理信息系统基本的空间操作功能之一。从 空间变换的观点看、缓冲区分析模型就是将点、线、面地物分布图变换成这些地 物的扩展距离图,图上每一点的值代表该点距离最近的某种地物的距离。实际 上,缓冲区就是地理目标或工程规划目标的一种影响范围。

缓冲区是地理空间目标的一种影响范围或服务范围在尺度上的表现。它是一种因变量,随所研究的要素的形态而改变。从数学的角度来看,缓冲区是给定 空间对象或集合后获得的它们的邻域,而邻域的大小由邻域的半径或缓冲区建 立条件来决定。因此对于一个给定的对象A,它的缓冲区可以定义为  $P = \{x | d(x,A) \leq r\}$  式中:d一般是指欧氏距离,也可以是其他的距离;r为邻域半径或缓冲区建立的 条件。

缓冲区分析包括缓冲区的建立及区域分析,它首先是对要素根据缓冲的条 件,建立缓冲区,然后将这个缓中区图层与其他图层进行诸如叠置分析、网络分 析服务设施查找等其他分析操作.得到所需要的结果,以便为某项分析或决 策 提供依据。缓冲区分析也称缓冲区操作,但有时也将缓冲区的建立称为缓冲区 操作,而建立后采用的分析过程称之为缓中区分析(图8、10)。

## 29、 网络分析的概念

在现实世界中,若干线状要素相互连接成网状结构,资源沿着这个线性网流 动,这样就构成了一个网络。在GIS中,作为空间实体的网络与图论中的网络不 同。它作为一种复杂的地理目标,除具有一般网络的边、节点间的抽象的拓 扑含 义之外,还具有空间定位上的地理意义和目标复合上的层次意义。具体说来,网 络就是指现实世界中,由链和节点组成的、带有环路,并伴随着一系列支配网络 中流动之约束条件的线网图形,它的基础数据是点与线组成的网络数据。

网络分析是通过模拟、分析网络的状态以及资源在网络上的流动和分配等, 研究网络结构、流动效率及网络资源等的优化问题的一种方法。对地理网络、城 市基础设施网络进行地理分析和模型化,是地理信息系统中网络分析功能的主 要目的。进行网络分析研究的数学分支是图论和运筹学,它的根本目的是研究、 筹划如何安排项基于网络数据的工程,并使其运行效果最好,如一定资源的最 佳分配,从一地到另一地的花费时间最短等,研究内容主要包括选择最佳路径, 选择最佳布局中心的位置,资源分配,节点弧段的遍历等。其基本思想则在于人 类活动总是趋向于按一定目标选择达到最佳效果的空间位置。这类问题在生 产、社会、经济活动中不胜枚举,因此研究此类问题具有重大意义。

## 九

## 30、 什么叫数字高程模型

数字高程模型(digital elevation model,简称DEM)是通过有限的地形高程数 据实现对地形曲面的数字化模拟(即地形表面形态的数字化表示),高程数据常 常采用绝对高程(即从大地水准面起算的高度)。DEM也常常称为DTM(digital terrain model)。“terrain”-词的含义比较广泛,不同专业背景对“terrain” 的理解 也不一样,因此,DTM趋向于表达比DEM更为广泛的内容。

数学意义上的数字高程模型是定义在二维空间上的连续函数 $H = f(x,y)$ 。

由于连续函数的无限性,DEM通常是将有限的采样点用某种规则连接成一系列 的曲面或平面片来逼近原始曲面,因此,DEM的数学定义为区域D的采样点或 内插点P按某种规则连接成的面片M的集合

$DEM = \{M = \{P | P(x_i, y_i, H) \in D\} \quad j=1, \dots, n, i=1, \dots, m\}$

## 31、 DEM建立的一般步骤

数字高程模型的建立过程是一个模型建立过程。从模型论角度讲,就是将 源域(地形)表现在另一个域(目标域或DEM)中的一种结构。建模的目的是对 复杂的客体进行简化和抽象,并把对客体(源域,DEM中为地形起伏)的研究

转 移到对模型的研究上来。

模型建立之初, 首先要为模型构造一个合适的空间结构( spatial framework)。空间结构是为把特定区域内的空间目标镶嵌在一起而对区域进行的划分, 划分出的各个空间范围称为位置区域或空间域。空间结构一般是规则的 (如格网), 或不规则的(如不规则三角网TIN)。

建立在空间结构基础上的模型是由n个空间域的有限集合组成。由于空间数据包含位置特征和属性特征, 而属性特征是定义在位置特征上的,因此每一个空间域就是由空间结构到属性域的计算函数或域函数。模型的可计算性要求有两点:一是空间域的数量属性域和空间结构是有限的:二是域函数是可计算的。

构筑模型的一般内容和过程为:

- ①采用合适的空间模型构造空间结构;
- ②采用合适的属性域函数;
- ③在空间结构中进行采样, 构造空间域函数;
- ④利用空间域函数进行分析。

当空间结构为欧几里得平面, 属性域是实数集合时, 模型为一自然表面。将 欧几里得平面充当水平的XY平面,属性域给出Z坐标(或高程) ,模型即为数字 高程模型。

对于数字高程模型而言, 空间结构的构造过程即为DEM的格网化过程(形成格网), 属性值为高程, 构造空间域函数即为内插函数的确定, 利用空间域函数进行分析就是求取格网点的函数值。

十一

### 32、 地理信息可视化概念

可视化的基本含义是将科学计算中产生的大量非直观的、抽象的或者不可见的的数据,借助计算机图形学和图像处理等技术, 以图形图像信息的形式,直观、形象地表达出来, 并进行交互处理。地图是空间信息可视化的最主要和最常用的形式。

在地理信息系统中,可视化则以地理信息科学、计算机科学、地图学、认知科学、信息传输学与地理信息系统为基础, 并通过计算机技术、数字技术、多媒体技术,动态、直观、形象地表现、解释、传输地理空间信息并揭示其规律,是关于 信息表达和传输的理论、方法与技术的一门学科。

地理信息系统中的空间信息可视化从表现内容上来分, 有地图(图形)、多媒体、虚拟现实等;从空间维数上来分有二维可视化、三维可视化、多维动态可视化等。本章侧重介绍空间信息可视化的基本形式和技术。

### 33、 什么是地理信息系统输出产品, 主要包括哪些类型

地理信息系统产品是指由系统处理、分析, 可以直接供研究、规划和决策人员使用的产品,

其形式有地图、图像统计图表以及各种格式的数字产品等。地理信息系统产品是系统中数据的表现形式, 反映了地理实体的空间特征和属性特征。

地图 图像 统计图表

### 34、 专题地图有哪些特点

专题地图具有下列3个特点:

- ①专题地图只将一种或几种与主题相关联的要素特别完备而详细地显示, 而其他要素的显示则较为概略, 甚至不予显示;
- ②专题 地图的内容广泛, 主题多样,在自然界与人类社会中,除了那些在地表上能见到的 和能直接进行测量的自然现



象或人文现象外,还有那些往往不能见到的或不能直接测量的自然现象或人文现象,均可以作为专题地图的内容;  
③专题地图不仅可以表示现象的现状及其分布,还能表示现象的动态变化和发展规律。

## 十二

### 35、什么是地理信息的网络服务

所谓地理信息的网络服务,就是指通过网络环境所提供的地理信息服务。

随着网络技术的发展,地理信息的网络服务已逐步拥有了改变GIS开发、访问和使用方式的潜力,并且借助普及的网络极大地拓展了GIS的应用面。另一方面,地理信息的网络服务很大程度上依赖网络技术的发展。

### 36、地理信息的网络服务内容

地理数据分发服务

制图服务

查询分析与辅助决策服务

基于位置的服务

### 37、什么叫基于位置服务,原理是什么

近年来兴起的基于位置的服务(Location-based service)成为地理信息服务的又一个新的领域。

LBS指的是在移动计算环境下,利用GIS技术、空间定位技术和网络通信技术,为移动对象提供基于空间地理位置的信息服务。

LBS的基本原理是:当移动用户需要信息服务或监控管理中心需要对某移动终端进行移动计算(跟踪监控、导航定位、搜索查询、实时调度等)时,首先移动终端通过内嵌的定位设备如GPS(全球定位系统)获得终端本身当前的空间位置数据,并实时地通过无线通信把数据(包括位置坐标、用户需求等)上传到网上的服务管理中心,中心GIS服务器根据终端的地理位置服务项目或运算要求进行空间分析,分析结果再下载到移动终端或监控中心的计算机上,并在屏幕上可视化地表示出来。

LBS服务目前主要分为查询、监控和导航3类。查询服务相对简单,例如,某人想找一家距离自己所在位置最近而又经济实惠的餐馆,他只要在随身携带的移动终端上输入最短距离及愿意承受的就餐价位,LBS即可根据用户的当前位置、就餐条件找出符合条件的餐馆并在终端上显示出行走路线简图,还可询问用户是否需要更高级的服务(如当出现多目标选择时,为用户提供智能决策等)。在LBS技术的支持下,以及现在GPS的普及化,监控和导航也成为非常平常的事情,甚至许多科学家将其用以检测野生动物。导航服务也不仅限服务于PDA等智能终端,随着手机定位技术的发展,“用手机拨打地图”的服务也已经可以满足相当精度的要求。