



信阳师范学院
Xinyang Normal University

地理信息系统概论

- 河南省一流本科建设课程
- 河南省课程思政样板课程
- 信阳师范学院课程思政建设课程

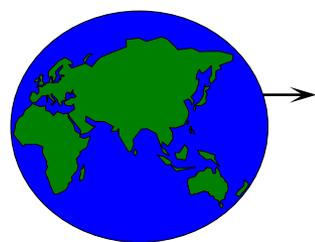
地理科学学院

牛继强

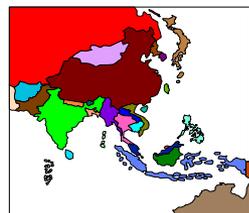


第五章：空间数据组织与管理

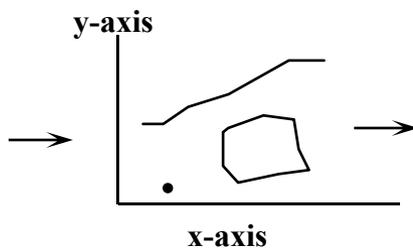
空间数据转化过程



现实世界



空间数据

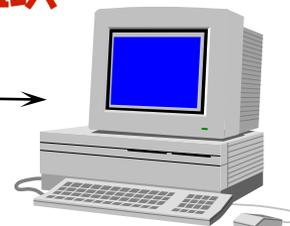


数据模型



数据结构

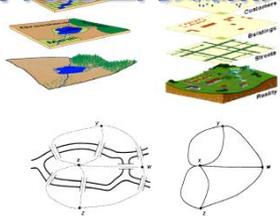
数据组织



空间数据库

空间数据组织 PK 空间数据管理

不同的空间数据模型



不同的空间数据形式



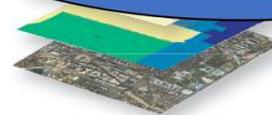
不同的空间数据管理
模式

关系数据库
对象-关系数据库
面向对象数据库



空间数据管理解决如何
“存”空间数据

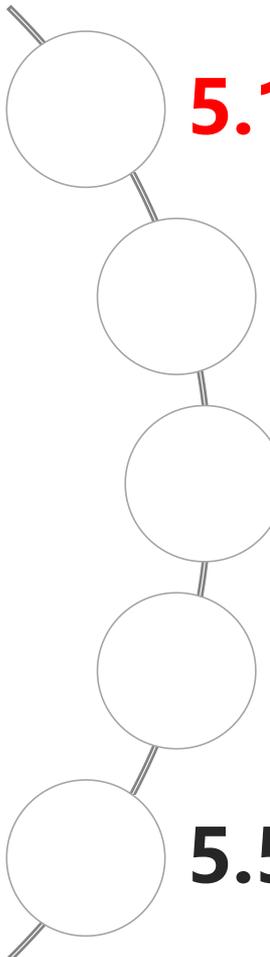
空间数据组织解决如何
有意义地“取”空间数据



R-ID	Owner	PurchDate	Pop.	Parcel#
10	033	09201974	10	34576
20	104	05131971	7	34599
30	006	09241060	2450	34712
40	999	99999999	0	34622
50	004	06061941	15	34227

其它组
织形式



- 
- 5.1 空间数据库概述**
 - 5.2 空间数据库设计**
 - 5.3 空间数据特征与组织**
 - 5.4 空间数据管理**
 - 5.5 空间数据检索**

5.1 空间数据库概述

当前大纲

5.1.1 数据库基础

5.1.2 空间数据库

5.1 空间数据库概述

5.1.1 数据库基础

□ 数据和数据库

数据 (data) 是描述事物的符号记录，可以是数字形式，也可以是文字、图形、图像、声音、语言等多种表现形式。人们收集并抽取出应用所需的大量数据后，将其保存起来以供进一步加工处理，抽取有用信息。

数据库 (database) 是长期储存在计算机内的、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。

5.1 空间数据库概述

5.1.1 数据库基础

□ 常用的数据模型

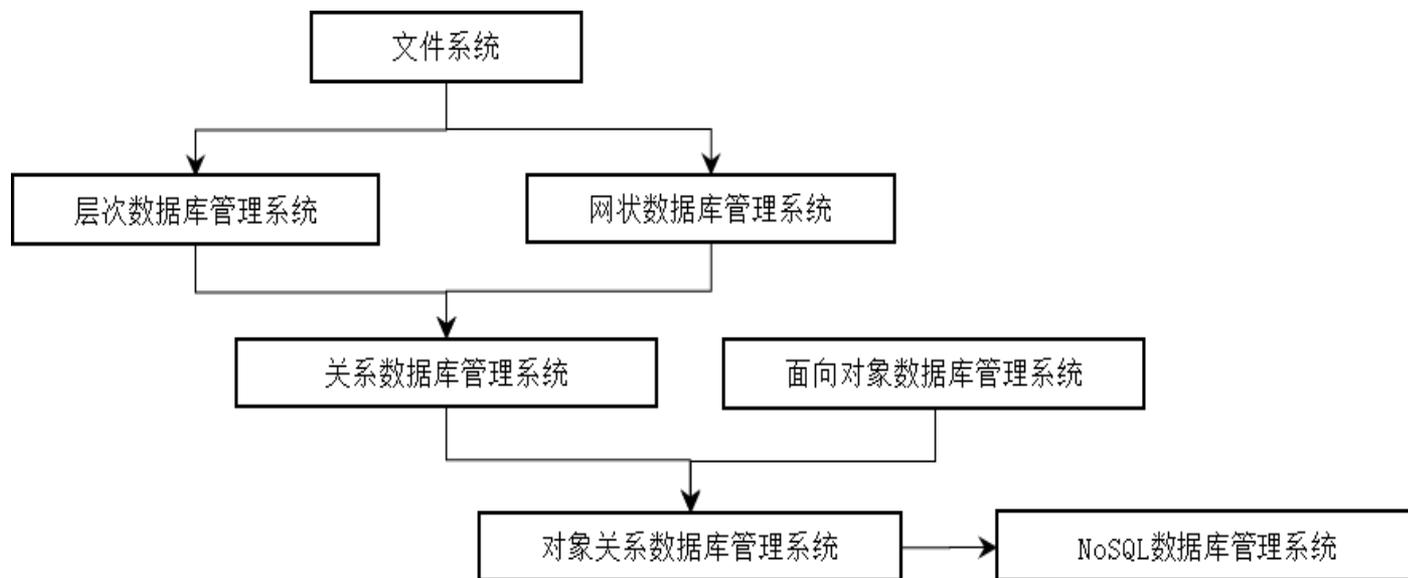
- 层次模型(hierarchical model)
- 网状模型(network model)
- 关系模型(relational model)
- 面向对象模型(object oriented model)。

其中层次模型和网状模型统称为非关系模型。非关系模型的数据库系统在20世纪70年代至80年代非常流行，在数据库系统产品中占据了主导地位，现在已逐渐被关系模型的数据库系统取代。

5.1 空间数据库概述

5.1.1 数据库基础

□ 数据库模型发展过程

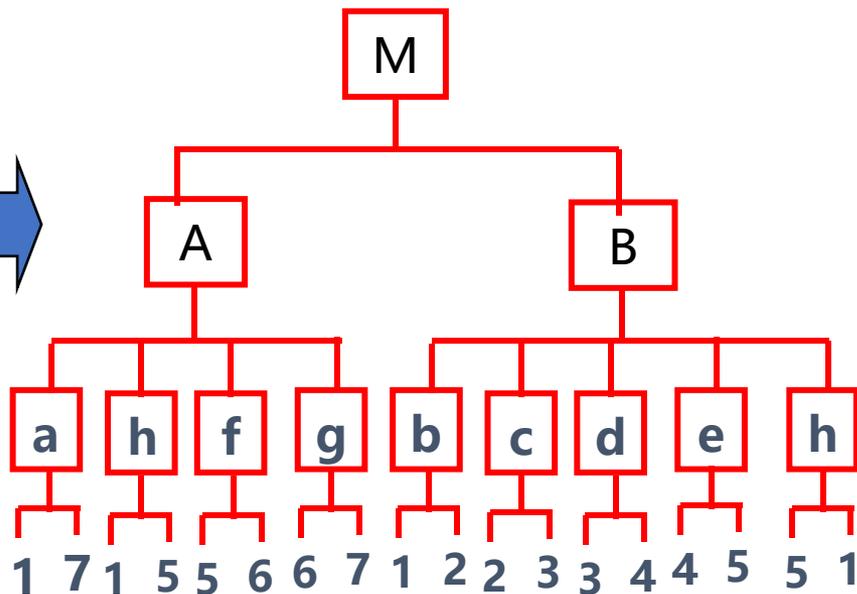
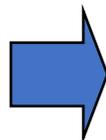
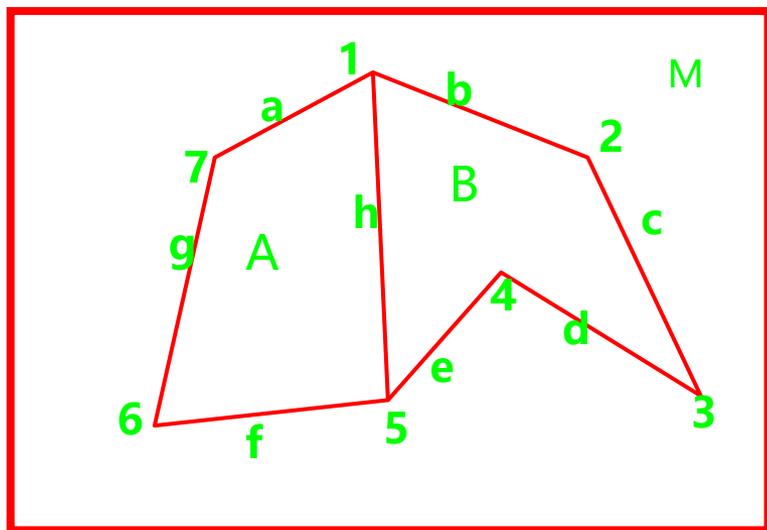


数据模型演化

5.1 空间数据库概述

5.1.1 数据库基础

□ 层次数据库模型



它的特点是将数据组织成一对多关系的结构。

层次结构采用关键字来访问其中每一层次的每一部分。

层次数据库结构特别适用于文献目录、土壤分类、部门机构等分级数据的组织。

5.1 空间数据库概述

5.1.1 数据库基础

□ 层次数据库模型的特点：

优点：

- ✓ 层次模型中各数据之间构成一种依次、从上至下的层次关系，层次清楚，数据存取、检索方便。

缺点：

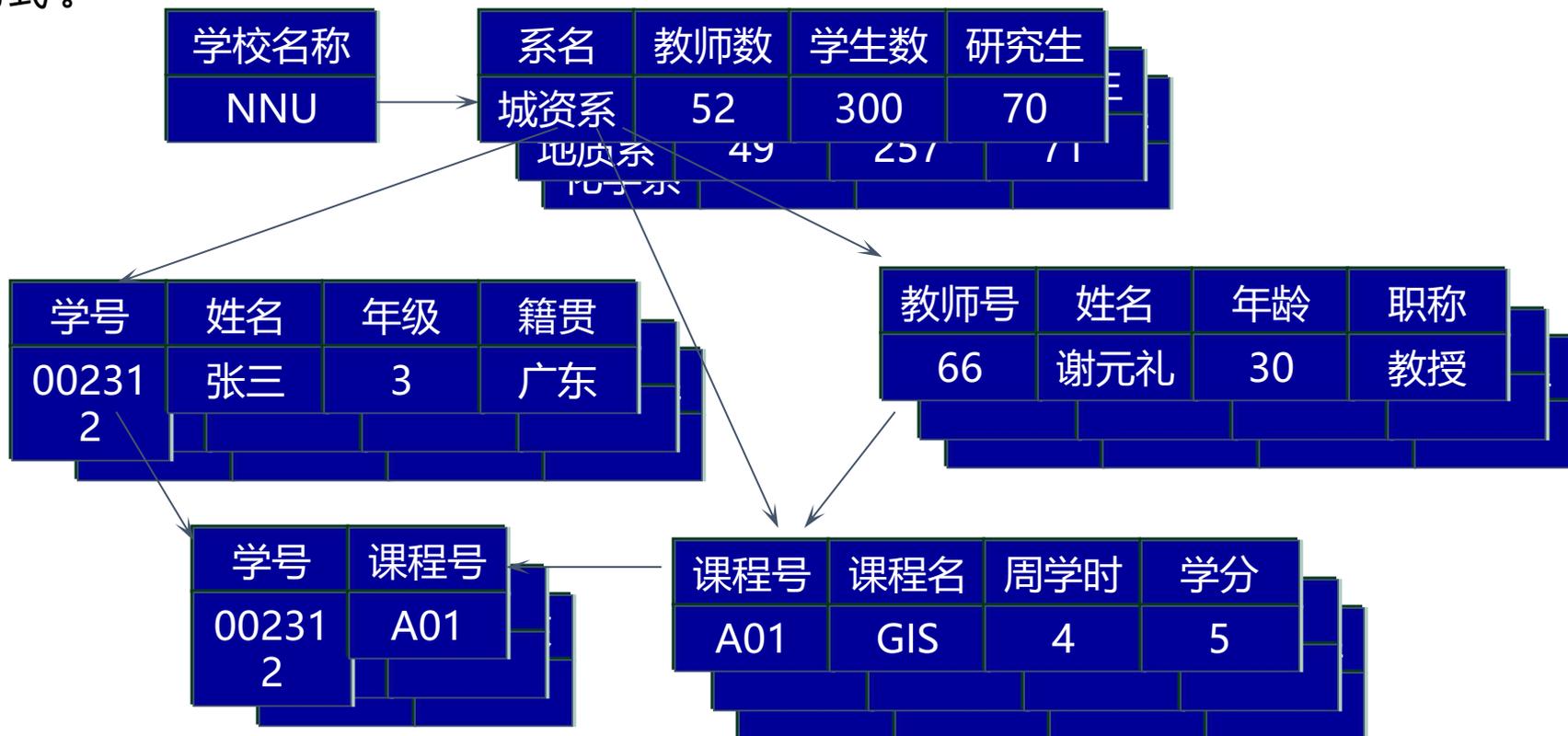
- ✓ 不能表示多对多的联系，结构缺乏灵活性；
- ✓ 在GIS中，不能顾及公共点、线数据共享和实体元素间的拓扑关系；
- ✓ 需要采用大量的索引文件；
- ✓ 实体元素重复存储，导致大量的冗余数据。

5.1 空间数据库概述

5.1.1 数据库基础

□ 网状数据库模型

网络模型用**连接指令**或**指针**来确定数据间的**显式连接关系**，是具有多对多类型的数据组织方式。



5.1 空间数据库概述

5.1.1 数据库基础

□ 网状数据库模型的特点：

采用双向指针来表示数据间的连接关系，可表示多对多关系。

优点：

- ✓ 网络模型 类似于拓扑数据模型，可表示相邻的图形特征；
- ✓ 在复杂的拓扑结构中搜索，双向环路指针很有效；
- ✓ 避免了数据冗余。

缺点：

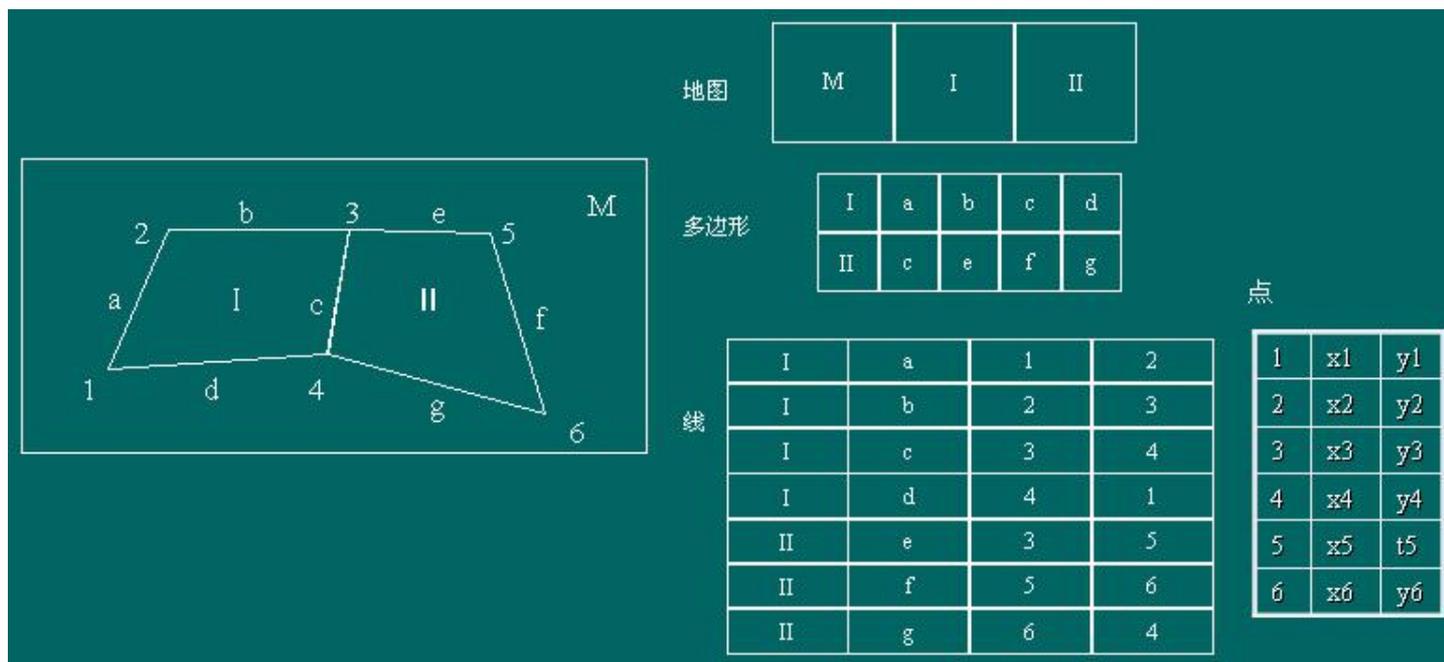
- ✓ 间接的指针占用大量存贮空间；
- ✓ 数据更新时，指针也需变化，其建立和维护困难。

5.1 空间数据库概述

5.1.1 数据库基础

□ 关系数据库模型

关系数据库模型是以记录组或**数据表**的形式组织数据，以便于利用各种地理实体与属性之间的关系进行存储和变换，**不分层也无指针**，是建立空间数据和属性数据之间关系的一种非常有效的数据组织方法。



5.1 空间数据库概述

5.1.1 数据库基础

□ 关系数据库模型的特点：

它把数据的逻辑结构归结为满足一定条件的二维表中的元素，关系表通过共同的特征联系起来。

优点：

- ✓ 对实体描述具有一致性，不需指针；
- ✓ 结构灵活，易于理解和维护，
- ✓ 可以用布尔逻辑运算和数学运算来操作数据；

缺点：

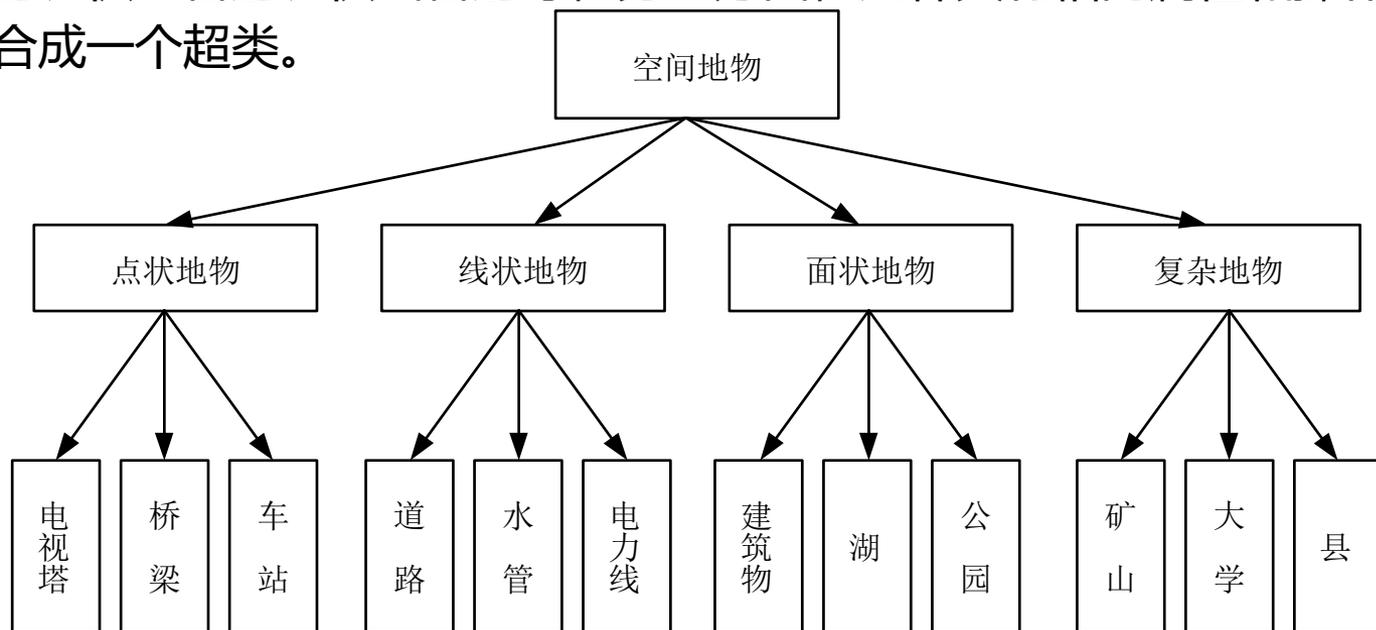
- ✓ 操作要求在文件中顺序查找满足特定关系的数据，当数据库庞大时，效率降低；
- ✓ 模型扩充困难。

5.1 空间数据库概述

5.1.1 数据库基础

□ 面向对象数据库模型

GIS中的地物可根据国家分类标准或实际情况划分类型。如一个大学GIS的对象可分为建筑物、道路、绿化、管线等几大类，地物类型的每一大类又可以进一步分类，如建筑物可再分成教学楼、科研实验楼、行政办公楼、教工住宅、学生宿舍、后勤服务建筑、体育楼等子类，管线可再分为给水管道、污水管道、电信管道、供热管道、供气管道等，另一方面，几种具有相同属性和操作的类型可综合成一个超类。

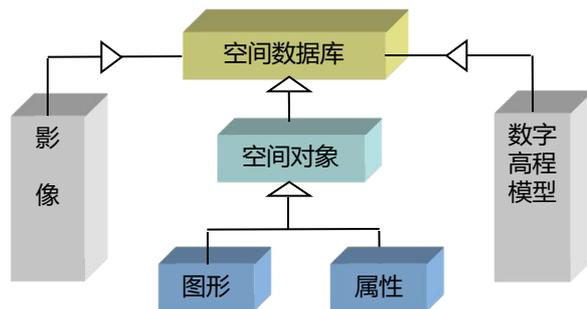


5.1 空间数据库概述

5.1.2 空间数据库

□ 数据和数据库

空间数据库 (Geodatabase) 是某一区域内关于一定地理要素特征的数据集合，是地理信息系统在计算机物理存储介质存储的与应用相关的地理空间数据的总和，一般是以一系列特定结构的文件的形式组织在存储介质之上的。



空间数据库组成

5.1 空间数据库概述

5.1.2 空间数据库

□ 空间数据库和一般数据库对比

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。
- (2) 数据结构复杂。
- (3) 数据关系多样。
- (4) 数据应用广泛。

5.1 空间数据库概述

5.1.2 空间数据库

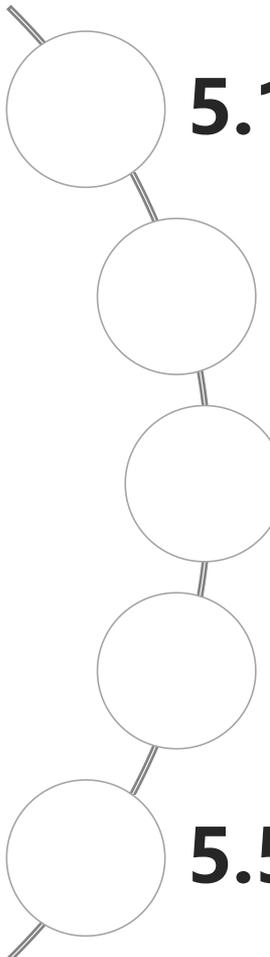
□ 空间数据库的组成

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）

本讲大纲

- 
- 5.1 空间数据库概述
 - 5.2 空间数据库设计**
 - 5.3 空间数据特征与组织
 - 5.4 空间数据管理
 - 5.5 空间数据检索

5.2 空间数据库设计

当前大纲

1.2.1 空间数据库设计内容

1.2.2 空间数据库连接步骤

5.2 空间数据库设计

5.2.1 空间数据库设计内容

数据库设计(Database Design)是指对于一个给定的应用环境，构造最优的数据库模式，建立数据库及其应用系统，使之能够有效地存储数据，满足各种用户的应用需求。

由于空间数据本身的特征，导致空间数据库的设计与传统数据库设计存在巨大的差异。良好的空间数据库设计，对于数据库的数据存储结构、存取效率等方面具有重要影响。因此对各GIS设计目标和方法有一个基本的了解至关重要。进行地理数据库设计，需要先确定要使用的数据专题，然后再指定各专题图层的内容和表现形式。

5.2 空间数据库设计

5.2.1 空间数据库设计内容

相比传统数据库，在空间数据库的设计阶段，主要包括以下几个方面的内容：

- (1) 选择数据模型与划分地理实体
- (2) 确定数据实体属性与空间结构
- (3) 实现丰富的地理实体行为
- (4) 属性关系及完整性约束。

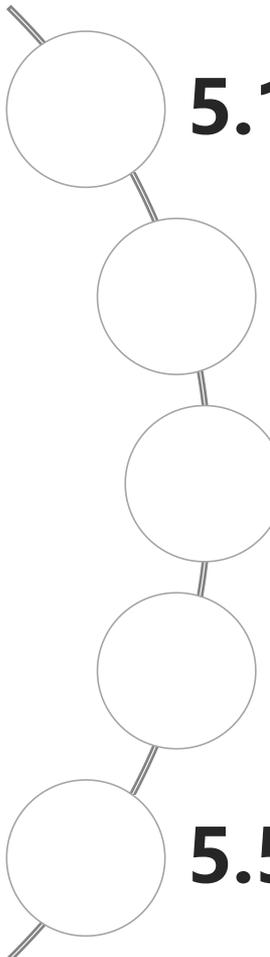
5.2 空间数据库设计

5.2.2 空间数据库设计步骤

空间数据库的设计是一个复杂的过程。其设计步骤主要包括以下几个方面：

- (1) 确定业务需求与目标信息产品。
- (2) 根据信息需求，确定主要数据专题。
- (3) 指定比例范围以及每个数据专题在每个比例下的空间表示。
- (4) 将各种表示形式分解为一个或多个地理数据集合。
- (5) 为描述性的属性定义表格型数据库结构和行为。
- (6) 定义数据集的空间行为、空间关系和完整性规则。
- (7) 构建可用的原型，查看并优化设计测试原型设计。
- (8) 记录地理数据库设计。

本讲大纲

- 
- 5.1 空间数据库概述
 - 5.2 空间数据库设计
 - 5.3 空间数据特征与组织**
 - 5.4 空间数据管理
 - 5.5 空间数据检索

5.3 空间数据特征与组织

当前大纲

5.3.1 空间数据的基本特征

5.3.2 空间数据组织

5.3.3 属性数据组织

5.3 空间数据特征与组织

5.3.1 空间数据的基本特征

1. 空间特征
2. 非结构化特征
3. 空间关系特征
4. 多尺度与多态性
5. 分类编码特征
6. 海量数据特征

5.3 空间数据特征与组织

5.3.1 空间数据的基本特征

由于空间数据存在非结构化特征、空间关系特征，使得通用数据库管理系统在管理空间数据时，面临较多问题。

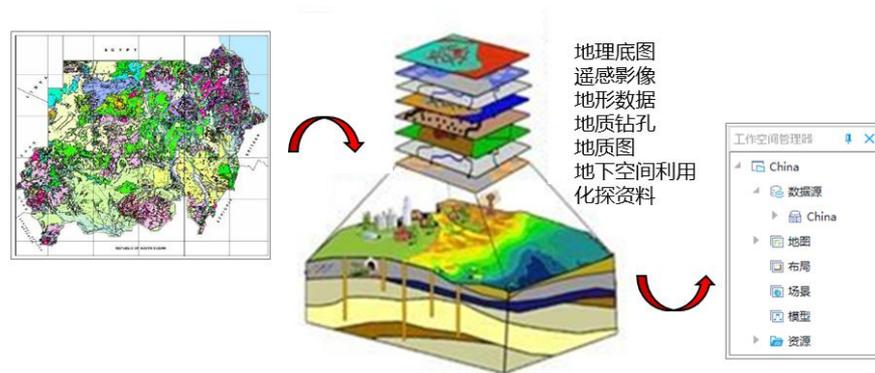
- (1) GIS需要一些复杂的图形功能，一般的DBMS不能支持。
- (2) DBMS一般都难以实现对空间数据的关联、连通、包含、叠加等基本操作。
- (3) 地理信息表达复杂，表达单个地理实体需多个文件、多条记录，一般的DBMS也难以支持。
- (4) 具有高度内部联系的GIS数据记录需要更复杂的安全性维护系统。一般的DBMS都难以保证的。
- (5) GIS中空间数据记录是变长的（存储的坐标点的数目随空间对象的变化而变化），而一般数据库都只允许把记录的长度设定为固定长度。

5.3 空间数据特征与组织

5.3.2 空间数据的组织

1. 空间数据的分层组织

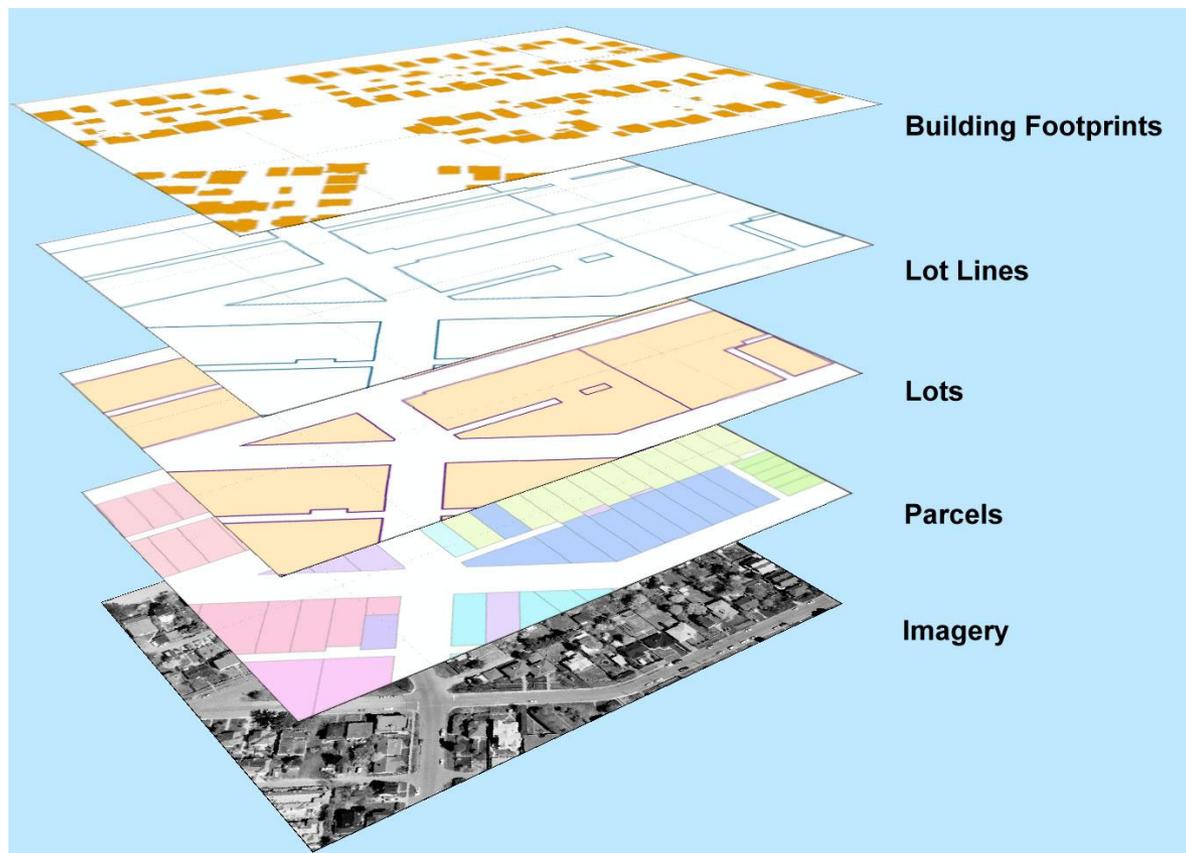
将表示同一地理范围内众多地理要素和地理现象的空间数据采用“分层”方式进行数据组织（图），这是一种起源于地图制图的空间数据组织方式。在分层数据组织中，图层（layer）可根据地理事物或地理现象的分类，按数据类型（矢量、栅格、影像等）、专题内容（theme）、要素几何类别（点、线和面）、时间次序等设定



5.3 空间数据特征与组织

5.3.2 空间数据的组织

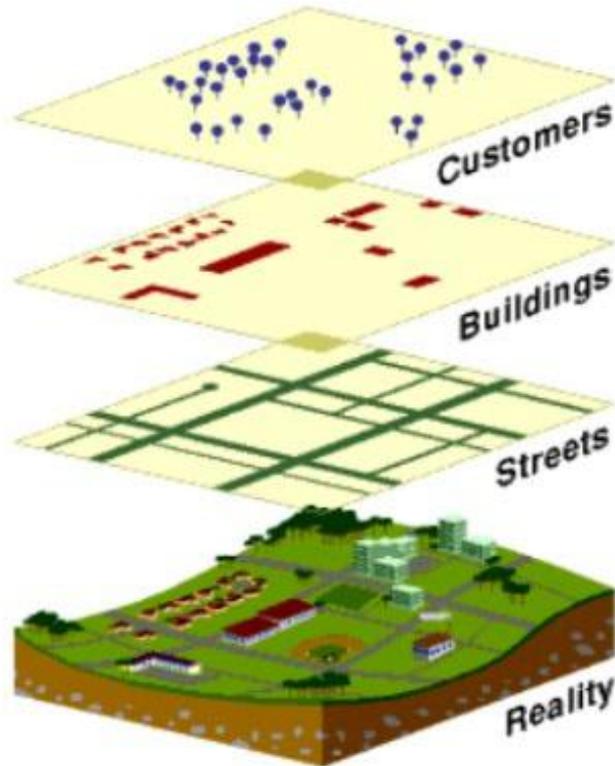
1. 空间数据的分层组织——按专题分层



5.3 空间数据特征与组织

5.3.2 空间数据的组织

1. 空间数据的分层组织——按专题分层



點

- 如氣候測站、交通事故地點、消防栓、地下水井、點污染源等。

面

- 如地籍坵塊、行政區域、土地利用、建築物、土壤等

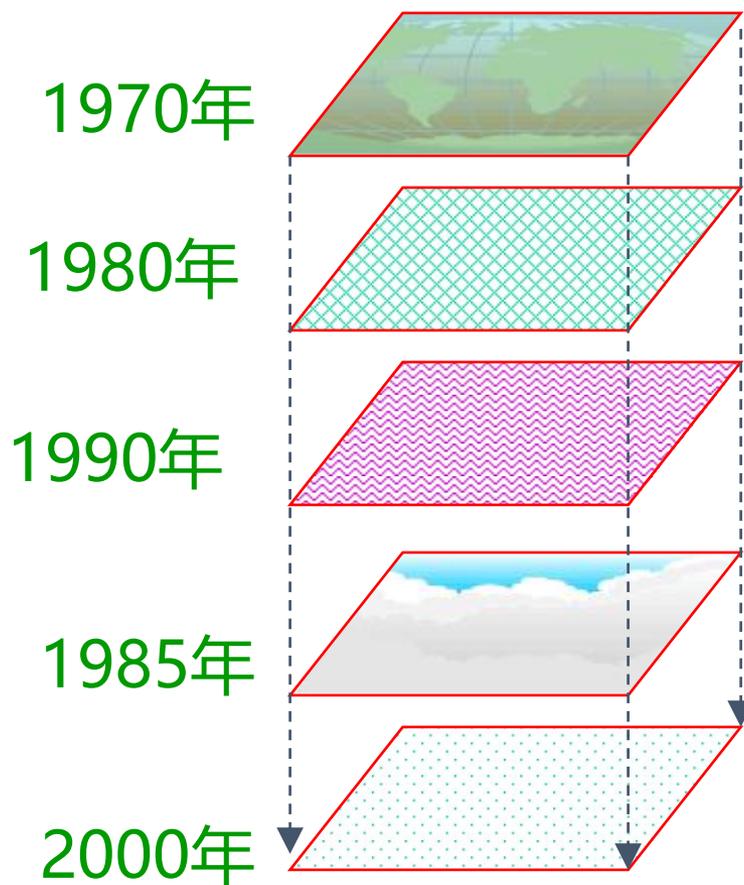
線

- 如道路、河流、地下管線、斷層線等

5.3 空间数据特征与组织

5.3.2 空间数据的组织

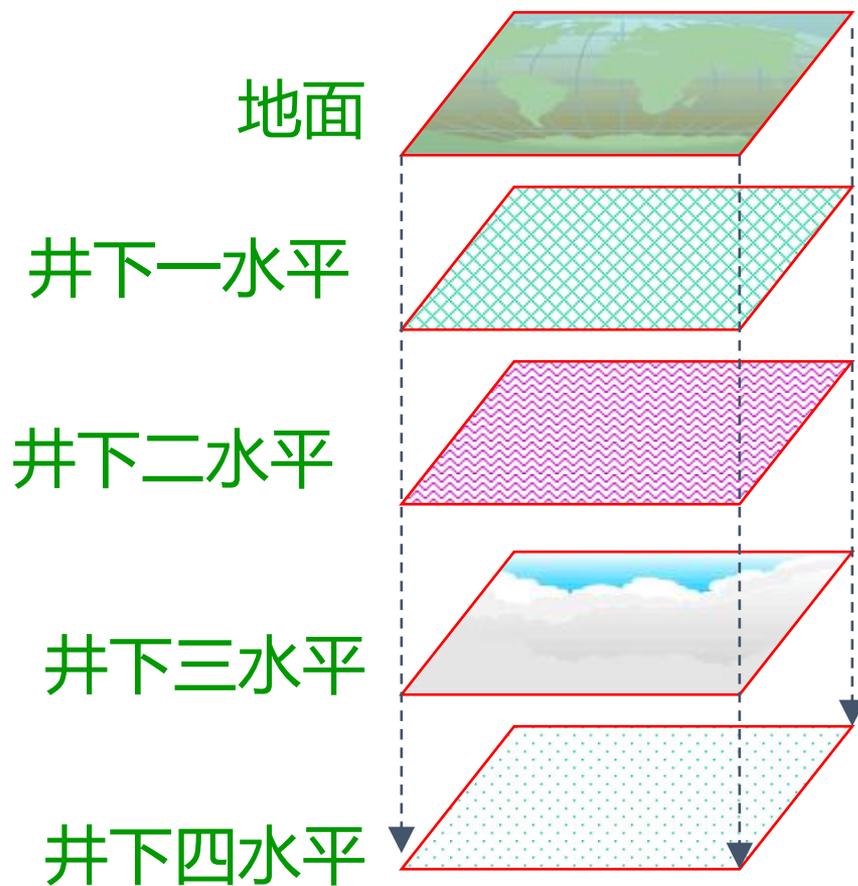
1. 空间数据的分层组织——按时间分层



5.3 空间数据特征与组织

5.3.2 空间数据的组织

1. 空间数据的分层组织——按垂直高度分层



5.3 空间数据特征与组织

5.3.2 空间数据的组织

1.空间数据的分层组织——优缺点

优点

- ✓ 这种方式是目前颇为普遍的数据组织方法，方便使用者选择合适的数
据，适合与栅格或矢量数据数据结构，目前大多数GIS软件采用这一
方法。

缺点

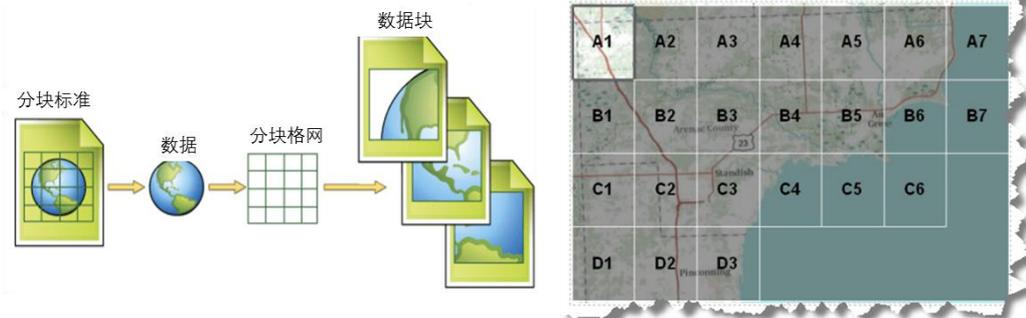
- ✓ 其缺点是层与层之间的数据必须经过层叠置(Overlay)处理才能关联在
一起，在叠置处理中，对栅格数据常需要大量存储空间来完成操作，
而矢量数据则需大量的计算处理。
- ✓ 同一图层内的图形数据的空间关系较为简单并易于处理，但不同图层
之间的空间关系难以处理。

5.3 空间数据特征与组织

5.3.2 空间数据的组织

2.空间数据的分块组织

当对大范围区域内众多类型空间数据进行存储和管理时，为了提高数据存储与管理的效率，可将空间数据所覆盖的区域范围分割为若干个块或分区，按块分别进行空间数据的组织。块可以是规则的，如遵照国家标准《国家基本比例尺地形图分幅和编号》（GB/T 13989-2012）的各级比例尺地形图图幅范围所划分的规则块，也可以是不规则的，如按照行政区边界范围进行不规则分块。



5.3 空间数据特征与组织

5.3.2 空间数据的组织

2.空间数据的分块组织——按图幅组织



BR012-5-2	BS012-5-1	BS012-5-2
BR012-5-4	BS012-5-3	BS012-5-4
BR011-5-2	BS011-5-2	BS011-5-2

5.3 空间数据特征与组织

5.3.2 空间数据的组织

2.空间数据的分块组织——优缺点

优点

- ✓ 分区式与分层式可同时采用，并不冲突。
- ✓ 分区式也是目前大部份商业软件所采用的方法，适合与栅格和矢量数据结构，在数据量大的系统中，分区方法可提高数据存取的效率。

缺点

- ✓ 图幅或区块间的衔接问题是分区法最大的困扰，尤其在空间数据查询、分析操作时更是这样。

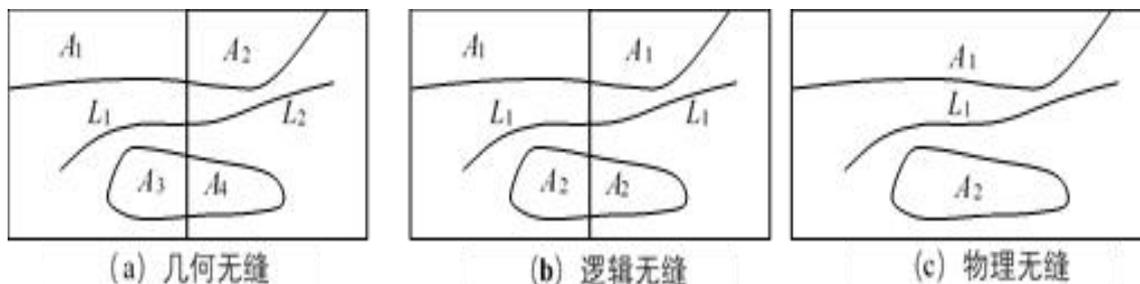
5.3 空间数据特征与组织

5.3.2 空间数据的组织

3. 空间数据的无缝组织

为了克服空间数据分幅或分块组织时，导致对跨越多个图幅或分块地理要素的割裂或不一致，从而难以查询和分析等问题，在涉及大范围、海量空间数据的数据组织时，通常采用连续、无缝的数据组织形式，以满足用户任意和透明地访问和操作数据的要求。

无缝空间数据组织有三种实现途径：**几何无缝**、**逻辑无缝**和**物理无缝**的各分幅或分块表示的地理要素都转换到该坐标框架下并进行几何接边处理。

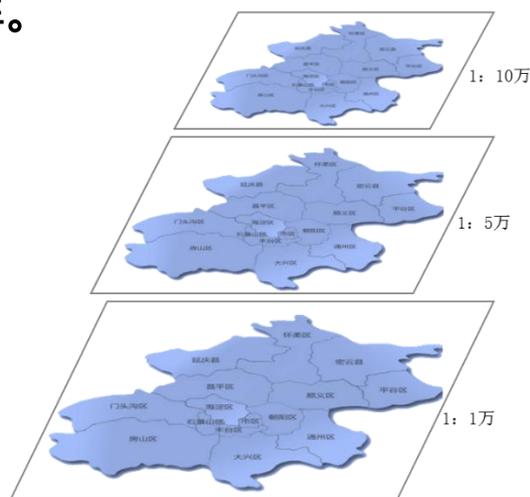


5.3 空间数据特征与组织

5.3.2 空间数据的组织

4. 多尺度空间数据组织

从理论上来说，在对现实世界的数字化表达中，不存在比例尺的概念，但从观察、理解及制图的角度来看，当涉及大范围区域时，往往需要从宏观到微观，以不同的层次细节来刻画地理要素，这就要求必须建立多尺度或多比例尺空间数据库。



5.3 空间数据特征与组织

5.3.3 属性数据的组织

属性数据由关系数据库管理系统管理，但它的文件组织方式也要服从上述工作层、工作区和图库的要求，以便于图形文件协调工作，共同组成工作区、工作层，并进行跨图幅操作。在不同的商业化软件中，属性文件组织方式各不相同。主要的3种方式如下：

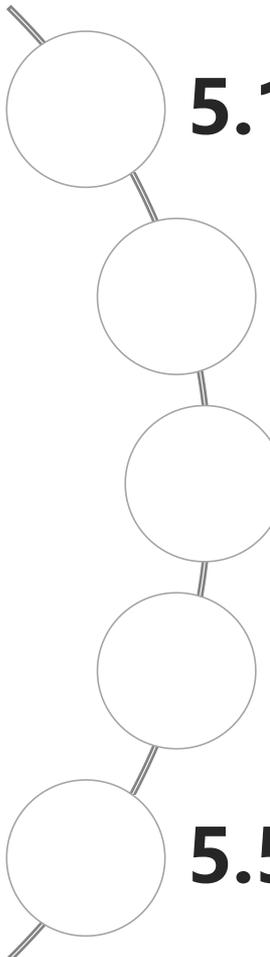
1. 与工作层对应的组织方式
2. 与地物类对应的组织方式
3. 混合方式



State	County	Rain	Total
Oklahoma	Atoka	1.80	10.16
Oklahoma	Kiowa	2.34	13.67
Oklahoma	Nowata	1.62	11.90

State	Count	Avg_Rain	Max_Rain
Ohio	88	3.21	4.50
Oklahoma	77	2.56	3.86
Oregon	36	5.66	7.92

本讲大纲

- 
- 5.1 空间数据库概述
 - 5.2 空间数据库设计
 - 5.3 空间数据特征与组织
 - 5.4 空间数据管理**
 - 5.5 空间数据检索

5.4 空间数据管理

当前大纲

- 5.4.1 矢量数据的管理
- 5.4.2 栅格数据的管理
- 5.4.3 时空大数据的管理
- 5.4.4 空间数据库引擎

5.4 空间数据管理

5.4.1 矢量数据的管理

1. 文件-关系数据库混合管理

由于空间数据的非结构化特征，早期关系型数据库难以满足空间数据管理的要求。因此，传统GIS软件采用文件与关系数据库混合方式管理空间数据，比较典型的是ArcInfo，有的系统也采用纯文件方式管理空间数据，如MapInfo；即用文件系统管理几何图形数据，用商用关系型数据库管理属性数据，两者之间通过目标标识或内部连接码进行连接

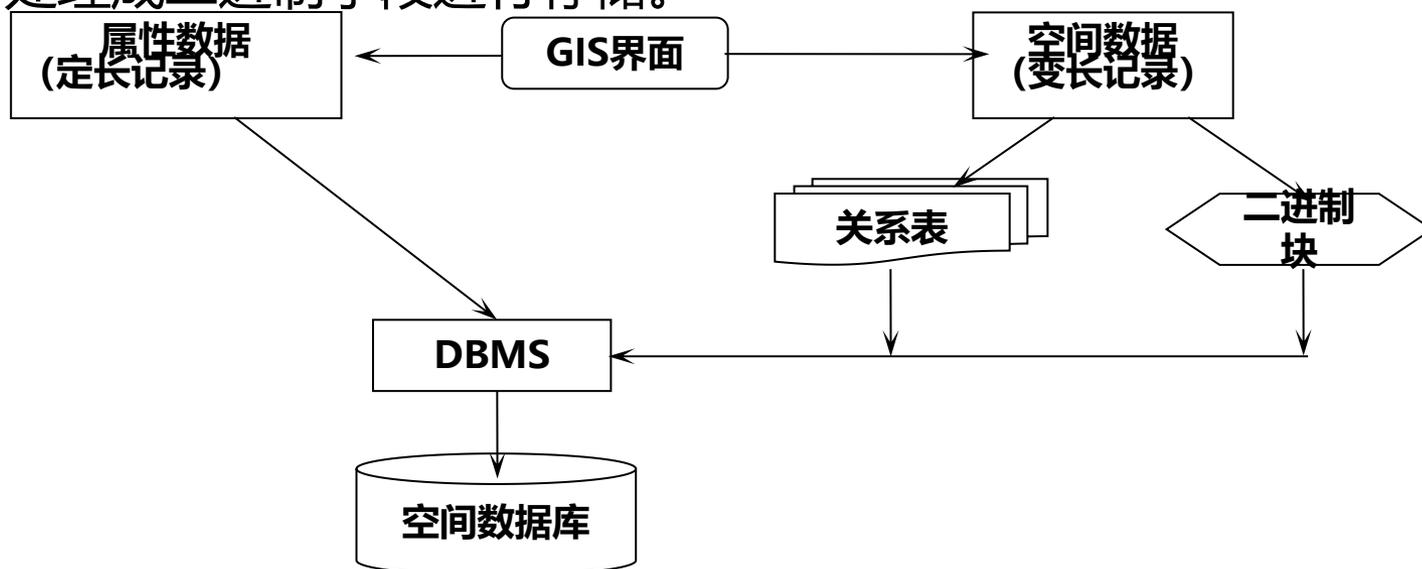
OID（目标ID 或内部连接码）	图形数据
	属性数据

5.4 空间数据管理

5.4.1 矢量数据的管理

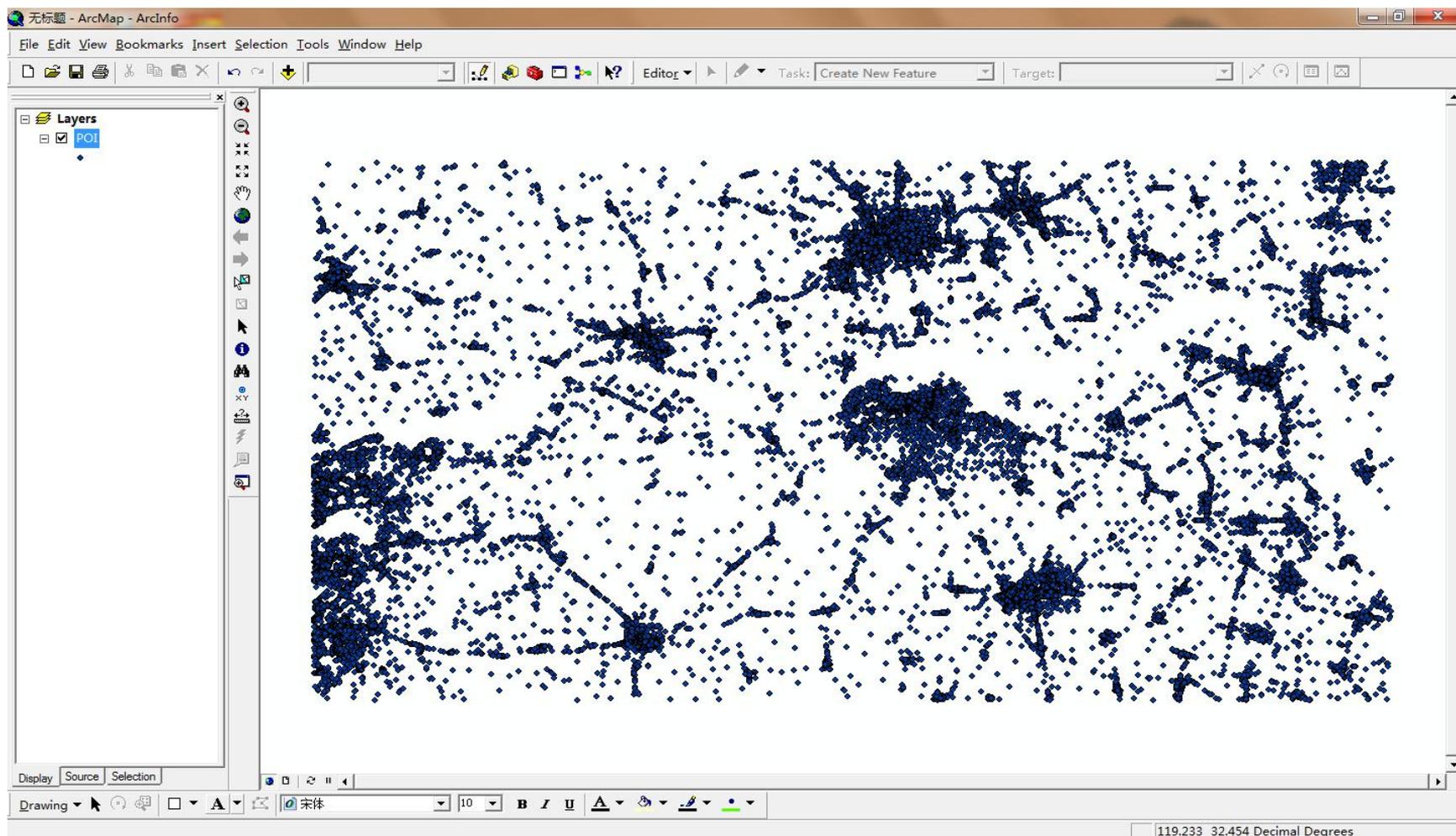
2. 全关系数据库混合管理

全关系数据库管理方式下，图形数据与属性数据都采用现有的关系型数据库存储，使用关系数据库标准连接机制来进行空间数据与属性数据的连接。对于变长数据采用定长记录的数据表或者处理成二进制字段进行存储。



5.4 空间数据管理

5.4.1 矢量数据的管理



POI数据在 Personal Geodatabase中的组织

5.4 空间数据管理

5.4.1 矢量数据的管理

OBJECTID	Shape	LAYER	FID_1	NAME	POPNAME	ENTRANCE	CLASS	DISTRICT	ADDRESS	ISPARKING
1	长二进制数据	Unknown Poi	0	古泉综合服务	古泉综合服务	0	0XB0FF	320115000	X426	0
2	长二进制数据	Unknown Poi	0	南京市江宁区	信用社	0	0XA180	320115000	X426	0
3	长二进制数据	Unknown Poi	0	南京东陶化工	南京东陶化工	0	0XBB60	320115000		0
4	长二进制数据	Unknown Poi	0	古泉联小	古泉联小	0	0X7400	320115000		0
5	长二进制数据	Unknown Poi	0	中共南京市江	古泉村总支部	0	0X1050	320115000	X426	0
6	长二进制数据	Unknown Poi	0	东方星娱乐园	东方星娱乐园	0	0X38C0	320115000		0
7	长二进制数据	Unknown Poi	0	高桥道班	高桥道班	0	0X5420	320115000		0
8	长二进制数据	Unknown Poi	0	古泉行政村	古泉行政村	0	0XFF86	320115000	X426	0
9	长二进制数据	Unknown Poi	0	古泉山庄	古泉山庄	0	0X2800	320115000		0
10	长二进制数据	Unknown Poi	0	侯家塘	侯家塘	0	0XFF86	320115000		0
11	长二进制数据	Unknown Poi	0	巴厘原墅	巴厘原墅	0	0XB900	320115000		0
12	长二进制数据	Unknown Poi	0	南京新天和制	南京新天和制	0	0XBA80	320115000	X426	0
13	长二进制数据	Unknown Poi	0	南京古泉集团	南京古泉集团	0	0XBA80	320115000	X426	0
14	长二进制数据	Unknown Poi	0	南京古泉塑料	南京古泉塑料	0	0XBB60	320115000	X426	0
15	长二进制数据	Unknown Poi	0	骨胶厂	骨胶厂	0	0XBB60	320115000		0
16	长二进制数据	Unknown Poi	0	汤山头	汤山头	0	0XFF86	320115000		0
17	长二进制数据	Unknown Poi	0	南京加特橡塑	南京加特橡塑	0	0XBB60	320115000	X426	0
18	长二进制数据	Unknown Poi	0	南京汤山彭润	南京汤山彭润	0	0XBA80	320115000		0
19	长二进制数据	Unknown Poi	0	南京双全科技	南京双全科技	0	0XBA80	320115000		0
20	长二进制数据	Unknown Poi	0	黄栗墅服务区	黄栗墅服务区	0	0X5090	320115000		0
21	长二进制数据	Unknown Poi	0	黄栗墅服务区	黄栗墅服务区	0	0X5090	320115000		0
22	长二进制数据	Unknown Poi	0	苏澳电子公司	苏澳电子公司	0	0XBA80	320115000	X426	0
23	长二进制数据	Unknown Poi	0	黄栗墅	黄栗墅	0	0XFF86	320115000		0
24	长二进制数据	Unknown Poi	0	黄栗墅服务区	黄栗墅服务区	0	0X5090	320115000	沪宁高速	0
25	长二进制数据	Unknown Poi	0	大凹	大凹	0	0XFF90	320115000		0
26	长二进制数据	Unknown Poi	0	中国石化加油	中国石化加油	0	0X5010	320115000	X426	0
27	长二进制数据	Unknown Poi	0	加油站	加油站	0	0X5010	320115000		0
28	长二进制数据	Unknown Poi	0	加油站	加油站	0	0X5010	320115000		0
29	长二进制数据	Unknown Poi	0	黄栗墅服务区	黄栗墅服务区	0	0X5090	320115000		0
30	长二进制数据	Unknown Poi	0	加油站	加油站	0	0X5010	320115000		0

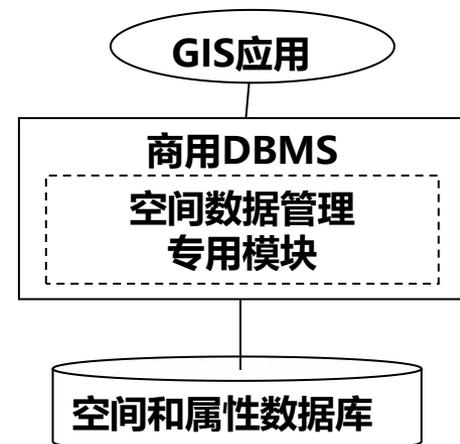
POI数据在 Personal Geodatabase中的组织

5.4 空间数据管理

5.4.1 矢量数据的管理

3. 对象-关系数据库管理

由于直接采用通用的关系数据库管理系统的效率不高，而非结构化的空间数据又十分重要，所以许多数据库管理系统的软件商在关系数据库管理系统中进行扩展，使之能直接存储和管理非结构化的空间数据（图 5.12），如Informix和Oracle等都推出了空间数据管理的专用模块，定义了操纵点、线、面、圆、长方形等空间对象的API函数。这些函数，将各种中间对象的数据结构进行了预先的定义。

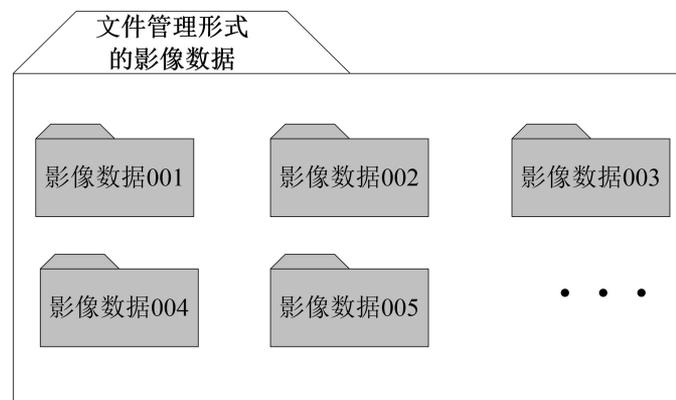


5.4 空间数据管理

5.4.2 栅格数据的管理

1. 文件管理方式

目前大部分GIS软件和遥感图像处理软件都是采用文件方式来管理遥感影像数据。由于遥感影像数据库并不是仅仅包含图像数据本身，而且还包含大量的图像元数据信息（如图像类型、摄影日期、摄影比例尺等），遥感图像数据本身还具有多数据源、多时相等特点，另外，数据的安全性、并发控制和数据共享等都将使文件管理无法应付。



5.4 空间数据管理

5.4.2 栅格数据的管理

2. 文件-数据库管理方式

为了改进文件方式管理影像数据的效率，一种新的管理方式被提出来：文件+数据库管理方式。实施这种方式管理影像数据时，影像数据仍按照文件方式组织管理；在关系数据库中，每个文件都有唯一的标识号（ID）对应影像信息，如文件名称、存储路径等。

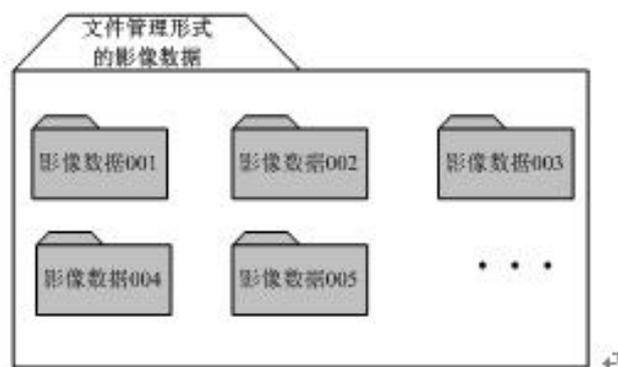


表 5.1 影像信息数据库表

影像名称	块号	...
Image-001	011001	...
Image-002	011002	...
Image-003	011003	...
Image-004	021001	...
Image-005	021002	...
...
...

5.4 空间数据管理

5.4.2 栅格数据的管理

3. 关系数据库管理

由于关系数据库发展成熟，具有良好的安全措施和数据恢复机制；目前关系数据库系统提供了存储复杂数据类型的能力，使利用关系数据库来管理影像数据成为可能。基于扩展关系数据库的影像数据库管理是将影像数据存储到二进制变长字段中，然后应用程序通过数据访问接口来访问数据库中的影像数据。同时影像数据的元数据信息存放在关系数据库的表中，二者可以进行无缝管理。

5.4 空间数据管理

5.4.2 栅格数据的管理

3. 关系数据库管理——特点

- ✓ ① 所有数据集中存储，数据安全，易于共享。
- ✓ ② 较方便管理多数据源和多时态的数据。
- ✓ ③ 支持事务处理和并发控制，有利于多用户的访问与共享。
- ✓ ④ 影像数据和元数据集成到一起，能方便的进行交互式查询。
- ✓ ⑤ 对Client/Server的分布式应用支持较好，网络性能和数据传输速度都有提高。
- ✓ ⑥ 影像数据访问只能通过数据库驱动接口访问，有利于数据的一致性和完整性控制，数据不会被随意移动、修改和删除。
- ✓ ⑦ 支持异构的网络模式，即应用程序和后台数据库服务器可以在不同操作系统平台下运行。现有商用数据库都有良好多的网络通讯机制，本身能够实现异构网络的分布式计算，使得应用程序的开发相对简单化。

5.4 空间数据管理

5.4.3 时空大数据的管理

时空大数据的来源

随着计算机技术、物联网、移动通信及遥感等技术的进一步发展和完善，人类进入了一个前所未有的数据大爆炸时代。许多与地理位置、时间相关的时空大数据在数据管理方面给传统的GIS数据管理带来了挑战和机遇，一些应对时空大数据管理的方法和技术应运而生，时空大数据主要来源于以下几个方面：

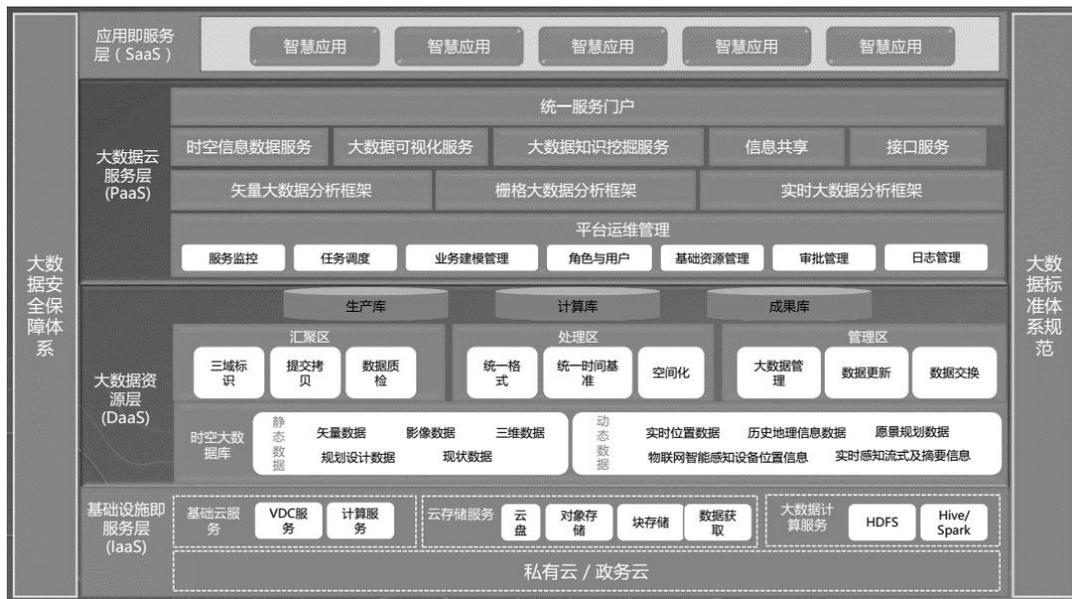
1. 基础测绘数据与专题数据
2. 遥感影像数据
3. 导航定位数据
4. 互联网及物联网数据

5.4 空间数据管理

5.4.3 时空大数据的管理

时空大数据的管理

时空大数据的存储管理，不能仅仅考虑或者重点考虑数据存储层面的问题。必须从整个时空大数据平台入手展开顶层设计。既要考虑多源异构大数据的接入，组织、提取，又要考虑分析过程数据、分析结果数据的协调和组织。

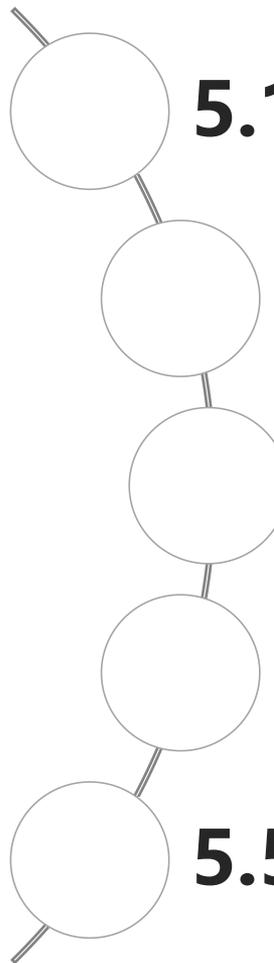


5.4 空间数据管理

5.4.4 空间数据引擎

之前采用关系数据库与文件混合管理模式的传统GIS数据库系统技术，在应用上取得了一定的成功，但不得不部分地采取文件方式管理，总体上无法达到数据库技术冗余度、独立性等要求，用现代数据库技术统一存放和管理空间数据与属性数据是GIS发展的必然趋势。1996年，ESRI公司与Oracle等数据库开发商合作，开发出一种能将空间图形数据也存放到大关系数据库中管理的产品，将其定名为“spatial database engine”，简称SDE，即为“空间数据库引擎”。

本讲大纲

- 
- 5.1 空间数据库概述
 - 5.2 空间数据库设计
 - 5.3 空间数据特征与组织
 - 5.4 空间数据管理
 - 5.5 空间数据检索

5.5 空间数据检索

当前大纲

5.5.1 空间数据索引概述

5.5.2 空间数据索引算法

5.5.3 空间数据库查询语言

5.5 空间数据检索

5.5.1 空间数据索引概述

索引是对数据库表中一列或多列的值进行排序的一种结构，使用索引可快速访问数据库表中的特定信息。

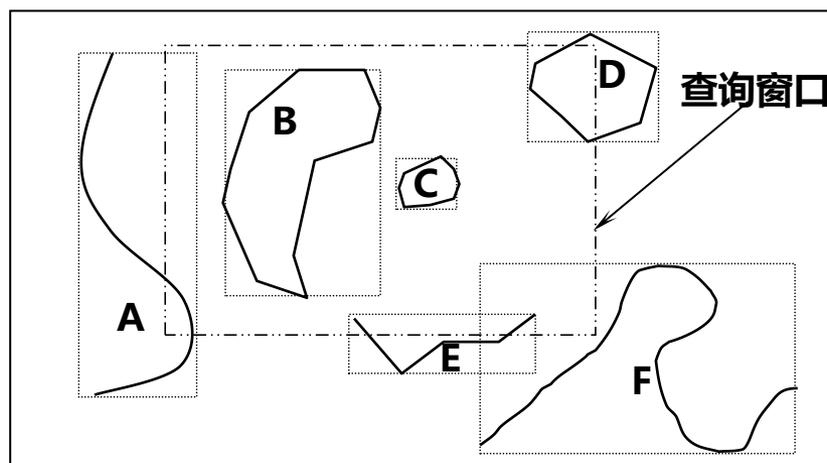
空间索引是指依据空间对象的位置和形状或空间对象之间的某种空间关系按一定的顺序排列的一种数据结构，其中包括空间对象的概要信息，如对象的标识、外接矩形及其指向空间对象实体的指针。作为一种辅助性的空间数据结构，空间索引介于空间操作算法和空间对象之间。

5.5 空间数据检索

5.5.2 空间数据索引算法

1.对象范围索引

在记录每个空间实体的坐标时，记录包围每个空间实体的外接矩形的最大最小坐标。这样，在检索空间实体时，根据空间实体的最大最小范围，预先排除那些没有落入检索窗口内的空间实体，仅对那些外接矩形落在检索窗口内的空间实体作进一步的判断，最后检索出那些真正落入窗口内的空间实体。



5.5 空间数据检索

5.5.2 空间数据索引算法

2. 格网索引

格网型空间索引思路比较简单，容易理解和实现。其基本思想是将研究区域用横竖线条划分大小相等和不等的格网，记录每一个格网所包含的空间实体。当用户进行空间查询时，首先计算出用户查询对象所在格网，然后再在该网格中快速查询所选空间实体，这样一来就大大地加速了空间索引的查询速度。

21	23	29	31	53	55	61	63
20	22	28	30	52	54	60	62
17	19	25	27	49	51	57	59
16	18	24	26	48	30	56	58
5	7	13	15	37	39	45	47
4	6	12	14	36	38	44	46
1	3	9	11	33	35	41	43
0	2	8	10	32	34	40	42

空间索引表

Peano 码	实体
7	B
14	F
15	F
25	A
26	F
32	D
33	D
35	D,G
37	F
38	D
39	F
48	F
50	F
54	C
55	C
60	C

实体索引表

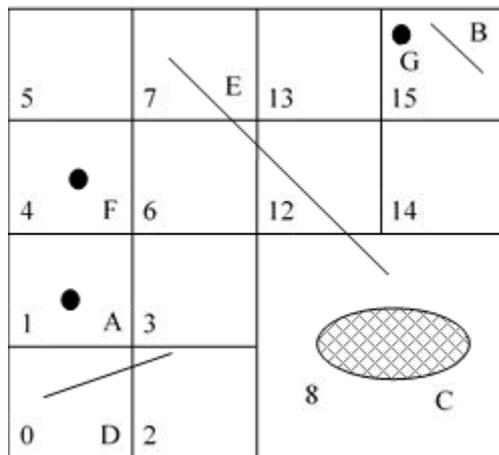
实体	Peano 码
A	25-25
B	7-7
C	54-55
C	60-60
D	32-33
D	35-35
D	38-38
F	14-15
F	26-26
F	37-37
F	39-39
F	48-48
F	50-50
G	35-35

5.5 空间数据检索

5.5.2 空间数据索引算法

2. 四叉树空间索引

四叉树作为一种有效的数据结构，不仅可以用来对栅格数据进行组织，它还可用于建立空间数据的索引。四叉树中的线性四叉树和层次四叉树都可以用于建立空间索引。



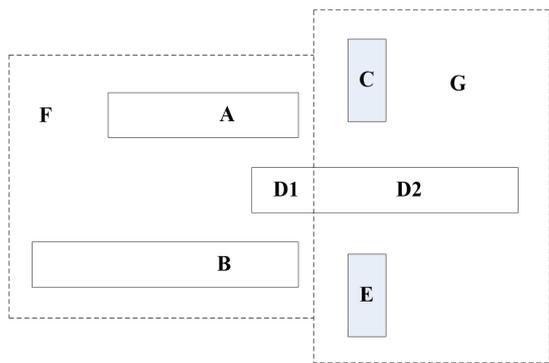
peano 码	边长	实体
0	4	E
0	2	D
1	1	A
4	1	F
8	2	C
15	1	B、G

5.5 空间数据检索

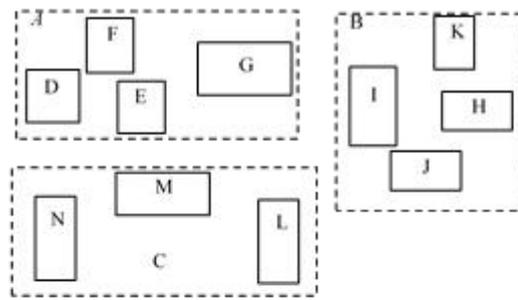
5.5.2 空间数据索引算法

2. R树和R+树空间索引

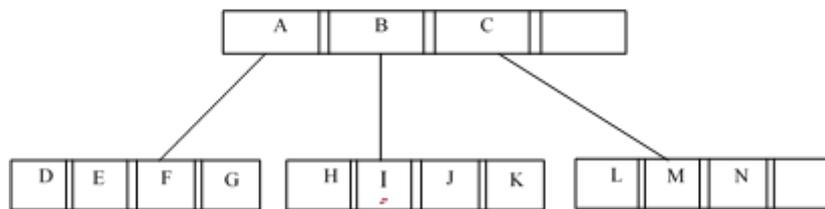
与实体范围索引类似，R树和R+树利用空间实体的外接矩形来建立空间索引。就R树而言，认为有N个实体被N个外接矩形（Rectangles, R）所包围。R+树的数据结构与R树的相同，但是，对于被分割的下层虚拟矩形或实体外接矩形，还要增加关系表达式。



R+树结构示意图



(a) 二层不重叠矩形



(b) 层状结构

R树结构示意图

5.5 空间数据检索

5.5.3 空间数据库查询语言

1. 标准查询语言

SQL是一种介于关系代数与关系演算之间的结构化查询语言，它不仅仅是查询，是一个通用的、功能极强的关系数据库语言。SQL语言是一个综合的、功能极强同时又简捷易学的语言。

SQL语言的动词	
SQL功能	动词
数据查询	Select
数据定义	Create, Drop, Alter
数据操纵	Insert, Update, Delete
数据控制	Grant, Revoke

5.5 空间数据检索

5.5.3 空间数据库查询语言

2. 扩展SQL处理空间数据

SQL不足之处是只提供简单的数据类型：整型、日期型、字符串型等。空间数据库（SDB）的应用必须能处理多点、线和多边形这样的复杂的数据类型。亟需对SQL语言进行空间扩展。

OGIS标准定义的一些操作

OGIS标准定义的一些操作		
基本函数	SpatialReference ()	返回几何体的基本坐标系统
	Envelope ()	返回包含几何体的最小外接矩形
	Export ()	返回以其他形式表示的几何体
	IsEmpty ()	如果几何体是空集则返回真
	IsSimple ()	如果几何体是简单的（即不自交）则返回真
	Boundary ()	返回几何体的边界
拓扑/集合运算符	Equal	如果两个几何体的内部和边界在空间上相等，则返回真
	Disjoint	如果内部和边界不相交，则返回真
	Intersect	如果几何体不相交，则返回真
	Touch	如果两个面仅仅是边界相交但是内部不相交，则返回真
	Cross	如果一条线和面的内部相交，则返回真
	Within	如果给定的几何体的内部不和另一个几何体的外部相交，则返回真
	Contains	判断给定的几何体是否包含另一个给定的几何体
空间分析	Overlap	如果两个几何体的内部有非空交集，则返回真
	Distance	返回两个几何体之间的最短距离
	Buffer	返回到给定几何体的距离小于或等于指定值得几何体的点集合
	ConvexHull	返回几何体的最小闭包
	Intersection	返回由两个几何体的交集构成的几何体
	Union	返回由两个几何体的并集构成的几何体
	Difference	返回几何体与给定几何体不相交的部分
SymmDiff	返回两个几何体与对方互不相交的部分	

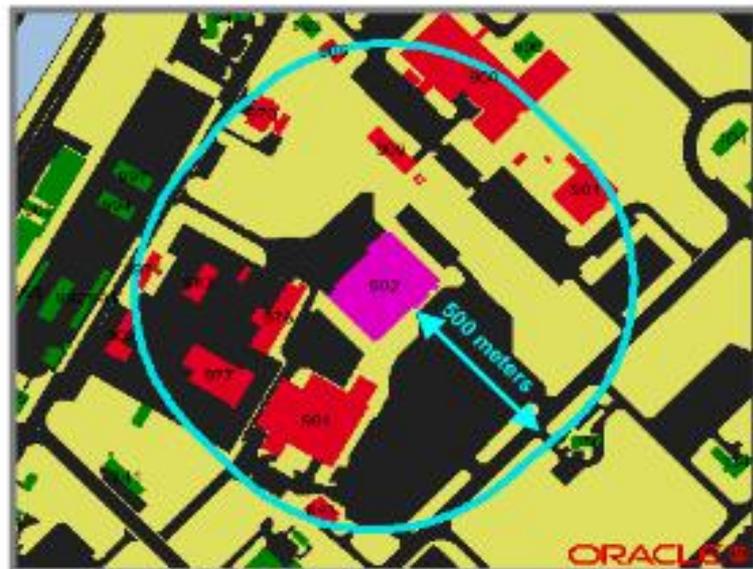
5.5 空间数据检索

5.5.3 空间数据库查询语言

通过 **SQL** 进行空间查询

找出 902号建筑物周围 500米内的所有建筑物

```
SQL> SELECT a.building_id
2>   FROM base_buildings a,
3>        base_buildings b
4>  WHERE b.building_id = 902
5>        AND MDSYS.SDO_WITHIN_DISTANCE(
6>          a.Location, b.Location,
7>          'distance=500 unit=meter')
8>          = 'TRUE';
```



专业术语

空间数据库、空间数据库设计、空间数据库引擎、空间索引、对象范围索引、格网索引、R树和R+树索引

复习思考题

一、思考题（基础部分）

- 1、什么是空间数据库，具有什么特点？
- 2、空间数据库的设计步骤有哪些？
- 3、矢量数据的管理方式有哪些，各有什么优缺点？
- 4、栅格数据的管理方式有哪些，各有什么优缺点？
- 5、数据库中空间数据是如何进行分幅分层组织的？
- 6、空间数据的索引方式有哪些，比较各种方法的优缺点。

复习思考题

二、思考题（拓展部分）

- 1、通过阅读资料，比较ArcSDE、Oracle Spatial等空间数据引擎的特点。
- 2、以江苏省为例，讨论建立江苏省省级GIS的数据管理思路和方法。
- 3、分析常用GIS软件的空间和属性数据管理的特点。
- 4、试比较ArcGIS、MapGIS和Supermap的属性数据管理策略之特点。