

文章编号: 1000-0585(2001)03-0364-08

# 综合省情地理信息系统地理 数据库的数据组织

马荣华, 黄杏元, 赵振斌

(南京大学城市与资源学系, 南京 210093)

**摘要:** 地理信息系统在政府机关的应用形成了 GIS 的专门研究领域, 综合省情地理信息系统是其研究的重要内容之一, 海量数据的组织与管理是这一领域走向成熟与应用的关键。文中采用的数据组织方法打破了图幅的界限, 把数据按行政单元进行存储管理, 对时态数据采用了一种改进的快照模型, 对多比例尺采用分别建库的方法, 对影像数据和 DEM 数据分别按块存储。这样对数据的查询检索可以按行政单元、比例尺、时间(时相)、地物类别以及按块进行, 大大提高了查询检索的速度, 满足为政府领导机关服务的目的。

**关 键 词:** GIS; 数据库; 数据组织; 查询检索

**中图分类号:** P208      **文献标识码:** A

信息化是当今世界经济与社会发展的大趋势, 地理信息系统(GIS)为社会信息化提供了通用的空间信息框架, 而其在政府机关的应用形成了 GIS 的专门研究领域, 即综合国情、省情、市情甚至县(市)情地理信息系统, 它依赖于大量甚至海量基础数据, 综合反映一个政区(国家、省、市、县)范围内的基础地理数据(地形地势、河流、湖泊等)、自然资源、经济和社会发展等信息, 其目的是为政府高层领导的宏观决策分析提供客观依据或进行辅助决策, 同时也是建设“数字地球”的必要信息积累。综合省情地理信息系统的建设, 地理数据库是其中的关键, 数据库中的数据是海量的, 如何组织这些数据, 进行有效的管理并投入应用, 是建立综合省情 GIS 数据库面临的挑战。

## 1 特点、内容

客观地掌握综合省情地理信息系统的要求, 准确地抓住其特点, 清楚地理解其要表达的内容, 是设计和建立地理数据库的前提和基础。

### 1.1 系统总体要求

(1) 系统必须体现出客观、综合的特点; (2) 在为领导提供全面、综合信息的前提下, 还必须能够适度反映某个地区或县(市)的详细情况, 以便领导能够进行从粗到细的信息查询; (3) 系统数据必须符合有关国家或行业标准, 符合信息共享和信息安全的要求, 为“数字省”甚至“数字中国”的建立作准备。

### 1.2 系统特点

系统的特别要求决定了系统数据库要有良好的结构框架, 它除具备一般 GIS 的基本特征外, 还具有以下特点:

收稿日期: 2000-12-15; 修订日期: 2001-04-08

作者简介: 马荣华(1972-), 男, 山东临沂人, 博士生, 主要从事地理信息系统设计与开发研究。

(1) 多尺度特征 综合省情 GIS 的特殊服务对象要求系统既能宏观地反映全省概况, 又能微观地反映个别地区较为详细的情况, 这需要多尺度空间数据的支持。

(2) 时态特征 政府领导不仅关心现实情况, 历史状况也是其关心的重要内容; 另外, 系统的预测分析功能也需要多时相数据的支持。

(3) 跨区域特征 省级范围很大, 以江苏为例, 全省 1: 100 万地形图有 4 幅, 1: 25 万有 17 幅。系统涉及地 (市) 县多级行政单元、众多图幅, 跨区域特征明显。

(4) 海量数据共享、最小冗余度特征 海量数据是指巨量的空间数据。“数字地球”要求空间数据共享, 共享可使其具有更广阔的应用前景。数据共享, 可为以后综合县 (市) 情空间数据库的建设打下坚实的基础, 从而进行综合省情和综合县 (市) 情之间的双向查询。为提高综合省情 GIS 空间数据库的运行效率, 应保证数据库的冗余度最小。

(5) 按行政单元进行查询检索的特征 综合省情反映的是一个省的综合状况, 是对本省各县 (市) 情况的综合, 有时为了详细了解有关情况, 需要查阅各县 (市) 的具体情况。因此, 数据库中的数据组织必须利于按行政单元进行查询检索。

(6) 数据安全性特征 综合省情 GIS 数据库的主要服务对象是省政府领导机关, 供领导直观地掌握全省概况, 起到辅助决策的作用, 部分数据属于机密。因此, 安全性是本系统需要考虑的重要问题。

### 1.3 内容

综合省情 GIS 地理数据库是一个多层次、综合性的空间信息数据库, 是系统的骨干和心脏。一个标准的能够为政府宏观决策分析提供客观依据或进行辅助决策的综合省情 GIS 数据库, 应该包括基础地理数据、国土信息数据、社会经济数据、资源环境数据、专题地图数据、政务数据以及卫星影像数据和 DEM 数据等。这些数据有些是空间 (地理) 数据, 有些是属性数据, 本文重点探讨前者。综合省 GIS 地理数据库由矢量数据库、影像数据库和 DEM 数据库三部分组成 (图 1)。

## 2 数据组织

综合省情 GIS 地理数据库具有多比例尺、多时段、利于按行政单元进行查询检索并且按行政单元上下隶属关系相互兼容的特点, 数据组织必须能够反映上述特点。

### 2.1 矢量数据库的数据组织

2.1.1 传统方法 受分幅纸质地图的影响, 地理信息也基本上采用分幅表达的形式。但基于这种分幅表达的传统 GIS 地理数据库的组织存在如下弱点: (1) 地理实体的完整性和一致性难以维护; (2) 不便于数据库的分布式管理; (3) 分幅管理对于数据共享和地理实体一级的安全性、一致性管理增加了难度; (4) 专题信息的分离管理不利于各种地理数据的有效获取。

为了克服上述弱点, 许多学者和 GIS 软件公司进行了大量的研究。李爱勤等 (1998) 对众多解决方案进行了分析, 认为它们大多采用的是一种逻辑上的无缝组织方案; 为此提出了一种逻辑和物理概念上真正的地理数据无缝组织模型, 这种数据组织方法虽然突破了传统图幅分块的概念, 但并不完全符合综合省情 GIS 按行政单元进行查询检索的特点。

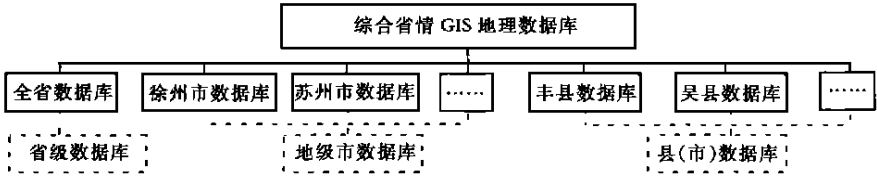


图 1 基于行政单元划分的数据组织

Fig. 1 Data organization based on administrative division

2. 1. 2 基于行政单元划分的无缝空间数据组织 按照行政单元进行快速查询检索是综合省情 GIS 地理数据库的重要特征之一。为此，我们采用了基于行政单元划分的无缝空间数据组织方案。行政单元的划分以县为最小单位，数据库是平行存放的（图 1，以江苏省为例），查询浏览时按行政单元上下隶属关系进行编码索引，实行编码索引元数据管理，索引结构的部分字段如下：

数据库名称	路径	上一级行政单元数据库名称	路径	下一级行政单元数据库名称	路径
-------	----	--------------	----	--------------	----

(1) 无缝图层的建立 综合省情 GIS 地理数据库的首要工作任务就是建立标准的分类分级和代码体系，用相同的分类分级标准确定空间地理要素的类别和级别，再根据地理要素的类别和级别产生具有唯一性的代码体系，然后参考建库比例尺、内容、性质等因素将地面沿坐标轴方向分割成若干层正交格网，层与层间格网大小不同，仔细选择相差倍数；此处格网的划分并非对空间的实际分割，只是为空间索引和数据组织服务。与地图分幅不同，无缝地理数据库中不存在分幅的概念，地物入库时，必须进行几何和逻辑接边（即属性数据的合并、统一），同一地物在数据库中合成为一个完整的实体对象，该对象记录地物的完整空间信息和属性数据。同一地物在相邻图幅拼接时，在各自的图幅中，系统分配给它一个唯一的标识号，按照事先建立的分级分类体系，用户又分配给它一个用户标识，这样在进行几何和逻辑接边时，两个图幅的系统标识号进行更改形成整个数据库中系统标识号唯一的完整无缝地物，但用户标识仍沿用原来的标识号。图 2 表示了无缝图层的建立过程。

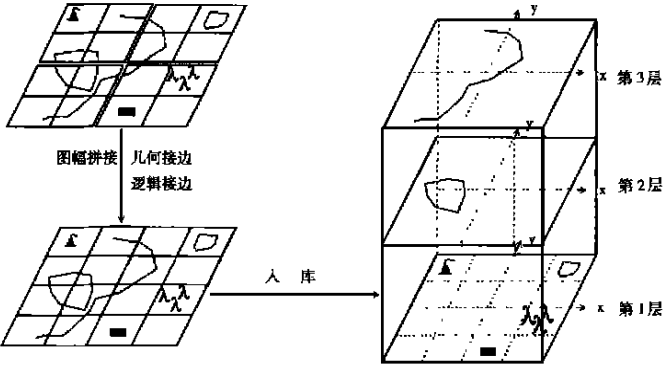


图 2 综合省情 GIS 地理数据库无缝数据组织

Fig. 2 Spatial represent for seamless geographical database in general situation GIS of one province

任一地物在整个数据库中是唯一的，而且它在所在层中最多只能跨越或覆盖四个格网大小，否则就得被提到上层或更高层中。这样，在每一格网中记录它所在范围内相应区域的地物系统标识码和用户标识

码，再对规则格网进行空间编码，就能建立起无缝地理数据库的空间索引。实际上，上述方法吸收和利用了面向对象技术和方法。

(2) 行政单元的划分 分为省级、地市级、县（市）级三个等级，分别建立无缝图层数据库。

(3) 多比例尺的处理 一般多比例尺的表现形式有两种：一种是一库多版本，即建立一个较大比例尺数据库，从中提取其他层次比例尺的数据；另外一种是多库多版本、独立对应于多级比例尺（或分辨率）的多个空间数据库。这种在 GIS 数据库中多比例尺共存的原因，是 GIS 本身的特性和目前自动制图综合技术还不够完善造成的。如果自动制图综合技术完善，可以建立一种较大比例尺的数据库，而系统对小比例尺数据的需求完全可以从大比例尺数据库中综合产生，这样无论是从系统的安全性、完整性，还是从减少系统数据的冗余方面都无疑是一种进步，遗憾的是目前地图自动综合还存在诸多问题，还不能提供这种快速准确有效的制图综合方法。即使已经建立了一种较大比例尺的空间数据库，可以在这个基础上综合产生其它层次的数据库，但为了满足对数据库快速的查询检索和应用需求，也还必须在系统中存在各种其他层次比例尺的空间数据库。另外，综合省情中也需要适当反映各县（市）的情况，由于行政区域的大小、地形地势不同，所反映信息的详细程度也不同，这也决定了不可能建立单一比例尺数据库，因此比较可行的方法是建立多比例尺数据库。同一专题（如土地利用现状）在不同的地区或按需反映不同的详细程度可能需要多种比例尺的支持，此时按比例尺分别建库，各比例尺数据库之间是平行的（图 3），数据库之间的关系通过关键重要地物的唯一编码来连接。

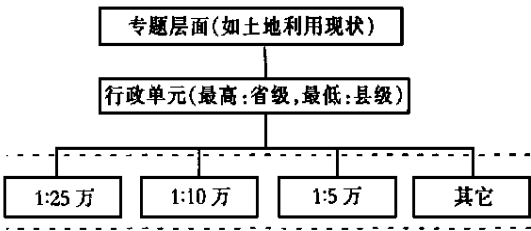


图 3 数据库比例尺的处理  
Fig. 3 Process to scale in database

(4) 时态性的处理 综合省情 GIS 地理数据库中的数据应该是随时间变化动态更新的，也只有实时动态更新，才能真实、准确地反映一个省的当前状况和历史状况。一方面，从客观实体发展的角度来看，其变化是在某一时间段内发生的，表现为一个过程，一种状态；另一方面，从客观实体变化的最终结果来看，其变化是在某一点上发生的，表现为一个时刻，一个事件。因此，时态数据库可分为基于状态的数据库和基于事件的数据库。

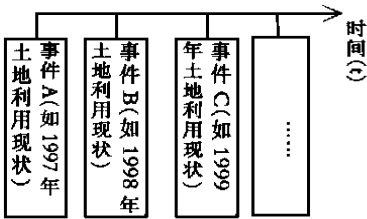


图 4 基于事件的数据库时间表示  
Fig. 4 Process to temporal in database



图 5 数据库时态性的处理  
Fig. 5 Process to temporal in database

综合的特点决定了数据库的时间量子应

该是年或月，时间区间较长；另外，根据系统所涉及到的各行业的性质，我们认为系统数据库结构采用基于事件的时间表达方法是较为适合的。这样，单从某一事件（某一时刻）看，数据是静态不变的，如 1997 年的土地利用现状；但从某一专题，从某类事件的连续性来看，数据又是动态的，如土地利用在 1997 年、1998 年等的现状，如图 4。目前所使用的数据模型大致有 5 种，各有优缺点（表 1），考虑到系统综合的特点以及前述分析，数据库对矢量数据和栅格数据的存储采用一种改进的快照模型，时态数据的总体组织如图 5 所示。

表 1 几种时空数据模型的简单比较

Tab 1 General comparison among five spatial-temporal data models

	矢量快照模型	矢量更新模型	混合模型	4D 模型	一体化模型
利用现有 GIS 建立	好	中	差	差	中
实现	好	差	中	中	差
数据输入	好	好	中	差	好
数据编辑	好	好	中	差	中
时态查询分析	差	中	中	好	好
数据可视化	差	中	中	好	中
便于利用	好	中	中	差	差

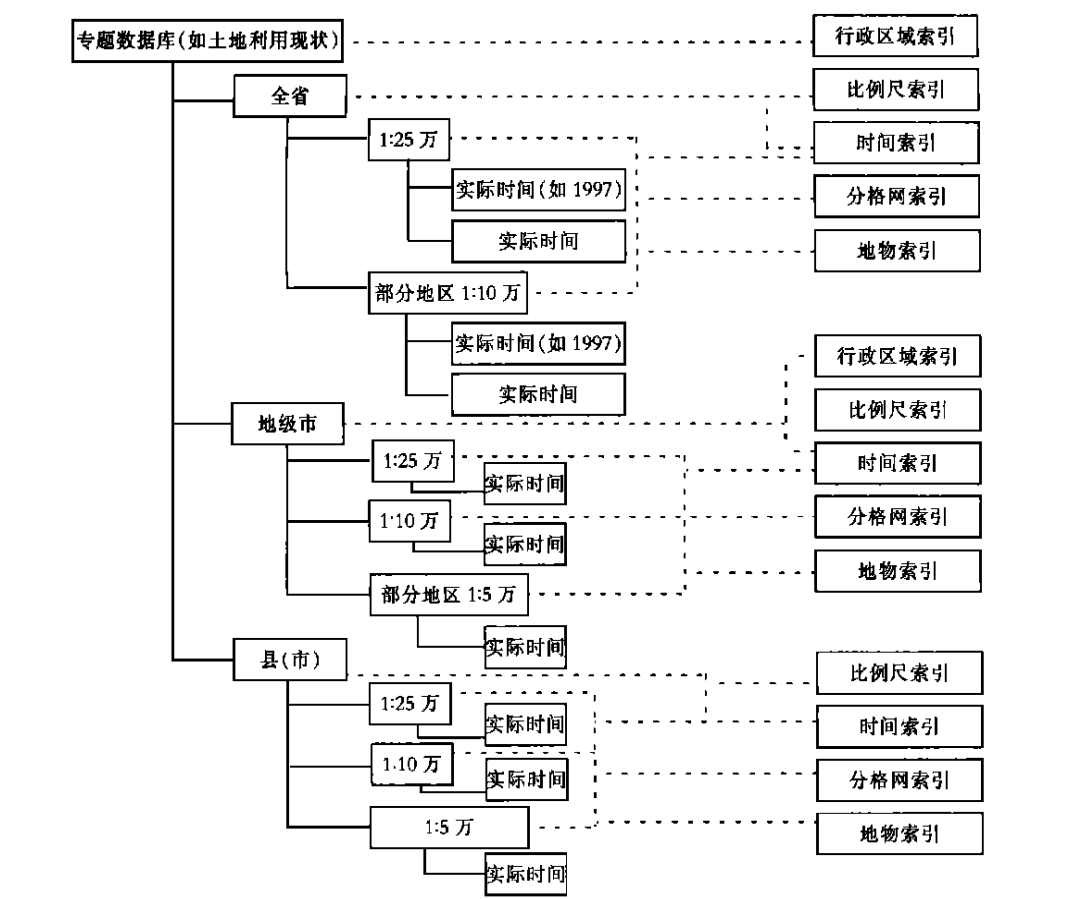


图 6 综合省情 GIS 地理数据库的整体组织

Fig. 6 Integrated organization of geographical database in general situation GIS of one province

(5) 数据库的整体组织 为了便于管理和索引，实行数据入库登记管理，即在数据库的顶层建立一张数据库管理索引表，表的部分字段结构如下：

表名称	存放路径	比例尺	实际时间	数据库时间	所在政区名称	所在数据库
-----	------	-----	------	-------	--------	-------

2.2 影像数据库的数据组织

建立卫星影像数据库的目的之一是为了更清楚、更真实地反映地表信息的状况，但需要矢量图形的辅助。因此在组织影像数据库时要考虑矢量数据和栅格数据的一体化问题。

卫星影像入库前的重要工作是影像数据和矢量数据之间的匹配，包括统一投影、比例尺匹配和坐标匹配。另外，还需把各级比例尺的卫星影像按行政区域进行裁剪，分别存放，以加快按行政区域进行查询浏览时影像显示的速度。一幅影像的数据量非常大，特别是一幅全省的卫星影像，实时调度有一定的困难。解决的办法就是分块存放（如  $512 \times 512$ ），建立索引记录的指针。这样，在影像漫游时，根据空间位置索引到指针，直接指向并调用数据块。另外，当比例尺缩小时，需要看到更抽象的影像，如果从底层调用数据库再抽取，速度太慢，解决的办法就是建立影像金字塔，根据不同的显示比例，调用不同金字塔层次上的数据。数据组织方式如图 7。

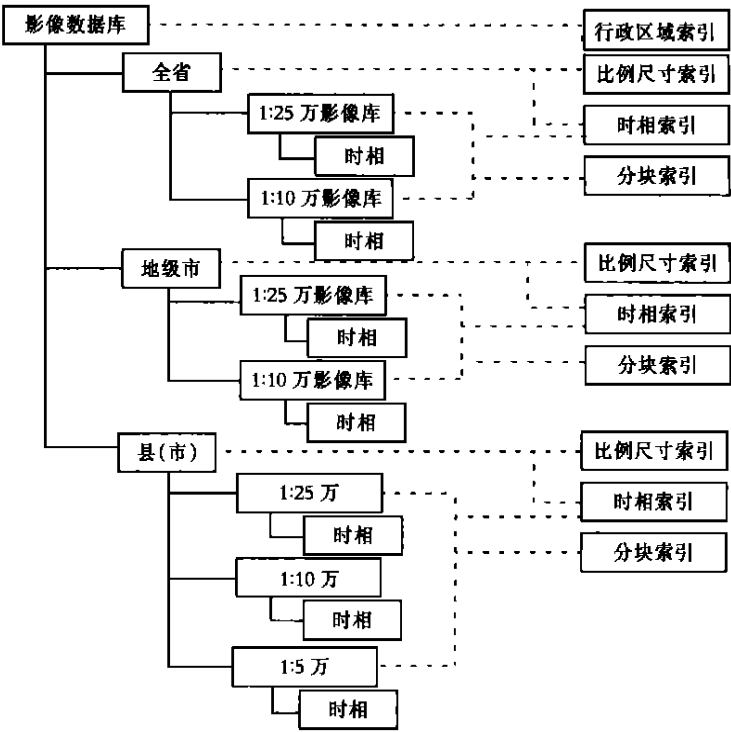


图 7 综使省情 GIS 影像数据库的数据组织

Fig. 7 Data organization of image database in situation GIS of one province

2.3 DEM 数据库的数据组织

DEM 数据库的存放同样采用金字塔式的数据结构，按行政区域和比例尺存放，这样可以将所有相关的数据组织起来，可以调度数据库中任意范围的数据，达到对某行政区域

内整个地形的无缝漫游。数据组织如图 8 所示。

### 3 结语

在综合省情 GIS 数据库的建设过程中, 我们认识到: (1) 大型 GIS 数据库的组织必须以“数字地球”为框架, 实现信息的共享和地球表层的三维恢复。只有这样才能提高数据的应用效率, 并最大限度地挖掘数据库的潜在信息。(2) 本文介绍的数据组织方法, 也非常适用于综合市情、县情等政府

GIS 数据库的数据组织。它打破了图幅的界限, 采用了基于行政单元的无缝空间数据组织方案, 将用户的注意力转移到行政区域单元的查询、检索、浏览与分析上, 能更好地为领导决策服务。(3) 由于数据按行政单元(省、市、县)和实际获取时间分别建库, 对时态数据的处理采用了快照模型, 数据库中存在一定的数据冗余, 但换取了系统查询、检索和浏览的速度, 对为政府领导服务的地理信息系统来说是值得的。

### 参考文献:

- [1] 陈述彭, 曾杉. 地球系统科学与地球信息科学[J]. 地理研究. 1996, 15(2): 1~ 11.
- [2] 郭建忠. 地理信息系统中比例尺的应用[J]. 解放军测绘学院学报. 1997, 14(3): 196~ 200.
- [3] 李爱勤, 龚建雅, 李德仁等. 大型 GIS 地理数据库的无缝组织[J]. 武汉测绘科技大学学报. 1998, 23(1): 57~ 61.
- [4] 王桥, 毛锋, 吴纪桃. GIS 中的地理信息综合[J]. 遥感学报. 1998, 2(2): 155~ 160.
- [5] Raper J., Livingstone D. Development of a geomorphologic spatial model using object-oriented design[J]. INT. J. Geographical Information System. 1995, 9(4): 359~ 384.
- [6] Worboys M F. A unified model for spatial and temporal information[J]. The Computer Journal. 1994, 37(1): 26~ 34.
- [7] Gail Langran. A review of temporal database research and its use in GIS applications[J]. INT. J. Geographical Information System. 1989, 3(3): 215~ 232.
- [8] Peuquet D J, Duan N. An event-based spatio-temporal data model (ESTDM) for temporal analysis of geographic data[J]. INT. J. Geographical Information System. 1995, 9(1): 7~ 24.
- [9] Kelmelis J. Time and space in geographic information: toward a four-dimensional spatio-temporal data model[D]. Unpublished Ph.D dissertation. Department of Geography, The Pennsylvania State University, 1991.
- [10] 宋广慧. 时空地理信息系统初步研究与实践[D]. 硕士学位论文. 南京: 南京大学, 2000.
- [11] 杜云艳等. 实用化的 GIS 数据库设计方法[J]. 地理研究. 1998, 17(增刊): 91~ 95.

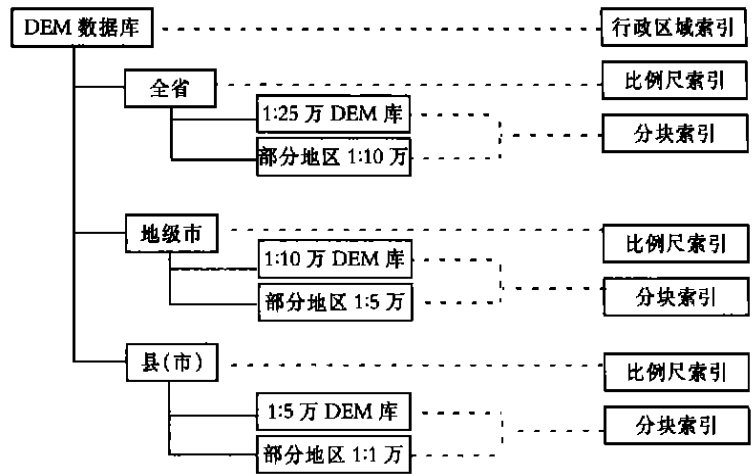


图 8 综合省情 GIS DEM 数据库的数据组织

Fig. 8 Data organization of DEM database in general situation GIS of one province

## Data organization of geographical database in general provincial situation GIS

MA Rong-hua, HUANG Xing-yuan, ZHAO Zhen-bin

(Dept. of Urban & Resources Science, Nanjing University, Nanjing 210093, China)

**Abstract:** GIS application in government agencies opens up a specialized research field and the General Provincial Situation GIS is one of the important research contents. Organization and management of billions of data with multi-scaled and multi-temporal characteristics, composed of vector data, image data and DEM data, is the key issue in giving the field a push to develop rapidly and be applied practically. And the system must make the strict demand on security of data. In order to provide consultation service for leaders to immediately make a strategic decision, the data organization in this paper breaks away from the constraint of the traditional map management, which has some evidently shortcomings that are unfavorable for querying and indexing information of some administrative units. As a result, the seamless spatial data organization scheme based on administrative units is put forward to organize the vector data and the snapshot model is used to process the temporal data. So all data are stored according to administrative units. And data of different scales are stored in different databases and the image data and DEM data are stored separately in line with blocks. Consequently, querying and indexing to all data can be done according to different administrative units, different scales, time (temporal resolution), different object classifications, image or DEM blocks, etc. and the speed to query and index is improved distinctly.

**Key words:** GIS; database; data organization; querying and indexing

---

### 会议讯息

中国科学院地理科学与资源研究所和德国 Essen 大学将于 2002 年 3 月底至 4 月初在北京联合召开“Urban Climate and Air Pollution Investigations in Beijing with special regard to Urban Planning, UCAPOBEIJ”国际学术讨论会。会议主要论题有:

- 1、城市气候
- 2、城市大气污染
- 3、城市气候因素与城市规划

会议语言为英语, 将不收取注册费, 并提供伙食; 交通、住宿费用自理。请欲参加会议者与董云社、齐玉春联系, 并准备英文论文和摘要, 会议将根据论文情况选择大会发言。

地址: 北京市安外大屯路 917 大楼 中国科学院地理科学与资源研究所

邮编: 100101; 电话: 010- 64856500; E-mail: dongys@igsnrr.ac.cn; qiy@c@igsnrr.ac.cn