

地震区划数据管理方法的研究

陈冠直 冯录刚

(河北省地震局)

本文阐述应用计算机图形学原理和工程数据库技术对地震区划数据的管理方法的研究。通过对区划数据的分类和存贮方法的确定,提出地震区划数据的存贮结构模型和区划数据库系统的概念化模型。对图形数据作分层管理,以块(Block)为描述区划图形数据的基本单位,在逻辑结构上采用固定页长的链页方式和双向链表结构,既提高了运行速度,也解决了数据库再组织问题。用常规数据库方法管理文本信息,并可自动生成指向图形实体的关键字,解决了图形信息和文本信息相互检索问题。文中还叙述了对区划数据存贮管理的功能设计,并介绍了在微机上的实验系统“地震区划数据库管理系统”的设计与实现。

一、引言

地震区划数据管理是编制新的地震区划图的一项研究课题*,它主要是应用计算机图形学原理和工程数据库技术研究区划基础资料在计算机内部的存贮结构和检索方法,以及有关图形和文本数据的输入和输出问题,并在微机上设计实现了一个实验系统——地震区划数据管理系统,它具有以下功能:

- 定义图形实体:块(Block),图形类型和图件;
- 将由数字化仪采集的图形数据文件经过座标转换处理装入图形数据库;
- 图形实体的修改、删除和合并;
- 对图中文字的插入、修改和删涂;
- 将图形实体生成图形在屏幕上显示或在绘图仪上绘出;
- 建立文本数据库模式,对文本数据装入,修改,删除和检索;
- 由图形数据检索出图形实体的有关文本参数;
- 检索文本数据生成相应的图形输出并可产生图件;
- 产生具有可选择的经纬度网格线,国境、省界、城市、河流、震中分布图等多种工作底图;
- 对图形和文本数据库作卸出脱机保存,以及后援文件装入相应数据库。

目前这个实验性系统已提供试用,装入了河北省部分区划基础资料。

*本课题的研究得到金学中,金严二位同志的关心,作者在此表示感谢。

二、区划数据的分类与存贮方法的确定

要研究区划数据的存贮管理方法,首先要分析区划数据的组成。在编制新的中国地震区划图的工作中,需要用到大量的基础资料数据,主要包括地震地质资料,地震资料,地球物理场资料和地震区地震带划分资料等。这此数据不仅种类繁多格式较复杂,而且目前尚无统一的技术规范。以便于计算机处理的角度,基本上可以分成下面两种:

一种是图形型数据,包括各种图,如新构造图,综合等震线图等等。构成图的可以是断裂、盆地或其它基本元素。这些基本元素是可以由点线(弧)等图素描述的二维数据,表示它们在二维空间存在的位置、形状等。这些特征作为图形表示准确而简单,但是我们用一般文字或数字来描述则很困难,因此我们称这种数据为图形型数据。

另一种数据则与前者相反,完全可以用数字或文字描述,如一条断裂的年代、名称甚至一张图的文字说明等等都属于这种数据。还有一些数据表现为图的形式,却很容易由通常文字或数字来描述,并且根据这些描述很容易转换成一张图,如震中分布图可以用地震目录来描述。对于这些数据我们广义地称之为文本型数据。

基于上述对区划数据的分类,不同的数据可作不同的处理。对于文本型数据采用通常的数据库技术作存贮管理,基本上可以完成。所谓数据库可视为存贮在一起的相关数据的集合,这些数据无有害的或不必要的冗余,数据的存贮独立于使用它的程序,对数据库插入新数据修改和检索原有数据均能按一种公用的和可控制的方法进行,数据被结构化,为今后的应用研究提供基础。在我们的区划数据中存在结构上完全公开的若干文本数据库,也就是说我们的系统包含一个数据库集合。在处理文本型数据中与通常数据库技术有所不同的是还需要解决它们与相应的图形型数据的联系问题,以便完成由图形型数据检索文本型数据和文本型数据检索图形型数据的操作。

对于图形型数据只是简单地应用通常的数据库技术就很难完成了。其主要原因是通常的数据库不适应图形型数据结构,一般二维图形型数据通常采用的数据结构有栅格结构和矢量结构,栅格结构是一种按二维正交坐标系划分的矩阵形的数据结构,这种数据结构适用于表示各种面状分布的数据,容易实现按位置选择,在图形处理时应用较为广泛。其中主要问题是数据量大,而且数据冗余度高。矢量结构一般用于描述线状分布的数据,也可描述面积型的数据,适合于图形数字化仪实现的数据采集。在这种数据结构中,所有数据组均以结点为界分割为弧,由弧联接成多边形,而弧则是一系列X,Y二维坐标所定义的线段。目前区划基础数据的图形型数据的数据来源主要是从已有的各种图件上采集。而像等深线以及地球物理场中的各种等值线本来很适宜采用栅格结构,但却缺少原始数据也只能从有关图件上采集,因此我们采用矢量结构为图形型数据的数据结构。

在存贮管理中图形型数据的基本单位即最小可划分的图形实体应该是什么呢?考虑到区划基础资料的特点,也为尽量减小系统开销,我们确定为块(Block),块是组合

成复杂图形的一组因素，一个块可以是一条断裂，一个盆地或一组温泉等等。而构成块的点线多边形等因素对用户来讲是透明的，即不提供管理功能，因为对区划数据来讲，管理到块这一级已经足够了。

三、区划数据的存贮结构模型

依据上述对区划数据的分析，可以得到如图1所示的区划数据的组成。其中文本数据中的每一个类型对应一个数据库模式，如断层参数数据库，盆地参数数据库等等，下一层的记录则成为相应数据库中的内容。这里采用关系型数据库模式，每一种类型可以定义成为一个关系模式，一个记录对应关系的一个元组，如对断层类型可视为一个关系，断层的每一种参数可对应应该关系的一个属性，而每一个具体的断层则是这个关系的一个元组。有若干种类型则建立若干个关系数据库，全部文本型数据可看作数据库的集合。

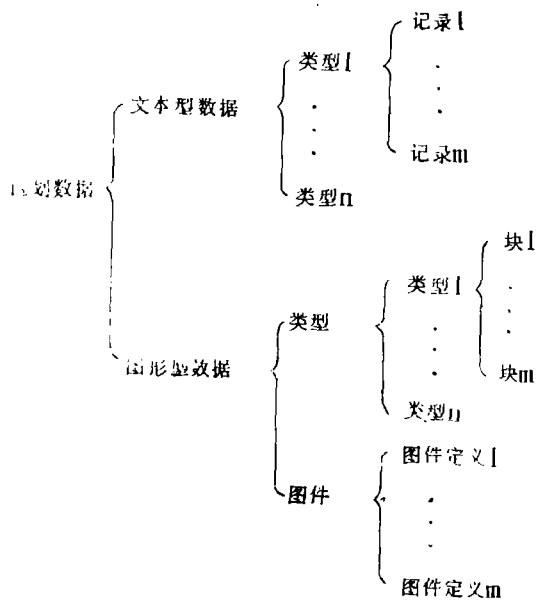


图1 区划数据的组成形式

图形型数据分为三个层次：图件、类型和块。关于块的定义前面已有描述，类型的含义与文本型数据中的类型基本相同，只是构成这个类型的不是文本数据记录，而是图形实体——块。一个块的数据有两部分，一部分（通常是大部分）是二维数据，另一部分是辅助数据如块名称、代码、经纬度范围、线型，颜色等等。图件的数据包括它描述这张图的有关参数，如图件名称、经纬度范围、座标转换参数、工作底图要求，以及构成该图件的所有图形实体类型名和块代码等等。每一个类型对应一个图形数据库文件，即具有相同类型的块的二维数据均存放在同一个图形数据库中。而图件的数据相对于块的数据量要小得多，则将图件数据集中存放在一个库文件内。

在图形数据中还有一种数据即图中的文字，如一个城市在图中由一个圆来表示，在它旁边还要注释这个城市的名字。这些文字可以出现在图形实体——块中，但却与块中的其它二维数据有所不同，因此要另行处理。在块的数据中只保留指针，对文字值及有关参数数据存放在一个专用的库文件中。

为了解决文本型数据和图形型数据的联系问题，在需要与图形数据连接的文本数据库模式中自动增加一个关键字属性，这个关键字值在元组中与相应的图形实体——块的

代码相对应，其作用类似于网状数据库中键的作用。

经过这一系列的处理，区划数据的存贮结构模型可由图2表示。

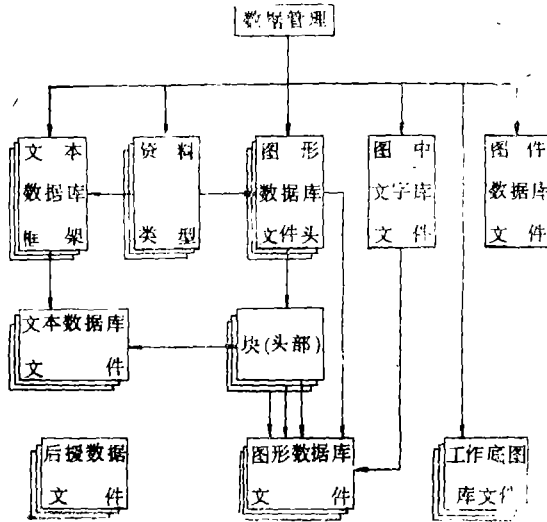


图2 区划数据的存贮结构模型

四、区划数据存贮管理的功能设计

根据用户的需求和实况，这个实验系统的计算机环境条件，对区划数据的存贮管理设计了以下几个方面的功能：

1、图形实体——块的建立。包括对块的定义和数据输入。对块的定义要求给出块的类型和名称，如该块属于断层类型，叫做太行山前断层。还要给出块成图时的线型和颜色，以及一些有关参数。数据输入是指对数字化仪从原始图件上采集来的该块的数据文件的读入，在系统找到该数据文件后，要求用户给出有关坐标转换的参数值，于是系统将数字化仪采集来的直角坐标的二维数据处理成地理坐标的二维数据存入相应的图形数据库中。

2、图形实体的数据检索。包括对指定块的数据列表输出；根据块的类型检索列表；检索指定经纬度范围的块的数据；删除一个块；修改一个块的有关参数；合并两个块的数据；将指定块的数据拷贝到一个ASCII码文件等功能。其中合并两个块这一功能是拼图工作设计的，如一个断层横贯两个省，各自成图，数据输入时形成两个块，这对检索和进一步生成新图和应用都不方便，而利用合并两个块这一功能则可解决这个问题。向ASCII码文件拷贝一个块的数据，是为了满足用户进一步对图形数据分析处理的需要，因为块数据在图形数据库中的内部形式对用户是透明的。

3、对图形实体——图件一级的管理。包括建立图件；修改图件；删除图件；图件数据列表等功能。其中建立图件要求给出图件名称、经纬度范围、中央经线、双标准纬度及比例尺等坐标转换参数、图件标题、图件文字说明等等；有关工作底图的数据如是否

需要经纬度网标线、国境、省界、河流、城市、震中分布图等；还要给出构成图件的所有图形实体——块的类型和代码等。

4、图中文字的管理。对图中的文字除了在块数据的采集和输入时可存入外，还提供了对图中文字的插入、删除、修改和查找等功能。插入操作要求首先给出图中文字所在块的类型名和块代码，然后再给出图中文字值及插入点的位置及字符高度和倾斜角度。

5、图形生成。包括对指定图形实体——块成图；对指定经纬度范围内的图形实体——块成图；对指定图件成图。前两者均要求用户给出有关座标转换参数，以便成图过程中的处理，对指定经纬度范围内的成图过程中，还可以根据用户的要求形成一个图件。

6、文本数据管理。包括定义文本数据库；文本数据的输入；修改一个数据记录；删除一个数据记录；文本数据库的检索；指定文本数据列表；显示文本数据库结构等。其中定义数据库时，系统提问是否要求与图形数据连接，若需要连接时，则将块的代码作为该数据库的关键字。

7、综合检索。包括由文本数据库检索成图；由图形数据检索文本数据库。前者根据对文本数据库的检索条件查找有关的图形实体——块并将这些块生成图形；后者根据图形实体的特征，先在给出的经纬度范围内查找出符合条件的图形实体，再对相应文本数据库查询有关的文本数据。另外，还提供了图形数据库和文本数据库的目录显示。

8、后援文件管理。包括指定图形数据库文件的卸出和装入；对指定文本数据库文件的卸出和装入；以及类型删除等。所谓卸出就是将数据库的结构和数据转贮到后援文件脱机存贮，反之所谓装入则将后援文件中的数据库结构和数据传输到系统中联机应用。这一点在微机的外存容量小的情况，尤为重要。

除了上述八个方面的主要功能外，还具有对图形数据的数字化采集和图形编辑功能以及系统初始化等等。

五、区划数据库管理系统的设计与实现

根据已确定的系统功能设计区划数据库的管理系统，首先要有一个概念化模型（图3）。由该模型可以知道，设计和实现这个区划数据库的管理系统，主要是区划数

据结构模型、区划数据管理程序、专用图形系统三个大部分。其中专用图形系统借用目前广为流行的AUTO CAD图形软件代行本系统的专用图形系统的职能。关于区划数据的存贮结构模型前面已作了描述，这里仅就有关图形数据结构作一点补充说明。在图形数据处理中，作为图形实体的块是描述区划图形数据的基本单位，而这些块本身由最基本的图素点、线

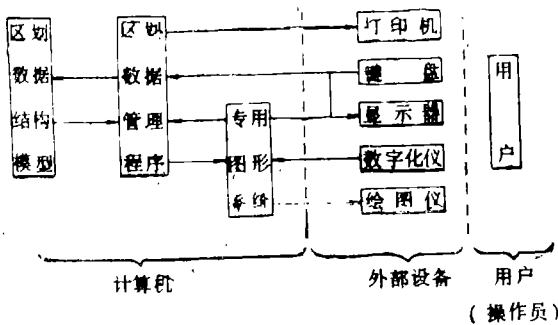


图3 区划数据库系统概念化模型

段、文字指针等组成。在一个块的图形数据中这些图素数量很大，形成了一条链，而且链的长度依这个块所描述的图形繁简不等，这不仅给链在内存中的处理带来困难，还会因访盘次数增加使处理时间很大。本系统为解决此问题，采用了固定长度的链页方式和双向链表结构（图4）。既减少了访盘次数，也为回收被删除的数据空间（废弃的链页及链表页）提供了方便，从而解决了数据库的再组织问题。

关于管理系统的程序设计，根据前面叙述的八个方面功能各自作为一个程序模块，另外还有系统初始化和总控制两个模块。其中系统初始化模块主要用于本系统第一次运行之前的系统安装任务，主要是建立必要的基本数据结构。有关数字化仪的数据采集和图形编辑则留作专用图形系统来完成，这里是由AUTO CAD软件包来完成的。构成本系统的十个程序模块均用FORTRAN-77语言书写，总程序量4000余行。考虑到本系统的用户均非计算机专业，因此采用多级菜单驱动提供功能选择，必要环节给出提示根据用户回答进行处理。在CCDOS下运行，菜单和提示均已汉化。限于本文的篇幅，对各程序模块的工作流程不作叙述，只从程序结构角度作一简单说明。本系统程序在研制时做

到模块化结构化，除主控模块为主程序模块，其余模块均以子程序形式出现，各模块又可分为若干个子模块。除一部分公共子程序可提供各模块调用外，大部分子模块在统一的公共数据结构基础上具有一定的独立性，这不仅对程序调试和维护有益，而且为系统功能的进一步扩充提供了方便。另外系统对用户

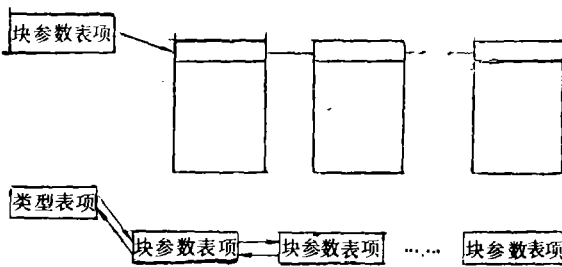


图4 链页及双向链表结构

键盘输入和图形数据输入转换等都作了检验，起到加强保护数据库的作用。还有系统的后援处理也提高了系统的安全性和可靠性。

六、结 论

经过对区划数据性质的分析，建立区划数据的存贮结构模型，对存贮管理系统的功能设计，直到在微机上实现“地震区划数据管理系统”，研究了区划数据的计算机存贮管理和检索方法。认为必须以计算机图形学原理和工程数据库技术处理区划数据的存贮管理，用矢量结构描述区划数据中的图形型数据，用块、类型、图件三个层次描述图形实体，用链页和双向链表等数据结构作为图形数据库文件组织形式，以及由通常数据库方式处理区划数据中的文本型数据的方法均是可行的，在微机上实现的实验系已具有一定的实用价值。

(1988年1月12日收到初稿)