

# 中华人民共和国行政区划单元 数据库的建立和应用

刘岳 梁启章 金学英

(中国科学院地理研究所)

## 提 要

中华人民共和国行政区划单元数据库是将全国2300个以上的县和县以上的行政区划的编码和分布位置进行存储、管理并提供应用它的一组程序。本文对数据库的建立方法和应用作简要介绍。

地理信息系统的核心部分是要建立一个具有统一地理基础,并对各种地理要素的空间信息进行了标准化和规范化的地理信息数据库。为了适应当前四化建设的需要,首先开展建立我国县界数据库的工作。经过应用,表明县界数据库可在全国性或区域性的多种分析与制图工作中发挥作用,本文对数据库的建立方法和应用作简要介绍。

## 一、数据库内容基本描述

我国县界数据库既是有关我国2300多个县的分布界线及其编码数字信息的集合,又包括应用它的技术方法。下面对数据库的目标设计、数据结构和文件组织、软件系统等加以描述。

### (一) 目标设计

县是我国的基本行政单位。县界的存储可组合成全国、大区、省、地区等多层次的地理基础,用于开展各种统计和地理区位分析、编绘不同类型的专题地图。为了满足多种应用需要,对数据库提出要求是:

- 1) 建立全面描述县级行政单位分布特征的各类数据文件。
- 2) 编码和数据格式标准化,以利为多种处理系统接受和调用。
- 3) 尽可能压缩数据存储量和合理组织文件,适用于小型计算机系统的处理能力。
- 4) 建立简便可行的管理系统,能灵活地实现数据检索和更新。
- 5) 配备基本的统计分析制图应用软件,便于用户利用。

### (二) 数据结构和文件组织

数据结构是数据在数据库中的组织形式。县界在二维平面上的图形是多边形,通常为非



对于上述四种数据集，在数据库物理存储中分别建立四种对应的文件。文件组织采用分区原则，即按省（自治区、市）分区，全国包括南海诸岛共划分三十个区域，共建立120个上述四种类型的文件。实验证实上述文件组织法可以充分利用RDOS操作系统资源并能在小型计算机功能下完成全国性大数据量的处理任务。由于数据库中采用了统一的文件命名原则，即每个文件名称用五位字符组成，其中前三位使用英文字母表示文件类型，后二位以数字表示全国省（市、自治区）的标准编号，便于记忆。在数据库总的描述文件中列出全部文件名称、文件的记录数、字符数以及物理存储地址等信息，便于用户查询。

### （三）软件系统

县界数据库的软件系统是在RDOS操作系统支持下建立的，这些软件保证有效地进行数据文件的组织管理，具有提供直接检索和更新数据的功能，而且包括了一组应用程序。软件系统约有30个程序构成，全部用FORTRAN IV语言完成。下面简述程序的设计思想和功能：

1）文件生成软件，从原始信息数字化到建立符合制图标准的数据文件需要经过一系列的编辑纠正处理，同时依据数据逻辑模式实现数据组织，生成上述四种类型的文件。这组程序的功能包括了原始数据输入、转换、建立磁盘文件，进行相邻省公共边界的传输，纠正边界数字化的几何误差，按着数据库要求，进行数据存储格式的变换，获取数据库的描述和检索信息。最终产生统一规格和标准，供系统检索和制图的“净化”文件和索引表。

2）更新维护软件，包含对数据库数据读写，输出出错信息，更新县界坐标和各级属性索引文件。还有数据结构再组织程序，对用户经常使用的数据项，可自己定义新的数据结构文件，它实际上是实现利用原有数据库文件扩展重新组成或提取，而生成新的文件。

3）检索软件，县界数据库的检索就其内容可分记录检索和文件检索；就其输出形式则有数据、图形和显示（CRT）三种。因此可提取县、地区、省、全国各级行政区域界线及其特征信息。检索程序可分为：弧段显示和打印输出程序，用户通过对文件名称的查寻有关文件可在CRT屏幕上和行式打印机上输出；多边形显示和打印输出程序，显示是按多边形单元逐一进行，其打印输出中还安排了错误检查功能；弧段绘图程序，是将检索数据转换为制图文件记录于磁带，在绘图机上绘出行政区划界线图；多边形绘图程序，具有弧段绘图程序相同的功能，在多边形文件绘图时县界均两次被重叠绘出。

4）统计分析与制图软件，为数据库应用程序，包括各种基本统计特征值计算，统计数据排序、画直方图、统计分级等程序。在统计制图程序中有绘制面状晕线表示的分级统计图，用不同结构图形符号表示的分区图表统计图和点描法表示的点数图，三角形法的类型结构图以及图幅注记，绘制图例等辅助性程序。

## 二、数据库建立和更新过程

建立全国县界数据库，合理完善的工艺是重要的技术保证。本工作的工艺过程见图3。

### （一）编辑准备

数据源的收集和分析是数据库建立的首要前提。本工作采用圆锥投影的全国县界分布

图, 复制在变形较小的聚脂薄膜上, 然后在县级行政区内标上统一编码或顺序编码, 县的特征编码在专门的表格中记录输入。在此原图上标明数字化坐标原点及纠正控制点等。

## (二) 数字化

进行县界数字化, 下述一些方法是行之有效的:

1) 根据文件数据结构, 以弧段为读取记录单元, 采用链式数据结构时必须读取链的首、尾结点坐标, 以及选取边界各特征点(折点曲率变化点), 保证数字化精确性。在结点多次数字化时其读取误差不得大于 0.5 毫米。

2) 要求以省为单元进行数字化且在相邻省的边界已经数字化时, 不再重复数字化。省内县界数字化应按编号顺序和顺时针方向进行。数字化可从任一弧段的结点开始, 但必须完成一个县的边界弧段数字化后, 再顺序进行下一个。数字化时弧段特征码记录输入必须将当前数字化的县的编码放在左边, 相邻县编码放在右边。

3) 数字化采用点记录方式。记录绝对坐标值, 并按规定的统一原点读取坐标。不论脱机和联机方式都要求实时核查, 发现错误, 立即改正或重新数字化。

经验表明, 上述方法适应了以省为单元组织文件, 而省界上进行一次数字化, 避免了节点匹配处理复杂性和减少数字化工量, 在脱机数字化方式中也便于检查和判别多种错误。

## (三) 编辑预处理

预处理是对原始数字化文件初步编辑, 发现和纠正数字化作业中的各种错误, 如边界特征码出错、弧段跟踪偏离了允许范围、线段遗漏、线段数字化重复等。系统的预处理是通过调用有关程序, 将原始数字化信息以省为单位输入系统和转换数据格式后, 恢复为图形形式, 首先由 CRT 放大显示和目视检查, 如发现错误再使用绘图机绘出图形, 并打印数据清单确定错误内容。

## (四) 县界坐标和弧段属性文件生成处理

建立县界数据库最先是进行县界坐标文件编辑, 而后进行弧段属性生成处理。它是通过调用经过编辑预处理的数字化文件, 来完成下列内容。

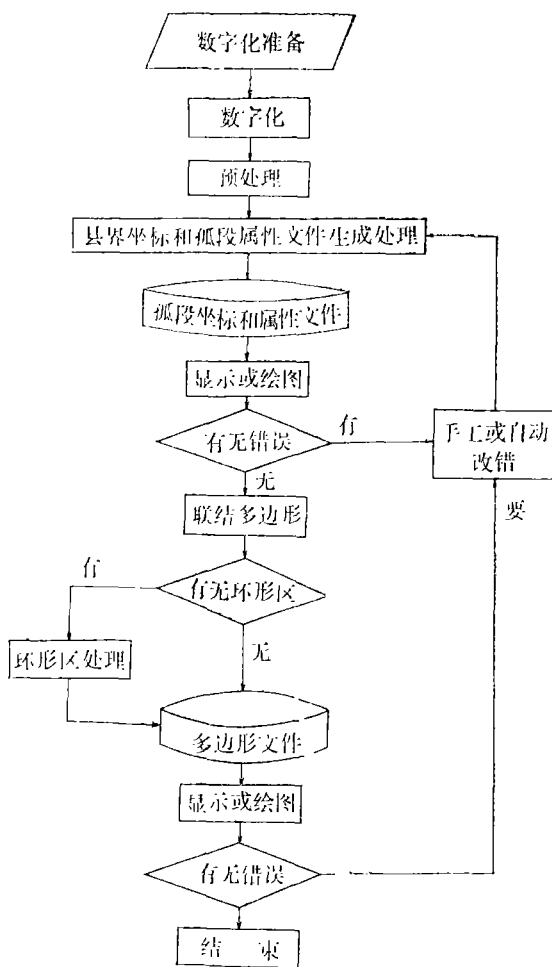


图 3 县界数据库建立过程

Fig. 3 Process establishment of county boundary data base

1) 公共省界传输, 在数字化过程中, 对公共省界可能存储在相邻省的数据文件中, 通过对相邻文件的查询将其公共边界坐标值读入, 并通过显示来检查。

2) 顶点匹配处理, 用于消除数字化中的读取误差, 最主要的是对弧段各结点由于多次数字化而产生不同坐标值的纠正。

3) 生成县的边界坐标文件, 同样以省为单位建立系统的统一命名的坐标文件, 该文件已是改正后的, 由县的边界XY坐标串组成。

4) 生成弧段属性文件, 如前所述, 按弧段文件数据结构和检索要求, 从生成的县的边界坐标文件中逐一提取并生成描述各弧段特征信息, 按规定格式写入磁盘; 并按系统规定的原则加以文件命名。

### (五) 多边形文件生成处理

多边形文件是在弧段文件基础上调用系统的软件完成以下几个主要方面的处理:

1) 多边形联结处理, 使用节点坐标将弧段联结成完整多边形的过程是通过弧段属性文件提取该多边形弧段记录号或全部弧段的坐标记录, 并通过节点判别和重新排序而实现。多边形文件生成过程中, 必须不断的进行编辑改正工作, 它是通过图形显示和打印错误信息指出下列错误: 1. 多边形不闭合; 2. 多边形闭合, 但未将该多边形全部弧段联结; 3. 联结了全部弧段多边形仍不闭合; 4. 该多边形编号不存在。对错误信息必须仔细核对有关输出清单, 通过修改弧段文件而重新进行多边形的联结处理。

2) 多边形环形区和“飞地”处理, 即在县范围内又出现了其他行政单元的占地(环形区)或同属于一个县的管辖而分别位于不联结的两个地区(飞地)的处理。环形区域或飞地的标志可以从原始数字化过程中加以说明, 或者通过多边形文件生成中发现。环形区或飞地的处理是由程序自动进行, 它是通过寻找两个内外多边形或相离的多边形间最短距离两对坐标值而加以联结, 并重新组织多边形的坐标串加以实现的。这个过程也需要通过两次显示进行目视检查。

3) 多边形属性文件生成。该文件是使用弧段属性文件或多边形坐标文件产生, 并按系统规定而命名。

为了适应我国行政区划的局部变动, 我国县界数据库的更新采用了下述工艺过程:

1) 修改县界图。首先将已经建立的县界文件绘于变形小的聚脂薄膜上同时在图上标明对应于数据文件的各县编码, 而后将局部更新的内容在原图上标出(以不同颜色区别), 这项工作是由人工完成的。

2) 局部界线数字化必须以同样格式和要求按弧段为记录单元进行, 并通过编辑处理而建立磁盘更新文件。

3) 县的边界坐标和弧段属性文件更新, 即调用原有的县界坐标和弧段属性文件, 提取有关记录信息并用局部更新内容加以代替(包括删除和增加), 更新后的文件同样需经过显示和输出而加以检查。

4) 多边形文件更新, 它是在县界坐标和弧段属性文件更新后, 使用程序生成更新的联结多边形的弧段号、多边形坐标记录, 然后将更新多边形记录写入(包括代替、删除和增加)原有的弧段联结多边形文件或多边形坐标文件中。而后修改相应多边形属性文件的部分

内容,最后由图形显示来检查。

### 三、数据库应用

县界数据库提供了使用计算机进行各种统计分析与制图所需要的行政区域的数字信息。应用县界数据库编制不同类型的专题地图的过程可简述如下:

1. 统计数据的存储和文件建立 各类数据(如人口、工农业、土地资源等调查统计数据)必须按数据库系统的要求组织存储。首先要以县为单元采集数据,每一个县的代码要和数据库中使用的国家标准一致,并以省为单位来建立文件。统计数据文件的存储结构可分为单要素和多要素两种方式,其格式可自由选择。建立文件应按五位字符的统一文件命名规则。

2. 统计数据特征分析 包括序列排队,计算平均值、绝对值、标准差、百分位数等特征值,打印指标分布直方图和统计单元指标频率分布直方图等。

3. 制图数据统计分级 这是编制专题地图时的重要环节。数据库系统的应用软件提供了多种分级方法,可根据统计数据特征分析和分布曲线的性质以及专业要求和制图表现的可能来选择恰当的分级方法。在执行这些分级程序时,采用了人机交互式的作业方法,以便充分发挥专业人员的决策作用。在确定了分级界限后生成制图用的分级文件,即每一个统计单元获得了应有的级别代码。

4. 分级代码的符号化处理 通常要调用多边形文件,将统计数据或者分级代码转化为多边形范围内的图形(如线、个体符号、点集、代码等),从而产生可用于输出图形的绘图文件,通常记录于磁带,供绘图系统使用。

5. 地图输出 由符号化处理得到的绘图文件只是解决了地图专门内容的表示,对基础底图(行政区划界线)需使用数据库的弧段文件生成,但是只需一次处理生成绘图文件,供以后反复使用。同时在地图输出时,用户可对输出比例、线划色彩、绘图介质等方面做出不同选择。

县界数据库系统已在我国地方病地图集、人口地图集等编制中应用,这些应用改进了传统的制图方法,具有快速、准确、高质量等优点,其社会效益十分明显。而且从资料存储,更新和维护,应用和输出的整个过程来看,如果它配备于不同的专业部门,能为各个部门提供分析与制图的先进手段,成为日常业务管理的有力工具。

### 参 考 文 献

- 1) 刘岳、梁启章:多边形软件系统,测绘学报,第2期,1982年。
- 2) GIRAS: A Geographic Information Retrieval and Analysis System for Handling Land Use and Land Cover Data, Geological Survey Professional Paper 1059, 1977.

## THE ESTABLISHMENT AND APPLICATION OF THE DATA BASE FOR ADMINISTRATIVE UNITS OF THE PEOPLE'S REPUBLIC OF CHINA

Liu Yue Liang Qi zhang Jin Xue ying

(Institute of Geography, Academia Sinica)

### Abstract

The Data Base for Administrative Units of the People's Republic of China has stored standard codes and coordinates of more than 2300 counties. It is divided into four data files, that is, arc records, coordinates data, polygon records and chaining from arcs to polygons, and it includes more than 30 routines, which are used for the establishment and management of the data base, as well as for statistical analysis and for making thematic maps. This data base is based on minor microcomputer and suitable for statistical analysis and mapping within the whole country, province or region. In the last two years it has been used in various statistical mapping such as population mapping, etc, and the efficiency is much better than by using traditional method.

This paper describes the aims, functions and utilizations for designing the data bank, its content, data capture method, data organization, and its application.