

全国 1:25 万地名数据库的设计与建立

狄琳, 欧阳宏斌

(国家基础地理信息中心, 北京 100044)

Design and Establishment of 1:250 000 Toponym Database of China

DILIN, OUYANG Hong-bin

摘要: 全国 1:25 万地名数据库是国家基础地理信息系统全国 1:25 万数据库的重要组成部分, 它是一个空间定位型的关系数据库。介绍了全国 1:25 万地名数据库的数据内容、数据的组织、数据库结构、建库的工艺流程和数据库应用。对将空间地理数据中的地名转到关系型数据库中的过程做了一次成功的探索。

关键词: 地名; 数据库; 地理信息系统; 设计; 建立

一、概述

全国 1:25 万地名数据库是国家基础地理信息系统全国 1:25 万数据库的重要组成部分, 它是一个空间定位型的关系数据库, 储存和管理 1:25 万地形图上的各类地名信息, 包括政区、居民地、河流、湖泊、岛礁、山脉、沙漠、盆地、自然保护区的名称及属性等。它既可以与全国 1:25 万地形数据库连接使用, 又可以作为一个独立的专业数据库单独运行。它的建立可以为国务院各部委、各专业部门提供准确的地名信息, 特别是为发生各类紧急突发事件的地区提供准确的空间定位信息, 做到快速反应。

二、全国 1:25 万地名数据库的设计

1 数据内容

数据内容包括 1:25 万地形图上的各类地名信息, 近 80 万条, 约 200MB。数据项如下: 接口码、现图上名称、原图上名称、汉语拼音、语种、类别、政区代码、归属、驻地名、高程、图幅号、图名、出版年代、更新日期、成图单位、经度、纬度、采集单位、总接口码、层名等 20 项。这 20 个数据项不是每个地名都有, 有些地名只有其中几项, 如政区代码、归属、驻地名只在点状和面状居民地地名中有, 高程只在山峰名中有。经度、纬度对于以点状采集的地名要素为图上点状符号的定位点坐标; 对于以面状采集的地名要素为多边形标识点坐标; 当多个多边形为同一地名时, 则取其中主要的一个标识点为定位点。

以上 20 个数据项中 1~6 项直接从 1:25 万地形图抄表, 并使用统一的软件录入到 Foxbase 数据库中; 政区代码、高程、定位点经纬度、归属等数据项从地形数据库中获得, 其他数据项在数据组织和处理时添加。

2 数据库设计

全国 1:25 万地名库实体建立在 SUN 1000E 服务器上, 采用 Oracle 7.X 关系数据库进行管理。数据库应用管理系统采用 Develop 2000 开发环境在 PC 机上开发, 数据库和用户具有 Client/Server 体系结构, 用户在 PC 机上实现对地名库的访问。

Oracle 数据库由表、视图、索引等基本元素构成, 将地名信息按照其结构进行设计组织可有效地管理、检索。地名库结构如下:

第 1 个表——与图幅有关的信息。字段: 图幅号、图名、出版年代、更新日期、成图单位、左下角经度、左下角纬度、采集单位。

第 2 个表——行政区划信息。字段: 政区名称、原政区名称、政区代码、汉语拼音、语种、类别、驻地名称、驻地汉语拼音、驻地语种、驻地经度、驻地纬度。

第 3 个表——地名信息。字段: 现图上名称、原图上名称、汉语拼音、语种、类别、高程、政区代码、经度、纬度、总接口码、层名。

第 4 个表——建立图幅与行政区域对应关系。字段: 图幅号、政区名称、政区代码。

上述表与表之间都有直接或间接的关联项, 为满足查询需要, 方便使用, 还应建立视图和索引, 以及相关字段的数据字典。

3 应用管理系统设计

根据全国 1:25 万地名数据库的建库目标和要求, 方便查询及结果输出, 应用管理系统功能如下:

1 输入功能。可以从键盘或以数据文件形式输入各种地名信息, 包括地名定位点坐标。

2 查询统计功能。可以对数据库中的任意数据项进行组合条件查询、模糊查询, 并对查询结果进行统计。

3 输出功能。查询结果可以按照所需格式输出, 如单条地名显示, 打印地名表, 地名卡片或输出数据文件。

4 数据库维护功能。用户按照所授予的权限使用数据库, 对数据库进行增、删、改、更新、备份等操作, 以保证数据安全。

地名数据库应用管理系统在 Oracle 提供的 PC 开发环境 Develop 2000 下按设计要求开发, 建立弹出式菜单界面, 界面友好, 使用方便。

三、全国 1:25 万地名数据库建立

1 地名数据采集流程

地名数据采集工作由各省测绘局按照统一的技术规定分别承担,成果交国家基础地理信息中心处理后入库(图 1)。

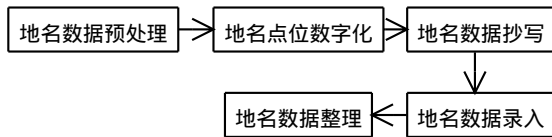


图1

以上每一步骤都要进行严格的正确性检查,发现问题及时纠正,以保证整个数据库的优良品质和可信度。

2 地名数据组织与处理

1:25 万地名数据在数据采集方案制定时,为了各省测绘局作业方便,减少采集工作量,规定将能够从地形数据中获取的地名数据项暂不采集,各省测绘局上交的地名数据只包括地名库中的 D 码,现图上名称,原图上名称,汉语拼音、语种、类别等 6 项,其他数据项需从地形库中直接或间接获取。因此,在建库前首先要对地名数据进行重新组织与处理,编制一些批处理程序,保证数据处理的质量与速度。可直接获取的地名数据项包括政区代码、山峰高程、地名定位点的经纬度、层名等;还有一些地名数据项可通过地形数据与地名数据的关系建立一些表格文件,间接获取数据,包括居民地的归属、驻地名等;其他如图幅号、出版年代、更新日期、成图单位、采集单位等项,建立一个以图幅为单位的数据文件,不加在每个地名记录中,以减少数据冗余。按照 1:25 万地名数据库的设计,用于数据库管理方面和加快检索速度的一些辅助数据文件和数据项,也将从地形数据中获得,如图幅左下角经纬度、政区中心点坐标等。

从地形库中获取地名数据内容以图幅为单位进行,利用 Arc/Info 软件的 Info 属性数据库管理功能,将地名数据与地形数据结合,完成数据获取。

3 地名数据检查修改及格式转换

在数据入库前,必须对重新处理后的地名数据作全面检查。在地名数据重新组织与处理过程中,牵扯到的内容较多,数据量较大,稍有不慎容易出现错误,因此可以编一些检查程序进行自动检查。

另外,必须按照设计的数据结构将整理好的数据文件转换成 ORACLE 数据库可以接受的格式。

4 地名数据入库

地名数据入库的步骤是:生成数据库(名称为 NGIS 25),生成表空间、定义表结构、定义视图、数据入库,建立索引等,数据入库主要采用文件插入方式(图 2)。

数据入库后分别按照图幅号、政区代码、地名类别建立索引。

在已开发库外汉字库的基础上,对原采集的地名数据中用代码表示库外字的地方替换成汉字和相应的汉语拼音。

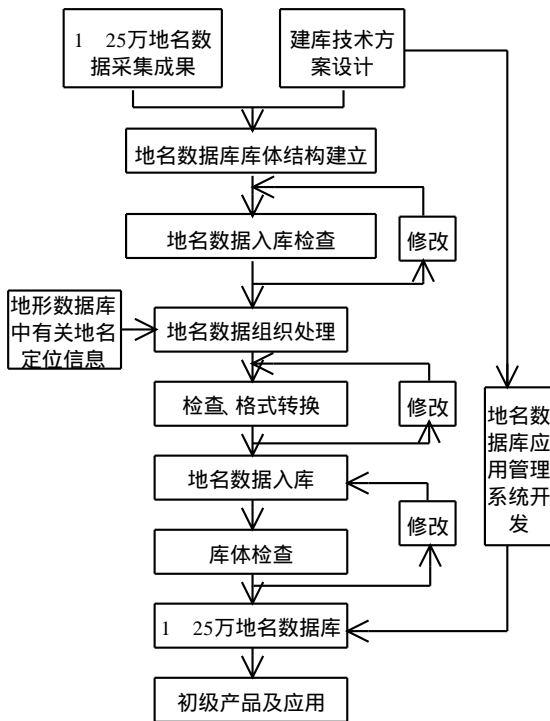


图2 建立全国 1:25 万地名数据库工艺流程图

四、全国 1:25 万地名数据库的应用

由于 1:25 万地名数据库检索速度快,又包含了空间定位信息,所以它在一些紧急突发事件的迅速定位方面发挥了很大作用,同时也弥补了现有 1:25 万地形数据库全库检索速度较慢的缺憾。我们已经为多家单位提供了全国或分省的 1:25 万地名数据,在数据提供过程中,我们也体会到目前 1:25 万地名数据库中包含的信息有限,特别是涉及到地名本身的信息如历史沿革、人口、工农业产值等经济信息都未包含,影响了它的使用面。随着今后信息共享的发展及全国 1:25 万地名数据库的补充和完善,我们应该相信 1:25 万地名数据库会在国民经济建设中发挥其应有的作用,得到更多用户的青睐。

五、结束语

由于关系型数据库的应用发展已经深入到广大的用户,其技术发展和广泛应用使其成为一种标准。使用关系型数据库技术解决空间地理数据库问题是当前 GIS 数据库发展的一个方向。全国 1:25 万地名数据库的设计和建立,对将空间地理数据中的地名转到关系型数据库中的过程做了一次成功的探索。也是使用关系型数据库技术解决空间地理数据库问题做了一次有益的尝试,为将来大型空间地理数据库的建设积累了丰富的经验。

参考文献:

[1] 曹瑜,胡光道 地理信息系统在国内外应用现状[J]. 计算机与现代化,1999,61(3).

[2] 吴信才 地理信息系统的基本技术与发展动态[J]. 地球科学,1998,23(4).