

Librarian 在数字黄土高原建设中的应用

韩琳, 杨勤科, 任洪玉, 梁伟

(中国科学院水利部水土保持研究所 西北农林科技大学, 陕西 杨凌 712100)

摘要: 黄土高原区水土流失与治理研究已积累了大量的空间数据, 如何有效地管理这些数据是一个亟待解决的问题。在描述 ARC/INFO 管理数据特点的基础上, 以黄土高原 1: 50 万专题数据为例, 应用 Librarian 建立了黄土高原专题数据库。分析表明建成的数据库具有数据提取、查询、显示及维护等功能。同时, 对数据库的元数据做了简单描述。

关键词: 黄土高原; 数据库; LIBRARIAN; 元数据

文献标识码: B

文章编号: 1000-288X(2003)03-0051-04

中图分类号: TP311.13

Employing Librarian Model of ARC/INFO for Data Management in "Digital Loess Plateau"

HAN Lin, YANG Qin-ke, REN Hong-yu, LIANG Wei

(Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, Shaanxi Province, China; Northwestern Sci-tech University of Agriculture and Forestry, Yangling 712100, Shaanxi Province, China)

Abstract: A lot of research has been carried out on soil and water conversion in the loess plateau. This resulted in the accumulation of a large quantity of spatial data. Improved management of these data well is becoming essential. Based on 1: 500 000 scale thematic data in the loess plateau, a method and process for building a database using the Librarian model of ARC/INFO is introduced. This database can be used to extract, query and display the spatial data. At the same time, the metadata of database is described.

Keywords: the loess plateau; database; Librarian model; metadata

黄土高原在世界范围内是一个十分特殊的地理区域, 深厚的黄土层记录了丰富的第四纪地质与环境(包括人类演化)信息, 强烈的水土流失和干旱沙化对我国乃至东亚地区环境变化构成强烈影响。自 20 世纪 50 年代以来, 黄土高原区水土流失治理得到了广泛重视, 为了制定合理的水土流失治理措施, 进行水土保持科学规划, 我国曾于 50 年代和 80 年代先后组织了 2 次全区综合科学考察, 积累了大量数据。近 10 a 以来, 有关调查研究成果数据的存储方式也从传统的纸质地图变为电子数据和矢量空间数据。

目前, 这种类型的数据多以行政为单元进行管理, 不利于数据量的合理分配, 使得数据应用困难; 并且由于增加了图形的边界长度, 给其维护等带来不便。LIBRARIAN 是 ARC/INFO 的基本模块之一^[1], 可管理大数据量的图形数据。它通过对图形分块, 并以建立空间索引的方式进行数据访问。

1 数据源分析

多年的科学考察, 自然资源调查所积累的成果数据, 比例尺范围从 1: 10 万到 1: 400 万不等。在数字黄土高原建设中, 已经收集并数字化了多种比例尺、多种专题和多种时期的矢量专题地图^[2]。本研究以 1: 50 万矢量数据为例, 说明黄土高原数据建库方式和过程, 其中专题数据有: 土壤图、土地资源图、土地利用图、侵蚀强度与侵蚀类型图; 基础数据有水系、道路、境界、地名图等; 地图投影均为 Albers 投影, 数据量共 108M。

2 数据库的建立

2.1 数据库设计

2.1.1 LIBRARIAN 数据库的简介 在逻辑上, LIBRARIAN 以 2 种方式组织区域数据: TILES 和

收稿日期: 2003-01-20

修回日期: 2002-03-03

资助项目: 中国科学院知识创新重要方向项目: 黄土高原水土保持的区域环境效应(KZCX3-SW-421)

作者简介: 韩琳(1978-), 女(汉族), 陕西渭南人, 在读硕士研究生。主要从事地理信息系统在水土保持中的应用研究。电话(029) 7012482, E-mail: hanlin kw@ hot mail. com。

LAYERS(图 1)。TILES 是对数据库覆盖的整个地理区域进行二维分割,数据在库中的存贮物理上是以 TILE 为单元的。TILE 可以基于规则格网、或不规则的自然地理边界等;在位置关系上,TILES 之间可连续或不连续,但不能出现相交或不相连的岛多边形。不同 TILE 之间不能用同一个名字。LAYER 是对数据的垂直分割,一个 LAYER 由某一种(点、线、面)或某些类型的 COVERAGE 构成,包含整个区域中的某类属性(如土地利用),用于存放数据具体内容。数据组织结构如图 1 所示。

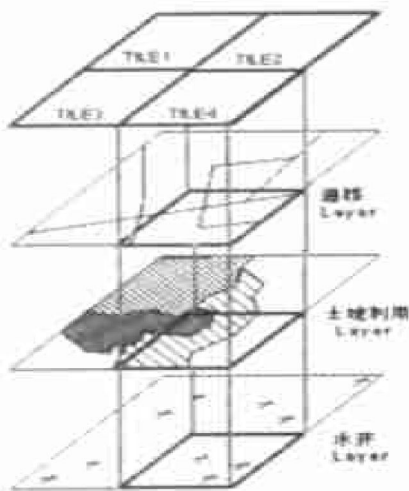


图 1 Librarian 数据结构

2.1.2 黄土高原数据库设计 数据库设计主要是 TILE, LAYER 和数据精度等的设计。设计时应考虑以下几方面问题:(1) 适当的数据量。考虑到数据存取速度和数据安全等问题,较大的数据量或许需要通过多个 LIBRARIAN 才能更有效地管理。(2) 用户对数据要求。面对不同用户,在数据范围、数据内容和数据精度等方面的要求有所不同。所以设计方法将有所不同。如对于某省级机关,以县为单元组织较好;对于测绘数据的管理,则适宜以图幅为单元管理。(3) 数据精度要求。数据库中所有的层(LAYERS)应当有相似的数据精度。(4) 空间坐标系。一个数据库系统应当使应用一个坐标系统,同一种投影方式。1:50 万黄土高原区数据库的建立,是根据国家基本比例尺地形图分幅和编号标准^[3],以 1:50 万比例尺的地形图划分方式,即在 1:100 万分幅的基础上,以经差 3°,纬差 2°的原则,将黄土高原区划分为 21 个 TILE 区域,投影方式选择 Albers,建立标准数据库。

2.2 数据库建立

2.2.1 INDEX COVERAGE 的建立 INDEX COV-

ERAGE 是数据库的索引文件,是由多个 TILE 边界组成的多边形 COVERAGE,其建立需经历以下几个步骤。

(1) 生成经纬网图。有 2 种方法,一是用 GENERATE 命令的子命令 FISHNET,生成 3*2 的格网多边形,投影并建立拓扑关系。这种方式简单易用,但是在 TIC 点的添加,线属性的注记等方面显得不能令人满意。二是利用 AML 编程方式生成,这种方式更加灵活,但是对操作者的要求更高。

(2) 以黄土高原边界为参考,选择与黄土高原区相交的格网区域,生成 INDEX COVERAGE(如图 2 所示),图中阴影部分为 TILE 区域。

(3) 用 ADDINDEXATT 命令给 INDEX.PAT 增加 2 个字段: tile-name 和 location;并给这 2 个字段分别赋值,tile-name 是根据国家对标准 1:50 万地形图的命名方式进行的命名的。

(4) 用 TOLERANCE 命令设置 INDEX COVERAGE 容限值。

(5) 用 PROJECTDEFINE 命令给 INDEX COVERAGE 定义投影文件。

INDEX COVERAGE 的数据精度要高于或等于将要插入的数据精度,它的投影方式代表了整个数据库的投影方式。因此,INDEX COVERAGE 投影与将要插入数据在投影方式和参数上一致。

2.2.2 建立 TILE 目录 用 LIBRARIAN 的 BUILD TILES 命令,根据 INDEX.PAT 属性表中的 2 个字段 TILE-NAME 和 LOCATION 的值来创建 TILE 工作空间。

LIBRARIAN: BUILD TILES< index> _COVER>

TILE 工作空间也可以通过 CREATE-Workspace 命令生成,用该命令生成的 TILE 工作空间必须在名字和位置上与 INDEX.PAT 属性表中 TILE-NAME 和 LOCATION 2 个字段值相同。



图 2 INDEX COVERAGE 结构

2.2.3 建立数据库框架 LIBRARIAN 的 CRE-ATEMAPLIBRARY 命令, 将根据 INDEX COVER-AGE 的数据库地址文件来建立数据库框架, 并在目录%ARCHOME%下的 tables 中增加一个记录, 以便在系统范围内查找数据库。

LIBRARIAN: CREATMAPLIBRARY LPB.DLG INDEX

LPB.DLG 是数据库名。运行该命令后, 计算机自动在当前目录下生成数据库参考目录 (database), 并在 Database 目录下生成数据库索引图 INDEX。数据的组织方式见图 3。此时, 用户自动拥有 administrator 权利。

2.2.4 定义 LAYER ADDLAYER 创建空的图层模板, 用该模板插入数据时, 数据在精度、投影方式上与 LAYER 模板一致。

Librarian:

ADDLAYER < layer> < cover> < format.cover>

- < layer> 是 database 目录下的 layer 名;
- < cover> 是在 tile 目录下存储 map section/ coverage 名;
- < format.cover> 是用来定义 layer 数据精度、属性项的 coverage 名称。

为了简化数据库结构, 在数据库建立过程中, 将 < Layer> 和< format.cover> 命名一致。在 INDEX.PAT 文件中, 添加了所有与 LAYER 同名的数据项; 黄土高原区 1: 500 000 共有 8 层数据, 再加上主控制点图层, 共 9 个数据模板(layer)。

2.2.5 数据入库 LIBRARIAN 在进行数据插入时, 有 2 种方式: PUTTILE 和 INSERT。当要插入数据图幅与 TILE 边界一致时, 用 PUTTILE 命令, 否

则用 INSERT 将数据分配到各个 TILE 中。本次建库用 INSERT 命令进行数据插入。

Librarian: INSERT < cover> < layer> {dropborder}

cover 是将要插入的图形数据, layer 是数据库中的数据模板。

2.2.6 主控制点(mastertic) 在 INDEX 结构建立过程中, LIBRARIAN 将自动在 TILE 区域范围的 4 个角点产生 4 个控制点, 为了保持图形原有的 tic 点, 在数据建库过程中, 必须生成含有原图所有控制点的 tic 图层, 并且将该图层放在数据库参考目录 database 下, 命名为 mastertic。

2.3 数据库适用性分析

建成的数据库具有数据存储和提取、更新、查询、显示以及数据安全性保护等功能。

2.3.1 数据存储和提取 数据存储是在 2 个层次上进行的: TILES 和 LAYERS。数据可以在水平和垂直 2 个方向上进行扩展。对于数据库区域范围覆盖度不够的, 可增加 tiles 区域; 否则直接对库进行数据插入。数据提取是在用户操作权限范围内, 根据用户设置的区域范围和专题信息进行的数据拷贝工作。

2.3.2 数据库更新 数据库更新包括 2 方面内容, 即数据库结构修改和数据的更新。LIBRARIAN 以分级形式组织数据, 该模型建立在 4 个层次上: 数据库地址文件、数据库参考目录 (Database)、TILE 工作空间、Map section/ coverage(如图 3)。这种库组织方式便于数据库的管理和维护。数据库结构的修改包括数据地址文件, 数据库框架结构, tiles, layers 等的修改; 数据修改是在 tiles 的选定范围内对 Map section/ coverage 修改, 修改的数据和原数据在属性项定义、数据精度和地图投影等方面一致。

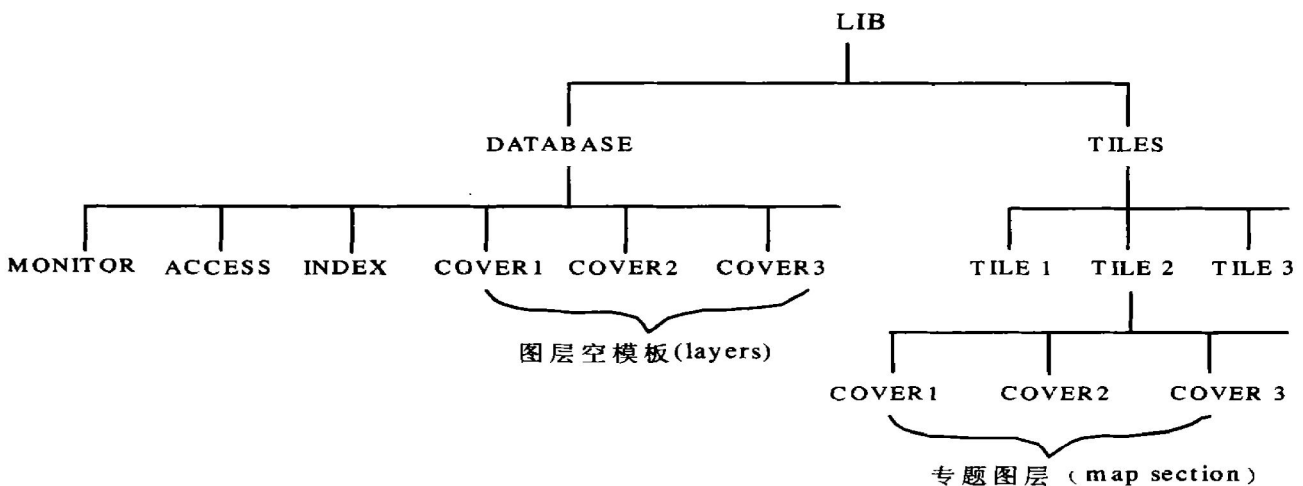


图 3 LIBRARIAN 数据目录结构

2.3.3 数据查询和显示 数据查询包括库查询、图形查询和属性查询及其显示等。库查询包括系统中所有的数据库名称的查询、当前数据库的访问权限、数据库地址以及 TILES 和 LAYERS 等的查询。数据查询是对某图属性表中某字段值的查询(IDENTITY)或是图上某些图斑面积和线段值的量算(MEASURE)等。

2.3.4 数据库安全性保护 为了保护数据安全性及其在维护上的一致性,需要设置数据库的访问权限。LIBRARIAN 数据库访问权限有 5 种类型: ADMINISTRATOR, MANAGER, OPERATOR, BROWSE, LOOK, NONE; 权限范围设置可在 LIBRARIAN, TILE 和 LAYER 3 个层次上进行。LIBRARIAN 的访问权限可通过 REVOKE 命令取消。

3 数据库说明文档

数据库说明文档即元数据文件。元数据是数据的数据,是关于数据内容、质量状况和其它有关特征的背景信息。是数据共享不可缺少的文件之一。

黄土高原 1:50 万数据库数据说明文档属于黄土高原水土保持元数据的一部分,是参考现有的元数据标准^[4]和一些空间数据库说明文件^[5]编制而成。该元数据说明是在 2 个层次上进行的。

(1) 区域级空间元数据,该级元数据是对黄土高原区 1:50 万空间数据库的总体说明,其内容主要包

括:标识信息数据库集成/数据分层信息、数据组织方式、数据库质量信息、数据覆盖范围、数据生产基础或数据来源信息、数据维护信息、数据共享信息等。

(2) 空间数据库级元数据,该级元数据是针对组成该数据库各图层的具体说明,其主要内容有:标识信息(所在集/库名、产品名称等)、数据质量信息、空间数据组织信息(数据格式、存储方式、路径)、空间参照系信息(坐标系统、投影信息)、共享信息(共享方式、发行信息等)、生产信息、属性数据编码信息(数据分类编码标准/原则、编码分类系统、图例系统)、以及专题信息(专题地理特征说明)等。标识信息是该级元数据和其相应的上级元数据标准联接的关键码。

[参 考 文 献]

- [1] 樊红编著. ARC/INFO 应用与开发技术[M]. 武汉测绘科技大学出版社, 1999.
- [2] 杨勤科. 论数字黄土高原建设的若干问题[J]. 水土保持通报, 2000, 20(4): 33—35.
- [3] 国家基本比例尺地形图分幅和编号. 中华人民共和国国家标准(GB/T 13989—92)[S]. 国家技术监督局, 1992—12—17 发布.
- [4] 基础地理信息数字产品元数据. 中华人民共和国测绘行业标准(CH/T 1007—2001)[S]. 国家测绘局, 2001—03—05 发布.
- [5] 全国 1:25 万地形数据库、数字高程模型数据说明[Z]. 国家基础地理信息中心, 1996 年 6 月. <http://nfgis.ns-di.gov.cn>.

(上接第 50 页)

(3) 须进一步对 SDEM 的适用范围与应用中的不确定性进行深入研究,并对不同比例尺、不同分辨率、不同地形的样区进行研究实验,为我国 DEM 的生产提供一种新的思路。

[参 考 文 献]

- [1] 余鹏,刘丽芬. 利用地形图生产 DEM 数据的研究[J]. 测绘通报, 1998(10): 16—18.
- [2] 唐新明,林宗坚,吴岚. 基于等高线和高程点建立 DEM 的精度评价方法探讨[J]. 遥感信息, 1999(3): 7—10.
- [3] 汤国安. 不同比例尺 DEM 提取地面坡度的精度比较研究[J]. 水土保持通报, 2001, 21(1): 53—56.
- [4] 赵牡丹,汤国安,陈正江,等. 黄土丘陵沟壑区不同坡度分级系统及地面坡谱对比[J]. 水土保持通报, 2002, 22(4): 33—36.
- [5] 汤国安,龚健雅,陈正江,等. 数字高程模型地形描述精度量化研究[J]. 测绘学报, 2001(11): 361—365.
- [6] 蒋忠信. 流域沟壑密度理论极值数学模式商讨[J]. 地理

研究, 1999, 18(2): 220—223.

- [7] 赵帮元,汤国安,马安利. 不同地貌类型区 1:25 万比例尺 DEM 的建立方法[J]. 水土保持通报, 2002, 22(2): 45—48.
- [8] Tang Guoan. A Research on The Accuracy of Digital Elevation Models[M]. Science Press, Beijing, 2000.
- [9] Li Zhilin. Sampling Strategy and Accuracy Assessment for Digital Terrain Modelling[M]. Ph. D. Thesis, The University of Glasgow, 1990.
- [10] Ackermann F. Techniques and strategies for DEM Generation, Digital Photogrammetry: An Addendum to the Manual of Photogrammetry[M]. ASPRS, 1996.
- [11] Wang Ping. Applying Two Demensional Kalman Filtering for Digital Terrain Modelling[M]. IAPRS, 1998.
- [12] Gao J. Resolution and accuracy of terrain representation by grid DEMs at a microscale [J]. International Journal of Geographical Information Science, 1997, 11(2): 199—212.