

文章编号:1671-5942(2013)03-0092-06

基于要素编码的基础地理数据库的构建及 DLG 自动 入库的实现^{* 1}

刘德儿 兰小机 袁显贵

(江西理工大学建筑与测绘工程学院, 赣州 341000)

摘 要 鉴于地理要素分类与编码在地理信息组织、数据转换等方面的应用作用,提出基于地理要素分类与编码的基础地理数据库构建策略,并借助 Geodatabase 按照现行分类原则在大类上构建要素数据集,在中类上构建要素类。在此基础上,建立要素分类编码与要素类的映射机制,将其应用于 DLG 数据入库软件开发中,并通过 ArcGIS Engine 和 ObjectArx 两者无缝结合开发数据转换接口嵌入 AutoCAD 进程中得到体现,从而实现 DLG 数据自动、高效地入库。

关键词 地理要素分类;要素编码;基础地理数据库;DLG;数据入库

中图分类号:P208 **文献标识码:**A

FUNDAMENTAL GEODATABASE DESIGN AND DLG DATA IMPORTING BASED ON FEATURE CODE

Liu Deer, Lan Xiaoji and Yuan Xiangui

(Faculty of Architectural and Survey Engineering, Jiangxi University of Science and
Technology, Ganzhou 341000)

Abstract According to the value of the classification and coding of geographic features in these aspects, such as organization of geographic information, data conversion, etc., a strategy was brought forward which could construct fundamental geographic database. In accordance with the principle of the present geographic feature classification, feature datasets were built in the major categories, and feature classes were built in the middle categories with geodatabase. On this basis, the mapping mechanism was established between feature classification code and feature class, and it was used to develop the DLG data software. Through the seamless integration between ArcGIS engine and ObjectArx, this software embedding in the process of AutoCAD was realized, and it can realize the warehousing of DLG data automatically and efficiently.

Key words: classification of geographic feature; feature code; fundamental geodatabase; DLG; data importing

1 引言

空间基础地理信息是 GIS 应用系统的重要组成部分,而空间数据入库是 GIS 工程项目开发的最基本也是最重要的环节。GIS 数据的来源具有多样

性,有矢量、栅格,甚至是视频数据。本文研究的 GIS 数据入库主要涉及的是 1:500、1:1000 比例尺的 dwg 格式的矢量地形图数据(DLG),目标地理数据库为 ESRI 的 Geodatabase。矢量地形图数据是城市基础地理数据库的主要支撑数据。当前,Auto-

* 收稿日期:2012-12-02

基金项目:国家自然科学基金(40971234);江西省教育厅科学研究项目(GJJ10489, GJJ13431);江西理工大学基金(jxxj11012)

作者简介:刘德儿,男,1976 年生,博士,副教授,主要从事 GIS 研究. E-mail:landserver@163.com

CAD 数据到 GIS 数据的转换主要的方式有直接转换和间接转换^[1]。直接转换的难度较大,需要非常熟悉两者的数据结构。间接转换主要将源数据格式转换成开放的中间格式,如 AutoCAD 的 DXF、Esri 的 Shape file、MapInfo 的 MIF 等格式^[1]。基于中间格式实现转换的途径主要有使用 FME 组件、宿主 GIS 软件二次开发组件或底层开发转换模块^[1-6]。通过中间格式转换方法不需要了解 dwg 数据结构,但是中间格式不能完全反应 dwg 数据的所有信息,例如图块信息不能得到完善保持。鉴于 ArcGIS Engine 组件支持嵌入式开发,本研究将其和 AutoCAD 二次开发组件 ObjectArx 实现了无缝衔接,使得开发的转换接口直接在 AutoCAD 进程中运行。再者,当前基础地理信息要素分类与编码在地理信息组织和转换中扮演着重要角色^[7],为了充分利用其作用,本文将其在数据库设计和入库软件模块开发过程中紧密结合,实现无需人工干预的自动化入库。

2 基础地理信息分类标准及基础地理数据库构建

2.1 基于地理信息分类与编码的基础地理数据库构建^[8,9]

地理数据库模型设计的优劣将直接影响到城市基础地理信息系统的运行效果,比如空间拓扑分析,管网模型、路网模型的建立等。根据笔者的实践经验,基于地理信息分类与编码的地理数据库构建应遵循的规则主要包括:

1)要素类不能过于综合,即构建要素类的分类级别不能太高,如果太高可能对低级别分类的要素属性会过于综合,不能很好地反应低级别分类间的属性差异,因为分类级别越高,抽象程度就越高;

2)构建要素类的分类级别不能过低。如果构建要素类的分类级别过低,那么地理数据库中包含的要素类就越多,这样将增加地理数据库维护成本以及占用更多计算机资源。笔者曾经在小类基础上建立要素类,每个类别分别构建点、线、面三种类型要素类,结果地理数据库中要素类达到了 708 个,在 ArcMap 中需要很长时间才能加载完毕。

根据实际工程经验得出在大类基础上构建要素数据集(如图 1(b)所示的管线要素数据集)、在中类基础上构建要素类是比较理想的选择。按照现行地理信息分类标准,整个基础地理数据库由八大专题构成,每个大类构成一个专题,并包含在一个要素数据集中,如图 1(a)所示共有 8 个数据集。按实际情况对某些类型分别构建点、线、面三种类型要素类,按此方式构建的地理数据库中将包含 138 个要素类,同时对于面状道路、河流还创建了中线要素类,要素类标准命名如图 1(c)所示,要素类的别名采用要素分类名称,用于图层名显示。为了满足图 1(c)要素类命名方式需求,本文对 GTB/T 13923—2006 地理信息分类编码增加了两位,用以标识地理信息特征码,它由图形实体码(第 7 位)+实体特征码(第 8 位)组成。

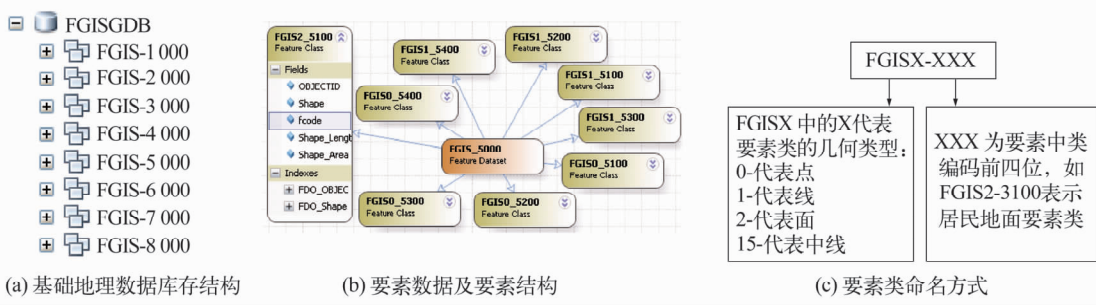


图 1 基础地理数据库及其内部组成
Fig. 1 The foundational geodatabase and its structure

2.2 基于 ArcSDE Geodatabase 的入库数据管理策略

为了能够使导入文件地理数据库的 DLG 数据实现共享、快速更新等应用,本文将入库后的数据统一存储在 Oracle11g 中的 ArcSDE Geodatabase 中,如图 2 所示,整个体系分为四层:客户应用层、业务逻辑层、数据连接层和数据层。由于 ArcSDE 本身具有海量数据存储、多用户并发访问、版本管理、长事

务处理等强大优势,Oracle11g 是成熟的关系数据库,基于 ArcSDE 和 Oracle 的空间数据管理策略是当前最佳选择之一。

3 基于要素编码的数字地形图入库流程

3.1 CAD 数字地形图预处理

数字地形图规则化的目的是要使 CAD 数据符

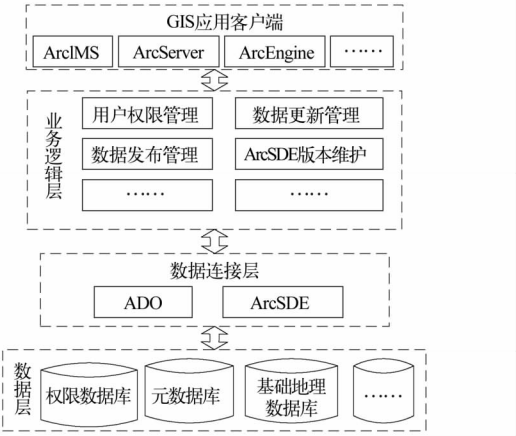


图 2 基于 ArcSDE geodatabase 的入库数据管理

Fig. 2 The management of warehousing data based on ArcSDE geodatabase

合 GIS 系统所需数据格式,在 GIS 中称现实对象为要素,它包括:几何属性和非几何属性。入库前数据预处理的主要工作有^[10]:1)对多边形进行封闭,如房屋、宗地块等;2)各图元注记整体化,如单位名称,如果是逐字注记,在导入时将生成多个点要素;3)线型图元的检查;4)道路、河流等用线描述的面要素的处理;5)地理要素编码和扩展属性赋予,本文将编码和属性数据写入地形要素的 XData 中;6)重复地形要素的清理;7)拓扑关系正确性判断,如房屋不能在道路上,道路不能和河流重叠等。

某些特殊情况的处理:

1)带“洞”多边形的处理:如图 3 所示,带“洞”房屋在 CAD 和 GIS 中的不同表达,数字成图时,在 CAD 中绘制两个多边形描述一个带“洞”的房屋,而在 GIS 中使用一个多边形描述,它由内环和外环组成,可视化时内环只显示边界。对于这样的情形,在数据预处理时将两个多边形分别标识为外环和内环,然后创建成块,最后将编码和属性赋给块引用对象。

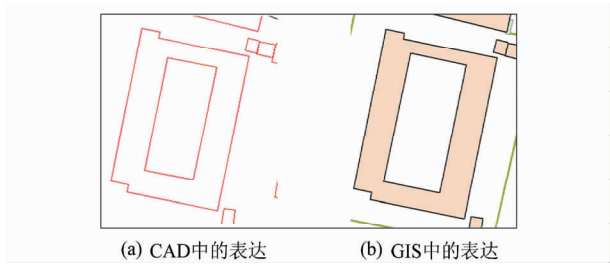


图 3 带“洞”多边形表达方式

Fig. 3 The polygon with holes

2)由多个岛屿组成对象的处理:如图 4 所示对象在 GIS 中的表达由多个外环和多个内环组成,而在 CAD 中由多个独立多边形表示,处理方式同 1)。

3)多路径对象的处理:如图 5 所示,一条道路

跨越一条河流,在 CAD 绘制时一般将其分成两段,中间用一桥梁进行连接,预处理时将这段组成一个块,将编码和属性赋给该块对象。



图 4 由多部分组成的对象

Fig. 4 The object composed of multi-part



图 5 多路径对象

Fig. 5 The multi-path object

3.2 CAD 和 ArcGIS 数据模型映射关系

CAD 数据模型和 GIS 数据模型之间存在较大差异,前者是为满足制图需要而设计,后者是为满足 GIS 各种空间分析和高级功能而设计。CAD 数据模型相对简单,是一种简单线框模型。如在 GIS 中,带岛多边形由外环和一个或多个内环构成的面,而在 CAD 中只是由封闭的多段线构成的多个空洞多边形,没有面的概念。为实现这两种数据模型的相互转换,本文建立这两者的属性及几何映射关系如表 1 和表 2 所示。

表 1 CAD 与 ArcGIS 模型属性映射关系

Tab. 1 The mapping of attribute between CAD and ArcGIS

CAD 数据模型		GIS 数据模型	
数据类型	说明	数据类型	说明
Text	文本	Text	文本
8-bit integer	8 为整型	Short integer	短整型
16-bit integer	16 位整型		
32-bit integer	32 位整型	Long Integer	长整型
64-bit integer	64 位整型		
Real	实数	Float	单精度浮点型
		Double	双精度浮点型
Binary chunk	二进制类型	Blob	大二进制类型

3.3 要素编码与要素类对照关系的建立

建立要素编码与要素类对照表的目的是为了提

高软件设计的灵活性和通用性。建立此表的好处主要体现在数据库模型的改变而对入库软件的影响不是很大,因为无论以什么原则来建立要素类,一个要素编码只对应一个要素类,所以数据库模型的改变只需对此表中的要素类项进行更新,而无需对入库软件做较大的修改。对照表(表3)使用 Access 数据库存储,管理和操作都比较简单。

表 2 CAD 与 ArcGIS 模型几何映射关系

Tab.2 The mapping of geometry between CAD and ArcGIS

CAD 数据模型		GIS 数据模型	
CAD 数据类型	几何接口	图元类型	几何接口
独立地物块	AcdbBlockReference	点	IPoint
单行文本	AcdbText		
多行文本	AcdbMText		
高程点	AcDbPoint		
直线	AcdbLine	多段线	IPolyLine
多段线	AcdbPolyline		
二维多段线	AcDb2dPolyline		
三维多段线	AcDb3dPolyline		
多条多段线组成的块	AcdbBlockReference	面	IPolygon
弧	AcdbArc		
表示多边形多段线	AcdbPolyline		
表示多边形闭合二维多段线	AcDb2dPolyline		
表示多边形闭合三维多段线	AcDb3dPolyline	面	IPolygon
多条闭合多段线组成的块	AcdbBlockReference		
圆	AcdbCircle		

表 3 要素编码与要素类对照表

Tab.3 The mapping between feature code and feature class

要素编码	要素名称	要素集	要素类	要素类型
80000000	植被与土质	FGIS_8000		
81000000	农林用地		FGIS1_8100	
81010010	地类界		FGIS1_8100	线
81010024	地类界		FGIS2_8100	面
81020010	田埂		FGIS1_8100	线
...

3.4 入库软件程序设计思路

入库软件设计采用要素到要素类的方法,由于一个要素编码唯一对应一个要素类,采用该方法所花时间少。例如,地理数据库中有 100 个要素类,一幅地形图中有 1000 个要素,按照由要素到要素类的方法只要循环一千次就可以入库完毕;而按照由要素类到要素的方法则需要循环 10 万次才能完成。入库软件设计思路如下:

- 1) 数字地形图预处理;
- 2) 对预处理好的数字地形图数据库对象进行遍历,逐一读取所有要素,并从 Xdata 中获取该要素的编码;
- 3) 根据编码从对照表 3 中找到对应要素类,或直接参照要素类命名规则根据要素编码生成要素类

名,如要素编码为 81010024,则要素类名为 FGIS2_8100;

- 4) 打开要素类,将要素几何和要素属性分别写入所属要素类中;
- 5) 数据写入 File Geodatabase 后,在 ArcMap 中进行再次检查,主要是充分利用 ArcMap 提供的拓扑工具来检测;
- 6) 将 File Geodatabase 中的数据导入到 ArcSDE Geodatabase 中。

3.5 入库软件程序的实现

本文充分利用 ArcGIS Engine 组件可嵌入式编程的特性,将其和 CAD 二次开发组件 Objects ARX2010 相结合利用 C#语言在 VS. NET2008 环境进行二次开发,然后在 CAD 下运行。这里要解决的关键问题是如何将 CAD 的几何模型转换成 ArcGIS 的几何模型,两种模型对几何的表达是不一样的,例如对多边形的表达,CAD 数字地形图中采用封闭的空洞的多段线来描述,而 GIS 中采用填充的面来表达。软件主要功能模块:几何数据提取与转换模块——主要负责将 CAD 实体几何转换成 GIS 几何,包括:点转换、线转换、面转换三个子模块;属性数据提取模块——主要负责从 CAD 实体的 Xdata 数据集中提取属性数据写入道到要素类中对应要素属性字段中。数字地形图入库前后对比效果如图 6 所示,本文开发的入库软件实现了无损转换。

数字地形图入库程序伪码执行过程:

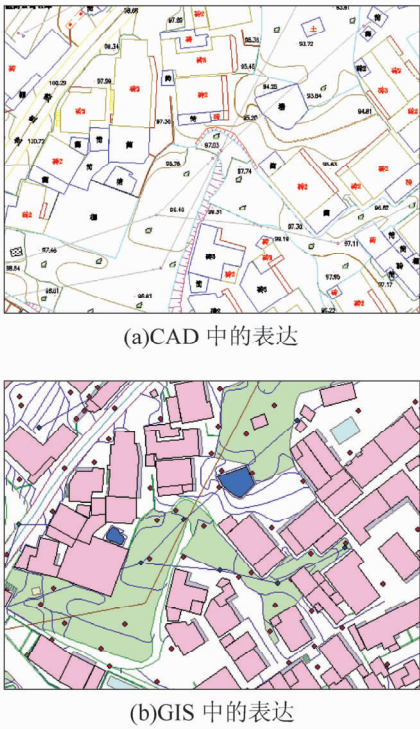


图 6 入库前后效果对比

Fig.6 The contrast between DLG and GIS data

```
Public void ConvertCADToFGDB()  
{ //得到当前活动的 dwg 文档  
    public Document pCurDoc = Autodesk. Auto-  
CAD. ApplicationServices.  
Application. DocumentManager. MdiActiveDocu-  
ment;  
    Database pCurDb = pCurDoc. Database;//得到  
当前活动 dwg 的数据库对象  
    using( Transaction acTrans = pCurDb. Transac-  
tionManager. StartTransaction() )  
    { BlockTable pBTab = acTrans. GetObject(  
pCurDb. BlockTableId, OpenMode. ForWrite ) as  
BlockTable; //打开块表  
        //打开模型空间块表记录  
        BlockTableRecord pBTABRec = acTrans. GetOb-  
ject(  
pBTAB [ BlockTableRecord. ModelSpace ], Open-  
Mode. ForRead ) as BlockTableRecord;  
        foreach ( ObjectId pId in pBTABRec )//遍历模型  
空间块表记录  
        {  
            Entity pEnt = ( Entity ) pTran. GetObject ( pId,  
OpenMode. ForWrite );  
            { 读取 CAD 实体扩展属性; }  
            { 根据要素编码得到对应要素类名; }  
            { 通过 ArcGIS Engine 组件打开要素类, 并创建  
一新要素; }  
            { 调用属性转换模块, 将 CAD 扩展属性映射为  
要素类对应属性; }  
            { 调用几何转换模块, 将 CAD 几何转换成 Arc-  
GIS 对应几何; }  
            { 将新建要素写入要素类; }  
        }  
        acTrans. commit();  
    }  
}
```

4 入库软件应用测试及评价

分别对 10 幅 CAD 格式的数字地形图由小到大进行测试, 被测试文件大小、包含实体个数、入库耗时见表 4。从表 4 可以看出: 1) 入库消耗时间随文件的增大而变长; 2) 入库耗时随实体个数增多而变长; 3) 入库软件总体效率高、稳定性好, 总体呈线性变化。同时, 从表 4 发现实验 8 的数据量大于实验 7 的数据量, 而耗时比试验 7 少很多, 引起这种结果的原因主要有两方面: 1) 实体顶点个数越多耗时越多; 2) 实体构造越复杂耗时越多。

表 4 实验数据及入库耗时

Tab.4 Test data and results

序号	文件大小(Mb)	实体个数	入库耗时(秒)
1	0.2	1 713	4
2	0.36	2 256	6
3	0.47	2 186	7
4	0.92	7 811	16
5	1.18	14 656	24
6	1.76	20 012	33
7	2.31	26 573	49
8	3.5	33 818	32
9	4.52	40 966	43
10	9.62	69 381	84

5 结束语

虽然现在对数字地形图导入 GIS 数据有很多研究, 但就地理要素分类与编码应用于其中还研究较少。从 GIS 工程实践效果可以看出, 本文基于地理要素分类与编码构建地理数据库以及基于要素编码的数字地形图入库软件开发优势体现在以下几点:

- 1) 在大类基础上构建要素数据集和在中类基础上构建要素类较好的保持了地理空间的专题特征, 也较好地符合了 Geodatabase 中设计“要素数据集”这个集合类型的初衷;
- 2) 在中类基础上构建要素类使得其数量大大减少, 从而大大减少了要素类 I/O 时间;
- 3) 基于要素中类编码的要素类命名规则, 可以比较方便地理要素与要素类之间建立映射关系, 在入库过程中可以直接由要素编码获得对应要素类, 从而提高入库软件运行效率;
- 4) 将 ESRI 公司的 ArcGIS Engine 组件和 AutoDesk 公司的 ObjectsArx 组件无缝结合进行嵌入式二次开发, 实现了数字地形图数据到 GIS 数据的直接转换, 转换后的数据具有较高的保真度。

参 考 文 献

1 陈能, 施蓓琦. AutoCAD 地形图数据转换为 GIS 空间数据的技术研究与应用[J]. 测绘通报, 2005, (8): 11 - 14. (Chen Neng and Shi Beiqi. Conversion technique of AutoCAD topographic map data to GIS spatial data[J]. Bulletin of Surveying and Mapping, 2005, (8): 11 - 14)

2 刘德儿, 兰小机. 基于 ArcObjects 的数字地形图导入地理数据库方法[J]. 工程勘察, 2008, (7): 37 - 39, 43. (Liu Deer and Lan Xiaoji. ArcObjects based method for DLG converted into geodatabase[J]. Geotechnical Investigation & Surveying, 2008, (7): 37 - 39, 43)

3 宋伟东, 符韶华. DLG 到 GIS 的数据转换方法研究[J]. 测绘通报, 2004, (3): 54 - 56. (Song Weidong and Fu Shaohua. A research on the data conversion of DLG into GIS[J]. Bulletin of Surveying and Mapping, 2004, (3): 54 - 56)