

DGN 数据格式转换的研究与实验

毛卫华,应荷香  
(浙江省测绘局,浙江 杭州 310012)

Format Transfer of DGN Data: Investigation and Test

MAO Wei-hua, YING He-xiang

摘要: DGN 格式的基础地理空间数据如何提供给其他 GIS 软件应用, 或用于建立基础地理空间数据库, 将涉及数据转换这一关键问题。以解决这一问题为目标, 结合实际研究和实验过程, 阐述了 Microstation 与其他 GIS 软件进行数据交换的一个行之有效的实现方法。

关键词: 外部数据交换; 空间要素; 图形信息; 属性信息

一、引 言

Microstation 是美国 Bentley 公司的 CAD/GIS 软件产品, 具有强大的 CAD 功能。在加载 Geovec/MSFC、Geographics 等 GIS 工程模块后, 即为一 GIS 软件平台。该软件在国内有很多的用户, 应用涉及测绘、规划、电力、环保等领域。浙江省测绘局在采集 1:1 万比例尺基础地理空间数据时, 主要采用了 Microstation 及 Geovec/MSFC、Geographics 等模块(以下简称“MS 系列”)。采集得到的空间数据主要为 DGN 格式。这些数据无论是作为基础地理数据提供给其他 GIS 软件应用, 还是用于建立基础空间地理数据库, 都将涉及进行数据转换的问题。

笔者对 Microstation 设计文件进行研究后, 设计了一个行之有效的数据通用交换格式(暂称为“ABC 格式”), 并使用 Microstation 的主要开发语言 MDL, 从底层开发出数据交换接口模块。利用该模块可实现 DGN 文件与纯文本文件相互间的输入与输出(I/O)转换。若在其他 GIS 软件平台上也开发了相似的接口, 则该软件与 Microstation 之间即可通过 ABC 格式实现数据的交换, 如图 1 所示。

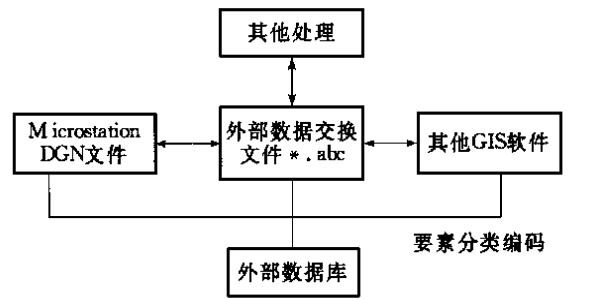


图1 通过外部数据实现数据交换

二、Microstation 中空间要素的表达及管理方法

使用“MS 系列”进行空间要素采集, 生成的数据文件中, 空间要素可视为由图形数据和属性数据两部分组成, 其中图形数据主要是要素的几何特征与空间定位信息, 属性数据则包括要素的图形属性(颜色、线型、图层等)、要素的拓扑

关系描述、要素上附着的其他属性信息(如分类编码等)。在我省采集的基础地理空间数据中, 属性数据尚未包含拓扑描述, 要素附着的属性信息主要为分类编码信息。

“MS 系列”对图形数据与属性数据是分开管理的。在采用 Microstation 和 Geovec/MSFC 组合时, 属性数据保存在特定的外部文件中, 由 MSFC 模块来统一维护和管理; 而采用 Microstation 和 Geographics 组合时, 属性数据通过 ODBC 用外部关系型数据库(如: Access 等)进行管理, 这也是目前大多数 GIS 软件采用的方法。外部数据库按以下方式组织属性数据: 在 category 表、feature 表、ugfeature 表中定义要素的分类、编码、图形属性(颜色、图层、线型、元素类型等); 在 mscatalog 表中登记和申明可使用的表; 在若干用户自定义属性表中记录要素的其他属性信息(如道路等级、最高水位、历年农产量等)。

图形数据与属性数据之间通过关键字段 MSLINK 等来建立链接。

三、数据交换格式的设计

按 Microstation 对空间要素表达及管理的方法, 结合我省基础地理空间数据的实际情况, 当数据交换文件提供为: ① 要素的空间定位数据, 即要素的几何类型、坐标; ② 要素的分类编码特征数据, 根据该数据可对外部数据库进行检索, 在数据库中记录了要素所采用的颜色、符号、粗细、所在的图层等信息时, 可以完整表达要素的图形与属性信息。

对要素几何类型的描述, 不同 GIS 软件之间存在差异, 这种差异既可能是类型名称差异, 也可能是几何特征表达差异。

以 Microstation 和 Geostar 为例, 几何特征表达差异表现为: 在 Geostar 中, 对于一根具有  $n$  个顶点的曲线(Curve), 只需用  $n$  个特征点来描述即可, 而在 Microstation 中, 同样一根曲线却含有 4 个隐藏点, 分别位于线串的首尾, 用于建立曲线端点的曲率, 进行描述时将需要  $n+4$  个特征点。

名称差异则如表 1 所示。

表1 Geostar 与 Microstation 几何类型使用名称的对应关系

Geostar	单点	折线	圆	曲线	面	文本
Microstation	Cell/ Point	Line / Lstring	Arc / Ellipes	Curve/ B-spline	Shape	Text

因此,在数据交换文件设计时,应综合考察所涉及的各GIS软件,尽可能完整描述要素的几何类型与几何特征,避免在数据交换的过程中产生数据丢失甚至错误的现象。

为了便于分析和使用,数据交换文件采用纯文本形式,文件由文件头、若干个数据块、文件结束标记3个部分构成。每个块对应设计文件(\*.dgn)内的一个Element块,其基本结构如下:元素类型(9种类型)、分类编码特征、几何空间特征描述(7种描述)。

元素类型与几何空间特征描述的对应关系见表2。

表2 元素类型及相应的几何特征描述表

元素类型	几何特征描述		
LINE_STRING_ELM	线串的顶点数目:	N	
	坐标串:	X <sub>1</sub>	Y <sub>1</sub>
		X <sub>n</sub>	Y <sub>n</sub>
LINE_ELM	线串的顶点数目:	2	
	坐标串:	X <sub>1</sub>	Y <sub>1</sub>
		X <sub>2</sub>	Y <sub>2</sub>
SHAPE_ELM	面域的顶点数目:	N	
	坐标串:	X <sub>1</sub>	Y <sub>1</sub>
		X <sub>n</sub>	Y <sub>n</sub>
ELLIPSE_ELM	长半轴:	Ellipse_2d.primary	
	短半轴:	Ellipse_2d.secondary	
	旋转角度:	Ellipse_2d.rotation	
	圆心坐标:	X <sub>0</sub>	Y <sub>0</sub>
	起始角度:	Arc_2d.startangle	
	终止角度:	Arc_2d.endangle	
	长半轴:	Arc_2d.primary	
ARC_ELM	短半轴:	Arc_2d.secondary	
	旋转角度:	Arc_2d.rotation	
	圆心坐标:	X <sub>0</sub>	Y <sub>0</sub>
	特征点数目:	2	
	采集点位:	X <sub>0</sub>	Y <sub>0</sub>
	方向点位:	X <sub>1</sub>	Y <sub>1</sub>
	线串的顶点数目:	N+4	
CURVE_ELM	坐标串:	X <sub>1</sub>	Y <sub>1</sub>
		X <sub>n+4</sub>	Y <sub>n+4</sub>
COMPLEX_STRING	可以由上述7种组合而成		
COMPLEX_SHAPE			

文件头与文件尾则较灵活也较容易实现,在文件头中可根据需要存储一些关于数据的说明,本文对它们将不再进行说明。以下为采用该数据交换文件格式的一个样本数据部分内容。

LINE\_STRING\_ELM 18001 38

3 340 344. 93 509 261. 44

...

3 340 810. 55 509 237. 10

CELL\_HEADER\_ELM 32020 2

3 340 930. 15 510 036. 90

3 340 927. 90 510 032. 44

SHAPE\_ELM 32 021 5

3 340 875. 16 510 551. 42

...

四、DGN 设计文件分析及数据提取

Microstation 设计文件采用 Intergraph 标准文件格式 (ISFF), 亦称 IGDS 格式, 其文件结构如图2。整个设计文件由文件头、若干个 Element 块、设计文件结尾标记 EOD( End Of Design file) 3 部分组成。Element 块与 DGN 文件内的元素一一对应, 并按元素的生成顺序在设计文件中依次存放。每个 Element 块内有一个 WTF 指针变量, 该指针指向下一个 Element 块或 EOD 标记。

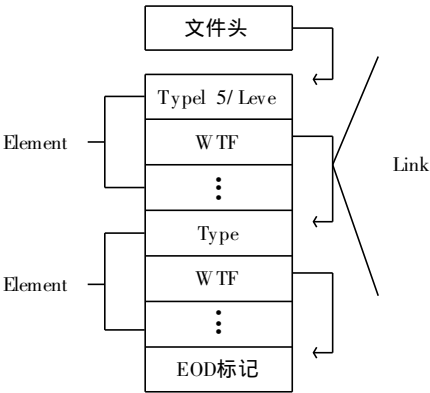


图2 DGN文件结构

Element 块实为一结构体(C/C++), 由固定部分和可变部分组成。

固定部分占用位于结构体头部的18个字, 定义了一些通用的数据项, 即第一个字定义为 Element header, 第二个字定义为 WTF, 第三至十四个字用于存放元素的范围等。每个字均为结构体, 以第一个字(Element header 为例, 其共占16位(1字=2字节=16位), C结构定义如下:

```
{
    unsigned level: 6; /* 元素所在图层, 占6位*/
    unsigned :1; /* 保留位, 一般为'0', 占1位*/
    unsigned complex: 1; /* '1'时为复杂元素的一部分, 占1位*/
    unsigned type: 7; /* 元素类型, 占7位*/
    unsigned deleted: 1; /* 是否为删除的元素, 占1位*/
}
```

其余17个字也有各自的C结构定义, 根据这些定义, 通

过 TCB 变量或 MDL 的内部函数即可获得当前元素的所在图层、颜色、线型、线宽、坐标范围、元素类型等基本信息。固定部分定义为

- 1 Element header
- 2 Number of W T F
- 3,4 X LOW
- 5,6 Y LOW
- 7,8 Z LOW
- 9,10 X HIGH
- 11,12 Y HIGH
- 13,14 Z HIGH
- 15 Graphics Group
- 16 Index to ATT.Link
- 17 Properties
- 18 Symbology

可变部分的内容则相当灵活,依元素类型而异,主要存放元素的几何特征等数据。对于复杂类型的元素,如 COM - PLEX\_STRING,COMPLEX\_SHAPE 等,可视为由一系列简单元素组成的双向链表,对该类元素信息进行提取时,可以使用元素描述符指针针对元素结构进行遍历,按检索到的简单元素分别逐个进行。

提取图形数据可以通过 MDL 提供的函数,如 mdlElement\_extract() 系列、mdlElement\_getSymbology()、mdlElement\_getProperties() 等。将元素的地址传递给它们后,将返回该元素的几何特征和图形特征。

由于数据交换文件不要求提供图形特征,而仅要求要素的分类编码,故还需要对元素上附着的属性数据进行提取,经过研究和测试,找出提取要素的分类编码的方法:将元素地址传递给函数 mdlMap\_firstElmdscrFeature(),返回一 int

数值,再传递该值给 mdlMapFeature\_indexFromMslink(),该函数可以从外部数据库中获取其分类编码在外部数据库中 RecordNo( int),根据该 RecordNo 利用 FCB 变量得到元素的分类编码特征。

五、实例说明

根据上述设计与分析,笔者采用 MDL 编制了一个数据交换接口模块,启动该模块后用户界面如图 3,用户可以选择输出的数据是否带高程,可对数据施加任意偏移操作,在指定目标文件名后,点击开始转入(转出)按钮即可完成 DGN 文件与外部数据交换文件之间的相互转换。外部数据交换文件可采用上述的自定义格式,也可采用 Mapinfo 的 MIF/MID 格式。



图 3 模块的对话框界面

使用 CPU 为 PIII 500,内存为 128 M 的微机对大小 6.3 M 的 DGN 文件做数据转换测试,将使用本模块的数据交换与使用 Microstation 自带的 DXF/DWG 的输入/输出( IO) 进行比较,结果如表 3,证明使用该模块进行数据交换具有速度快、线型完整、数据量小、分类编码特征不会丢失等优点。

表 3 比较结果表

方法	转换时间/( )	线型完整性	数据量/ M	分类编码特征是否丢失
使用本模块	15	完整	6.2	未丢失
通过 DXG( I/O)	220	破碎	46.2	丢失

六、结 论

采用要素的分类编码特征作为数据交换文件的内容,要素的属性数据主要以分类编码为主,尚未考虑拓扑关系,如不能满足要求时,可考虑选用国家地球空间数据交换格式

( CNSDTF)。在掌握了 Microstation 设计文件数据结构,并明确所使用的数据交换文件的格式和内容后,通过编程具体实现 Microstation 的外部数据交换不再困难了。

声 明

为适应我国信息化建设需要,扩大作者学术交流渠道,本刊已加入《中国学术期刊( 光盘版)》和“万方数据——数字化期刊群”,提供网络信息服务,作者著作权使用费与本刊稿酬一次性给付。如作者不同意将文章编入该数据库,请在投稿时说明。

《测绘通报》编辑部