

有关 ArcInfo 标准交换文件格式的探讨

喻 华

(沈阳市勘察测绘研究院,辽宁 沈阳 110015)

A Study of ArcInfo's Standard Exchange File Format

YU Hua

摘要: ArcInfo 是美国 ESRI 公司开发的大型地理信息软件包。E00 文件是 ArcInfo8.1 软件包的标准交换文件。本文阐述 E00 文件中的信息和记录及文件交换格式。

关键词: ARC; CNT; PAL; ANNO

一、前 言

ArcInfo 适合于地理信息方面的数据采集及数据挖掘,并可以将地理数据进行拓扑处理并与大型关系型数据库以 ODBC 的方式进行连接,以便于让决策者不必理解复杂的概念就能很容易地对拓扑后的地理数据进行查询和分析。ArcInfo 在 NT 平台上是标准的 COM(组件对象模型)结构,支持面向对象的数据结构及其管理,允许用户根据需要利用可视化的 CASE(计算机辅助软件工程)工具对数据类进行定义和扩展。支持内置的 VBA(Visual Basic for Application)开发环境,可对 ArcInfo 进行在线的客户化和二次开发。

二、E00 文件的格式

ArcInfo 软件包的基础文件格式采用二进制格式,不宜进行直接转换。因此,ESRI 公司推出了一种标准交换文件——E00 文件。它采用的是文本格式,以便于将各种自定义的数据格式导入到 ArcInfo 的系统当中。只要转换以后的 E00 文件格式正确,就可以使用 ArcInfo 工具包中所提供的软件插件将 E00 文件批量地转换成 ArcInfo 正常的文件格式。因为在 E00 文件中不仅记录了地物、地形的空间坐标,而且还记录着所有的属性信息以及拓扑关系信息,所以其格式有一定的复杂性。在 E00 文件的文本记录中,每行记录都不超过 80 个字符(第一行 81 个字符除外),如果其中某个信息超过 80 个字符,则必须写到下一行。一般来说,非空间坐标类数据从第十列开始,每个数据之间有 9 个空格。空间坐标

类数据从第二列开始,每个数据之间有一个空格。标识类(arc, cnt 等)数据从第一列开始。E00 文件信息分为若干个段落,分别为 exp, arc, cnt, lab, pal, tol, sin, eox, log, eol, mtd, eod, inf, eoi, anno 等段落。下面就 arc, cnt, pal, anno 等主要段落为例,对 E00 文件的格式加以详细阐述。

1. Arc 段

Arc 段记录着弧段的详细信息。它包括序号、弧段号、起点、终点、左多边形号、右多边形号、顶点数和具体的 x, y 坐标。我们以下面的内容为例:

ARC 2

1 14830002 2 1 2 1 2

3.5500000E+04 3.6470156E+04 3.5494785E+04 3.6471941E+04

2 14830243 2 3 1 2 2

3.5500000E+04 3.6470156E+04 3.5500000E+04 3.6456773E+04

3 14830003 1 3 2 1 4

3.5494785E+04 3.6471941E+04 3.5488742E+04 3.6474004E+04

3.5484660E+04 3.6462035E+04 3.5500000E+04 3.6456773E+04

4 14830004 4 4 3 1 5

3.5481824E+04 3.6454133E+04 3.5483289E+04 3.6459031E+04

3.5479602E+04 3.6460133E+04 3.5478117E+04 3.6455277E+04

3.5481824E+04 3.6454133E+04

.....

收稿日期:2002-07-23; 修回日期:2002-08-05

作者简介:喻华(1971-),男,辽宁沈阳人,工程师,现主要从事 ArcInfo 的开发工作。

第一行的“ARC 2”是段落的起始标识。第二行的第一个“1”为序号,“14830002”是弧段号,以后的数字依次为:起点号、终点号、左多边形号、右多边形号和顶点数。第三行是具体的 x, y 坐标,根据弧段顶点数的不同,具体坐标的数目也不相同。

2. Cnt 段

Cnt 段定义了多边形的重心。0 号为全体多边形的重心(以多边形面积为权作加权平均计算得到),0 号多边形以后的是每个多边形重心的坐标。我们以下面的内容为例:

CNT 2

```
0 3.5463996E + 04 3.6391227E + 04
1 3.5493301E + 04 3.6465418E + 04
1
1 3.5480711E + 04 3.6457141E + 04
2
1 3.5481719E + 04 3.6443445E + 04
3
.....
```

第一行的“ARC 2”是段落的起始标识。第二行是全体多边形的重心。第三行是具体的多边形的重心坐标,每个坐标均以 1 和一个空格开始。下面一行是多边形的编号。

3. Pal 段

Pal 段记录了多边形填充的详细信息。分为总体信息和详细信息两个部分。我们以下面的内容为例:

PAL 2

```
44 3.5318816E + 04 3.6256277E +
04 3.5500035E + 04 3.6474004E + 04
0 0 0 1 2 2
3 1 2 -2 3 2
0 0 0 4 4 3
.....
3 3.5484660E + 04 3.6456773E +
04 3.5500000E + 04 3.6474004E + 04
2 2 1 -3 3 1
-1 1 1
1 3.5478117E + 04 3.6454133E +
04 3.5483289E + 04 3.6460133E + 04
-4 4 1
3 3.5475586E + 04 3.6436770E +
04 3.5489266E + 04 3.6451082E + 04
7 6 5 -8 7 1
-5 5 1
```

```
2 3.5482523E + 04 3.6440648E +
04 3.5491371E + 04 3.6449152E + 04
-7 7 4 -6 6 1
.....
```

第一行的“PAL 2”是段落的起始标识。第二行分别为:多边形个数乘 2、所有多边形 x, y 坐标的最小值、所有多边形 x, y 坐标的最大值。再以下为多边形的排列关系:序号、结点 1、左多边形号、序号、结点 1、左多边形号。如果多个多边形相连,则填充多边形时按一个整体进行填充且其前序列皆为 0。如果整体的多边形在弧段的右侧,则相应弧段的序号为负。若与其他多边形有公共边,则舍去公共边那条记录,只记录最外围的弧段。每个弧段只记录一个结点,结点沿着整体多边形的最外侧依此列出,以第一个结点的前一个结点结束。

省略号以后的内容是多边形填充的详细信息。分别为:弧段数、弧段 x, y 坐标的最小值、弧段 x, y 坐标的最大值、弧段信息。弧段信息的排列方式为:弧段序号(负值)、结点 2、右多边形号、弧段序号(负值)、结点 2、右多边形号。若按此方法排列以后,得出的右多边形是其本身,则按弧段序号(正值)、结点 1、左多边形号的方式排列。

4. Anno 段

Anno 段记录着文字信息的坐标、文字大小、旋转等内容。旋转内容以坐标形式给出,以文字的坐标为基点,以文字大小的值 \times 旋转角度的 \sin, \cos 值得出坐标增量再加上基准坐标。我们以下面的内容为例:

ANNO

```
2.0000000E + 00 0.0000000E + 00 0.0000000E
+ 00
2.6550246E + 04 2.4542674E + 04
2.6550246E + 04 2.4542674E + 04
.....
```

第一行的“ANNO”是段落起始标识。其中:2.0000000E + 00 0.0000000E + 00 0.0000000E + 00 代表字符大小。如果字符大小大于 10,则表示为 2.0000000E + 01。2.6550246E + 04 2.4542674E + 04,代表字符的坐标。如果有旋转,则在坐标后面加上旋转内容。

三、结束语

通过前面的描述,我们就可以采用通用的开发工具如 VB, VC 等用纯文本的形式将自定义的文本格式转换为 E00 文件格式。