

基于 ArcGIS Engine 的数据格式转换

康 燕

(浙江有色测绘院信息中心, 绍兴 318000)

摘 要:本文主要讨论了基于 ArcGIS Engine 的数据格式转换方法, 此方法在海洋水色遥感中的等值线产品数据转换中取得了很好的应用。

关键词: ArcGIS Engine; 数据格式转换

1 引言

由于数据的来源、数据的获取方式和数据的处理等不同, 造成了目前 GIS 数据格式的多样性, 给数据的共享与互操作带来困难。Shape 文件作为 ESRI 公司 ArcGIS 系列软件的通用数据格式, 也是目前多数 GIS 软件产品能够直接读取或导入的一种数据格式。本文以中国海洋水色遥感等值线数据的转换来说明利用 ArcGIS Engine 进行数据格式转换的方法。目前的水色遥感等值线产品是以栅格图像文件的形式来存储和显示的, 这种形式不能很好地体现矢量数据的优点, 研究海洋水色遥感电子海图的数据格式转换, 是将等值线数据格式转换为 ArcGIS 的 shape 格式文件以便数据的后续分析应用。

2 数据格式介绍

2.1 水色遥感等值线数据格式

目前的水色遥感等值线数据文件中等值线的数据格式解译后可以表示如下:

总共的等值线的个数 n

第一条等值线 等值线的数值 颜色值
经度, 纬度

.....

END

第二条等值线 等值线的数值 颜色值
经度, 纬度

.....

END

第 n 条等值线 等值线的数值 颜色值
经度, 纬度

.....

END

END

从以上等值线数据的格式可以看出, 它的等值线是由一系列的点构成的, 每条等值线包括编号、

等值线的值、等值线在 bmp 图像上的颜色信息三个属性。

2.2 shape 文件的数据格式

Shape 文件是二进制文件, 能够同时存储空间图形数据和属性数据, 至少由 shp, dbf, shx 三个文件组成, 分别存储空间、属性和前两者的关系, 是 GIS 中比较通用的一种数据格式。

2.2.1 .shp 文件

.shp 文件格式也称为主文件, 包括一个固定长度的头文件和一个变化长度的记录。每一个变长记录又有一个固定长度的记录头和变长记录组成。

固定长度的头文件长度为 100bytes, 在一个 shape 文件中的 shape 必须为 Point、Arc、Polygon、MultiPoint 四种基本类型中的一种。变长记录头文件有固定的长度, 为 8byte。

2.2.2 .dbf 文件

.dbf 文件包含想要得到的特征属性或其他能被连接的属性的关键值表。它的形式是一个标准的 DBF 文件, 有以下四个要求:

(1) 文件名必须与 shape 主文件和索引文件一致, 而扩展名是 .dbf;

(2) 表中必须包含每一个 shape 特征的一条记录;

(3) 记录的顺序必须与在主文件中 shape 特征记录的顺序一致;

(4) dbf 文件头中的年份值必须晚于 1900。

2.2.3 .shx 文件

.shx 文件(索引文件)是由一个长度为 100 字节的头文件和一系列长度为 8 字节的记录组成的。索引文件的头文件组织形式与上面的主文件文件头的描述一样, 索引文件中的第 n 个记录存储着第 n 个记录在主文件中的偏移量和内容长度。

3 数据格式转换

3.1 数据之间的对应关系

Shape 文件是按照地物类型来组织的。地物类

型可以简单地分为点(Point)、线(Line)、面(Polygon)三种类型,一个 Shape 文件中只能包括一种类型的地物,即只能包括点、线或面状地物中的一种。点状地物存储的是一组坐标(X,Y),线包括简单线和多段线,存储的是一系列的坐标串($X_1, Y_1, X_2, Y_2 \dots$),面状地物存储的是一系列封闭的坐标串($X_1, Y_1, X_2, Y_2, \dots, X_1, Y_1$)。

原始文件中每条等值线上的一串坐标对应转化后的一条等值线,相应的等值线的编号、数值以及颜色值等信息将作为转化后 shape 文件中每一条等值线的属性值。

3.2 数据格式转换的实现

初始化——检查 license,初始化应用程序。用 Engine 开发包创建的 Engine,必须要有 ArcGIS Engine 的运行许可,初始化的目的就是使开发的应用程序可以到达 license 许可的功能,也能保证它在分发的任何机器上成功地运行。

创建 shape 文件,主要包括:设置一个文件,并在此文件中创建特征空间,然后对特征空间进行设置。设置的内容包括特征空间的坐标系、范围,特征的类型(这里是 Point 和 PolyLine)以及特征的属性值的相关信息(比如属性的数据类型、属性名称以及所占用的空间大小等信息)。然后逐行读入原始数据,每读完一条等值线上的点,便将这些点收

集成一条等值线,按规定的特征空间将数据写入对应的 .shp 文件,并将等值线等对应的属性数据写入 .dbf 文件,以及将其对应关系的信息写入 .shx 文件。每个文件的数据处理完后,释放内存空间,然后处理下一个文件,可进行批量处理。

4 结果实例

将转换好的 shape 文件插值成以 shape 范围为参考的平面内连续变化的栅格数据。转换前后表现形式的对比如图 1 所示。原等值线产品的表现形式是一幅 bmp 图像,它是一点点地将数据画在栅格图像上,不同的颜色代表不同的线值,但这种画在图像上的等值线并不是真正意义上的矢量等值线,因为它并不是将每一条等值线作为一个对象来存储的。从放大的图像上我们可以看出,它是按像素来显示的,这样不能体现出等值线作为矢量数据的优越性。经过格式转换后的 shape 文件是将空间对象按照点、线、面的方式进行存储的,shape 文件中的等值线是以每一条等值线为对象来存储和显示的,因此,从放大的图像上看它仍然是一条曲线,而不像在 bmp 图像上看到的那样是一个个的像素。将转换好的数据按照邻近插值的方法进行插值,成为在 shape 文件的最大最小范围内连续变化的 GRID 栅格数据。

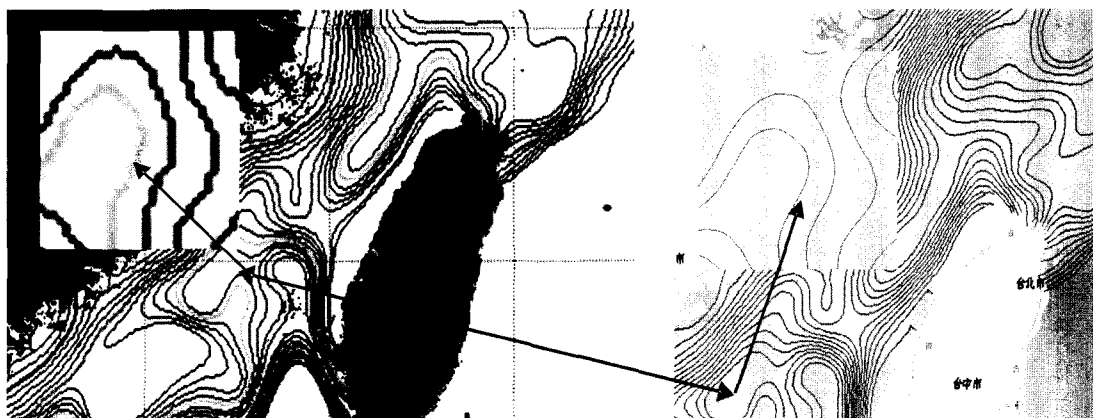


图 1 等值线数据格式转换前后局部放大

5 结束语

本文介绍了等值线矢量数据格式与 ArcGIS 的 shape 文件格式,并将海洋水色遥感原等值线矢量数据格式转换为 ArcGIS 的 shape 文件格式的程序流程及具体实现过程,给出了转换前后文件的表现

形式。

在此研究的数据格式转换是没有信息损失的格式转换,转换后的数据格式比原有的数据格式有着更加丰富的表现形式。