

面向 MapGIS 和 ArcGIS 的跨平台数据转换研究

文 缔 尧, 周 娅, 秦 兴 国, 吴 清 霞, 贺 正 红

(桂林电子科技大学 计算机科学与工程学院, 广西 桂林 541004)

摘 要: 深入讨论了 MapGIS 和 ArcGIS 两种 GIS 软件的数据格式, 并且分析 ArcGIS Geodatabase 格式和 shapefile 格式的差异, 设计并实现了从 MapGIS 到 ArcGIS 平台的数据无损自动转换的方案, 有效地解决了原有数据转换方案中信息丢失、注记无法渲染等问题. 通过对比实验结果表明, 该方案在从 MapGIS 到 ArcGIS 平台的 Geodatabase 数据转换相比于传统 shapefile 文件格式具有很高的准确率和较好的时间性能.

关键词: MapGIS; ArcGIS; Geodatabase; shapefile; 无损自动转换

中图分类号: TP311. 51

文献标识码: A

文章编号: 1000-7180(2016)10-0097-05

Cross-Platform Data Transformation Rresearch for MapGIS and ArcGIS

WEN Di-yao, ZHOU Ya, QIN Xing-guo, WU Qing-xia, HE Zheng-hong

(College of Computer Science and Engineering, Guilin University of
Electronic Technology, Guilin 541004, China)

Abstract: This paper deeply discusses the data format of MapGIS and ArcGIS software, and analyze the differences of ArcGIS Geodatabase format and shapefile format, designs and implements the scheme of nondestructive and automatic conversion of data from the MapGIS platform to ArcGIS platform, effectively solves the problem in original data conversion scheme that information is missing and note can't render. Through the contrast experiment shows that the scheme has higher accuracy and better time performance from MapGIS to ArcGIS Geodatabase of data transformation process than the traditional shapefile format.

Key words: MapGIS; ArcGIS; Geodatabase; shapefile; nondestructive automatic conversion

1 引言

在地理信息系统中, 我国国土资源部门中绝大多数单位使用国产 MapGIS 软件, 很多 GIS 数字资料的数据加工、建库工作都是基于 MapGIS 平台来加以实现^[1]. 然而在数据存储管理、空间分析、数据共享及发布等功能上与国际主流 ArcGIS 平台还存在一定的差距. 如今, 许多国土资源信息的数据共享、数据应用都是在 ArcGIS 平台中进行的, 因此这两个 GIS 平台间的数据无障碍共享必不可少^[1]. 如何充分利用现有积累 MapGIS 平台数据自动生成符

合标准规范的 ArcGIS 图件成为亟待解决的问题.^[2]

现有的主要数据转换方法有: ①通过利用 E00、DXF 等中间格式文件^[3]; ②通过直接读取 MapGIS 加密二进制格式文件^[4]; ③通过半自动转换技术^[5]; ④通过第三方商业转换软件, 如 FME^[6]. 然而这些转换方法常常需要繁琐的人工操作, 并且经常会出现图形参数、属性、注记丢失, 注记无法渲染等问题, 很难达到理想效果^[7].

本文考虑到读取 MapGIS 9 接口的丰富和利用 ArcGIS 全新的空间数据库 Geodatabase 的优

收稿日期: 2016-01-05; **修回日期:** 2016-02-25

基金项目: 广西教育厅高校科技项目(2013YB095); 广西信息实验科学中心重点项目(20130111); 广西教育厅一般资助项目(2013YB051); 桂林电子科技大学研究生创新项目(GDYCSZ201465); 广西教育厅科研项目(201106LX191)

势^[7]. 结合 MapGIS K9 全新的二次开发接口和 ArcGIS 空间数据库 Geodatabase, 实现 MapGIS 数据文件向 ArcGIS 空间数据库 Geodatabase 和工程文件 MXD 的无损转换, 达到数据共享的目的, 在保证图件准确度的同时最大限度地减少工作量. 与此同时, 本文采用 shapefile 文件格式进行存储对比, 更加充分证明 GeoDatabase 文件存储形式不仅在图面渲染上还是在时间效率上均具有独特的优势.

2 平台数据分析

在地理信息系统中, MapGIS 和 ArcGIS 这两个 GIS 平台的数据组织方式差别较大, 要实现跨平台数据无障碍共享, 首先需要对两种 GIS 平台下数据存储形式及表达方式进行分析. MapGIS 的数据文件包括点文件、线文件、面文件, ArcGIS 的数据文件包括点文件、注记文件、线文件、面文件, 不同的 GIS 平台对数据的底层存储形式也具有较大差异, 存储模型如表 1 所示.

表 1 MapGIS 和 ArcGIS 的格式对照

文件格式		存储内容
MapGIS	MPJ 工程文件	*, wt, (点文件) *, wl(线文件), *, wp(区文件)的管理文件
	图层文件	*, wt, (点文件) *, wl(线文件), *, wp(区文件)数据信息
ArcGIS	MXD 地图文档	存储载入的图层名称/路径, 图层的显示属性等索引信息
	Layer 图层文件	存储图层的渲染信息与 gdb 或 shp 文件配合使用
	Shapefile 文件	主文件(. shp)、索引文件(. shx)、数据文件(. dbf) ^[7]
	Geodatabase 文件	一种空间数据模型, 对数据标准要求严格

两个 GIS 平台都集成了面向对象思想, 对不同的文件类型进行分类管理, 形成不同的文件和要素类. 在底层存储结构中 MapGIS 的空间数据和属性数据采用二进制的形式进行存储, 相应的 ArcGIS 的 shapefile 文件也采用二进制形式进行存储, 但其对属性数据的管理采用 DBF 数据表的形式进行存储. 其中本文采用的 ArcGIS GeoDatabase 数据模型是一种“关系-对象”模型, 引入空间数据库的机制对数据进行组织管理^[8].

3 关键转换技术分析

3.1 系统框架

本系统的整体过程, 首先解析 MapGIS 文件数据, 其次将图层数据创建为 ArcGIS 要素文件保存

在 Geodatabase 中, 最后利用 Geodatabase 文件中的字段渲染图层, 生成最终的 MXD 工程文件. 具体过程如图 1 所示.

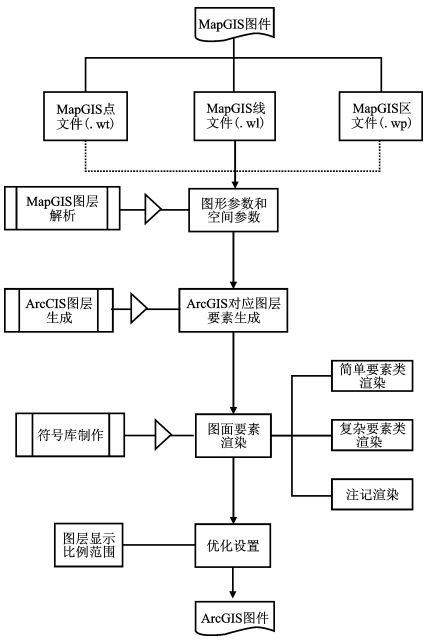


图 1 MapGIS 向 ArcGIS 数据转换框架

3.2 数据转换

3.2.1 MapGIS 数据读取

数据转换主要针对于空间数据及属性参数的转换, 数据无损转换的根本在于读取 MapGIS 源数据的准确性. 由于 MapGIS 中数据的索引信息均存储于 MPJ 工程文件, 因此根据地图类 XMap 读取 MPJ 工程文件, 得到地图的名称和组图层数目, 遍历地图中的组图层. 其中, 点组图层包含有注记图层 ann 和标准点图层 point, 线和区组图层包含有一个与之对应的线和区图层. 基于 MPJ 中的索引信息分别读取 MapGIS 中图层信息, 各类图层文件均包含有图层的属性参数、图形参数和空间参数. 当然, 实体类型不同, 上述几方面的数据就有不同的特点. 点实体: 空间数据为 (x, y) 和 D_DOT, 图形参数为 PNT_INFO, 缺省属性为 ID; 线实体: 空间数据 $(x_0, y_0) \cdots (x_n, y_n)$, 图形参数 LIN_INFO, 缺省属性为 ID 和长度; 区实体: 空间数据为边界弧, 图形参数 REG_INFO, 缺省属性为 ID、周长、面积. 根据 MapGIS 提供的二次开发接口就能够解析各类型的数据.

3.2.2 ArcGIS 数据存储

(1) shapefile 文件格式存储

在 ArcGIS 平台下, 由于 Shapefile 文件数据格式开源, 基于 Shapefile 文件的各种转换方案比较成

熟,所有之前的大量转换研究都是使用该文件作为目标文件^[9]. MapGIS 的每一个图层都对应一个 Shapefile 文件,这样生成的 Shapefile 文件也对应点、线和区三种类型. Shapefile 中没有与 MapGIS 对应的图形参数的概念,而是通过属性来处理,因此需要将 MapGIS 的图形参数转换为 Shapefile 的属性,与 MapGIS 下原有属性数据一同存储. 对空间参数的存储,先需要根据解析的数据构建子图对应的空间参数对象,点需要构建坐标,线需要构造组成线的各线段,区则需要构建围成区的所有闭合线,最后还需要对生成的图层做渲染^[10]. 创建 Shapefile 文件生成可分为属性信息存储、空间信息存储和子图渲染三部分处理. 详细算法流程如下所示.

算法 1 MapGIS 源数据存入 Shapefile

输入:MapGIS 图层信息
输出:Shapefile 图层要素类
①初始化 shapefile 工作空间和 ArcGIS 字段 IFields
②设置图层空间参考和对应类型并创建要素图层
③通过使用 IFeatureCursor 和 IFeatureBuffer 插入对应字段值
④返回要素类
算法 2 图层空间信息
输入:MapGIS 中图层对应的空间信息
输出:Shapefile 图层的空间信息类
①if 点或注记:设置 PointClass 坐标点
②else if 线类型:将相邻点构成线段 Line,最终形成线类
③else if 区类型:根据点构建围成区的各条线并加入区类型空间坐标(区分内圈和外圈)
④End if:返回图层空间空间信息类

算法 3 Shapefile 图层渲染

输入:图层要素类 IFeatureClass
输出:渲染后的要素类
①生成渲染对象 IUniqueValueRenderer
②判断图层类型并生成渲染符号(点 IMarkSymbol、线 ILineSymbol、面 IFillSymbol)
③符号相关参数设置(颜色、大小、角度、style 匹配等)
④为图层设置渲染对象并保存
⑤返回渲染后的要素类
(2)GeoDatabase 文件格式存储

在 ArcGIS 平台,空间数据库 Geodatabase 是用来存储 GIS 数据的最新技术,由于其在数据存储、共享、完整性检查方面具有各种优势,因此本文采用 Geodatabase 来管理地理数据. 数据转换中,在 Geodatabase 存储的是 MapGIS 解析的属性参数、图形参数和空间参数. 与 Shapefile 文件存储类似,将图形参数通过属性来处理,与 MapGIS 下原有属性数据一同存储. 对空间参数的存储,先需要根据解析的数据构建子图对应的空间参数,点需要构建坐标,线

需要构造组成线的各线段,区则需要构建围成区的所有闭合线. 与 Shapefile 文件存储的区别在于对注记的处理,在 Shapefile 文件中将注记作为子图图层处理,在 Geodatabase 中有特定的注记图层 IAnnotationLayer 来处理,这样在注记的显示时,就会更加的完善. 详细算法流程如下所示.

算法 1 MapGIS 源数据存入 Geodatabase

输入:MapGIS 图层信息
输出:GDB 中图层要素类
①初始化 GDB 工作空间和 ArcGIS 字段 IFields
②If 点、线、面图层:设置空间参考和类型并创建要素类
③else if 注记图层:创建 IAnnotationLayer 图层并创建对应要素类
④通过 IFeatureCursor 和 IFeatureBuffer 插入字段对应值
⑤返回要素类

算法 2 Geodatabase 图层渲染

输入:图层要素类 IFeatureClass
输出:渲染后的要素类
①If 点、线、面要素,创建渲染对象并生成对应渲染符号,设置符号参数并存储
②else if 注记要素,转换为 IAnnotationFeature 注记要素类并创建文本符号 ITextSymbol,设置文本符号相关参数(通过和指示上下标),为图层设置渲染对象并存储
③返回带渲染的图层要素类
(3)MXD 文件格式生成

存储在 Geodatabase 与 Shapefile 中的仅为图层数据,ArcGIS 需要 MXD 工程文件对图层文件进行组织管理,根据给定的文件名和地图文档类 IMapDocument 创建 ArcGIS 地图工程文件 MXD,此时地图工程中就生成了一幅默认地图,得到该工程对应的地图类 IMap,最后将所有图层加入到地图类 IMap,保存得到完整的图幅,按照给定的名称给工程文件命名,存储整个地图文件.

3.3 地图设置调优

ArcGIS 中全图复位显示时比较复杂,需要根据实际的地质图坐标系和比例尺进行选择性的优化调整.

Shapefile 文件格式,由于注记图层初始打开并不会显示,需要通过图层设置标注后才会显示注记信息,此时的注记为默认,同时还会出现部分注记丢失,直接影响了地图的整体效果,这也是 Shapefile 文件的不足之处.

Geodatabase 图层在 ArcMap 中显示会出现部分图元丢失的现象,而图层的属性表中包含有完整的图元信息,同时将整幅图打开会出现图层偏移的情况. 出现这些情况是由于地图的坐标系不正确引

起,由于不同的坐标系对应的范围不同,当图层的范围大于设置的坐标系范围就会出现图元丢失,不同的图层也会出现坐标偏移最终导致地图部分移位.因此只要设置对应坐标系则能使数据无损还原显示.

由于 Geodatabase 支持注记要素类,生成对应的注记图层,因此就可以对注记的大小进行固定,同时也可以对每个注记进行单独设置,例如出现压盖的调整、字体上下标的处理、某些字体旋转角度、横竖排的效果等.不同的注记图层以不同的比例尺固化,这样就能克服地图文档中只能设置单一参考比例尺的缺点.

4 实验结果对比分析

4.1 实验 1(图幅准确性对比)

4.1.1 实验数据

该系统的工程性应用由 MapGIS 转换至 ArcGIS 的数据超过 5 000 幅,其中包含各种大小的文件,主要集中在 10 MB 至 100 MB 之间.本次对比主要选取特征比较明显,图幅相对比较复杂的作为实验结果验证.

4.1.2 结果分析

利用最终完成的转换中间件,将项目合作地矿局提供的 MapGIS 地质图转换到 ArcGIS 平台下.由于水文图信息丰富,数据比较密集,下面以水文图为例说明,该地图经过合作地矿局工作人员处理,不会危及到地矿部门的信息安全.

MapGIS 水文原图部分截图:选取的信息比较丰富的图件部分和注记部分.如图 2 所示.

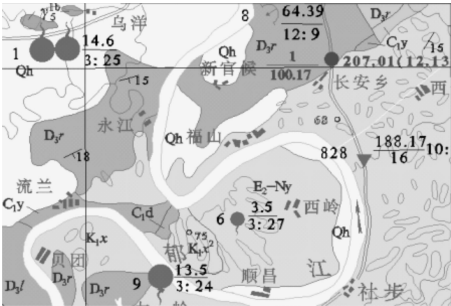


图 2 MapGIS 图件与注记部分

实验 1:转换成目标文件为 Shapefile 格式.通过加载水文原图和其对应的子图库信息,利用完成的中间件转换.对注记图层通过标注来显示注记信息,得到整体的地图效果.截取与原图对应的图件部分和注记部分,如图 3 所示.

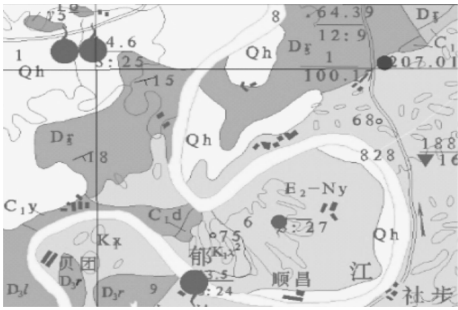


图 3 转换后 ArcGIS shapefile 图件与注记部分

实验 2:转换成目标文件为 Geodatabase 格式.通过加载容县文件原图和子图库信息,利用完成的中间件完成该工程文件的转换.截取与原图对应的图件部分和注记部分,如图 4 所示.

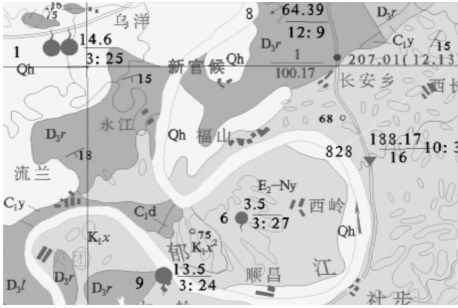


图 4 转换后 ArcGIS Geodatabase 图件与注记

将转换为 Shapefile 文件与转换为 Geodatabase 文件图件部分的精确度进行对比,得到如下的结论.

(1)图件中 Shapefile 格式图件与 MapGIS 原图相比,出现了注记丢失、子图偏移的问题.

(2)注记中 Shapefile 格式图件与 MapGIS 原图相比,注记的字体、颜色、排列方式(横竖排)都是默认,并且注记的位置出现偏移.

(3)图件中 Geodatabase 格式图件与 MapGIS 原图相比非常接近,注记显示的完整、子图位置偏差较小,图层的填充颜色与线型与原图一直.与 Shapefile 格式图件相比,图件的整体效果更接近原图,基本达到数据的无损转换与还原.

对比得到 Geodatabase 格式转换比 Shapefile 格式转换效果好,原因在于 Geodatabase 是一种采用标准关系数据库技术来表现地理信息的数据模型,对数据的约束更加完整,最重要的是对注记的支持.解决了 shapefile 文件对注记的不可控问题,并且 Geodatabase 支持对单个要素的编辑,这样就能够对图层中的子图单独处理,固定其位置,这样就解决了在 Shapefile 文件下子图偏移的问题.

4.2 实验 2(时间性能对比)

4.2.1 实验数据

由于地理数据的密集性,在数据转换过程比较耗时.本次实验的机器配置为双核 Intel(R) Xeon (R)、2.53 GHz,内存为 8 GB.在工程性应用转换中的图幅中随机选出 50 幅作为时间性能测试数据,其中的数据均为在 MapGIS 下手工绘制得到的,这些工程数据大小集中在 10 MB 到 100 MB 之间.

4.2.2 结果分析

分别采用 shapefile 文件和 GeoDatabase 文件格式进行转换,得到最终时间效率值.总结得出 Geodatabase 格式转换方案与 Shapefile 格式转换方案的时间对比.如图 5 所示.

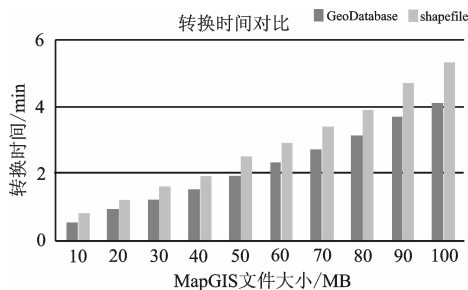


图 5 不同文件类型转换时间性能对比

从图 5 中得知,随着文件大小增大,Geodatabase 转换的时间效率比 Shapefile 越来越明显,出现这种情况的原因:Shapefile 文件在是以磁盘文件形式进行存储,而 Geodatabase 文件是以数据库形式进行存储,空间地理数据一般存放在大型的 binary object 中,关系型数据库对数据插入、读取性能优于文件的 IO 读写.

5 结束语

本文通过 GIS 平台提供的二次开发接口,结合两个 GIS 平台的特点,开发完成了平台数据的精确转换中间件,完成了地质数据的共享和跨平台利用.

ArcGIS 目标文件选择的是全新的空间地理数据库 Geodatabase,因为 gdb 文件对数据的完整性检测更加完善,最重要的是对注记的支持完善.在 Shapefile 文件格式中,不支持注记,使用标注来处

理注记,这样就会出现元素丢失、旋转角度、颜色等无法渲染的问题.使用全新的 Geodatabase 后能够很好地解决这些问题.目前使用 Geodatabase 是转换中的先例,对以后的读者会提供更多的参考.

本方案在对地图渲染时,需要提前准备好在 ArcGIS 平台下与 MapGIS 地图对应的符号库与颜色库.目前符号库与颜色库没有统一的标准,不同地图对应的符号会存在重复,未来的研究可以提供一个统一规范的符号库和颜色库,并同时管理这些现有的符号和颜色数据.

参考文献:

[1] 范晓明. 跨 GIS 平台数据格式与地图符号转换[D]. 北京: 中国地质大学,2009.

[2] 郭豫宾,何政伟,许辉熙. 应用 MapGIS 与 ArcGIS 构建矿产规划数据库[J]. 测绘与空间地理信息,2009,31(5):127-129.

[3] 麦柳妍. MapGIS 向 ArcGIS 数据转换的若干问题初探[J]. 测绘标准化,2005,20(4):13-16.

[4] 冯杭建,麻士华,刘伟宏,等. MapGIS 到 Geodatabase 数据自动批量转换实践研究[J]. 测绘科学,2007,32(3):181-185.

[5] 张龙,汪新庆,刘岳,等. 一种 MapGIS 向 ArcGIS 地图数据转换的方法[J]. 测绘科学,2014,39(7):146-150.

[6] 陈影,程耀东,闫浩文. 利用 FME 进行 GIS 数据的无损转换[J]. 测绘科学,2007,32(2):75-77.

[7] 方坤,吴信才. GIS 矢量数据跨平台转换研究[J]. 测绘通报,2009,54(8):60-63.

[8] Michael Zeilar. ESRI Modeling Our World[M]. America: Esri Press,2004.

[9] 彭晶晶,戴爱德. MAPGIS 数据向 ARCGIS 转换研究[J]. 国土资源信息化,2008,8(1):19-25.

[10] 黄云. MapGIS 向 ArcGIS 数据格式转换方法探讨[J]. 地矿测绘,2010,26(1):33-35.

作者简介:

文锦尧 男,(1992-),硕士研究生.研究方向为数据挖掘、地理信息系统及应用. E-mail:jazze1114@163.com.

周 娅 女,(1966-),教授.研究方向为数据库、地理信息系统及应用等.