

区域地质图空间数据建库质量控制的 MapGIS 分析

马瑜宏, 王纪存, 谢颂诗

(山东省地质调查院, 济南 250013)

摘要: 空间数据库质量的优劣直接影响到 GIS 应用分析结果的准确性与可靠性。因此, 质量控制是区域地质图数据库建设的内在要求, 而由于所有的空间基础信息在产生时都存在误差, 所以, 空间数据库的质量控制必须严格贯彻在数据库建设流程的每一个环节。本文以 1:5 万山东省区域地质图空间数据库的建设流程为例, 通过对基于 MAPGIS 建设区域地质图空间数据库过程中的图件预处理、矢量化、误差校正、拓扑重建、属性输入和投影变换等过程中可能产生的质量问题进行总结、分析, 探讨应采取的解决办法和预防措施, 以实现数据库建设整体质量控制, 最大限度地减少空间数据误差, 提高作图效率与空间数据的精度和准确度, 更好地为社会公益性应用提供服务。

关键词: 空间数据库; 质量; MapGIS

DOI: 10.3724/SP.J.1047.2011.00758

1 引言

地质图空间数据库, 系利用 GIS 技术将以图件为基础的地质信息即传统的文字报告及图件数字化, 从而为基础地质研究、国土资源开发利用、矿产资源评价、国民经济建设、区域规划制定、地质环境保护等应用领域提供有效的数字化信息服务, 实现全国基础地质数据信息的社会化共享, 提高地质信息的利用程度和使用价值。

但是, 空间数据库的质量控制, 关系到空间数据库建设的成败。由于所有的空间基础信息在产生时都存在误差, 鉴此建立的空间数据库在蝴蝶效应影响下误差就可能会被进一步放大, 因此, 空间数据库的质量控制必须严格贯彻在建库流程的每一个环节。下面就地质图空间数据库建库流程中应注意的一些关键问题予以解析, 以实现渐趋完美的质量控制。

2 数据建库中质量控制问题分析

2.1 数据源质量问题

完整、齐全而又有效的第一手资料是建立区域地质空间数据库的前提, 也是对区域地质空间数据进行空间分析的基础。

地图是最主要的数据源, 收集到的基础地质图一般为纸质折叠图。纸质图由于折叠和保存时间长, 折痕可能会非常深刻, 而且整个图件可能会产生变形, 甚至破损, 这些都会影响到所采集信息的质量。扫描图件时, 如果选择的分辨率太低, 也不能保证栅格图像的精度要求, 不利于操作, 从而影响数据的建库质量。

2.2 数据矢量化中的精度问题

用 MapGIS 软件进行矢量化的过程中, 影响数据精度的因素有:

(1) 系统参数设置。结点/裁剪搜索半径的参数设置会影响拓扑关系的正确性。拓扑错误检查后, 有些线或区发生变化, 就跟这个参数值的设置有关。拓扑会把长度小于半径值的线或弧删除掉, 线或弧上两点距离小于半径的点也会进行结点平差合成一点。自动剪断线后, 剪断后的线与原来的线有偏移, 也与这个参数值有关。结点/裁剪搜索半径设置过大会导致线上定位点一定程度的偏移。

(2) 图像放大倍数。矢量化过程中, 图像的放大倍数会直接影响到点与线的采集精度, 放大倍数过小会导致采集数据与原始数据之间发生较大的偏移, 影响精度, 而放大倍数过大也不利于数据的采集。

收稿日期: 2011-07-12; 修回日期: 2011-12-01.

作者简介: 马瑜宏(1981-), 女, 本科, 工程师, 主要从事地质编图与地理信息系统应用工作。

E-mail: wendy11400219@163.com

(3) 结点关系的建立。线遇结点必须断开,单线河遇双线河、双线河的左右岸遇地质界线时要建立结点关系。没有正确的结点关系,拓扑造区之后会出现线弧不一致或线面不套合的错误。

(4) 原图错误。原始图件由于填图人员的操作误差和当时条件所限,难免存在错误,例如高程点与周围等高线值不符,推测断层表示为实测断层,地质代号与图案花纹表示不符等,如果照搬原图,不将错误修正,就会将原图错误延续到数据库中。

2.3 拓扑误差问题

拓扑关系是空间分析的基础,能清楚反映实体之间的逻辑结构关系,有助于空间要素的查询,对于数据处理和 GIS 空间分析具有重要的意义。只有建立了拓扑关系,才能进行空间分析和统计等。拓扑查错是拓扑关系处理的关键步骤,只有数据规范,无错误后,才能建立正确的拓扑关系。

拓扑错误主要有重叠坐标点、悬挂弧段、弧段自相交、重叠弧段、结点不封闭等,这些问题主要是由矢量化不严格造成的。如图 1 所示,图中的悬挂弧段,在拓扑成区之前都应删除。

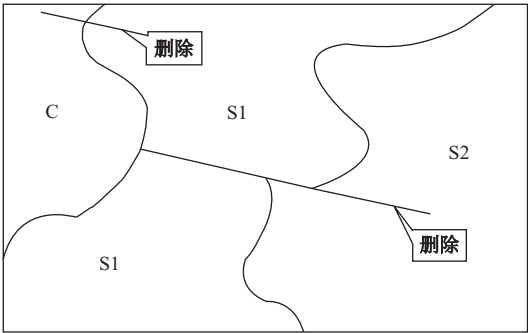


图 1 悬挂弧段
Fig. 1 Map of dangling arcs

2.4 误差校正中的质量问题

在数字化的过程中,由于原图纸张变形、扫描仪精度和手工操作的失误等因素,使输入后的图形位置与实际图形所在的位置往往存在偏差。对于这些图元偏差,有的经过简单编辑、修改后即可满足精度要求,而有的由于位置偏移较大,即使经过简单的编辑、修改也很难达到实际要求的精度,此时即产生了所谓的变形或畸变。对于出现变形或畸变的图形,就必须经过误差校正,才能使之满足实际要求。误差校正时影响数据质量的因素有:

(1) 校正控制点的精度。精度过低会降低数据质量。

(2) 误差校正实际控制点的采集。控制点的分布不合理或者控制点太少,会产生较大的误差变形,同时,实际控制点与理论控制点的采集顺序号不一致,也会导致校正后的文件发生变形。

(3) 利用 MapGIS 误差校正功能进行校正时,选择不同的校正方法,也会使校正后所形成的文件之间存在一定偏差。

2.5 属性数据输入误差问题

属性数据是空间数据的重要组成部分,描述空间实体的性质。地理部分的属性内容比较简单,可以在 MapGIS 软件下录入,而地质部分的属性内容相对复杂,可以在 EXCEL 中编辑并保存为 .dbf 文件,再通过关键字段与图形数据实现连接。

属性错误主要指属性结构错误和属性内容错误。由于区域地质空间数据库的图层文件和属性结构条目很多,在录入过程中难免会有错漏,字段名称、字段长度和类型可能会输入错误。属性内容错误包括全半角不分、地质代号大小写错误、图形与属性不一致、图元编号不唯一等。

2.6 投影变换的误差问题

目前,比较常用的地图投影方法有墨卡托投影(等角正切圆柱投影)、高斯-克吕格投影、UTM(等角横割圆柱投影)和兰伯特等角圆锥投影等。不同的投影方法实现不同图件的需要,因此,图形数据经常需从一个地图投影坐标系统转换到另一个投影坐标系统,这就需要对数据进行投影变换。

投影变换误差的产生是由于没有拷贝 TIC 点或投影参数设置错误。MapGIS 提供了 TIC 点的操作功能,用来确定用户坐标系和投影坐标系的转换关系。TIC 点实际上是一些控制点,即用户已知其理论值的点。在进行文件投影变换时,至少得输入 4 个 TIC 点,否则结果文件的位置或图形参数将出现错误。

3 建库质量控制中的 MapGIS 分析应用

3.1 数据源质量控制

对数据源资料进行预处理是数据源质量控制

的有效方法。

为控制误差,扫描前首先应将纸质地质图的折痕予以处理,用玻璃板或其他平板工具将图件压平整,尽量减少图件变形对数据质量的影响。

而在扫描时,图件扫描范围应覆盖全图,无漏洞和扫描带错位;图像应清晰,颜色偏差小,断线和粘连少,原图控制点、公里网和经纬网交点清晰。由于扫描精度越高,文件就越大,运行起来速度就越慢,不利于操作,因此,一般选择 300dpi 的分辨率,既能确保栅格文件的精度要求,又便于操作。对于栅格文件,应保存为 tif 格式,若超差则重扫新描。扫描后,再用 photoshop 软件将其旋转水平,以便于矢量化。

对原始图件中发现的明显的标记错误、线划错误及缺失、设色错误以及地质体时代或属性标注错误等非观念认识差异造成的错误,经地质专家认定后予以修改并在日志中记录。

3.2 矢量化的精度控制

矢量化过程中的精度控制既包括技术控制,也包括技巧控制。其中技术控制主要体现在结点/裁剪搜索半径的参数设置、图像放大、建立结点关系和修改原图错误等方面,而技巧控制主要体现在注意矢量化顺序和图例板的运用等方面。

(1)设置合适的系统参数。一般将结点/裁剪搜索半径设置为 10^{-9} ,由于数据库一般是多人操作共同完成,因此,每个作业人员的系统参数设置必须一致。系统参数设置如图 2 所示。

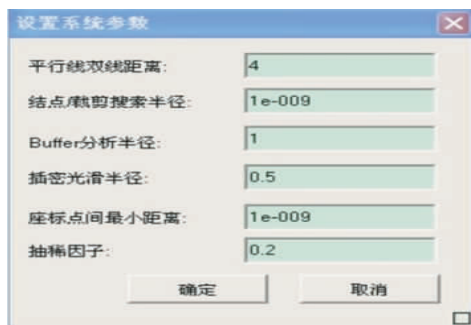


图 2 设置系统参数

Fig. 2 Parameters input in the data processing

(2)选择恰当的图像放大倍数。矢量化时一般宜将图像放大到 80-100 倍左右,同时,图像放大处理时还应严格按光栅中心走线,尽量保持线条光滑。

(3)建立正确的结点关系。结点关系的建立一律采用 F12 捕捉线头的方式,使结点处尽量与实际相符,同时还应尽量避免端点回折,不要产生超过 1 毫米的无用线段。

(4)修改原图错误。矢量化过程中应将原图错误改正,并作好记录。

(5)注意矢量化顺序。由于图件信息包括地理、地质和整饰 3 部分内容,为便于修改与管理,矢量化时应尽量遵循一定的顺序原则。一般来讲,对于地理、地质和整饰 3 部分内容宜分别进行分层、分要素,然后按照图层顺序从左到右、从上到下进行矢量化。其中地理内容可按照图幅基本信息→等高线→水系→交通→居民地→境界→其他的顺序,地质内容可按照断层→脉岩→第四系地层→侵入岩→火山岩→非正式地层→其他地层的顺序进行矢量化。而对方向性有特殊要求的水系等地理内容则应严格按照从上游到下游的方向进行矢量化。

(6)运用图例板。由于地质图空间数据库一般都有统一规定的图元参数和属性,涉及到的点、线、区图元参数内容较多,因此,矢量化之前,先按照要求建立完备的工程图例,可避免矢量化时重复设置图元参数,保证采集的点和线要素的图层与参数正确,提高建库效率。

同时,技术上还要注意,对于靠近图框处的线,矢量化时要沿其延伸方向向图框外多矢量化一段,以备将来拓扑造区时使用;技巧上还要做到经常压缩存盘,这样可以把删除的图元彻底删除,减少冗余,节省数据存储空间。而在矢量化完成后,还应对照底图逐格检查,主要检查是否丢失图元和内容,以确保矢量化内容与原图完全一致。具体可先以原图为依据进行计算机内自互检与修改,然后再输出全要素彩图以原图为依据进行图面自互检与修改。

3.3 数据拓扑误差的处理

由于数据矢量化过程中难免有些错误,在建立拓扑关系前,应先进行查错处理,进而提高数据的准确性,提高拓扑建立的效率。

进行整体拓扑处理前,应先将那些与拓扑关系无关的线删掉。为了纠正矢量化错误,减少数据误差,可以顺序执行系统中的自动剪断线→清除微短线→重叠线检查→清线重叠坐标及自相交→线拓

扑错误检查等功能。发现错误马上修改,如此反复,直至准确无误。若存在需要成区的线不封闭的情况,就利用线结点平差的功能将线封闭。确定没有任何拓扑错误再将线转成弧段,生成区文件,进行拓扑重建。最后再利用相关的检查软件进行线弧的套合性检查,以免产生误差点,进而提高数据的准确性。图 3 所显示的是线弧不一致错误,该错误产生的原因一般是系统参数设置过大或是线与线相交的地方没有处理好结点关系。

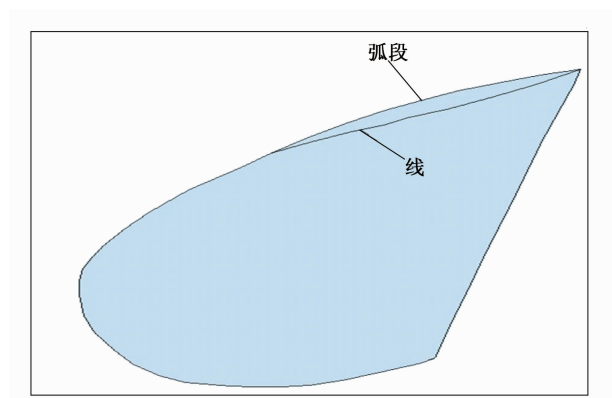


图 3 线弧不一致

Fig. 3 The line disaccord with the arc

多次执行“自动剪断线”会形成微小线段,应避免多次使用,一般以不超过 3 次为宜。

3.4 误差校正的精度控制

为了确保误差校正的精度,要从以下 3 点加以注意:

(1)将校正控制点的精度由系统默认的 2 位精度改为 6 位。

(2)误差校正形成的误差主要是由于控制点分布设置不合理而使校正后的图形发生了扭曲变形,因此,采集误差校正实际控制点时除 4 个图幅角点外,应增加公里网点作为实际控制点。同时,实际控制点与理论控制点的采集顺序号必须一一对应,从而保证校正后的文件不发生变形。

(3)尽量选择同一种校正方法。

另外,在保存文件时要注意区分原文件和校正后的文件,避免保存错误。以点文件校正为例,校正转换后,工作区中会出现 YWJ. WT 文件和 NEWLIN. WT 两个文件,其中, YWJ. WT 为原文件, NEWLIN. WT 为校正后的文件,保存时必须注意区分。

3.5 属性数据输入误差的 MapGIS 检控

属性录入前应认真检查各文件的属性结构、字段长度及类型是否有误,并根据要求进行修改和补充。利用 MapGIS 提供的“根据参数赋属性”等功能,对可赋属性的图层进行属性挂接。属性内容中用到的逗号“,”、分号“;”、连字符“-”、百分号“%”等符号一律用半角表示。若输入属性后需要重新拓扑,可利用“生成 label 点文件”的功能提取属性数据,再通过“label 与区合并”将原有的属性赋给新的区文件。

属性输入完毕,用以下 3 种方法检查属性数据的录入错误:

(1)利用 MapGIS 的工作区属性检查功能对关键字段进行检查,或输入条件进行条件检索;

(2)把属性数据打印出来人工校对;

(3)在 MapGIS 属性库管理功能中,导出 EXCEL 格式的属性数据,利用数据筛选功能,过滤导出属性中的重复数据,对照地质专业人员填写的属性表格进行检查。

建库人员和负责属性采集的地质人员,分别对原始数据进行两次校对,以确保属性与图元一一对应。

3.6 地图投影变换的 MapGIS 转换分析

在投影变换过程中,一定要确定文件有 TIC 点。文件间拷贝 TIC 点后,一定要点“确定”后再保存文件,不然刚拷贝了 TIC 点的文件还是没有将 TIC 点保存。投影参数要设置正确,从而保证数据转换的质量。

4 结语

空间数据库质量的优劣直接影响 GIS 应用分析结果的可靠性。在整个地质图空间数据库的建设过程中,影响数据库质量的因素还有很多,有的是技术方面的。数据库的质量更多程度上取决于建库者的责任和细心,因此,对建库过程中的每个环节进行认真仔细地检查,尽量减小或避免各程序环节上出现误差,从而提高数据库的质量和可使用性才是建设误差小、准确度高的高质量的地质图空间数据库的保证。随着地质调查工作的逐步深入,新数据、新成果不断出现,必须对数据库进行定期更新与维护,使之更好地为社会经济活动的各种规

划和管理决策服务。

参考文献:

- [1] 中国地质调查局. 地质图空间数据库建设工作指南(2.0 版)[S]. 2001.
- [2] 吴信才, 等. 地理信息系统原理与方法[M]. 北京: 电子工业出版社, 2002.
- [3] 中国地质调查局发展研究中心. 1:5 万区域地质图空间数据库(分省)建设实施细则(2009 版)[S]. 2009.
- [4] 白玲, 王小平. 1:20 万数字水文地质图空间数据库建库的质量控制[J]. 资源环境与工程, 2006(4): 450 - 454.
- [5] 赵庆英, 王国良. 运用 MapGIS 软件在进行误差校正中几种常见错误的剖析及改正措施[J]. 吉林地质, 2001(3): 73 - 76.
- [6] 范爱民, 景海涛. 地图数字化质量问题研究[J]. 测绘通报, 2000(4): 1 - 3.
- [7] 陈爱明, 柯育珍, 周录英. MapGIS 地质图空间数据库建设常见错误与分析[J]. 资源环境与工程, 2008(5): 543 - 546.
- [8] 徐卫, 张艳玲. 关于地理信息系统(GIS)数据精度与质量探讨[J]. 吉林地质, 2004(3): 98 - 101.
- [9] 顾蓬蓬. 浅析基于 GIS 的土地利用数据库的平台选择[J]. 内蒙古石油化工, 2007(2): 60 - 63.
- [10] 尹淮新. 图件数字化的质量控制[J]. 西北煤炭, 2005(3): 42 - 44.
- [11] 曹爱民, 汪晓萍. 应用 MapGIS 进行图形投影变换方法探讨[J]. 东北水利水电, 2009(10): 60 - 62.
- [12] 霍改兰, 冯宝爱. 浅谈 1:20 万数字水文地质图空间数据库建库工作方法 & 流程[J]. 水文地质工程地质, 2003(2): 65 - 68.
- [13] 马秋斌, 王道远. 1:25 万南京市幅数字地质图制作方法浅析[J]. 地质学刊, 2009(3): 303 - 309.

Analytical Application of MapGIS for Quality Control in Geological Map Spatial Database Constructing

MA Yuhong, WANG Jicun, XIE Songshi

(Shandong Institute of Geological Survey, Jinan 250013, China)

Abstract: The emphasis of Geological Map Spatial Database constructing is the quality control throughout each step. For the quality of the data directly affects the accuracy and reliability of the results of the analysis in using GIS, and for deviations exist objectively in all the basic spatial geological information, quality control is the inherent demand in constructing Geological Map Spatial Database. Through the analysis of quality problems which may arise from building of the regional geological map spatial database based on MapGIS, especially in the process of map pretreatment, digitization, vectorization, error correction, topological reconstruction, attribute input and projection transformation, the authors try to find the solutions and preventive measures, so as to implement the control of quality, most minimize the deviations of spatial data, improve the efficiency of mapping and precision of spatial data, and provide better service for the public application fields of spatial database, such as the constructing of national economy, the developing of regional planning and the protecting of geological environment.

Key words: space database; quality; MapGIS