

AutoCAD 地形图数据转换为 GIS 空间数据的  
技术研究与应用

陈 能,施蓓琦  
(上海师范大学,上海 200234)

Conversion Technique of AutoCAD Topographic Map Data to GIS Spatial Data  
CHEN Neng,SHI Bei-qi

**摘要:**为了高效地建立地理信息数据库,经常要将各类数字地形图数据转换为符合 GIS 要求的数据。以 AutoCAD 数据为例,介绍批量修改 GIS 基础图形数据的方法,分析 CAD 与 GIS 之间数据转换的各种模式,提出应用 GIS 数据中间件来实现 CAD 与 GIS 数据的无缝转换。并提出数据转换过程中的质量控制的办法。

**关键词:** AutoCAD;GIS;数据转换

一、引 言

在 CAD 的数据模型中,点、线、面等几何要素以二进制形式保存于文件中,相关的注记、颜色、线形等属性也跟几何数据放在一起,故其具有强大的绘图功能和处理矢量图形的能力,广泛地被应用在工业设计、机械设计、建筑设计、城市规划之中。尤其是在我国早期的地图制图系统中,多数以 AutoCAD 作为软件的开发平台,因此,目前很大一部分地理信息数据仍是以 AutoCAD 的数据格式(\*.dwg)存储的。但随着计算机的迅速发展,从空间地理信息系统的角度来看,CAD 缺乏较强的对空间数据信息的描述和分析功能。而地理信息系统(GIS)却在空间信息分析方面发挥着巨大的作用,能对已存在的空

间数据进行建模、分析和管理。因此已有的 AutoCAD 数字地形图只能作为 GIS 数据库建立的数字化形式的基础数据源,而并非是 GIS 概念中的空间信息数据库,是以供图为目的的。所以,为充分利用已有的数据资源,就必须有一个切实可行的处理方法,将已有的 AutoCAD 数字地形图数据转换为适用的 GIS 数据,以便更高效地建立城市基础地理信息数据库。

二、CAD 图形数据检查

一般 CAD 数字地形图要素主要分建筑物、道路、水系、电力线、高程点、等高线、地貌、植被、独立符号、垣栅、汉字注记、控制点等要素,其表现形式如表 1。

表 1 数字地形图要素的表现形式

地形要素	表现形式	图形及属性是否满足 GIS 要求
建筑物	单线(LINE)、复线(POLYLINE)	图形部分闭合,无属性描述
道路	单线(LINE)、复线(POLYLINE)、弧(ARC)	边线不连续,无属性描述
水系	单线(LINE)、复线(POLYLINE)、弧(ARC)	图形部分闭合,无属性描述
电力线	块(BLOCK)、线(LINE,POLYLINE)、圆(CIRCLE)、形(SHAPE)	电线不连续,电杆不是点状要素,无属性描述
高程点	块(BLOCK)、线(LINE,POLYLINE)、圆(CIRCLE)、点(POINT)	高程点有的不是点状要素,有的无高程值描述
等高线	单线(LINE)、复线(POLYLINE)	不连续,无属性描述
植被	线(LINE,POLYLINE)、弧(ARC)	有的无完整的面状,无属性描述
地貌	块(BLOCK)、线(LINE,POLYLINE)	没有按照统一的要求生成线形,无属性描述
独立符号	块(BLOCK)、形(SHAPE)	块符号满足 GIS 要求,形符号不能满足要求
汉字注记	注记(TEXT)、块(BLOCK)、线(LINE,POLYLINE)	注记没有按照 GIS 要求进行属性分层
控制点	块(BLOCK)、形(SHAPE)	块符号满足 GIS 要求,但无属性描述,形符号不能满足要求

从表 1 的分析数据可以看出 CAD 数据格式地形图要素的表现形式有多种,而且其面状地物如建筑物、水系也不一定完全闭合;线状地物如道路、陡坎等碰到软地物如高程点、汉字注记有断开;独立符号、汉字注记表示不是很清晰,不能满足 GIS 要求。因此大部分的空间数据仍需重新编辑,有的需程序处理,有的需人工干预。

虽然现有的 GIS 软件都提供了较强的图形编辑功能,但是 AutoCAD 具有更强大的编辑功能,尤其是在大批量的修改编辑方面。况且在 AutoCAD 的数据转换为 GIS 数据格式中总会存在部分数据的丢失,这就会给 GIS 的数据编辑带来很大的工作量,而且也会使数据的质量精度难以得到保证。因此在 AutoCAD 中对原有数字化图形做编辑修改等规范化工作比用 GIS 软件提供的编辑功能省力省事,且精度高。

### 1. 图形数据的分层检查

为了满足 GIS 空间信息数据库的要求,首先对建筑物、道路、河流等地形要素及其对应的注记层作分层的检查,以确保不同地物放置在相应图层,弥补相邻要素接边的几何裂缝或逻辑裂缝,删除悬挂线。图 1 包含了几种需要纠正的图形情况,图 1 中列举的线交叉、节点未闭合的情况是属于几何裂缝;中央无节点的情况是指两线交叉但没有交叉点,称为 T 形线;左侧的悬挂线是指不属于某个实体的多余线。对于较小的几何裂缝只要用捕捉方式连接节点即可。逻辑裂缝是指某一空间实体的组成元素位于不同图层,看似一个整体,实际存在逻辑错误,这种情况通过分层查看的方法可以检查出来。

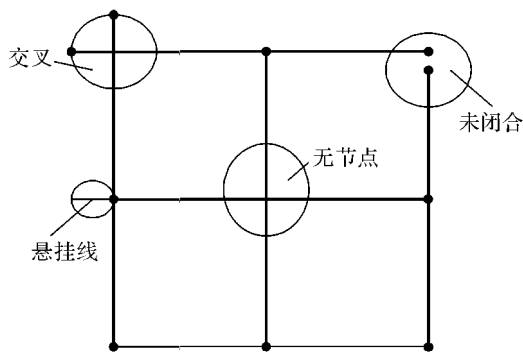


图 1 几种错误的图形情况

### 2. 图形数据的程序检查及处理

通过分层检查后,必须将不符合 GIS 拓扑结构要求的实体挑选出来,再选择集中过滤掉。以建筑物层数据为例,其不符合拓扑结构的实体大致分为两类。

1. 非多边形类型的实体:包括类型不是 Polyline

的实体和顶点个数小于 3 个的 Polyline 类型的实体。

2. 不封闭的多边形实体:先检查实体的封闭属性,如果是 False,则需查看其起始点坐标是否相同,若相同则将其封闭属性设为 True,否则即是不封闭多边形。

在选择集中将这些不符合要求的实体过滤,在图上用其他颜色表示,以便再次复查。在 AutoCAD Map 2000i 中要完成实体过滤,只需使用 Quick Select 命令即可执行此功能操作。

经过实体过滤之后,选择地形图中所有要素,删除重复对象,将重叠的图形元素清除成单一元素。并对图形中 T 型线进行处理,使之产生必要的节点,使图形入库时能顺利构成拓扑关系。以上的处理均可运用 AutoCAD Map 2000i 中强有力的图形图面清理功能来处理。但若出现如图 2 所列举的多边形的不相邻的重复顶点情况时,则不能判断出来。图 2 中的多边形是在起始点重合之后又多走出一条短线,这种情况下多边形的封闭属性为 True,不会被过滤掉。此时,有必要在重复顶点删除之后再进行一次重复点判断,即多边形的每一顶点和其他顶点进行比较,找出上述情况的多边形。由于这种重复顶点的情况较为复杂,故采用人机互动的方法以询问的方式来修改。

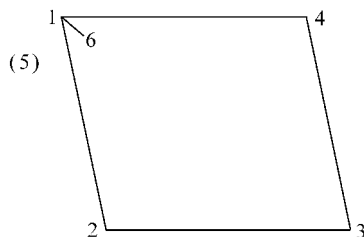


图 2 多边形的不相邻的重复顶点

## 三、CAD 与 GIS 的数据转换

在完成 CAD 图形数据的检查、修改之后,即要进行 AutoCAD 与 GIS 的数据转换,以建立 GIS 数据库。

一般来说,不同数据格式之间的数据转换方法有两种:直接转换和间接转换。通常意义上的直接转换是指两个系统不同数据格式数据之间的转换是直接转换、读写。这种方式的效率较高,但首先得必须知道相互转换的两个系统内部的数据结构才可进行,因此难度较大。而间接转换则是指不同数据格式的数据集之间的转换通过中间数据格式进行的,如 AutoCAD 的 DXF, ArcView 的 Shape, MapInfo 的 MIF 格式等。通过交换格式可以实现不同软件之间的数据转换,但是由于缺乏对空间对象统一的描

述,不同数据格式描述空间对象时采用的数据模型不同,因而转换后不能完全准确地表达原数据的信息,经常性地造成一些信息丢失,而且用户必须详细地掌握交换格式的细节及应用系统数据格式的细节时,才能进行转换程序的设计工作,对用户的编程能力及工作量要求都很高,不利于无缝 GIS 快速获取多源数据。而在 CAD 与 GIS 的数据转换时,其转化方式也是这两种,但在此界定直接转换与间接转换的标准是指 CAD 与 GIS 二者之间是否通过特制程序接口进行相互转换、读写。

1.CAD 与 GIS 的数据直接转换方式

数据的直接转换要求 CAD 与 GIS 二者之间不通过特制程序接口就能够相互转换或直接读写。以 ESRI 的系列 GIS 软件与常用的 CAD 系统数据交换为例,通常有 3 种技术方式,即文件转换方式、直接访问方式和数据库共享接口方式,如图 3 所示。其中,文件转换方式是指数据从一种文件格式转换到另一种文件格式(如 DXF 到 SHP)。直接访问方式是不经过中间格式,系统直接把另一种系统的数据读入内存,如 ArcView 可直接读入 AutoCAD 的 DWG 文件,并可像操作其他 GIS 数据一样进行显示、查询与打印。相比前两种方式而言,数据库共享接口方式是一种新型的方式,它通过在系统中嵌入应用程序接口(API),访问其他系统的数据。ESRI 的 ArcSDE 就具有 CAD 客户端扩展程序,允许 AutoCAD 用户在 DBMS 存储或提取 CAD 要素或 GIS 特征。同时,ArcSDE 的数据访问 API 也可嵌入到 CAD 系统中。

由于 CAD 与 GIS 系统描述同一空间对象的数据模型与语义存在一定的差别,以上方法虽然在某种程度上可以实现数据交换与共享,但在转换过程中容易导致信息丢失。如在使用直接访问方式时,可在 ArcView 中直接读入 CAD 专题,但在 CAD 转化实体到 ArcView 要素时,会有明显的差异。除此之外,因为 CAD 专题具有与其他专题不同的特殊属性,所以表现在 AutoCAD 中的 Block 块会在 CAD 专题中消失,要素和属性是不可编辑的,而且在缺省情况下,当 CAD 专题添加入视图时仅当前的活动图层可见。而在使用另两种方式将 AutoCAD 转换到 ESRI 的 Shapefiles 或 Coverage 时,有时就会丢失 AutoCAD 中定义的图形对象扩充属性。因而,直接转换难以顾及 GIS 中的属性数据,无法实现数据无缝转换。

2. 间接转换

在实际的应用中,为真正实现由 CAD 到 GIS

数据无缝转换,应该应用 GIS 数据中间件,通过间接转换的方法来解决这些问题。所谓 GIS 数据中间件是指能够嵌入各类 GIS 系统的软件插件,这类插件由各类 GIS 软件开发人员与用户各自独立完成。其原理类似于即插即用设备的驱动程序设计,即 GIS 软件平台开发者规定系统内部数据的读写接口(简称 GDIO),这些接口操作本平台内部或交换数据结构,接口与平台间的通信对数据源是个黑匣子,对于不同源头的空间数据,数据提供者编写 GDIO 接口内的操作代码,编译注册后,GIS 软件平台即可操作该类空间数据。具体实现如图 4 所示。

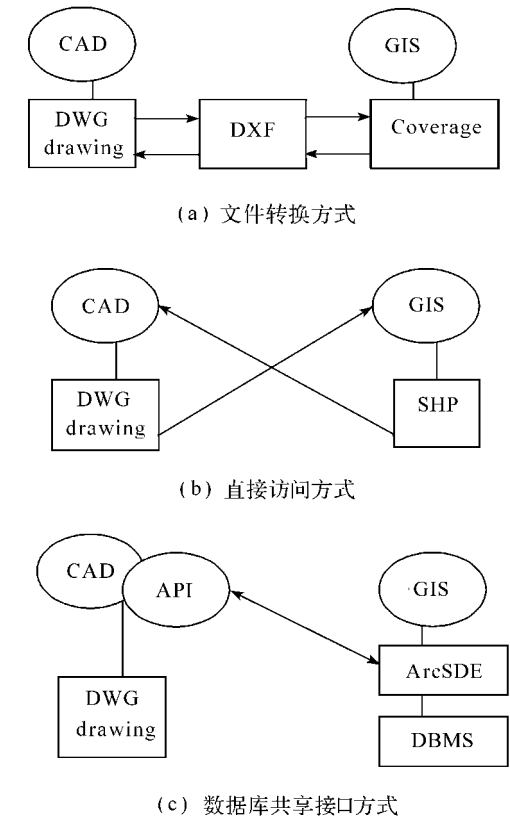


图 3 CAD 与 GIS 数据直接转换的 3 种方式

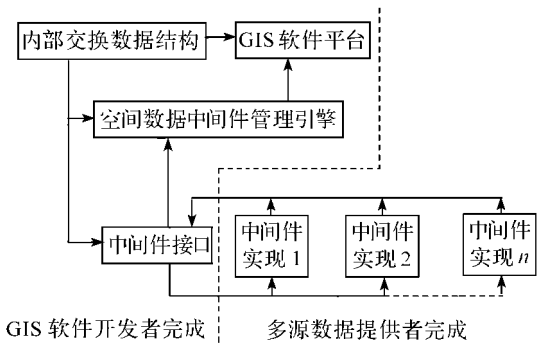


图 4 数据中间件模式

而从时间段上 GIS 数据中间件的任务分两个阶段完成,第 1 阶段,某类 GIS 软件开发人员根据本类 GIS 的特点规定 GIS 数据中间件的数据输入输出(I/O)接口,并完成相应与 GIS 底层直接通信的部分;在第 2 阶段,GIS 用户根据自身需要及所处理的空间数据源的特点完成中间件的数据源解释部分,经过简单的编译实现可运行的 GIS 数据中间件,注册进系统,GIS 系统就完成了对用户数据的支持。

数据中间件主要由内部数据写入 API、数据中间件接口、具体的用户数据中间件 3 部分组成。其数据中间件全部继承本接口,待接口的内容完成后,注册进系统,完成用户中间件的实现过程。

由此可见,采用数据中间件技术,开发者与数据提供者任务明晰,通过设计的数据中间件接口,不同类别 GIS 提供的空间数据经由不同的数据中间件接收并转换为 GIS 无缝数据中的一块,既可快速、并行完成多源空间数据装载,又可后续通过扩充增加新的数据源。

#### 四、数据转换过程中的质量控制

数据转换过程中,数据质量是指点、线、面数据几何和属性逻辑一致性的可靠程度。例如,在一种 GIS 软件中的线数据,由于交换模块功能上的缺陷被写成了点,出现几何不一致性的错误;在数据交换过程中,由于各种原因,有时会造成地物属性的丢失,出现逻辑不一致性的错误。

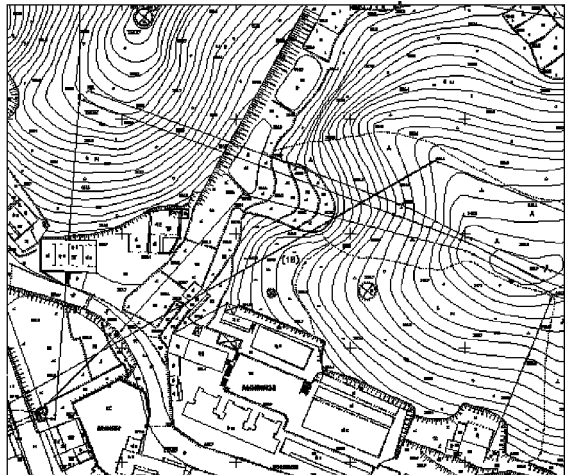
数据质量控制应根据外部数据文件的格式和内容采取相应的方法。首先通过地物编码定义文件,然后按照点、线、面数据的一般特点进行检查。对特殊的点、线、面数据除了作一般性的检查外,还需根据编码对其数据的特殊性进行检查。如对有向点符号及两点定位符号,它们后面只能有两个坐标对,对桥、室外楼梯等特殊规则地物,根据其符号化的要求,对其应有的坐标对数据进行检查,对有属性的地物还应对其应有的属性项数据进行检查。

做完检查后,通过程序对检查结果给出统计信息,即被检查的总记录数及出错记录数,对出错记录给出其对应编码及其在文件中的位置,以便启动原数据进行对照检查、修改。这样就为顺利实现数据交换提供了可靠的保障。

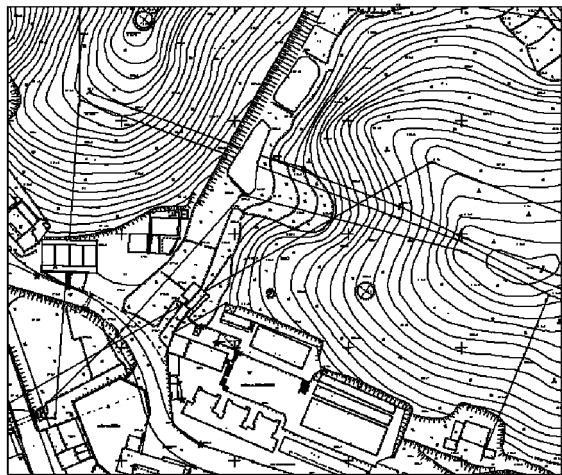
#### 五、实例分析

县域基础地理信息系统是目前“数字泰宁”工程的最基础的项目。其中的空间基础地理信息是指在一定尺度下,能完整地描述城市自然和社会形态的

地物地貌信息、管理境界信息以及它们相对应的基本属性信息。可见,空间基础地理信息中包含了大量的地形信息。因此,将原有的全要素数字地形图更新、修改,并转为符合 GIS 要求的数据格式是一项相当繁琐的工作,涉及大量的图形编辑与属性录入。考虑到 AutoCAD 具有强大的图形编辑功能,并能挂接属性,实现属性与图形能以面向对象的方式操作,便于数据的更新、修改,故采用 AutoCAD 实现县域基础地理信息系统的数据采集与更新功能。首先对泰宁县城 1:500 CAD 格式的地形图作数据检查、修改,使其图形数据能满足 GIS 格式的要求,然后应用数据中间件技术将 CAD 数据格式转换为 GIS 的数据格式,如 MapInfo 的 MIF/MID 格式, ArcView 的 SHP 格式。同时,对转换过程中的数据进行质量控制。图 5 为由 CAD 数据转成的 GIS 数据格式。



(a) 泰宁县城 1:500 CAD 格式的地形图



(b) 泰宁县城 1:500 ArcView SHP 格式的地形图

图 5 由 CAD 数据转成的 GIS 数据

数据等采用无损压缩(即 LZ77 算法),否则采用有损压缩(即 JPEG 算法)。以高分辨率为底层,通过逐级抽取,建立不同分辨率的金字塔结构,逐步形成较低分辨率的遥感影像数据。该方法通常会增加 20%左右的存储空间,但却可以提高影像数据的显示速度。背景数据采用格网索引方式,通过试验确定合适的格网级数和单元大小。

### 三、系统功能及开发

塔河流域遥感影像数据库系统是在统一的总体规划和详细设计完成后,经过模型开发、专家论证和行政批准,分阶段实施完成的。其实施步骤为:建库试验—完善方案—试建库—重新完善方案—规模建库—模型开发—试运行—规模开发—完善系统—系统测试—完善系统—投入运行—逐步完善系统。

1. 数据入库,根据比例尺、图幅号、时相、影像类型等多种方式,通过入库程序(基于 ArcObjects 开发的 DLL)将遥感影像数据、空间地理等数据自动导入到数据库中;遥感数据入库后,利用 ArcSDE 工具创建影像金字塔以提高系统检索速度。

2. 数据查询,可通过图幅号、经纬坐标、居民地、地名、线缓冲、圆形区域、矩形区域、多边形区域等条件查询显示背景数据和影像数据;可将查询的背景数据实时转换到用户自定义的地图投影;可将背景数据按照设定比例尺进行显示;具有鹰眼功能,能够任意开窗放大、缩小、漫游等操作;用户可根据向导、或输入 SQL 语句进行如包含、相邻等条件的空间查询分析。

3. 元数据查询,可通过图幅结合表点击查询视图中任意点的元数据信息,也可通过数据类型、比例尺等多种类型直接查询或通过组合条件查询主要的元数据项。

(上接第 14 页)

### 六、结 论

由于 CAD 与 GIS 在数据与功能上的互操作性与集成应用已显得越来越重要,因此,首先通过对 CAD 数字化的地形图数据全面的检查后,再应用中间件技术将 CAD 数据无缝地转换成符合 GIS 数据要求的格式,并同步地进行数据转换的质量控制,这样才能保证高质量的地理空间数据,从而提高 GIS 的整体开发的效率,降低成本。故这种转换技术具有更好的应用价值。

4. 数据更新,可通过入库程序自动导入和删除分幅、分层数据,并可对元数据进行编辑修改。

5. 数据维护,包括数据库的备份和恢复、用户权限管理、数据库权限管理等。其中数据备份需根据系统的可用性、可接受的停工时间和数据损失制定备份策略,并对备份策略进行不断测试,以保证数据库能从各种故障中及时有效地恢复。

### 四、结束语

综合使用遥感、地理信息系统、数据库等技术,建立基于动态监测与管理的塔河流域生态环境信息系统,具有明显的工程、技术优势。主要表现在:① 分析数据快速、直观,可快速反映当前塔里木河流域的生态环境状况;② 可对历年来的生态环境状况进行叠加分析,为政府和有关部门进行长期的环境保护和生态规划提供科学依据;③ 采用当前流行的数据仓库管理技术,将空间数据-属性数据一体化存储管理、多源数据无缝拼接,以及基于 C/S 和关系数据库的 3 层体系结构设计,有效地管理了跨带的遥感、地理、元数据等基础数据;④ 整个数据基于 ArcObjects 组件开发,便于系统以后的移植、升级。

### 参考文献:

- [1] 乔 木. 新疆耕地资源卫星遥感调查数据分析研究[J]. 干旱区地理, 2002, 20(2): 309-314.
- [2] 陈述彭, 鲁学军, 周成虎. 地理信息系统导论[M]. 北京: 科学出版社, 2000.
- [3] 方 涛, 龚健雅, 李德仁. 影像数据库建立中的若干关键技术[J]. 武汉测绘科技大学学报, 1997, 22(3): 266-269.
- [4] 赵蕴冬. 数据仓库及其建设方法[J]. 油气田地面工程, 2003, 22(2): 74-75.

### 参考文献:

- [1] 王 卉. 一种解决 GIS 多源数据无缝集成的方法[J]. 测绘工程, 2003, (2).
- [2] 郑丽敏, 等. 土地房产 GIS 基础数据采集的质量控制[J]. 工程勘察, 2001, (6).
- [3] 张雪松, 等. AutoCAD 环境中组织 GIS 数据的方法[J]. 测绘通报, 2003, (11).
- [4] 施一军. 数字地形图向 GIS 数据转换处理方法的探讨[J]. 江苏测绘, 2000, 23(4).
- [5] 闾国年, 等. 地理信息系统集成原理与方法[M]. 北京: 科学出版社, 2003.