

利用 AutoCAD 数据建立地籍数据库方法研究

张源, 杨锦华, 常原飞, 乔彦友

(中国科学院遥感应用研究所, 北京 100101)

Research on an Approach for Cadastral Database Construction Using AutoCAD Data

ZHANG Yuan, YANG Jin-hua, CHANG Yuan-fei, QIAO Yan-you

摘要:提出手工编辑与程序自动处理相结合的数据转换方法,该方法相当程度地保证了转换的正确性,且具有较高的自动化程度,已运用此方法成功建立了山西省阳城县地籍空间数据库。

关键词:GIS;CAD;数据转换;宗地;地籍数据库

一、引言

CAD 类软件具有操作简单、绘图功能和图形编辑功能强大、精度高、出图图面整洁美观、可二次开发等优点^[1,2],是我国早期的地图制图系统所采用的主要开发平台。因此,现在仍有很大一部分地理信息数据以 CAD 数据格式存在。CAD 软件虽然在图形处理方面具有极强的优越性,但由于 CAD 数据中没有拓扑关系、不包含属性数据导致其在空间数据建模与管理、空间分析、空间查询等方面有着很大的欠缺^[3];而 GIS 以其集采集、存储、查询、分析于一体的优势以及强大的空间数据管理功能^[3,4],逐渐成为空间数据应用和建立各种空间数据库的主要手段。将大量现存的 CAD 数据科学合理地转化为 GIS 数据并建立 GIS 空间数据库能够降低建库成本,缩短建库周期,提高原有 CAD 数据的利用率。因此,如何科学合理地将 CAD 数据转为 GIS 数据,就成为一个值得研究的现实问题。

目前 CAD 数据转为 GIS 数据的方法^[1-5],主要分为两大类:一是直接数据转换方式,即 GIS 软件直接读取 CAD 数据或 CAD 格式文件不经过特制程序接口直接转换为 GIS 格式文件,这种方法在某种程度上可以实现数据交换和共享,但是由于 CAD 与 GIS 两种数据的数据存储、图元定义等存在一定差别,因此这种直接转换方法很容易造成信息丢失、图形要素变形等问题,不易实现数据无缝转换;二是根据 CAD 和 GIS 两种图形数据的存储结构,进行二次开发以实现数据转换,但是这是在 CAD 数据完全正

确和标准的前提下设计的数据转换路线,自动化程度很高,却存在实用性不强的缺点。

本文所采用的方法是手工编辑与程序自动处理相结合,既避免了直接转换所产生的数据丢失、图形要素变形等问题,又避免了二次开发转换方法由于 CAD 数据质量参差不齐所产生的一系列错误。

二、建立地籍空间数据库流程

建立地籍空间数据库的流程图如图 1 所示。

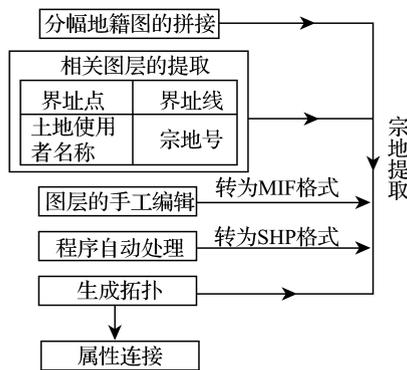


图 1 流程图

具体步骤为:

1. 分幅地籍图的拼接:在 AutoCAD 中将分幅地籍图拼接好,将其移至正确地理坐标。
2. 相关图层的提取:从拼接好的地籍图中提取界址点图层、界址线图层、土地使用者名称图层和宗地号图层。
3. 图层的编辑:针对 AutoCAD 数据各图层中存在的问题进行手工编辑。

4. 程序自动处理:用 VC 编程实现对各图层的处理。由于 MIF 文件(MapInfo 的文本交换格式文件)为 ASCII 码文件,文件形式比较简单,便于程序读取和处理,而且 MIF 文件是用文本记录和描述几何要素及其坐标,绝对忠实于原始 CAD 文件,不会造成几何要素的丢失或变形,因此先将各图层 CAD 文件转为 MIF 文件;经程序处理生成正确的界址点、界址线 MIF 文件,并对土地使用者名称 MIF 文件和宗地号 MIF 文件进行必要的编辑,以便之后可以顺利地转变为带有属性数据的点图层;程序处理完成后所产生的 MIF 文件完全符合 GIS 数据结构和拓扑结构,因此在转为 GIS 格式数据时不会有图形要素丢失和变形的问题,此时再将新生成的 MIF 文件转为 GIS 格式。

5. 生成拓扑:为界址线图层建立拓扑关系,生成属性表中包含“土地使用者名称”和“宗地号”字段的面状要素宗地 GIS 数据。

6. 属性连接:使用 ArcGIS 编写 VB 脚本,将宗地 GIS 数据属性数据表和地籍数据库表中的宗地号和土地使用者名称字段进行关联,建立地籍空间数据库,为地籍信息系统的构建做好准备。

下面就本方法的两个重点步骤手工编辑和程序处理生成宗地地块的算法进行详细分析和解释。

1. 手工编辑

(1) 土地使用者名称图层的手工编辑

对于土地使用者名称图层来说,下列情况中的注记(见图 2)需要进行手工编辑。

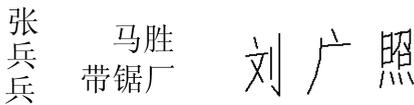


图 2 错误的注记

图 2 中的注记为了配合地物的延伸方向、形状和大小,本应作为一个注记输入的却被分为两个或多个注记输入。如“张兵兵”分为“张”、“兵”、“兵”输入;“马胜带锯厂”分为“马胜”、“带锯厂”输入;“刘广照”分为“刘”、“广”、“照”输入。这种情况对于今后的属性连接影响很大,但却只是少数情况,因此,只要进行手工编辑将其改为一个完整的注记即可。

其次,有一小部分非人名类注记与数据库文件中的不一致,如 AutoCAD 文件中的注记为“人民政府”,而数据库文件中为“镇政府”;再如,“供销合作社”与“合作社”,“**水电公司”与“**镇水电

公司”等。这种情况也需要在 AutoCAD 中进行编辑,使其与数据库文件保持一致。

(2) 界址点、界址线图层的手工编辑

为了顺利生成拓扑,需要对界址点、线图层中的图形进行纠正;而像交叉、未闭合、悬线、悬点、重复顶点等情况都可以通过程序自动处理进行纠正,因此,需要手工编辑的仅有以下两点:

1. 对界址点图层来说,需要处理的是无顶点的情况,即需要在界址线连接处没有界址点的地方手工加入界址点,如图 3 所示。

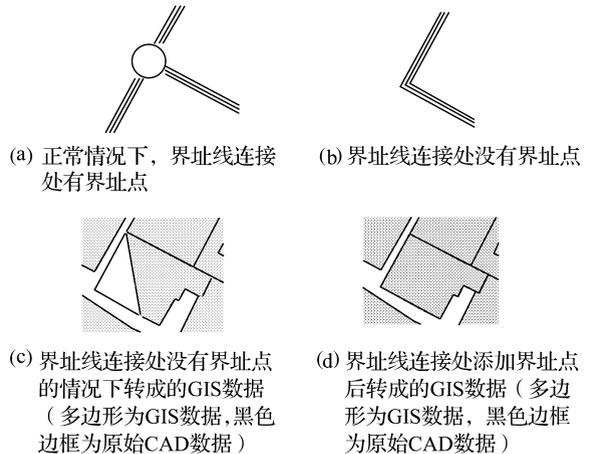


图 3

2. 对于界址线图层来说,由于作业人员的不规范作业,界址线图层中混入了一些线状地物,这些线状地物大部分会在程序自动处理时去除掉,但某些穿过宗地的线状地物会影响宗地多边形的构建,导致多边形构建错误,如图 4 所示。图 4 中,左边本该是一个宗地却被分成了两个宗地,而右边的宗地由于线状地物的影响缺少了一角。像这样的线状地物应该通过手工编辑去掉。

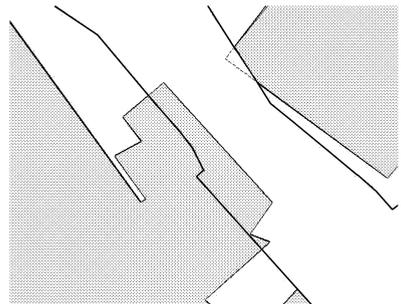


图 4 线状地物影响宗地构建

2. 空间数据提取和处理

界址点图层的处理:首先读入界址点、界址线图层 MIF 文件,经程序处理后,去除非界址点,生成新

的界址点 MIF 文件。主要算法如下:

```

读入界址点 MIF 文件
读入界址线 MIF 文件
设定阈值 MINDIST
for ( int PointNum = 总点数, PointNum > 0,
PointNum—)
{
    for ( int LineNum = 总线数, LineNum > 0, Line-
Num—)
    {
        if (点与线起始点之间的距离 < MINDIST)
            break;
        if (点与线终止点之间的距离 < MINDIST)
            break;
    }
    if (LineNum == 0)
        删除第 PointNum - 1 个点
}

```

生成新的界址点 MIF 文件

界址线图层的处理:由于 CAD 文件中每条界址线是由多条线组成的(如图 3(a)、图 3(b)所示),因此,界址线图层的处理需要解决的一个主要问题就是如何将多条线组成的界址线变为一条惟一的正确的界址线。由于上一步界址点图层的处理已经生成了包含所有正确界址点的界址点 MIF 文件,因此,只要将每一条线都用离它的起始点和终止点最近的两个界址点的连线代替,再将重复的线删除即可。

根据界址点重新生成界址线的主要算法如下:

```

for ( int LineNum = 总线数, LineNum > 0, Line-
Num—)
{
    FirstDIST = MINDIST;
    LastDIST = MINDIST;
    tempLine = 第 LineNum - 1 条线;
    for ( int PointNum = 总点数, PointNum > 0, Point-
Num—)
    {
        if (界址点与 tempLine 起始点的距离
< FirstDIST)
        {
            FirstDIST = 界址点与线起始点的距离;
            FirstINDEX = 该界址点序号;
        }
        if (界址点与 tempLine 终止点的距离
< LastDIST)

```

```

{
    LastDIST = 界址点与线终止点的距离;
    LastINDEX = 该界址点序号;
}
}
if ( FirstDIST < MINDIST)
{
    tempLine 起始点 = FirstINDEX;
}
if ( LastDIST < MINDIST)
{
    tempLine 终止点 = LastINDEX;
}
用 tempLine 替代第 LineNum - 1 条线
}

```

三、属性数据提取和处理

土地使用者名称图层的处理:为了能够在建立拓扑关系时为宗地赋属性值,将土地使用者名称图层 MIF 文件由注记类型 MIF 文件转为点类型 MIF 文件,土地使用者名称作为点的属性。

宗地号图层的处理:与土地使用者名称图层一样,宗地号图层 MIF 文件也需要转为点类型 MIF 文件;但是与之不同的是,宗地号图层中的注记不需要全部提取。在 AutoCAD 地籍图中,宗地号是以分式的形式表示的,例如 136/266 中,分子 136 表示宗地号,分母 266 表示地类号;在本文中只需提取宗地号,因此相对于土地使用者名称图层,宗地号图层的处理需要进行一次判断,只提取宗地号而舍弃地类号。由于数据读取过程中注记是被连续读取的,即同一个宗地号中的宗地号和地类号是被连续读取的,因此,只要判断相邻两个注记 y 坐标的大小,保留 y 坐标大的注记,即实现了只提取宗地号的目的。

四、应用实例

使用上文中介绍的数据转换和地籍空间数据库建立方法,利用山西省阳城县 2005 年 CAD 格式分幅地籍图以及 dBase 形式的地籍数据库,成功建立了山西省阳城县地籍空间数据库。其中宗地正确提取率达到了 99% 以上,仅有个别宗地未被提取或提取错误;属性连接正确率达 98% 以上,存在小部分由于土地使用者名称不统一或宗地提取问题而产生的错误。以山西省阳城县东冶镇为例,CAD 数据与转换后的 GIS 数据如图 5 所示。

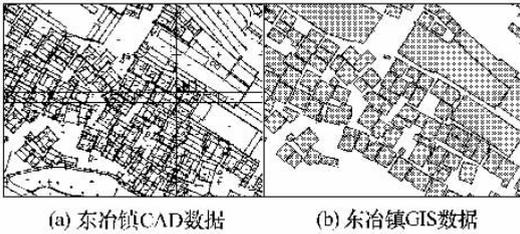


图5

五、结论

随着 GIS 的高速发展和广泛应用, CAD 数据到 GIS 数据的转换已经成为一种必然的趋势。本文提出了一种将手工编辑和程序处理相结合将 CAD 格式地籍数据转换为 GIS 格式数据的方法, 并利用 GIS 数据与地籍属性数据库进行关联, 建立了地籍空间数据库。在实际中的应用表明该方法省时省力、操作简单、精度和正确率高, 具有很好的效果。

而且, 该方法仍有进一步发展的空间: 首先, 一个注记被分成两个或多个注记输入的情况, 在本文中是采用手工编辑, 在这种情况下大量出现的情况下, 也可以采用编程的方法自动处理, 即设定两注记间的距离阈值, 当两注记间距离小于该值时, 就将前一个注记缀于后一个注记尾部, 坐标中心点取两注记

坐标中心点的平均值^[5]; 其次, 关于几何要素图层的处理仍未达到高度自动化, 手工编辑是 CAD 数据转 GIS 数据的过程中必不可少的重要一步, 但仍应尽量减少手工编辑的工作量, 使数据转换过程向高度自动化的方向发展。除此以外, 继续深入研究这一方法, 将该方法应用于其他类型 GIS 数据库建库过程也是接下来的研究方向之一。

参考文献:

- [1] 张雪松, 张友安, 邓 敏. AutoCAD 环境中组织 GIS 数据的方法[J]. 测绘通报, 2003, (11): 45-48.
- [2] 周勇前, 陈 军. AutoCAD 与 Arc/Info 的数据转换[J]. 测绘通报, 1995, (3): 20-22. 34.
- [3] 陈 能, 施蓓琦. AutoCAD 地形图数据转换为 GIS 空间数据的技术研究与应用[J]. 测绘通报, 2005, (8): 11-14. 34.
- [4] 史 军, 邱 昕, 崔林丽, 等. CAD 图形数据转换为 GIS 空间数据的技术研究及应用[J]. 东北林业大学学报, 2001, 29(6): 83-85.
- [5] 韩 勇, 陈 戈, 甘宇亮, 等. 基于 CAD 数据建立 GIS 文本数据库方法探讨[J]. 测绘通报, 2004, (5): 22-24.
- [6] 王艳东, 龚健雅, 黄俊韬, 等. 基于中国地球空间数据交换格式的数据转换方法[J]. 测绘学报, 2005, 29(2): 142-148.

(上接第 64 页)

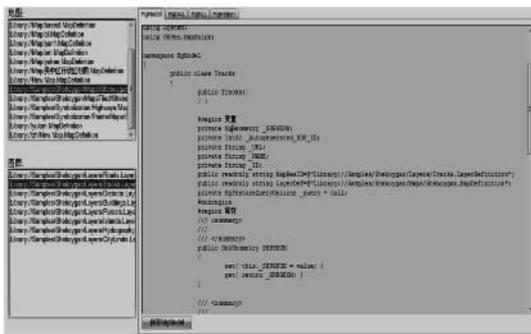


图3 两层一体代码生成器

四、结论

本文在 ORM 三层结构的工作原理基础上, 以 MapGuide 地图服务实例为对象提出 OSM 概念, 并对这种 OSM 模式进行结构设计与实现探索, 在基于 MapGuide 的电子地图应用系统中使用该模式可以

提高开发效率、减少维护难度。

分层式设计可以达到分散关注、松散耦合和逻辑复用的目的, 但也存在一定的缺点, 如降低了系统的性能, 有时甚至会导致自上而下的级联性修改。因而在系统设计开发时需根据具体业务需求确定是否使用这种 OSM 框架结构。

参考文献:

- [1] 秦洪现, 崔惠岚, 孙 剑, 等. Autodesk 系列产品开发培训教程[M]. 北京: 化学工业出版社, 2008.
- [2] FOWLER M. Patterns of Enterprise Application Architecture[M]. New York: Addison-Wesley Publishing Company, Inc., 2002.
- [3] Open GIS Consortium. OpenGISReferenceModel[EB/OL]. <http://www.opengis.org>, 2003-07-09.
- [4] 林大地, 江 南, 等. Java 环境下基于 Hibernate 的 GIS 数据库访问策略[J]. 地理与地理信息科学, 2006, 22(2): 119-120.