

地形图数据规整及变化要素自适应更新方法



张新长^{1,2}, 高金顶¹, 何显锦¹

(1. 中山大学地理科学与规划学院, 广东广州 510275;

2. 广东省城市化与地理环境空间模拟重点实验室, 广东广州 510275)

摘要: 地形图数据作为一种重要的空间数据, 其更新一般包括数据规整和更新入库 2 个过程。提出了一套地形图数据规整及自适应更新方法。该方法在 AutoCAD 环境下对地形图数据进行规整, 包括地物编码、拓扑检查编辑、数据转换等; 在数据规整的基础上, 进行要素的变化检测, 提取出变化的要素, 再对变化要素进行更新入库。实验表明, 该方法能系统、高效地实现地形图数据的更新。

关键词: 地形图; 数据规整; 变化检测; 增量更新

中图分类号: P208

文献标志码: B

文章编号: 1672-4623 (2016) 03-0001-04

空间数据更新是维护空间数据库现势性的主要手段^[1], 是当前 GIS 的前沿研究课题。目前, 我国大多数城市的地形图测量数据以 AutoCAD 格式存储, 而在进行空间数据存储、空间分析等方面 CAD 格式数据相对于 GIS 数据存在很大的不足。为充分利用现有的数据资源, 一般是将已有的 AutoCAD 数据转换为相应的 GIS 数据^[2]。目前, 对 AutoCAD 数据向 GIS 数据转化的研究较多^[3,4], 但由于 AutoCAD 数据在生产方式和数据标准方面存在较大差异, 造成数据本身存在不规范性, 使得这些方法的转换结果并不理想^[5]。本文主要研究将 AutoCAD 数据进行规范整理并无损转换为 GIS 数据的方法。

空间数据规整后, 如何将数据快速更新到现势数据库, 以维持数据库的现势性, 是一个关键的技术问题。目前, 常用的地形图数据更新方法有 2 种: 基于范围的更新方法和基于要素的更新方法^[6]。基于范围的更新是将旧数据按相应的范围挖空后填入新数据, 然后进行要素接边操作^[7,8]; 基于要素的更新是对新旧数据进行变化检测, 只针对变化的要素进行更新^[9-11]。基于范围的更新方法需要较多的接边操作, 在接边过程中会出现接边错位现象^[12], 并且整体替换加大了数据的存储量, 降低了数据的更新效率, 因此, 本文采用基于要素的更新方法。

1 地形图数据规整及自适应更新方法

1.1 地形图数据规整关键技术流程

地形图数据规整一般包括 2 个阶段: 将初始地形图数据进行标准化处理; CAD 数据与 GIS 数据的转换^[13], 本文提出了筛选特定的图层并进行批量转换的方法。

1.1.1 地形图数据标准化

只有符合技术规程及相关作业标准的数据才能正确进行数据转换, 因此首先对地形图数据进行标准化整理。AutoCAD 地形图标准化步骤主要包括地物编码赋值、编码检查、拓扑检查与编辑、属性编辑等, 具体标准化流程如图 1 所示。

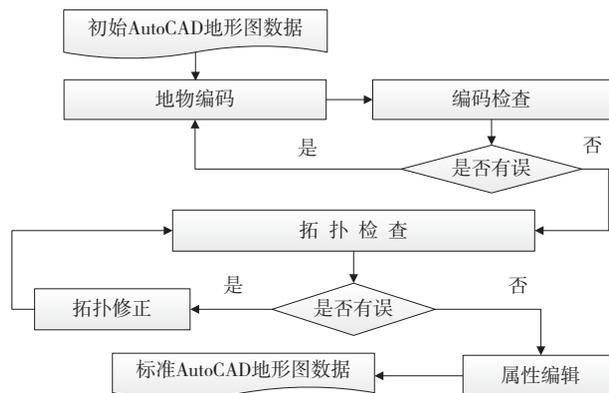


图 1 地形图数据标准化流程图

地形图要素编码即给图形赋予唯一的编码值^[14], 编码是标准化的核心, 赋值过程中若仅使用目视方法进行逐要素判断需要花费较长时间。通过对实测地形图的调查研究发现, 不少数字地形图由 CASS 软件编辑而成, 且一般存储了 CASS 编码和编码意义。本文提出了批量转换编码的方法, 可快速将 CASS 编码转换为对应的要素编码, 加快了转换速度。

在编码完成后可以对要素的拓扑信息进行相应的检查与修正, 主要检查常见的拓扑错误, 包括图形的自相交、图形的悬挂点、图形的重叠等。本文提供了面要素的自动闭合、重叠要素的删除等处理方法来解决相应的拓扑错误。

最后对要素进行相应的属性编辑与赋值, 采用图

收稿日期: 2015-12-09。

项目来源: 国家自然科学基金重点项目资助项目 (41431178); 高等学校博士点专项科研基金资助项目 (20120171110030)。

形的扩展字典方法,将名称和对应的值分别以成对的方式存储在字典的 Name 属性和 Data 值中,这样既可快速进行存储和调用,又增加了属性值的可读性。

1.1.2 地形图数据转换

在完成 AutoCAD 数据的标准化之后,要进行 AutoCAD 数据与 GIS 数据的转换,以建立 GIS 空间数据库,主要包括图形转换和属性转换。在分析了 2 种数据的基础上,采用直接转换的方式,将 AutoCAD 地形图数据转换为 ArcGIS 格式的标准 GIS 空间数据,具体步骤包括:

1) 建立图层间的映射关系,建立各个专题的点、线、面、注记等图层,并依据入库标准添加相应的属性字段。

2) 实现 CAD 数据到 GIS 数据间的图形转换。图形转换是转换流程中最容易产生要素丢失的环节,这里提取相应的坐标,并对坐标进行组合,组合成相应的要素类型。常见的 CAD 实体图形包括点、直线、圆、弧线、样条曲线、多线段和文字^[15]。针对以上各种类型,设置相应的提取方法。直线仅提取起点和终点;圆实体提取圆心的位置和半径;而弧线不仅需要提取圆心和半径,还需要提取对应的圆心角;样条曲线提取起点和终点,并根据该曲线的复杂度选择中间控制点的数量;多线段将其进行拆分成相应的简单直线和圆弧后再进行提取操作;为了实现文字的朝向和大小,需提取其旋转角度和字体高度。采用以上提取方式后,根据相应的还原方式,连接成对应的要素。

3) 在图形转换后,要进行相应的属性转换和赋值,这里就从图形的扩展字典中读取并存储到对应图层的相应属性字段中。

1.2 基于要素的更新方法

1.2.1 总体思路

基于要素的更新方法主要包括 2 个阶段:第一阶段是变化信息检测,即通过新旧数据的对比,检测出变化的要素,将变化的要素存放在临时库中;第二阶段是更新数据的入库,即根据临时库中变化要素的类型(新增或删除),对数据库中的数据进行相应的更新操作。总体更新流程如图 2 所示。

1.2.2 变化信息检测

变化信息检测是基于要素更新的关键步骤,其目的是为了发现变化的要素信息^[16]。目前的变化检测研究大都是利用新旧数据间的匹配处理,检测空间对象的变化信息^[17],匹配处理采用双向匹配方法,以获取所有的变化类型。

传统方法需逐一比较来进行判读,需要较多的空

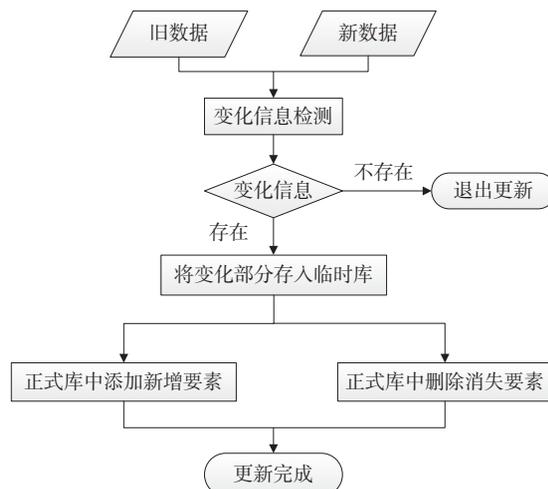


图 2 基于要素的更新流程图

间和属性对比,为了快速找到对应的要素进行对比,本文采用了格网搜索定位的方法^[16]。该方法可以快速实现定位,具体如下:

1) 分别对新旧数据增加特征点坐标属性字段(Center_X、Center_Y)以及属性汇总字段(TotalStr)。

2) 遍历要素进行新增字段的计算,这里的特征点坐标代表要素所在的位置,点要素直接取其坐标,线要素取其中点,面要素取其质心。属性汇总信息是将属性字段的字段值进行拼接,每个字段值之间用特殊符号分割,如式(1)。

$$\text{Totalstr.value} = \text{Field}(1).\text{value} + "|" + \text{Field}(2).\text{value} + "|" + \dots + "|" + \text{Field}(\text{count}).\text{value} \quad (1)$$

3) 将检测范围(X_{\min} , Y_{\min} , X_{\max} , Y_{\max})划分为 $m \times n$ 个格网,根据需要格网的宽(G_{width})和高(G_{height})可以自己设定,划分公式为:

$$\left. \begin{aligned} m &= (X_{\max} - X_{\min}) / G_{\text{width}} \\ n &= (Y_{\max} - Y_{\min}) / G_{\text{height}} \end{aligned} \right\} \quad (2)$$

4) 对新旧数据按特征点坐标进行排序,并按特征点坐标将要素匹配到相应的格网。格网编号与特征点坐标匹配方法按式(3)计算,其中 INT() 为向下取整。将格网内要素的特征点坐标、几何信息、属性汇总信息,按照排序结果依次累加到相应的网格中,如式(4)。线要素和面要素的几何信息为要素的弧段长度,点要素不需累加几何信息。

$$\left. \begin{aligned} p &= \text{INT}((\text{Center_X} - X_{\min}) / G_{\text{width}}) \\ q &= \text{INT}((\text{Center_Y} - Y_{\min}) / G_{\text{height}}) \end{aligned} \right\} \quad (3)$$

$$\left. \begin{aligned} \text{gridcenx}(p, q) &= \sum_{i=1}^k \text{Center_X}_i \\ \text{gridceny}(p, q) &= \sum_{i=1}^k \text{Center_Y}_i \\ \text{gridlen}(p, q) &= \sum_{i=1}^k \text{length}_i \\ \text{gridstr}(p, q) &= \sum_{i=1}^k \text{Totalstr}_i \\ \text{gridcount}(p, q) &= k \end{aligned} \right\} \quad (4)$$

式中, $\text{gridcenx}()$ 、 $\text{gridceny}()$ 分别表示网格中要素特征点 X 坐标及 Y 坐标的和; $\text{gridlen}()$ 表示网格中要素的弧段长度和; $\text{gridstr}()$ 表示网格中要素属性值字符串拼接; $\text{gridcount}()$ 为行号网格中要素个数; k

$$\left. \begin{aligned} \text{change cenx}(p, q) &= |N(\text{gridcenx}(p, q)) - O(\text{gridcenx}(p, q))| / O(\text{gridcenx}(p, q)) \\ \text{change ceny}(p, q) &= |N(\text{gridceny}(p, q)) - O(\text{gridceny}(p, q))| / O(\text{gridceny}(p, q)) \\ \text{change len}(p, q) &= |N(\text{gridlen}(p, q)) - O(\text{gridlen}(p, q))| / O(\text{gridlen}(p, q)) \end{aligned} \right\} \quad (5)$$

式中, $\text{change cenx}()$ 为网格中要素特征点 X 坐标和的变化率; $\text{change ceny}()$ 为网格中要素特征点 Y 坐标和的变化率; $\text{change len}()$ 为网格中弧段长度的变化率; $N()$ 与 $O()$ 分别代表新旧数据。当新旧数据中对应网格中以上 3 个指标发生变化时, 说明该网格中的要素存在变化, 需对网格内的要素进行逐一变化检测。

1.2.3 变化数据检测入库

在找到对应要素后, 以新旧要素的缓冲区重叠度 (Index) 来判断要素是否变化。

$$\text{Index} = \frac{\text{Area}[\text{Buffer}(A_i) \cap \text{Buffer}(B_i)]}{\text{Area}[\text{Buffer}(A_i) \cup \text{Buffer}(B_i)]} \quad (6)$$

式中, A_i 为新数据中的要素; B_i 为旧数据中与该要素相对应的旧要素; $\text{Buffer}(A_i)$ 表示新要素 A_i 生成的缓冲区; $\text{Buffer}(B_i)$ 表示旧要素 B_i 生成的缓冲区; $\text{Area}[\text{Buffer}(A_i) \cap \text{Buffer}(B_i)]$ 表示 A_i 与 B_i 相交的几何面积; $\text{Area}[\text{Buffer}(A_i) \cup \text{Buffer}(B_i)]$ 表示 A_i 与 B_i 相并的几何面积。在理想情况下, 如果新数据中的要素和旧数据中相比没有发生变化, Index 等于 1, 其他情况下即发生了变化。但由于数据精度等原因, 这里一般设定阈值 Q , 在阈值外数据即为变化要素。在变化检测完成后, 即可将变化要素入库。这里将变化要素分为新增要素和删除要素, 根据相应的 ID 值进行插入和删除操作, 同时将删除要素插入到对应的历史数据库中, 方便进行历史回溯查看。

2 实验与分析

为了验证提出的数据规整和增量更新方法, 本文在 Windows 环境下, 以 CAD 和 Visual Studio 2010 为开发工具, 集成 ArcGIS Engine 开发包研制了数据规整与更新原型系统, 实现了从数据规整到基于要素的地形图数据更新入库的流程, 以 1:2 000 地形图数据进行实验。

首先对 CAD 格式的地形图数据 (如图 3 所示) 进行编码、拓扑检查处理操作, 将其转化为标准数据, 然后将该数据批量转换为 GIS 数据, 转换结果如图 4。对比转换后的结果数据, 数据无缺失, 证明该数据规整程序具有可行性。

为网格中要素的数量。

5) 对新旧数据中编号相同网格的特征点 X 坐标、 Y 坐标、弧段长度、属性值拼接字符串进行对比。

按式 (5) 计算网格中几何特征的变化率。



图3 规整转换前 CAD 数据



图4 规整转换后 GIS 数据

在数据规整后即可进行数据的更新操作。图 5 为更新前数据, 通过对新旧数据进行变化检测, 发现新数据中的居民地面要素发生了变化 (如图 6 所示), 图中对变化要素以粉红色进行标注, 可以准确检测出细小的变化, 并进行相应替换。



图 5 更新前 GIS 数据



图 6 更新后 GIS 数据

实验结果显示,该系统可以快速、无损地将地形图 CAD 数据转换为 GIS 数据,且满足数据标准的要求。在基于要素的更新过程中,本文提出的自适应更新方法可高效准确地完成测量数据的更新操作。

3 结 语

本文结合修补测地形图数据进行了数据规整和增量更新的实践,设计了 CAD 数据规整的技术流程,实现了 CAD 数据向 GIS 数据的无损转换;提出了基于快速定位的变化检测方法,实现变化数据的快速更新

入库,保证了地形图数据库的现势性。实验结果表明,提出的地形图数据规整及自适应更新方法能系统、高效地实现地形图数据的更新。

本文提出的技术路线具有较好的应用效果,但由于 CAD 数据图形赋值,在数据转换时仍要进行多次检查编辑操作,且有些不常见拓扑错误需要人工筛选编辑。鉴于此,一些新的拓扑修正方法和质量控制研究将是下一步的工作任务。同时基于要素的更新方法在检测到对应网格内有变化要素存在时,仍要对网格中要素进行逐要素对比查看,在变化要素较多时更新效率会下降,因此快速定位变化要素的方法仍需进一步研究。

参考文献

- [1] Briat M, Monnot J, Kressmann T. Incremental Update of Cartographic Data in a Versioned Environment[C].22nd ICA Conference Proceedings, A Coruña, 2005
- [2] 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
- [3] 王波,张亮,孙霞.CAD 向 GIS 数据的转换入库[J].地理空间信息,2011,9(3):24-26
- [4] 袁源琳,张新长,黄健锋,等.AutoCAD 地形图数据规整入库的研究与应用[J].测绘通报,2013(5):84-88
- [5] 张叶,孙毅中,陈年松.CAD 城市基础数据到 GIS 转换的有关问题探讨[J].测绘与空间地理信息,2007,30(1):94-97
- [6] 付仲良,吴建华.多比例尺空间数据库更新技术研究[J].武汉大学学报:信息科学版,2007,32(12):1115-1118
- [7] 梁史进,张新长,郭泰圣.基于图幅的地形图数据自适应更新实现[J].地理信息世界,2014,21(6):77-81
- [8] 陈年松.基于图幅的城市基础地理信息更新研究[C].江苏省测绘学会 2009 年学术年会,南京,2009
- [9] 罗国玮,张新长,齐立新.顾及地理要素变化过程的数据增量更新方法[J].中山大学学报:自然科学版,2014(4):131-135
- [10] 张新长,郭泰圣,唐铁.一种自适应的矢量数据增量更新方法研究[J].测绘学报,2012,41(4):613-619
- [11] 杨文杰,张新长,罗国玮,等.基于要素的地形图数据更新方法研究[J].地理信息世界,2014,21(6):12-16
- [12] 孙英杰.基于变化信息文件的增量更新方法研究[D].长沙:中南大学,2008
- [13] 王蔚,陈明锐.AutoCAD 数据向 GIS 空间数据转换机制的研究[J].热带农业科学,2010,30(11):48-52
- [14] 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):44-46
- [15] 杨娜娜,张新长,黄健锋.CAD 规划成果数据 GIS 建库的技术与研究[J].测绘通报,2015(6):44-48
- [16] 罗国玮,张新长,齐立新,等.矢量数据变化对象的快速定位与最优组合匹配方法[J].测绘学报,2014,43(12):1285-1292
- [17] 吴建华,付仲良.数据更新中要素变化检测与匹配方法[J].计算机应用,2008,28(6):1612-1615

第一作者简介:张新长,教授,博士生导师,主要研究方向为数字城市理论与方法。