

长江河道 AutoCAD 数据转换为 GIS 数据的方法

安如,张琴,丁贤荣,王船海

(河海大学水文水资源学院,江苏南京 210098)

摘要:在由 CAD 数据转入 GIS 数据库时属性信息的正确获取成为当前研究数据转换的重点.探讨了利用河道地形高程测点数据生成的不规则三角网为 CAD 中不连续等高线自动赋高程值的方法:首先基于最短距离和最小方向约束进行断线连接;然后利用三角网进行计曲线赋值;再利用计曲线与首曲线的相互位置关系对首曲线进行高程赋值;最后利用等高线上的特征点加密 TIN 网,并按高程值再次连接等高线.运用该方法实现了等高线高程值的自动赋值,为 CAD 等高线数据转入 GIS 系统提供了新的研究思路.试验及工程实践表明,该方法在少量人工质量控制的前提下,实现了高效率的 CAD 不连续等高线数据向 GIS 数据的自动转换.

关键词: AutoCAD; GIS; 数据转换; 等高线赋值; 断线连接; 不规则三角网; ArcEngine

中图分类号: P208

文献标识码: A

文章编号: 1000-1980(2009)05-0523-06

在传统地图和地理信息系统的地貌表达中,等高线占有很重要的地位^[1].而在生产实践中,等高线数据的获取占用了大量的人力,很多情况下都采用矢量化地形图的方法,工作量较大.由于目前 CAD 技术在工程勘测、设计和规划管理部门得到了广泛的应用,充分利用其已有的图形数据可以大大减少工作量.但过去满足制图要求的 CAD 数据因其软件功能限制或制作过程不够严格,其要素的空间表达形式无法满足 GIS 中空间分析的要求,给转换工作带来了很大困难.如何利用计算机技术,减少人工干预,尽量实现数据的自动转换,成为目前在数据建库中的重要研究课题.

CAD 数据与 GIS 数据本身的特点决定了两者在转换过程中存在很多困难. CAD 数据强调图形信息的传输,其空间信息和属性信息都用可视化的形式表达在地图上;而 GIS 数据的属性信息是通过与空间信息相关联的数据库来表达的^[2],它不仅表达图面信息,更重要的是要能进行空间分析,这就需要将 CAD 通过图形直接表达的属性信息转换成与空间信息关联的数据库.目前文献中关于这 2 种数据之间转换的文章很多,但主要是根据 2 种图形数据的存储结构,研究制定图形数据的转换技术路线,是在基于 CAD 数据是完全正确和标准的前提下,通过读取 CAD 图形的 DWG 格式或 DXF 格式文件,将其空间坐标信息和以标注形式存储的属性信息分别读取出来,在 GIS 系统下开发数据接口程序读入文件,生成 GIS 数据库,但没有对 CAD 数据中存在的问题进行分析和处理,也没有充分利用 GIS 的空间分析功能.

对于等高线数据的转换,其空间信息的转换比较容易,但高程值的获取则成为难点.已有的算法一般都是基于闭合等高线之间的拓扑关系来自动赋值.但由于现有地形图上的各类要素相互交叉及覆盖,等高线中含有大量的断点,且等高线稀疏的地方会产生断线,在线状要素稠密的地方也会产生粘连,无法通过常规算法赋值^[1].本文以江苏省长江河道监测管理项目为研究背景,在已经获取水下高程测点数据的基础上,利用测点数据和 ArcGIS 提供的数据分析和数据管理功能,对不连续的等高线进行半自动赋值,供河道地形分析使用.采用的开发语言为 ArcGIS 的嵌入式 GIS 组件库 ArcGIS Engine,开发环境为 VC++6.0.

1 数据源及存在的问题

试验提供了长江河道测区范围内所有测点的文本格式数据. ArcGIS 中提供了由点数据自动生成等高线的功能,可以利用测点数据在 ArcMap 中直接生成等高线.但直接生成的等高线是将计算的 2 点之间理论上应该通过的等高线个数等间距地分配给每根等高线(图 1),因而形状平直生硬,不能正确表达水下地形复杂

收稿日期: 2009-04-08

作者简介: 安如(1963—),女,江苏淮安人,教授,博士,主要从事遥感与地理信息系统研究.

的地区,而 CAD 格式的等高线经过了样条插值和人工识别修正,形状平缓圆滑(图 2),走向更符合实际地形情况^[3],因此试验中希望能将现有的 CAD 格式等高线数据原样转入 ArcGIS,进行空间分析研究。

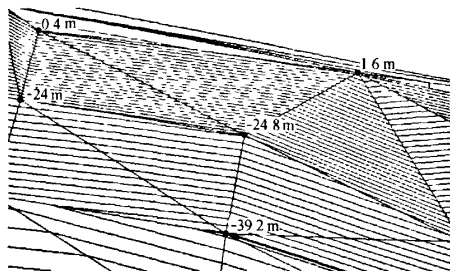


图 1 由 ArcMap 自动生成的等高线形态

Fig. 1 Contours automatically generated by ArcMap

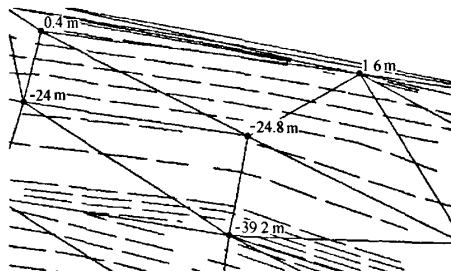


图 2 CAD 中等高线的实际形态

Fig. 2 Contours in CAD format

由于 CAD 地图没有严格按标准生成, CAD 数据源存在较多问题,给自动转换为 GIS 数据带来了困难。

a. 数据分层不规范. 按规范每个图层只对应 1 种要素,但在实际生成过程中制图人员可能会比较随意地添加图层或在不对应的图层上进行修改,造成数据分层的混乱. 1 种要素有可能分放在不同的图层中,而 1 个图层中也会有多种要素出现,给数据整理工作带来困难。

b. 存在多种线型. CAD 强大的制图功能为其提供了多种线型,用户可以根据需要自己定义线型,有些线型转入 GIS 系统后会丢失或变形^[4-6]。

c. 等高线制作不规范. 为了达到图幅的整体美观效果, CAD 在进行图幅整饰时没有考虑到数据之间的空间逻辑关系,因此图中存在大量手工添加的无意义短线,等高线出现多处自相交、重复画线等现象,不符合 GIS 数据的要求。

d. 数据冗余量大. 由于其他要素压盖、等高线过密等原因,等高线在矢量化后被打断成若干个目标^[7],且线段之间间隔不等,造成数据大量冗余,这些断线放入不规则三角网中赋值会影响数据的正确性。

e. 存在混合等高距. 在河道地形平缓的地方等高距为 1 m;在地形陡峭的地方等高距为 5 m,且 1 m 等高距和 5 m 等高距等高线的分布毫无规律。

2 CAD 等高线数据转换为 GIS 数据的自动赋值

2.1 连接断线

水下地形 CAD 数据中等高线线型用虚线表达,理论上应完整且闭合的线对象被打散成许多间隔不等的短线(图 3),造成大量数据冗余,且这些断线转入 GIS 系统后建立的空间拓扑关系意义不大,给空间分析造成较大困难。

为了减小数据冗余,尤其是为了提高赋值的正确性,首先需要对不连续的等高线进行断线连接. 现有的断线连接方法有基于最小距离或方向差异最小的局部连接法、基于费曼码的断线连接、基于数学形态学膨胀与收缩原理的断线连接法、基于等高线空间关系结构约束的最大集团图搜索法以及顾及拓扑关系的连接方法等^[8]. 本文主要目的在于提高等高线赋高程值的正确性,方法如下:

- 以每根线对象终点为圆心,一定长度阈值为半径做缓冲,计算落入缓冲区内其他线对象的个数^[9];
- 若落入缓冲区的线对象多于 1 根,通过减小角度阈值找到沿搜索线延线方向角度变化最小的线;
- 若没有线对象落入缓冲区内则增大搜索半径,循环直到找到唯一一根线对象落入缓冲区内;
- 循环终止条件:循环中设置一变量 k 用于计算满足连接条件的线的条数,另一变量 m 用于计算落入缓冲区的线的条数,当 k 为零而 m 不为零时跳出循环,结束搜索过程,程序从下一根线开始搜索。

ArcGIS Engine 中连接断线主要用到的接口为 IPointCollection. 通过搜索线图层对应的 FeatureClass 中的每一个 Feature, 获得 Feature 对应的 IGeometry 类型,再将 IPolyline 类型的线对象通过 Coclass 转变为 IPointCollection 类型. 做缓冲时用到 IRelationalOperator 接口中的 Contains 方法,判断连接线的起点是否包含在



图 3 不连续的等高线

Fig. 3 Discontinuous submarine contours

缓冲区内,当搜索到唯一一根连接线时,IPointCollection 中提供的 AddPointCollection 方法可以将连接线上的所有点集加入到搜索线的点集中,即可将 2 条线合并成 1 条。

试验结果表明,角度和距离阈值的变化是相互制约的,当角度一定,增大缓冲区半径时,搜索到的线对象可能并不是实际应该连接的对象。而在距离一定的情况下,角度增大时也会出现连错的情况,见图 4。

2.2 利用不规则三角网进行计曲线赋值

首曲线又称为基本等高线,即按相应比例尺地图规定的基本等高距,由基准面起算而测绘的等高线。计曲线又称加粗等高线,即为了读图时判读等高线高程的方便,由基准面起算,每隔 4 根首曲线加粗描绘的首曲线。

试验中采用的三角网由 ArcMap 中根据河道测点数据生成,在给线对象赋高程值的过程中,只有当 CAD 的等高线与由 ArcMap 生成的等高线形状相符时才可以正确赋值(图 5),而在很多情况下由于两者算法不同,且 CAD 数据经过人工修正,因此形态不会完全一致。在两者不能重叠的情况下,赋值的错误率相对较高(图 6)。

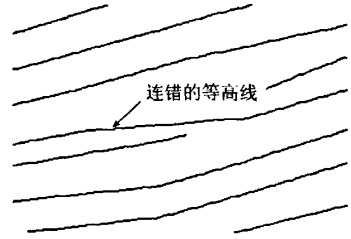


图 4 角度约束和距离约束的矛盾

Fig. 4 Results of broken-line connection based on constraint of minimum distance and angle

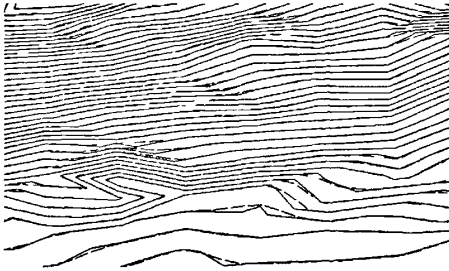


图 5 位置相符的情况下可以正确赋值

Fig. 5 Correct value assignment of two similar contour shapes

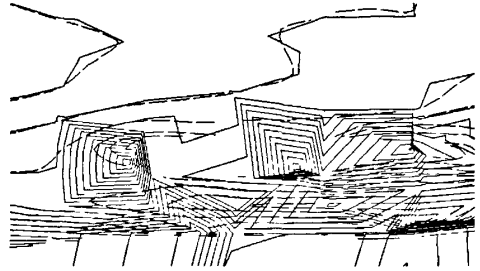


图 6 位置不符的情况下会赋错值

Fig. 6 Wrong value assignment of two dissimilar contour shapes

实际计算过程是将线对象放入三角网中,通过获取线上每个结点在 TIN 中的高程之和求平均后赋给线对象,作为此等高线的实际高程值^[10]。

等高线经过断线连接后可以由多个高程控制点控制其值,提高了赋值的正确率,但仍然存在一些问题:当一根线的起点和终点距离跨度太大,而线对象上节点个数较少时,有可能由于某些高程异常点使得整根等高线受异常值的影响较大而偏离正确值。针对这些问题,可以通过计算等高线上每个节点的高程均方差得到每个点的高程偏离平均高程值的程度,去除线上无效点对高程值的影响;对于等高线长度较长、节点过少的情况则采用加密点算法,在线上等距离地加上控制点,获得这些点的高程再求平均,从而提高赋值的正确率。

赋值的过程主要利用 ITinSurface 接口下的 GetElevation 方法获取线上每个节点的高程值,求和后取平均即为线对象的高程。

2.3 利用计曲线和 TIN 三角网边线判断首曲线高程

1:10000 地形图等高距通常为 1 m,若数据较规则,可以通过给定起始等高线高程,按等高线空间拓扑关系顺序自动赋值^[11]。但实际数据相当复杂,地势较陡的地方等高线密集,部分等高线被省略,等高距由通常的 1 m 变为 5 m,直接利用不规则三角网赋值,出错率较高。为了尽可能减小手工改正量,本文研究了利用计曲线和首曲线的空间逻辑关系判断首曲线高程的方法,即每 2 条计曲线之间应该通过 4 根首曲线,若多于或少于 4 根则不判,在正好通过 4 根首曲线的情况下依次赋值。算法思路如下:

a. 分别提取计曲线和首曲线。

b. 将计曲线放入三角网中赋值,用四舍五入的方法将计曲线高程值控制为 5 的整数倍,对提取出的首曲线进行连线后复制到计曲线图层中。

c. 在 TIN 图层中对不规则三角网的每条边线进行搜索,在等高线图层中找到与此边线相交的所有线对象,计算相交线的个数。

d. 判断与 TIN 边线相交的线对象中有无计曲线,若有,则将 TIN 边线以计曲线为界分为几个部分,分别计算相交线的个数,若正好等于理论上应该相交的线个数,则从最低点开始依次赋值,对重新赋值的等高线做标记,程序下次搜索时对已经判过的线不再重新赋值。

经过以上处理,还存在少部分处于地形较陡区域较短的断线无法与 TIN 边线相交,这时考虑利用已经正确赋值的计曲线上的端点及特征点再构造三角网,可以提高原三角网的精度和密度^[12],使得剩下的断线可以与 TIN 边线相交,从而利用原算法进行正确判读。

2.4 按高程值再次进行连线

当所有等高线都由 TIN 赋值并通过以上算法进行正确性判别后基本可以正确赋值,但仍存在一些无法判读高程的线对象,此时需要人工检验。在确保高程属性正确后可以通过高程属性相同这一搜索条件再次进行断线连接,使得等高线对象尽量完整。算法仍采用前述方法,加上高程值相等这一约束条件即可。转换技术流程见图 7。

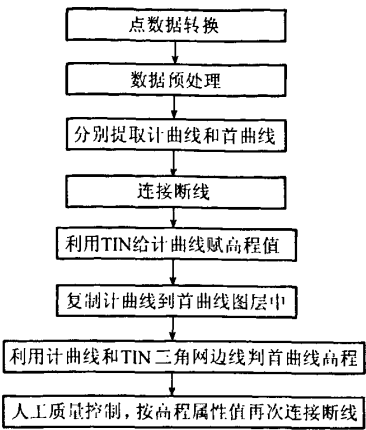


图 7 等高线赋值流程

Fig. 7 Flow chart of value assignment of contours

3 试验与讨论

试验在连接不连续等高线的过程中研究了最小距离约束和方向约束相结合的方法,在线与线之间拐角不大、距离不超过阈值的情况下取得了较好的效果。图 8 为等高线原图,连线结果见图 9。



图 8 等高线原图

Fig. 8 Original map of contours



图 9 等高线连线结果

Fig. 9 Results of connected contours

利用 TIN 边线和等高线计曲线与首曲线之间的关系对赋值错误的等高线进行了错误检验,纠正了可以与 TIN 边线相交的等高线高程值,如图 10 和图 11 所示。

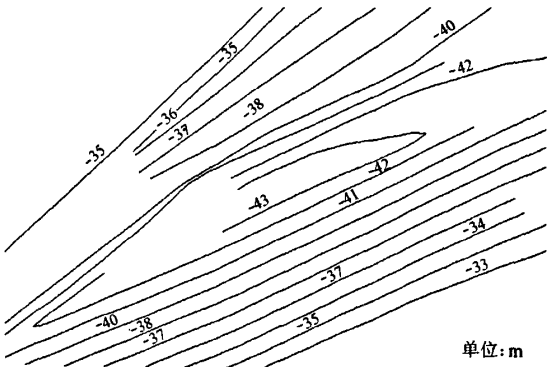


图 10 计曲线判首曲线原图

Fig. 10 Original map of using index contours to assign elevation values to intermediate contours

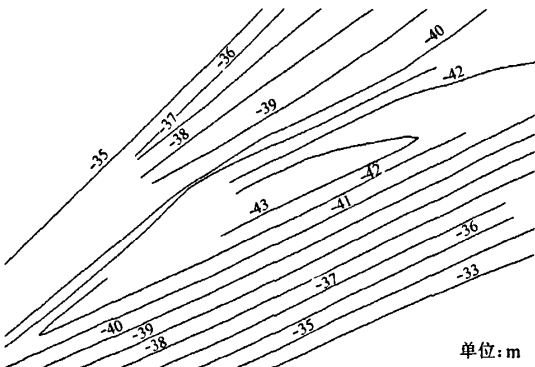


图 11 计曲线判首曲线结果

Fig. 11 Resulting map of using index contours to assign elevation values to intermediate contours

等高线经过以上操作后基本可以正确赋值,但仍需要少量人工检验的过程,最后按高程属性连线后的等高线效果较好,线对象已基本完整,可以满足 GIS 中空间分析的需要,如图 12 和图 13 所示。



图 12 转入 ArcGIS 的等高线原图

Fig. 12 CAD map before being converted to ArcGIS

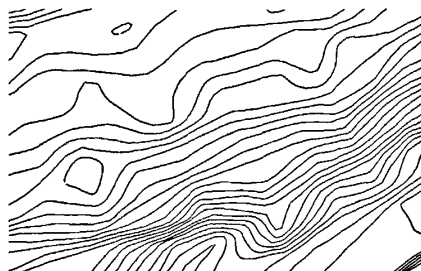


图 13 经属性连线的最终结果

Fig. 13 Final map after connection using elevation attribute values

4 结 论

本文以江苏省长江河道监测管理项目为研究背景,对全长江河道的 CAD 类型测点及等高线数据进行了批量处理,获得高程属性后转入 GIS 数据库,供河道地形分析使用。试验过程已全部投入实际生产操作并通过检验,试验结果证明,该方法在少量人工质量控制的前提下,可以获得比较理想的效果。

由于数据本身的复杂性,批量处理时只能针对普遍存在的问题提出解决方案,对于一些特殊情况则还需依靠人工检验来确保数据的正确性。本文提出,距离和角度约束取得最佳值的方法是在控制出错率的前提下尽可能多地连线,但落在 TIN 网外部的等高线,以及无法与 TIN 边界线相交的短线如何检验并正确赋值的问题还没有较有效的解决方法,有待于进一步研究。

参考文献:

- [1] 王涛,毋河海.等高线拓扑关系的构建以及应用[J].武汉大学学报:信息科学版,2004,29(5):438-442. (WANG Tao, WU He-hai. Construction and applications of topological relation among contour lines[J]. Editorial Board of Geomatics and Information Science of Wuhan University, 2004,29(5):438-442. (in Chinese))
- [2] 王铁军,麻玺琳,卜昭铭.从 CAD 到 GIS 的数据转换[J].东北测绘,1999,22(3):12-13. (WANG Tie-jun, MA Xi-lin, BU Zhao-ming. Data transformation from CAD to GIS[J]. Northeast Surveying and Mapping, 1999,22(3):12-13. (in Chinese))
- [3] PARK H, JUNG H B, KIM K. A new approach for lofted B-spline surface interpolation to serial contours[J]. The International Journal of Advanced Manufacturing Technology, 2004, 23(11):889-895.
- [4] 周小成,焦道振.基于 Geodatabase 的 CAD 数据到 GIS 的解决方法[J].现代测绘,2004,27(6):1056-1063. (ZHOU Xiao-cheng, JIAO Dao-zhen. Study on format transformation from CAD to GIS based on of geodatabase[J]. Modern Surveying and Mapping, 2004, 27(6):1056-1063. (in Chinese))
- [5] 胡胜华,何宗宜,陶利佳. AutoCAD 与 GIS 数据转换的研究[J]. 测绘通报, 2007(12):44-47. (HU Sheng-hua, HE Zong-yi, TAO Li-jia. Research on data conversion between AutoCAD and GIS[J]. Bulletin of Surveying and Mapping, 2007(12):44-47. (in Chinese))
- [6] 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):4-17. (CHEN Neng, SHI Bei-qi. Conversion technique of AutoCAD topographic map data to GIS spatial data[J]. Bulletin of Surveying and Mapping, 2005(8):4-17. (in Chinese))
- [7] 周献中,郑华利.基于可变形模型及区域流向分析的等高线自适应矢量化算法[J].计算机学报,2004,27(8):15-17. (ZHOU Xian-zhong, ZHENG Hua-li. Automatic vectorization of contour lines based on deformable model and field flow orientation[J]. Chinese Journal of Computers, 2004, 27(8):15-17. (in Chinese))
- [8] 陈丹,龙毅.基于知识的等高线断点连接方法[J].测绘信息与工程,2003,28(3):31-33. (CHEN Dan, LONG Yi. A knowledge-based breakpoint processing in contour extraction[J]. Whusm Bulletin of Science and Technology, 2003, 28(3):31-33. (in Chinese))
- [9] 陈争光,吴裕树,王玉芳.一种新型的地形等高线矢量化方法[J].计算机工程与应用,2004(3):84-86. (CHEN Zheng-guang, WU Yu-shu, WANG Yu-fang. A new schema of vectorization of contour[J]. Computer Engineering and Applications, 2004(3):84-86. (in Chinese))

- [10] 张海堂,罗睿,郭建星.基于三角网渐进式简化的等高线多尺度综合[J].测绘信息与工程,2004,29(5):11-13.(ZHANG Hai-tang, LUO Rui, GUO Jian-xing, et al. Multi-scale generalization of contour lines based on progressive simplification of tin[J]. Wtasm Bulletin of Science and Technology, 2004, 29(5):11-13. (in Chinese))
- [11] 张琳琳,武芳,王辉连.等高线空间关系的确定及应用[J].测绘通报,2005(8):19-22.(ZHANG Lin-lin, WU Fang, WANG Hui-lian. Representation of spatial relationships of contour lines and it's application[J]. Bulletin of Surveying and Mapping, 2005(8):19-22. (in Chinese))
- [12] MARC V K. Algorithms for triangulated terrains[J]. Lecture Notes in Computer Science, 1997, 1338:19-36.

Data conversion from AutoCAD to GIS for watercourse of Yangtze River

AN Ru, ZHANG Qin, DING Xian-rong, WANG Chuan-hai

(*Department of Geographical Information Sciences, College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China*)

Abstract: During the conversion from CAD data to GIS database, the obtaining of precise attributes of spatial data is a hot topic in the current researches on data conservation. The method of automatic assignment of elevation values to CAD discontinuous contours was discussed by use of the triangulated irregular network (TIN) generated based on the measured data at submarine elevation points. The method is described as follows: 1) the broken-line is connected based on the constraint of the minimum distance and angle; 2) the assignment of elevation values to discontinuous submarine index contours are performed by use of the TIN; 3) the elevation values of intermediate contours are determined by use of the relationship of positions between the index contours and the intermediate contours; 4) the TIN is encrypted by means of the characteristic points of the contours, and the contours are again connected according to their correct values of elevation. The automatic value assignment of attributes to the elevation of contours was realized using the present method. A novel thought was given for the data conservation from CAD contours to GIS. The tests and engineering practice indicate that the present method is efficiently of automatic conversion from CAD discontinuous contour data to GIS data under the prerequisite of a little manual control upon data quality.

Key words: AutoCAD; GIS; data conversion; value assignment of contours; broken-line connection; triangulated irregular network; ArcEngine