

自动矢量化后等高线处理研究

张晓远¹, 赵晓琳²

(1. 中山大学遥感与地理信息工程系, 广东 广州 510275; 2. 辽宁地质工程职业学院, 辽宁 丹东 118008)

摘要:等高线是地形图中反映地貌形态特征的基础地理信息数据, 其自动矢量化后普遍存在毛刺和断点连接等问题, 该文提出基于毛刺生长点的邻域搜索剔除毛刺法和复合连接法等解决方法。在 CAD 环境下通过 VBA 编程对技术方案进行验证, 结果表明该方法具有连接速度快、准确率高等优点。

关键词:等高线; 自动矢量化; 毛刺; 断点连接; 复合连接

中图分类号: P208 **文献标识码:** A **文章编号:** 1672 - 0504(2008)03 - 0028 - 04

等高线是反映地形图中区域地貌形态的基础地理信息数据, 其数据量约占整幅图的 40 % 以上, 对其矢量化的工作量相当大。随着 GIS 技术的发展和应用的深入, 大量的地形图亟待数字化, 传统的数据输入方式在速度和精度等方面已不能适应这一发展趋势, 对栅格图进行自动矢量化则是一种非常有效的方法。但由于图纸自身质量、扫描质量、矢量化算法和其他地物的干扰等因素, 如因等高线有毛刺、不光滑造成数据存储空间的冗余, 由于图像各类要素相互交叉、覆盖及图像颜色不均匀等导致等高线自动矢量化后出现断线等, 因此需对地形图自动矢量化后的数据进行后续处理。

由于等高线矢量化后处理研究针对性强且应用领域相对狭窄, 国内外相关研究较少。但矢量化后处理作为扫描矢量化数据向 GIS 数据转化的过渡阶段, 不仅可打印输出高精度数字地图, 也可为 GIS 提供高精度的前端数据。本文侧重于等高线自动矢量化后毛刺剔除、断线连接及接边检查等方面, 并基于 AutoCAD 制图软件, 利用 VBA 二次开发工具实现了研究方案。

1 等高线矢量化后数据的预处理

毛刺主要是由于地形图上地貌符号与等高线相交、线条不光滑等原因产生的, 在图像细化时线条边缘也可能产生毛刺, 给断裂等高线的连接带来困难。地形图中的毛刺通常线条短小、分布零散, 剔除毛刺的方法较多, 但都不成熟。本文提出基于毛刺生长点的邻域搜索剔除毛刺法, 通过自动剔除和人工交互式剔除两种方法编程实现。

1.1 基于毛刺生长点的邻域搜索剔除毛刺法

在整幅地图内搜索所有等高线的端点, 包括伪

端点(即误将毛刺作为端点), 记录搜索的各端点位置。每搜索到一个端点后需从这一像素开始搜索毛刺生长点, 步骤如下^[1]:

(1) 利用八邻域寻找与此端点像素相连的像素位置。在判断是否为端点时必须满足图 1 的条件, 即如果一个像素被认为是端点, 以其为中心的八邻域必须是图 1 中情况之一。其中“C”代表中心像素, “1”代表地图地物像素, “0”代表地图背景像素。

1	0	0	0	1	0	0	0	0	0	0	0
0	C	0	0	0	C	0	0	C	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0

(a) 模板 1 (b) 模板 2 (c) 模板 3 (d) 模板 4

0	0	0	0	0	0	0	0	0	0	0	0
0	C	1	0	0	C	0	0	C	0	0	0
0	0	0	1	0	0	0	1	0	0	0	1

(e) 模板 5 (f) 模板 6 (g) 模板 7 (h) 模板 8

图 1 用于寻找端点的八邻域模板
Fig. 1 Eight neighboring palette to be used to seek endpoint

由端点的定义可知, 每个端点像素在其八邻域内只能寻找到一个相邻地图地物像素点, 针对此点判断其是否为毛刺的生长点。判断方法为: 以此像素点为中心取八邻域, 剔除搜索到的前一个点, 然后利用式(1)计算八邻域像素中地图地物像素到地图背景像素的变化次数, 变化次数大于 1 说明此时找到的点就是毛刺的生长点。

$$N(8) = \sum_{k=C}^8 (f(x_k) (1 - f(x_k))) \quad (1)$$

式中: x_k 的位置如图 2 所示, $x_1 = x_0$, $C = \{1, 2, 3, 4, 5, 6, 7, 8\}$ 。

(2) 若利用式(1)计算的值等于 1, 并不说明已找

收稿日期: 2008 - 01 - 03; 修订日期: 2008 - 03 - 25

作者简介: 张晓远(1983 -), 男, 硕士研究生, 研究方向为 GIS、RS 应用及水利信息化。E-mail: zxyuan1@126.com

到毛刺生长点,还需利用四邻域模板继续搜索判断。如果以此像素为中心的四邻域中,除搜索过的点外还存在两个以上地图背景像素,则说明此中心像素点就是毛刺生长点,此时四邻域中除搜索过的点外有且仅有一个像素为地图地物像素,记录此像素为毛刺生长点。

x_4	x_3	x_2
x_5	x_0	x_1
x_6	x_7	x_8

图 2 公式(1)中 x_k 的位置

Fig. 2 The explanation of x_k 's position in formula(1)

(3) 从步骤 2 中找到的像素剔除搜索过的像素,再利用式(1)及步骤 2 判断其是否毛刺的生长点。毛刺的长度一般不超过 4 个像素,所以运用步骤 1 到步骤 3 最多搜索 3 个像素,如果在下一个像素内没有找到毛刺的生长点,则说明搜索到的端点是图像中的真正端点,不需做任何处理。如果在 3 个像素内找到了毛刺生长点,则说明搜索到的这个端点是毛刺产生的“伪端点”,此时应将“伪端点”及其与毛刺生长点之间的像素删除,即将毛刺删除;然后在全图范围内对所有搜索到的端点进行上述处理,便可将自动矢量化生成的等高线上产生的毛刺剔除。

通常将长度小于指定阈值的线段作为毛刺剔除,因此在剔除毛刺的过程中,阈值的设定是关键环节。如果阈值偏小,则达不到剔除的效果,结果仍然存在很多小段线;如果阈值偏大,则可能会剔除一些连续的断线,造成断点间的距离过大,影响缓冲区的构建,从而导致接边时判断失误。阈值的选取应经过反复检验。

1.2 具体实现方法

(1) 自动剔除。实验中进行自动剔除的地图要素是:短小且分布零散的曲线;与等高线相交的地貌符号,如林地、格网干扰线等。对于前者,可以根据其特点,遍历地形图上所有的曲线,将线长小于指定阈值、线段趋势与邻近等高线趋势相差较大者剔除。以林地为例,其特点是在地形图中分布广且零散,而且很多线条与等高线相交,如不剔除,则会影响接边的质量。林地符号一般在起点和终点重合或相距较近,根据其特点,遍历所有曲线,将起点与终点距离小于指定阈值的作为毛刺剔除;对于格网干扰线,其斜率一般接近 90° 或 0° ,且格网间距离相差 1 km,先从格网线的起点到终点画一条红线,凡是与该红线

斜率(或斜率的倒数)相近或重合的线都应剔除。

(2) 人机交互式剔除。有些与等高线相交的地貌符号没有规律可循,且其在地形图中数量相对较少,最好采用人机交互式剔除,包括高程注记和房屋等。

2 断线连接及接边处理

连接处理的关键是寻找需要连接的等高线,通过对地形图中的等高线进行分析,发现需要连接的等高线端点附近的曲线大多走向相同或相近,且需要连接的两端点的间隔不大。根据最小点对法思想^[2],通过端点趋势和最短距离两种方法可实现等高线的断线连接,但两者都有其局限性,尤其是对断线进行批量连接时会出现多处误连接的现象,准确率较低。本文对最小点对法进行改进,提出一种复合连接方法,利用缓冲区缩小连接范围,结合端点趋势法和最近距离法在缓冲区内层层筛选对断线进行批量连接,然后对等高线进行接边检查,大大提高了连接的准确率。

2.1 建立矩形缓冲区

构建缓冲区可缩小连接线段范围、缩减连接处理的运算量,以达到加快接边速度的目的。遍历所有等高线寻找需要连接的等高线时,采用构建矩形缓冲区进行约束判断的方法筛选,即在缓冲区内等高线中随机指定一条等高线,用分布轴线法中的线缓冲区边界生成算法^[3]求出距端点最近的 3 个点(包括端点)拟合的直线,再以该等高线端点在拟合直线上的投影点为基点,通过极坐标算法生成一个矩形缓冲区,把起始端点落入矩形缓冲区的其余等高线作为连接处理的参考曲线,排除没有落入该缓冲区的曲线,缩减可连接线段范围及连接处理的运算量,达到加快连接速度的目的。应注意的是,断点的间距不定,因此缓冲区的范围在实际操作中可适当改变。

(1) 求曲线上 3 点拟合直线的斜率。矩形缓冲区是沿着曲线的走向建立的,即在建立矩形缓冲区前首先求出随机指定曲线上距端点最近(包括端点在内)的 3 个点拟合直线的斜率。分别对曲线的首部 3 个点和尾部 3 个点同时拟合直线,相当于一次考虑 3 条以上的曲线,即同时对 1 条曲线建立两个缓冲区,减少不必要的重复循环,从而加快处理速度。离散点群的空间分布趋势(走向)通过分布轴线法计算,原理如下:

对于离散点群 $P_i(x_i, y_i)$ ($i = 1, 2, \dots, n$),可以拟

合 1 条直线 $L: ax + by + c = 0$, 点群相对于 L 的距离反映了离散点群的离散程度, 而 L 的走向则描述了点群的总体走向, 其中:

$$\frac{a}{b} = \frac{(S_{xx} - S_{yy}) \pm \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}}$$

$$c = -a\bar{x} - b\bar{y}$$

式中: $S_{xx} = (x_i - \bar{x})^2$, $S_{yy} = (y_i - \bar{y})^2$, $S_{xy} = (x_i - \bar{x})(y_i - \bar{y})$ 。若 $S_{xx} > S_{yy}$, 直线 $ax + by + c = 0$ 的斜率的绝对值小于 1, 应有 $|a| < |b|$, 其保证了所求直线代表分布椭圆的长轴方向。一般令 $b = 1$, 则可给出一组 (a, b, c) 。

可以看出, 以点群拟合的直线必经过点群的几何中心, 但不一定通过点群中的点。

(2) 构建缓冲区。应用极坐标法, 根据曲线的已知点求出矩形缓冲区的 4 个顶点坐标^[4]。如图 3 所示, 选择已有线段 AB , (x_A, x_B) 和 (y_A, y_B) 分别为两端点的坐标, d_x, d_y 分别为线段两端点的 x, y 轴间距离; 以 AB 为极轴, 输入角度 α 和长度 d , 求点 P 坐标。

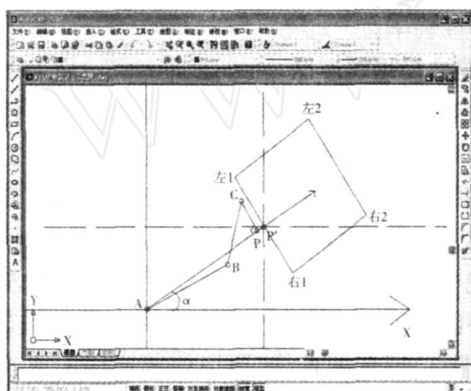


图 3 建立缓冲区

Fig. 3 Construction of buffer area

第一步: 计算有向线段 AB 的长度 L :

$$L = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}$$

第二步: 根据有向线段 AB 计算坐标:

$$d_x = x_B - x_A, d_y = y_B - y_A$$

第三步: 以点 A 为基点旋转有向线段 AB , 则:

$$d_x = d_x \cos \alpha - d_y \sin \alpha, d_y = d_x \sin \alpha + d_y \cos \alpha$$

第四步: 求取 P 点坐标:

$$x_P = x_A + d_x * d/L, y_P = y_A + d_y * d/L$$

据上述方法求曲线上 3 点拟合直线。通过 C 点作拟合直线的垂线, 垂足为 P , 以 P 为基点沿该直线延长 1 m 至点 P , 然后以 P 为矩形缓冲区一条边上的中点生成一个距 P 左右各 30 m 、沿直线方向 70 m 的矩形缓冲区。延长 AP 至 P 作矩形缓冲区是为了防止将指定等高线本身选中导致不必要的循环和判断, 由此可选中落入缓冲区内的所有参

考曲线。

缓冲区的端点按顺时针顺序依次为左 1、左 2、右 2、右 1 (图 3)。根据极坐标算法依次求出矩形缓冲区各角点坐标, 在线段的首尾分别生成缓冲区。按照上述原理编程运行后, 用其中 1 条线段进行实验, 缓冲区显示效果如图 4。

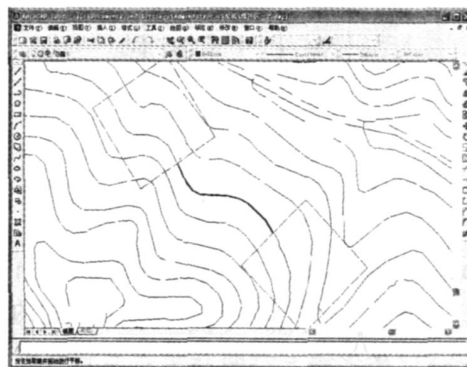


图 4 缓冲区的显示

Fig. 4 The display of buffer area

2.2 断线连接筛选

距离最小法和方向差异最小法具有如下缺点: 容易误连接断点, 尤其是在断点距离较大时, 误连接的可能性更大^[5]。因此, 本文将这两种方法结合, 以提高连接的正确率。在矩形缓冲区中选中的参考等高线与随机指定的等高线进行连接还需经过趋势比较和距离判断。

(1) 趋势比较。在地形图复杂的等高线关系分析中, 发现无论等高线如何断裂, 都沿一定的趋势存在, 即等高线曲线走向基本一致。基于这种思想, 文中对端点落入矩形缓冲区内的参考连接曲线进行再次筛选。首先, 用分布轴线法对指定曲线和所有参考曲线分别以端点附近的 3 个点 (包括端点) 拟合 1 条直线并计算其斜率 (即走向); 其次, 需分别计算指定曲线拟合的直线与参考曲线拟合的直线的斜率之差, 运用冒泡法提取与指定曲线斜率差相对较小的参考曲线, 这样的曲线可能不止 1 条, 对其进行进一步筛选。

(2) 距离判断。分别计算落入矩形缓冲区内的参考曲线端点与随机指定曲线端点间的距离, 分别进行比较判断, 运用冒泡法提取与随机指定曲线端点间距离最小的参考曲线作为需要连接处理的等高线曲线, 设定适当的阈值进行连接。

通过建立缓冲区、趋势比较和距离判断进行层层筛选, 得到符合条件的参考曲线, 使之与指定曲线连接, 最后实现等高线的批量连接。这种方法加快了接边的速度, 提高了接边的准确率。经过上述计算处理后, 图 5 为同一区域连接处理前和断线连接

后示意图,可以看出,本文提出的矢量化后处理方法达到了理想的效果。

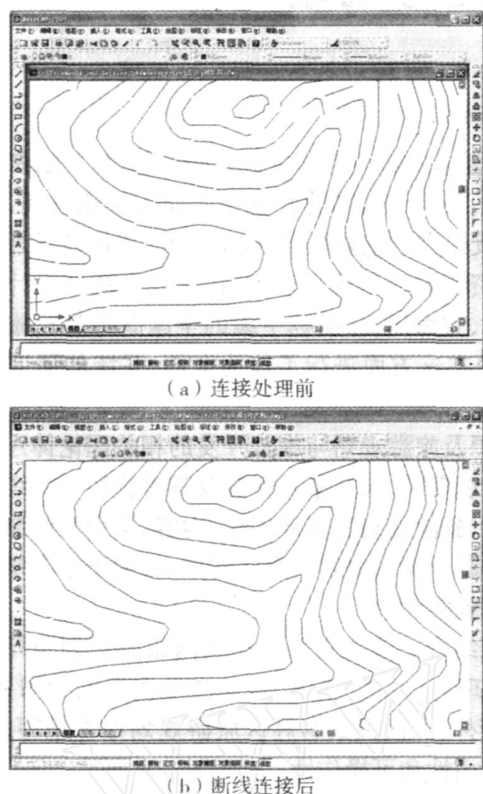


图 5 连接处理前和断线连接后示意

Fig. 5 The sketch map before and after disposal of joining

需强调的是,在 AutoCAD VBA 开发环境中,曲线都有起点和终点。无论是进行趋势比较、距离判断还是最后的连接处理,都涉及曲线的起始方向。对于随机指定的曲线和参考连接曲线,其相应位置和关系通常属图 6 中显示的 4 种情况。

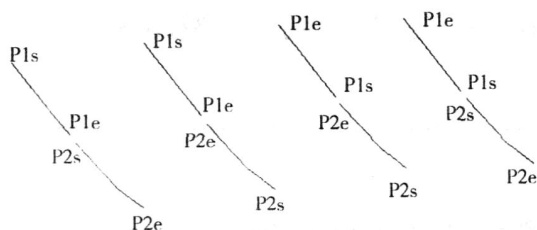


图 6 曲线之间的 4 种关系

Fig. 6 Four relations between the curves

图 6 中下标为 s 表示曲线的起点,下标为 e 表示曲线的终点。由于存在上述 4 种情况,在趋势判断、距离比较以及最后的断线连接处理过程中需要分别予以考虑。整个连接过程在 AutoCAD VBA 开发环境中进行,其开发流程如下:建立等高线选择集;输入阈值;随机指定 1 条等高线作为基准连接;将其余等高线设为参考连接曲线;建立矩形缓冲区进行第一次筛选;比较端点趋势走向进行第二次筛选;比较距离进行第三次筛选;通过筛选,将最终得到的参

考等高线与指定等高线连接;反复执行以上步骤,连接所有断线。

2.3 连接后的接边检查

等高线自动接边后的理想状态应该为:闭曲线或边界上没有可连接等高线的开曲线^[6]。接边后对地形图中的开曲线进行检查,凡是端点没有落在图幅边上的曲线都进行高亮显示,并查明原因。

经实验可知,除了端点落在边界上未连接的曲线,以下原因造成断开的等高线尚未连接:1)等高线与地物或高程注记难以区分,自动矢量化时都作为等高线,等高线走向发生异常;2)自动矢量化软件(如 R2V)本身的局限性使得少数断开的等高线之间构成一个整体,用任何接边方法都无法实现断线连接;3)由于缓冲区不够大,使断开距离较大的等高线难以连接,需要扩大缓冲区以达到连接的目的;4)由于本文提出的断线连接方法不适于角度大的拐弯处的断线连接,因此图 5 中山脊、山谷等复杂地形区没有实现断线连接。

3 结语

本文对等高线自动矢量化后的数据预处理、断线连接和接边检查等方案进行研究,在数据预处理的毛刺剔除方面,提出基于毛刺生长点的邻域搜索剔除毛刺法;在断线连接及接边处理部分,提出端点趋势和最短距离相结合的连接方法,将符合要求的断裂等高线连接,避免了常规连接方法的盲目性。实验证明曲线连接效果较好,具有良好的通用性;缺点是对最小点对法进行改进实现断线连接,虽减少了误连接率,但在断点距离较大时,同样存在误连的可能。如何改进算法,实现更高效的断线连接,有待进一步研究。

参考文献:

- [1] 安然,张少军. 字符识别中毛刺的去除方法[J]. 计算机技术与发展, 2007, 17(9): 136 - 137.
- [2] 罗以宁,李毅,杨凯. 金相图像分析中一种改进的断点连接方法[J]. 四川大学学报, 2002, 39(6): 1028 - 1029.
- [3] 张宏,温永宁,刘爱利,等. 地理信息系统算法基础[M]. 北京: 科学出版社, 2006. 262 - 264.
- [4] 于真. 极坐标法在高速铁路测量中的应用[J]. 科技咨询导报, 2007(17): 62 - 63.
- [5] MARDLE D F. Some thoughts on the integration of spatial analysis and geographic information systems[J]. Geographic Systems, 2000(18): 107 - 110.
- [6] 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.

(下转第 32 页)

简评《旅游资源开发与规划》

王益谦

(四川大学,四川 成都 610065)

任何产业的发展,首先遇到的是资源问题。旅游资源是旅游业赖以生存与发展的物质基础,旅游资源的合理开发和有效保护对旅游业健康、持续发展有着决定性意义。近年来,随着我国旅游业的快速发展,旅游资源的开发与规划日显重要,旅游资源与开发、旅游规划与开发研究的重要性、实践迫切性、市场需求性日益凸显。如何使高等院校旅游专业学生掌握最新的国内外旅游资源开发与规划理念,在实践中能够设计出更好的规划作品,是旅游教育中面临的一个前沿问题。由陈兴中、方海川、汪明林主编的全国高等院校旅游系列规划教材《旅游资源开发与规划》(2005 年科学出版社出版)在这方面进行了较为全面、系统的研究和探索,为旅游资源开发与旅游产业发展提供了一个全新的视角。

该书令人耳目一新的是运用现代景观学理论阐述旅游资源及旅游规划与旅游开发的相关理论体系。全书从旅游资源的基本概念和分类出发,全面介绍了自然旅游景观资源、人文景观资源及其对旅游可持续发展的意义,在对旅游资源调查、评价的基础上,对旅游资源的开发、旅游规划的基础理论、内容、类别和编制方法以及旅游规划与开发的重点进行了系统阐述,从而构建旅游资源开发和规划的总体框架,有助于学者掌握旅游资源开发和规划的基本理论和基本方法,并将其较好地运用到实践中。

该书作为高校教材,另一个鲜明特色是注重理论与实际相结合。该书联系当前我国旅游规划与开发实践,列举了丰富的案例,方便读者理解旅游资源规划与开发的理论、方法、技术,从而激发读者的学习兴趣。该书全面介绍了室内旅游地图的判读、区域旅游景观资源的野外调查、风景名胜区旅游规划报告编写、相关旅游规划地图和模型制作以及专业实验室建设,具有很强的实用性和可操作性。

总之,该书注重理论与实际相结合,突出旅游开发的前瞻性和应用性,是一本富有特色和新意的旅游管理教材。该书既注重系统的理论学习,又注重实际工作的需要,具有实用性特点。读者无论是高等学校旅游管理类的师生,还是相关行业的从业人员,都有机会从中获益。其最大的价值在于,有理论追求但不唯理论,有实践经验但不囿于经验,较好地解决了旅游资源开发与规划中理论与实践、创新与运用的辩证关系,弥补了以往研究的不足,具有一定的建设性和前瞻性。

该书是作者多年从事旅游管理与规划教学和科学研究的结晶,其内容丰富、资料翔实、方法科学、论证充分,是探讨新时期旅游资源开发与规划的力作,为从更深的理论层面理解和把握旅游资源开发与规划提供了启迪,对指导旅游业健康、有序发展必将产生重要作用。目前该书已经广泛运用于全国高校旅游管理及相关专业的教学,是一本不可多得的好教材。

(上接第 31 页)

Research of Processing Contour Line after Automatically Vecteded

ZHANG Xiao - yuan¹, ZHAO Xiao - lin²

(1. Department of Remote Sensing and GIS Project, Sun Yat - Sen University, Guangzhou 510275;

2. Liaoning Geology Engineering Profession College, Dandong 118008, China)

Abstract: The contour line is the basic geographic information data that reflects the physiographic symbol and characteristics. After automatically vectored of contour line, there are still such problems as burrs and breakpoint connection, and this paper aims to give solutions for these problems. On the one hand, it proposes a neighborly search method to eliminate the burrs based on the burr growing point during the process of data preprocessing. On the other hand, it puts forward a composite connection method to connect the breakpoints after deeply analyses the causes and types of the disconnection of contour lines. These technological methods mentioned above have been proved in CAD environment by using VBA program, which shows that the composite connection method is not only practicable but also fast and accurate. Therefore, the methods proposed in this paper are both helpful to perfect and improve the data accuracy of contour line after the automatically vectored and beneficial to its research.

Key words: contour line; automatically vectored; burr; breakpoint connection; composite connection