

基于结构分析的曲线概括方法*

张青年¹, 廖克²

(1. 中山大学城市与资源规划系, 广东 广州 510275;

2. 中国科学院地理科学与资源研究所, 北京 100101)

摘 要: 线状要素往往是地图的主要构成要素, 因此在概括过程中常常需要进行对曲线进行概括处理. 目前已存在的曲线概括方法主要是一种几何上的逼近, 概括处理后难以保持其原有形态特点. 以曲线形态分析为基础, 提出基于复合弯曲分析的曲线概括方法.

关键词: 曲线概括; 结构单元; 复合弯曲

中图分类号: P283.7 **文献标识码:** A **文章编号:** 0529-6579 (2001) 05-0118-04

1 曲线概括的主要方法及存在问题

曲线形态概括是地图概括中研究最多的问题. 经过长期研究, 已经提出了许多自动概括算法. 其中较为的常见方法是 nth 算法、垂距算法、Douglas-Peucker 算法等. 近年来, 随着分形、小波等新方法引入到地图学领域, 又出现了基于分形和小波分析的曲线概括方法. 对于面的边界线的简化, 还可以采用数学形态学方法. 当然, 新的曲线概括方法仍在不断涌现. 不过, 尽管存在大量的曲线概括算法, 曲线概括问题仍没有得到圆满的解决.

以 nth 算法为例, 它纯粹是一种对曲线的点序列的机械的取舍, 即简单地每隔若干个保留一个点, 而不管这些点序形成的曲线段的形态如何, 因此概括后的曲线形态可能产生严重变形. 又如垂距算法和 Douglas-peucker 算法, 它们以垂距为准则对曲线上的顶点进行取舍, 能够在一定程度上保持曲线的形态. 还有人认为 Douglas-Peucker 算法是迄今最成功的曲线概括算法. 不过, 随着曲线形态的复杂化, Douglas-Peucker 算法简化效果变差. 而且该算法本身还存在曲线自交等难以解决的问题.

分形作为描述不规则曲线的工具, 被誉为刻画大自然的几何学. 许多制图曲线, 例如河流、城市边界等都具有一定程度的分形特征. 许多研究者探讨了曲线简化的分形方法. 其中较简单的情况是把制图曲线当作同质曲线处理^[1], 按 Douglas-Peucker 算法简化时曲线长度与阈值的对

应关系确定分维数, 因此仍需借助 Douglas-Peucker 算法. 较复杂的情况是把曲线当作非同质曲线处理, 其处理结果是 Douglas-Peucker 算法的一种近似, 只是比 Douglas-Peucker 算法额外地多选取一些点^[2].

按小波分析进行曲线简化则是另外一种思路, 通常是基于多分辨率分析的方法^[3], 见图 1. 小波分析方法用于曲线概括有利于保持曲线轴线延伸的趋势, 而把轴线两侧小的摆动作为噪声消除掉. 这种方法在消除小摆动(噪声)方面有可取之处, 但噪声剔除之后剩余的趋势轴线仍然有变形, 也就是离真正的趋势轴线还有一定差距.

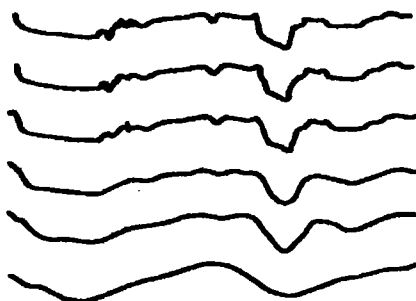


图 1 基于多分辨率分析的曲线概括

Fig.1 Line generalization based on multiresolution analysis

曲线自动概括并不是一个简单机械的顶点取舍过程. 现有的一些算法往往把曲线概括作为一个纯粹的几何问题处理, 即在较低分辨率上逼近原曲线, 自动处理程序中没有考虑曲线的认知特点.

显然, 根据曲线的形态特点进行概括处理, 将得到更符合曲线认知规律的结果. 人们已认识

* 收稿日期: 2000-11-30; 作者简介: 张青年 (1968-), 男, 讲师, 博士.

到曲线是一个具有几何形态的实体，而不仅仅是一个点的序列，因此在曲线概括过程中应该保持曲线的几何特点。有人认为，人对线的形状的判断基于两个标准：线的方向和线的基本弯曲度。因此，不同自然实体的曲线可被分成3组：相对光滑的、非常弯曲的、多方向的。也有专家假定一条线可以被分成同质的段，并提出了描述曲线段的许多参数，然后针对不同的曲线段的带宽、一致性和其它参数来选择特定的算法和阈值进行概括处理。总之，不同的曲线具有不同的形态特点，同一曲线的不同部分仍具有不同的形态特点。根据曲线的形态变化进行结构分析，以反映曲线各部分的形态特点，将能取得更好的概括效果。

2 基于复合弯曲的曲线结构分析

曲线的不同部分具有不同的形态特点，据此可以将曲线划分为一些具有不同形态特点的结构单元，再针对各个结构单元的具体特点采取适当的方法进行概括处理。人们最熟悉的曲线结构单元是各类易于被视觉识别的弯曲，例如圆形弯曲、菌形弯曲等。此外，自动概括研究中也提出了一些新的曲线结构单元划分方法，例如边界区域、曲线分支（branch）等^[4]。显然，新的曲线结构单元划分方法将继续涌现。相对于其它各种曲线结构单元而言，弯曲更易于被人们所理解和接受，按弯曲单元进行曲线划分是一种符合认知规律的方法^[5]。不过，自然曲线（例如海岸线）往往具有弯曲嵌套的特点，简单地按弯曲单元划分曲线将不能有效地反映其复杂的结构形态特点。因此，本文提出以复合弯曲作为曲线结构单元，将曲线数据组织为复合弯曲层次，在复合弯曲基础上进行曲线概括。采取的方法是首先识别基本弯曲，然后在基本弯曲的基础上识别复合弯曲。

2.1 基本弯曲的识别

弯曲是曲线的基本构成单元。也就是说，一条曲线是由一系列前后串接的弯曲组成的。对于数字化曲线而言，曲线是一个顶点的序列。因此，基本弯曲的识别实际上是将顶点的序列转换为基本弯曲的序列。我们根据曲线上各个顶点处的夹角来识别各个弯曲。显然，除曲线端点之外的任何一个顶点都有前驱线段和后继线段，后继线段方向与前驱线段方向的差即为该顶点的夹

角. 基本弯曲包含一个顶点与后继顶点的序列, 它总是按顺时针或逆时针方向绕转. 当曲线由左向右数字化时, 将逆时针的角度标志为正, 顺时针的角度标志为负. 那么, 在同一个基本弯曲内, 各个顶点处的夹角总是同为正 (顺时针) 或同为负 (逆时针), 而弯曲的两个端点的夹角符号与弯曲内部顶点的符号相反. 根据曲线顶点的夹角的符号正负值, 可以识别曲线上的各个基本弯曲.

2.2 复合弯曲的识别

自然界中的曲线包含着复杂的弯曲，某个基本弯曲可能是一个更复杂的弯曲的一部分。也就是说，由相互邻接的若干个基本弯曲可以组合成一个复合弯曲，见图 2。在图 2 中，复合弯曲 A 包含 a, b, c, d, e, f 等基本弯曲，复合弯曲 B 也包含若干个基本弯曲。类似地，若干个相邻的复合弯曲可以构成一个更高层次的复合弯曲，即复合弯曲具有层次性。复合弯曲的存在表现了弯曲的嵌套性质。识别复合弯曲之后，可以在更高层次上进行概括处理。

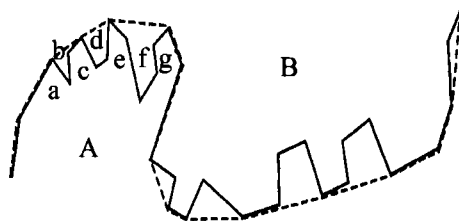


图 2 复合弯曲及其包含的基本弯曲

Fig.2 Compound bends and basic bends

识别复合弯曲的目的是用复合弯曲代替基本弯曲（或低层次的复合弯曲），从而使曲线形状得以简化。例如，图 2 中，用复合弯曲 A、B 代替其内嵌的一系列弯曲之后，曲线形状将有效简化。这里的关键问题是识别基本弯曲所在的复合弯曲。考察基本弯曲与复合弯曲的关系，将其分为三种情形，即基本弯曲在复合弯曲内侧、基本弯曲在复合弯曲外侧、基本弯曲跨越复合弯曲。

对于前两种情形（位于内侧或外侧），一般可以通过切除基本弯曲而得到其所在的复合弯曲，也就是用基本弯曲的底边代替基本弯曲。有时，基本弯曲还有一个底座，这时需要将底座也切除掉。某弯曲的底座是包含该弯曲的底边的一个基本弯曲，可以按基本弯曲识别方法来识别它。对于第三种情形（跨越），一般可以通过连

接相邻两个基本弯曲的外侧端点而得到其所在的复合弯曲,另一种方法则是将其转化为前两种情形进行处理。下一节将结合曲线概括说明其具体方法。

3 基于复合弯曲的曲线概括方法

如前所述,曲线是由相互嵌套的弯曲构成的。对于不同嵌套层次的弯曲,其视觉分辨率不同。一般说来,位于最低层次的基本弯曲最细微而不易辨别;高层次上的弯曲(例如曲线的趋势线)则相对较大而显著。当比例尺由大变小时,位于底层的基本弯曲往往由于过小而无法辨认,因此需要进行概括处理。本文采取的方法是用复合弯曲取代小于阈值的基本弯曲,从而实现曲线概括。这涉及弯曲的删除、合并、夸大等多种操作。

3.1 弯曲的面积阈值

弯曲的面积是决定其取舍的重要依据。当弯曲面积小于阈值时,通常要考虑删除该弯曲、或者夸大该弯曲、或者将该弯曲合并到其它弯曲。基本弯曲的大小由该弯曲及其底线围成的多边形的面积定义;复合弯曲的大小由其凸壳面积定义。

不过,相同面积的弯曲在视觉感受上并不总是相同的,因为不同形状的弯曲存在视觉效果上的差异。与圆形的弯曲相比,狭长的弯曲与扁平的弯曲在视觉上更突出,因此应该更易有机会被保留。同时,狭长的弯曲在比例尺减小时较少引起冲突。弯曲的形状可由紧凑度(cmp)描述。

$$\text{cmp} = \frac{\text{弯曲多边形的面积}}{\text{与弯曲多边形等周长的圆的面积}}$$

弯曲取舍以其大小为主要依据,但应该考虑形状的影响,因此应该使用大小与形状的综合指数。一种方法是用紧凑度的倒数来调整弯曲的面积,然后使用调整后的大小作为弯曲选取的主要标准。若以半圆(紧凑度为0.75)为标准形状,则公式为:调整面积 = 面积 × (0.75/cmp)。

在执行曲线简化之前,应该由用户指定最小弯曲的面积阈值,作为弯曲取舍的标准。实际上只需指定最小半圆形弯曲的直径。

3.2 弯曲的概括处理

概括操作的顺序是:寻找孤立的弯曲以便夸大;判断两个相邻弯曲是否相似以便合并;寻找局部最小弯曲作删除处理。这一过程重复执行,

直到所有的弯曲都大于面积阈值,这时就完成了曲线的概括。

(1) 删去操作。这是指删去小于阈值的局部最小弯曲并以其所在的复合弯曲的边线代替。分两种情况。一种情况是该弯曲无底座,这时只需由弯曲底边取代该弯曲,也就是删除该弯曲除两端点以外的其它顶点,弯曲沿底边被切除。另一种情况是该弯曲有底座,这时可能要同时切除该弯曲及其底座。在切除弯曲后判断其底座的面积是否小于阈值。若小于阈值,再判断它是否为局部最小弯曲。切除属于局部最小弯曲的底座。无底座和有底座的划分适用于基本弯曲位于复合弯曲内侧、基本弯曲位于复合弯曲外侧、基本弯曲跨越复合弯曲三种情形,见图3。图3右为两个基本弯曲跨越复合弯曲,删除一个基本弯曲后,其底座不是局部最小弯曲而暂被保留,基本弯曲跨越复合弯曲的关系随之消失。

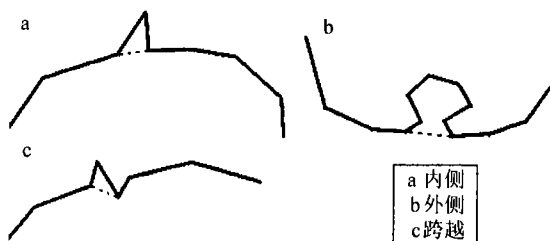


图3 删去小弯曲

Fig.3 Bend excision

切除操作是按迭代方式进行的,通过若干次循环来切除所有小于阈值的弯曲。每一次循环中切除局部最小弯曲及其底座(若底座为局部最小弯曲),使曲线由基本弯曲简化到复合弯曲的层次。删去操作迭代进行,曲线概括到新的复合弯曲层次,直到曲线上的每一个弯曲都大于给定的面积阈值。

(2) 合并操作。对于相邻出现的若干个小弯曲,通常需要将其合并为一个或多个大弯曲,或者夸大保留其中的某几个弯曲。本文采取合并小弯曲的方法。例如,对于图4中紧邻的两个小弯曲A与B,应合并为一个大弯曲,否则它们都将被切除。为避免弯曲被删除,应该将其合并。方法是切除两弯曲之间的弯曲C及其底座,则合并产生一个大弯曲D,见图4。

(3) 夸大操作。曲线上的一些特征弯曲常常需要保留或强调,本文将孤立弯曲作为重要的特征弯曲。孤立弯曲是指弯曲两侧的延长线远大



图4 相似弯曲的合并

Fig.4 Combination of similar bends

于弯曲的底线长度的小弯曲。当孤立弯曲小于阈值时,进行适当程度的夸大(例如夸大1.2倍),以增加其被选取的机会。可以使用高斯分布作夸大处理,计算弯曲各顶点的新坐标值。弯曲底线的中点作为移位中心,外围的点的移位距离由中心到边沿按高斯分布逐渐衰减。按高斯分布夸大后,弯曲存在一定的变形,扁平的弯曲将变高,而狭窄的弯曲将变宽。

3.3 若干实例

按上述方法编制了曲线自动概括程序,进行了曲线概括的一些试验,并与 Douglas-Peucker 方法比较,自动概括的部分结果如图5所示。它反映了三条曲线简化的结果。左面一条曲线是大弯曲套小弯曲,中间一条曲线是小弯曲围绕大弯曲轴线波动,右面一条曲线则是复杂的弯曲嵌套。概括结果表示为上下两行,上行为基于复合弯曲的概括结果,下行为 Douglas-Peucker 方法的概括结果。对于这三种情形,上述基于复合弯曲的概括方法都得到了比较满意的结果,概括后的曲线保持了其曲折波动的基本走势。相比之下,D-P方法的概括结果则存在较大的变形。

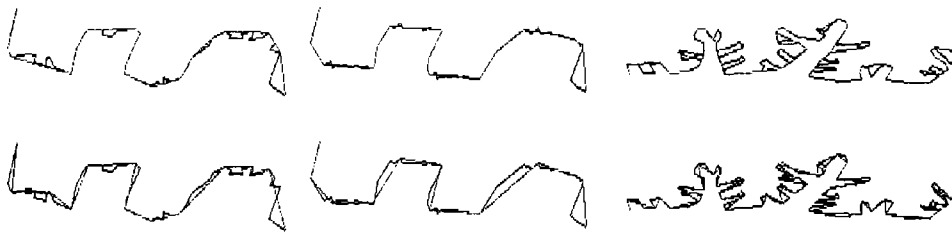


图5 曲线概括示例

Fig.5 Examples for line generalization

参考文献:

- [1] 王桥, 吴纪桃. 一种新分维估值方法作为工具的自动制图综合[J]. 测绘学报, 1996, 25(1): 10-16.
- [2] NORMANT F, WALLE A V D. The sausage of local convex hull of a curve and the Douglas-Peucker algorithm[J]. Cartographica, 1996, 33(4): 25-34.
- [3] 吴纪桃, 王桥. 小波分析在 GIS 线状要素图形简化中的应用研究[J]. 测绘学报, 2000, 29(1): 71-75.
- [4] POORTEN P, VAN der JONES, C B. Customisable Line Generalisation using Delaunay Triangulation[C]. In: Proceedings of the 19th ICA, Ottawa, 1999.
- [5] WANG Z, MULLER J C. Line Generalisation based on analysis of shape characteristics[J]. Cartography and Geographic Information Systems, 1998, 25(1): 3-15.

Line Generalization Based on Structure Analysis

ZHANG Qing-nian¹, LIAO Ke²

(1. Department of Urban and Resource Planning, Zhongshan University, Guangzhou 510275, China;

2. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

Abstract: Line generalization is an important issue receiving extensive attention in the field of map generalization. In the past 30 years, numbers of algorithms were designed for line generalization. Most of those algorithms are based on distance, area and other parameters, which result in generalized lines approximating the original lines geometrically. The generalized lines usually lost their original shape in some way. In order to maintain the shape of original lines, lines can be divided into structural units. Among all possible structural units, compound bend is a reasonable one for line division. Based on compound bend analysis, methods for line generalization are designed, including bend deletion, combination and exaggeration. These three generalization operators are iteratively processed so that small basic bends are replaced by compound bends larger than threshold. A case study is carried out to test the method, and some examples show that satisfactory results can be obtained by this method.

Keywords: line generalization; structure unit; compound bend