

矢量曲线数据的网络渐进传输

操震洲^{1,2} 李满春¹ 程亮¹ 陈振杰¹

(1 南京大学地理与海洋科学学院,南京市汉口路 22 号,210093)

(2 南京工业大学测绘学院,南京市中山北路 200 号,210009)

摘要:提出了一种保持拓扑一致性的曲线多尺度表达方法,并基于此方法实现了曲线数据的网络渐进传输。该方法首先将曲线结点的重要程度、公共约束点和单调链等信息组织成单调线性 BLG 树结构,然后基于该结构按需实时生成多尺度曲线,利用优化的单调链求交算法维护拓扑一致性。曲线数据网络渐进传输实验验证了此方法的有效性。

关键词:渐进传输;单调线性 BLG 树;多尺度曲线;线化简;拓扑一致性

中图分类号:P208

矢量数据网络渐进传输概念由 Bertolotto 于 2001 年提出^[1]。渐进传输是指在传输前先将矢量数据分解为多个不同分辨率数据,服务器先传低分辨率数据,再传高分辨率数据,用户在接收到低分辨率数据后即可对数据进行浏览、查询等操作,从而减少等待数据的时间。一个理想的矢量数据渐进传输系统应具有动态性、实时性和拓扑一致性。动态性指服务器端对矢量数据的多尺度分解是根据当时的网速、用户设备、数据量等因素动态规划的;实时性指服务器从接收用户数据请求到生成多尺度数据的过程是即时完成的;拓扑一致性指要素间相离、相邻、相交等空间位置关系保持不变。

1 多尺度曲线模型分析

矢量渐进传输技术的难点是矢量数据的多尺度表达,而后的关键则是多尺度曲线模型,目前主要有基于线化简方法和基于小波算法的两类。小波算法将曲线视同信号波,利用频谱分析和滤波方式来化简曲线^[2],该方法的缺点是时间代价很高,且图形拓扑一致性较难维护。通过线化简算法生成多尺度曲线是目前的主要方式,Douglas-Peucker 算法^[3](简称 DP 算法)是所有线化简算法中公认最优秀的,其原理是递归通过在当前

基线上寻找最远点的方式来选择特征点,但其时间复杂度为 $O(n^2)$ 。为提高时间性能,Oosterom 提出用 BLG 树结构来存贮 DP 算法的搜索结果,通过遍历 BLG 树可得到任意尺度的曲线^[4]。DP 算法存在着父结点偏离量小于子结点的异常情况^[5],艾波等提出将子结点偏离量修改为父结点偏离量的方法来解决,并采用线性表结构组织 BLG 树^[6]。

线化简过程会产生拓扑异常,表现为曲线自身或与周围曲线的空间位置关系变化。如图 1 所示,本文把化简中出现的拓扑异常分为三种:第一种是因为公共点的删除导致曲线间已有的相邻、相交位置关系的改变;第二种是因为非公共点的不合理删除导致曲线间已有的相离位置关系的改变;第三种是因为不合理的结点删除导致曲线出现自相交。目前,对拓扑一致性的维护方法有两类:一是以文献^[7]为代表的预先将曲线分割为单调的且其凸包互不重叠的子曲线,然后在各子曲线内化简以避免异常发生;二是以文献^[8-9]为代表的在综合中实时进行拓扑异常判断与消除。前者时间开销较小,但由于引入约束点会导致图形冗余,后者能保证图形的精简,但时间开销较大。

Bertolotto 从制图综合角度提出了包含 7 种综合算子的概念框架,但鉴于自动综合的实际难度,该概念框架难以实现。文献^[10]提出了面向

收稿日期:2013-03-15。

项目来源:国家自然科学基金资助项目(41001238);国家科技支撑计划资助项目(2012BAH28B02)。

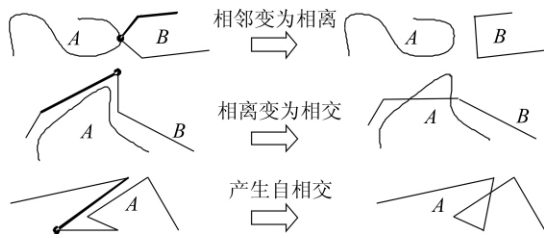


图1 拓扑异常的类型

Fig. 1 Types of Topological Exception

多边形渐进传输的空间数据变化累积模型,但未解决多边形之间的拓扑一致性问题。文献[6]提出了河网多尺度数据结构用于网络渐进传输,同样未考虑河流化简中的拓扑一致性问题。文献[11-12]基于VW算法^[13]与拓扑规则提出了“点删除”的在线化简方法,并在此基础上实现了矢量地图的网络渐进传输。该方法能有效解决多尺度矢量数据生成与拓扑一致性问题,但时间效率不是最优。本文从时间优化和保持拓扑一致性角度提出了一种基于单调线性BLG树结构的曲线多尺度表达方法,该方法预先计算曲线的单调链、约束点及结点偏离量等信息并将其组织成单调线性BLG树结构,通过该结构实现对要素的快速选取、概括和拓扑一致性维护,其化简和拓扑维护的时间更短。

2 基于单调线性BLG树的曲线多尺度表达

2.1 单调线性BLG树

本文针对的是无拓扑模型的矢量曲线数据,预先分析曲线并建立单调线性BLG树结构。对于BLG树存在的父结点偏离量小于子结点异常,本文约定将父结点偏离量修改为子结点的值。分析拓扑异常类型可知,第一类拓扑异常是由于删除公共结点而引起的,通过修改公共结点的偏离量能避免其发生。本文把曲线首尾端点、多要素的公共结点统称为约束点。图2列出了约束点的三种存在形式:第一种约束点为曲线端点,曲线A大小用其最小外接矩形表示,设当前分辨率为R(R为长度指标),以 $\max(W_A, H_A) \geq R$ 作为曲线选取条件。因此,在单调线性BLG树结构中,设置曲线A的端点(即第一种约束点)的偏离量为 $\max(W_A, H_A)$,这样可根据它的偏离量决定该曲线的取舍。第二种约束点为多线共有的单交点或单邻接点,对曲线A而言, P_i 点仅当曲线B在当前尺度下不显示时,才可以删除,否则不能删。设

P_i 点在曲线A中的偏离量为 ΔP_{iA} ,规定 P_i 点在曲线A的单调线性BLG树中的偏离量值为 $\max(\Delta P_{iA}, W_B, H_B)$ 。第三种约束点为多线共有的连续多点(即共边)。对于公共边的首、尾结点如 P_j 和 P_{j+5} ,按第二种约束点来处理;而对于中间点,需要保证它们在曲线A、B的单调线性BLG树中拥有相同的偏离量,否则,在概括时公共边界处有可能产生裂缝。规定中间点取其在两相邻曲线上偏离量的最大值,如 P_{j+1} 点的偏离量为 $\max(\Delta P_{j+1A}, \Delta P_{j+1B})$ 。

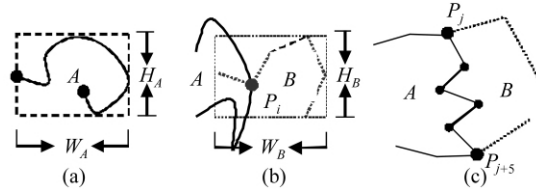


图2 约束点的类型

Fig. 2 Types of Constraint Points

第二、三类拓扑异常表现为线相交和自相交,文献[14]提出的扫描线算法是判断平面线段集相交的经典算法,文献[15]在扫描线算法的基础上,根据曲线上直线段首尾相连的特点,提出了基于单调链的求交算法,时间得以优化。如果对单调链构建最小凸包,当两凸包不相交时,其对应的单调链也不会相交。本文据此将单调链分为两类:独立单调链(其凸包与所有其他单调链凸包不相交)和重叠单调链(其凸包与某单调链凸包有相交)。若预先将单调链划分为重叠单调链和独立单调链,然后基于重叠单调链进行求交计算,可提高算法速度。如图3所示,以X轴为单调链方向,曲线A沿X轴方向可分为1、2两个单调链,曲线B为一个单调链3。其中,链1由 $P_1 P_2 P_3 P_4 P_5 P_6$ 组成,为独立单调链;链2由 $P_6 P_7 P_8 P_9$ 组成;链3由 $P_{10} P_7 P_{11} P_{12} P_{13}$ 组成,链2、3都是重叠单调链。

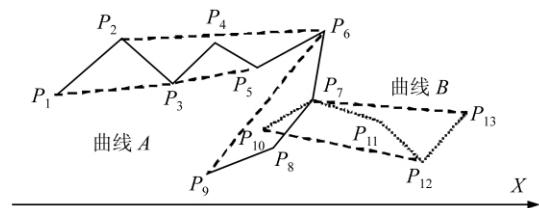


图3 曲线分解为单调链

Fig. 3 Curve is Decomposed into Monotone Chains

把两个相邻单调链的连接点称为单调拐点,如图3中的 P_6 点。删除单调拐点外的点不会改变单调链类型,而删除单调拐点可能改变单调链类型。当单调拐点连接的都是重叠单调链时,此

时应维持该单调拐点左、右两单调链的类型不变; 当单调拐点连接有独立单调链时, 本文采用的方案是: 以独立单调链上结点的最大偏离量值来重新设置该单调拐点的偏离量值, 这样可保证单调拐点与独立单调链一同出现或消亡, 确保在化简中独立单调链的类型不变。

单调线性 BLG 树的构造方法为: ① 逐要素执行 DP 算法, 建立单调线性 BLG 树, 按结点连接顺序存贮, 按 DP 算法结果初始化各结点的偏离量; ② 若有父子结点偏离量异常, 将父结点偏离量修改为子结点的偏离量; ③ 将各要素按 X

方向分解为单调链, 并构建最小凸包, 划分为重叠单调链和独立单调链; ④ 将单调链信息记录到单调线性 BLG 树, 修改独立单调链左、右端点的偏离量值; ⑤ 获取各要素约束点及要素最小外接矩形, 更新单调线性 BLG 树中约束点的偏离量值。

图 3 中曲线 A 所对应的单调线性 BLG 树的生成过程如图 4 所示。其中, 重叠单调链用非零值的整数表示, 正负号表示单调链的增减性, 数字为单调链编号; 而独立单调链记为 0; 每条单调链从其左端点开始标记单调信息, 右端点一般不标记(仅当曲线终点时标记)。

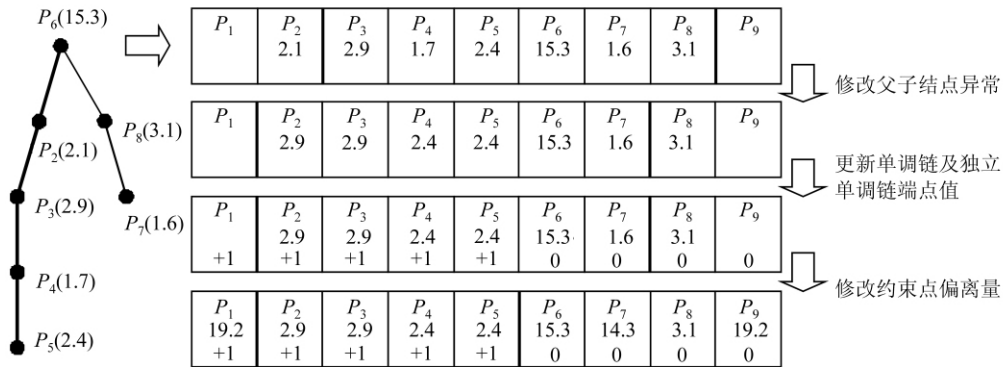


图 4 单调线性 BLG 树

Fig. 4 Monotonous Linear BLG Tree

此单调线性 BLG 树结构的特点有 4 个: ① 以结点偏离量为依据, 支持对要素的选取和概括操作, 其中要素选取以曲线首结点的偏离量为依据; ② 通过修改约束点的偏离量, 避免了化简中第一类拓扑异常的发生; ③ 记录了单调链及其类型信息, 支持快速求交计算, 为判定化简中的第二、三类拓扑异常服务; ④ 处理了 BLG 树中父子结点偏离量异常情况。

2.2 保持拓扑一致性的多尺度曲线生成方法

在建立起所有要素的单调线性 BLG 树后, 根据要素首结点偏离量的大小对所有单调线性 BLG 树按从大到小排序, 建立纵向索引, 让尺寸大的曲线居前, 而尺寸小的居后。若当前选取分辨率、概括分辨率分别为 R_1 、 R_2 ($R_1 > R_2$), 则该尺度下矢量曲线数据的生成方法为: ① 根据 R_1 值, 查询纵向索引, 确定选取要素集的所有单调线性 BLG 树; ② 遍历选择集的单调线性 BLG 树, 选取偏离量 $\geq R_2$ 的所有结点, 同时检索出化简后的重叠单调链, 此过程时间复杂度为 $O(n)$ (n 为结点数); ③ 对上一步得到的所有重叠单调链, 执行基于重叠单调链的求交算法, 搜索相交线段; ④ 对所有相交线段执行拓扑一致性维护工作; ⑤ 输出最终的多尺度曲线数据。

步骤④工作是消除步骤③所发现相交线段, 可采用回溯法实现, 即在相交线段对应的单调线性 BLG 树上查找并添加该线段间偏离量最大的结点来协调, 直到不相交为止。

3 曲线数据的渐进传输实验

基于 VS 2005 平台, 本文开发了矢量曲线数据渐进传输实验系统。服务器端程序预先分析曲线数据并建立线性单调 BLG 树结构, 在接收客户端数据请求后, 根据客户端设备、网速及数据量, 动态规划出层数及各层分辨率, 从单调线性 BLG 树中快速检索出轮廓数据及各层增量数据, 然后渐进传输。客户端程序功能有: 发送数据请求, 完成数据接收, 重构大尺度数据并显示。图 5 为渐进传输模式下客户端的数据视图, 各层数据保持了拓扑一致性, 随着数据传输的继续, 客户端的曲线图形越来越详细。

为测试重叠单调链求交算法在拓扑维护方面的时间优势, 将其与另外两种方法进行了比较, 方法一是以线段两两判断法, 方法二是单调链求交算法。采用普通台式机(主频 1.5 GHz, 内存 1.2 G)进行了实验, 实验数据为河网数据, 共 3 157 个

要素。建立单调线性 BLG 树结构后,单调链、重叠单调链、直线段数目分别为 3 945、2 372、

67 016,三种方法比较结果见表 1。

拓扑维护包括线段相交、自相交异常发现与

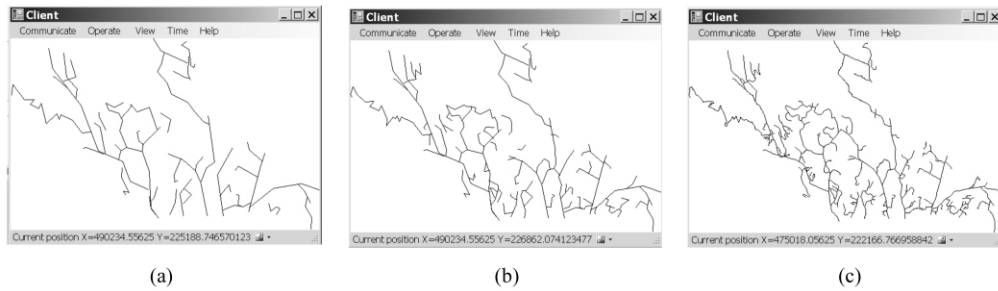


图 5 渐进传输模式下的客户端数据视图

Fig. 5 Data View of Client Under Progressive Transmission Mode

表 1 三种拓扑维护方法的时间比较

Tab. 1 Time Comparison of Three Methods for Maintaining Topological Relationship

方法	项目			时间/ ms
	直线段 (67 016 个)	单调链 (3 945 个)	重叠单调链 (2 372 个)	
线段两两判断法	✓			325
单调链法		✓		87
重叠单调链法			✓	53

异常消除两个阶段。异常消除阶段工作包括:首先,在相交线段上添加该线段间偏离量最大的结点;然后,在添加结点后线段之间再判断是否相交,直到不相交为止。实验发现,在绝大部分情况下,相交线段数只占总线段数的小部分,而多次反复相交线段数更少,拓扑维护的时间主要由异常发现阶段决定。

由表 1 可知,相对于线段两两判断法,采用单调链求交算法能大幅度缩短拓扑维护时间,而采用重叠单调链求交算法,则能将拓扑维护时间进一步缩短 40%左右。

4 结 语

本文将曲线化简中的拓扑异常分为三类,通过修改约束点偏离量避免了第一类异常,利用优化的单调链求交算法快速处理第二、三类异常,比较系统地考虑了曲线拓扑一致性问题。通过预先将曲线划分为独立单调链和重叠单调链,然后以重叠单调链进行求交计算,可优化单调链求交算法的时间。本文提出的单调线性 BLG 树结构既支持对要素的选取和概括操作,又支持对要素的拓扑一致性维护。其中,选取和概括通过结点偏离量进行,拓扑维护通过约束点和单调链进行。但对要素选取本文仅简单地以尺寸作为依据,未考虑要素的地理含义,未来将进一步进行研究。

参 考 文 献

- [1] Bertolotto M, Egenhofer M J. Progressive Transmission of Vector Map Data over the World Wide Web[J]. *Geoinformatic*, 2001, 5(4):345-373
- [2] 吴纪桃,王桥. 小波分析在 GIS 线状数据图形简化中的应用研究[J]. *测绘学报*, 2000, 29(1):71-75
- [3] Douglas D H, Peucker T K. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Character[J]. *The Canadian Cartographer*, 1973, 10(2):112-123
- [4] van Oosterom P. The Reactive-tree: a Storage Structure for a Seamless, Scaleless Geographic Database[C]. *The Auto-Carto*, Baltimore, U S, 1991
- [5] 毋河海. 基于多叉树结构的曲线综合算法[J]. *武汉大学学报·信息科学版*, 2004, 29(6):479-483
- [6] 艾波,艾廷华,唐新明. 矢量河网数据的渐进式传输[J]. *武汉大学学报·信息科学版*, 2010, 35(1):51-54
- [7] Mantler A, Snoeyink K. Safe Sets for Line Simplification[C]. *The Tenth Annual Fall Workshop on Computational Geometry*, New York, 2000
- [8] Alan S. Topologically Consistent Line Simplification with the Douglas-Peucker Algorithm[J]. *Cartography and Geographic Information Science*, 1999, 26(1):7-18
- [9] da Silva A C G, Wu S T. A Robust Strategy for Handling Linear Features in Topologically Consistent Polyline Simplification [J]. *Geoinformatics*, 2006(1):19-34
- [10] 艾廷华,李志林,刘耀林,等. 面向流媒体传输的空间数据变化累积模型[J]. *测绘学报*, 2009, 38(6):514-526
- [11] Yang Bisheng, Purves R, Weibel R. Efficient Transmission of Vector Data over the Internet[J]. *International Journal of Geographical Information Science*, 2007, 21(2):215-237

- [12] 杨必胜, 李清泉. World Wide Web(WWW)上矢量地图数据的多分辨率传输算法[J]. 测绘学报, 2005, 34(4):355-360
- [13] Visvalingham M, Whyatt D. Line Generalization by Repeated Elimination of Points [J]. The Cartographic Journal, 1993, 30(1):46-51
- [14] Bentley J L, Ottmann T A. Algorithms for Reporting and Counting Geometric Intersections [J]. IEEE Transactions on Computers, 1979, 28(9):643-647
- [15] Park S C, Shin H. Polygonal Chain Intersection [J]. Computers & Graphics, 2002, 26(2):341-350

第一作者简介: 操震洲, 讲师, 博士生, 从事多尺度空间数据的组织与表达、地图综合等方面的研究。

E-mail: czz@njut.edu.cn

Progressive Transmission of Vector Curve Data over Internet

CAO Zhenzhou^{1,2} LI Manchun¹ CHENG Liang¹ CHEN Zhenjie¹

(1 College of Geographic and Oceanographic Science, Nanjing University,
22 Hankou Road, Nanjing 210093, China)

(2 College of Geomatics Engineering, Nanjing University of Technology,
200 Zhongshan Road, Nanjing 210009, China)

Abstract: We propose a method for multi-scale representation of curve which keeps topological relationship unchanged and elaborates the process of progressive transmission of curve data over the Internet based on this method. Firstly, the importance degree of nodes and the information about public constrain points and monotone chains are stored in a monotonous linear BLG tree structure. Then the multi-scale curve is generated in real time based on this structure and the topological relationship is maintained by an optimized monotone chains intersection algorithm. Finally, the method was used in the experiment of progressive transmission of curve data over the Internet and verified its effectiveness.

Key words: progressive transmission; monotonous linear BLG tree; multi-scale curve; line simplification; topological consistency

About the first author: CAO Zhenzhou, lecturer, Ph.D candidate, majors in organization and representation of multi-scale spatial data, and map generalization.

E-mail: czz@njut.edu.cn