

# 适用于网络渐进传输的多分辨率曲线生成算法

操震洲<sup>1,2\*</sup>, 李满春<sup>1</sup>, 程亮<sup>1</sup>, 陈振杰<sup>1</sup>

(1. 南京大学 地理与海洋科学学院, 南京 210093; 2. 南京工业大学 测绘学院, 南京 210009)

(\* 通信作者电子邮箱 czz@njut.edu.cn)

**摘要:** 针对渐进传输系统在多分辨率矢量数据生成过程中存在的计算费时、拓扑不一致问题, 提出一种适用于网络渐进传输的多分辨率曲线生成算法。该算法通过预先存储的节点偏离量化简曲线, 利用优化的单调链求交算法维护曲线拓扑一致性, 从而支持多分辨率曲线的快速生成和拓扑一致性维护。基于该算法开发了曲线数据渐进传输实验系统, 实验结果表明, 多分辨率曲线数据保持了拓扑一致性, 且其生成时间与数据量大小呈近线性的关系, 证明了算法的有效性。

**关键词:** 渐进传输; 地理信息系统; 多分辨率曲线; 曲线化简; 拓扑一致性

**中图分类号:** TP301.6 **文献标志码:** A

## Algorithm of generating multi-resolution curves for progressive transmission over the Internet

CAO Zhenzhou<sup>1,2\*</sup>, LI Manchun<sup>1</sup>, CHENG Liang<sup>1</sup>, CHEN Zhenjie<sup>1</sup>

(1. College of Geographic and Oceanographic Science, Nanjing University, Nanjing Jiangsu 210093, China;

2. College of Geomatics Engineering, Nanjing University of Technology, Nanjing Jiangsu 210009, China)

**Abstract:** Concerning the problems of high time complexity and topological inconsistency existing in the multi-resolution representation of curve for progressive transmission, an algorithm of generating multi-resolution curves for progressive transmission over the Internet was proposed in this paper. By using pre-stored vertex deviation to simplify curves and using an optimized monotone chain intersection algorithm to maintain topological consistency, the algorithm can quickly generate topologically consistent multi-resolution curves. The algorithm was used in the experiment of progressive transmission for curve data, and the experimental results show that the multi-resolution curve data maintain topological consistency and the generation time changes linearly with the amount of data. The effectiveness of the algorithm has been verified in the experiment.

**Key words:** progressive transmission; Geographical Information System (GIS); multi-resolution curve; line simplification; topological consistency

## 0 引言

随着空间信息服务的发展, 空间数据在网络上的应用日益普及, 而海量的空间数据和有限的网络带宽常常导致数据传输时间过长, 成为制约空间信息服务发展的一个技术瓶颈。空间数据网络渐进传输根据数据量大小、网络传输速度、客户端显示设备等因素将待传输数据分解为多个层次的不同分辨率的数据视图, 每一层数据可以和上层动态重构, 形成更详细的数据轮廓。服务器先传粗轮廓数据, 再传细轮廓数据, 逐层进行直到完成。用户在收到第一层数据后, 就可以对数据进行浏览、查询等操作。在网络繁忙时, 渐进传输可显著减少用户等待数据的时间。目前, 地形数据、影像数据网络渐进传输已取得较丰富的研究成果, 应用也很深入。矢量数据网络渐进传输概念最早由 Bertolotto 等<sup>[1]</sup>于 2001 年提出, 国内外学者对其做了大量研究<sup>[2-6]</sup>, 但由于矢量数据结构的复杂性, 矢量数据网络渐进传输技术仍不够成熟。该系统原理如图 1 所示。

矢量数据网络渐进传输系统的技术难点有两个: 一是如何快速生成多个不同分辨率数据; 二是如何保证各分辨率数据中的空间要素拓扑关系(即相离、相交、相邻等空间位置关

系)一致性。Bertolotto 从制图综合角度提出矢量地图多分辨率表达的概念框架, 但鉴于自动综合的实际难度, 该概念框架难以具体实现。通过线化简方法可实现从原始矢量地图到较低分辨率矢量地图的变换, 但是线化简法也存在着两个不足: 一是图形精度高的化简算法时间复杂度都较高(如 Douglas-Peucker 算法<sup>[7]</sup>、VW 算法<sup>[8]</sup>时间复杂度都为  $O(n^2)$  ( $n$  为节点数), 不适合大数据量的实时化简; 二是化简中会产生拓扑异常。文献 [2-3] 基于 VW 算法与拓扑规则, 提出“点删除”的在线化简方式, 能有效解决多尺度矢量数据生成与拓扑一致性问题, 但该方法化简和拓扑维护的时间开销都较大。文献 [4-5] 针对多边形数据、河网数据的渐进传输, 分别提出基于变化累计模型和线综合二叉(Binary Line Generalization,

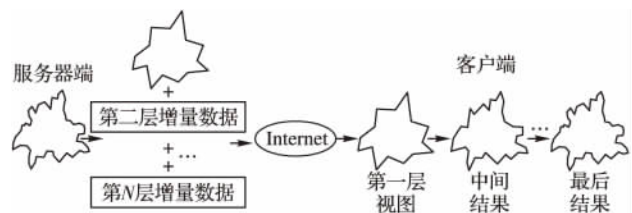


图 1 矢量数据的网络渐进传输

收稿日期: 2012-09-26; 修回日期: 2012-10-31。

作者简介: 操震洲(1977-), 男, 安徽怀宁人, 讲师, 博士研究生, 主要研究方向: 多尺度空间数据的组织与表达; 李满春(1964-), 男, 江苏常州人, 教授, 博士, 主要研究方向: 地理信息系统、土地利用/覆被变化; 程亮(1978-), 男, 江苏南京人, 副教授, 博士, 主要研究方向: 激光雷达数据处理; 陈振杰(1974-), 男, 江苏南京人, 副教授, 博士, 主要研究方向: 地理信息系统。

BLG) 树结构的多尺度表达方法,但都未考虑多边形、河流数据的拓扑一致性问题。文献 [8-10] 提出化简后检测和消除拓扑异常的拓扑维护策略,但其检测拓扑异常的时间成本较高。拓扑异常主要表现为出现不合理的相交和自相交现象, Bentley 等<sup>[12]</sup> 提出的扫描线算法是判断直线段相交的经典算法。Park 等<sup>[13]</sup> 针对曲线上直线段首尾相连特点,提出了单调链求交算法,其时间效率较扫描线算法更高。若基于单调链来处理拓扑异常,可缩短拓扑一致性维护的时间。

本文提出一种适用于网络渐进传输的多分辨率曲线生成算法,通过事先计算并存储曲线的节点偏离量、单调链、权重等信息,该算法能在更短的时间内生成保持拓扑一致性的多分辨率曲线。其时间优势体现在两个地方:一是根据事先存储的节点偏离量筛选节点,将 Douglas-Peucker 算法<sup>[6]</sup> 的时间复杂度从  $O(n^2)$  降为  $O(n)$ ;二是在执行单调链求交算法搜索相交线段前,通过排除不相交的单调链,缩短了拓扑异常的处理时间。

### 1 多分辨率曲线生成算法

曲线可用曲线上若干个特征点表示,本文的曲线具体指无自相交的一系列首、尾相连的直线段集合。曲线节点  $k$  用  $\overline{V}_k$  表示,  $\overline{V}_k$ 、曲线  $L$ 、化简后曲线  $L'$  可以被描述为:

$$\begin{cases} L = \{ \overline{V}_0 \overline{V}_1, \overline{V}_1 \overline{V}_2, \dots, \overline{V}_{n-1} \overline{V}_n \} \\ L' = \{ \overline{V}_0 \overline{V}_1, \dots, \overline{V}_{k-1} \overline{V}_{k+1}, \dots, \overline{V}_{n-1} \overline{V}_n \} \\ L' = L - \{ \overline{V}_k \} \\ \overline{V}_k = \{ x_k, y_k \} \end{cases} \quad (1)$$

算法框架如图 2 所示。由于计算曲线的节点偏离量、单调链、权重等耗时较大的工作都在预处理阶段完成,在化简阶段,可根据这些信息对曲线集快速进行选取、概括和拓扑一致性维护。

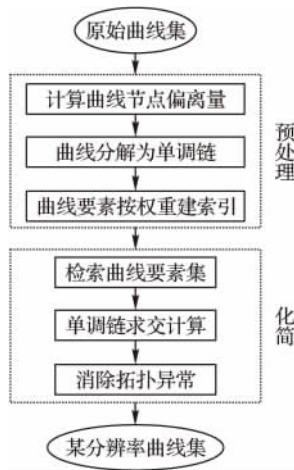
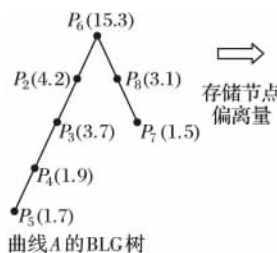


图 2 算法框架



$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$	$P_8$	$P_9$
0	4.2	3.7	1.9	1.7	15.3	1.5	3.1	0
$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$	$P_8$	$P_9$
15.3	4.2	3.7	1.9	1.7	15.3	15.3	15.3	15.3

图 4 节点偏离量与单调链的设置过程

#### 1.1 计算曲线节点偏离量与单调性

Douglas-Peucker 算法的搜索结果为 BLG 树,偏离量大的节点通常位于树的上层,节点偏离量大小体现了节点的重要程度。首先执行 Douglas-Peucker 算法以获取曲线各节点偏离量并以线性表结构存储,然后根据线性表里节点偏离量大小来选取节点,从而实现曲线化简,这样可将 Douglas-Peucker 算法的时间复杂度从  $O(n^2)$  降为  $O(n)$ 。

化简中出现的拓扑异常是因为不合理的节点删除,对多要素公共节点的删除将改变要素间的邻接关系,对非公共节点的删除也可能导致曲线出现相交和自相交。称曲线首尾端点、多要素的公共节点为约束点,将其分为两类:第一类包括曲线首尾端点、单邻接点以及邻接边的首尾端点;第二类是邻接边上除首尾端点外的其他公共点。拓扑一致性维护方法为:对于第一类约束点,规定不可删除,设置其偏离量为曲线上所有节点偏离量的最大值;对于第二类约束点,由于它在每个邻接要素里都有一个偏离量,每个约束点就有若干个不同的偏离量,规定其偏离量取这若干个值的最大值;对于非约束点,在删除后通过优化的单调链求交算法来搜索相交线段。对单调链构建最小凸包,根据凸包间是否相交将单调链分为两类: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)。通过重叠单调链进行求交计算,可缩短拓扑异常的处理时间。

图 3 为两条邻接关系的曲线,粗、细实线分别为曲线 A、B,虚线为单调链的凸包边界。以 X 轴为单调链方向,曲线 A 沿 X 轴方向可分为 1 和 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  $P_{10} P_8 P_7 P_{11} P_{12} P_{13}$  组成,链 2 和链 3 都是重叠单调链。为确保独立单调链在化简中的类型不变,约定以独立单调链上节点的最大偏离量值来重新设置该独立单调链左、右端点的偏离量值,这样可保证独立单调链左、右端点与独立单调链一同出现或消亡。

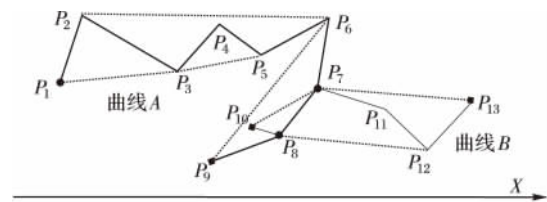


图 3 曲线分解为单调链

节点偏离量与单调性的设置方法为:执行 Douglas-Peucker 算法计算并记录各节点偏离量;将曲线按 X 方向分解为重叠单调链和独立单调链,记录单调链类型;修改约束点偏离量;修改独立单调链左、右端点偏离量。图 4 为图 3 中曲线 A 的节点偏离量与单调性的生成过程。重叠单调增、减链分

别用 1 和 -1 表示; 独立单调链记为 0; 单调链类型标记于单调链左端点。

### 1.2 要素按权重建立索引

当曲线集在某分辨率下显示时,对于尺寸小于该分辨率的曲线以及在该分辨率下重要性较小的曲线则不需要显示。将要素尺寸和重要性综合量化为权重指标,以该指标作为曲线选取标准,并基于曲线权重建立索引。如图 5 所示,要素权重初始值为其最小外接矩形长、宽值的较大者,然后根据要素类型及其重要性对其做适当调整,在调整后的权重值上建立索引。

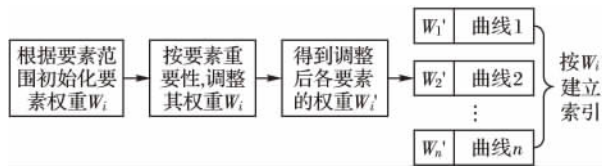


图 5 曲线要素索引的建立方法

### 1.3 多分辨率曲线的生成

多分辨率矢量曲线的生成方法为(记当前选取与概括分辨率分别为  $R1, R2$ ):

- 1) 根据  $R1$  检索要素,选取满足  $W_i' \geq R1$  的所有曲线要素集。
- 2) 遍历选择集内所有曲线,选取偏离量  $\geq R2$  的所有节点,同时得到化简后的重叠单调链,此过程时间复杂度为  $O(n)$  ( $n$  为节点数)。
- 3) 对上一步得到的所有重叠单调链,执行单调链求交算法,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,可采用回溯法实现,即在两相交线段对应的原曲线段上查找并添加该线段间偏离量最大的节点来协调,直到不相交为止。
- 5) 输出当前分辨率下保持拓扑一致性的曲线数据集。

单调链求交算法原理是:沿扫描线方向从左向右扫描,判断与当前扫描线相交的重叠单调链在  $Y$  方向的次序关系,若两单调链相交,则在交点  $I$  前、后处,它们在扫描线方向上的次序关系会发生改变。在算法进行前,根据各重叠单调链左端点的  $X$  值大小建立单调链在  $X$  方向的次序关系,此过程采用堆排序,其时间复杂度为  $O(p \log p)$  ( $p$  为重叠单调链个数)。采用堆结构存储当前扫描线上的单调链次序关系,从左向右扫描,待判断节点和交点的总数目为  $q + k$  ( $q$  为所有重叠单调链上的节点数,  $k$  为交点数),每次堆查找的时间复杂度为  $O(\log p)$ ,故本求交算法的时间复杂度为  $O((q + k) \log p)$ 。Bentley 提出的基于直线的求交算法时间复杂度为  $O((n + k) \log n)$  ( $n$  为所有曲线上的节点数)。Park 提出的基于单调链的求交算法时间复杂度为  $O((n + k) \log m)$  ( $m$  为单调链数)。很明显  $q < n$  且  $p < m < n$ ,故基于重叠单调链的求交算法时间要优于 Park 的方法,更优于 Bentley 的方法<sup>[12]</sup>。

## 2 曲线数据网络渐进传输实验

基于 .NET 平台和本文算法开发了曲线数据网络渐进传输实验系统,数据源采用 ShapeFile 公开交换格式。服务器端程序在传输前先分析曲线数据,计算节点偏离量,划分单调

链,确定要素权重并建立索引。在接收到客户端数据请求后,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,快速生成保持拓扑一致性的轮廓数据及各层增量数据,然后分层传输。客户端程序功能有发送数据请求、接收服务器所传数据、重构大尺度数据并显示、与服务器端程序简单交互等。图 6 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。

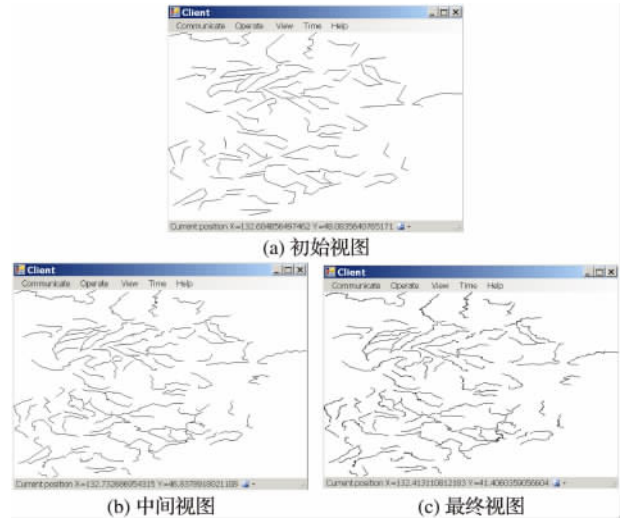


图 6 渐进传输模式下的客户端数据视图

为测试算法的时间性能,采用普通 PC 测试了不同容量的原始曲线数据到某固定分辨率曲线 ( $R1 = 25\text{ m}, R2 = 5\text{ m}$ ) 的生成时间,实验结果统计在表 1 中。

表 1 不同数据量下的多分辨率曲线生成时间

数据量 /Kb	时间 /ms	数据量 /Kb	时间 /ms
103	46	410	170
253	115	580	312

由于 Douglas-Peucker 化简算法和传统拓扑维护方法的时间复杂度都为  $O(n^2)$ ,在化简和拓扑维护两阶段所耗的时间会随数据量增大而呈非线性变化。而本文算法在此两阶段的时间复杂度分别为  $O(n)$  和  $O((q + k) \log p)$ ,从表 1 可看出,曲线的生成时间随着数据量增大呈接近线性的变化,本文算法适用于大容量曲线数据的网络渐进传输。

## 3 结语

矢量数据网络渐进传输技术的难点是矢量数据的多分辨率表达,而曲线的多分辨率表达则是矢量数据多分辨率表达的核心。本文从时间优化和保持拓扑一致性角度对曲线数据网络渐进传输进行了研究,提出一种多分辨率曲线生成算法,该算法支持对曲线要素的快速选取、概括和拓扑一致性维护。基于该算法开发了矢量曲线数据的渐进传输实验系统,验证了算法的有效性。

参考文献:

[1] BERTOLOTTO M, EGENHOFER M J. Progressive transmission of vector map data over the World Wide Web [J]. Geoinformatica, 2001, 5(4): 345 - 373.

观评价的相关性来分析,无参考的 DGS 模型均优于 PSNR 模型和文献 [1] 中提出的 BLIINDS 模型的结果。由于篇幅有限,文献 [1] 的散点图就不再列举,其实验对比效果见表 2。

表 3 LIVE2 不同失真类型无参考质量评价算法性能比较

失真类型	客观质量评价方法	CC	RMSE	SROCC	KROCC
JPEG2000	DGS 模型	0.882 50	11.482 00	0.962 70	0.827 00
	PSNR 模型	0.865 00	12.008 00	0.935 70	0.784 30
	文献 [1] 模型	0.832 00	11.887 00	0.913 00	0.793 00
GBlur	DGS 模型	0.880 08	11.426 00	0.954 60	0.812 60
	PSNR 模型	0.821 16	13.249 00	0.916 30	0.754 40
	文献 [1] 模型	0.832 00	12.793 00	0.917 00	0.788 00
Fastfading	DGS 模型	0.898 41	10.721 00	0.894 70	0.716 10
	PSNR 模型	0.858 34	12.069 00	0.869 70	0.682 40
	文献 [1] 模型	0.732 00	13.079 00	0.813 00	0.612 00

由表 3 可看出: DGS 模型对上述三种失真类型的图像评价准确度要明显优于 PSNR 模型和文献 [1] 中提出的 BLIINDS 模型。DGS 模型对图像的梯度进行奇异值分解作为图像的主要结构特征信息,能更好地反映图像的边缘和纹理的变化,符合人类视觉系统对图像的变化区域敏感的特点。

#### 4 结语

本文针对现有的无参考图像质量评价方法存在专用型但实际上无法预知图像的失真类型等缺点,提出了一种新的无参考图像质量评价方法 DGS,提取图像的梯度作为结构信息,并对梯度的相关性进行奇异值分解以更好地符合人眼的主观视觉特性,从而对失真图像的质量进行准确的客观评价。通过对 TEXAS 大学的 LIVE2 图像库中的三种不同失真类型的 575 幅图像(其中每种失真类型包含 29 幅原始图像)进行实验,充分证明了本文提出的 DGS 无参考图像质量评价模型要明显优于最广泛使用的简单有效的 PSNR 模型。后续工作中,我们会更进一步对其他失真类型的子图像库进行实验并与更多算法的性能全面比较,使得本文算法更具有通用性及实际应用价值。

#### 参考文献:

- [1] MICHELE A S, ALAN C B, CHARRIER C. A DCT statistics - based blind image quality index [J]. *IEEE Signal Processing Letters*, 2010, 17(6): 583 - 586.
- [2] 楼斌,沈海斌,赵武锋,等. 基于自然图像统计的无参考图像质量评价[J]. *浙江大学学报:工学版*, 2010, 44(2): 248 - 252.
- [3] HORITA Y, SAZZAD Z M P, KAWAYOKE Y. No reference image quality assessment for JPEG2000 based on spatial features [J]. *Journal of Image Communication*, 2008, 23(4): 257 - 268.
- [4] MOORTHY A K, BOVIK A C. A two-step framework for constructing blind image quality indices [J]. *IEEE Signal Processing Letters*, 2010, 17(5): 513 - 516.
- [5] WANG Z, HAMID R S, ALAN C B. No-reference perceptual quality assessment of JPEG compressed images [C] // *International Conference on Image Processing*. Bellingham, WA: SPIE, 2002, 1: 477 - 480.
- [6] 王正友,黄隆华. 基于对比度敏感度的图像质量评价方法[J]. *计算机应用*, 2006, 26(8): 1857 - 1859.
- [7] WANG X, TIAN B F, LIANG C, *et al.* Blind image quality assessment for measuring image blur [C] // *IEEE 2008 Congress on Image and Signal Processing*. Piscataway, NJ: IEEE Press, 2008, 4: 467 - 470.
- [8] CRETE F, DOLMIERE T, LADRET P, *et al.* The blur effect: perception and estimation with a new no-reference perceptual blur metric [C] // *Proceedings of 2007 SPIE Conference on Human Visual and Electronic Image XII*. Bellingham, WA: SPIE, 2007: 6492.
- [9] MARZILIANO P, DUFAUX F, WINKLER S, *et al.* A no-reference perceptual blur metric [C] // *Proceedings of IEEE International Conference on Image Processing*. Washington, DC: IEEE Computer Society, 2002: 57 - 60.
- [10] JAIN R, KASTURI R, SCHUNCK B G. *Machine vision* [M]. New York: McGraw-Hill, 1995.
- [11] JAHNE B, HAUBECKER H, GEIBLER P. *Handbook of computer vision and applications* [M]. Cardiff: Academic Press, 1999.
- [12] TAKEDA H, FARSIU S, MILANFAR P. Kernel regression for image processing and reconstruction [J]. *IEEE Transactions on Image Processing*, 2007, 16(2): 349 - 366.
- [13] Final report from the video quality experts group on the validation of objective models of video quality assessment, phase I [DB/OL]. [2012-05-20]. <http://www.its.bldrdoc.gov/vqeg/downloads.aspx>.

(上接第 690 页)

- [2] YANG B S, PURVES R, WEIBEL R. Efficient transmission of vector data over the Internet [J]. *International Journal of Geographical Information Science*, 2007, 21(2): 215 - 237.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 杨必胜,李必军. 空间数据网络渐进传输的概念、关键技术与研究进展[J]. *中国图象图形学报*, 2009, 14(6): 1018 - 1023.
- [7] 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.
- [8] VISVALINGHAM M, WHYATT D. Line generalization by repeated elimination of points [J]. *The Cartographic Journal*, 1993, 30(1): 46 - 51.
- [9] ALAN S. Topologically consistent line simplification with the Douglas-Peucker algorithm [J]. *Cartography and Geographic Information Science*, 1999, 26(1): 7 - 18.
- [10] da SILVA A C G, WU S T. A robust strategy for handling linear features in topologically consistent polyline simplification [EB/OL]. [2010-10-10]. <http://www.geoinfo.info/geoinfo2006/papers/p69.pdf>.
- [11] CORCORAN P, MOONEY P, BERTOLOTTO M. Line simplification in the presence of non-planar topological relationships [C] // *Lecture Notes in Geoinformation and Cartography: Bridging the Geographic Information Sciences*. Berlin: Springer, 2012: 25 - 42.
- [12] BENTLEY J L, OTTMANN T A. Algorithms for reporting and counting geometric intersections [J]. *IEEE Transactions on Computers*, 1979, 28(9): 643 - 647.
- [13] PARK S C, SHIN H. Polygonal chain intersection [J]. *Computers & Graphics*, 2002, 26(2): 341 - 350.