

# 矢量电子地图的线目标在线简化\*

李伟生

(武汉测绘科技大学信息工程学院,武汉市珞喻路39号,430070)

**摘要** 为了适应实时 GIS的变比例尺显示与网络 GIS的快速矢量数据传输,提出了数字地图的在线综合概念,讨论了对线目标的在线简化算法。提出的两个在线简化算法都是对 DP简化算法进行改造,使算法对时间的复杂度从  $\theta(n) \sim \theta(n^2)$  的乘法运算改进为  $\theta(1) \sim \theta(n)$  的赋值运算,并讨论了减少算法所涉及的空间复杂度的可能性。

**关键词** 算法分析;二叉树算法;在线综合

**分类号** P283. 7; TP391

## 1 在线综合的意义

RS GPS与 GIS的结合,加速了 GIS在各应用领域的渗透。GIS的广泛应用引出了一个越来越不容忽视的问题:快速清晰地显示各种分辨率的地图。

与纸上地图相比,电子地图有以下几个特点:(1)显示媒体(屏幕)的大小是固定的,只有小比例尺地图才能显示出全貌,而大比例尺地图只能显示部分,无法满足细节丰富与总览全貌共存;(2)某一固定比例尺的数字地图可在屏幕上放大缩小,但当缩小超过一定限度时,地图目标会堆积、拥挤在一起,使得目标无法辨认<sup>[1,2]</sup>。电子地图常用的漫游与缩放操作无法满足清晰显示的要求。

为了使电子地图更容易辨认,一些系统采用了将几个不同比例尺的电子地图统一组合到一个单一的空间结构的方案。这一结构自动地根据当前显示分辨率,选择合适的比例尺对地图进行显示。这种多比例尺地图的结构明显地改善了地图显示效果。但是,应注意到,比例尺层次太多,会增加存贮、显示、查询的难度,也对计算机的性能要求增高;如果比例尺的层次太少,会使每一个比例尺的地图担负过大范围的尺度的显示,目标拥挤的现象必然发生。

让一固定比例尺的地图能最大限度地缩小而不影响显示清晰程度具有十分重要的意义。这样,可以极大地减少 GIS系统中所需不同比例尺层次。对于这一要求,自然会想到利用地图制图的综

合技巧。但是,常见综合技巧无法快速地实现综合目标,而实时地实现综合才是电子地图所需的。

从下面的例子中,可以感受到实时综合的不容忽视的意义:

某一线目标在数据库中由 1 000个点表示,把它缩小到某个需要的尺寸显示在屏幕上,在屏幕上只要 100个点就能清楚地描述它。如果能快速地挑出这 100个点,就没有必要将那 1 000个点逐一进行坐标变换并全部堆在屏幕上。这既可节省时间,还使得显示出的线目标更清晰。这种思想若用于地图在网络中的传输则更有意义,能大为减少网络传输的负担。

不妨称在显示和网络传输中所进行的综合为在线综合。在线综合力求以最快的方法实现最清晰的显示和最少量的传输。以计算机算法分析为基础,以地图与制图综合技巧为借鉴,是实现在线综合的保障。在线综合的基本对象是点、线、面及其集合,其中线的综合在矢量表示的地图中占有重要地位。本文只讨论线的在线简化问题。

## 2 线目标的简化

线状目标简化常用于消除地图数字化时的噪声,也用于从大比例尺地图产生小比例尺地图。基本目标是从表示线目标的点序列中去掉多余的点,尽量地用较少的点表现符合所需精度的视觉效果。目前最为流行的线简化算法当属 DP算法<sup>[3]</sup>。为了比较,下面简要地描述 DP算法:

设线目标由一个点序列表示。

收稿日期: 1996-10-07. 李伟生,男,36岁,副教授,现从事 GIS与数据库研究

\* 由测绘遥感信息工程国家重点实验室开放研究基金资助项目,编号 WKL(96)0303

(1)对线目标的起始点与终结点构成的直线段,计算该目标上所有其它点到这一直线段的距离,并求出产生最大距离的点  $A$  和它相对应的距离  $D$ ;

(2)如果最大距离  $D$  小于一个指定的阈值  $d_0$ ,则去掉点序列中除端点以外的点,以两端点的连线代替原线目标,否则以  $A$  点分割线目标,对分割出的两个线目标分别重新考虑以上两步骤。

以上算法以线目标分割为基础,每一次分割都要计算当前线目标除端点外的各点到端点连线的距离。用递归的方法和栈的方法都可实现上述算法。事实上,递归的方法通过编译后,其内部实现仍然是借用栈。对于算法的时间复杂度而言,最好的情况下,一次分割都不做,此时仅检查了一次最大距离,复杂度为  $\theta(n)$ ,其中  $n$  是线目标中点的个数。最坏的情况下,分割不断进行,一直到各线目标两端点间再无其它结点,此时,时间复杂度为  $\theta(n^2)$ 。

在 DP 算法中,最费时间的操作是求最大距离  $A$  不妨称  $A$  为最凸点,其对应的距离称为最凸距。为了赢得时间,我们采用以空间换时间的策略,预先将可能的分割点求出并存贮起来,以求大大减少所需线目标简化时间,使这一操作可以满足用户的实时需求。

### 3 空间换时间

根据 DP 算法特点,采用分治策略,使用二叉树结构,对每一线目标构造一索引,并将二叉树结构的索引以顺序结构存贮起来。在显示或传输线目标时,根据索引仅挑选部分点实施精度允许范围内的简化操作。这就是空间换时间的在线简化思路。称这样的索引结构为线目标简化索引。

#### 设计二叉树结构

设线目标由点的序列构成,  $line = \{P_1, P_2, \dots, P_n\}$ , 其中  $P_i$  是线目标的结点。以下用点的序号构造一索引树:

(1)以  $n$  为根, 为其右子结点

(2)求出  $P_1, P_n$  连线的最凸点  $P_i$  和最凸距  $d_i$ , 以  $i$  为  $n$  的左子结点。

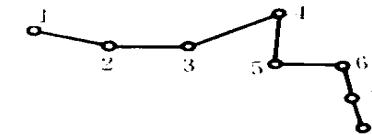
(3)对  $P_i$  分割出的左、右两个线目标分别求出最凸点和最凸距, 作为  $i$  的左子结点和右子结点, 并对每一次新分出的线目标重复这个操作, 直到不可再分。

(4)每一个结点有 3 个信息段: 左子、右子、最凸距。终点序号为  $n$ , 定义  $n$  的最凸距为某个足够

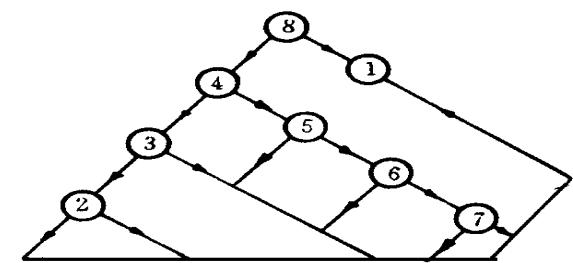
大的常数  $D$ , 定义起点的最凸距为 0

(5)为了控制计算终止, 让每一个叶结点的左、右子都是 1, 补充每一个只有一个儿子的结点的另一个儿子也为 1

图 1 中 (a) 的简化索引即为 (b)。



(a) 线目标



(b) (a) 的索引树

图 1 线目标及其索引树

Fig. 1 Line Object and Its Index Tree

在上述树结构中,  $D$  是最大可能阈值。当最大凸距大于  $D$  时, 都将其定为  $D$ 。考虑这个树的结构, 从根结点的左子开始中根遍历, 恰为该线目标去掉端点后的排列顺序, 不妨称这一索引结构为中根索引树。

以下讨论实现上述索引树的顺序结构、建树操作和基于这个结构的简化算法。

为了存贮索引结构到外存贮设备, 顺序结构是必需的。

借用 C 语言结构, 定义二叉树中的结点如下:

```
typedef struct Nodetag {int r; // 右子树序号
                        int l; // 左子树序号
                        float d; // 当前凸距
} Node;
```

对线目标  $line = \{P_1, \dots, P_n\}$  定义其线目标索引树为:

Node ST[n];

每一个点  $P_i$  所对应的结点为  $ST[i]$  这样,  $n$  点线目标的简化索引树的空间占有量为:

(2)  $\text{sizeof}(\text{int}) + \text{sizeof}(\text{float}) \times n$

以下讨论建树操作。

用  $A$  表示栈,  $A \leftarrow E$  表示  $E$  压栈,  $E \leftarrow A$  表示从栈中弹出一个值给  $E$ 。initST( ) 表示初始化

ST[ ]使所有元素的  $l$  分量和  $r$  分量都为 1,除  $n$  以外的结点  $d$  分量为 0,但  $n$  结点的  $d$  分量为常数  $D$  initStack()表示使栈  $A$  空。函数

int MaxDistance(S, E, d, t, Line, n)

表示对  $n$  点线目标 Line 求第  $S$  个点到第  $E$  个点之间的最凸点  $t$  和最凸距  $d$ 。如果存在这样的点,返回 1,否则返回 0。图 2 描述了建树过程。

这一算法对线目标找出最凸点,并以之分割线目标,直到完全分割。分割点按中根索引树的定义结构赋值给 ST[ ]。由于 MaxDistance() 操作,其算法难度为  $\theta(n^2)$ 。但是这个算法只在建库时使用一次,以后的显示与网络传输操作都将由此获益。

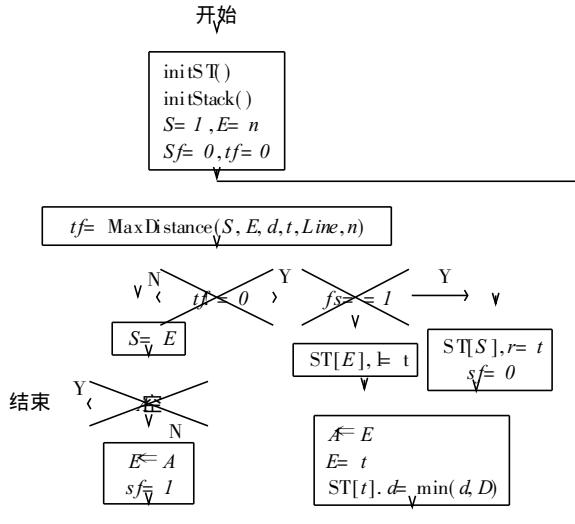


图 2 建简化索引树程序图

Fig. 2 The Program for Making Index Tree

将这个算法产生的 ST[ ] 存于磁盘,一旦有显示或传输操作,直接将之读入内存,据此实施简化操作。以下讨论基于简化索引的简化运算。

设 AddToDrawLine( $P$ ) 表示把点  $P$  接到简化了的线目标点列的后面。图 3 表示了简化算法程序流。

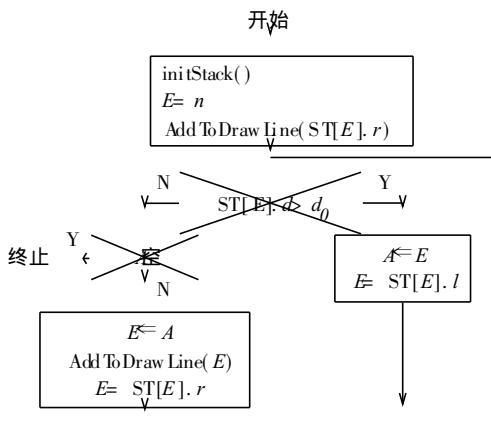


图 3 在线简化算法

Fig. 3 The Algorithm for On-line Simplification

这个算法基于中根条件遍历,即当按中根方式遍历到结点  $P$  时,如果  $ST[P].d < d_0$ ,则认为  $P$  是叶结点,不再考虑其子孙,否则深入到  $P$  的子孙。对于图 1,如果结点 2 和 5 的凸距都小于  $d_0$ ,实际遍历部分和简化结果分别如图 4 的 (a)、(b) 所示。

现在分析其复杂度。最复杂情况下,所有线目标中点都保留。此时,遍历了整个简化树,因而算法难度为  $\theta(n)$ ,这与 DP 算法最简单情况一样。实际上还要略胜一筹,因为在 DP 算法中算法的次数是乘法次数,而本算法中没有乘法运算,因而更优。最简单情况下,只需遍历一个点,即简化结果为连接两端点。此时算法复杂度为  $\theta(1)$ 。

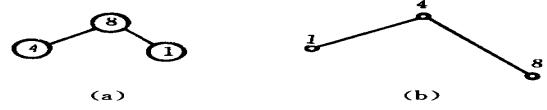


图 4 图 1 的实际遍历部分和简化结果

Fig. 4 The Actual Visit Parts and the Simplified Result of the Fig. 1

实际上,如简化运算从  $n$  个点中选出了  $m$  个点,则算法复杂度为  $\theta(m)$ 。对于本文开头的例子,仅仅只需  $\theta(100)$  次运算而不是  $\theta(1000)$ ,更不是 DP 最大复杂度  $\theta(1000^2)$ 。这使得快速实时地简化线目标有了保障。

由以上讨论,我们通过预先计算贮存分割点的二叉树,用  $(2 \cdot \text{Sizeof(int)} + \text{Sizeof(float)}) \cdot n$  的空间实现了大大改善时间复杂度的算法。这一改进可使矢量电子地图的显示和传输得以极大地优化。显然,除了线性系数可能有一些改变外,不可能在时间复杂度上有比之更优的算法了。新的问题是,有没有可能用较小的空间代价来换取同样的时间优势呢?结论是肯定的。现在就此作出进一步分析。

#### 4 更小空间代价的算法

为了达到减少空间消耗的目的,修正上节定义的简化树结构。

前 4 步骤不变,第 5 步骤改为:

(5) 将生成的二叉树按先根遍历顺序排序,并依此序将线目标的结点列重排  $Line' = \{P_1', \dots, P_n'\}$

(6) 将二叉树按新排序重新编号。为控制程序结束,给序号  $n$  加右子结点  $n+1$ 。如序号  $i$  没有左子和右子,令其右子为结点  $n$ ;如  $i$  有左子但无右

子,令其右子为  $n+1$  结点  $n$  和  $n+1$  的  $d$  值都为 0

由树的构造,不妨称之为先根索引树

图 1 所示线目标的先根索引树如图 5 线目标的点列重排后为  $Line' = \{P_8, P_4, P_3, P_2, P_5, P_6, P_7, P_1\}$

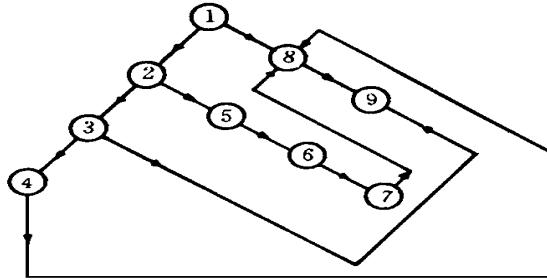


图 5 图 1(a)的先根索引树

Fig. 5 The Preorder Index Tree for Fig. 1

以上结构的顺序结构可用 C 语言定义如下: 对结点

```
typedef struct SNodeTag {int r; //右子树
                           float d; //凸距
                           } SNode;
```

对  $n$  结点的线目标,先根索引树就定义为

```
SNode SST[n+1];
```

与中根索引树相比,每个结点少了一个表示左子的分量,但总的多了一个结点,其空间需求为  $(\text{Sizeof}(\text{int}) + \text{Sizeof}(\text{float})) \cdot (n+1)$ , 比前一节的少用  $\text{Sizeof}(\text{int}) \times (n-1) - \text{Sizeof}(\text{float})$  字节。开始

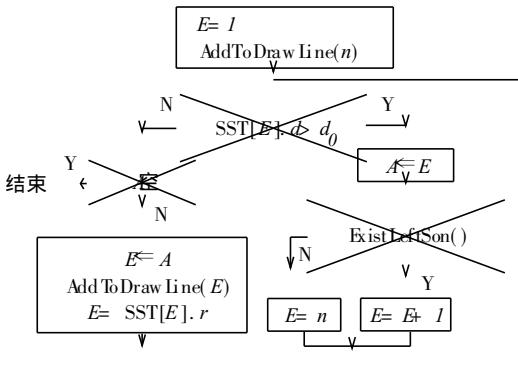


图 6 基于先根索引树的算法

Fig. 6 The Algorithm Based on Preorder Index Tree

在以上的结构里没有左子分量,它如何表示左子呢?对于  $i$ ,如果  $SST[i].r$  不是  $i+1$ ,也不是  $n$ ,则其左子是  $i+1$ ,否则无左子。以下用过程 `ExistLeftSon()` 按上述原则测试是否有左子。显然这个过程的复杂度为  $\theta(1)$ 。

建立先根索引树的算法与建立中根索引树的算法差别不大,复杂度也是  $\theta(n^2)$ ,现在着重讨论基于这一结构的简化算法。算法如图 6

这一算法的本质仍与前节的一样,复杂度最好情况下为  $\theta(1)$ ,最坏情况下为  $\theta(n)$ 。

## 5 讨 论

第 3 节和第 4 节的算法都显著地改进了 DP 算法。它们的算法复杂度都是线目标中实际显示或传输的点数,在实际操作中实际处理点数很可能大低于线目标原始数据数。这对在屏幕上放大缩小地图或在网络上传输变了比例的地图都有很大的用处。

中根索引上的在线简化算法比先根索引上的算法需要更大的索引空间,但它保留线目标原来的点顺序,即不必重新安排线目标数据库。这样,不影响基于线目标的其它操作。

先根索引上的在线简化算法以空间消耗较少而言更优,但它要求对每一线目标,将其结点按先根索引树的先根顺序重排。这样会影响线目标存储结构,也就可能影响其它算法。如果主要的应用需求是显示或传输,则此算法明显优于中根索引上的算法;否则,则应考虑这个结构对其它算法的影响。

## 参 考 文 献

- 1 李伟生.数字环境下的地图传输.武汉测绘科技大学学报,1994,19(4):359~362
- 2 李伟生.地图制图综合中相邻关系的概念框架及例子.测绘学报,1995,24(23):231~238
- 3 Douglas D H, Peucker T K. Algorithms for Reduction of the Number of Points Required to Represent a Digitized Line or Its Caricature. The Canadian Cartographer, 1973, 10(2): 112~122

## On-Line Simplification for Digital Map

*Li Weisheng*

(School of Information Engineering, W TU SM, 39 Luoyu Road, Wuhan, China, 430070)

**Abstract** For zooming in or out the map in a real time GIS application, and for quick transportation of the vector map in a network, this paper proposes the on-line map generalization concept. It discusses the on line simplification to the line object in detail. Two algorithms are proposed. Their Complex are all from  $\theta(1)$  to  $\theta(n)$ . It is a great progress comparing the normal DP algorithm whose complex is from  $\theta(n)$  to  $\theta(n^2)$ . The discussion about the reducing the spatial complex is also involved in the paper.

**Key words** computer algorithm analysis; binary tree; on-line simplification

(上接第 145 页)

甚至于地图综合原则的改变,如地物选取的阈值、统计数据分级、背景层面筛选量。制图与读图融合的程度是一个值得研究的课题,对于普通用户,地图在内容表达方面的设计应预先制作固定,而在形式表达方面可以让用户参与。

### 参 考 文 献

1 MacDougall E B. Exploratory Analysis. Dynamic Statistical Visualization, and Geographic Information Systems. *Gartography and GIS*, 1992, 19(4)

- 2 Bertin J. *Semiology of Graphics*. Wisconsin: The University of Wisconsin Press, 1983.
- 3 McGranahan M. A Cartographic View of Spatial Data Quality. *Cartographic*, 1993, 30(2, 3)
- 4 郭仁忠. 电子地图(集)与电子制图系统. 地图, 1995(2)
- 5 江斌, 胡毓钜. 地图视觉化——现代地图学的核心. 地图, 1995(2)
- 6 王选. 彩色出版系统的若干技术发展趋势. 今日印刷, 1994(4)

## Map Design Based on Modern Electronic Technology

*Ai Tinghua*

(School of Land Science, W TU SM, 39 Luoyu Road, Wuhan, China, 430070)

**Abstract** Nowadays the rise of Desk Top Publishing system and electronic map is an important representation of modern cartography technology. The electronic environment provides cartographers with powerful tools, such as color adjustment, graph layer intergration, sheet layout, dynamic symbol generlization and statistic chart automation. The paper illustrates the DTP mapping system component and its characteristics, also discusses the change in map design compared with traditional manual design techniques. In the latter, the paper investigates the properties of new kind of map——electronic map or atlas, which involves dynamic, interaction and hypermedia structure. Some special principles for electronic map design are presented in the end of paper: extending visual variable of map symbol, multiple sheet single task map replacing single sheet multiple task map and leading map users participating in map design during the process of implementing map.

**Key words** map design; electronic map; DTP system