借助梯形分割的点面包含分析

崔　璨[1],王结臣[2],沈定涛[3]
CUI Can[1],WANG Jie-chen[2],SHEN Ding-tao[3]

南京大学 地理信息科学系,南京 210093
Department of Geographic Information Science,Nanjing University,Nanjing 210093,China
E-mail:cc0029@163.com

CUI Can,WANG Jie-chen,SHEN Ding-tao.Point-in-polygon analysis based on trapezoidal decomposition.Computer Engineering and Applications,2009,45(29):49-51.

Abstract: Point-in-polygon analysis has a broad range of application in many research fields.Firstly,some existing methods determining whether a point is in a polygon have been analyzed,and then,a new algorithm that based on the trapezoidal decomposition has been proposed.This algorithm decomposes the polygons into a series of trapezoidal units,so that the point-in-polygon analysis has been transformed into the problem of determining whether a point is in a trapezoidal unit.This method has already been applied to a GIS platform software,and the test results prove that this algorithm is of great stability and reliability.Moreover,it can be adapted to any complex polygons(with inner hole),with no need to process the abnormal conditions.
Key words: trapezoidal decomposition;point-in-polygon analysis;spatial analysis;geographic information system

摘　要:在诸多研究领域中,判断点是否在多边形内是一个非常基本的问题。在分析了解决这一问题的传统方法基础上,提出了一种基于梯形分割的点面包含算法。该算法将多边形分割成若干个梯形网格单元,将判断点是否位于多边形内的问题转化为判断点是否位于梯形网格单元中这一问题。算法已用于 GIS 平台软件相应的包含分析模块,试验结果证明该算法稳定可靠、适用于任意复杂多边形而无需对奇异情况进行单独处理。
关键词:梯形分割;点面包含分析;空间分析;地理信息系统
DOI:10.3778/j.issn.1002-8331.2009.29.014　**文章编号**:1002-8331(2009)29-0049-03　**文献标识码**:A　**中图分类号**:TP391.4

1　引言

判断点是否位于多边形内在计算几何、计算机图形学、地理信息学、模式识别、以及医学图像处理等众多领域是一个非常基本且十分重要的问题[1];尤其在地理信息系统中,“平面上点与多边形包含关系的计算”是一个使用频率很高且计算量很大的过程,它是许多算法(例如裁剪、相交和包含等)的基础。点的位置判断正确与否及效率直接影响到后续算法的好坏进而影响整个软件的计算速度和软件质量,因此,对于这个过程的研究一直受到学者们的关注[2]。

迄今为止,对于平面内点在多边形内外的判断,已有了深入研究。目前常用的算法有:射线法[3]、旋转角度之和法[4]、叉积判断法[5]、夹角之和检验法[5]、有向三角形面积之和符号判断法[6]、拓扑映射判断法[7]等。其中射线法应用最为广泛,射线法原理简单,过待测点作一条水平射线,求射线和多边形的交点,若交点个数为奇数,则点在多边形内;否则,点在多边形外。由于射线法存在大量的求交运算,而且对于临界情况(当射线穿过多边形的顶点时)必须做特殊处理,特别是当多边形的边数较多时,射线法存在较大的计算量。旋转角度之和判断点在多边形内外需要计算三角函数,因此计算量较大,效率低下,使其实用价值不大。FEITO 等提出采用有向三角形面积之和的符号来确定点是否在多边形内,但此判别法需要事先确定多边形的方向,计算量相对较大,并且难以处理自相交多边形的情况。

提出了一种基于梯形分割的判断点是否在多边形内的方法。其算法思路为:将多边形划分为若干梯形状的网格单元,遍历判断待定点是否位于梯形网格单元中,从而判断待定点是否位于多边形内。

2　算法原理

2.1　多边形基本概念

多边形是一个首尾相连的多边线,它可以用点序列 $P_0P_1P_2\cdots P_n$ 表示,$P_0P_1,P_1P_2,\cdots P_{n-1}P_n,P_nP_0$ 称为多边形的边,$P_0,P_1,P_2,\cdots,P_n$ 为多边形的顶点,且边数和顶点数相等[8]。

多边形在计算机图形学中有着广泛的应用,它们一般要求具有以下性质[9]:

(1)封闭:任何一条边有且只有两个端点,每一个端点都是两条边的交点;

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.40401046);国家基础科学人才培养基金(No.J0630535)。
作者简介:崔璨(1987-),女,硕士,主要研究领域为 GIS 理论与应用;王结臣(1973-),男,博士,副教授,现主要从事 GIS 理论与应用研究;沈定涛(1983-),男,硕士,主要研究领域为 GIS 理论与应用。
收稿日期:2008-11-11　**修回日期**:2009-01-19

（2）不自交：任何两条边只有在相邻的情况下才相交，并且交点就是边的端点；

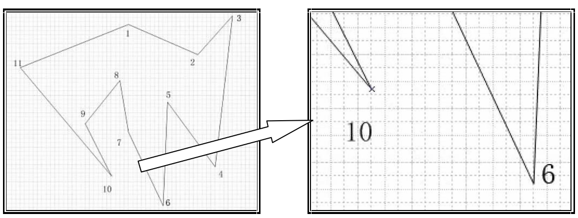
（3）有向：任何一条边都有方向，并且边的方向是一致的。规定多边形沿逆时针方向为正，顺时针方向为负。

所讨论的多边形是简单多边形，即满足三个约束条件：（1）所有的顶点均不相同；（2）多边形的任何两条非相邻边都不相交；（3）多边形的边是直线段。

2.2 梯形分割

空间数据有两种基本的表示形式：矢量数据和栅格数据[10]。矢量数据是一种面向实体的表示方法，它以坐标的形式表示点、线、多边形等具体的空间物体，其优点是数据精度高，但物体越复杂，描述就越困难，涉及的空间分析算法也较为复杂。栅格数据是一种面向空间的表示方法，它将几何空间作为整体进行描述，具体空间物体的复杂程度不影响数据量的大小。一般来说，栅格数据结构要比矢量数据结构简单得多，基于空间的几何分析也相对容易。事实上，利用栅格方法进行有关矢量分析，已有不少成功的示例[11]。

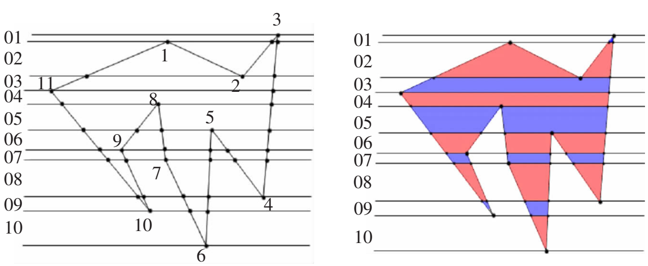
传统的栅格场是将一定空间范围划分为大小均匀的正方形（或长方形），栅格单元的大小取决于精度要求和计算机内存容许两个因素。为了保证栅格化后对多边形的分辨能力，栅格单元的大小一般小于 $L/3$，其中 L 为图幅全部多边形中边长的最小值[1]。但栅格单元越小，栅格数量越大，所需计算机内存就越多。既便如此，在对矢量数据进行栅格化时，多边形顶点不可能恰好都位于栅格单元边界处，这样就不可避免会产生精度丢失的问题（如图 1）。



（a）传统栅格数据 （b）局部放大图

图 1 传统栅格化

综合考虑到以上因素，提出对矢量多边形数据进行梯形分割[12]。即过多边形顶点引入一组平行于 y 轴的平行线，将多边形所在平面分割成若干个水平的条带（图 2.a），这组平行线与多边形各边相交，将多边形的边打断，即把多边形分割成若干个梯形网格单元（图 2.b）。每个梯形网格单元的上下底为由多边形某顶点引出的水平线上的一段，两条斜边为多边形某边的一段。采取该种梯形分割的方法，既能保证多边形端点坐标值精度不损失，又最大限度地解决矢量多边形栅格化后数据冗余的问题。



（a）扫面线分割 （b）梯形面片

图 2 梯形分割

3 算法实现过程

基于上述对矢量多边形进行梯形分割的思想，在算法实现过程，需将多边形按照梯形网格单元的方式进行存储。对任一待定点是否位于多边形内的判断，即可转换为判断待定点是否位于每个梯形网格单元中。

将多边形所有顶点（假设有 n 个顶点）按照 Y 坐标值的大小进行排序，并存入一个数组 Polygon_Vertex_Y[]中；该组水平分割线将多边形区域分割成 n-1 个条带。每行中，多边形被划分为一个或多个梯形网格单元（多边形顶端和尾端会出现三角形状的网格单元，可将其视为特殊梯形，其中一个底边退化为一个点），每个梯形网格单元的顶点都是多边形的顶点或者是水平分割线与多边形边界的交点。

构建一个行数为 n-1 的梯形网格场来表示和存放这些梯形网格单元，即创建一个大小为 n-1 的行梯形网格数组。行梯形网格数据结构 TrapGrid_Line 包含一个变量 TrapGrid_Num，表示一行中梯形网格单元的个数，以及首尾梯形网格单元指针 *Begin、*End，分别指向存放首尾梯形网格单元的梯形网格结构。梯形网格数据结构 TrapGrid_Line_Unit 包括：该梯形网格单元上下底端点的 X 坐标（即梯形四个顶点的 X 坐标）UpLeft、UpRight、DownLeft、DownRight，该梯形网格单元所属多边形的标识号 PolygonID，和指向同行中下一个梯形网格单元的指针 *next。

```
struct TrapGrid_Line
{
  int TrapGrid_Num;   //一行中梯形网格个数
  TrapGrid_Line_Unit*Begin,*End;//首尾指针
};
struct TrapGrid_Line_Unit
{
  double UpLeft, UpRight;          //左上、右上点 X 值
  double DownLeft, DownRight;      //左下、右下点 X 值
  int PolygonID;                   //所属多边形编号
  TrapGrid_Line_Unit*next;         //下一个梯形网格指针
};
```

对矢量多边形进行梯形分割后，根据待定点的 Y 坐标，在事先已排序的多边形顶点坐标数组 Polygon_Vertex_Y[]中定位，找到待定点所在梯形网格场的行号；循环遍历该行各梯形网格单元，判断待定点是否位于其中，一旦判断出待定点位于任一梯形网格单元中，即跳出循环，得出点位于多边形内部的结论；否则直到遍历完该行所有梯形网格单元，判断出点不位于多边形内。

关于判断点 P 是否位于梯形网格单元 $ABCD$ 中，采用如下算法：之前已将待定点 P 定位于梯形网格场某行中，则点 P 的 Y 坐标必定在 Polygon_Vertex_Y [LineNo-1]、Polygon_Vertex_Y [LineNo]区间范围内，即在第 LineNo 个条带中，线段 AB、CD 将条带分割形成三个区域（如图 3）；若要判断点 P 是否位于梯形 $ABCD$ 中，则可转化为判断点 P 是否位于有向线段 BA 和 DC 左侧的问题；通过判断向量 $\boldsymbol{BA}$ 到向量 $\boldsymbol{BP}$ 之间的旋转方向即可判断出点 P 是否位于有向线段 BA 的左侧，同理判断点 P 是否位于有向线段 DC 的左侧，如果答案都是肯定的，则点 P 必定在梯形 $ABCD$ 范围内。由于梯形网格数据结构中记录了梯形网格单元所属多边形的标识号 PolygonID，如点在多边形范围内，则返回结果还可反映出点所在面的标识号。

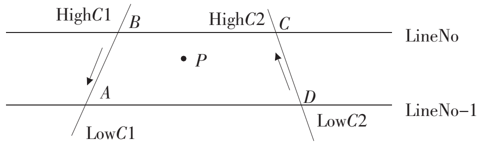


图 3　判断点是否位于梯形中的示意图

综上所述，基于梯形分割的点面包含算法的程序流程图如图 4 所示：

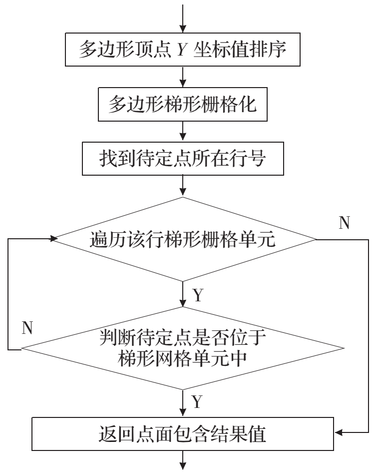


图 4　点面包含算法程序流程图

算法涉及以下几个主要计算过程：（1）对多边形所有顶点的 Y 坐标值进行排序；（2）对多边形进行梯形分割，遍历多边形各顶点，在有序数组 Polygon_Vertex_Y[]中进行二分查找，将多边形边界分隔为小段，即梯形网格单元的两边；（3）对每个点，遍历梯形网格场中的梯形网格单元，判断是否位于其中。

上述是单点是否位于多边形内的判断算法，在实际应用中，往往是对点集与多边形图层的包含关系进行分析判断，则只需在上述基础上，将单个多边形进行梯形分割扩展为对多边形图层中所有多边形进行梯形分割至同一个梯形网格场，接着遍历点集中各点，判断其是否位于梯形网格场中的梯形网格单元即可。

4　算法分析及评价

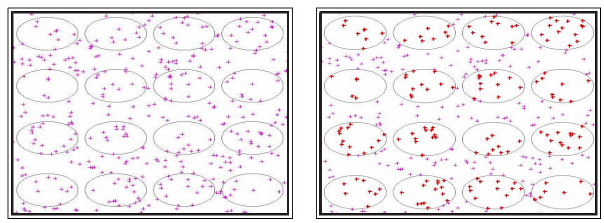
提出的点面包含算法引入了梯形网格结构这一概念，把多边形内外点的判断转化为梯形单元内外点的判断。相比传统矢量栅格化的算法，大大减少了栅格数据的存储空间；相对射线法、角度和等矢量方法，其计算过程较为简单快速，效率得到了提高。

下面来看该算法所占的存储空间：假设多边形图层中所有多边形顶点数为 n，首先，需要一个 double 型大小为 n 的数组 Polygon_Vertex_Y[n]存放多边形各顶点的 Y 坐标值，占用的空间为 O（n）；接着，需申请大小为 n-1 的行梯形网格结构数组 TrapGrid_Line[n-1]存放多边形梯形网格场，一个行梯形网格结构中有 1 个表示每行中梯形网格单元个数的整型变量以及两个指向首尾梯形网格单元的指针，占用的空间大小为 O（n-1）；每个梯形网格单元 TrapGrid_Line_Unit 所占存储空间为其数据结构中所有成员所占存储字节之和，包括：4 个 double 型变量，1 个整型变量，还有一个梯形网格单元指针。而传统的栅格存储量级是 $O(n^2)$，相对来说，提出的算法大大节约了进行点面包含分析时所占的计算机内存资源。

在执行效率上，首先对多边形所有顶点的 Y 坐标值进行排序，这里采用快速排序法，理想情况下时间复杂度为 O（$n$log$n$）；接着对多边形进行梯形分割，遍历多边形各顶点，在有序数组 Polygon_Vertex_Y[]中进行二分查找，时间复杂度为 O（log$n$）；最后，对每个点，遍历梯形网格场中的梯形网格单元，判断是否位于其中，最坏情况下遍历次数为梯形网格单元个数。对算法进行时间测试，测试数据面图层大小为 8.44 M，图层中有 2 500 个多边形，点图层的点数为 6 400 时，进行包含分析耗时 1.7 s，结果较为理想。且由算法实验结果可知，随着多边形的顶点数增多、复杂度增大，算法的执行效率优越性愈加突显。

将多边形分割成梯形网格单元来进行包含关系的判断，整个过程几何意义明显，便于理解和实现。该算法已成功应用于 GIS 平台软件的点面包含分析模块中，证实了该算法的可靠稳定和快速。

图 5 为点面包含判断的结果示意局部放大图（大红色（深色）的点在多边形内）：



（a）点面包含判断前局部放大图　（b）点面包含判断后局部放大图

图 5　点面包含算法结果示意图

5　结论

点面包含算法在地理信息系统中是一个使用频率很高且计算量很大的过程，它是许多算法，如裁剪、相交和包含等空间分析的基础。该问题的研究具有重要的理论价值和工程实际意义[1]。基于梯形网格这一结构，提出了判断点在多边形内外的算法，该算法首先对多边形进行梯形分割，在此基础上，通过判断点是否位于梯形网格单元中进而判断其是否位于多边形内。试验结果表明方法具有以下特点：对于环状及非相邻边之间相交等复杂多边形同样适用；回避了传统射线算法所遇到的奇异性问题；梯形网格的引入使得算法执行效率在工程上可以接受。

参考文献：

[1] 夏仁波，刘伟军，王越超.点在平面多边形内外的判断方法[J].机械工程学报，2006，42（3）：130-134.

[2] 沈陈华.平面上点与多边形包含关系的 Q 算法[J].扬州大学学报，1999，2（4）：24-26.

[3] Taloy G.Point in polygon test[J].Survey Review，1994，32（254）：479-484.

[4] Balbes R，Siegel J.A robust method for calculating the simplicity and orientation of planar polygons[J].Computer Aided Geometric Design，1991，8（4）：325-327.

[5] 孙家广.计算机图形学[M].3 版.北京：清华大学出版社，2000：414-418.

[6] Feito F，Torres J C，Urena A.Orientation，simplicity，and inclusion test for planar polygons[J].Computers & Graphics，1995，19（4）：595-600.

[7] 温星，陆国栋，李基拓.基于拓扑映射的点集在凸多边形内外判断算法[J].中国图象图形学报，2003，8（4）：468-471.

（下转 124 页）