

利用双邻点判断法优化拓扑关系自动生成算法

杨海宏 李天宏

薛安

(北京大学遥感与地理信息研究所,北京 100871) (北京大学环境工程研究所,北京 100871)

摘要 拓扑关系自动生成算法的效率直接影响地理数据空间关系的建立和查询等操作的性能。作者在实际的软件设计过程中,发现双邻点判断法可以在算法至关重要的2个环节处大大减少运算量,显著提高算法效率。这2个环节就是多边形的区域归属判断以及点与多边形包含关系的判断。

关键词 双邻点判断法,拓扑关系自动生成,点与多边形包含关系判断

1 引言

地理信息系统不仅存储管理地理实体的空间位置及其自然社会经济等属性信息,而且建立并维护地理实体之间的拓扑关系(连通性、邻接性、包含性等)。对于采用矢量数据结构的地理信息系统而言,产生并维护拓扑空间关系是空间数据表达的关键环节。在二维数据平面上,矢量数据结构的基本数据类型为点、线、多边形。不同类型数据之间以及相同数据类型之间存在着点点重合、点线共线、点被多边形包含、线线重合、线线相交,线被多边形包含、线穿越多边形、多边形与多边形相邻、多边形被多边形包含等种种空间关系。本文着重于点与多边形、多边形与多边形的空间关系的判断。对于这些空间关系的判断,虽然已有成熟的算法,但是作为空间数据管理的关键环节,算法的效率直接影响着系统的性能。尤其在进行空间数据查询和叠加分析等操作时,需要对多边形的空间关系进行动态或局部的判断与更新^[1],其算法的效率更成为直接制约操作性能的因素。因此,对于拓扑关系自动生成算法的优化一度是人们研究的热点。

在早期的地理信息系统中,人们往往采用手工赋值的方式为多边形建立空间关系。这种人为赋予弧段左右码相应多边形序号的方法,既耗费时间,又容易因左右码缺失或左右码调换造成错误,从而影响高层的空间运算(多边形查询,图层叠加,缓冲区

分析等)。后来人们设计出拓扑关系自动生成算法,可以使用户在输入空间数据时不必考虑空间实体的输入次序和相互关系,而由系统自动完成空间拓扑关系的建立。要正确地进行拓扑关系的自动建立,一般需经过弧段的相交处理和空间数据的完整性处理2个预处理阶段,消除悬挂结点以及弧段自相交或弧段互相交现象,即平面图化参与拓扑关系自动建立的空间数据^[1]。

2 多边形拓扑关系自动生成算法

多边形的拓扑关系自动生成一般采用旋转搜索区域边界法。在拓扑关系生成过程中,需要建立几个链表结构,分别管理弧段、结点、以及搜索过程中识别出来的多边形与岛等信息。

算法基本约定如下:

区域约定为:一个连通的平面域,由0到多个闭合边界围成;

负区域约定为:一个区域的一个闭合内边界围成的平面域,一般称为岛;

多边形约定为:区域的闭合边界,一个区域由0~1个外多边形和0~n个内多边形围成。区域的内多边形所围成的另一区域被称为本区域的岛;

多边形的方向约定为:逆时针方向为多边形的正方向。区域的外边界用逆时针的多边形表示,区域的内边界用顺时针的多边形表示;

多边形的区域归属判断约定为:判断一个多边形是一个区域的外边界还是另一个区域的内边界;

多边形拓扑关系自动生成约定为:采用搜索区域边界弧段法,生成多边形,判别多边形的区域归属,求出多边形的左右区域编码,并可在此基础上生成区域。

弧段的数据结构主要成分约定为:

```

struct ARC{
short int  m_nCode[2];
//m_nCode[0]-左区域码,m_nCode[1]-右区域码
//左区域为弧段前进方向的左侧区域
short int  m_nPoints;//构成弧段的平面点数
POINT * m_pPoints;//构成弧段的平面点};
结点的数据结构主要成分约定为:struct NODE{
POINT m_ptPos; //结点的位置
LINKARC m_linkFirst;
//本结点的连接弧段链表的首弧段};

```

连接弧段的数据结构主要成分约定为:

```

struct LINKARC{
BYTE m_bFlag;
//结点在弧段 m_pArc 的头或尾,0 - 头,1 - 尾
double m_lfAzimuth; //方位角
ARC * m_pArc;//所连接的弧段
LINKARC * m_pAnother;
//与 m_pArc 弧段相关的另一个连接弧段
LINKARC * m_pNext;
//结点的连接弧段表中的下一个连接弧段};

```

算法基本过程如下:

多边形拓扑关系自动生成算法

输入:平面图化处理后的弧段链表

输出:正确建立拓扑关系的弧段链表(弧段左右码已明确记录其左右侧区域的序号)

过程:

步骤 1 初始化弧段链表中的所有弧段。

给所有弧段的左右区域码赋初值,一般取-1,表示一个无效的编码值。

步骤 2 遍历弧段链表,搜索所有弧段的首尾端点(即结点),建立结点链表。

结点链表中的每一个结点元素不仅记录结点位置(x,y),而且其中建立有相应的连接弧段链表(请与上述弧段链表相区别)。连接弧段链表中的每一元素记录结点位于相连弧段的部位(头/尾)。所有连接弧段按其于结点的连接方位角在连接弧段链表中从大到小排序。连接方位角的确定原则是:以结点为原点,以 Y 轴正方向为正北方向,取值范围为[0, 360)。如图 1 所示,结点链表中以 N 为代表的结点

元素不仅记录了 N 的坐标位置,而且管理一个连接弧段链表,其中有 3 个连接弧段元素,在链表中的先后次序为 A₁, A₂, A₃。

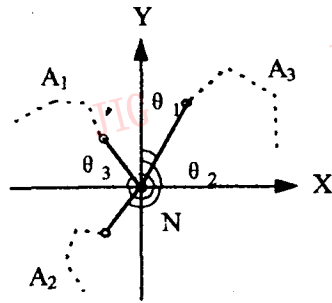


图 1 连接弧段按方位角排序

步骤 3 从弧段链表中顺次选择起始弧段进行多边形边界搜索。

若当前弧段左码为-1,将其作为起始弧段,选择其尾结点为当前结点,继续步骤 4。否则,若其右码值为-1,也将其作为起始弧段,选择其头结点为当前结点,继续步骤 4。若当前弧段左右码均非-1,从弧段链表中选择下一弧段,继续本步骤。若所有弧段的左右码已经均非-1,表明所有弧段的两侧都已被搜索,继续步骤 7。

步骤 4 搜索一个完整的多边形边界。

在结点链表中定位当前结点,按逆时针方向(也可按顺时针方向,为叙述方便,我们选择逆时针方向做讨论)在其连接弧段链表中寻找起始弧段的后继弧段,把后继弧段的另一端点设为当前结点,重复进行后继弧段寻找工作,直到返回到起始弧段,表明一个完整的多边形边界已经整理完成。如图 2 所示,从弧段 ab 前进方向左侧开始搜索,形成区域内边界 B_{abca};从弧段 ab 前进方向右侧开始搜索,则形成区域外边界 B_{adba}。

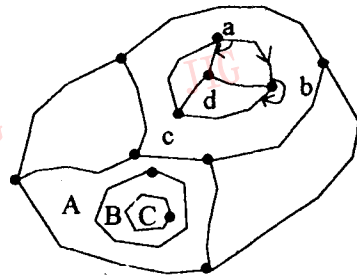


图 2 搜索多边形边界

步骤 5 判断多边形边界的区域归属。

通过步骤 4 形成的多边形可能是区域的外边

界,也可能是另一区域的内边界。由于本算法按逆时针方向搜索多边形边界,故而外边界必为逆时针多边形,内边界必为顺时针多边形。通过判断多边形的方向,即可判断多边形边界的区域归属。若形成的是逆时针多边形,往区域链表中加入一个区域元素,按递增顺序赋予区域序号(从 1 开始),并把多边形边界弧段所属区域侧的区域码赋予区域序号。若形成的是顺时针多边形,往区域链表中加入一个负区域元素,按递减顺序赋予负区域序号(从 -1 开始),把多边形边界弧段所属区域侧的区域码赋予 0 值。

步骤 6 返回步骤 3。

步骤 7 判断负区域多边形的边界弧段区域码。

此时,所有区域的内外多边形边界已经建立完成,而且任一区域的外边界多边形的所有边界弧段的本区域侧编码已经被赋予有效的区域编号。然而,形成区域的内边界多边形的所有边界弧段的本区域侧编码还没有赋予相应区域编号,这需要进行归属关系判断。

在步骤 5 中,把区域的内边界多边形用区域编号为负值的负区域表示出来。从负区域链表中顺次选出一个负区域,与正常区域做判断,求出包含该负区域的最小正常区域,再给该负区域的所有边界弧段外侧码赋予该正常区域编号。如图 2 所示,包含 C 的正常多边形是 B,而不是 A。继续这一过程,直到全部负区域的归属关系判断完成为止。

步骤 8 所有弧段的左右区域码都已赋予有效的区域编号,基本的拓扑关系建立完成。

3 多边形的区域归属两点判断法

多边形的拓扑关系自动生成算法中,步骤 5 涉及到对多边形的区域归属的判断,等价于多边形方向的判断。常规采用面积取值正负判断法,其算法基本思想如下。如果按逆时针旋转法搜索完成多边形边界后,沿搜索顺序计算该边界所包围面积,采用公式如下(图 3):

$$S = \sum_{i=1}^n \Delta S_i; \quad \text{其中,}$$

$$\Delta S_i = (y_i + y_{i-1})(x_i - x_{i-1})/2$$

面积计算完成以后,如果其结果为正值,表明多边形是逆时针方向,属于区域的内边界(图 3(a));否则多边形为区域的外边界(图 3(b))。

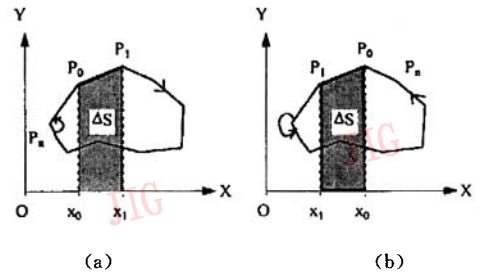


图 3 多边形方向面积判断法

面积判断方法的缺点是需要对所有边界顶点进行计算,运算量较大,时间复杂度较高。作者设计的双邻点判断法,只对极值点的前后邻点方位角进行判断,就可确定多边形方向,从而极大地提高了判断效率。算法思想如下:

对于待判边界,找出任何一个横坐标值最小点 Q(也可以是横坐标值最大点,或纵坐标值最小点或最大点),分别计算 Q 点前后 2 个邻接顶点 P 和 R 以 Q 为中心的方位角,取值方式如前所述,设 θ_P 为 Q 点前邻接点 P 的方位角, θ_R 为 Q 点后邻接点 R 的方位角,然后进行如下判断:

若 $\theta_P > \theta_R$,多边形为顺时针方向(图 4(a))。若 $\theta_R > \theta_P$,多边形为逆时针方向(图 4(b))。此方法只需要计算极值点前后 2 个邻接顶点的方位角,运算量极小,在实际的软件设计中,往往并不需要使用反三角函数进行实际方位角的计算,而只需进行一定的比值运算就可达到目的。

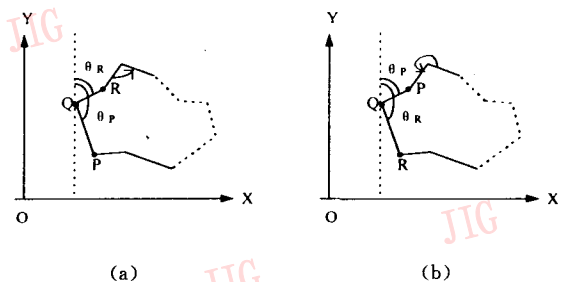


图 4 多边形方向两点判断法

4 点与多边形包含关系两点判断法

多边形的拓扑关系自动生成算法中,步骤 7 在判断负区域多边形的边界弧段区域码时,需要用到点与多边形的包含关系算法。虽然可以采用最小矩形框(Minimum Bounding Box)^[2]滤除不包含负区域的大多数正常区域,但特殊情况仍然需要采用点

与多边形的包含关系判断,即选取负区域多边形边界的任一端点与正常区域多边形进行包含关系判断。点与多边形的包含关系算法已有较多论述,一般有定边算法^[2],其局限在于要求待判多边形先剖分为凸多边形,并且难于处理待判点与边界共线情况。另外一种方法是遍历角度法^[2],其缺点是需要计算所有相邻边界点与待判点的夹角,运算量较大。目前普遍使用的是定向射线法^[2,3]。定向射线法的难点是对于如图5所示边界点与射线共线几种特殊情况的处理。

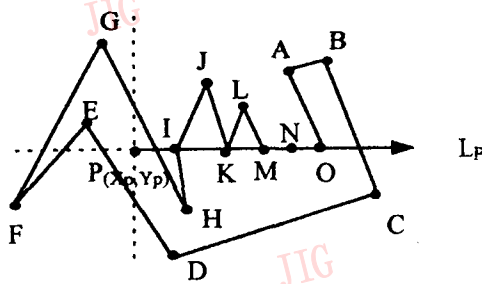


图5 边界顶点与射线的位置关系

一般处理方式是射线上下移动法^[3],对于边界坐标采用整型数值表示时,一般将从待判点所引出的射线沿与射线垂直方向移动一个微小距离,从而避免出现边界顶点与射线共线的情况。对于边界坐标采用非整型数值表示时,射线平移法的缺陷是可能需要平移多次,不但有时会增加交点计算量,而且还会因平移距离不适当,造成待判点与多边形的实际包含关系发生变化(图6)。

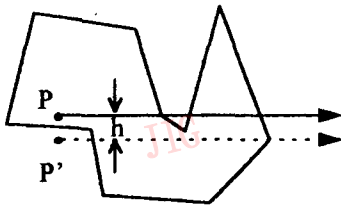


图6 下移端点射线的缺陷

为解决这一问题,作者设计了一种快速简便的算法,采用双邻点进行判断,较好地解决了这一问题。算法基本思想如下:

点与多边形包含关系判断算法

输入:待判点 P 与待判多边形的位置信息。

输出:待判点 P 与待判多边形的包含关系(真/假)。

过程:

步骤1 从待判点 P 引出射线 L_P ,并为包含关系判断算子赋初值。

从待判点 P 引出的射线 L_P 可以任意取向,为计算方便,一般沿 X 或 Y 轴正方向。我们以沿 X 轴正方向来讨论。对包含关系判断算子,若取布尔型变量 $insideFlag$,则 $insideFlag = FALSE$ 。

步骤2 选择当前边界点及其邻点。

顺次取多边形边界的顶点作为当前边界点 P_1 ,其前邻接点作为 P_0 ,后邻接点作为 P_2 。

步骤3 判断当前边界点 P_1 与射线 L_P 的位置关系(图5)。

(1) 一般情况(边界顶点与射线不共线)

这时, $y_1 \neq y_P$ 。当 $y_0 > y_P$ 且 $y_1 < y_P$,或者 $y_0 < y_P$ 且 $y_1 > y_P$ 时,即当前点 P_1 与其前邻接点 P_0 组成的线段跨越射线 L_P 所处直线时,继续进行下列判断。否则,线段 P_0P_1 与射线 L_P 无交点,返回步骤2。如图5中线段 AB 、 CD ,可以不做求交点运算。对于线段 P_0P_1 跨越射线 L_P 所处直线的情况,做如下判断:

当 $x_0 < x_P$ 且 $x_1 < x_P$ 时,线段 P_0P_1 位于射线 L_P 沿 P 点的垂线左方,与射线 L_P 无交点,返回步骤2。如图5中线段 EF 、 FG 。

当 $x_0 > x_P$ 且 $x_1 > x_P$ 时,线段 P_0P_1 与射线 L_P 肯定相交。对 $insideFlag$ 进行逻辑非操作,返回步骤2。如图5中线段 BC 。

当 $x_0 > x_P$ 且 $x_1 < x_P$,或者 $x_0 < x_P$ 且 $x_1 > x_P$,即线段 P_0P_1 跨越射线 L_P 沿 P 点的垂线时,线段 P_0P_1 与射线 L_P 可能相交,也可能不相交。这时候需要求线段 P_0P_1 与射线 L_P 所处直线的交点坐标 (x', y_P) 。若 $x' < x_P$,线段 P_0P_1 与射线 L_P 无交点,返回步骤2。如图5中线段 DE 。若 $x' > x_P$,线段 P_0P_1 与射线 L_P 有交点,对 $insideFlag$ 进行逻辑非操作,返回步骤2。如图5中线段 GH 。

(2) 特殊情况(边界顶点与射线共线)

这时, $y_1 = y_P$ 。如图5中,边界顶点 I 、 K 、 M 、 N 、 O 都属于此种类型。

若 $x_1 < x_P$,边界顶点 P_1 不在射线 L_P 上,返回步骤2。

不失一般性,如果边界顶点 P_1 在射线 L_P 之上时,其前后邻接顶点的相对位置关系如图7所示,其中 P_0 与 P_2 位置可以交换。

这时,我们只要顺次进行下列2次比较即可(依次判断前后点是否位于射线上方)。

当 $y_0 > y_P$ 时,对 $insideFlag$ 进行逻辑非操作;

当 $y_2 > y_P$ 时,对 $insideFlag$ 进行逻辑非操作。

(也可以依次判断前后点是否位于射线下方,即:

当 $y_0 < y_P$ 时,对 insideFlag 进行逻辑非操作;

当 $y_2 < y_P$ 时,对 insideFlag 进行逻辑非操作。

虽然可以只选择一方进行判断,但在整个算法实现时,选择方必须固定。其中原因,留待读者思考。)

步骤4 返回步骤2,对下一个边界点进行判断,直到所有边界点都已参与判断。

步骤5 根据 insideFlag 取值判断待判点与多边形的包含关系。判断 insideFlag,若为真,表明待判点 P 在多边形之内,否则,待判点 P 在多边形之外。待判点 P 落在多边形边界上的情况,可以在上述各个坐标比较的过程中发现,包含与否,由设计人员自行约定。

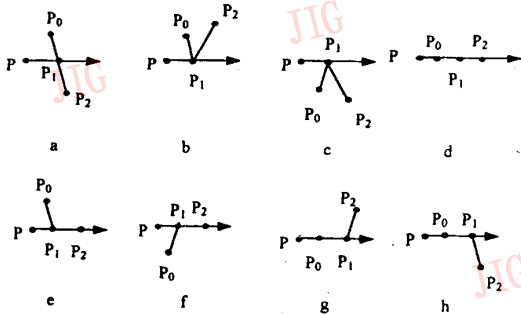


图7 边界顶点与射线共线情况

本算法彻底解决了待判点所引射线与多边形顶点共线的特殊情况,不但避免了射线平移可能引起的错误,而且省略了求交点计算,提高了整个算法的效率。使得本来复杂的情况变成了最容易处理的情况,这是本算法的优势所在。本算法不仅在拓扑关系

自动生成时可显著提高判断负区域多边形的边界弧段区域码的效率,而且可完全在选取多边形进行定位查询时所采用。

参考文献

- 1 卢开澄,卢华明. 图论及其应用. 北京:科学出版社,1995.
- 2 Worboys M F. GIS: A Computing Perspective. New York: Elsevier Science Inc., 1995.
- 3 Bonham-Carter G F. Geographic Information Systems for Geoscientists: Modelling with GIS. New York: Elsevier Science Inc., 1994.

杨海宏 1990年毕业于北京大学城市与环境学系,1993年于北京大学遥感与地理信息系统研究所获得理学硕士学位,主要从事空间数据库和网络地理信息系统的研究工作。1997年8月赴美留学。



李天宏 1992年毕业于北京大学城市与环境学系。1995年于北京大学遥感与地理信息系统研究所获得理学硕士学位,现为该所博士生,主要从事地理信息系统和遥感方面的研究和系统开发工作。



薛安 1991年毕业于南京大学大地海洋科学系。1994年获得工学硕士学位,1997年获理学博士学位,现在北京大学环境工程研究所任教,主要从事地理信息系统和环境工程结合方面的科研和教学工作。



Optimization on Automatic Topological Relation Generating Algorithm with Double Neighboring Point Judgement Method

Yang Haihong, Li Tianhong

Xue An

(Institute of RS and GIS, Peking University, Beijing 100871)(Institute of EE, Peking University, Beijing 100871)

Abstract Automatic topological relation generating algorithm influences directly the performance of such spatial operations in GIS as spatial relation establishment and query. Double neighboring point judgement method can be adopted so as to decrease computing complexity greatly in two key stages of this algorithm - determining both direction of polygon and point-in-polygon relation.

Keywords Double neighboring point judgement method, Automatic topological relation generating