

闭合坐标链多边形的拓扑关系快速构建

罗 芳¹ 艾廷华¹ 王 洪¹

(1 武汉大学资源与环境科学学院, 武汉市珞喻路 129 号, 430079)

摘 要: 基于格网化的思想, 提出针对该数据快速构建拓扑关系的算法。该算法有效利用了闭合坐标链多边形数据本身具备的多边形几何信息。试验证实, 该算法提高了运算速度。

关键词: 多边形; 拓扑关系; 网格; 算法

中图分类号: P208

“空间数据是否需要建立拓扑关系并在数据库中显式地存储记录”一直是 GIS 研究领域争论的问题, 因为空间关系可以基于空间数据的几何特征推算出来。基于拓扑结构的空数据模型, 在针对单个地理实体的操作中(如编辑更新多边形边界的某一弧段)会导致复杂的后续一致性维护^[1,2]。随着空间数据应用领域的不断拓宽, 图形可视化(网上浏览)等非空间分析功能的需求占有越来越大的比重, 这类应用不需要拓扑关系作支撑。无拓扑结构表达多边形目标采用闭合坐标链, 格式简单, 便于交换, 数据更新维护方便, 在图形系统中被普遍采用, 如当前一些无拓扑的交换数据格式 ESRI 的 shape 文件、AutoCAD 的 dxf 文件、MapInfo 的 mif 文件等。目前成为 GIS 概念数据模型主流的面向对象整体数据模型^[3], 也采用闭合坐标链存储多边形对象(包括简单多边形和简单多边形组合成的多边形集)。于是在空间数据处理中提出了为闭合坐标链多边形数据建立高效率的拓扑关系构建算法的需求。本文对该问题进行研究, 建立一种基于网格处理的拓扑关系快速构建算法。

1 基于分类网格构建闭合坐标链多边形数据的拓扑关系

1.1 算法过程

闭合坐标链多边形数据记录多边形边界点坐标, 数据无缝、无自相交、无结点、内点、弧段信息

描述。直接的方法是采用矢量线求交过程来对弧段分解, 进行弧段重组, 获得结点信息, 但是大量的点比较运算, 其计算量大, 而且丢弃了数据本身具有的多边形几何信息。基于此, 提出了利用多边形几何信息的分类网格构建闭合坐标链多边形数据的拓扑关系算法。

本文采用 Shape 格式数据作为试验数据。Shape 闭合坐标链多边形数据具有以下特点: 多边形以首末点重合的点坐标串来描述, 点在其中有序, 即沿坐标点的方向行进, 多边形的内部始终在行进方向的右边, 多边形的外部始终在行进方向的左边。在整个图层中, 两个相邻多边形的共享边(即共享点串)被存储两次^[4]。

根据上述特点, 建立搜索两个多边形共享边的算法, 生成弧段和多边形、结点拓扑关系, 主要步骤如下。

1) 矢量数据栅格化: 在栅格场内以多边形为单位, 依次判断边界点所在的栅格号(重合点处理两次, 结点可能处理三次或以上)。

2) 提取多边形 I 的边界点经过的所有栅格 $G = \{g_1, g_2, \dots, g_n\}$, 对该栅格集合元素分组: $G = G_{i, j_1} \cup G_{i, j_2} \cup \dots \cup G_{i, j_m}$, 其中 G_{i, j_m} 为多边形 I 及其相邻多边形 J_m 的栅格集合, 其内包含多边形 I 和 J_m 的边界点。

3) 对 G_{i, j_m} 内的诸栅格, 通过栅格内点之间的欧氏距离计算判断点重合的情况。将栅格分为三种类型, 如图 1 所示。类型 I: 该栅格内的所有的点都可以找到它的重合点, 即在该栅格内的多

边形 I 和 J_m 的坐标链完全重合; 类型 II: 部分点可以找到重合点, 即在该栅格内的两个多边形的坐标链有重合、有分叉; 类型 III: 所有点都找不到重合点, 即在该栅格内的两个多边形的坐标链完全不重合。

II 类栅格内有结点, 取 II 类栅格进行点的匹配, 找出进入重合状态的起始点 m , 退出重合状态的终止点 n , m 、 n 之间的点便构成了两相邻多边形 I 和 J_m 之间的共享边 Arc_{i, j_m} 。

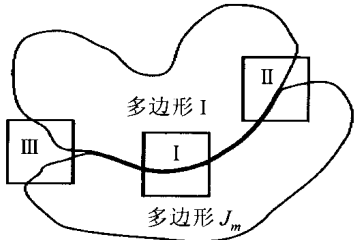


图 1 三类栅格

Fig. 1 Three Kinds of Grids

4) 建立弧段 Arc_{i, j_m} 与多边形 I 、 J_m 之间的左、右多边形关系, 以及弧段和起、终结点 m 、 n 的关系。

5) 对所有多边形按以上方法处理一遍, 多边形 1 与其后 $n-1$ 个多边形比较搜索共享边, 多边形 2 与其后 $n-2$ 个比较(为加速判断共享边, 以最小外接矩形 MBR 作过滤, 即两多边形的 MBR 不相交, 则一定可以排除多边形边界的共享^[5]), 依次比较, 直到全部多边形比较完成。

1.2 栅格化

栅格场范围为图幅范围加上外扩值(以免多边形边界与栅格边界重合)。本算法栅格化的目的是避免大量的点之间的欧氏距离计算, 所以栅格边长不能太小。在多边形栅格化的过程中, 以多边形为单位依次将闭合坐标链放入栅格内, 有可能出现闭合链跨越栅格的情况(即闭合链上有一段直线跨越该栅格, 无点落在本栅格内), 所以需将闭合链与栅格边界的交点求出并记录下来, 以增加必要的出入点信息。

1.3 共享边提取

对多边形 I 、 J_m 在栅格内的点串分别排序, 计算点之间的欧氏距离, 判断点重合的情况, 找出 II 类栅格。ISO 标准定义的拓扑原型中规定: 多边形在左边为弧段正向, 即弧段逆时针为正向^[6], 而 Shape 中定义外边界点按顺时针排序。由此可以得出, 在 II 类栅格内, 点由重合状态变为不重合状态的过渡点即为共享弧段的始结点, 由不重合变到重合的过渡点为该弧段的终结点, 且

多边形 I 为该弧段的左多边形, 多边形 J_m 为右多边形。

在判断栅格类型时, 其特殊情况如图 2 所示。图 2(c)是图 2(b)的全局示意图。图 2(a)、2(b)中加粗的 A 、 B 两点之间是两个多边形的坐标链在栅格内的重合部分, 图 2(a)和 2(b)都是部分点重合(即 II 类栅格), 图 2(b)中线 AB 是两个面的共享边的中间一部分。图 2(a)中点 B 是结点, 2(b)中没有结点, 故在搜索结点时要要把图 2(b)的情况排除掉。有很多方法可以区别图 2(a)与图 2(b), 如通过出入点的比较或是根据点在多边形边界中按顺时针排序的性质都可以排除图 2(b)的情况。

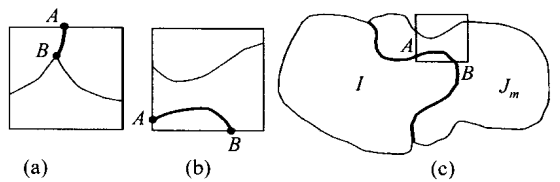


图 2 判断 II 类栅格时的一种混淆情况

Fig. 2 Confusing Cases in the Judgment of Sort II Grid

从始结点开始, 在多边形边界点串中依次搜索到终结点, 即形成共享弧段, 记录其始、终结点和左、右多边形, 即得到弧段与结点、多边形的关系。

根据以上描述, 建立搜索共享边算法伪代码如下。

在整个图幅范围内构建栅格;

for $i = 0; i < \text{多边形总数 } n - 1; i++$

{

取出比较多边形;

扫描比较多边形闭合链上的每个点, 为其所在的栅格赋值;

for $j = i + 1; j < \text{多边形总数 } n; j++$

{

取出被比较多边形;

If (两多边形外接矩形不相交)

Continue; (*两多边形没有共享边, 执行和下一个多边形的比较*)

扫描被比较多边形闭合链上的每个点, 为其所在的栅格赋值;

for (每一个栅格) (*判断栅格类型, 得到始、终结点*)

{

分别将栅格中两个多边形的出入点加入总点串中, 并排序;

for (比较多边形闭合链在栅格中的每个点)

}

}

```

{
    for(被比较多边形闭合链在栅格
    中的每个点)
    {
        比较该点的重合状态;
    }
    记录重合状态变化的过渡点; (* 如果
    没有过渡点,以多边形在栅格内点串的最后
    一个点代替 *)
    if (是 II 类栅格)
    {
        记录由重合变为不重合的过渡点
        为本弧段的始结点;
        记录由不重合变为重合的过渡点
        为本弧段的终结点;
        记录比较多边形为左多边形,被
        比较多边形为右多边形;
    }
}

```

原闭合坐标链多边形数据经过提取共享边后弧段被分割开,并得到结点和左右多边形信息。如图 3 所示,图 3(a)为原始数据,显示为闭合坐标链组成的多边形的边界;图 3(b)为执行上述算法后的结果,显示多边形由多个弧段组成。

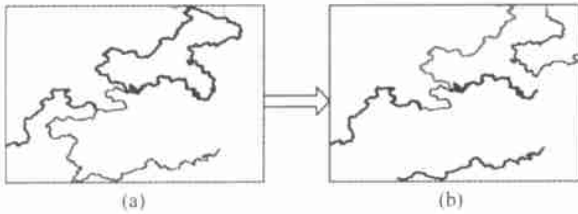


图 3 闭合坐标链多边形数据提取共享边后结果
Fig. 3 Shared Arcs Result Extracted from the Closed Coordinate Chain Data

1.4 拓扑信息存储

全部多边形处理完毕后,得到整个图幅的弧段文件,以此为基础,即可得到多边形-弧段文件、结点-弧段文件。此方法目前已比较完善,文献 [7] 提出了较详细的算法。需指出的是,本文提出的搜索算法得到的关于多边形 I 的所有弧段集合 $Arc_{i,j_1} \cup Arc_{i,j_2} \cup \dots \cup Arc_{i,j_m}$ 不一定能使多边形 I 闭合,因为本算法只搜索两个多边形的共享边,而没有搜索到多边形与空集(可设想为无限远的区域)的共享边和环状多边形的内边界,故还需添加弧段,补充其起、终结点和左、右多边形信

息,将其加入整个图幅的弧段文件才构成完整的弧段文件。基于拓扑编辑算法,提出补充外边界弧段算法控制流程图,如图 4 所示。

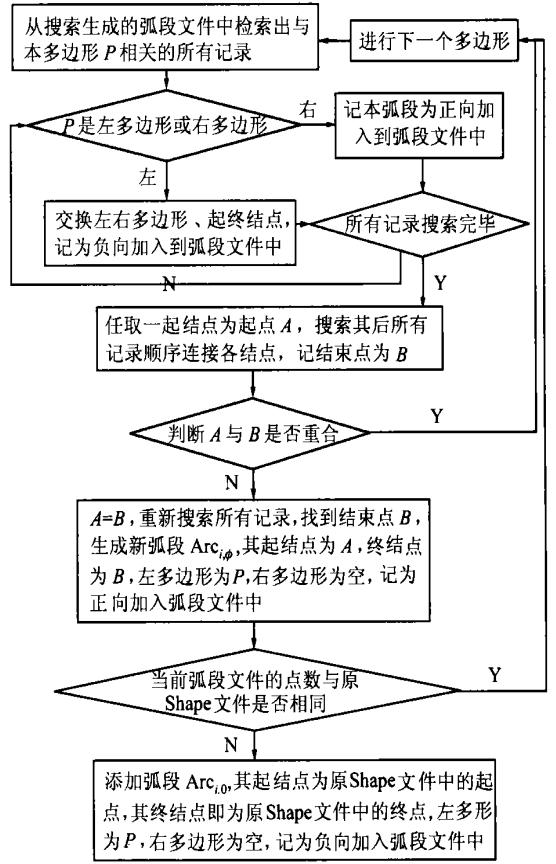


图 4 拓扑信息完善流程图
Fig. 4 Workflow of the Topology Completion

2 试验

本研究使用 VC6.0 语言开发工具,在交互式地图缩编软件 Domap 的平台上编制程序实现本算法。分别以南海市 1:1 万土地利用现状数据、全国 1:400 万省界和全国 1:25 万县界数据进行试验,均取得了较好的效果。图 5 是全国 1:25 万

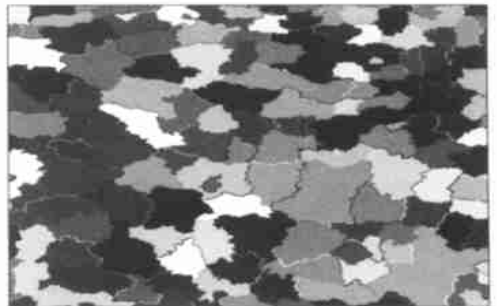


图 5 全国 1:25 万县界 Shape 数据的试验结果(部分)
Fig. 5 Experiment Based on China 1:250 000 County Boundary Shape Data

县界试验结果图的一部分, 图中使用不同灰度来表示多边形和各弧段。

3 结 语

本文提出了通过在特定类型网格中(II类格网)搜索结点的算法, 实现了闭合坐标链多边形数据拓扑关系的快速构建。与文献[8, 9]比较, 本文算法具有如下优点。

1) 采用了快速判断的优化策略, 通过判断外接矩形是否相交排除了大量多边形比较。网格化的处理则将矢量点的重合判断限定在局域范围内, 避免了全局性盲目比较, 从而提高运算速度。

2) 传统算法将获取结点与构建弧段-多边形拓扑关系分开进行, 本算法在搜索有效栅格获取结点的同时生成弧段, 确定起、终结点和左、右多边形, 所有过程一步完成, 简洁紧凑, 对于环状多边形的拓扑关系判断也一起进行, 不用再单独提出考虑。

3) 有效利用了闭合坐标链多边形数据本身具备的多边形几何信息。

本算法有待改进的地方在于栅格大小的自适应判定。栅格太小占用内存, 栅格化花费时间太长; 栅格太大意味着栅格内点较多, 通过在栅格内比较点来减少计算量就失去了意义, 可以通过反复试验建立适当的格网, 也可以基于四叉树的思想建立动态格网来优化本算法。

参 考 文 献

- 1 van Oosterom P, Stoter J, Quak W, et al. The Balance Between Geometry and Topology. SDH, Berlin, 2002
- 2 Quak W, Stoter J, Theo T. Topology in Spatial DBMSs. Digital Earth, Brno, Czech Republic, 2003
- 3 肖乐斌, 钟耳顺, 刘纪远, 等. 面向对象整体 GIS 数据模型的设计与实现. 地理研究, 2002(1): 35~44
- 4 ESRI Shapefile Technical Description, An ESRI White Paper. <http://www.esri.com>, 1998
- 5 张锦明, 何成. 利用分区思路优化拓扑关系自动生成算法. 测绘学院学报, 2000, 17(2): 119~122
- 6 OGC: The OpenGIS Guide. An Introduction to Interoperable Geo-processing. The OGC Project Technical Committee of the OpenGIS Consortium, Wayland, Mass. USA, 1998
- 7 黄杏元, 马劲松, 汤勤. 地理信息系统概论. 北京: 高等教育出版社, 2002
- 8 王杰臣. 多边形拓扑关系构建的栅格算法. 测绘学报, 2002, 31(3): 249~254
- 9 齐华. 自动建立多边形拓扑关系算法步骤的优化与改进. 测绘学报, 1997, 26(3): 254~260
- 10 艾廷华. 空间数据多边形拓扑结构的建立. 武汉测绘科技大学学报, 1995, 20(11): 82~87
- 11 艾廷华, 陈涛. 基于三角网的“种子法”多边形生成. 武汉大学学报·信息科学版, 2004, 29(1): 14~19

第一作者简介: 罗芳, 硕士生. 研究方向为 GIS 应用技术、土地信息技术。

E-mail: lula-ll@163.com

An Algorithm to Create Topological Relationship Based on Closed Coordinate Chain Data

LUO Fang¹ AI Tinghua¹ WANG Hong¹

(1 School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: This paper presents an algorithm to construct topology based on the grid handling. The algorithm uses geometric information of closed coordinate chain data available. Examination results show that it improves the speed of operation greatly.

Key words: polygon; topological relationship; grid; algorithm

About the first author: LUO Fang, postgraduate. Her research interests include application of GIS and LIS.

E-mail: lula-ll@163.com

(责任编辑: 晓平)