

# 地理信息系统中多边形叠合算法的研究\*

高显连 杨国勇 孟献策 田永林

(中国林业科学研究院资源信息研究所)

**摘要** 本文探讨的多边形叠合算法是在开发牡丹江林管局信息管理系统时提出并完成的。结果表明:采用将两幅专题图并成一张新图,重建空间数据之间的拓扑关系,重新生成多边形,重新生成标志点,并利用标志点进行属性回填的方法,可圆满解决专题图叠合后拓扑关系复杂、属性不易回填的问题,成功地实现两幅专题地图的叠合。

**关键词** 地理信息系统 多边形叠合 多边形生成 标志点判别 属性回填

地理信息系统又称空间数据管理系统,是目前国际上计算机应用领域中较为活跃的部分,广泛应用于地质、地理、环境等科学领域,及林业、农业、水利等国民经济部门。在资源清查、城乡规划、灾害监测、宏观决策、地籍管理等方面,地理信息系统已得到十分成功的应用。地图图形信息的采集、存贮、处理在地理信息系统和专题制图中有着重重要的作用。地理信息系统通过图形数字化,将森林分布、土地利用现状、土壤类型、行政区划、水系、道路、居民点等专题图形信息变成计算机的内部表示,在此基础上进行图形的多种操作,如把相邻的两幅相同专题类型的图拼接成一幅图,对某一区域进行剪裁,或进行多条件查询等。本文重点讨论图形操作中的多边形叠合操作的算法。

## 1 空间数据的数据结构和文件结构

专题地图所表示的空间实体可分为三种形式:面状、线状、点状。例如林相图是面状要素,可称之为多边形结构的专题图。交通或水系图为线状结构的专题图,居民点等为点状结构的专题图。各种空间图形要素之间均存在拓扑关系,这主要指:

**拓扑邻接:**指存在于空间图形的同类元素之间的拓扑关系。例如结点邻接关系  $N_1/N_2$ ,  $N_1/N_4$ ,  $N_3/N_4$ ……。

**拓扑关联:**存在于空间图形的不同元素之间的拓扑关系。例如结点与弧段的关联关系  $N_1/C_1, C_3, C_6$ ,  $N_2/C_1, C_2, C_5$ , ……。

**拓扑包含:**存在于空间图形的同类、但不同级的元素之间的拓扑关系,如  $P_2$  包含  $P_4$  等。

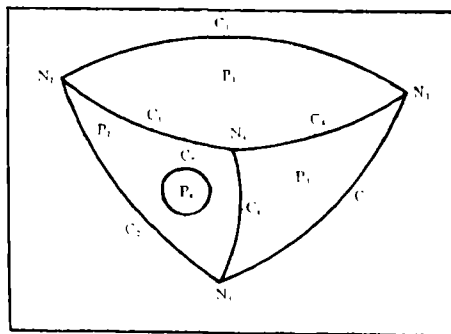


图1 空间数据的拓扑关系

1992-03-20收稿。

\*本算法已由程序实现并投入运行,本研究得到资源信息所人工智能室全体同仁的大力支持,在此表示感谢!

在具体实现中, 对不同的空间实体作分层次存贮, 对点状图记录其地理位置和属性信息。对线状图记录线的起始点和终止点, 线段起点和终点之间的多个中间点(称为折点)。线段起点或终点又称为结点, 每个结点的图形信息包括该点的地理位置, 与该结点相连接的线段。对线状图还要记录各线的属性信息。对面状图除了记录线状图的图形信息外, 还要包括每个多边形的组成线段, 多边形的属性信息, 多边形的面积, 多边形中线段的左右多边形信息, 多边形的标志点(内点), 以及多边形内包含的子多边形(岛多边形)等图形信息。

由于实际中的一幅专题图可能包含点线面的全部信息, 所以对依层次存贮的图形进行叠合, 不同类型的专题图之间有时也需要叠加, 例如林相图与土壤图的叠合。各种叠合操作包括: 点与点, 线与线, 面与面, 点与线, 点与面, 线与面等, 其中最复杂的操作是面与面(多边形与多边形)的叠合。

## 2 面状图与面状图的叠合操作

叠合操作的前提, 是两幅图控制在相同的地理范围内, 如果两幅图的比例尺不一样, 应先进行预处理, 使两幅图的数据得到匹配。叠合算法的基本思想是: 假设有两幅图(图形 1 与图形 2), 先将两幅图的数据组织到一张新图上, 然后建立新图的拓扑关系, 重新生成多边形和标志点(内点), 利用标志点进行属性回填。现将该算法分别解释如下。

### 2.1 新图的建立

(1) 将图形 1 的数据复制到新图。

(2) 以(1)建立的新图为基础图, 将图形 2 的数据作相应的处理。如图形 2 中, 表示线段起始位置的数据需加以改变, 同时由于图形 2 与图形 1 的结点可能重合, 所以还应做结点的重合处理。将处理后图形 2 的数据添加到新图中。

### 2.2 新图重建拓扑关系

由于图形 1 与图形 2 各自的拓扑关系在叠合前已经建立, 故只需考虑图形 1 与图形 2 之间可能产生的新拓扑关系即可。图形 1 与图形 2 之间拓扑关系的建立, 可简化为图形 1 中的一条线段与图形 2 建立拓扑关系的重复。判断图形 1 的一条线段可能与图形 2 的哪些线段相交, 一条线段的所有点(结点, 折点)的最大最小( $x, y$ )坐标确定了该线段的范围, 因为两条线段相交的必要条件是它们的范围必须重叠, 所以利用线段范围的比较, 可以排除大部分无关的线段。图 2 中  $AB$  线段为图形 1 中的线段,  $CD$  为图形 2 中的线段, 当两条线段相交时交点为  $O$ , 则在图形 1 中  $AB$  由  $AO$  取代, 同时添加  $OB$  线段, 即将  $AB$  分成  $AO$  与  $OB$  两条线段, 图形 2 中的  $CD$  线段也作相应的处理。两条线段相交, 产生了 4 条线段和 1 个结点。 $AO$  线段继续与图形 2 中的线段比较, 此过程一直重复下去, 直到图形 1 中的所有线段都处理完, 新的拓扑关系建立完毕。可能出现的复杂情况是相交的两条线段有重叠(图 3), 处理的方法是将  $AB$  分成  $AE$ 、 $EF$ 、 $FB$ , 将  $CD$  分成  $CE'$ 、 $E'F'$ 、 $F'D$ 。由于  $EF$  与  $E'F'$  是两条重叠的线段, 所以删掉其中的一条线段  $E'F'$ , 则  $AB$  与  $CD$  相交, 最终产生了  $AE$ 、 $EF$ 、 $FB$ 、 $CE'$ 、 $F'D$  5 条线段和 2 个结点。这一步只是建立了点线之间的拓扑关系, 涉及到面的拓扑关系则由下一步骤完成。

### 2.3 多边形生成的算法

所谓多边形的生成就是面的生成, 建立与面有关的拓扑关系。将顺时针方向定为多边形

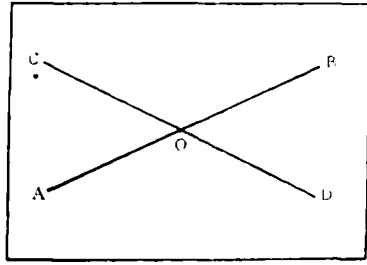


图2 线段相交

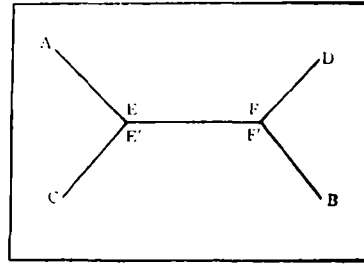


图3 相交的两条线段有重叠

的方向，对某一线段递归地求其后继线段，直到某线段的后继线段为最初的开始线段，就得到一个多边形及其组成线段的图形信息。与每条线段相连的后继线段有两条，一是与该线段的起点相连的线段，另一条是与该线段的终点相连接的线段。对一幅专题图的所有线段都作上述处理后，新图中涉及到面的拓扑关系也就建立起来了。对于岛多边形要做一些特别处理，不但要生成该岛的多边形，还要生成包括该岛的主多边形。包围该岛的多边形中最里面的一个多边形定义为该岛的主多边形(图4)。算法的具体过程是：①对每个结点所连的线段，按其出入该点时与X轴的夹角从小到大排序。②求每条线段的后继线段，即线段起点或终点上，比该线段方位角大的线段中方位角最小的一个线段。③对每条线段，递归地求其后继线段，得到组成一个多边形的所有边。④处理岛多边形的情况，求多边形面积，填写多边形文件，并填写线段的左右多边形信息。

### 2.4 标志点生成的算法

标志点是指多边形内的一点，它具有如下性质：如果一个点是某一多边形的标志点，则它就不是另一多边形的标志点，因此标志点只能在多边形内，而不可能在边界上。2.3节中生成的多边形包含多边形的范围信息，设P为生成的某个多边形，Max y 为P中坐标点(x, y)的y最大值，Min y 为P中坐标点(x, y)的y最小值，作水平线  $y = \text{Min } y + 1/2(\text{Max } y - \text{Min } y)$ ，则该线必与P有两个交点(图5)，令  $x = x_1 + 1/2(x_2 - x_1)$ ，则(x, y)即为多边形P的标志点。

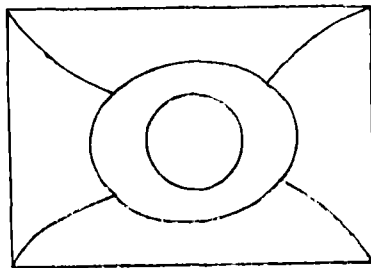


图4 岛多边形

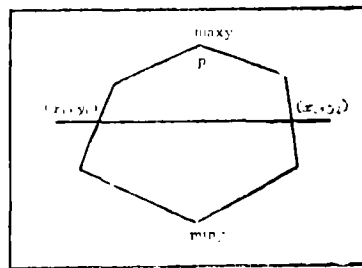


图5 求多边形的标志点

较为复杂的情况是所作的水平线可能与多边形的一条或几条线段重叠，或者与多边形有多个交点，但对算法稍做细化，问题即可解决。图6表示了几种较复杂的情况。

### 2.5 标志点的判别算法

判断一个点是否为多边形的标志点，可先判断该点是否在组成该多边形的线段上，即是否在该多边形的边界上。如果在多边形的边界上，则该点不是该多边形的标志点。在判别时，

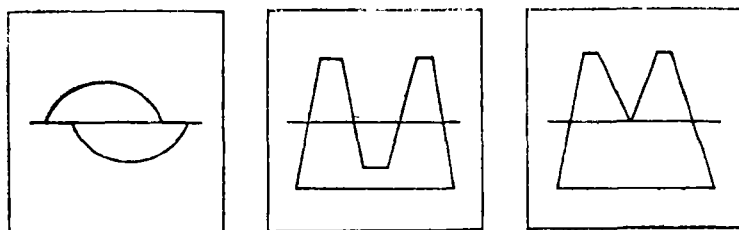


图 6 标志点的特例情况

利用线段范围来确定线段是否可能与该点重合, 可大大减少判别工作量。当确定该点不在多边形边界上时, 可用图 7 所示的铅垂线法来判别。具体做法是: 作通过该点的 Y 轴平行线, 计算出该线段与要判别多边形的交点, 如交点为偶数, 则不是该多边形的标志点, 如果是奇数, 则是该多边形的标志点。要考虑的复杂情况是可能出现线段的重叠, 或交点与多边形的凸凹有关的情况。

### 2.6 属性回填

所谓属性回填是指建立起来的新图应具有图形 1 与图形 2 的属性。例如土壤图与森林分布图叠合, 生成新图的任一个多边形应同时具有土壤种类和树种的信息。图形 1 中的一个多边形与图形 2 中的一个多边形叠合到一起, 假设其相交生成了三个新面(图 8), 这三个新面的属性可通过判断标志点来确定。例如 A 多边形与 B 多边形相交生成 A'、B'、O 三个面, 在判断 A' 面时, 如 A' 的标志点也可以作 A 的标志点, 则 A' 应包含 A 的属性信息, 当 A' 的标志点也可以作 B 的标志点时, 则 A' 中也应包含 B 的属性信息。对新图中所有的多边形都作上面的处理后, 则可将图形 1 与图形 2 中所有信息都回填到新图中去, 至此, 图形 1 与图形 2 相叠合所产生的一张新图就建立起来了。

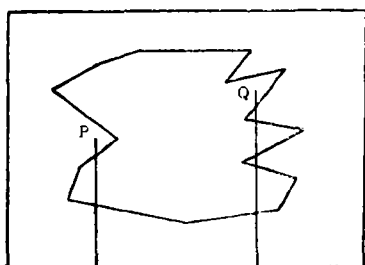


图 7 用铅垂线法进行标志点判别

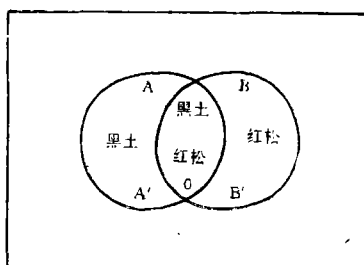


图 8 面相交

## 3 分析与讨论

评价地理信息系统中算法的优劣, 主要是看运用算法解决实际问题时所占用的时间与空间资源的多少。本文所讨论的算法, 实际运行中只增加了叠合后的新图数据文件所占用的空间, 及少量临时文件所占用的空间, 故在此不讨论算法的空间参数, 只对时间参数进行分析。

专题图叠合所耗费的时间主要是对数据文件进行读写操作的时间。现假设图形 1 包含  $m_1$  条线、 $n_1$  个结点、 $s_1$  个多边形; 图形 2 包含  $m_2$  条线、 $n_2$  个结点、 $s_2$  个多边形; 新图包含

$m_3$  条线、 $n_3$  个结点、 $s_3$  个多边形, 现对算法的各步骤分别讨论如下:

(1) 建立新图时, 必须先调整图形 2 的线段号和结点号, 需要对图形 2 的线段文件和结点文件各遍历两次, 即要访问  $2m_2$  条线段记录和  $2n_2$  条结点记录。处理结点重合时, 对图形 1 的每个结点都要进行考虑, 最差情况是对图形 2 的结点文件遍历  $n_1$  次, 即共要访问  $n_1 + n_1 \times n_2$  条记录。

(2) 重建拓扑关系时, 如果最终建立的新图有  $m_3$  条线段, 那么图形 1 与图形 2 相结合就要新产生  $m_3 - m_1 - m_2$  条线段。设图形 1 产生了  $K_1$  条新线段, 图形 2 产生了  $k_2$  条新线段, 则  $k_1 + k_2 = m_3 - m_1 - m_2$ , 那么至少要访问  $(k_1 + m_1) \times (1 + k_2 + m_2)$  条线段记录。对每个产生的新结点, 要考虑是否是旧结点的情况, 因此要对新图结点文件进行访问。

(3) 生成多边形时, 先对结点文件所连线段进行排序, 需访问  $2m_3$  条线段记录和  $n_3$  条结点记录。每条线段均存在左右多边形, 当多边形生成完时, 每条线段又被访问了 2 次, 同时要生成多边形文件和多边形的组成边文件, 相当于对这两个文件各遍历一次。

(4) 生成标志点时, 对每个多边形的组成边, 均要访问一次, 同时对每相应边(可能与水平线相交的边)的坐标文件也要进行访问, 以求出水平线与多边形的两个交点, 总共要访问  $2m_3$  条线段记录和  $s_3$  条多边形记录。

(5) 判断一个点是哪个多边形的标志点, 要访问多边形文件, 平均要访问  $1/2s_3$  条多边形记录。若可能是某一多边形的标志点, 则要访问该多边形的所有边以便判断, 对符合范围要求的边还要访问坐标文件, 以便确定交点。

(6) 回填属性时, 由于新图的每个多边形的属性均要回填, 即要判断出每个多边形在图形 1 与图形 2 内的属性情况, 平均要访问  $1/2s_3 \times (s_1 + s_2 + 2)$  条多边形记录。

## *Study on Algorism for Polygon Overlay in Geographic Information System (GIS)*

Gao Xianlian Yang Guoyong Meng Xiance Tian Yonglin

(The Research Institute of Forest Resources Information Technique CAF)

**Abstract** This paper discusses the algorism for polygon overlay. The result shows: The problem of complicated topology relation after coverage overlay and difficulty of feature attribute refilling can easily be resolved in the way of merging two map into a new one, regenerating the topology relation of spatial data, regenerating polygons, regenerating label points and refilling feature attribute with label points. The coverage overlay can be implemented successfully.

**Key words** geographic information system polygon overlay polygon generation label point judging feature attribute filling