ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.19, No.11, November 2008, pp.3053−3060
DOI: 10.3724/SP.J.1001.2008.03053

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

计算线段集合的相交直线及其最大存在范围*

汪嘉业[1], 杨承磊[2+], 张彩明[2]

[1](山东经济学院,山东 济南 250100)

[2](山东大学 计算机科学与技术学院,山东 济南 250100)

Computing Lines Intersecting with Set of Line Segments and the Maximum Range

WANG Jia-Ye[1], YANG Cheng-Lei[2+], ZHANG Cai-Ming[2]

[1](Shandong Economic University, Ji'nan 250100, China)

[2](School of Computer Science and Technology, Shandong University, Ji'nan 250100, China)

+ Corresponding author: E-mail: chl_yang@sdu.edu.cn

Wang JY, Yang CL, Zhang CM. Computing lines intersecting with set of line segments and the maximum range. *Journal of Software*, 2008,19(11):3053−3060. http://www.jos.org.cn/1000-9825/19/3053.htm

Abstract: For a given set *S* of line segments, finding a straight line intersecting with all the line segments in *S* is studied in this paper. If an intersection restriction is satisfied by the set, the algorithm presented is to answer whether there is a straight line intersecting with all the line segments in *S*. If the straight lines exist, the algorithm finds a maximum range, where every straight line located in the range intersects with all the line segments in *S*. The time complexity of the algorithm is $O(n\times\log n)$. The algorithm can be used in pattern marching and so on.
Key words: line segment set; query intersection; convex hull; intersecting straight line; arrangement of line segments

摘 要: 对给定的一个直线段集合 *S*,研究求与 *S* 中所有直线段都相交的直线的问题.设 *S* 中的线段满足一定的不交性假设,算法可回答是否存在与 *S* 中所有线段均相交的直线的问题.如果该直线存在,则求出这样的直线的最大存在范围——位于该范围内的每条直线都与 *S* 中的所有直线段相交.该算法的时间复杂性为 $O(n*\log n)$,应用背景是模式匹配等领域.
关键词: 线段集合;求交查询;凸包;相交直线;线段排列
中图法分类号: TP391 **文献标识码**: A

在计算几何中关于线段簇及其求交问题的讨论有很多.线段的整理(arrangement of line segments)[1]就是讨论线段簇之间的交点、由此衍生的线段以及由线段相交生成的多边形的有关问题.线段的整理在可见图、隐藏面消除等方面有很多应用[2].由于它广泛的应用价值,使其成为计算几何中的重要结构之一.

在计算几何中,多边形求交也是一个重要的问题.此外,涉及相交的还有检索问题.线段相交检索的一个典型的问题是:给定平面上一个直线段集合 *S*,对给定的询问直线段(query segment),需要报告 *S* 中有多少条直线段与询问直线段相交.这方面的结果可参阅文献[3,4].

* Supported by the National Natural Science Foundation of China under Grant Nos.60573181, 60703028 (国家自然科学基金)
Received 2007-10-09; Accepted 2008-01-29

在模式匹配中经常会遇到平面上两个对象集合要通过空间排序方式建立两个对象集合的对应关系的问题.这两个对象集合往往可以用两个不相交的直线段集合 S_1,S_2 来代替.这时在平面上需要有一个共同的方向,以便沿这个方向对两组直线段分别作排序,然后按次序确定两组对象的对应关系.如果存在指向同一方向的两条直线 l_1,l_2,它们分别与一个集合内的所有直线段均相交,则这个排序方向的一个合理选择就是这两条直线 l_1,l_2 的共同指向(如图 1 所示).

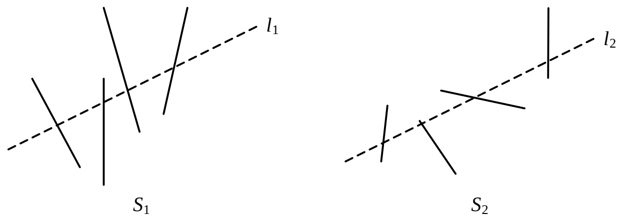

Fig.1 Two line segment sets intersect with two lines that have the same direction
图 1 两个直线段集合分别和两同方向直线相交

下文把与一个直线段集合 S 中的所有直线段均相交的直线 l 称为 S 的相交直线.

为了找到这两条直线 l_1,l_2 的共同指向,可对每一个直线段集合求其相交直线存在区域.如果这两个相交直线的存在区域在方向上存在非空的交,则该交中任一方向即可取为排序方向.

本文假定集合 S 中的线段是不相交的,而且要求集合 S 中任意两线段所在直线的交点也不在线段内.对上述直线段集合 S 我们给出一种算法,以判断是否存在与 S 内所有直线段相交的直线,如果存在,则求出该直线最大允许存在区域.该问题的一个比较显然的算法时间复杂性为 $O(n^3)$,本文给出的算法时间复杂性为 $O(n\times\log n)$.

本文第 1 节中给出算法的说明,第 2 节作算法复杂性分析,最后是总结.

1 算法描述

如果直线段集合 S 存在相交直线,则一定存在这样的相交直线,它和 S 中的两条或多条直线段的交点是这两条或多条直线段的端点,而且这两条线段位于该相交直线的两侧,如图 2 中的两条虚线 AB 和 CD.其中,AA'和 BB'与 AB 相交在端点 A 和 B,AA'和 BB'位于 AB 的两侧,且 AB 和 S 中每一线段均相交.这样的相交直线称为极端相交直线.不难证明,只要存在相交直线,就一定存在极端相交直线,因为我们可以在保持相交直线和 S 的所有直线段都相交的条件下,通过移动和旋转相交直线而找到极端相交直线.由于极端相交直线也是相交直线,所以相交直线存在的充要条件是存在极端相交直线.

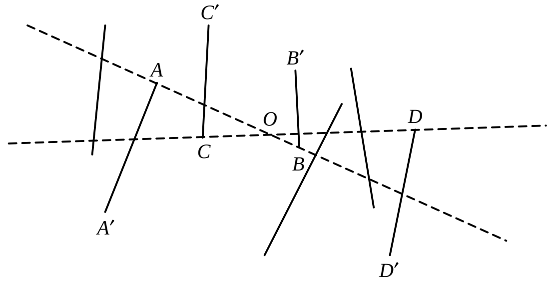

Fig.2 Special intersecting lines of a line segment set (dashed lines)
图 2 线段集合的极端相交直线(虚线)

因此,一个比较显然的算法是用穷举法去求极端相交直线:对 S 中线段的每一顶点和 S 中另外所有线段的每个顶点连线,判断该直线是否和 S 中所有线段相交,如果答案是肯定的,那么 S 存在相交直线,否则 S 不存在相交直线.设 S 中有 n 条直线段,极端相交直线的候选直线可多达 $(2n)\times(2n-2)/2$, 每一候选直线要去测试是否与

另 n–2 条直线段相交,时间复杂性为 $O(n^3)$.这仅仅是确定是否存在相交直线和求出所有极端相交直线,并未包含求出所有相交直线存在的范围.

在下文的讨论中假定集合 S 中的线段是不相交的,而且要求集合 S 中任意两线段所在直线的交点也不在两线段内.我们称其为不交性假设.

现在我们提出一种算法以确定是否存在相交直线,如果存在相交直线,则求出相交直线的最大允许存在区域.下面以图 3 中所有实线段的集合 S 为例来说明算法的思想.

算法首先求出所有顶点的凸包,如图 3 中多边形 $E'A'F'D'DKCE$,然后按凸包上顶点的次序把过该顶点的直线段的另一顶点依次连接,这就得到图 4 中的多边形 $EAFDD'K'C'E'$.我们称其为该凸包的附属多边形.

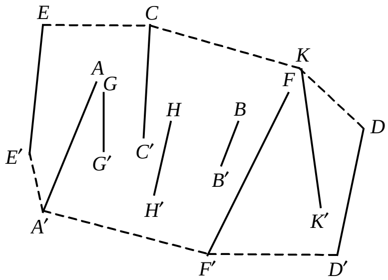

Fig.3 Compute convex hull of all the vertices
图 3 计算所有顶点的凸包

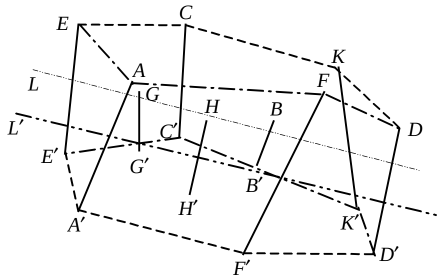

Fig.4 Accessorial polygon of the convex hull
图 4 凸包的附属多边形

设过凸包顶点的所有 S 的直线段集合为 S'(它是 S 的子集).若直线段集合 S'存在相交直线,则在不交性假设下,附属多边形有以下两个性质:

(1) S'中与相交直线相交的交点在最外侧的两条直线段(如图 4 中 DD'和 EE')一定在 S'的线段顶点的凸包上,它们也是附属多边形的两条边.

对此可作如下说明:如果 E 不在凸包上,则 EE'的延长线要交在线段 AA'或 CC'的内部,这与不交性假设相悖.

(2) 附属多边形在相交直线两侧的两条链是不自交的.

如图 4 中 $EAFD$ 和 $D'K'C'E'$是不自交的.否则,像图 5 那样,若 $ABCD$ 自交,则 C 要在 AA'的右侧,由于 CC'和 AA'还要与相交直线 L 相交,这时 CC'的延长线要交在线段 AA'内部,不符合不交性假设.

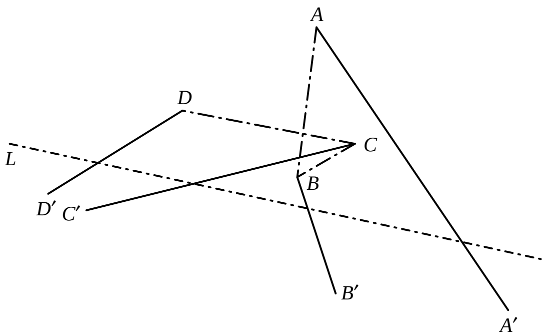

Fig.5 The case that $ABCD$ is self-intersection
图 5 $ABCD$ 自交情况

引理 1. 直线段集合 S'存在相交直线的必要条件为其凸包的附属多边形是简单多边形,且其走向与凸包走向相反.

证明:由于上述性质 2,附属多边形在相交直线两侧的两条链是不自交的(如图 4 中 $EAFD$ 和 $D'K'C'E'$).它们又被相交直线 L 隔开,所以它们不会相交,而这两条链被凸包上的边 DD'和 EE'连接起来,DD'和 EE'不会与凸包

内的两条链相交,因而附属多边形是不自交的,它是简单多边形.凸包被相交直线分成连续的两部分,如图 4 中凸包 $E'A'F'D'DKCE$ 被相交直线 L 分成 $E'A'F'D'$ 和 $DKCE$ 两部分.因相交直线与 S' 的所有线段相交,对 S' 中的直线段来说,它们的两个端点分别处于 L 的异侧,如 E,A,F,D 和 E',A',F',D' 在 L 的异侧,同样,D',K',C',E' 和 D,K,C,E 也处于 l 的异侧.这样,从 L 上的附属多边形内任意一点为中心来看,$E'A'F'D'DKCE$ 与 $EAFDD'K'C'E'$ 的走向正好相反.前者为逆时针方向,后者为顺时针方向.因此,附属多边形与凸包走向相反,证毕. □

由引理 1 可知,如果凸包的附属多边形与凸包走向一致,或者虽然走向不一致,但自交,则 S' 以至 S 均不会有相交直线.

但引理 1 的条件并不是线段集合 S' 存在相交直线的充分条件,也就是说,满足引理 1,线段集合 S' 未必一定有相交直线.要确定 S' 是否存在相交直线,我们必须对附属多边形进行计算和测试,测试 S' 是否存在相交直线,如果存在,则求出其相交直线的存在区域.

由凸包附属多边形的构造可知,与 S' 中全部直线段相交的直线 L 只能穿过附属多边形.

引理 2. 设直线段集合 S' 存在相交直线.在相交直线和附属多边形边界的交点中,把相距最远的两个交点分别称为进点和出点,相交直线上位于该两点之间的部分均属于该附属多边形的闭集.

证明:若进点和出点之间存在一点不属于附属多边形,则相交直线上有一段线段在附属多边形的外部,而其相邻两段则在附属多边形的内部,如图 4 中直线 L'.L' 上在附属边形外的线段的两个端点,是 L' 和附属多边形边界的两条不同边的交点,在这两个交点之间存在附属多边形的顶点(如图 4 中点 C'),显然,通过这些顶点上的 S' 的边(如图 4 中点 CC')与直线 L' 是不相交的.这与 L' 是相交直线相矛盾,引理得证. □

引理 2 并未排斥在进、出两点之间,相交直线与附属多边形边界存在某些交点,但这些交点只能是附属多边形的向内凹的顶点.

由上述性质 1 可知,S' 中有两条直线段(如图 4 中 DD' 和 EE')一定是附属多边形的两条边.在附属多边形上也只能存在两条 S' 的边.因为从附属多边形的构造方法可知,附属多边形的边中不可能存在 S' 中第 3 条直线段也是它的边.相交直线也必须穿过上述两条边才能与 S' 中所有直线段相交.反之,若有一条直线穿过这两条边且完全在附属多边形上通过,则它一定与 S' 的所有直线段相交,它是一条相交直线.综上所述,我们得到定理 1.

定理 1. 若直线段集合 S' 满足不交性假设且存在相交直线,其凸包的附属多边形是走向与凸包走向相反的简单多边形,在该附属多边形上存在且仅存在两条边是 S' 的边.一条直线是 S' 的相交直线的充要条件是它穿过这两条边,且直线的这两点之间的部分完全在附属多边形上.

由定理 1 可知,根据附属多边形上 S' 的两条边(如图 4 中 DD' 和 EE'),可把附属多边形 $EAFDD'K'C'E'$ 的边界分成两条链,即 $EAFD$ 和 $D'K'C'E'$.当且仅当在两条链 $EAFD$ 和 $D'K'C'E'$ 中间穿越的直线才是相交直线.为了找到该相交直线可能存在的区域,我们对两条链 $EAFD$ 和 $D'K'C'E'$ 分别求出其向中间凸的凸包折线,我们称其为附属凸链.为了形成附属凸链,只要依次检查两条链是否有向中间凹的点,如果有,则把凹点去掉.如链 $EAFD$ 上 F 点是向两条链中间凹的,去掉 F 点后,$EAFD$ 就变成一条附属凸链 EAD(如图 6 所示).这里求附属凸链与 GRAHAM 求平面点集凸包算法采用的方法[1]是一样的(在求多边形凸包时,也是把边界链变凸的过程,当边界点走进由多边形边界形成的凹的口袋时,必须作适当处理[5].不难证明,在不交性假设下,这两条链不会出现文献[5]中的情况).

如果求得的两条附属凸链相交,则由引理 1 可知,S 和 S' 都不存在相交直线,算法结束.当两条附属凸链不相交时,我们求这两条附属凸链 EAD 和 $D'K'C'E'$ 的公切线 MN' 和 $M'N$.求公切线的方法参见文献[1]中求凸包的分而治之算法.易知,存在与 S' 中的线段都交的相交直线段 L,当且仅当 L 的两个端点必须分别在多边形 $MPP'M'$ 和 $NDTN'$ 内,相交直线段也必须完全属于多边形 $MPP'M'$,$NDTN'$ 和 $PADTC'P'$(如图 7 所示).这也是 S' 的相交直线的最大存在区域.

也可要求相交直线段的两个端点分别落在直线段 PP' 和 DT 上,同时,整条相交直线段均属于多边形 $PADTC'P'$(如图 8 所示).

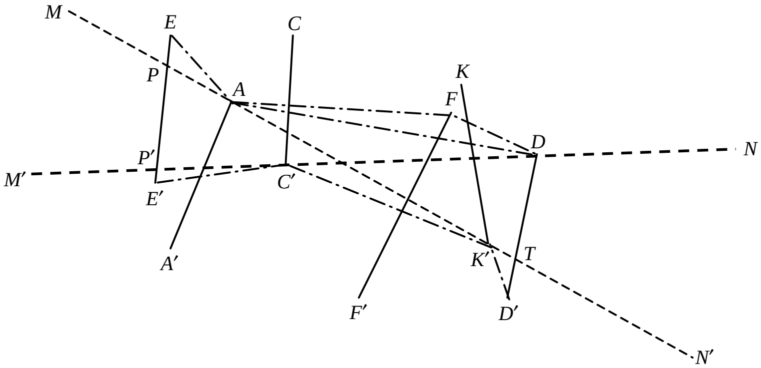

Fig.6　Two special convex polygon lines *EAD* and *D′K′C′E′*, and their tangent lines *MM′* and *NN′*

图 6　两条附属凸链 *EAD* 和 *D′K′C′E′*,及公切线 *MM′*和 *NN′*

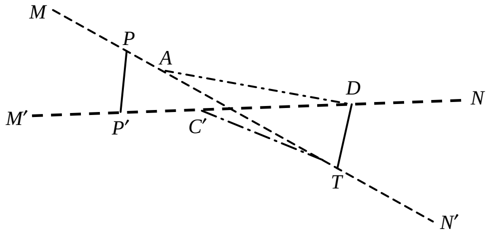

Fig.7　Compute the maximum range of intersecting lines of *S′*

图 7　求 *S′*的相交直线最大存在区域

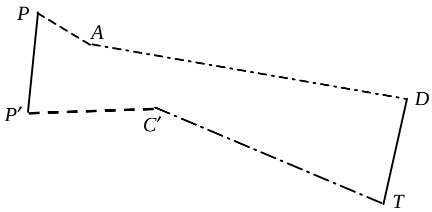

Fig.8　The maximum range of intersecting lines of *S′*

图 8　*S′*的相交直线最大存在区域

上面求得了线段集 *S′*的相交直线的存在区域,*S′*仅是 *S* 的子集.下面接着去求线段集 *S–S′*(如图 3 中的{*GG′*,*HH′*,*BB′*})的相交直线.

同样地,首先求 *S–S′*线段顶点的凸包 *G′H′B′BG*(如图 9 所示),其附属多边形为 *GHBB′G′*,得到两条链 *GHB* 和 *B′G′*,它们本身已是向两链中间凸的附属凸链.由此可求得公切线 *UQ′*和 *QU′*.可知集合{*GG′*,*HH′*,*BB′*}的相交直线的两个端点应在多边形 *UGG′U′*和 *BQQ′B′*内,相交直线应完全属于多边形 *UGG′U′*,*BQQ′B′*和 *GHBB′G′*(如图 10 和图 11(a)所示).

如果还有 *S* 中的线段端点未包含在两次凸包的顶点内,则上述求凸包和公切线的过程还要进行下去.如果 *S–S′*中只剩 1 条线段的情况(如图 11(b)中的 *BB′*),则得到相交直线的最大存在区域,如图 11(b)所示.相交直线的两个端点分别在 *BB′*的两侧半平面内,而且与线段 *BB′*相交.

在图 3 所示的例子中,由于第 2 次凸包顶点涉及的线段包含了 *S–S′*所有的线段,因此 *S* 的所有线段被分成了两组,*S′*和 *S–S′*,分别得到了 *S′*和 *S–S′*相交直线最大存在区域,如图 8 和图 11(a)所示.这两个相交直线最大存在区域的交便是 *S* 的相交直线的最大存在区域(这个最大存在区域必须是连通的.如果不连通,则不存在相交直线),如图 12 和图 13 所示.这里,要计算的相交直线的两个端点分别落在直线段 *P″P′*和 *DT* 上,同时,相交直线段

的全部点均属于多边形 $P''GHBDTC'P'$(如图 13 所示).在得到图 13 时,我们把 $P''P'$和 DT 中间的两条链又作了凸化处理.这样做的目的除了便于计算公切线以外,还在于随后求两个相交直线最大存在区域的交时,只要对凸链求交,两条凸链求交可在线性时间内完成.对链作凸化处理的时间复杂性也是线性的.

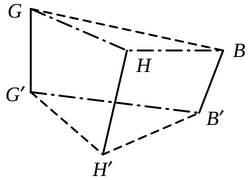

Fig.9 Convex hull and special convex polygon lines of the vertices of S–S'

图 9 S–S'顶点的凸包及附属凸链

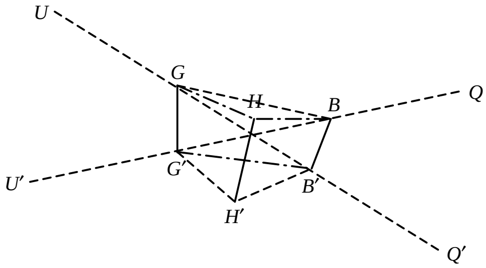

Fig.10 Tangent lines of two special convex polygon lines

图 10 两凸链的公切线

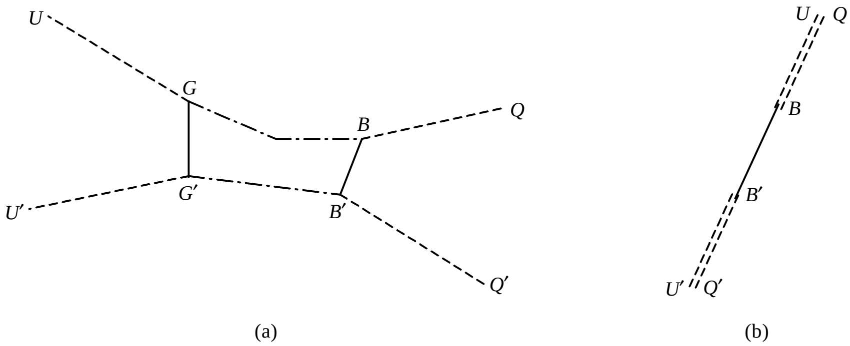

Fig.11 The maximum range of intersecting lines of S–S'

图 11 S–S'的相交直线最大存在区域

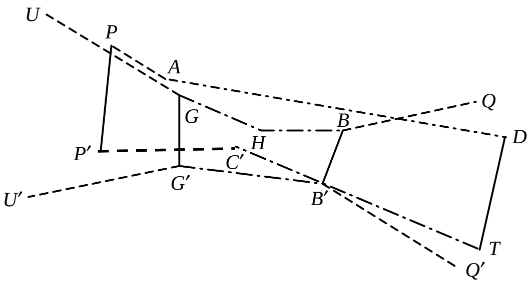

Fig.12 Intersection of two maximum ranges of intersecting lines

图 12 两个相交直线最大存在区域求交

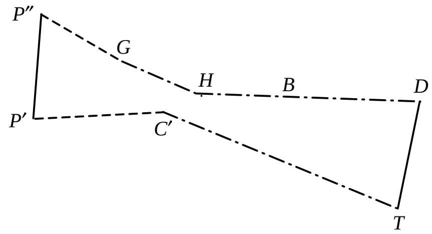

Fig.13 The maximum range of intersecting lines of S

图 13 S 的相交直线最大存在区域

由上述算法不难看出,如果 S 中线段的顶点均在一个凸包上,那么只需要 1 次计算凸包就可算出 S 相交直线最大存在区域.这是算法效率最高的情况.另一种极端的可能情况是每次求凸包都只有少数线段顶点(如两条线段)在凸包上,这样,最后可能有 $O(n)$个相交直线存在区域要求交,其中 n 为 S 中直线的段数.相交直线存在区域求交的算法对整个算法的时间复杂性影响很大.为了降低时间复杂性,先对这 $O(n)$个相交直线存在区域两两求交,得到 $O(n/2)$个交的区域,再对这 $O(n/2)$个交的区域两两求交,如此重复 $O(\log n)$步,即可求得所有相交直线存在区域的交.两个区域求交也就是区域的两条凸链求交,求得的结果不一定是凸链,但在下次求交之前我们还要对它们作变成向中间凸的处理.下一节将证明本文提出的上述算法总的时间复杂性为 $O(n\times\log n)$.

上述算法步骤大致描述如下:

算法. 用凸包法计算直线段集合 S 的相交直线及最大存在范围.

S:输入的直线段集合.

R:S 的相交直线允许存在的最大范围.

i:当前处理步骤.初始值为 1.

1) 计算 S 的凸包.设过凸包顶点的所有直线段集合为 S_i.
2) 计算凸包的附属多边形 P_i.
3) 根据处于 P_i 边界上的两条 S_i 的线段,计算凸包的两个附属凸链.
4) 如果两个附属凸链相交或 P_i 与 S 的凸包方向一致,则可以确定肯定不存在 S 和 S_i 的相交直线;算法结束.
5) 否则,计算 S_i 的相交直线允许存在的最大范围 R_i.
6) 令 $S=S-S_i$,如果 S 非空,则转 1),进行类似的处理;否则,转 7).
7) 用分而治之方法对集合$\{R_i\}$求交集 R.

在求交过程中对每次求交结果的两条链作向内凸化的处理,一旦发现两集合的交为空,或两集合的交不连通,或两条凸链相交,则可以确定不存在原输入的直线段集合 S 的相交直线,算法结束;否则,存在其相交直线,且该交集 R 为其相交直线允许存在的最大范围.

我们提出的第 2 种算法在想法上更简单,即代替求 S 的凸包,而是把 S 分成 $n/2$ 个组,每组有两条线段.如果 n 为奇数,则可以把余下的一条线段作为一组.对每一组求出与组内线段相交直线的存在区域,最后求所有存在区域的交,也就是 S 的相交直线的最大存在区域.下一节将说明这种算法的时间复杂性也是 $O(n\times\log n)$.其缺点是要对 $O(n)$个相交直线存在区域计算交集,并要在最后求交过程中才能确定是否存在相交直线.

2 算法复杂性分析

第 1 种算法要对 S 中所有线段凸包一层一层剥离,并每次求得内部的凸包.文献[6]可在 $O(n\times\log n)$时间内完成 n 个点的所有层的凸包.设计算过程共求了 m 次凸包才把 S 的所有线段都涉及到,集合 S 也相应地被分成集合$\{S_i\}$,$i=1,2,\ldots,m$,设 S_i 有 n_i 条线段,其顶点的凸包有 $O(n_i)$个顶点,则计算和集合 S_i 相关凸包的附属多边形及凸链可在 $O(n_i)$时间内完成.由于所有集合 S_i 凸包上的顶点数的总和为 $O(n)$,所以求所有集合 S_i 的相交直线存在区域及其凸链过程的时间总和为 $O(n)$.

对所有集合 S_i 的相交直线存在区域的求交,每一步我们采用两两求交的办法,共进行 $O(\log n)$步.由于对原

始的及中间结果的相交直线存在区域所得的两条链作了凸化处理,也即相交直线存在区域是由两条凸的链组成,两条凸链的求交过程与两个凸多边形求交过程类似,可在线性时间内完成.由于每一步所有存在区域的顶点数总和都是 $O(n)$,因此每一步所有存在区域两两求交可在 $O(n)$时间内完成,总共 $O(\log n)$步,因此总的求交时间为 $O(n\times\log n)$.

综上所述,计算 S 的相交直线的最大存在区域算法的时间复杂性为 $O(n\times\log n)$.

由于第 1 种算法包含了第 2 种算法的各个步骤,因而第 2 种算法的算法复杂性也为 $O(n\times\log n)$.

3 结 论

本文提出了线段集合的相交直线最大存在区域算法,算法复杂性为 $O(n\times\log n)$.第 1 种方法的优点是有可能较早发现 S 不存在相交直线的情况.如果存在相交直线,则也有可能减少计算多边形求交的次数,从而减少计算量.第 2 种算法从概念上看十分简单,但计算量可能更大.

该问题是否可以在线性时间内完成以及给出其算法还是一个值得讨论的问题.如果不作不交性假设,则相交直线存在区域可能不止一个.可举出一个有 3 条直线段的集合存在 3 个相交直线存在区域的例子.对不满足不交性假设情况的算法进行研究也是有意义的.

References:

[1] O'Rourke J. Computational Geometry in C. 2nd ed., New York: Cambridge University Press, 2005. 63−96, 193−218.

[2] McKenna M. Worst-Case optimal hidden-surface removal. ACM Trans. on Graphics, 1987,6(1):19−28.

[3] Goodman JE, O'Rourke J. Handbook of Discrete and Computational Geometry. New York: Chapman & Hall/CRC, 2004. 857−876.

[4] Mortensen CW. Fully-Dynamic two dimensional orthogonal range and line segment intersection reporting in logarithmic time. In: Littleton R, ed. Proc. of the 14th Annual ACM-SIAM Symp. on Discrete Algorithms. Edmonton: ACM Press, 2003. 618−627.

[5] Wang JY, Liu DY. A new linear algorithm for finding convex hull of 2D simple polygon. Chinese Journal of Computers, 1989,12(1): 38−43 (in Chinese with English abstract).

[6] Preparrata FP, Shamos MI. Computational Geometry: An Introduction. New York: Springer-Verlag, 1985. 165−168.

附中文参考文献:

[5] 汪嘉业,刘鼎元.对平面简单多边形求凸包的线性时间算法.计算机学报,1989,12(1):38−43.

汪嘉业(1937－),男,上海人,教授,博士生导师,主要研究领域为计算机图形学,计算几何.

张彩明(1955－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为几何与可视计算.

杨承磊(1972－),男,博士,副教授,CCF 高级会员,主要研究领域为虚拟现实,计算几何,计算机图形学.