

一种新的图件扫描影像细化算法

许捍卫 潘松庆

(河海大学土木工程学院 南京 210098)

摘要 针对传统的细化算法如经典细化算法、跟踪剥皮算法等在细化时需花费大量的时间这一问题,提出了一种基于最短距离边缘跟踪技术的细化算法,此算法通过对线宽两边相应边缘点中点的求取来细化并同时矢量化,能有效地提高细化速度,此算法已在作者自编的 DCMS 系统中得到应用。

关键词 细化;边缘跟踪;屏幕数字化;图件扫描;地理信息系统

当今地理信息系统(GIS)市场迅速扩展,而在建立 GIS 时花费最多的工作就是数字数据的获取。GIS 中数字数据的来源主要是对地图图纸图像的数字化。为了充分利用图纸图像来快速建立 GIS,首先需建立有关扫描图件矢量化与识别的商品化软件系统。近几年来,市场上已出现了相当数量的此类软件,且各具特色,亦能基本满足生产需要,但其自动化程度偏低,运行速度偏慢,离实用还有一定的距离。

扫描图像首先需进行图形与符号、文字注记分离,分离后的图形细化矢量化建库是建立 GIS 库或建立数字化地图的核心内容之一,对于数字制图意义更大。

细化指的是在保持拓扑性质不变的同时提取线状影像结构的骨架,并且细化后的线宽严格保持为一个像元。细化目的是除去大量冗余的数据,为下一步的栅格数据矢量化准备必要的条件。

关于细化算法的研究,如用 4 个匹配模板进行边缘点去除的并行细化算法和基于并行模板的二值图串行细化算法^[1]等,其主要处理过程可分为串行和并行两种,串行处理是逐步去除目标边缘上的点,并保留形成图形骨架的最终组成点。在这类方法中,每次迭代只能削去一个目标层,因此目标的细化很费时间,如经典细化算法、跟踪剥皮算法等。对于并行处理,图中各个像元的新值是其诸相邻像元的局部函数,而所有像元都被独立地同时处理,能有效地提高细化速度。本文提出了一种能有效提高细化速度并矢量化的算法,即基于最短距离的边缘跟踪技术,它属于并行算法。

1 算法的基本内容

对于线型图形来说,线宽应该基本固定,此线宽

值即为最短距离 $d^{[2]}$ 。

图符分离后的图形文件,其实体由线和多边形组成,而多边形则可看成起点与终点重合的线,是线的特例,如图 1 和图 2 所示。



图 1 单一的目标线段

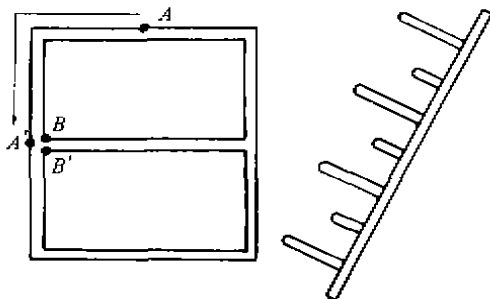


图 2 多边形及丁字型实体

在图 1 中,首先检索到的是 A 点,找到 B 点,AB 线的中点即为细化后的保留点,A 点沿着边缘点往一方向移动一步距到 A',B 点沿着边缘点逐点移动,称之为 B',可求得 A',B' 两点之间的距离 $S_{A'B'}$,B' 一直移动,直至为最小,此时的 A'B' 的中点即为细化后的保留点。同理,可求得这一方向上到线段终点的所有细化后的保留点。此方向处理结束后,从 A 点沿着另一方向同样处理,整条线段细化即告完成。

对线段来说,算法简单,容易实现。但实际上,图形中往往含有一定的丁字形实体,如图 2 所示。A 点按箭头方向移动到 A' 点时,可求得 B,当 A' 继续移动,B 却不能前行,此时 $S_{A'B'}$ 就不断增大,当 $S_{A'B'} - d$ 大于限值(如 3 个像素)时,即可判断 B 点为一分

支,把此分支压入堆栈, A' 继续前行并重复开始工作,直至线段终点或回到 A 点。

把压入堆栈的数据弹出,重复上面的工作,直至堆栈为空。

2 边缘跟踪

2.1 边缘跟踪方法

在最先搜寻到线划影像边缘时,此点即为线段一边的一个边缘点,当 A 点不是线段终点时, A 点有两个边缘跟踪方向。在下面算法中,我们采取第一个边缘方向跟踪到终点以后返回,然后还是从 A 点出发跟踪第二个边缘方向,直至整条线段跟踪结束。

如图 3 所示,从 0 方向开始的方向顺序,是我们采用的方向码。当边缘跟踪时,按一定顺序(例如顺时针)检测其八邻域的灰度。在第一边缘方向跟踪时检测出八邻域中某一像元灰度为 1,而其后像元灰度为 0 时,就是我们所需要的下一边缘点。当反方向边缘跟踪时,其检测出八邻域的条件变为此像元灰度为 0,而其后像元灰度为 1 时,反方向的下一边缘点也就找到了。

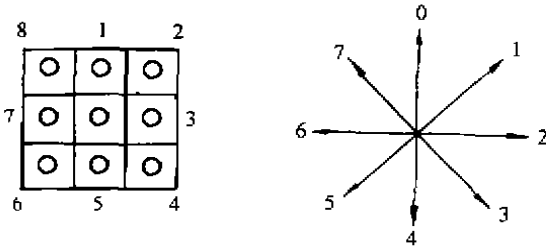


图 3 像素点的八邻域位置和方向

2.2 最短距离 d 和线宽另一边边缘点 B 的确定

继续上面的第一个边缘方向跟踪,直至走过一步距(2或3个边缘点),求出方向如图 4 所示的 f_1 ,逆时针转 90° 得方向 f_2 ,从 A 点出发,沿方向 f_2 前进,每次前进 1 个像元,当再往前走,其像元灰度值为 0 时,找到了线宽另一边边缘点 B 。此时的 B 点不一定对应于最短距离 d ,还需要对 B 点两个边缘方向进行追踪,追踪到一个边缘点,就可求出 AB 之间的距离,当找到最短距离 d 时结束 B 点边缘追踪。此时的 d 即为细化前的线宽,对应于 d 的边缘点 B 为线宽另一边边缘点追踪的起点。

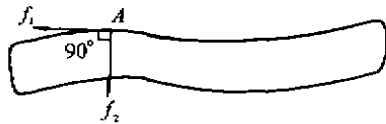


图 4 线宽另一边边缘点 B 的求解

3 全自动细化的算法实现

3.1 基本思路

首先,从图形文件每行每列中读出像素,如为 0,则继续前进,如为 1,则为第一个检索到的边缘点,据此边缘点就可对一实体进行处理,处理的同时,对此实体做上标记。此实体处理完后,又开始检索边缘点,对有标记的实体跳过检索,直至整个图形文件都检索完后结束。

3.2 算法实现

a. 找到第一个边缘点,判断第一个边缘点是否刚好为线段终点,如为线段终点,则只能往一个方向前进一步距,求出细化前的线宽 d 。如不为线段终点,则直接求出线宽 d ,并求出中点 P_x, P_y 。

b. 沿着边缘点前进一步距,求出此时的最短距离 $S_{\min}$,如果 $S_{\min} - d$ 小于限值,那么求出此时的中点 P_x, P_y 。否则把对应于 $S_{\min}$ 的另一边缘点压入堆栈,沿着边缘点前进一步距,求出另一最近边缘点,把此点也压入堆栈。

c. 重复步骤 b,直至结束。

d. 如第一个边缘点不为线段终点,则另一分支重复步骤 b,直至结束。

e. 弹出堆栈数据两个,对第二个求出其另一边边缘点,与第一点进行相邻性匹配,如相邻,转步骤 b,如不相邻,暂存第一点,转步骤 b,完成后,以第一点转步骤 b。

f. 当堆栈为空时,一实体结束。

g. 对此实体做上标记,转步骤 a,直至找不到实体边缘点。

3.3 堆栈结构

在上面的步骤 b 中,有一堆栈结构,以记录断点位置。断点堆栈结构如表 1 所示,其压入的数据除断点位置外,还有边缘跟踪方向值 True 或 False,以决定堆栈数据弹出转步骤 b 时执行哪一个边缘方向跟踪。

表 1 断点堆栈结构

X 坐标	Y 坐标	边缘方向
X_{i+1}	Y_{i+1}	True or False
X_i	Y_i	True or False
$\vdots$	$\vdots$	$\vdots$
X_1	Y_1	True or False

在断点压入堆栈时,连续压入一对断点,如图 2 所示的 B, B' 两点。这样,弹出时也连续弹出两个点,如步骤 e。其中步骤 e 的相邻性匹配的目的是决定在此点附近的局部实体形状是丁字形还是十字形,相邻为丁字形,不相邻则为十字形。

4 屏幕细化算法实现

屏幕数字化细化算法与全自动细化算法基本一致,不过比全自动细化算法简单,不存在对已细化后的栅格图像又一次被细化的问题,不同的仅是对边缘点的拾取问题,在程序中可以用循环语句对当前鼠标所在点上下左右一定范围内进行搜索,直至找到边缘点为止。

5 矢量化

矢量化既可在细化的同时进行,也可在全部细化完后进行。经过细化后保留下来的中心点,是一些离散的点,其离散程度依据所给的步距而定。由于是一些离散的点,这一定程度上又加速了矢量化数据压缩的过程,至于矢量化数据的压缩方法,可采用如垂距法、偏角法、道格拉斯-普克(Douglas-Peucker)法^[3]等等,也可采用文献^[4]中的“四点法”。矢量化后,还需根据需要决定是否进行结点匹配。

6 结 语

图纸图像为现阶段数字化地图和 GIS 的主要信息来源之一,其图形细化矢量化自有重要意义。本文提出的细化算法与传统的细化算法相比,除提高细化速度外,也使后续矢量化时的数据冗余量大为减少,另外本文算法在细化自动化、矢量化时结点求取上也具有一定的优越性。

参考文献

1 郑宏,袁宇正.基于并行模板的二值图串行细化算法.武测科技,1996(2):18~21
2 李新友,唐泽圣,张凤昌等.TH-DAIMS 3.0 清华图纸自动输入与管理系统的计算机辅助设计与图形学学报,1997(5):450~455
3 徐庆荣,杜道生,黄伟等.计算机地图制图原理.武汉:武汉测绘科技大学出版社,1993
4 许捍卫,周卫娟.一种新的等值线数据压缩算法.江苏测绘,1997(4):5~7

(收稿日期:1997-12-22 编辑:熊水斌)

·小资料·

世界上的大型抽水蓄能电站

自本世纪 60 年代以来,抽水蓄能电站在全球有了很大的发展。据统计,截至 1994 年,全世界已建成抽水蓄能电站 290 座,总装机达 8280 万 kW,其中装机在 100 万 kW 以上的有 44 座,可见抽水蓄能电站正日益向大型化发展。据笔者统计,目前装机在 150 万 kW 以上的已建及在建的抽水蓄能电站共有 19 座,见表 1。

表 1 大型抽水蓄能电站

序号	电站名称	国 名	装机容量/万 kW	台数/台	单机容量/万 kW	水头/m	开始发电年份
1	广州	中国	240	8	30	531	1992
2	德涅斯特	摩尔多瓦	226.8	7	32.4	155	在建中
3	巴斯康蒂	美国	210	6	35	390	1984
4	路丁顿	美国	205.8	6	34.3	108	1973
5	希望山	美国	204	6	34	810	在建中
6	丁诺维克	英国	180	6	30	541	1982
7	大屋	法国	180	12	15	955	1987
8	卡拉彦	菲律宾	180	12	15	290	1982
9	天荒坪	中国	180	6	30	667	1998
10	塔什别克	乌克兰	180	10	18	84	在建中
11	奥清津	日本	160	4 2	25 30	470	1978
12	明潭	中国台湾	160	6	26.7	400	1992
13	凯夏多尔	立陶宛	160	8	20	112	在建中
14	葛野川	日本	160	4	40	714	在建中
15	卡斯泰克	美国	156.6	6	26.1	328	1973
16	腊孔山	美国	153	4	38.3	310	1979
17	图穆特	澳大利亚	150	6	25	162	1973
18	奥美浓	日本	150	6	25	484	1994
19	萨米特	美国	150	6	25	682	在建中

在上述 19 座电站中,我国广州抽水蓄能电站名列世界首位,分两期建设,第一期工程为 4 台 30 万 kW,已于 1994 年建成,第二期工程仍为 4 台 30 万 kW,首台机组也于 1998 年投入。另外,天荒坪电站装机 180 万 kW,并列世界的第 6 位,第一和第二两台机组也在 1998 年投产运行。

(吴荣高 供稿)