

# 一种二值图像连通区域标记快速算法

曹长虎 李亚非

(南京工业大学电子与信息工程学院, 南京 211816)

**摘要** 由于处理时间和存储空间的限制, 常规的区域标记算法往往无法满足实时图像处理的需要。针对这些不足, 提出了一种新的二值图像区域标记快速算法。该算法通过结合轮廓跟踪技术, 经过一次图像扫描即可完成所有连通区域的标记, 避免了大多数改进算法都必须处理的标记冲突问题; 此外, 算法不受所标记的区域的形状和面积的影响, 能够准确标记任意形状的连通区域, 表现出良好的鲁棒性。最后与现存的多种算法进行了比较, 实验结果表明该算法是快速和高效的。

**关键词** 连通区域 区域标记 二值图像 轮廓跟踪 快速算法  
**中图分类号** TP391.41; **文献标志码** A

二值图像的连通区域标记是从仅由“0”像素(通常表示背景点)和“1”像素(通常表示前景点)组成的一幅点阵图像中, 将相互邻接(4-邻域或8-邻域)的“1”值像素集合提取出来。其目的就是要寻找图像中所有的目标对象, 并且将属于同一目标对象的所有像素用唯一的标记值进行标记。通过处理位于一个完整目标对象(也即具有相同整数标记值)内所有像素的特性, 从而获得被研究目标对象的整体性能和参数。这种预处理操作在图像处理和模式识别的许多应用领域中被采用, 如文本识别中的图文分割、工程图识别中的图形/标注符号分割、生物医学领域中的自动细胞分类计数、流水线上自动产品质量检测等。

## 1 算法回顾

### 1.1 基于像素点的区域标记算法

基于像素点的区域标记算法通常可以归纳为如下两类: 顺序法<sup>[1]</sup>和区域增长法<sup>[2]</sup>。

顺序法作为二值图像区域标记的传统算法需

要对图像进行两次扫描, 对于未作标记的像素标记值可以从它的邻域中获得, 对于邻域中出现多个不同标记的情况, 则通过构造一个标记冲突表来解决。这个冲突表会记录所有遇到的邻域中具有不同标记的情况。寻找冲突表中不同标记之间的对应关系需要运行一个类似于图遍历的过程。在第二次扫描图像的过程中, 使用冲突表中不同标记之间的对应关系更新标记矩阵中每个像素的标记。冲突表的长度受到连通区域形状的影响, 如果输入图像具有很多“U”型或“E”型连通区域, 该算法的性能会严重下降。此外, 该算法需要两次扫描图像, 增加了算法的运行时间。

区域增长法采用种子点填充的原理, 仅需对图像进行一次扫描, 可以分为递归法<sup>[3]</sup>、深度优先搜索法<sup>[4]</sup>和广度优先搜索法<sup>[4]</sup>。递归法需要消耗大量的堆栈空间, 仅适用于小面积区域的标记, 如果图像中含有大面积连通区域则会导致系统堆栈溢出, 算法运行失败。深度优先搜索法和广度优先搜索法都需要反复检查每个像素的8-邻域, 因此效率较低, 不适于应用在对实时性要求较高的场合。此外, 深度优先搜索法对内存空间的使用量也比较大。

### 1.2 基于线的区域标记算法

线标记法<sup>[5]</sup>将目标段(目标物体同一行中两个边界点像素集合, 包括边界点)作为连通体检测的

2010年8月27日收到

第一作者简介: 曹长虎(1986—), 男, 江苏省宿迁市人, 南京工业大学硕士研究生, 研究方向: 计算机应用技术。E-mail: caochang-hu0001@163.com。

基本单元,第一次扫描时记录冲突的标记并形成等价对,解决标记冲突后,在第二次扫描时用等价标记中最小的标号赋予所有等价标号对应的连通体。该算法与顺序法类似,唯一的不同是把连通体而不是像素点作为检测的基本单元。该算法性能优于顺序法,但仍需两次扫描图像,并解决标记冲突。

文献[6]提出了一种结合线标记法和区域增长法的改进算法。该算法与深度优先搜索法类似,不同点在于把连通体而不是像素点作为检测的基本单元。该算法仅需扫描一次图像,不会出现标记冲突的现象,性能优于普通线标记算法。但该算法受到深度优先搜索思想的制约,需要反复搜索每个连通体的邻域,降低了运行效率。当图像中含有很多点状或竖线状连通区域时,该算法会退化成深度优先搜索法。

## 2 本文算法

### 2.1 算法概述

本算法首先在待标记图像  $I$  边界上增加一圈像素进行扩展,新增的像素被设置为背景像素(如图1和图2所示),这样可以使本算法能够处理一些连通区域位于图像边界的特殊情况。

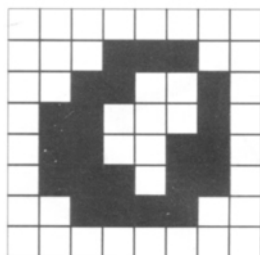


图1 扩展后的待标记图像  $I$

0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0
0	0	1	1	0	0	1	0
0	1	1	0	0	0	1	0
0	1	1	1	0	1	1	0
0	0	1	1	1	1	0	0
0	0	0	0	0	0	0	0

图2 与  $I$  对应的标记矩阵  $LM$

如同顺序标记法,采用8邻域连通,从上向下、从左向右扫描图像  $I$ ,对于图像中的每个像素点都要考虑如下四种情况。

情况 A: 由一个背景像素点向一个未标记前景像素点  $A$  移动,这意味着像素点  $A$  出现在一个新区域外侧的边缘上。于是分配一个新的标号给像素

点  $A$  (区域标号从 2 开始),然后调用轮廓跟踪过程  $\text{TraceContour}(A, \rho)$  对经过像素点  $A$  的区域外轮廓进行顺时针跟踪并标记为与像素点  $A$  相同的标号,所有与外轮廓邻接的背景像素都被标记为 -1 (如图3和图4所示)。

情况 B: 由一个已标记的前景像素  $B$  向未标记的前景像素移动,则把像素点  $B$  的标号传递给未标记的前景像素(如图5和图6所示)。

情况 C: 由一个前景像素点  $C$  向未标记的背景像素点移动,这意味着像素点  $C$  出现在一个区域的内轮廓的边缘上。于是调用轮廓跟踪过程  $\text{TraceContour}(C, 1)$  对经过像素点  $C$  的区域内轮廓进行逆时针跟踪并标记为与像素点  $C$  相同的标号,所有与内轮廓邻接的背景像素都被标记为 -1 (如图7和图8所示)。

其他情况: 搜索下一像素点。

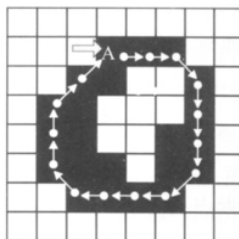


图3 情况 A

0	0	-1	-1	-1	-1	-1	0
0	-1	-1	2	2	2	-1	-1
-1	-1	2	1	0	0	2	-1
-1	2	1	0	0	0	2	-1
-1	2	1	0	0	1	2	-1
-1	-1	2	2	2	2	-1	-1
0	-1	-1	-1	-1	-1	-1	0

图4 情况 A 对应的标记矩阵

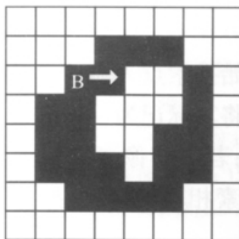


图5 情况 B

0	0	-1	-1	-1	-1	-1	0
0	-1	-1	2	2	2	-1	-1
-1	-1	2	2	0	0	2	-1
-1	2	1	0	0	0	2	-1
-1	2	1	0	0	1	2	-1
-1	2	1	1	0	1	2	-1
-1	-1	2	2	2	2	-1	-1
0	-1	-1	-1	-1	-1	-1	0

图6 情况 B 对应的标记矩阵

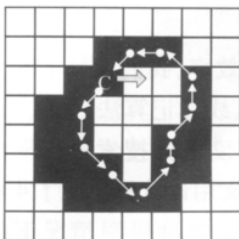


图7 情况 C

0	0	-1	-1	-1	-1	-1	0
0	-1	-1	2	2	2	-1	-1
-1	-1	2	2	-1	-1	2	-1
-1	2	2	-1	-1	2	2	-1
-1	2	1	2	-1	2	2	-1
-1	-1	2	2	2	2	-1	-1
0	-1	-1	-1	-1	-1	-1	0

图8 情况 C 对应的标记矩阵

表 1 二值图像连通区域标记算法运行效率的试验比较

图序号	连通区域个数	标记时间/ms						
		递归法	深度优先搜索法	广度优先搜索法	传统的顺序标记法	线标记法	文献[6]提出的改进算法	本文提出的算法
a	1 000	10.7	10.2	10.5	33.1	39.3	17.6	21.1
b	165	17.5	34.2	38.2	386.0	790.0	976.8	79.7
c	12	10.2	17.2	25.0	20.5	15.4	14.9	14.1
d	4	∞	333.0	862.5	282.0	106.3	107.2	87.9

2.2 算法流程

2.2.1 主算法

Labeling( *I* )

- (1) 扩展二值图像 *I*;
- (2) 创建图像对应的标记矩阵 *LM*;
- (3) 按从上向下从左向右的顺序扫描图像 ,根据可能出现的四种情况处理标记矩阵;
- (4) 返回标记矩阵。

2.2.2 轮廓跟踪算法

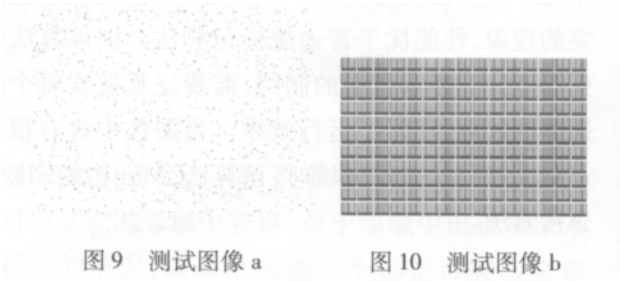
TraceContour( *X* , *D* )

- (1) 将 *X* 记为起始搜索点 *X<sub>s</sub>*;
- (2) 若 *D* 为 0 则设置搜索方向 *direction* 为 0° , 否则设置 *direction* 为 45°;
- (3) 沿着搜索方向搜索当前像素点的相邻像素;
- (4) 若相邻像素为起始搜索点 *X<sub>s</sub>* 则算法结束; 若相邻像素为背景像素则将其标记为 -1 ,搜索方向角增加 45° ,返回步骤 (3); 若相邻像素点为前景像素则将其标记为与当前像素相同的标记 ,搜索方向角减小 90° ,返回步骤 (3) 。

3 实验与结果分析

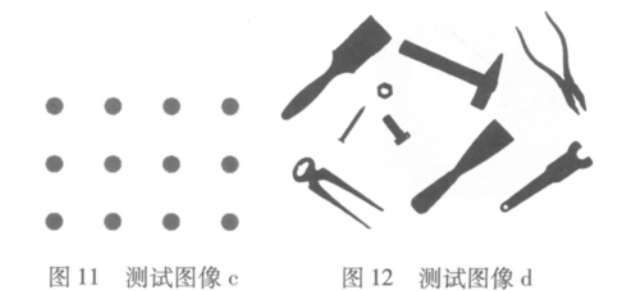
为了评价本文算法的效率 ,我们设计了本文算法与传统的顺序标记算法、线标记算法、文献 [6] 提出的改进算法、递归法、广度优先搜索法和深度优先搜索法的对比试验。本文用计算机生成了 4 种二值图像 ,用于测试各算法在各种极限情况下的效率 ,如图 9—图 12 所示。测试结果如表 1 所示。所有算法均运行在 P4/2.0 GHz/1.0 GB DDR2 的 PC

环境中。



结论:

(1) 图像 a 是尺寸为 400 × 300 的二值图像 ,含有 1 000 个单点连通区域。由于连通区域小 ,前三种采用区域增长技术的算法速度较快 ,传统的顺序标记法速度较慢 ,线标记算法没能发挥其优势 ,速度反而比顺序标记法慢。本文算法和文献 [6] 提出的改进算法速度适中。



(2) 图像 b 是尺寸为 400 × 300 的二值图像 ,含有 165 个形状极度不规则的“U”型连通区域 ,需要处理标号冲突的顺序标记法和线标记法效率严重下降 ,文献 [6] 提出的线标记改进算法虽然不用处理标记冲突但由于多数目标段都很短 ,又要进行邻域搜索 ,因此效率严重下降 ,但采用区域增长技术的递归法、深度优先搜索法和广度优先搜索法由于

其对连通区域形状的鲁棒性,效率基本不受影响。本文算法由于不受连通区域形状的影响,效率较高。

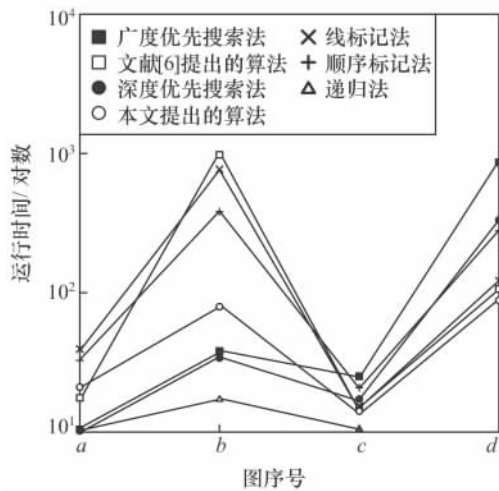


图 13 七种二值图像区域标记算法的效率比较

(3) 图像 c 是尺寸为  $400 \times 300$  的二值图像,含有 12 个  $30 \times 30$  像素的圆形连通区域。此时顺序标记法由于不会产生标记冲突,因此速度较快。线标记算法及其改进算法发挥了以连通体为扫描单位的优势,比传统的顺序标记算法效率高。本文算法的处理速度仅次于递归法。

(4) 图像 d 是尺寸为  $1\,212 \times 836$  的二值图像,含有 9 个面积较大的不规则连通体。递归法会导致系统堆栈溢出,因而算法运行失败。由于某些连通体面积较大,深度优先搜索和广度优先搜索法效率严重下降。顺序标记法效率较低,线标记法及其改进算法以及本文提出的算法效率都比较高。

综上所述,采用区域增长技术的递归法、深度优先搜索法、广度优先搜索法对连通区域的形状不敏感,但对连通区域的面积十分敏感,只能应用于小面积连通区域的标记问题。传统的顺序标记算法由于需要两次扫描图像,并且需要处理标记冲突,

因此效率较低,很难满足实时性应用的要求。线标记算法由于仍需处理标记冲突,因此对连通区域的形状敏感。文献[6]提出的线标记法的改进算法对连通区域的形状仍然敏感,但其性能比普通线标记算法有所提高。本文提出的算法对连通区域的面积和形状敏感度较低,具有很好的鲁棒性。从图 13 可以看出,本算法的平均效率远高于其他 6 种算法,因此适用于对实时性要求较高的应用场合。

## 4 结束语

本文通过结合轮廓跟踪和区域标记,提出了一种二值图像连通区域标记的新方法。本算法只需扫描一次图像便可以标记图像中所有连通区域,其运算效率不受连通区域形状和面积的影响,具有很好的鲁棒性,真正做到了快速、有效,同时也能满足实时性要求。本算法的输出是二值图像的标记矩阵,扫描一次标记矩阵便可以测量出各连通区域的面积、周长、重心、轴长、离心率、方向向量等特征参数,用于后续的更高层次的操作。

## 参 考 文 献

- 1 Rosenfeld A, Pfaltz J L. Sequential operations in digital picture processing. Journal of the ACM, 1966; 13(4): 471—494
- 2 陈柏生. 一种二值图像连通区域标记的新方法. 计算机工程与应用, 2006; 42(25): 50—51
- 3 喻杰, 许化溪. 一种易于实现的适于细胞图像连通区域的标记算法. 江苏大学学报, 2005; 15(2): 152—155
- 4 Burger W, Burger M J. Digital image processing. 北京: 清华大学出版社, 2010
- 5 张树生. 一种基于线的标号传播二值图像连通快速检测方法. 计算机研究与发展, 1994; 31(10): 51—54
- 6 高红波, 王卫星. 一种二值图像连通区域标记的新算法. 计算机应用, 2007; 27(11): 2776—2777

(下转第 8180 页)

## 参 考 文 献

- 1 杨 群. 支盘桩试桩极限承载力的二次趋势曲线预估法. 科学技术与工程, 2009; 9(14): 4230—4233
- 2 谢耀峰, 吴芳银, 王云球. 试桩未达极限状态时桩的极限承载力预测. 水运工程, 2002; (9): 5—8
- 3 赵明华, 胡志清. 预估试桩极限承载力的调整双曲线法. 建筑结构, 1995; (3): 47—52.
- 4 许小健, 钱德玲, 黄小平. 差异演化算法用于单桩承载力指数曲线模型优化. 岩土工程学报, 2009; 31(2): 265—270
- 5 张怡权, 夏柏如. 多分支承载力盘桩单桩竖向静载荷试验极限承载力的判定. 探矿工程, 2000; (5): 27—28
- 6 赵 岚, 王国体, 许小健. 微分进化算法在单桩极限承载力灰色优化预测中的应用. 合肥工业大学学报(自然科学版), 2009; 32(2): 245—248
- 7 钱德玲. 变截面桩与土的相互作用机理. 合肥: 合肥工业大学出版社, 2003; 58—61

## Comparison of Predicting Method for Ultimate Bearing Capacity of Single Squeezed Branch Pile

GONG En-de<sup>1</sup>, DENG Zu-bao<sup>1</sup>, XU Xiao-jian<sup>2</sup>

(Anhui Institute of Engineering Reconnaissance, Hefei 230011, P. R. China;

Anhui Geotechnical Engineering Co., Ltd.<sup>1</sup>, Hefei 230011, P. R. China;

Wuhu Geotechnical and Survey Design Institute<sup>2</sup>, Wuhu 241000, P. R. China)

**[Abstract]** To discuss the adaptability of curve fitting method for predicting ultimate bearing capacity of single squeezed branch pile, five kinds of prediction method, which are the hyperbolic method, adjusted hyperbolic method, quadratic trend forecasting method, exponential curve method and GM (1,1) prediction method, were applied to  $Q-s$  data-fitting calculations of practical engineering. The results show that the hyperbolic method may result in large deviations, however, the exponential curve method, GM (1,1) prediction method and the adjusted hyperbolic method are consistent with the measured data and indulge in safety. Therefore these three kinds of method has a good adaptability for engineering application.

**[Key words]** squeezed branch pile      ultimate bearing capacity of single pile      curve fitting method

(上接第 8171 页)

## Fast Algorithm for Connected Region Labeling of Binary Image

CAO Chang-hu, LI Ya-fei

(Electronic and Information Engineering Institute, Nanjing University of Technology, Nanjing 211816, P. R. China)

**[Abstract]** Common region labeling algorithms often cannot meet the requirement of processing image at real-time because of the limitation of processing time and memory space. A new fast region labeling algorithm is proposed in view of these shortcomings. Through combining region labeling and contour tracing, this algorithm labels all connected regions in a single scan and the problem of label collision which most improved methods have to deal with is avoided. In addition, the algorithm could label connected regions accurately independent of the shape and area of regions. At the end, comparative experiments with several existing algorithms are made. The experimental results show that the proposed algorithm is more efficient and effective.

**[Key words]** connected region      region labeling      binary image      contour tracing      fast algorithm