

菌群细胞图像分离算法研究

刘相滨¹, 邹北骥², 孙家广³

(1. 湖南大学计算机与通信学院, 湖南长沙 410082; 2. 中南大学信息科学与工程学院, 湖南长沙 410083; 3. 清华大学软件学院, 北京 100084)

摘要: 颗粒图像分析中聚堆目标的分离对目标的计数及特征的提取非常重要. 现有分离算法都要求聚堆目标粘连处凹陷比较明显, 并且/或者存在灰度局部最小边缘. 菌群聚堆细胞大小不一、聚堆形态各异, 多数聚堆细胞并不具备上述条件. 文中提出了一种基于聚堆区域轮廓跟踪、删除的分离算法, 依据跟踪“虫”在跟踪过程中遇到已跟踪过轮廓点的情况判断分离的进行. 实验结果表明算法对菌群聚堆细胞分离的成功率要高于现有算法 10% 以上.

关键词: 聚堆; 分离算法; 8-连通; 轮廓跟踪; 菌群图像

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2005) 06-1056-04

A New Algorithm for Separating Cell in Bacteria Image

LIU Xiang-bin¹, ZOU Bei-ji², SUN Jia-guang³

(1. College of Computer and Communication, Hunan University, Changsha, Hunan 410081, China;

2. College of Information Science & Engineer, Central South University, Changsha, Hunan 410083, China;

3. School of Software, Tsinghua University, Beijing 100084, China)

Abstract: Separating objects touching and overlapping with each other into single ones is very important for the counting and feature extracting. Most of the existing separating algorithms are based on two hypotheses: (1) the touching objects tend to form an acute angle on the contour where objects touched, and/or (2) there would always be a lower optical density at touching points. For colony image, the bacteria cells vary in sizes and touching manners. Many touching cells are short of above hypotheses. A new separating algorithm based on 8-connected contour tracking is proposed in the paper. The boundary of touching-cell was tracked and deleted layer by layer, and judgment was made by the tracking result to determine whether the touching-cell should be separated. The experiment results show the success ratio to separate bacteria touching-cell is 10% higher than the existing algorithms.

Key words: touching and overlapping; separation algorithm; 8-connected; contour tracking; colony image

1 引言

随着图像分析技术的不断提高, 细胞、植物种子等各种颗粒图像的自动分析成为近年来研究的一个热点^[1-3, 8-10], 在医学诊断、病理研究、植物选种、生活卫生检测及工业矿物颗粒分析等应用领域均有重要的意义.

颗粒图像的自动分析, 图像分割是其中的一个关键步骤. 首先需要将颗粒目标从背景中分割出来, 然后统计目标个数并提取目标的有关参数, 例如面积、周长、直径、中心矩等, 最后对其进行分析和归类. 图像分割技术是计算机视觉和人工智能领域中一项意义重要而又颇为艰巨的研究工作, 属图像信息工程的一大经典难题, 长期以来一直是困扰和制约模式识别发展的瓶颈问题和技术难点. 众多国内外学者对此进行了深入、广泛的研究, 推出了不少应用算法, 但当前没有一种方法对所有测试图像分割效果均为最佳^[4-7]. 特别是在颗粒图像分析中, 不论采用何种分割算法, 目标总会产生聚堆现象, 即目标粘连、重叠在一起. 为了提高分析结果的准确性, 需要设计一种有效的分离算法将它们自动分离为单个目标.

现有聚堆目标的分离算法主要有基于边缘检测和基于数学形态学两类. 基于边缘检测的算法来自以下事实: 粘连、

重叠目标由于在三维空间中深度的不同可能会导致其在二维图像上的投影不同, 从而具有灰度差异, 表现在粘连处存在灰度局部最小、梯度差边缘线, 因此可以结合凹点的分析来检测这条边缘线, 从而将其分离^[1]. 而在实际应用中, 由于光照及深度原因, 很多场合下目标在三维空间中的深度差异很小并且光源充分, 其二维图像中聚堆目标间并不存在灰度差异. 因此, 分离算法主要是解决聚堆目标间不存在灰度差异的分割问题, 因而研究最多的是基于数学形态学的分离算法, 例如采用腐蚀、距离变换操作并结合边缘搜索、分水岭等方法进行分离^[2, 3, 8-10]. 本文在分析和比较了这些算法的基础上, 提出了一种根据跟踪聚堆目标区域轮廓所获得的信息进行分离的算法. 该算法对聚堆目标区域的轮廓进行逐层跟踪、删除, 根据跟踪过程中跟踪“虫”对轮廓分段的情况来判断分离的进行, 从而克服了对凹陷性明显及具有边缘线的要求. 实验结果表明该算法对菌群聚堆细胞进行分离的成功率要高于现有算法 10% 以上.

2 基于数学形态学的分离算法

基于数学形态学的分离算法一般分两个步骤: 首先利用高斯滤波、腐蚀或距离变换得到单个目标的数目及各目标的

收稿日期: 2004-06-11; 修回日期: 2004-12-06

基金项目: 湖南省自然科学基金 (No. 03JJ Y6025)

核,然后采用分水岭、条件膨胀或测地重建等算法实现分离.主要有以下三类典型方法:

(1)分离点搜索^[2]. 阈值化后的二值图像通过连续的腐蚀,使得聚堆目标区域分裂为一些小区域,这些小区域分别对应于单个目标.连续的腐蚀过程中,每次腐蚀操作后,判断是否有区域分裂发生,如发生,则在未经腐蚀的区域的边缘轮廓上搜索分离点,从而达到分离目标的目的.

(2)分水岭算法^[3,8,9]. 阈值化后的二值图像通过高斯滤波变换或极限腐蚀得到一个距离变换图,图中的局部最大值对应于各目标的核,从而可以获知单个目标的数目及各目标的核.如果将距离图视为一拓扑表面,则距离图中的局部最大值对应于山峰,局部最小值对应于山谷.假设有一颗水珠落在拓扑表面,对于极限腐蚀的一个给定点,水珠会滚过的所有的点的集合被称为该给定点的流域.山谷就是流域间的分界线,连接流域分界线即可分离聚堆目标.算法的实现是以各目标的核为基础,进行测地膨胀运算,同时避免各测地膨胀领域相连,直到重建原图,则各测地膨胀领域的相邻处,即形成分界线.

(3)水坝算法^[10]. 水坝算法直接对灰度图进行灰度形态学运算,不须阈值化处理过程.该算法基于聚堆目标存在的局部梯度差的梯度信息,以类似区域生长的方式从种子点开始搜索各区域的边界.其基本原理是:先对灰度图像进行形态梯度运算,得到梯度图,然后将经过边缘增强的灰度图像看作一种自然的地质表面,像素的亮度值对应地理高度,而地质表面的山脉对应颗粒图像的边缘、重叠边界,低谷、盆地则对应灰度图像中的背景或目标区域;假设在该地质表面的盆地打一些小孔,使水从这些孔中溢出,不断地淹没各流域,当不同流域的水面不断升高到要相互汇合时,就筑起一道堤坝,防止不同流域中的水相互汇合;水面不断上涨,最终,仅仅高出水面的堤坝部分可见,而其他部分均被淹没;这时,可见的堤坝部分,即对应目标的边缘和目标间的分界线.

分离点搜索算法原理简单,实现方便.由于在连续腐蚀过程中,可能会有不止一个的区域在某一次腐蚀后发生分裂,且个数事先未知.要在此基础上从边缘轮廓上搜索分离点,分离点的个数、分布形式就变得很难确定,因此,该算法对于聚堆情况较复杂的目标很难实现.分水岭算法使用了数学形态学中的测地学理论,分离的准确度相对要高,但由于需要跟踪每次腐蚀的操作运算过程,运算量大,对计算机软、硬件要求较高.此外,这两类算法都用到了腐蚀操作,对于凹陷不太明显的聚堆目标,如果腐蚀操作不能最终将其分裂为小区域,则算法不会成功.距离变换是计算二值图像中每个目标像素到背景像素的最短距离,理论上需要对图像进行全局操作运算,即计算该像素与所有背景像素的距离,再取最小值.除非这幅数字图像的尺寸非常小,不然这种全局操作的计算量是非常大的.考虑到算法效率问题,实际应用中的距离变换实现算法是采用倒角(chamfer)或类似于腐蚀的快速算法,得到的是近似的、而不是精确的欧氏距离,因此存在同样的不足.

水坝算法基于聚堆目标存在局部梯度差的梯度信息进行分离,要求灰度图中聚堆目标粘连处存在灰度梯度差.而

且在很多情况下,对颗粒图像形态梯度化提取出的边界梯度差曲线是不连续的、非闭合的,还需要复杂的边缘连接算法才能真正分割图像为闭合区域,增加了算法的复杂性.另外,值得注意的是种子点的选定,同一区域取不同位置的种子点,可能会影响该区域边界的划定.

在菌群细胞图像分析中,细胞聚堆区域粘连情况严重、聚堆形态各异.如图 1 为一幅直径 9cm 的国际标准细菌培养皿培养的菌群细胞图像,经分析,具有如下特性:

(1)细胞面积差异很大.不考虑很小的细胞,单个细胞面积大约在 5 - 1900 个像素之间,因而无法根据区域面积的大小来判定该区域是否属于聚堆区域,以及该聚堆区域所包含的单个细胞数.

(2)大多数细胞粘连很紧.粘连很紧的细胞,粘连处的凹陷很小,而且很多聚堆区域粘连的细胞几乎没有灰度差异,因此也不存在灰度梯度差.

考察其中的两个聚堆区域,均由两个细胞粘连而成,如图 2、图 3 所示,图中(a)分别为原图像、动态阈值二值化分割图、本文算法分离结果图,(b)为动态阈值二值化分割后的标号矩阵,图 3(c)为该聚堆区域较小细胞被腐蚀掉的前一次腐蚀结果.由于图 2 中粘连的两个细胞大小、灰度差异都不大且粘连很紧,所以粘连处凹陷很小、也没有梯度差边缘线.通过对标号矩阵的分析可知,两个聚堆区域的整个腐蚀过程中都不会产生区域分裂,因此上述分离算法都不能使其分离,而应用本文算法分离成功.

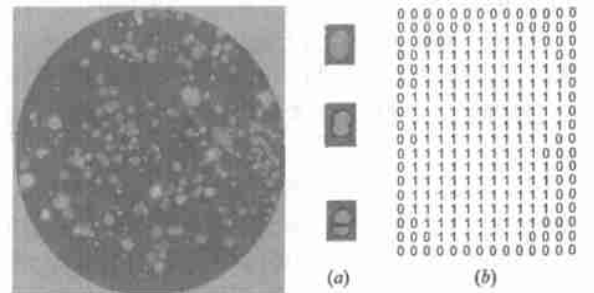


图 1 原始细胞图像 图 2 一聚堆细胞及其标号矩阵

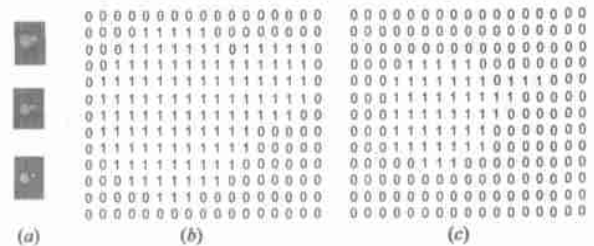


图 3 一聚堆细胞及其标号矩阵

3 基于轮廓跟踪的分离算法

3.1 算法思想

如图 4 所示聚堆目标区域,带“+”的像素为区域的 8-连通轮廓.

区域宽度 目标区域在某点的宽度是指从区域在该处的一个轮廓像素出发穿过区域到达另一边轮廓的最短路径的长度(像素数),这条路径完全包含在区域内并且将区域一分为

二. 如图 4 中聚堆区域在 p_5 、 p_6 、 p_{13} 和 p_{14} 处的宽度均为 2.

8-连通轮廓跟踪 从区域左上角轮廓像素出发,按逆时针如图 4 中箭头所示方向对区域的 8-连通轮廓像素进行逐点跟踪.跟踪结果得到一个轮廓像素点序列: $p_0 p_1 p_2 \dots p_n$,其中, p_0 为跟踪起点、 p_n 为终点.易知,如果区域为孤立点,则序列中只有 p_0 一个点,而单像素宽位置的轮廓点,在序列中将重复出现.

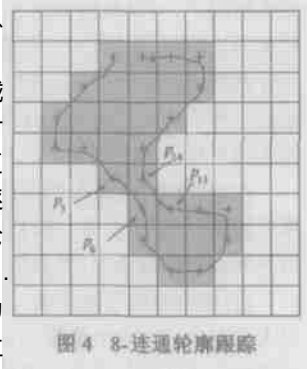


图 4 8-连通轮廓跟踪

从图 4 中可以看出,跟踪时,如果区域某处的宽度小于等于 2 个像素,那么跟踪“虫”在第二次经过该处时会遇到原先跟踪过的轮廓点,例如跟踪到 p_{13} 和 p_{14} 时,会分别遇到已跟踪过的轮廓点 p_6 和 p_5 .一般情况下,这正是两个目标的粘连处,如果此时将跟踪到的轮廓序列点从区域删除,则聚堆区域会在此处分裂为两个目标.进一步,即使两个目标粘连处的宽度大于 2 个像素,只要图像中的单目标满足凸集特性,那么两个目标的粘连处或多或少会存在一些凹陷,这样,如果能等宽度地不断跟踪、删除区域轮廓,则聚堆区域终会在粘连处分裂.实际上,这也是现有基于腐蚀或距离变换分离算法的思想^[2,3,8,9],然而,致使这些算法失效的根本原因在于腐蚀运算或采用的距离变换操作对聚堆目标轮廓的处理并不是真正等距的,从而使区域收缩不均匀,无法保留较小的凹陷信息.为了提高分离的成功率,要么设计一种能够真正体现等距性的快速、精确欧氏距离变换算法,要么就不能仅依据区域的分裂来进行分离.

为此,本文从两方面加以改进:(1)根据人的视觉特性,8-连通轮廓线比 4-连通轮廓线更能体现等距性,而通常的腐蚀运算每次腐蚀掉的是 4-连通轮廓;(2)对于凹陷不明显的聚堆目标,即使采用 8-连通轮廓跟踪、删除,最后也不会产生分裂.本文算法并不根据区域的分裂与否来实施分离,而是根据跟踪“虫”在跟踪过程中遇到已跟踪过的轮廓点时轮廓线分段的情况决定是否实施分离,从而克服现有算法的不足.

3.2 算法描述

如图 5 所示,以一个包含两个目标的聚堆区域为例,算法的具体步骤可描述如下:

跟踪过程中,对跟踪过的轮廓像素设置已跟踪标志,在跟踪“虫”回到跟踪起点、结束本层轮廓线的跟踪前.

(1)如果当前轮廓像素为已跟踪过像素(区域宽度为 1 时),或当前轮廓像素的 4 邻域为已跟踪过像素(区域宽度为 2 时),则说明跟踪“虫”进入了一个分离段,该聚堆目标可能应在此处分离.设此时遇到的已跟踪过轮廓像素为 A,当前轮廓像素为 B,其中,A、B 可能为同一轮廓像素(区域宽度为 1 时),也可能不是同一轮廓像素

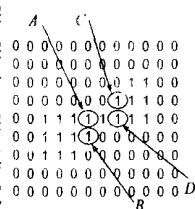


图 5 分离段示意图

(区域宽度为 2 时),但它们在跟踪得到的轮廓像素点序列中处在两个不同的位置.

(2)继续跟踪.当下一轮廓像素及其 4 邻域都是未跟踪过像素时,则说明跟踪“虫”离开该分离段,设此时遇到的最后一个已跟踪过像素为 C,当前轮廓像素为 D.

(3)距离计算.分别计算 A 和 B、C 和 D 在跟踪得到的轮廓像素序列中的距离.设 AB 间距离为 l_{ab} , CD 间距离为 d , CD 间反向距离为 l_{cd} ,即 $l_{cd} = L - d + 2$,L 为聚堆区域轮廓周长.如在图 4 中,跟踪得到的轮廓像素序列为 $p_0 p_1 p_2 \dots p_{20}$,A 为 p_6 、B 为 p_{13} ,C 为 p_5 、D 为 p_{14} .因此, $l_{ab} = 13 - 6 + 1 = 8$, $L = 21$, $d = 14 - 5 + 1 = 10$, $l_{cd} = L - d + 2 = 21 - 10 + 2 = 13$.

(4)分离判断.如果 l_{ab} 和 l_{cd} 均超过设定的阈值 T_1 ,则认为这个分离段是聚堆目标的分离处,转(5)实施分离操作,否则不是,转(6).

(5)实施分离.首先在本层轮廓上确定局部分离点对,局部分离点对的确定需要考虑分离段出现的情形.分离段出现的位置包含两种情况,一种是如图 4、图 5 所示的聚堆区域出现分裂,另外一种为聚堆区域不出现分裂,小目标以毛刺的形式出现在聚堆区域一端,如图 3(c)所示.对第一种情况,优先选取分离段中区域宽度为 1 部分的中点,如果分离段中区域宽度均为 2 则取分离段的中点作为局部分离点对;对第二种情况,遵循靠近大目标进行分离的原则,取分离段中靠近剩余区域的一端作为局部分离点对.然后在原图像上,从局部分离点对出发、由方向控制分别向分离段两侧搜索分离线,最后沿分离线将聚堆区域一分为二.为提高分离的准确性,可以增加分离的合理性判断,包括对灰度梯度差、凹凸性、区域整体形状参数等信息的综合判断.

(6)将本层轮廓像素序列从聚堆区域中删除,然后转(1)继续对剩余区域的轮廓进行跟踪,直至剩余区域的面积为 0 或小于设定的阈值 T_2 ,算法结束.

对于多目标聚堆区域,处理方法类似.即使在一层轮廓的跟踪过程中出现多个分离段,也可根据由分离段隔开的各个子区域轮廓的周长情况,确定分离段的取舍及分离操作的实施.

关于阈值 T_1 、 T_2 的设定: T_1 是根据实验获得的,设得过小则会出现“过分离”、设得过大则会出现“欠分离”的情况,其值和图像中各目标的总体面积有点关系,面积大其值也应大一些,一般取值 5~10; T_2 和图像中目标的最小面积有关,应设为图像中目标的可能最小面积值.

4 实验结果与讨论

采用 Visual C++ 6.0 实现了上述算法,并对实际菌群细胞图像进行了实验,选取其中两幅具有代表性的图像,将其实验数据列于表 1,部分分离效果如图 6 所示.图 6 中的三幅图像分别为原细胞图像、动态阈值二值化分割图、本文算法分离结果图.

从表 1 中数据可以看出,本文算法分离该类菌群聚堆细胞的成功率比现有算法提高了 10%以上,直观上从图 6 也可看出,有些聚堆细胞在粘连处凹陷及灰度变化均不明显,采用现有分离算法是无法实现分离的,而本文算法基本上实现了正确分离.

表 1 各算法对聚堆细胞分离结果比较

图 像	分离前		分离后							
	聚堆细 胞区域	包含单 细胞数	文献/5/算法		文献/6/算法		文献/7,8/算法		本文算法	
			细胞数	成功率 %	细胞数	成功率 %	细胞数	成功率 %	细胞数	成功率 %
1	37	89	67	75.3	69	77.5	72	80.9	83	93.3
2	46	112	83	74.1	89	79.5	87	77.7	105	93.8

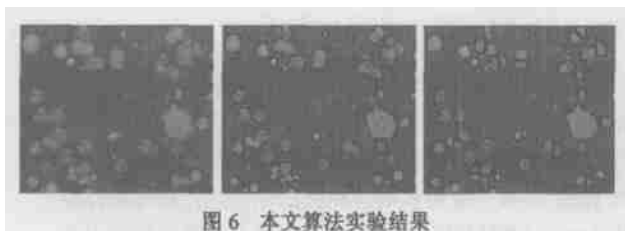


图 6 本文算法实验结果

本文算法可适用于所有单目标满足凸集特性并具有—定圆度的颗粒图像中聚堆目标的分离,其特点主要有两个:

(1) 采用 8-连通轮廓跟踪算法对聚堆目标区域轮廓进行跟踪、删除,模拟腐蚀、距离变换操作。一方面,8-连通轮廓跟踪、删除比现有算法中采用的腐蚀、距离变换操作更能体现等距性;另一方面,在每层轮廓跟踪过程中能立即知道是否应进行分离,因而算法效率要比基于腐蚀、距离变换的算法高。基于腐蚀的分离算法首先要根据腐蚀结果搜索分裂情况,然后再去搜索分裂点、分离线。基于距离变换的算法要结合分水岭算法来完成分离,而分水岭算法的运算量相当大。

(2) 和基于腐蚀、距离变换的分离算法有根本区别。基于腐蚀、距离变换的分离算法只有在聚堆目标区域经过腐蚀、距离变换操作后发生区域分裂才能进行分离,而本文算法即使聚堆区域不发生分裂,也能实现分离,因而提高了算法分离的成功率。

当然,由于图像信息的特殊性,和图像的分割算法一样,任何一种分离算法都不可能适合所有情况下的聚堆目标分离,本文算法同样具有局限性:

(1) 阈值 T_1 的设定很关键。过小则会出现“过分离”,过大则会出现“欠分离”,当图像中聚堆目标数量多且大、小细胞的面积差异较大时,将很难设定合适的阈值 T_1 使其适合所有聚堆目标的分离,如果 T_1 能够做到自适应,则算法的通用性会得到提高。

(2) 本文算法不适合狭长目标(如染色体细胞)及单目标形状不规则的颗粒图像中聚堆目标的分离。形状的不规则会使得聚堆区域在轮廓跟踪、删除过程中产生不可预测的分裂从而导致错误分离,但是,如果单目标的形状不规则只是由于具有或多或少的毛刺,或者单目标形状的不规则程度不是很厉害,则可设计光滑算法或直接应用数学形态学的开运算首先对目标轮廓进行光滑,消去毛刺、分支等,再应用本文算法进行分离。

参考文献:

- [1] Liang Ji. Intelligent splitting in the chromosome domain[J]. Pattern Recognition, 1989, 22(5): 519 - 532.
- [2] 陆建峰,杨静宇,等. 重叠细胞图像分离算法的设计[J]. 计算机研究与发展, 2000, 37(2): 228 - 231.
Lu Jianfeng, Yang Jingyu, et al. Design of a separating algorithm for overlapping cell images[J]. Journal of Computer Research & Development, 2000, 37(2): 228 - 231.
- [3] Talukder Ashit, et al. Modified binary watershed transform for segmen-

tation of agricultural products[J]. SPIE, 1998, 3543: 53 - 64.

- [4] 景晓军,李剑峰,刘郁林. 一种基于三维最大类间方差的图像分割算法[J]. 电子学报, 2003, 31(9): 1281 - 1285.
Jing Xiaojun, Li Jianfeng, Liu Yulin. Image segmentation based on 3D maximum between cluster variance[J]. Acta Electronica Sinica, 2003, 31(9): 1281 - 1285.
- [5] 王艳清,卫保国,蔡铁珩,等. 一种基于先验知识的自动舌体分割算法[J]. 电子学报, 2004, 32(3): 489 - 491.
Wang Yanqing, Wei Baoguo, Cai Yiheng, et al. Acknowledge based arithmetic for automatic tongue segmentation[J]. Acta Electronica Sinica, 2004, 32(3): 489 - 491.
- [6] 邢延超,谈正. 基于多阈值融合的图像分割[J]. 计算机学报, 2004, 27(2): 252 - 256.
Xing Yanchao, Tang Zheng. Multi threshold fusion based image segmentation[J]. Chinese journal of computers, 2004, 27(2): 252 - 256.
- [7] 叶齐祥,高文,王伟强,等. 一种融合颜色和空间信息的彩色图像分割算法[J]. 软件学报, 2004, 15(4): 522 - 530.
Ye Qixiang, Gao Wen, Wang Weiqiang, et al. A color image segmentation algorithm by using color and spatial information[J]. Journal of Software, 2004, 15(4): 522 - 530.
- [8] 袁天云,姜志国,孟如松. 目标分割图中粘连对象的自动切割和分离[J]. 中国体视学与图像分析, 2003, 8(1): 40 - 43.
Yuan Tianyun, Jiang Zhiguo, Meng Rusong. Automatic splitting and separating overlapped objects in target-segmented image [J]. Chinese Journal of Stereology and Image Analysis, 2003, 8(1): 40 - 43.
- [9] 聂守平,王鸣. 球状重叠区域图像分割与参数统计算法研究[J]. 中国激光, 2003, 30(11): 1036 - 1040.
Nie Shouping, Wang Ming. Algorithm study of image segmentation and parameters calculation for spherical overlapping region image [J]. Chinese Journal of Lasers, 2003, 30(11): 1036 - 1040.
- [10] 陆振晖,范影乐,庞全. 基于数学形态学的重叠细胞分离方法及比较研究[J]. 计算机工程与应用, 2004, 40(6): 57 - 59, 104.
Lu Zhenye, Fan Yingle, Pang Quan. Comparison of several algorithms based on mathematical morphology of overlapped cells segmentation [J]. Computer Engineering and Applications, 2004, 40(6): 57 - 59, 104.

作者简介:



刘相滨 男, 1970 年 12 月生于江西吉安, 2000 年 7 月湖南大学计算机应用专业硕士毕业, 现为湖南大学计算机应用专业博士生, 湖南师范大学副教授, 硕士生导师, 研究领域为数字图像处理与分析、模式识别。Email: xbluifrank@hunnu.edu.cn