

# 基于边界剥离的细胞图象分离算法

刘相滨<sup>1)</sup> 邹北骥<sup>2)</sup> 胡峰松<sup>3)</sup>

<sup>1)</sup>(湖南师范大学图象识别和计算机视觉研究所, 长沙 410081)

<sup>2)</sup>(湖南大学计算机与通讯学院, 长沙 410082)

<sup>3)</sup>(湖南大学多媒体教育学院, 长沙 410082)

**摘要** 细胞图象的自动判读中,经常遇到聚堆细胞的问题,因而需要采用一种有效的分离算法把它们分离为单个细胞.目前存在的大多数分离算法都要求聚堆细胞连接处的凹陷性比较明显,或者要求在细胞连接处存在局部最小灰度值,否则难以实现正确分离.为了解决这一问题,提出了一种基于边界剥离的分离算法,该算法首先对聚堆细胞区域进行层层剥离,然后根据剥离结果判断是否发生了细胞分裂,进而完成分离,这样就避免了对凹陷性及局部最小灰度值的要求,实验结果表明算法是有效的.

**关键词** 细胞图象 聚堆 分离 凹点

中图分类号: TP391.41 Q2-33 文献标识码: A 文章编号: 1006-8961(2002)03-0234-06

## Separating Algorithm for Cell Image Based on Boundary-Stripped

LIU Xiang-bin<sup>1)</sup>, ZOU Bei-ji<sup>2)</sup>, HU Feng-song<sup>3)</sup>

<sup>1)</sup>(Institute of Image Recognition & Computer Vision, Hunan Normal University, Changsha 410081)

<sup>2)</sup>(College of Computer and Communication, Hunan University, Changsha 410082)

<sup>3)</sup>(College of Multimedia Education, Hunan University, Changsha 410082)

**Abstract** The clustering phenomenon often appears in the cell image auto-reading system, i. e., some cells overlap or touch together to form a big area. It is necessary to design an effective algorithm to separate the clustering cells into single ones. Most of the existed separating algorithms are based on two hypotheses: 1) at points where cells touch, the cluster boundary tends to form an acute angle; 2) at points where the cells touch, the optical density is relatively low, otherwise it will be very difficult to separate them properly. But in the actual cell images, the two hypotheses are very difficult to be met. To solve this problem, an algorithm based on boundary-stripped is presented in the paper. The boundary of clustering cells is stripped layer by layer in the algorithm. During stripping, judgement is made to determine whether the splitting has happened, and then the actual separation is taken in the original cell image. It does not require the clustering cells should necessarily meet the two hypotheses. The experiment result shows that the algorithm is effective.

**Keywords** Cell image, Clustering, Separating, Concavity

## 0 引言

细胞图象自动判读系统的设计是近年来研究较多的一个内容<sup>[1~3]</sup>.这种自动判读系统和人工判读相比,既可以减少主观上的干扰,以提高判读精度,又可以减轻技术人员的繁重劳动负担,而提高工作效率,因而在细胞生物学研究中具有非常重要的意义.

细胞图象自动判读系统的主要任务是对细胞图象进行自动分析,并将细胞群分割成单个细胞,以便统计细胞个数,并计算各单个细胞的有关参数,例如面积、周长、直径、中心矩等,其中最关键的技术为图象分割,分割质量的好坏将直接影响后续的一系列工作.关于图象分割,目前已有许多成熟的算法,但由于细胞图象的特殊性,不论采用何种分割算法,其分割结果图象中,总会存在一些细胞聚堆的现象,即

细胞粘连、重叠在一起的现象,因而影响细胞的判读.有些判读系统采用人工交互方式来对聚堆细胞进行分离<sup>[1,2]</sup>,但速度较慢,且常受人因因素的干扰,为了提高系统的自动化程度,必须设计一种有效的分离算法,以便把它们自动分离为单个细胞.目前存在的大多数分离算法都要求聚堆细胞连接处的凹陷性比较明显<sup>[3~5]</sup>,有的还要求聚堆细胞连接处有一条灰度值局部最小的连接线<sup>[4]</sup>,否则难以实现正确分离,然而在实际细胞图象中,很多聚堆细胞在连接处并没有明显的凹陷特性,甚至没有灰度值局部最小连接线,从而给分离算法的设计带来了很大困难.针对这一问题,本文提出了一种基于边界剥离的分离算法,该算法首先对聚堆细胞区域进行层层剥离,然后根据剥离结果判断是否发生了细胞分裂,进而完成分离,从而避免了分离算法对凹陷性要明显及要具有连接线的要求.实验结果表明,该算法的分离效果是比较理想的.

## 1 问题分析

大家知道,把粘连、重叠在一起的细胞分离成单个细胞,关键在于如何快速、准确地找到分离点.所谓分离点,实际上就是最终用来进行细胞分离的一个点对,因为连接点对的一条曲线段可以把该区域比较合理地一分为二,故这条曲线段通常称为分离线.通过分析发现,分离点应该处于细胞的连接处,由于聚堆的细胞在连接处一般呈凹陷状,因此,很多分离算法都是基于凹点进行的<sup>[3~5]</sup>.

这种基于凹点的分离算法,其思想是:首先找出聚堆细胞边界上所有的凹点,然后对凹点进行配对,以得到分离点,配对的原则是分离点对间的直线距离局部最小,并且分别处于中轴的两侧,分离点确定后,就可根据灰度值最小路径或边缘信息来搜索分离线,进而进行细胞的分离.分离算法的思想虽然很简单,但实际应用起来却复杂得多,由于聚堆细胞的形态各异,因而如果直接按照上述思想设计算法,将存在如下一些问题:

(1) 真正的凹点很难确定 目前,计算凹点的算法主要是通过计算 $K$ 个轮廓点所组成的夹角来判断,这种方法计算量大,计算结果与 $K$ 的取值以及设定的夹角大小有关,且容易受噪声的影响,极有可能把噪声点当作了凹点,却把真正的凹点丢失了.

(2) 凹点不能配对或很难正确配对 一种情况是,两个细胞的连接处两边应各有一个凹点,但有时

因为各种原因,致使细胞连接处的凹陷并不明显,搜索时,其中一个凹点,甚至两个凹点都没有找到,这时就无法配对;另一种情况是,对于多细胞区域,即使正确找出了所有凹点,但由于事先并不知道聚堆细胞区域内细胞的具体个数,仅仅根据分离点配对原则的两个条件是很难进行正确配对的.

(3) 分离线很难搜索 如果配对不正确的分离点,则无论如何处理,所得到的分离线肯定是错误的.对于无法配对的凹点,有时可能搜索到正确分离线,有时则搜索不到.例如,当两个细胞的连接处只有一个凹点时,如果这两个细胞的连接处存在灰度值局部最小的一条“沟渠”,或者存在一条灰度值突变线,即边缘线,那么都很容易进行分离;但是,如果两个细胞的灰度值差别不大,并且粘连很紧,则粘连处灰度值变化很小,不仅不存在灰度值局部最小的连接线,也不存在灰度值突变线,此时,分离线就很难搜索到.

因此,在设计分离算法时,不能完全依靠凹点,必须考虑其他的因素.为此,文献<sup>[4]</sup>和文献<sup>[5]</sup>对基于凹点的分离算法进行了改进,并在一定程度上改善了分离效果,但仍然存在一些不足.

文献<sup>[4]</sup>提出,首先从聚堆细胞边界上找到的凹点中选择一个夹角较锐的凹点作为搜索起点,然后从该点出发,沿凹点与前后两个相邻的第 $K$ 个边界点所成直线的夹角顶点方向,向边界另一方搜索分离线,搜索原则是灰度值局部最小,并且局部距离最短.尽管该算法对搜索到的分离线还要进行合理性判断,如果符合要求,则算法结束,否则再选择一个凹点进行搜索,直至所有凹点都搜索完,但由于该算法毕竟采用了灰度值局部最小原则,因此虽有一定的适用性,但对于没有明显灰度值差异,也没有明显凹陷的聚堆细胞来说,将仍然是个问题.

文献<sup>[5]</sup>提出基于数学形态学的分离算法,文中采用 $3 \times 3$ 结构元来对聚堆细胞进行连续的腐蚀运算操作,使得聚集在一起的细胞,分离为一些小的区域,这些小区域分别对应于单个细胞,在此基础上,再从未经腐蚀的区域上找到分离点,从而达到分离细胞的目的.但是,由于腐蚀运算会使聚堆细胞的边界收缩速度很快,并且收缩厚度不均匀,因此对一些粘连很紧、粘连处凹陷不明显的聚堆细胞,经过几次腐蚀后,将造成凹陷消失,并使聚堆细胞最终不会产生分裂,从而导致算法失败.

如图1所示,其中图1(a)中3幅图象分别为原图象、动态阈值法二值化分割效果图、本文算法分离

结果图;图 1(b)为动态阈值法二值化分割后的标号矩阵.从图中可以看出,由于聚堆的两个细胞的灰度值几乎没有差异,并且粘连很紧,没有边缘线;细胞连接处凹陷程度也不大,故计算凹点时,如果  $K$  值设得过小,则找不到凹点.实际上,由于通常的细胞图象中细胞数量多,聚堆细胞情况差异大,因此即使对整幅图象在  $K$  值和夹角阈值均设得比较合适时,文献[4]算法也不能分离图 1 聚堆细胞.从图 1(b)中则可以看出,经过两次腐蚀运算之后,聚堆细胞区域的凹陷消失了,最后收缩成一条直线,所以文献[5]算法也不能实现图 1 聚堆细胞的分离.

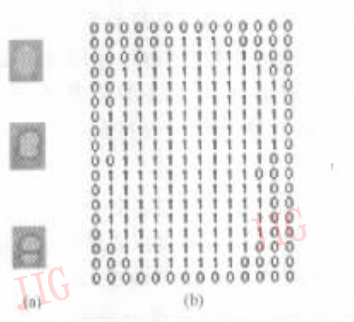


图 1 一聚堆细胞及其标号矩阵

通过采用“虫跟踪”算法来跟踪区域的外边界时发现,当两个细胞连接处的厚度为 1 时,在“虫”回到搜索起点之前,会路经原先走过的一段路,而这段路恰好是两个细胞的连接处.经过进一步的分析发现,即使两个细胞连接处的厚度不为 1,只要细胞图象中的单细胞满足凸集特性(实际大部分细胞都满足凸集特性),则两个细胞的连接处或多或少会有一

点凹陷,这样,只要能等厚度地不断剥离区域外边界,外边界将首先在连接处断开,基于这个事实,本文提出了一种基于边界剥离的细胞图象分离算法.

## 2 算法设计

本文算法的基本思想是:首先对聚堆细胞区域边界进行一次剥离,并判断是否发生了细胞分离,若发生了,就在刚剥离的那一层边界上,搜索局部分离点,进而在原图象上搜索分离线,以完成细胞分离.这样的剥离过程不断重复下去,直到单个区域的面积小于设定的阈值为止.根据该思想,算法必须解决如下 3 个关键问题:①如何判断是否有细胞分离出来了?②如何在刚剥离的边界上,确定局部分离点?③如何在原图象上确定分离线?

(1) 一般情况下,只需简单地根据在细胞区域剥离过程中,是否有新区域产生来判断是否有细胞分离出来.但在实际工作中,可能存在这样的情况,即当一个较大的细胞和一个很小的细胞粘连在一起时,两细胞连接处的凹陷并不明显,并且随着细胞边界的不断剥离,始终不会产生新的区域.如图 2 所示,其中,图 2(a)的 3 幅图象分别为原图象、动态阈值法二值化分割效果图及本文算法分离结果图,图 2(b)为动态阈值法二值化分割后的标号矩阵,图 2(c)为剥离过程中产生细胞分离前的标号矩阵.从图中可以看出,在细胞边界剥离过程中,虽始终不会有新区域产生,但由于剥离前与剥离后区域的宽度会发生一次较大的变化,因此,在剥离过程中,如果剥离前与剥离后细胞区域的高度或宽度发生较大变化,且超过设定阈值  $T_1$  时,则认为可能有新的细胞分离出来了,但因为一些矩形细胞在最后分离结果为一个点之前,会发生一次区域高度或宽度的较大



图 2 剥离前、剥离后区域的宽度发生较大变化时,表明可能有细胞分离

变化,实际上有时并没有新细胞分离出来,故还必须进行进一步的判断.经过这样的处理,本文算法即可对细胞进行正确的分离.

(2) 在跟踪区域外边界时,约定以左上角边界点为搜索起点,并按逆时针方向搜索.搜索过程中,对搜索过的边界点设置已搜索标志,同时保存搜索边界点的全序列.搜索“虫”在回到搜索起点前,如果下一边界点为已搜索过的点(边界厚度为 1 时),或下一边界点的邻域为已搜索过的点(边界厚度为 2 时),则表明搜索“虫”进入了一个分离段,但由于这个分离段可能是细胞的分离处,也可能不是,故必须进一步判断.判断的方法是,在分离段中选取代表点  $A$ 、 $B$ ,这里需要注意的是: $A$  和  $B$  可能为同一点(边界厚度为 1 时),也可能不是同一点(边界厚度为 2 时),但由于它们在搜索边界点全序列中,肯定处在不同的两个位置,因此可通过计算  $A$ 、 $B$  所在两个位置之间相隔的边界点数来判断是否是分离点对,如果超过设定的阈值  $T_2$ ,则认为  $A$ 、 $B$  是一个真正的局部分离点对,否则这个分离段不是细胞的分离处.

其中,根据是否有新区域产生,将分离段中代表

点的选取分为下列两种情况(如图 3 所示,图中  $S$  为搜索起点):(1)如果有新区域产生,则分离段必处于两个细胞的连接处,此时取分离段的中点为代表点,如图 3(a)所示;(2)如果属于因剥离前与剥离后区域的高度或宽度变化很大而产生的细胞分离,则遵从靠近大细胞进行分离的原则,若起点  $S$  在分离段内,则取进入分离段时的位置作为代表点,如图 3(b)所示,否则,取退出分离段时的位置作为代表点,如图 3(c)所示.

(3) 确定了局部分离点后,就可以回到原图象进行分离线的搜索了.通过上一节的分析已经知道,分离线应是分离点对间之直线距离局部最小的一条曲线段,并且点对分别处于中轴的两侧.这样只要从局部分离点对  $A$ 、 $B$  出发,由方向控制分别向分离段两侧原图象的边界搜索,然后取直线距离局部最短的边界点即可.考虑到细胞连接处一般具有凹陷性,为了找到比较真实的分离线,可以对每次剥离的边界都进行计算、并保存其一阶凹点信息(差分链码),这样在搜索分离线时,就可沿着凹点线进行,并优先取凹点位置.



图 3 分离段中代表点的选取情况

### 3 算法描述

限于篇幅,在此只给出算法的主控部分.

对于细胞图象中每个细胞区域,执行分离

Separate()操作:

Separate(int Label)

```

{
    if (面积 <  $T_3$ )
        返回;
    获得该区域的标号矩阵;
    while(待分离区域数 > 0)
    {
        剥离一次边界,保存该边界,并计算凹点;
        重新标号,计算区域数,并计算各区域的高度和宽度;
        if( 区域数增加 || (区域数不变且有新旧对应区域的高

```

度或宽度差的绝对值 >  $T_1$ ))

```

{
    GetLocalSplitPoint(int &A, int &B);
    //在刚剥离的那层边界上,搜索局部分离点
    A,B;
    if(A,B 找到)
    {
        参考刚剥离边界上的凹点信息,在原图象中搜索
        分离点对 P,Q;
        根据 P,Q 在原图象中分离细胞;
    }
}
删除面积 <  $T_3$  的区域;

```

## 4 实验结果及讨论

为了验证本文算法的效果,用 Visual C++ 6.0 实现了上述算法,并用实际细胞图象进行了实验,为

说明问题,抽取了其中两幅具有代表性的细胞图象,将其实验结果数据列于表 1 及表 2,部分实验结果如图 4 所示.其中,表 1 为人工及各算法判读得到的细胞数比较,表 2 为各算法对聚堆细胞区域进行分离后的结果数据比较.

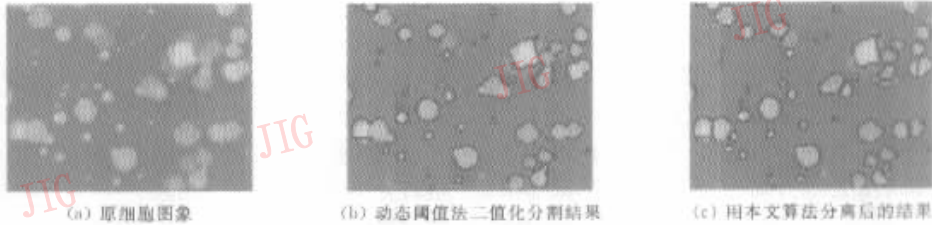


图 4 本文算法实验结果

表 1 人工及各算法判读得到的细胞数对比表

图象编号	人工判读	文献[4]算法		文献[5]算法		本文算法	
		细胞数	判读率(%)	细胞数	判读率(%)	细胞数	判读率(%)
1	345	314	91.0	322	93.3	329	95.4
2	387	349	90.1	358	92.5	372	96.1

表 2 各算法对聚堆细胞分离结果数据对比表

图象编号	分离前		分离后					
	聚堆细胞区域	包含单细胞数	文献[4]算法		文献[5]算法		本文算法	
			细胞数	成功率(%)	细胞数	成功率(%)	细胞数	成功率(%)
1	49	126	97	77.0	104	82.5	115	91.3
2	44	149	113	75.8	121	81.2	138	92.6

在这两幅图象中,最大的细胞其面积达 389 pixels,而最小的细胞面积仅为 5 pixels,其中一幅图象的上细胞粘连现象严重,粘连细胞数最多的聚堆细胞区域包含 13 个单细胞.从表中数据可以看出,本文算法效果要优于文献[4]、[5]算法,细胞判读精度达 95%以上,已达到实用性要求,此外,还可以看出,细胞图象的判读结果不仅和细胞分离算法有关,而且和图象分割算法有关,如果采用的分割算法不好,则在分割后就会丢失一些细胞,从而影响判读精度,限于篇幅,至于如何综合运用各种分割算法及其他措施,来提高判读精度,将另外撰文论述.从图 4 可进一步直观看出,聚堆细胞区域基本上实现了正确分离.

当然,由于图象信息的特殊性,细胞分离算法和图象的分割算法一样,没有哪一种算法能够适合所有情况下的细胞分离,本文算法也不例外,本文算法也存在如下一些问题:

(1) 当细胞图象中聚堆细胞数量较多、聚堆形态各异,并且大、小细胞的面积相差较大时,将很难

设定合适的阈值  $T_1$ 、 $T_2$ ,使其适合所有细胞的分离,如果  $T_1$ 、 $T_2$  能够做到自适应,则算法的通用性会得到提高.

(2) 本文算法的前提是单个细胞必须呈凸集特性,并且具有一定圆度.在此前提下,对那些非重叠、非粘连的单细胞,通过本文算法处理后,它们会收缩至一点(单细胞形状呈近似圆形)或一条短线条(单细胞形状呈近似等宽长形),却不会产生分裂,从而不会影响算法的分离效果.同时,算法为了处理粘连很紧、粘连处凹陷不明显的聚堆细胞而增加了“如果剥离前与剥离后区域的高度或宽度发生较大变化,则认为可能有新的细胞分离出来”的处理,由此可见,本文算法不适合狭长细胞(如染色体细胞)的分离,如果要处理狭长细胞,则必须去掉这个处理,并对算法作适当调整.

(3) 本文算法原则上也不适合于单细胞形状不规则的细胞图象,形状不规则将使得细胞区域在边界剥离过程中会产生不可预测的分裂,从而导致错误分离.但是,如果单细胞的形状不规则,而只是具

有或多或少的毛刺,或者单细胞形状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来首先对细胞边界进行光滑,消去毛刺、分支等后,再用本文算法进行分离。

### 参考文献

- 1 张斌,赵卫东,马世雄等. 人类显带染色体的图象分析与识别系统. 中国图象图形学报,1996,1(5,6):428~432.
- 2 俞帆,欧阳明,陈钢. 人类染色体组型图像分析与识别系统. 云南大学学报(自然科学版),2000,22(5):343~345.
- 3 刘泉,傅祖芸. 一种植物染色体图像核型分析系统. 计算机工程与应用,2000,36(3):71~74.
- 4 Liang Ji. Intelligent splitting in the chromosome domain. Pattern Recognition, 1989,22(5):519~532.
- 5 陆建峰,杨静宇,唐振民. 重叠细胞图像分离算法的设计. 计算机研究与发展,2000,37(2):228~231.

刘相滨 1970年生,讲师,2000年获湖南大学计算机与通讯学院计算机应用专业工学硕士学位. 主要研究方向为图象理解、模式识别和计算机视觉.

邹北骥 1961年生,教授,湖南大学计算机与通信学院副院长. 1984年获清华大学计算机系工学硕士学位,2001年获湖南大学工学博士学位. 主要研究方向为计算机图形学、数字图象处理及多媒体技术.

胡峰松 1969年生,讲师,湖南大学多媒体教育中心教学主任. 1999年获湖南大学计算机与通讯学院计算机应用专业工学硕士学位,现为湖南大学计算机与通讯学院博士研究生. 主要研究方向为数字图象处理及多媒体技术.