

一种重叠红细胞图像的分离方法

尤育赛 于慧敏

(浙江大学信息与电子工程学系, 杭州 310027)

摘要 细胞图像中经常出现重叠(或者粘连)现象,利用红细胞是近似圆形的这一先验知识,用粒度测量来获取红细胞半径 R ,对重叠区域轮廓上的每个非凹陷点,估计其所对应的圆心点,这些圆心点聚集成各个细胞的中心区域,对每一个中心区域,采用半径为 R 的圆形结构元进行一次膨胀操作,然后和原二值图像进行交集运算,将其结果作为对各个细胞形状的估计。实验结果表明,该方法能得到比较满意的分离效果,亦适用于其他近似圆形的细胞或者颗粒的重叠区域图像的分离。

关键词 细胞图像 图像分割 重叠 分离

中图分类号: TP391.41 Q2-33 **文献标识码**: A **文章编号**: 1006-8961(2005)06-0736-05

A Separating Algorithm for Overlapping Red Blood Cell Images

YOU Yu-sai, YU Hui-min

(Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027)

Abstract Overlapping appears frequently in all images. The algorithm based on the knowledge that red blood cells are disk-shaped. Granulometry is employed to estimate cell's radius R . Every contour point corresponds to center point of a circle with radius R estimated with the information of radius and chain code, which will assemble central areas of every single cell. Then, morphological dilation with a disk-shaped structuring element of radius R is applied to every central area. The intersection of the result of dilation operation and the original binary image is used to estimate the shape of every single cell. Experiments show the algorithm can obtain acceptable separating result. It can also be used for images of other kind of disk-shaped cells and granules.

Keywords cell image, image segmentation, overlapping, separating

1 引言

图像处理在生物学和医学领域正得到越来越多的应用,细胞图像自动识别分析系统便是其中一个比较典型的例子,它的主要任务就是自动地识别和分析细胞图像,提取出一个一个细胞,进而统计细胞数目和获取细胞特征信息。在实际处理中,经常会出现多个细胞粘连或者重叠在一起的现象,如何把它们分离成单个的细胞,是这些图像处理系统必须解决的问题。

常用的分离算法,很多是通过构造分离线(包括直线形式的和曲线形式的分离线)来隔开重叠或

者粘连的细胞。一些算法通过搜索凹点,然后进行配对构成分离点对,用连接分离点对的线段隔开相邻细胞^[1]。有一些分离方法采用距离变换和分水岭算法^[2]相结合^[3,4],把被处理图像二值化,进行距离变换得到距离灰度图,再用分水岭算法进行图像分割。还有一些方法采用数学形态学工具,比如:用连续腐蚀或者边界剥离来分裂重叠细胞区域,然后寻找分离点,构成分离线^[5,6];通过粒度测量获取细胞的尺寸,再计算重叠细胞图像的形态梯度,然后提取细胞的轮廓^[7,8]。

前述方法均是用分离线来隔开重叠或者粘连细胞的,这些基于构造分离线的方法有时会损害甚至丢

基金项目:浙江省科技厅资助项目(2003C11022)

收稿日期:2003-11-24;改回日期:2004-12-23

第一作者简介:尤育赛(1977~),男,1999年于大连海事大学获学士学位,现为浙江大学信息与电子工程学系硕士研究生。主要研究方向为图像与视频处理。E-mail:youyusai@sina.com

失细胞原来的一些形态信息。为此,本文算法不去构造分离线,而是尝试估计重叠红细胞图像中每一单独细胞在不发生重叠的情况下可能的原始形状。

2 算法设计

2.1 问题分析

算法利用了红细胞是近似圆形的这一先验知识。图 1 所示是一个半径为 R 的圆,从圆周上的任何点出发在圆内画一条和该点位置的切线相垂直并且长度为 R 的线段,则该线段的另一个端点必是圆心。根据这一特性,如果从红细胞的轮廓点出发画一条与轮廓点切线相垂直并且长度和红细胞半径相当的线段,那么该线段的另一个端点会在细胞中心位置附近。对每一个轮廓点都执行这样的操作,便

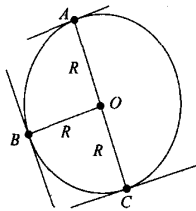


图 1 圆周点和圆心点关系示意图

Fig. 1 Relation of points on circle and circle's center

可以获得一定数量的点,它们能够勾画出各个单个细胞的中心区域。从这些中心区域的点出发来画圆,便可以估计出细胞的形状。这里,需要解决以下两个关键性问题:如何获取细胞的半径;如何从轮廓点出发来估计圆心。这两个问题,可以分别采用粒度测量和圆的几何关系运算来解决。

2.2 粒度测量

粒度测量(*granulometry*)^[7-9]是用来获取图像中颗粒尺寸分布的有效工具。通过采用尺寸逐渐增大的结构元对原图像进行一系列数学形态学开操作,计算出每一次开操作后的图像和原图像之间的差异,再对这些差异进行归一化处理,从而得到尺寸分布的直方图。

用 $s(k)$ ($k = 1, 2, \dots$) 表示一系列尺寸逐渐增大的结构元, $r_{s(k)}(f)$ 表示用结构元 $s(k)$ 对原始图像 f 进行开操作的结果, $A(k)$ 表示图像 $r_{s(k)}(f)$ 中物体像素点数目。

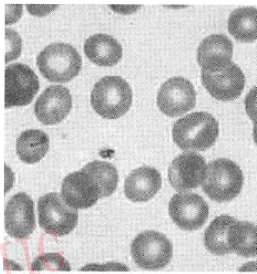
$$A(k) = \text{sum}(r_{s(k)}(f)) \quad (1)$$

对 $A(k)$ 进行归一化处理

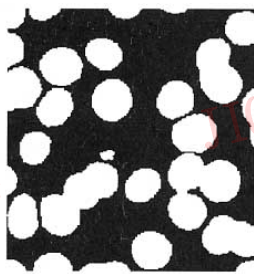
$$N(k) = 1 - \frac{A(k)}{\text{sum}(f)} \quad (k = 1, 2, \dots) \quad (2)$$

$$G(k) = N(k + 1) - N(k) \quad (k = 1, 2, \dots) \quad (3)$$

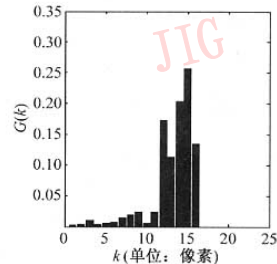
$G(k)$ 可以看成是图像结构尺寸的离散密度函数,根据它可以画出尺寸分布直方图,如图 2 所示。



(a) 红细胞灰度图像



(b) 预处理后图像



(c) 尺寸分布直方图

图 2 粒度测量示意图

Fig. 2 Illustration of Granulometry

图 2(c) 是采用半径 ($k = R$) 逐渐增加的圆形结构元对图 2(b) 进行粒度测量所得到的尺寸分布,横轴表示的是圆形结构元半径大小,从图中可以看出这幅图像中红细胞半径的取值范围为 12 ~ 16 像素。

2.3 估计圆心

为了从轮廓点来估计与之相对应的圆心点,需要获取该轮廓点的切线和法线的方向,可通过对链码的处理来实现。这里假定链码是通过顺时针方向跟踪轮廓点而获得的。假设轮廓点数目为 N , 用 C_i

表示起始点开始第 i 个轮廓点的链码,得到链码序列 C_1, C_2, \dots, C_N (其中 $0 \leq C_i \leq 7$)。用 D_i 表示第 i 个轮廓点的切线方向,其值用该轮廓点的链码 C_i 和前一个轮廓点的链码 C_{i-1} 的矢量和来近似(如图 3(a)), 第 1 个轮廓点的切线方向 D_1 采用 C_1 和 C_N 的矢量和表示。 D_i (其中 $i = 1, 2, \dots, N, 0 \leq D_i \leq 15$) 会把整个圆周分成间隔为 22.5° 的 16 等份,如图 3(b) 所示。由于采用顺时针方向跟踪链码,所以轮廓点所对应的圆心必定在切线方向的右侧,从切线方向 D_i

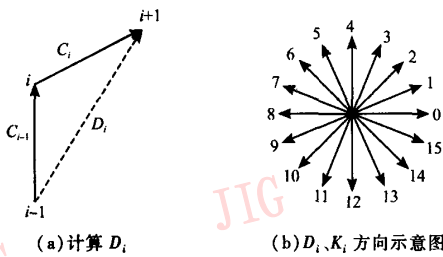


图 3 D_i 和 K_i 方向示意图

Fig. 3 Directions of D_i and K_i

再顺时针转 90° 便可以得到由轮廓点指向圆心的方向 K_i (图 4),

$$K_i = \begin{cases} D_i - 4 & 4 \leq D_i \leq 15 \\ D_i - 4 + 16 & 0 \leq D_i \leq 3 \end{cases} \quad (4)$$

和 D_i 一样, K_i (其中 $i = 1, 2, \dots, N, 0 \leq K_i \leq 15$) 也把圆周分成图 3(b) 所示的 16 等份。

用 $P_i(x_i, y_i)$ 表示第 i 个轮廓点, 其对应的圆心点为 $O_i(x_{oi}, y_{oi})$ 。如果以图像左下角为坐标轴原点, 建立图 4 所示直角坐标系, 则根据几何关系可以得到

$$(x_i - x_{oi})^2 + (y_i - y_{oi})^2 = R^2 \quad (5)$$

$$(y_i - y_{oi}) / (x_i - x_{oi}) = \tan(2\pi K_i / 16) \quad (6)$$

$(K_i \neq 4, 12)$

$$\begin{cases} x_{oi} = x_i \\ y_{oi} = y_i + R \end{cases} \quad (K_i = 4) \quad (7)$$

$$\begin{cases} x_{oi} = x_i \\ y_{oi} = y_i - R \end{cases} \quad (K_i = 12) \quad (8)$$

由式(5)和式(6)又可以推导出

$$\begin{cases} x_{oi} = x_i + \frac{R}{\sqrt{1 + \tan^2(K_i \times 2\pi/16)}} \\ y_{oi} = y_i + \frac{R \times \tan(K_i \times 2\pi/16)}{\sqrt{1 + \tan^2(K_i \times 2\pi/16)}} \end{cases} \quad (9)$$

$(K_i = 0, 1, 2, 3, 13, 14, 15)$

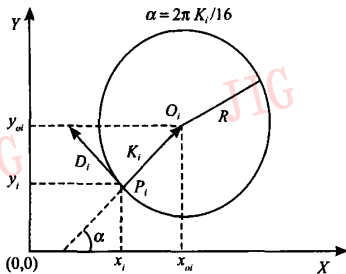


图 4 D_i 和 K_i 关系示意图

Fig. 4 Relation of D_i and K_i

$$\begin{cases} x_{oi} = x_i - \frac{R}{\sqrt{1 + \tan^2(K_i \times 2\pi/16)}} \\ y_{oi} = y_i - \frac{R \times \tan(K_i \times 2\pi/16)}{\sqrt{1 + \tan^2(K_i \times 2\pi/16)}} \end{cases} \quad (10)$$

$(5 \leq K_i \leq 11)$

用式(7)~式(10)便可以估计出每一个轮廓点所对应的圆心坐标。

事实上不需要对每一个轮廓点都来估计它的圆心, 而且对那些可能是凹点的轮廓点也不应该去执行这一操作。在采用顺时针方向跟踪链码的情况下, 对第 i 个轮廓点, 如果它的链码 C_i 相对于前一个轮廓点的链码 C_{i-1} 逆时针旋转了 $45^\circ \sim 135^\circ$, 则认为它可能是凹点, 从而不去估计它所对应的圆心点。

这种计算切线方向和判断凹点的方法, 尽管精度不是非常高, 但是实现起来简单, 计算量小, 而且在本文算法中能得到的结果。如图 5 所示, 对图 5(a) 采用该方法得到的圆心点和原图轮廓叠加后得到的效果如图 5(b), 颗粒的中心区域已经被勾画了出来。



(a) 二值图

(b) 轮廓和圆心点

图 5 圆心估计效果示意图

Fig. 5 Result of estimating circle's center

2.4 算法原理

用粒度测量来获取图像中红细胞的半径 R , 跟踪重叠细胞区域的轮廓并提取链码; 根据半径和链码来估计轮廓点所对应的圆心点, 这些点会集中在各个细胞的中心区域, 对它们进行分类聚合并加以标记, 可以得到由一定数量的圆心点聚集而成的各个单个细胞的中心区域(一般会包含了多个圆心点); 用半径为 R 的圆形结构元对获取的每一个中心区域进行一次膨胀操作, 然后和原重叠区域图像进行一次交集运算, 其结果作为对每个细胞形状的估计。

2.5 算法实现流程

(1) 对红细胞图像进行二值化处理, 填充细胞内部的空洞, 然后采用小的圆形结构元进行开操作, 从而完成去噪和圆滑细胞轮廓的目的, 得到二值图像 f_1 。

(2) 对 f_1 进行粒度测量, 获取细胞半径 R 。

(3) 从 f_1 中提取出所有重叠区域构成重叠区域

二值图像 f_2 , 用半径为 R_1 的圆形结构元对 f_2 进行腐蚀得到 f_3 。 R_1 的值根据 R 来确定, 比 R 稍小。

(4) 对 f_2 中的每一个重叠区域, 执行下面的步骤 5~9。

(5) 跟踪重叠区域轮廓并提取链码。

(6) 根据链码和半径 R , 求取对应轮廓点的圆心点, 聚集成各个细胞的中心区域。

(7) 圆心点构成的图像和 f_3 进行交集运算, 从而删除那些偏离细胞中心较远的圆心点, 得到图像 f_4 。

(8) 设置阈值 T , 对 f_4 中的圆心点进行分类聚合并加以标记。对每个圆心点检查以它为圆心半径为 T 的圆形区域中被标记点的数目。如果存在唯一一个被标记的点, 则圆心点采用与该点相同的标记, 表示属于同一个细胞; 如果没有找到被标记的点, 则为圆心点分配一个新的标记, 代表一个新的细胞; 如果发现多个被标记的点, 则采用距离最近的标记点的标记。从而, 得到由一定数量的圆心点聚集而成的各个细胞中心区域。在所有的圆心点被处理完之后, 便可以知晓细胞的数目。阈值 T 根据半径 R 来选择, 可以选择一个和 R 相当的值, 或者直接令 $T=R$ 。

(9) 对每个细胞中心区域, 用半径为 R 的圆形结构元进行一次膨胀操作, 然后与 f_2 作交集运算, 其结果便作为对该中心区域所对应细胞的形状估计。

3 实验结果及讨论

由图 2 提取出重叠细胞区域图像如图 6 所示。图 7(a) 是图 6 中的一个重叠区域, 提取它的链码估计圆心点, 得到图 7(b) 的结果, 在此细胞半径 R 选择为 12 像素, R_1 为 10 像素。令 $T=R$, 从图 7(b) 中可以确定 3 个细胞的中心区域, 最后分离出图 7(c)、图 7(d)、图 7(e) 3 个单独的红细胞。图 8~图 10 是图 6 其他 3 个重叠区域的分离结果。图 11 所示为使用 Matlab 工具箱, 对图 6 进行距离变换, 再用分水岭算法进行图像分割的结果^[3,4]。相比较而言, 本文算法所得到的分离结果是比较满意的。

为了克服噪声的影响, 使得到的圆心点尽量集中在细胞中心, 算法中采用了以下的措施:

(1) 在预处理阶段, 采用小的圆形结构元对二值图像进行开操作, 达到去噪和圆滑细胞轮廓的目的。

(2) 估计轮廓点所对应圆心点的操作, 并不是

对所有的轮廓点均执行, 而是先判断是否凹点, 如果可能是凹点, 则不去估计圆心点。从而使得估计出



图 6 重叠细胞区域图

Fig. 6 Image of overlapping cells

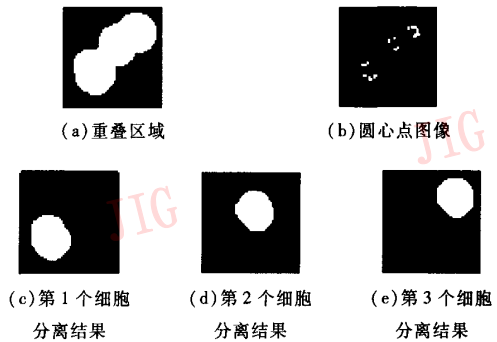


图 7 第 1 个重叠区域的分离效果

Fig. 7 Results of the first overlapping area

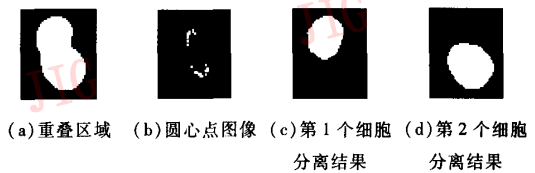


图 8 第 2 个重叠区域的分离效果

Fig. 8 Results of the second overlapping area

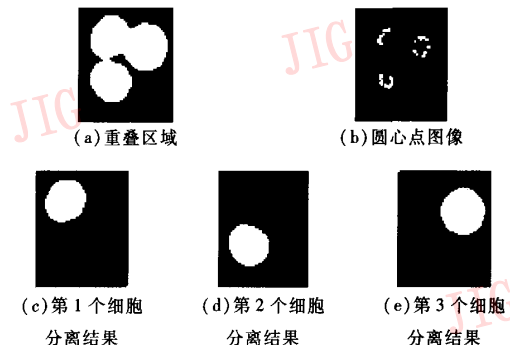


图 9 第 3 个重叠区域的分离效果

Fig. 9 Results of the third overlapping area

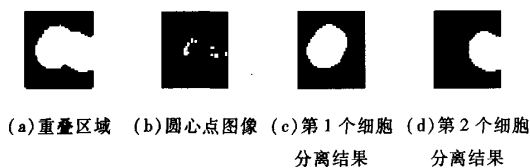


图 10 第 4 个重叠区域的分离效果

Fig. 10 Results of the fourth overlapping area

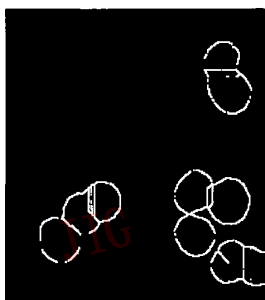


图 11 基于距离变换和分水岭算法的分割结果

Fig. 11 Result of segmentation using watershed algorithm

来的圆心点尽量聚集在细胞的中心。

(3)通过 2.5 节步骤 7 的交集运算,可以进一步删除那些偏离细胞中心较远的圆心点。

本文算法的关键是估计轮廓点对应的圆心,当重叠区域由比较多的细胞构成时,只要细胞的外轮廓有一段处在重叠区域的外轮廓中(如图 6 所示),便可以通过轮廓点来估计圆心。对某一个细胞来说,实际的分离效果跟能够获取的该细胞的轮廓长短有关,当获取的轮廓点数目越多时,一般会得到较好的效果。如果细胞全部处于重叠区域外轮廓之内,则无法获取该细胞的任何轮廓点,从而该算法失效。

在前面的讨论中是假设被处理图片中细胞大小比较均一,而实际应用中可能存在图像中包含几种大小不同的细胞。此时,采用粒度测量可以得到多个半径值 $R^{[9]}$,每一个 R 对应一种细胞,这时,可从最大的 R 值开始,对每一个 R 值执行 2.5 节中的步骤 5~9,当某一个 R 值被处理完之后,从图像中删除该 R 值所对应的细胞,然后去处理更小的 R 值,直到所有的 R 值均被处理完为止。

4 结 论

针对红细胞图像中可能存在细胞重叠的现象,本文提出了一种分离方法,估计重叠区域中各个红细胞

在不发生重叠的情况之下可能的原本形状。利用了红细胞是近似圆形这一先验知识,用粒度测量来估计红细胞的半径,然后利用圆的几何关系来估计轮廓点所对应的圆心,从而得到由一定数量的圆心点聚集而成的各个细胞的中心区域,最后用一个和细胞大小相当的圆形结构元对每个中心区域进行一次膨胀操作,把其结果和原来图像进行交集运算,从而得到对各个细胞形状的估计。实验表明,该方法能够得到比较满意的分离效果,其不仅对红细胞图像有效,对其他近似圆形的细胞或者颗粒的图像亦适用。

参考文献 (References)

- Lu Zong-qi, Tong Tao. The application of chain code sum in the edge form analysis[J]. Journal of Image and Graphics A, 2002, 7(12): 1323 ~ 1328. [陆宗琪,童韬. 链码和在边界形状分析中的应用[J]. 中国图象图形学报(A版), 2002, 7(12): 1323 ~ 1328.]
- Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(6): 583 ~ 598.
- Liu Sheng-hao, Zeng Li-bo, Liu Bing, et al. Separating algorithm for overlapping granule images [J]. Computer Engineering, 2002, 28(2): 198 ~ 210. [刘生活,曾立波,刘斌等. 重叠颗粒图象的分离[J]. 计算机工程, 2002, 28(2): 198 ~ 210.]
- Castleman K R. Digital Image Processing [M]. Beijing: Tsinghua University Press, 1998: 475 ~ 477. [Castleman K R. 数字图像处理(影印版)[M]. 北京:清华大学出版社, 1998: 475 ~ 477.]
- Lu Jian-feng, Yang Jing-yu, Tang Zhen-min, et al. Design of a separating algorithm for overlapping cell images [J]. Journal of Computer Research & Development, 2000, 37(2): 228 ~ 232. [陆建峰,杨静宇,唐振民等. 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 ~ 232.]
- Liu Xiang-bin, Zou Bei-ji, Hu Feng-song. Separating algorithm for cell image based on boundary-stripped [J]. Journal of Image and Graphics(A), 2002, 7(3): 234 ~ 239. [刘相滨,邹北骥,胡峰松. 基于边界剥离的细胞图象分离算法[J]. 中国图象图形学报(A版), 2002, 7(3): 234 ~ 239.]
- Ruberto C D, Dempstr A, Khan S, et al. Segmentation of blood images using morphological operators [A]. In: Proceedings of IEEE 15th International Conference on Pattern Recognition [C], Barcelona Spain, 2000, 3: 401 ~ 404.
- Ruberto C D, Dempstr A, Khan S, et al. Analysis of infected blood cell images using morphological operators [J]. ELSEVIER Image and Vision Computing, 2002, 20(2): 133 ~ 146.
- Gonzalez R C, Woods R E. Digital Image Processing (2nd ed) [M]. Beijing: Publishing House of Electronics Industry, 2002: 559, 620 ~ 622. [Gonzalez R C, Woods R E. 数字图像处理(第二版,英文版) [M]. 北京:电子工业出版社, 2002: 559, 620 ~ 622.]