

一种粘连血细胞图像分割新方法

印 勇, 李阿琼

YIN Yong, LI A-qiong

重庆大学 通信工程学院, 重庆 400044

College of Communication Engineering, Chongqing University, Chongqing 400044, China

E-mail: yongyin@cqu.edu.cn

YIN Yong, LI A-qiong. New approach for segmentation of overlapping blood cell images. *Computer Engineering and Applications*, 2009, 45(35): 173-175.

Abstract: To segment the overlapping blood cells, an approach based on the characteristic that the blood cells are disk-shaped is proposed. A distance map and the edge of image can be obtained by morphological erosion, while the regional centers and centers of circle can be deserved with grayscale reconstruction algorithm. After all of these, the edge of the blood cells is searched based on the circular characteristics to find the separating points and separating line, then the overlapping blood cells are segmented into singles accurately. The experimental results show that the proposed approach is feasible and effective.

Key words: overlapping cells; image segmentation; disk-shaped characteristic; separating points; separating line

摘 要: 提出一种基于血细胞近圆形特征的粘连血细胞分割方法。用形态学腐蚀得到距离图和图像的边缘, 并利用灰度重建得到细胞的中心区域和圆心, 在此基础上基于细胞圆形特征对细胞的边缘进行搜索, 得到粘连细胞的分离点, 从而求得分离线, 完成对粘连血细胞的精确分割。实验表明, 该分离方法是有效可行的。

关键词: 粘连细胞; 图像分割; 圆形特征; 分离点; 分离线

DOI: 10.3778/j.issn.1002-8331.2009.35.052 **文章编号:** 1002-8331(2009)35-0173-03 **文献标识码:** A **中图分类号:** TP391.4

1 引言

随着计算机模式识别技术及人工智能研究的不断发展, 计算机在医学图像处理中发挥了越来越重要的作用, 它为准确、方便、定性、定量地识别不同类型的血细胞提供了新的手段, 使得图像分析逐步成为细胞学定量分析研究的有力工具。很多临床症状表现为白细胞数目增多、肿大或者白细胞中细胞核形状与数目发生变化, 所以对白细胞数目和形状的研究有着重要意义。但是由于染色条件、涂片制备、图像来源、采样光照条件的差异以及细胞间相互重叠、粘连情况的发生, 使得对白细胞的计数和识别等后续分析变得困难。为此, 将粘连细胞分割成为单个细胞, 就成为医学图像处理中必须解决的问题。

目前研究细胞粘连分割的方法较多, 游迎荣, 范影乐等^[1]利用距离变换与分水岭方法相结合, 把处理过的二值图像转化为依据一定规律变化的灰度图, 再进行分水岭分割, 分割速度快, 但只适用于粘连不是很严重的情况。马东, 曹培杰等^[2]采用数学形态学中的膨胀、腐蚀等对图像进行处理, 先对图像进行腐蚀得到中心区域, 再选择合适的结构元进行膨胀, 从而得到分离结果, 实现起来比较容易, 但只考虑了灰度信息, 未能利用空间分布信息、纹理信息等, 存在误分割。阳波, 王卫星等^[3]用细化算法提取细胞图像的骨架, 并计算骨架各点的边界距离值, 利用骨架边界距离函数波谷的特性将相互粘连的细胞分离, 该算法

提高了细胞分离的精度, 并且具有速度快、抗干扰性强等特点, 但对于距离函数的极小值点的精确求取比较复杂。总的来说, 每一种方法都有其优势和局限性。

通常, 血细胞呈圆形出现, 并且在粘连细胞的粘连处出现凹陷。提出基于细胞的这种特征, 利用相交圆的特性, 查找细胞的凹陷点, 也即分离点, 对粘连细胞进行精确分离。

2 求细胞圆心及半径

2.1 求距离图和图像边缘

对于一幅原始的血细胞图像, 通常包含白细胞、红细胞和背景这三部分, 选用 HIS 模型中的 S 分量和 RGB 中的 G 分量以及区域生长方法得到白细胞, 并对白细胞用面积判断是否为粘连, 选取粘连白细胞进行处理。这时的处理对象即是一个以粘连白细胞为对象的二值图像。

数学形态学是一个非线性的图像处理和分析理论, 它既可以用于二值图像, 也可以用于灰度图像, 并且特别适于图像的形状分析^[4]。在数学形态学中, 膨胀和腐蚀是最基本的运算, 其他许多形态学算法都是以这两种原始运算为基础的。 A 和 B 是 Z^2 中的集合, A 被 B 膨胀定义为: $A \oplus B = \{z | (B)_z \cap A \neq \Phi\}$, A 被 B 膨胀是所有位移 z 的集合, 这样, \hat{B} 和 A 至少有一个元素是重

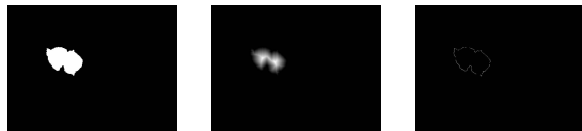
基金项目: 重庆大学“211 工程”三期创新人才培养计划建设项目(No.S-09102)。

作者简介: 印勇(1963-), 男, 博士, 副教授, 研究方向为图像处理与识别; 李阿琼(1984-), 女, 硕士研究生, 研究方向为图像处理与识别。

收稿日期: 2009-07-14 修回日期: 2009-09-16

叠的。对 Z 中的集合 A 和 B , 使用 B 对 A 进行腐蚀, 用 $A \ominus B$ 表示, 并定义为: $A \ominus B = \{z | (B)_z \subseteq A\}$, 这个公式表明, 使用 B 对 A 进行腐蚀是所有 B 中包含于 A 中的点 z 的集合用 Z 平移^[5]。

使用单位圆形结构元 B 对二值图像(图 1(a))进行层层腐蚀, 同一层上的点与背景的距离应该是相等的, 不断腐蚀, 用第 n 次的腐蚀结果减去第 $n+1$ 次的腐蚀结果, 再乘以 n , 即为这一层上的距离值, 距离值不断增加, 直至区域为空^[6]。由此, 可得到图像的距离图(图 1(b)), 在距离图上找到距离值为 1 的点, 即为图像的边缘(图 1(c))。



(a) 粘连细胞二值图像 (b) 距离图 (c) 粘连细胞边缘
图 1 距离图和粘连细胞边缘

2.2 灰度重建

将距离图像 IM 的灰度值减去一个幅度 N 得到图像 JM , 即原图的每个点均在定义域范围内减去幅度 N , 然后用 JM 代替 IM 进行灰度重建, 得到的图像可以保持 IM 原有的形态, 而消除幅度小于 N 的微小差异。灰度重建还可以消除微小波动和噪声^[7]。

这里, 利用灰度重建的部分原理, 得到粘连细胞各自的中心区域。如图 2 所示。在此基础上在各自区域上求距离灰度最大值, 即为所要求的单个细胞的半径。同时, 在各自区域上求得坐标的均值, 即可看做两个圆的圆心。



(a) 圆心区域 1 (b) 圆心区域 2
图 2 圆心区域

3 近圆形特征求细胞分离

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。对于分离点, 一般应满足以下性质: 处于粘连细胞的连接处; 它们分别处于中轴的两侧; 分离点直接连线的距离是局部最小的^[8]。为此, 利用血细胞的近圆形特征对分离点的位置进行分析。

在数学中, 两圆相交(如图 3 所示), 其交点之间的连线即为两圆的公共弦 L_1 , 理论上, 两圆的交点即为两圆的分离点, 公共弦 L_1 即为分离线。设两圆圆心之间的连线为 L_2 , 半径 R_1, R_2 已知, 圆心 O_1, O_2 坐标也已求出, 分别为 $(x_1, y_1), (x_2, y_2)$ 。设 L_1 与 L_2 交点 P 的坐标为 (x, y) , 由圆及公共弦的特性, 可以得到以下公式:

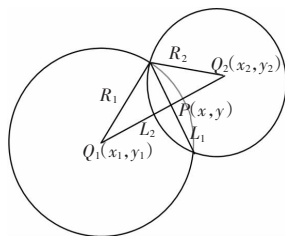


图 3 相交圆几何关系示意图

$$R_2^2 - (\sqrt{(y_2 - y)^2 + (x_2 - x)^2})^2 = R_1^2 - (\sqrt{(y_1 - y)^2 + (x_1 - x)^2})^2 \quad (1)$$

$$\frac{y_2 - y}{x_2 - x} = \frac{y - y_1}{x - x_1} \quad (2)$$

由公式(2)得:

$$y = \frac{y_1 \times (x - x_2) - y_2 \times (x - x_1)}{x_1 - x_2} \quad (3)$$

将(3)带入公式(1), 可得:

$$k = (y_1 - y_2)^2 + (x_1 - x_2)^2$$

$$x = \frac{k \times (x_1 + x_2) - (R_1^2 - R_2^2) \times (x_1 - x_2)}{2 \times k}$$

$$y = \frac{k \times (y_1 + y_2) - (R_1^2 - R_2^2) \times (y_1 - y_2)}{2 \times k}$$

由此求出 L_1, L_2 交点 P 的坐标 $P(x, y)$, 且 P 点必位于圆心 O_1, O_2 之间。

对于血细胞边缘点, 使用 L_2 直线可以将其分为两部分, 记为 $edge_1$ 和 $edge_2$ 。分别遍历 $edge_1$ 和 $edge_2$ 中的每一个边缘点, 计算其与点 P 的距离, 则 $edge_1$ 和 $edge_2$ 中必有且分别只有一个点满足距离最小, 则这两个点就是所要求的分离点。

使用该方法, 产生两个位于圆心连线 L_2 两侧的分离点, 克服了传统的分割方法中凹点不能精确求取、凹点对困难等问题, 因而该方法分割精确有效。

下一步采用 Bresenham 算法生成两分离点之间的直线段, 如图 4(a), 此直线段取反再与原二值图像进行与运算即可得出所要的结果, 如图 4(b)所示。



(a) Bresenham 生成直线段 (b) 分割结果
图 4 分割线及分割结果

4 算法步骤

根据上述分析与讨论, 要完成该算法已经很容易了, 限于文章篇幅, 简述如下:

- (1) 原始细胞图像经过处理后产生白细胞二值图, 对象为 1, 背景 0;
- (2) 形态学腐蚀, 层层剥离产生距离图, 并可得到细胞边缘;
- (3) 灰度重建, 产生两细胞核心区域, 并计算出半径 R_1, R_2 , 圆心坐标 O_1, O_2 ;
- (4) 数学运算, 求得交点 P 的坐标, 遍历 L_2 两侧细胞边缘像素, 求使得点距离最小的点, 即为要求的分离点;
- (5) 用 Bresenham 算法做两分离点之间的分离线, 与原图进行逻辑运算即可得到分离了的血细胞图像。

5 实验结果及讨论

与使用基于距离变换的分水岭算法进行比较, 距离图的求取是一样的。由于分水岭算法对目标区域的微弱边缘比较敏感, 使得它能够精确定位粘连细胞的微弱边缘的同时也容易造成过分割, 另外对种子点的选取也不好确定, 对同一区域取不同的种子点, 得到的分割结果可能就有很大差别。而采用提出的分割方法, 直接在血细胞近圆形特征的基础上得到分离点和

分离线,分离点对存在且唯一,不存在凹点不能精确配对等问题。典型的分割效果对比如图5所示,其中,(a)为原二值图,(b)为采用分水岭对粘连细胞分割的结果,(c)为采用该方法分割的结果,经过比较可以看出算法在分割粘连细胞的效果上比使用分水岭的方法要好。

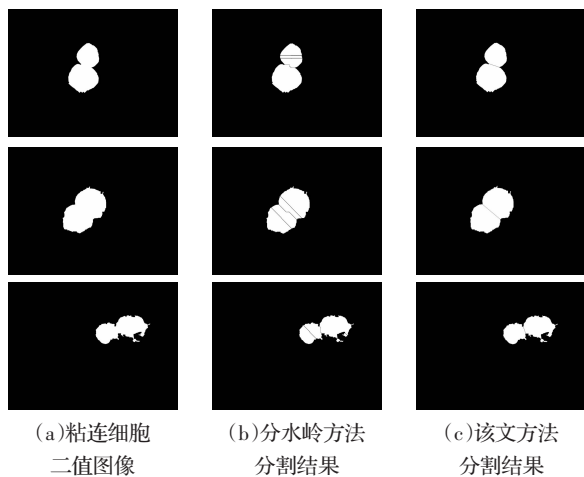


图5 粘连细胞分割结果

6 结论

分水岭算法很大程度上依赖于如何进行图像变换以得到更适合分水岭分割的图像^[9]。而由于图像粘连情况的多样性,要找到这样一种适合所有情况的图像变换就很困难了。该方法提出结合细胞的近圆形特征,利用圆的边缘和相交弦的关系,精

(上接 152 页)

4 仿真实验

实验采用了UCI机器学习知识库中的Heart Disease Database和Breast Cancer数据集以及一个实际的蒸汽涡轮机故障诊断的样本集^[9](<http://ftp.ics.uci.edu/pub/machine-learning-databases/>),具体描述如表1所示。属性约简处理采用由挪威科技大学计算机与信息科学系和波兰华沙大学研究所合作开发Rosetta软件包。支持向量机处理采用Libsvm软件包。

表1 实验数据集具体描述

数据集	样本数	条件属性	类别
Heart Disease	303	13	2
Breast Cancer	699	9	2
Fault Diagnosis	21	11	2

分别进行3组实验T1、T2和T3。T1使用单一的SVM进行训练和测试,T2首先对训练集进行属性约简,利用所有得到的约简集构成成员支持向量机,并通过多数投票策略构成集成系统,T3是利用该文设计的选择性集成方法。表2给出了各实验所得结果。

表2 不同方法的实验结果 (%)

Dataset	Single SVM(T1)	Ensemble all SVM(T2)	Method in this paper(T3)
Heart Disease	82.61	83.91	86.21
Breast Cancer	87.23	95.55	96.54
Fault Diagnosis	73.30	78.10	81.70

实验中核函数统一采用高斯核函数 $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$,参数的选取采用网格搜索获得。

确定位出分离点,再在此基础上做出分离线,从而实现粘连细胞的较精确分割,原理简单,运算速度快,且由于使用了腐蚀、灰度重建等过程,使得算法受噪声影响小,具有一定的鲁棒性。

后续的研究将进一步考虑如何对粘连情况严重的细胞,尤其是两细胞中心距离小于细胞半径的情况进行分割,使其具有更好的鲁棒性和普适性。

参考文献:

- [1] 游迎荣,范影乐,庞全.基于距离变换的粘连细胞分割方法[J].计算机工程与应用,2005,41(20):206-208.
- [2] 马东,曹培杰,潘凯丽,等.分割重叠细胞核的方法及比较研究[J].北京生物医学工程,1999,18(3):142-147.
- [3] 阳波,王卫星,李彦文.基于骨架的血细胞图像分离算法[J].计算机工程与应用,2003,39(16):94-96.
- [4] 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.
- [5] Gonzalez R C, Woods R E. Digital image processing[M]. Beijing: Publishing House of Electronics Industry, 2000.
- [6] 卢艳芝,刘相滨.一种改进的重叠细胞分离算法[J].计算机工程与应用,2008,44(24):172-174.
- [7] 丛培盛,孙建忠.分水岭算法分割显微图像中重叠细胞[J].中国图象图形学报,2006,11(12):1781-1783.
- [8] 陆振晔,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.
- [9] 丁伟杰,范影乐,庞全.一种改进的基于分水岭算法的粘连分割研究[J].计算机工程与应用,2007,43(10):70-72.

5 结论

论文通过对学习样本的属性约简,获得了多个变化的成员分类器。同时对各成员分类器的差异性和分类精度进行了对比研究,有效地构造了一种选择性集成系统。实验仿真结果表明该方法能够有效提升支持向量机的推广能力。

参考文献:

- [1] Cristianini N, Shawe-Taylor J. An introduction to Support Vector Machines and other kernel-based learning methods[M]. [S.l.]: Cambridge University Press, 2000.
- [2] Dietterich T. Machine-learning research: Four current directions[J]. The AI Magazine, 1998, 18: 97-136.
- [3] 张文修,吴伟志.粗糙集理论与方法[M].北京:科学出版社,2005.
- [4] Zhou Z H, Wu J, Tang W. Ensemble neural networks: Many could be better than all[J]. Artificial Intelligence, 2002, 137(1/2): 239-263.
- [5] Sanchez A V D. Advanced support vector machines and kernel methods[J]. Neuro-Computing, 2003, 55: 5-20.
- [6] Hsu Chih-Wei, Chang Chih-Chung, Lin Chih-Jen. A practical guide to Support Vector Classification[EB/OL]. <http://home.eng.iastate.edu/~julied/classes/ee547/Handouts/SVMuseguide.pdf>.
- [7] Kuncheva L, Whitaker C. Measures of diversity in classifier ensembles and their relationship with ensemble accuracy[J]. Machine Learning, 2003, 51(2): 181-207.
- [8] Brow G, Wyatt J, Harris R, et al. Diversity creation methods: A survey and categorisation[J]. Journal of Information Fusion, 2005, 1: 1-28.
- [9] Li Ye, Cai Yun-Ze, Li Yuan-Gui, et al. Rough set method for SVM data processing [C]/Proceeding of the 2004 IEEE Conference on Cybernetics and Intelligent System Singapore, 1-3 December, 2004.