

文章编号: 2095-2163(2020)05-0114-04

中图分类号: TP391

文献标志码: A

基于图像处理的异常细胞检测方法研究

陈紫薇, 原松梅

(哈尔滨工业大学 建筑学院, 哈尔滨 150001)

摘要: 将计算机图像处理技术应用于生物医学领域,能极大地提高病患病情诊断的效率和精准性,对病患的治疗及其生活质量的提高都具有现实意义。医学细胞图像处理是一个热门且现实的课题,本文提出应用形态学开闭运算对给定的医学图像进行滤波,运用大津法对图像进行分割,通过区域生长算法实现对连通区域编号,对掩模部分限定的数据进行处理,最终得到细胞核占整体的比例,进而对异常细胞进行判定。

关键词: 图像处理; 医学细胞图像; 细胞检测; 图像分割

Research on abnormal cell detection method based on image processing

CHEN Ziwei, YUAN Songmei

(School of Architecture, Harbin Institute of Technology, Harbin 150001, China)

[Abstract] The application of computer image processing technology in biomedical field can greatly improve the efficiency and accuracy of disease diagnosis, which is of practical significance to the treatment of patients and the improvement of their life quality. Medical cell image processing is a hot and realistic subject, this article put forward the application of morphological opening and closing operations for a given medical image filtering, using OTSU method of image segmentation, by region growing algorithm for connected area number, limit on a mask part of data processing, finally obtained the nucleus, the proportion of the total, in turn, to determine abnormal cells.

[Key words] Image processing; Medical cell images; Cell detection; Image segmentation

0 引言

细胞是组成人体结构的基本单位,在医学检测中,细胞的检测结果是诊疗病情的重要依据。目前的细胞检测方法有使用薄层制片技术的核 TBS 诊断系统,但是更多的还是依靠专家阅片,判断标准往往根据医生的医术水平和个人意见而定,因此具有较强的主观性和不确定性^[1]。细胞核是控制细胞生命活动的中心,异常细胞的细胞核一般存在形态变化,核占比会发生变化。

计算机图像处理技术的快速发展使其应用延伸到各个领域,将计算机图像处理技术应用于生物医学领域能极大地提高病情诊断的精准性,提高医用显微技术处理分析的效率^[2]。医学细胞图像处理是医用图像处理领域一个热门的课题。本文提出通过形态学运算、图像分割等图像处理相关操作,计算细胞核占比,进而判断细胞是否存在异常。

1 图像处理相关算法

1.1 形态学运算

腐蚀操作与膨胀操作是形态学的两个基本运算,腐蚀与膨胀组合又会形成开运算和闭运算。腐

蚀能够造成图像边缘收缩,可以用来消除小的目标物。腐蚀的原理如图 1 所示。

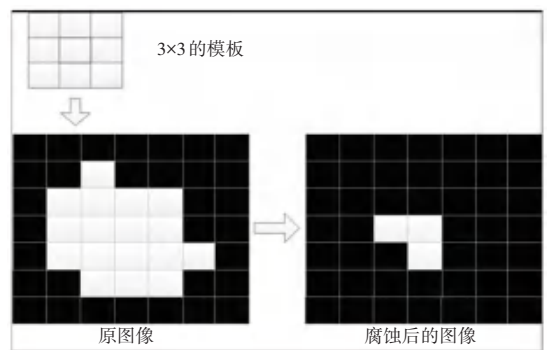


图 1 腐蚀操作原理图

Fig. 1 Schematic diagram of corrosion operation

使用白色 3×3 模板,对白色目标区域进行遍历操作,背景黑部分为(0),背景白部分为(1),模板与覆盖区域做“与”操作,若结果全为 1,则腐蚀后像素点为白,否则为黑,黑色的部分更易获得,最终表明白色区域收缩。

膨胀能够将目标的边缘向外扩张,可用于填补区域孔洞及消去区域的细小颗粒噪声。

膨胀的原理如图 2 所示。

作者简介: 陈紫薇(1998-),女,本科生,主要研究方向:图像处理与计算机视觉;原松梅(1960-),女,副教授,硕士生导师,主要研究方向:工程美学与数字仿真的应用研究、新媒体人机界面的相关研究。

收稿日期: 2020-03-13

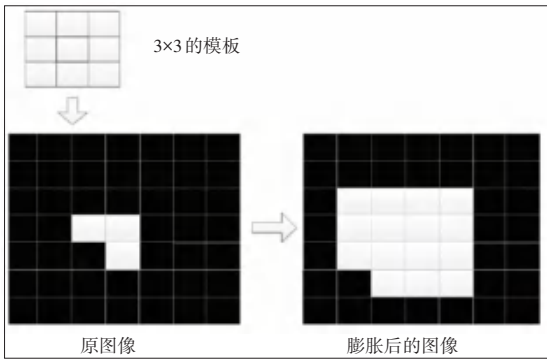


图 2 膨胀操作原理图

Fig. 2 Schematic diagram of expansion operation

膨胀操作同样也是使用 3×3 的白色模板对白色目标区域进行操作遍历操作,背景黑部分为(0),背景白部分为(1),模板与覆盖区域做“与”操作,若结果不全为 0,则膨胀后相应点为 1,最终表明白色目标区域明显扩大了。

开运算是先对图像做腐蚀操作再做膨胀操作的组合运算。开运算能放大空洞,消除小物体。开运算的原理如图 3 所示。

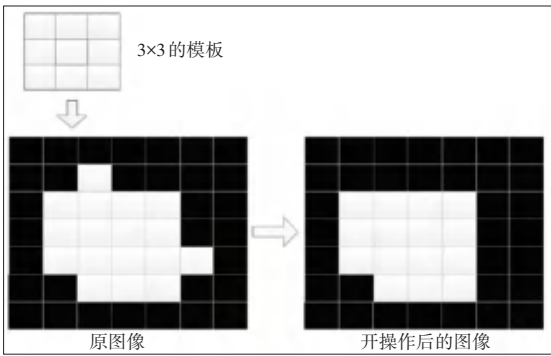


图 3 开运算操作原理

Fig. 3 Operating principle of open operation

闭运算是先对图像做膨胀操作再做腐蚀操作的组合运算。闭运算能填补小型空洞,联通两个区域^[3]。闭运算的原理如图 4 所示。

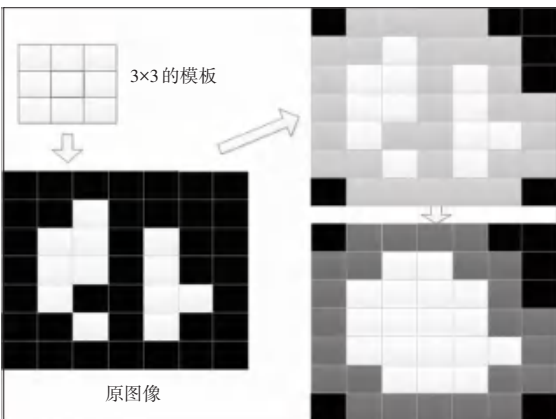


图 4 闭运算操作原理

Fig. 4 Operation principle of closed operation

顶帽操作是基于开运算的操作,使用开运算将峰顶消去,再用原图与开运算后的操作结果做异或运算,进而得到被去掉的部分。顶帽操作就是开运算消去图像中的较亮区域,即封顶,因此也叫白色顶帽。

1.2 图像分割算法

图像分割就是将图像分割成不同区域,特性包括:灰度、颜色和纹理等,最终获得一个或多个感兴趣区域的技术。图像分割方法可分为:阈值分割,边缘分割和区域分割^[4]。

基于阈值的分割是依据阈值对不同物体进行分割。最大类间方差法,也称大津法,是公认的无需手动选取阈值的最佳方法。大津法的基本思路是,首先设一个能将图像分为两组的阈值 K ,能够使两组类间方差最大的 K 值即为所求分割阈值。大津法的具体原理为:对于 $M \times N$ 大小的图像 $f(x, y)$,灰度级取值范围为 $[0 \sim L - 1]$,记 $p(k)$ 为灰度值 k 的出现频率,则有:

$$p(k) = \frac{1}{MN} \sum_{f(i,j)=k} 1, \quad (1)$$

设用灰度值 t 为阈值分割前景与背景,设属于前景的像素占比为 $w_0(t)$,则前景部分所占比例,即灰度级 t 之前的所有灰度级百分比之和为:

$$w_0(t) = \sum_{0 \leq i \leq t} p(i), \quad (2)$$

设前景像素个数 N_0 ,总图像像素个数与前景部分所占比例之积为:

$$N_0 = MN \sum_{0 \leq i \leq t} p(i), \quad (3)$$

设背景部分像素占比为 $w_1(t)$,则背景部分所占比例,即灰度级 t 至 m 之间的所有灰度级百分比之和为:

$$w_1(t) = \sum_{t < i \leq m-1} p(i), \quad (4)$$

设背景部分像素个数为 N_1 ,为总图像像素个数与背景部分所占比例之积:

$$N_1 = MN \sum_{t < i \leq m-1} p(i), \quad (5)$$

前景灰度均值为 t 之前所有灰度级的加权平均值为:

$$\mu_0(t) = \sum_{0 \leq i \leq t} ip(i)/w_0(t), \quad (6)$$

背景灰度均值为灰度级 t 至灰度级 m 之间所有灰度级的加权平均值:

$$\mu_1(t) = \sum_{t < i \leq m-1} ip(i)/w_1(t), \quad (7)$$

总灰度均值为前景灰度均值与背景灰度均值的

加权平均值:

$$\mu = w_0(t)\mu_0(t) + w_1(t)\mu_1(t). \quad (8)$$

因此,最佳阈值为:

$$g = \text{Max}[w_0(t) (\mu_0(t) - \mu)^2 + w_1(t) (\mu_1(t) - \mu)^2]. \quad (9)$$

由公式(9)可知,组间方差越小,则组内的相约越相似,组间方差越大,则两组差别越大,组间方差与组内方差之比越大,则分割效果越好。因此使组间方差与组内方差之比最大的 t 值即为所求阈值。大津法的缺点是不能反映图像的空间结构,因此有时其分割结果与人眼判断存在不同^[5]。

基于边缘的分割算法主要依靠边缘检测算法,通过获取并计算图像边缘灰度级或是结构突然改变之处进行图像的分割。边缘检测算法的基本思想是:先检测不连续的点,然后将点连接成边,先确定边缘像素,并连接这些像素,进而构成边界。实现边缘检测主要依靠各种算子,如 sobel 算子和 canndy 算子。

基于区域的分割则是将每个像素划分到相应的区域之中。区域生长算法作为基于区域的分割算法代表,目前应用较为广泛。区域生长算法的原理是:将性质相似的像素集成区域,从一组生长点开始,将生长点邻域性质相似的像素点与生长点合并成新的生长点,迭代生长点至不能生长为止。实际应用区域生长算法时需要解决3个关键问题:(1)选取一组能代表区域的种子生长点;(2)确定生长准则,使相邻像素能在生长过程中与生长点合并;(3)制定生长过程停止条件。

2 异常细胞检测方法设计与实现

异常细胞检测方法思路为:从本地读入图像,对图像进行预处理操作,包括将彩色图像转换为灰度图像,并对灰度图像做形态学滤波,使用大津法获得灰度阈值,生成二值图像。使用手动选择的方法确定生长点,并将目标细胞存入数组。对每个生长点使用生长算法,得到每个联通的细胞区域。对于每个细胞,计算总体面积,对细胞和掩码图做顶帽操作,利用大津法确定分离细胞核和细胞质的灰度阈值,获得二值图像,计算细胞核面积以及核占比。

异常细胞检测方法的实现使用 matlab,运行环境为 windows 10。

3 实验结果

对于输入的细胞图像,本方法能较好的分离细胞核并计算核占比,详细的实验结果如图 5~9 所示:

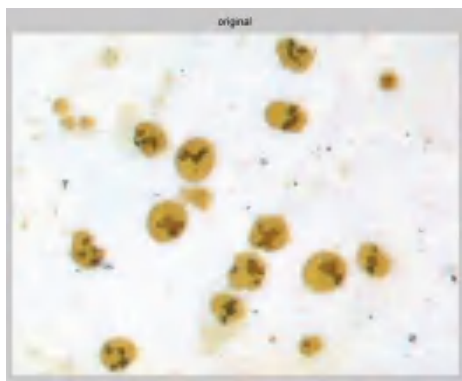


图 5 原图

Fig. 5 Original figure

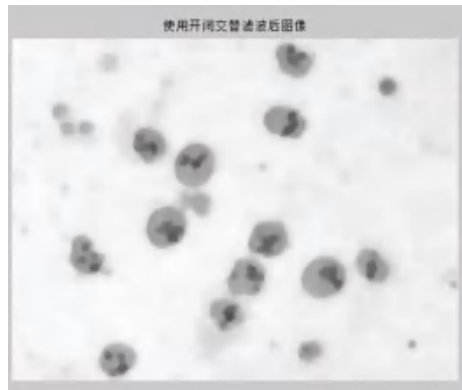


图 6 形态学操作滤波结果

Fig. 6 Results of morphological operation filtering

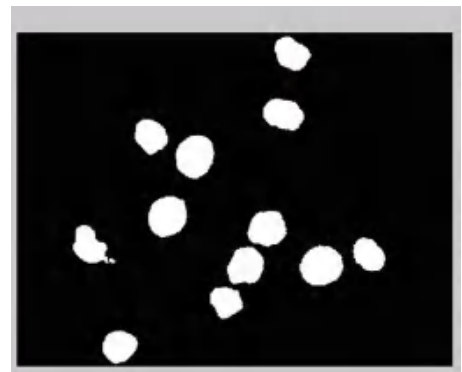


图 7 二值化图像

Fig. 7 Binarization image



图 8 细胞核图像

Fig. 8 Nuclear image

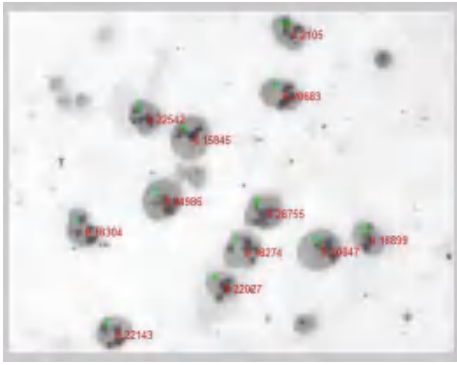


图 9 细胞核占比计算

Fig. 9 Nuclear proportion calculation

4 结束语

本文根据图像分割等原理,结合细胞检测这一实际需求,提出了使用图像处理相关算法检测异常细胞的方法,利用形态学运算、大津阈值分割法和区域生长算法实现了细胞分割,通过掩模区域相关操作,计算出细胞核的占比,进而判断细胞是否异常。

(上接第 113 页)

可以预测到,当 $Re = 160$ 时的回流长度为 $f(160) = 11.786m$ 。表 5 比较了 $Re = 80, 120, 160$ 时回流长度计算值和拟合值的误差,误差均在 1.23% 左右。可见该拟合公式能较好地预测雷诺数与回流长度的关系。

表 5 回流长度计算值与拟合值比较

Tab. 5 Comparison of calculated value and fitted value of reflux length

Reynolds number	80	120	160
计算值	6.479	9.066	11.643
拟合值	6.559	8.954	11.786
相对误差(绝对值)	1.23%	1.24%	1.23%

3 结束语

通过对三维长流道的简化得到了该模型的二维特征面,使用前处理工具 ICEM 对其进行网格划分,应用商业软件 FLUENT 对本案例进行计算,使用后处理软件 TECPLOT 绘制压力云图、速度云图和流线图。为了尽可能降低网格对于计算结果的影响,对 $Re = 100$ 时的网格进行了独立性检验,结果表明当最大网格尺寸取为 0.05 时,能兼顾计算精度和计算效率。为了保证在出口处的流动已经充分发展,对于流动雷诺数较大的情况,适当增加出口通道长度。通过提取求解结果中的压力梯度值,分析得到流动的分离点和回流点,进而获取回流长度。通过雷诺数 $Re = 50, 100, 150, 200$ 时的回流长度值,拟

合了回流长度关于雷诺数的三次方程公式,并预测了 $Re = 160$ 时的回流长度。通过研究,得出点结论:

(1) 使用先膨胀后腐蚀的方法可以省去去噪步骤; (2) 阈值分割后,会残留许多噪声,使用开运算可以去除; (3) 大津法是基于最小二乘法的阈值分割法,计算量很大,可用分水岭算法代替,也可使用梯度算子检测边缘。

参考文献

[1] 胡羽. 异常腺细胞形态学检测方法及应用[D]. 南昌航空大学, 2017.
 [2] 刘翠芳. 计算机图像处理技术应用分析[J]. 数字技术与应用, 2019, 37(10): 76-77.
 [3] 杨玉娥, 马志雯, 刘娟. 基于形态学的医学图像分割方法研究[J]. 中阿科技论坛(中英阿文), 2019(4): 42-47+55-61.
 [4] 王秋萍, 张志祥, 朱旭芳. 图像分割方法综述[J]. 信息记录材料, 2019, 20(7): 12-14.
 [5] 李了了, 邓善熙, 丁兴号. 基于大津法的图像分块二值化算法[J]. 微计算机信息, 2005(14): 76-77.

合了回流长度关于雷诺数的三次方程公式,并预测了 $Re = 160$ 时的回流长度。通过研究,得出点结论:

- (1) 最大静压出现在入口处,在水平阶段,随着流动发展,静压逐渐降低,当流动到达 Arc1 拐角处时,静压降至最低,此后流动继续发展,静压又逐步升高,最后达到稳定状态;
- (2) 随着流动发展,流体速度先增加再减小,处于管道中部的流体速度高于靠近管道壁面的速度,速度最大值出现在管道 1/2 径向处;
- (3) 随着雷诺数的增大,入口处的压力逐渐降低;
- (4) 雷诺数越大,拐角分离点处的角度越大;
- (5) 分离点随雷诺数增加而越来越接近;
- (6) $Re = 150, 200$ 时,流动分离前,压力梯度会出现震荡;
- (7) $Re = 100, 150, 200$ 时,在回流发生之前,压力梯度均有类似的震荡现象出现;
- (8) 拟合得到的公式能较好的预测其他雷诺数对应的回流长度。

参考文献

[1] ANSYS, Theory Guide. ANSYS Help Viewer. Release 16.0-© SAS IP, Inc. 2014.
 [2] TU Yijuan, Guan Heng Yeoh, LIU Chaoqun. Computational Fluid Dynamics: A Practical Approach[J]. Artificial Organs, 2012, 33(9): 727-32.