

文章编号: 1005—0523(2005)05—0082—04

目标特征识别技术在医学图像处理中的应用

胡晓燕, 蒋先刚, 郑国阳

(华东交通大学 交通信息工程及控制研究所, 江西 南昌 330013)

摘要:通过对目标连通区域标记及紧凑度、重心位置计算的方法来实现数字图像中特征目标自动识别的功能,并将其应用到医学图像中统计分析血液细胞的特征和物性确定细胞是否由病变.实验结果表明,该方法在血液分析中实用有效.

关 键 词:连通区域;区域标记;紧凑度

中图分类号:TP301.6

文献标识码:A

1 引 言

用图像处理方法做目标检测的一般顺序是:图像预处理、边缘检测、阈值分割、区域标记、形状判断分析.其过程如图 1 所示.用阈值分割得到的二值图像通常包含多个区域,而且有的区域可能由几个不连通的子区域组成,因此,有必要检测每个区域的连通情况,然后给每个区域标记后进行形状判断分析并将它们分别提取出来.几何特征分析是形状分析最便捷的方法和手段.在各种几何特征中,紧凑度和重心则是常用最重要的特征之一.紧凑度是在面积和周长的基础上,计算目标(或区域)的形状复杂程度的特征量,其参数为 $e=\frac{4\pi\times\text{面积}}{(\text{周长})^2}$,利用紧凑度参数,可以将目标物体从背景中提取出来.

本文通过对血液细胞图像的中细胞几何特征的统计分析,来确定送样是否存在病变,并通过实验来证明其有效性.

2 细胞特征统计算法分析

2.1 图像预处理

在进行图像识别之前,有以下预处理工作是必须的:

1) 二值化.设定一个赋值 T ,将图像分为两部分: $g(x,y)=\begin{cases} 1 & f(x,y)\geq T \\ 0 & f(x,y)< T \end{cases}$ 其中 $f(x,y)$ 为输入灰度图; $g(x,y)$ 为输出的二值图.

2) 细胞分割.将粘连的细胞分开,数学上满足两个区域的交集为空.多次统计发现目标图像中的粘连的细胞所占的比例较小,所以采用数学形态学中的开运算实现.对于轻度粘连的细胞,经开运算后,细胞之间分开了,大部分细胞没有明显的变化.

3) 孔洞填充.填充封闭区域内部的空洞,凡数学上满足单连通条件,采用种子填充算法即可实现.

4) 去除噪声.采用小波去噪,利用 Mallat 多分辨率分解,通过对得到的不同频带内的小波系数进行硬值域滤波的方法实现去除噪声.这种方法的去噪步骤为:由 Mallat 算法的分解公式:
$$\begin{cases} V_{j+1}=H\cdot V \\ W_{j+1}=G\cdot W_j \end{cases} j=0,1,2\cdots J-1$$
其中 H 和 G 为滤波器系数矩阵, V_0 为原始图像的采样值, V_j 和 W_j 分别为尺度 j 上的逼近系数和小波变换系数,将含有噪声的采样值在某个尺度下分解到不同的频带内,

收稿日期:2005—08—25

作者简介:胡晓燕(1980—),女,浙江永康人,华东交通大学在读硕士研究生,主要研究领域:工业测控,多媒体技术.

然后再将噪声所处的频带置零(强制消噪处理),再利用相应的重构公式: $V_j = \hat{H} \cdot V_{j+1} + \hat{G} \cdot W_{j+1}, j = J - 1, \dots, 1, 0$ (其中 $\hat{H}$ 为 $\hat{G}$ 综合滤波器,且满足 $\hat{H} \cdot \hat{H} + \hat{G} \cdot \hat{G} = 1$)进行小波重构,从而达到去噪的目的.

2.2 细胞区域标记

区域标记是指给连接在一起的像素赋上相同的标记,不同细胞的连接部分附上不同的标记.在

二维图像中,假设目标像素的灰度为 1,背景像素灰度为 0,它周围有 $m(m \leq 8)$ 个相邻的像素,如果这 m 个像素中的某一个点 A 灰度也等于 1,则称该像素与点 A 具有连通性,给 A 点赋一个与该像素相同的像素值.为了能更好的区分不同的细胞体,本文采用颜色差值区域标记法如图 2 所示,给不同的连通体附上不同的像素值,而且其颜色差值比较大,从图 3 我们可以很清楚的区分每个细胞.

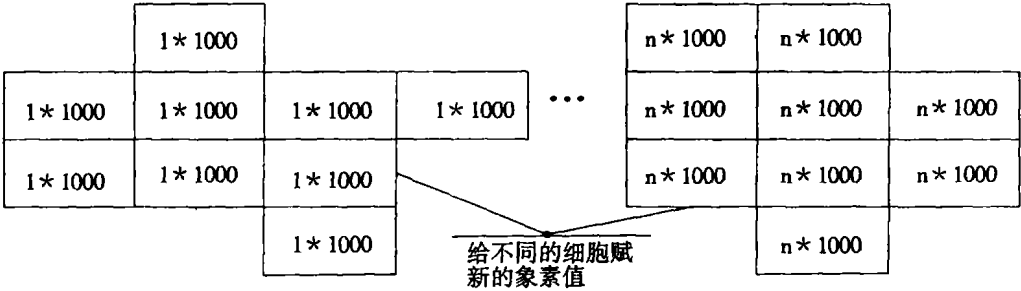


图 2 颜色差值区域标记法

图 2 为颜色差值区域标记法给不同细胞标记的过程,通过该处理每个细胞体呈不同的颜色,根据像素值的不同,可以进一步方便计算每个细胞的形状特征.

2.2.1 区域标记的算法及算法实现

图像处理中有很多不同的标记算法,本文主要讨论颜色像素标记法和二值图像像素快速标记法这两种实用的方法,它们能够标记出符合连通性质的所有连通区域.

2.2.2 颜色像素标记法

像素标记法是搜索整幅二值化图像,当搜索到灰度为 1 的像素时开始标记并搜索与其连通的点,赋上相同的像素值,当遇到没有加标记的目标像素时,给赋新的像素值.其标记算法步骤如下:

1) 从左到右、从上到下进行扫描,遇到没有加标记的目标像素 P 时,给这个点的像素赋一新值. $BloodImg.Canvas.Pixels[x,y] = C * 1000$; 其中 x,y 为目标像素在图像中对应的坐标, C 是一个累加值,给像素赋值时让每个细胞之间的像素颜色值之差为 1000,这样在图片显示的效果和进一步分析时更加容易区分不同的细胞体.

2) 给与 P 连接在一起(即相同连接成分)的像素添加相同的标记.

3) 进一步给所有与加标记像素连接在一起的像素添加相同的标记.

4) 直到同一连通域的像素全部被添加标记.这

样,一个连通区域就被添加了相同的标记,即赋了相同的像素值.利用 Connect 函数过程,通过 8 连通区域判断来实现以上功能,并返回每个连接成分的像素点,主要算法代码实现如下:

```
function TComponentReco.Connect(BMP: TBitmap; c, x, y: integer): integer;
begin
    count := count + 1; //连通物体的像素总数累加
    bloodImg.Canvas.Pixels[x,y] := c * 1000; //给不同连通物赋上不同的像素值
    if (BloodImg.Canvas.Pixels[x + 1,y] = clWhite) then //判断(x+1,y)方向点像素
        Connect(BloodImg,c, x + 1, y);
    if (BloodImg.Canvas.Pixels[x + 1,y - 1] = clWhite) then
        Connect(BloodImg,c, x + 1, y - 1);
    if (BloodImg.Canvas.Pixels[x + 1,y+1] = clWhite) then //
        Connect(BloodImg,c, x + 1, y+1);
    if (BloodImg.Canvas.Pixels[x,y - 1] = clWhite) then //
        Connect(BloodImg,c, x, y - 1);
    if (BloodImg.Canvas.Pixels[x,y + 1] = clWhite) then //
        Connect(BloodImg, c,x, y + 1);
    if (BloodImg.Canvas.Pixels[x-1,y - 1] = clWhite) then //
        Connect(BloodImg, c,x-1, y - 1);
    if (BloodImg.Canvas.Pixels[x-1,y] = clWhite) then //
        Connect( Blood,c, x-1, y);
    if (BloodImg.Canvas.Pixels[x - 1, y+1] = clWhite) then
        Connect(BloodImg,c, x - 1, y+1);
    //
```

```
Result := count; //连通物体的像素总数
end;
```

5) 返回到第(1)步,继续查找新的没加标记的像素,重复以上各个步骤.

6) 图像全部扫描后,处理结束.若图像中有两个及以上的目标物体,经过像素标记后,每个细胞颜色是不同的,如图 3 所示.

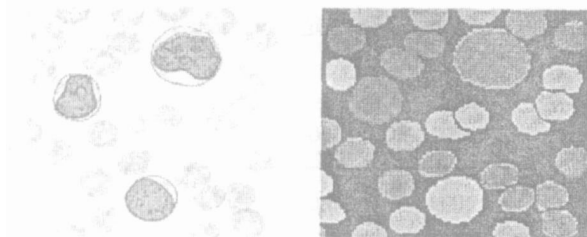


图 3 标记前后的图像

通过颜色像素区域标记法可以以不同色彩等级准确地标记出图像中各个连通区域,并得出图像对应每个目标物体的总像素点和图像中物体的个数.通过扫描特定的像素值就可以得到特定的连通区域,从而对这个连通区域作出更多的判断,如周长、形状、直径等等.但是,这种算法进行的是全局搜索,有些已经标记过的像素点可能被重复扫描,因此需要大量的堆栈空间,需要消耗大量的存储空间和处理时间.所以在物体比较多的数字图像中且实时性要求比较高的情况下,可以采用以下的算法.

2.2.3 二值图像像素快速标记法

这种标记法是采用基于扫描像素线检测连通域的方法,以直线段作为连通检测的基本处理单元,首先需要对二值图像进行逐行扫描,每扫描出当行的一条直线段(由连续的“1”像素点组成),则与上一行已检测出的直线段进行连通检测.算法步骤如下:

输入:二值化图像 $f(i, j)$

输出:中间标记图像 $g(i, j)$, 分析表 tab , 标记总数为 $mark_num$.

1) 令 $mark_num := 0$, $tab[mark_num] := 0$.

2) 从上到下逐行扫描图像,若在第 j 行中有一段值为 1 的连续像素 $f(i, j) \sim f(k, j)$, 则执行 3), 否则返回(2)进行下一行扫描.

3) 检查 $f(i-1, j-1) \sim f(k+1, j-1)$ (若超出图像边界,则边界点略去)

(3.1) 若 $f(i-1, j-1) \sim f(k+1, j-1)$ 的值全为 0, 则令 $mark_num := mark_num + 1$, 并讲 $mark_num$

赋给 $g(i, j) \sim g(k, j)$, 再令 $tab[mark_num] := mark_num$, 然后返回(2).

(3.2) 若 $f(i-1, j-1) \sim f(k+1, j-1)$ 有 L 个像素段, 则设与之对应 $g(i-1, j-1) \sim g(k+1, j-1)$ 中的 L 个标记为 $mark(1), mark(2), \dots, mark(L)$, ($L \geq 1$), 令 $mark$ 等于这几个值中最小的值, 将 $mark$ 赋值给 $g(i, j) \sim g(k, j)$, 同时记录连通关系. 返回(2).

该方法可检测出各种复杂类型的连通区域, 而且避免了基于像素标记方法重复扫描一些像素点的现象, 因此可提高处理速度.

2.3 获取细胞特征参数

1) 紧凑度. 其计算是在面积和周长的基础上, 其方程为 $e = \frac{4\pi \times \text{面积}}{(\text{周长})^2}$. 面积等于细胞(或区域)中包含的像素数, 对每个细胞像素点的统计通过连通物体计算函数 Connect 函数中的像素总数 $count$ 来实现. 细胞的周长是指轮廓线上像素间距离之和, 像素间距离有两种情况, 并列方式可以是上、下、左、右四个方向, 这种并列像素间的距离是 1 个像素, 倾斜方向连接的像素有左上角、左下角、右上角、右下角四个方向, 这样倾斜方向像素间的距离是单位像素. 在进行周长测量时, 需要根据像素间的连接方式, 分别计算距离. 如图 4 是一个周长测量示例, 其周长值为. 周长算法步骤如下:

(1) 扫描图像, 顺序调查图像上各个像素的值, 寻找没有扫描标志 a 的边界点.

(2) 如果 a 周围全为黑色像素, 说明 a 是个孤立点, 停止追踪.

(3) 否则按图 4 的顺序寻找下一个边界点. 用同样的方法, 追踪其他的边界点.

(4) 到了下一个边界点 a , 即证明已经围绕物体一圈, 终止扫描.

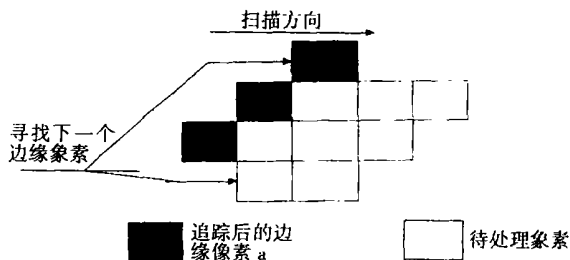


图 4 轮廓追踪与边缘计算

2) 重心位置. 就是求连通区域中像素坐标的平均值. 若其像素的坐标为 (x_i, y_i) ($i=0, 1, 2, \dots, n-1$), 其重心坐标 (x_0, y_0) 可以由下式求得: $(x_0, y_0) =$

$$\frac{1}{n} \sum_{i=0}^{n-1} x_i, \frac{1}{n} \sum_{i=0}^{n-1} y_i)$$

3 血液细胞特征分析实验

以某血液图为例,截取了其中一部分,判断是否有细胞病变.其原图及标记后图像如图 3 所示.根据紧凑度及重心位置公式,可以得到下表 1(记录部分细胞的参数).

表 1 细胞的特征参数及分析结果

细胞序号	面积	周长	紧凑度	重心坐标	细胞归类
1	253	75	0.565207	122,41	正常
2	6022	415	0.438069	86,29	病变
3	250	76	0.543905	42,72	正常
4	262	73	0.617512	17,82	正常
5	235	71	0.585519	59,88	正常
6	589	149	0.333390	28,52	病变
7	589	156	0.303988	69,108	病变
8	245	73	0.577744	91,95	正常
9	105	36.35	0.298875	10,50	边界细胞
...	...	...	...	...	...

从表 1 可以看出,正常细胞的像素面积在 250 左右,周长在 75 左右,病变后细胞的像素面积会变大,同时其周长也变长,而边界点的细胞,面积偏小.由表可见各类细胞面积和周长存在一定规律性,紧凑度在 0.55 以左右的为正常细胞,0.50—0.30 之间的为病变细胞,0.30 以下的为边界细胞,所以表中 2、6、7 号细胞为病变后细胞,9 为边界细胞,其他为正常细胞.该血液图不同细胞分布概率如图 5 所示由图 5 可知,在这幅血液中,通过概率分布等统计计算知有 8% 的血液细胞已经病变,80% 的细胞为正常细胞,其中 12% 细胞为边界细胞,可以对之作进一步得补偿分析,进

而分析这些细胞是否有病变.综上,紧凑度能够很好地表示出细胞的形状特征,这对于病变后的细胞和正常细胞的分类有很重要的意义,而重心位置为下一步病变细胞的定位及平面分布状态提供了进一步检验参数.

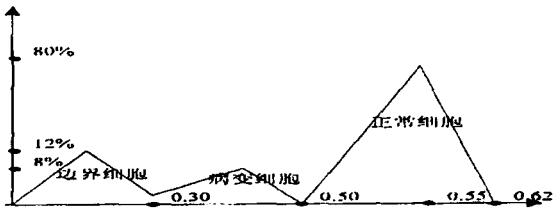


图 5 细胞概率分布图

4 结束语

血液细胞分析实验表明采用二值图像像素快速标记法能检测出任意形状的连通体,并正确标号,而且速度比像素标记法快,所需内存资源少.而对细胞的分析,很大程度上涉及到细胞的具体分布及几何特性的统计与分析,能对血液图像进行自动诊断,具有实用价值.

参考文献:

[1] 陈兵旗,等. Visual C++实用图像处理专业教程[M]. 北京:清华大学出版社,2004.
[2] David F. Rogers. Procedural Elements for Computer Graphics [M]. 北京:机械工业出版社,2002.
[3] 李强,张桂林. 二值图像标记的一种快速算法[J]. 计算机学报,1997, (3): 238—240.
[4] 济芳. Visual C++小波变换技术与工程实践[M]. 北京:人民邮电出版社,2004.

Application of Automatic Recognition Based on Image Feature in Iconological Image

HU Xiao-yan, JIANG Xian-gang, ZHENG Guo-yang

(Institute of Transportation Info—engineering and Control, East China Jiaotong Univ., Nanchang 330013, China)

Abstract: This paper presents how to automatically recognize components in a digital iconological image by labeling connected component and calculate its compactness and centriod. Then this method was applicated in iconological image processing for analyzing blood cell. The experiment results show that the method is effective for analyzing blood samples.

Key words: connected component; component labeling; compactness