

基于稀疏轮廓点模型的彩色重叠细胞图像分割

关涛<sup>1</sup> 周东翔<sup>1</sup> 樊玮虹<sup>1</sup> 刘云辉<sup>2</sup>  
<sup>1</sup>(国防科学技术大学电子科学与工程学院 长沙 410073)  
<sup>2</sup>(香港中文大学机械与自动化工程学系 香港 999077)  
(gt\_mike2003@126.com)

Segmentation of Color Overlapping Cells Image Based on Sparse Contour Point Model

Guan Tao<sup>1</sup>, Zhou Dongxiang<sup>1</sup>, Fan Weihong<sup>1</sup>, and Liu Yunhui<sup>2</sup>  
<sup>1</sup>(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073)  
<sup>2</sup>(Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong 999077)

**Abstract** Based on the analysis of cell contour structure, a sparse contour point model, which can describe the characteristics of the cell contour, is proposed in this paper. In the sparse contour point model, cell contour is divided into 2 parts, namely light contour and dark contour, respectively; and then the cell contour is approximately described as a set of sparse contour points. Based on this model, the color and grayscale image segmentation techniques are combined to locate the basic contour, which lies between the cell and the background. Then, a circular dynamic contour searching method is proposed to search for the dark contour that lies in the overlapping cell region along the basic contour. Contour points located by the searching method are arranged to construct the initial contour of the gradient vector flow (GVF) Snake model. Then the GVF Snake model is performed to obtain the final accurate segmentation result of the cell image. Various cell images containing single cell, overlapping cells of similar colors and overlapping cells of different colors have been tested to show the validity and effectiveness of the proposed method. The proposed techniques are useful for the development of automatic cervical cell image analysis systems.

**Key words** medical image processing; overlapping cell images; color image segmentation; sparse contour point model; gradient vector flow (GVF) Snake model

**摘要** 根据对细胞轮廓结构的分析,提出了描述细胞轮廓特征的稀疏轮廓点模型.该模型将细胞轮廓划分为强轮廓和弱轮廓2部分,并用一系列稀疏的轮廓点来近似表示细胞轮廓.在该模型的基础上,算法综合了彩色图像处理与灰度图像处理的方法,首先提取细胞与背景交界处的强轮廓作为基准轮廓.然后提出了一种环形动态轮廓搜索算法,沿着基准轮廓搜索位于重叠细胞区域的弱轮廓.经过轮廓搜索算法获得细胞的稀疏轮廓点组成初始轮廓,采用GVF(gradient vector flow)Snake模型进行细胞的精确分割.通过对2个细胞图像数据集的图像样本进行的分割实验验证了算法对于分割单细胞图像、同种颜色重叠细胞图像和不同颜色重叠细胞图像的准确性.

**关键词** 医学图像处理;重叠细胞图像;彩色图像分割;稀疏轮廓点模型;GVF Snake 模型  
**中图法分类号** TP391

显微镜细胞图像分割是医学图像处理领域的研究热点之一。通过图像处理、模式识别和计算机视觉技术对细胞图像进行精确的分割,从而获得细胞的形态参数,这样可以辅助医生对细胞的病变状态进行判别。例如在宫颈癌的预防筛查中,宫颈涂片筛查方法被认为是预防宫颈癌最有效的手段之一。但是由于人的主观性、劳累、经验等因素影响,在临床上人工阅片造成的误诊对于宫颈癌的早期发现有着严重的制约性。因此,计算机自动筛查宫颈涂片技术可以为宫颈癌的预防和筛查带来很大的益处。宫颈细胞图像的自动精确分割是实现这一自动筛查技术的关键一步。由于宫颈癌是危害女性的最严重、发生率最高的恶性肿瘤之一,因此计算机自动筛查宫颈涂片技术的实现对于人类的健康具有积极的意义<sup>[1]</sup>。

在宫颈细胞图像分割领域,国内外学者已经进行了大量的研究<sup>[1]</sup>。细胞图像分割的目的主要是得到细胞核与细胞浆的轮廓、计算核浆的形态参数、判断细胞的病变状态,从而为宫颈癌的诊断提供重要依据。按照分割图像的特点划分,相关研究主要有2类:

1) 针对显微镜低倍物镜采集的低分辨率大图像的分割,这类分割主要是实现细胞核、实质细胞区域的定位。Plissiti 等人<sup>[2]</sup>采用形态学重建的方法分割多细胞图像的细胞核。Bamford 等人<sup>[3]</sup>采用双主动轮廓模型分割细胞核。Genctav 等人<sup>[4]</sup>采用多尺度分水岭方法分割细胞核与细胞浆区域。

2) 针对显微镜高倍物镜采集的高分辨率细胞图像,这类细胞图像(如 Herlev 数据集<sup>[5]</sup>中的图像)既包含单个无重叠的细胞图像,也包含重叠细胞图像。这类分割算法主要是实现单细胞、重叠细胞图像的精确分割,主要难点是解决细胞边缘相对模糊和细胞重叠的问题。相关的算法有的是基于边缘的分割算法<sup>[6]</sup>,主要针对边缘模糊的问题,通过采用一些灵活的方法对细胞的轮廓进行边缘增强,从而获得细胞的轮廓;有的算法<sup>[7]</sup>针对杂质干扰和边缘模糊的问题,将图像的边缘与区域信息结合来得到细胞的轮廓。例如 Li 等人<sup>[7]</sup>提出的 RGVF(radiating gradient vector flow) Snake 算法首先对图像采用 spatial  $k$ -means 聚类算法进行粗分割获得 Snake 演化的初始轮廓,然后提出了一种基于径向边缘检测的边缘提取方法,最后利用 Snake 模型进行细胞的精确分割。但是,这一类研究目前还仅能实现针对单个无重叠细胞的分割。对于重叠细胞图像,尤其是重叠程度比较复杂的细胞图像的分割,目前的算法效果不好。正如文献<sup>[7]</sup>所述,重叠细胞分割是一项具有挑战性的

难题,这方面的算法还有待大量深入的研究。本文的研究就是针对这种高分辨率图像中重叠细胞的分割。

在重叠细胞分割领域,有的算法<sup>[8]</sup>首先标记细胞(或者细胞核)中心的位置,然后采用改进的分水岭方法分割。但是这种方法得到的细胞分割结果与真实细胞轮廓相差较大。比较复杂一些的方法<sup>[9]</sup>首先对待分割的目标采用精确的数学模型进行形状建模,获得目标的先验形状信息,然后采用参数估计方法得到分割结果。这种方法利用了目标的形状特征,采用数学模型精确的描述目标的几何形状。但是由于细胞轮廓的形状是非刚性、不规则的,难以用精确数学模型进行建模,所以采用基于已知轮廓建模估计未知轮廓的方法是不可取的。在细胞图像中既包含有同种颜色重叠的细胞,又包含有不同颜色重叠的细胞。本文作者曾经提出一种基于色差向量场的重叠细胞图像分割算法<sup>[10]</sup>,但是该算法仅仅利用了图像的颜色信息,而没有充分利用图像的亮度(灰度)信息,对于同种颜色重叠细胞图像的分割效果不好。

GVF(gradient vector flow) Snake 模型<sup>[11]</sup>是图像分割领域的研究热点之一,它解决了传统 Snake 模型<sup>[12]</sup>作用距离小和无法处理轮廓凹陷的问题。由于 Snake 模型本身是一个弹性形状模型,对于形状复杂的目标分割问题,GVF Snake 能够取得好的分割结果。本文研究的分割方法结合了 GVF Snake 模型。通过对细胞轮廓的几何结构的深入分析,本文提出一种描述细胞轮廓的稀疏轮廓点模型。基于该模型,本文将彩色图像处理与灰度图像处理方法结合,提出了一种基于环形动态搜索的细胞轮廓定位方法。首先,本文采用形态学重建算法<sup>[2,13]</sup>提取细胞核。根据细胞核的位置与细胞轮廓的几何关系,采用梯度分解的方法实现了图像中主体细胞(中心细胞)的边缘增强。在通过该方法增强的细胞图像中,中心细胞轮廓的边缘得到有效增强,虚假细胞轮廓和杂质的边缘得到有效的抑制。为了提取细胞区域的粗轮廓,本文利用文献<sup>[10]</sup>的彩色细胞图像分割法进一步提取细胞与背景交界处的细胞轮廓。然后,通过对重叠细胞轮廓特点的分析,本文提出了先定位细胞轮廓中的一部分明显轮廓作为基准轮廓,然后采用环形动态轮廓搜索算法定位细胞整体轮廓。最后将上述搜索算法得到的轮廓作为 GVF Snake 模型的初始轮廓,通过轮廓演化最终得到细胞的精确分割结果。由于在本文研究的重叠细胞图像分割中,主要的难点在于整体细胞(细胞浆)的分割,细胞核的分割难度较小,而且前人的算法对于细胞核的

分割已经取得了很好的效果,所以本文算法研究的是重叠细胞图像中整体细胞(细胞浆)的分割. 本文算法对于其他类型细胞图像的分割也有很好的潜在应用价值.

1 细胞的稀疏轮廓点模型

在实际采集的细胞图像中,经常会出现重叠细胞. 图 1 给出了 2 幅重叠细胞图像. 在重叠细胞图像中,既有同种颜色细胞重叠,也有不同颜色细胞重叠. 在细胞重叠区域,细胞轮廓的边缘相对模糊,重叠细胞的轮廓相互交错,较难分辨. 以上这些因素对重叠细胞图像的分割提出了较大的挑战. 单一的彩色图像分割方法或灰度图像分割方法很难实现对复杂重叠细胞的分割. 为了能够提高算法对于细胞图像的普适性,应该将彩色图像分割方法与灰度图像分割方法结合.

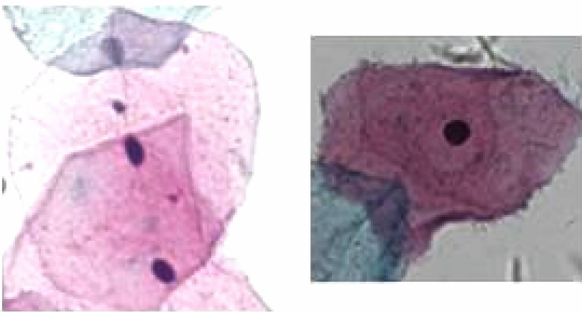


Fig. 1 Two Color overlapping cervical cell images.  
图 1 两幅彩色重叠宫颈细胞图像

在细胞图像分割研究领域有很多算法采用 GVF Snake 模型进行细胞图像的精确分割,但是 GVF Snake 模型仍然存在一定的局限性,即初始轮廓的自动设置问题和边缘增强的问题. 解决这 2 个问题的主要目的是让 Snake 轮廓越过虚假边缘的干扰,最终收敛到目标的真实轮廓. 虽然有大量研究试图解决这 2 个问题,但是针对重叠细胞图像的分割,这些方法仍然不能够得到好的分割结果. 要么是初始轮廓定位不准,使得 Snake 轮廓的演化很难越过假边缘的干扰;要么是在削弱假边缘的同时也抑制了大量的真边缘,致使目标的轮廓信息丢失,分割结果不准确. 因此应用 GVF Snake 模型分割细胞图像,关键是要解决 Snake 模型初始轮廓的自动定位和细胞图像的边缘增强. 本文试图解决这 2 个问题,从而提出一种结合 GVF Snake 模型的重叠细胞图像分割算法.

通过对细胞几何结构的分析,本文根据细胞轮廓的特点提出一种描述细胞轮廓的模型——稀疏轮廓点模型,该模型可以概括为以下 4 点定义:

1) 假设细胞轮廓由一系列轮廓点组成,这些轮廓点环绕在细胞核的周围. 从细胞核沿着径向引出一条射线,这条射线与细胞轮廓交于 1 点,即细胞在这个径向方向的轮廓点. 以细胞核中心作为原点  $O$ ,以极坐标  $(\rho, \theta)$  表示细胞轮廓点的位置坐标.  $\rho$  是轮廓点到细胞核心的距离,  $\theta$  是轮廓点所在半径方向的角度. 则细胞轮廓可以由一系列点的集合  $P$  来表示

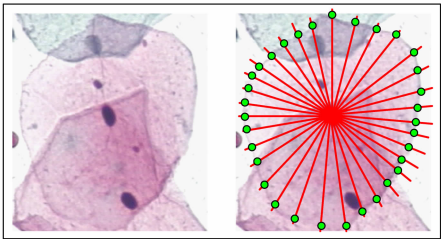
$$P = \{p_i \mid (\rho_i, \theta_i) \text{ 是点 } p_i \text{ 的坐标值}\}.$$
 (1)

2) 由于近似描述细胞轮廓的形状并不需要细胞轮廓中的所有点,所以本文将轮廓点集合  $P$  定义为一系列稀疏的轮廓点的集合. 也即相邻的 2 个轮廓点  $(\rho_i, \theta_i)$  和  $(\rho_j, \theta_j)$ , 其对应的角度相差  $\Delta\theta = \theta_i - \theta_j$ . 要求  $\Delta\theta$  为较小的数值,且该数值与细胞图像的尺寸有关,本文第 4 节将会详细讨论  $\Delta\theta$  的取值. 这样细胞的轮廓就可以由一系列稀疏的轮廓点来描述,如图 2(a)所示.

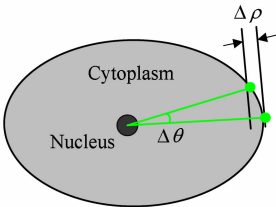
3) 本文将细胞轮廓分为了 2 部分,一部分是细胞与背景交界处的轮廓,用  $P_q$  表示,称为强轮廓. 这一部分轮廓点明显可辨,对应的边缘较清晰. 另一部分是位于重叠细胞区域内部的轮廓点,用  $P_r$  表示,称为弱轮廓. 这一部分轮廓点处的边缘较模糊. 则细胞的轮廓点集:

$$P = P_q \cup P_r.$$
 (2)

4) 图 2(b)为细胞的近似模型,2 条从细胞核出发的射线的角度相差  $\Delta\theta$ ,其对应的轮廓点距离圆心



(a) Sparse contour point model of a cell



(b) Characteristics between neighboring contour points

Fig. 2 Analysis of the cell structure.  
图 2 细胞的轮廓结构分析

的长度差值  $\Delta\rho$  为

$$\Delta\rho = k(\Delta\theta), \quad (3)$$

其中,  $k$  是非线性函数, 其具体形式我们并不关心, 但是要求  $\Delta\theta \rightarrow 0$ , 则  $\Delta\rho \rightarrow 0$ . 也即对于 2 个角度相差很小的轮廓点, 其对应的位置相近.

根据上述定义的模型, 本文将细胞轮廓提取的过程视为确定细胞的稀疏轮廓点的过程. 为了定位细胞的稀疏轮廓点, 本文提出了基于环形动态搜索的细胞轮廓定位算法. 算法分为 3 大步骤: 1) 定位细胞的强轮廓点; 2) 以强轮廓点为基准轮廓, 沿着基准轮廓进行环形搜索以定位弱轮廓点, 最终确定细胞的整体轮廓; 3) 采用 GVF Snake 模型提取精确的细胞轮廓.

## 2 预处理

### 2.1 细胞核提取

细胞核在图像中的主要特征是: 颜色深、形状近似为圆形、面积在一定的范围内, 本文参考文献[2]中的方法, 首先将彩色细胞图像转换为灰度图像, 采用形态学重建方法<sup>[13-14]</sup>提取图像中的局部颜色较深(对应灰度图像中的局部区域极小值)的区域. 但是提取的这些区域包含一部分杂质, 本文根据作者之

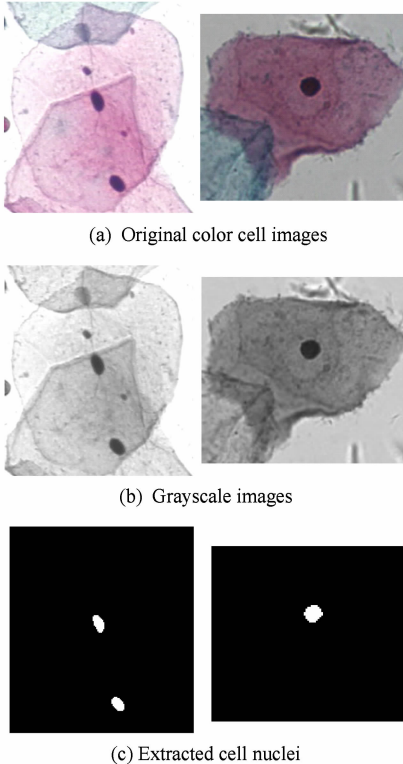


Fig. 3 Cell nuclei extraction process.

图 3 细胞核提取过程

前的工作<sup>[15]</sup>中采用的方法, 提取细胞核区域的面积、周长、区域面积与外接凸多边形面积比、长宽比等参数对分割后得到的细胞核区域进行筛选, 一方面剔除杂质, 另一方面从中提取出有效的细胞核区域. 采用形态学重建算法提取的细胞核区域如图 3 所示.

由于细胞图像中可能含有多个细胞核, 而图像中的主体细胞所包含的细胞核位置靠近图像的中心点, 所以我们选择最靠近图像中心点的细胞核, 该细胞核中心的位置  $(x_{cen}, y_{cen})$  为

$$\begin{cases} x_{cen} = \frac{1}{N} \sum_{i \in J} x_i, \\ y_{cen} = \frac{1}{N} \sum_{i \in J} y_i, \end{cases} \quad (4)$$

其中,  $J$  为细胞核区域,  $(x_i, y_i)$  为细胞核区域内像素点的位置坐标,  $N$  为细胞核区域内像素点总数.

### 2.2 中心细胞的边缘增强

图 4 所示是一幅细胞重叠示意图, 中心细胞左侧的部分细胞浆与邻近细胞发生部分重叠. 可以发现, 由于细胞的轮廓一般是凸起的, 而且图像中重叠细胞浆的灰度比周围非重叠细胞浆和背景的灰度低. 所以属于同一个细胞的细胞轮廓点的梯度方向与细胞核心径向方向的夹角一般较小, 而对于不属于该细胞的细胞轮廓点, 这个夹角将会比较大. 在图 4 中, 以中心细胞作为主体细胞来分析, 细胞轮廓点  $p_1$  和  $p_2$  位于重叠区域, 分别属于中心细胞和邻近细胞. 图中的箭头代表细胞轮廓点处的梯度方向, 虚线代表由细胞核心向细胞轮廓点引出的射线. 由上述分析可知  $\alpha_1$  的值较小, 而  $\alpha_2$  较大.

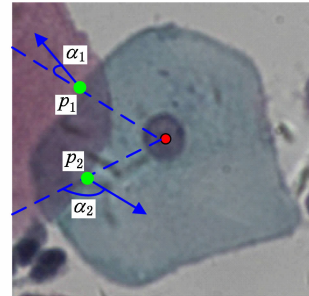


Fig. 4 Illustration of the difference between the gradient vectors at different cell contour points.

图 4 不同细胞轮廓点处的梯度方向差异分析

本文先采用 Sobel 算子对细胞图像  $f$  进行梯度检测. 令图像的像素点  $(i, j)$  对应的窗为  $W(i, j)$ , 梯度向量为  $g(i, j)$ , 则:

$$g(i, j) = (M_x * W(i, j), M_y * W(i, j)), \quad (5)$$

$$|g(i,j)| = \sqrt{(M_x * W(i,j))^2 + (M_y * W(i,j))^2}, \tag{6}$$

其中, \* 表示卷积运算,  $M_x$  和  $M_y$  分别为 Sobel 算子的 2 个模板.

设增强后的图像为  $f_{enhance}$ , 得到最终的边缘增强图像为

$$f_{enhance}(i,j) = |g(i,j)| \times \frac{\cos\alpha + |\cos\alpha|}{2}, \tag{7}$$

其中,  $\alpha$  是梯度向量  $g(i,j)$  与射线方向的夹角. 为了防止微弱的虚假边缘对后续处理的干扰, 这里本文对增强后的图像作进一步的弱边缘抑制:

$$f_{enhance}(i,j) = \begin{cases} f_{enhance}(i,j), & f_{enhance}(i,j) > \\ \mu \max\{f_{enhance}\}, & \\ 0, & \text{others,} \end{cases} \tag{8}$$

其中,  $\mu$  为加权参数,  $\max\{f_{enhance}\}$  为  $f_{enhance}$  中的最大像素值. 参数  $\mu$  控制了弱边缘抑制的程度, 在本文算法中我们取值为 0.1, 能够得到较好的效果.

图 5 给出了本文方法得到的边缘增强结果与直接采用 Sobel 算法进行边缘提取结果的对比. 可以看出, 在本文方法的到的边缘增强结果中, 中心细胞轮廓的边缘得到了一定程度的增强, 而不属于中心细胞的假边缘得到了有效的抑制. 同时也可以看到, 中心细胞的一部分边缘并没有得到有效的增强. 这主要是由于细胞的轮廓中可能有部分并不完全是凸状的, 但这并不会对本文的分割算法有很大影响. 此处边缘增强的目的主要是在抑制不属于中心细胞的

虚假边缘的同时, 对属于中心细胞的边缘进行一定的增强. 边缘增强是为后续轮廓搜索算法作必要的准备, 以便于后续轮廓搜索算法定位到更多细胞轮廓点. 少量细胞轮廓边缘没有被增强并不会对后续分割算法产生很大的影响.

2.3 基于色差向量场的细胞轮廓粗提取

我们在之前所做的彩色细胞图像分割研究中, 提出了一种基于色差向量场的彩色细胞图像分割算法<sup>[10]</sup>, 本文提出的算法用到了该方法. 在由 RGB 表示的彩色图像中, 我们提出采用归一化色差向量  $(u,v)$  来表征细胞图像的颜色信息:

$$u = \frac{R - G}{\sqrt{(R - G)^2 + (R - B)^2}},$$

$$v = \frac{R - B}{\sqrt{(R - G)^2 + (R - B)^2}}. \tag{9}$$

基于式(9)定义的色差向量, 文献[10]中分割方法的大致思路为: 首先定位图像中的细胞核, 根据细胞核位于细胞内部、被细胞浆环绕这一特性, 选取细胞核周围的一部分像素点所对应的色差向量均值作为模板. 然后采用一种循环匹配的方法实现图像中主体细胞的分割. 算法的具体细节详见参考文献[10].

这种分割方法对于单细胞图像和不同颜色重叠细胞图像能够取得较好的分割结果, 但是对于同种颜色重叠细胞图像的分割效果不好. 本文利用该算法提取细胞区域的边缘作为细胞轮廓的粗提取结果. 设文献[10]中算法的分割结果为  $f_{seg}$ , 最终提取的细胞区域边缘图像为  $f_{ce}$ , 则:

$$f_{ce} = f_{seg} \oplus SE - f_{seg} \ominus SE, \tag{10}$$

其中, SE 代表结构元素,  $\oplus$  代表形态学膨胀运算,  $\ominus$  代表形态学腐蚀运算. 通过这种方法提取的细胞边

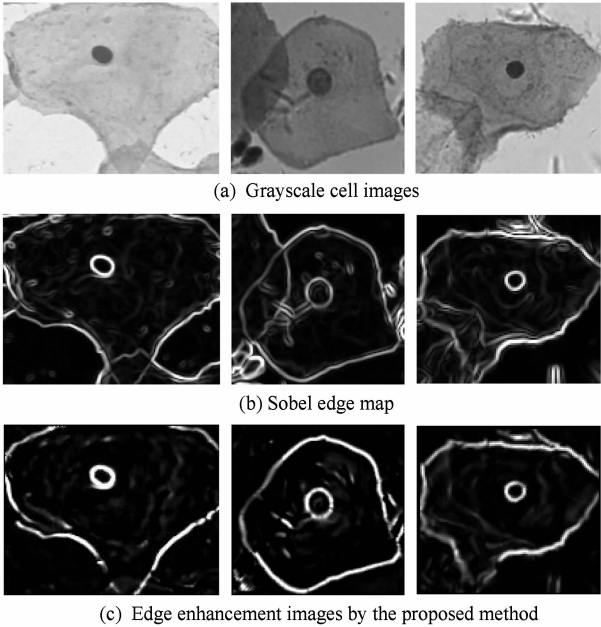


Fig. 5 Edge enhancement result.

图 5 边缘增强结果

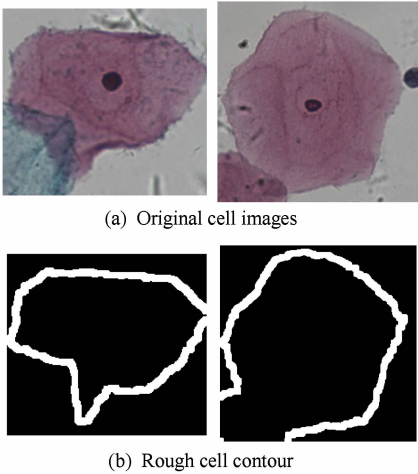


Fig. 6 Rough extraction of the cell contour.

图 6 细胞轮廓粗提取结果

缘图像特点为:该图像为二值边缘图像,细胞边缘为细胞与背景交界处的边缘,或者为不同颜色重叠细胞交界处的边缘,如图 6 所示。SE 的尺寸决定了所提取细胞边缘的宽度,其取值以能够覆盖细胞轮廓的边缘为准,在本文中 SE 选为圆形结构元素,直径尺寸选为 4 个像素点。

### 3 重叠细胞分割

#### 3.1 定位基准轮廓

重叠细胞分割的主要难点在于细胞重叠区域的正确分割,相关研究中有的采用对图像粗分割得到的细胞区域外轮廓作为 Snake 演化的初始轮廓<sup>[7]</sup>,有的则采用几何分析的方法,用一个包含细胞的圆作为初始轮廓<sup>[16]</sup>,但是这些方法所获得的初始轮廓与细胞的真实轮廓距离较远,都容易受到杂质的干扰,尤其是受到细胞重叠区域的虚假轮廓的干扰。本文第 2 节提出了细胞的稀疏轮廓点模型,将细胞轮廓划分为 2 部分:1)强轮廓(明显的轮廓),位于细胞与背景交界处的非重叠细胞轮廓;2)弱轮廓(不明显的轮廓),位于重叠细胞区域的细胞轮廓。本文要定位的基准轮廓点是细胞轮廓中较明显的轮廓点,也就是强轮廓中的轮廓点。细胞的强轮廓一般具有这样的特点:1)在 3.2 节的边缘增强结果中,强轮廓中的大部分得到一定的程度的增强(其边缘幅值相对较高);2)强轮廓位于 3.3 节提取的细胞边缘图像中。

设最终确定的基准轮廓图像为  $f_{cb}$ ,则其像素点  $(i, j)$  的像素值为

$$f_{cb}(i, j) = f_{enhance}(i, j) \times f_{ce}(i, j), \quad (11)$$

其中,  $f_{enhance}$  为 3.2 节进行边缘增强后的边缘图像,  $f_{ce}$  为 3.3 节通过彩色图像处理获得的细胞边缘。细胞的基准轮廓定位结果如图 7 所示:

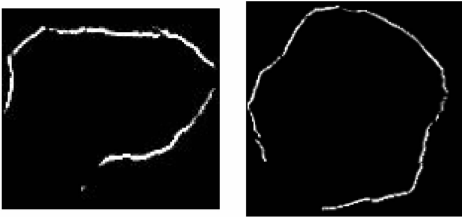


Fig. 7 Located basic contour of the two cells in Fig. 6.

图 7 对图 6 中 2 幅细胞图像的基准轮廓定位结果

#### 3.2 环形动态轮廓搜索

在定位基准轮廓之后,本文算法从基准轮廓出发,以细胞核为中心,环绕细胞核进行环形搜索,以

尽可能找到细胞的大部分轮廓点。通过观察细胞的边缘图像,可以发现:以细胞核心为原点,向细胞轮廓引一条射线,则这条射线将会与细胞的轮廓线交于一个轮廓点。本文定义位于这条射线上的边缘值为径向边缘值。如果取这条射线上对应的径向边缘值组成 1 维信号,则细胞轮廓点位于该信号的一个峰值点,如图 8 所示,该细胞图像为经过 3.2 节方法确定的边缘增强图像。

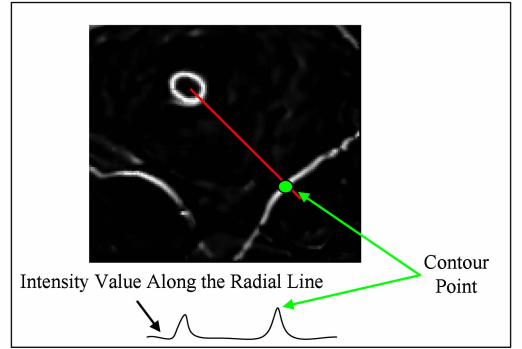


Fig. 8 The characteristics of the cell contour point in the radial gradient.

图 8 径向边缘中细胞轮廓点的特点分析

根据上述规律和第 2 节对细胞模型提出的假设,本文的轮廓搜索算法就是在每条这样的射线上定位一个细胞轮廓点,其中每个角度的径向边缘值从边缘增强图像  $f_{enhance}$  中获取。因此,算法首先根据基准轮廓图像在基准轮廓对应的角度方向定位基准轮廓点。然后以基准轮廓点作为已知的细胞轮廓点,采用环形搜索的方法逐步定位细胞的所有轮廓点。也即在每个以细胞核为中心、角度为  $\theta_i$  的方向,确定细胞轮廓点与细胞核心的距离  $\rho_i$ ,轮廓搜索完成后,细胞轮廓点随之确定完毕。

轮廓点搜索的流程为:算法从已知轮廓点所在的角度出发,环绕细胞核的每个角度进行扫描。每扫描到新的角度,首先估计该角度方向的轮廓点位置,确定轮廓点位置的可能范围作为搜索区间,然后在该方向的径向边缘数据中搜索,确定轮廓点位置,更新已知轮廓点。之后扫描到下个方向进行新的轮廓点搜索。根据第 2 节提出的稀疏轮廓点模型中的定义 4,对于 2 个角度相差很小的轮廓点,其对应的位置相近。所以算法每次扫描到新的角度,将前一个扫描角度所确定的轮廓点位置作为当前扫描角度轮廓点的初始估计位置。

为了避免虚假轮廓点的干扰,提高算法的鲁棒性。本文在搜索过程中引入环形动态的搜索策略。

按照稀疏轮廓点模型的定义,为了提高算法的搜索效率,本文在搜索轮廓点时,不是每个角度都做一次搜索,而是每次步进一个角度  $\Delta\theta$ ,也即角度扫描的过程是步进式扫描的.最后搜索得到的并不是全部轮廓点,而是一系列稀疏的轮廓点.角度  $\Delta\theta$  选取的过大会使得细胞轮廓点过于稀疏,导致轮廓定位过于粗糙;而选择的过小,将会导致算法的计算量剧增、稳定性差.本文根据对大量细胞图像的实验结果选择的步进角度为

$$\Delta\theta = \frac{\pi}{100}. \tag{12}$$

由图 5 可知,细胞图像经过边缘增强后,虽然大部分轮廓得到增强,但是也有一部分真实细胞轮廓被抑制掉,这就造成细胞轮廓上的边缘其实是断裂的.因此在进行轮廓搜索的过程中,可能会出现找不到轮廓点的情况.这时,算法引入参数  $\omega$ ,用来控制当在某个角度找不到符合条件的轮廓点时继续跳至下个角度搜索的次数.由第 2 节中的定义 4 可知相邻角度的轮廓点的位置  $\rho$  相近.因此本文引入参数  $\tau$  来控制当在某个角度沿着径向射线搜索轮廓点的区间范围.本文采用的环形动态轮廓搜索算法分为 2 个阶段:第 1 阶段采用小范围( $\omega$  较小)、固定尺寸( $\tau$  较小、尺寸固定)的搜索区间,搜索基准轮廓周围明显的轮廓点;第 2 阶段采用大范围( $\omega$  较大)、尺寸动态调整( $\tau$  动态变化)的搜索区间,搜索不明显、微弱的轮廓点.算法先执行第 1 阶段的小范围搜索,再执行第 2 阶段的大范围动态搜索.在每个阶段,算法做 2 次循环搜索,第 1 次顺时针方向,第 2 次逆时针方向,再将 2 次搜索得到的属于同一个角度方向的轮廓点位置取平均.由于当轮廓发生断裂时, $\omega$  用来控制算法继续进行搜索的角度区间范围,而  $\tau$  用来控制在某个角度方向沿着径向射线搜索的径向区间范围.因此,当轮廓发生断裂时由  $\omega$  和  $\tau$  共同决定的搜索区间是近似扇形的区域.

每个阶段的具体流程为:

1) 第 1 阶段小范围搜索.

已知搜索的初始角度为  $\theta_0$ ,当在某个方向找不到轮廓点时,继续搜索的次数为  $\omega_{\text{small}}$ ,径向搜索区间尺寸为  $\tau$ .搜索过程采用循环扫描的方式,从  $\theta_0$  角度开始,每次步进角度  $\Delta\theta$ ,直到环绕细胞核扫描 1 周,完成一次循环.每扫描一个角度的具体搜索流程如下.

Step1. 步进角度  $\Delta\theta$ ,如果该角度方向的轮廓点已经确定,则更新轮廓位置寄存器,跳过本次搜索,

计数器  $c$  置 0(记录未找到轮廓点的次数),进行下个角度的搜索;否则,将上次扫描确定的轮廓点位置作为当前扫描角度方向的轮廓点估计位置  $\rho_c$ ;

Step2. 沿着该角度方向的径向边缘,设定径向搜索区间为  $[\rho_c - \tau, \rho_c + \tau]$ ,在该区间内,如果只有一个边缘强度的峰值点,则该点作为该角度方向的轮廓点,计数器清 0;如果发现多个峰值点,则取距离位置估计  $\rho_c$  最近的那个点作为该角度方向的轮廓点,计数器清零;如果没有找到峰值点,则该角度没有最佳轮廓点,计数器  $c$  加 1,跳过本次循环继续搜索下个角度;

Step3. 如果计数器  $c$  达到了  $\omega_{\text{small}}$  次仍然找不到轮廓点,说明此处的边缘非常微弱,停止从该角度方向继续搜索,而转到下个已经确定轮廓点位置的角度方向,从这个角度继续搜索.

2) 第 2 阶段大范围动态搜索.

与第 1 阶段搜索过程相似,不同之处在于参数的设置改为大范围动态变化的参数.最大步进次数设置为较大的数值  $\omega_{\text{large}}$ ,径向搜索区间尺寸:

$$\tau_d = \tau \times \text{round}\left(\frac{c}{2}\right), \tag{13}$$

其中  $\text{round}$  表示取整.这时径向搜索区间为  $[\rho_c - \tau_d, \rho_c + \tau_d]$ ,由式(13)控制的径向搜索区间随着  $c$  的变化而动态变化.

由于在经过 3.2 节方法进行增强的细胞增强图像中,细胞轮廓中可能出现断裂的情况.这将会导致在算法得到的细胞轮廓点中,在一部分角度区间未找到合适的轮廓点.此时算法根据该角度区间两端处对应角度的轮廓点位置,采用线性插值的方法将缺失的轮廓点补齐.由于在算法的具体实现中,本文是以细胞核心为坐标原点,将细胞图像变换到极坐标下进行的.在极坐标下采用线性插值的方法将缺失的轮廓点补齐,相当于在原始图像中用圆弧线将断裂的细胞轮廓进行连接.因此本文搜索算法得到的细胞轮廓点将与真实细胞轮廓非常接近.图 9 中给

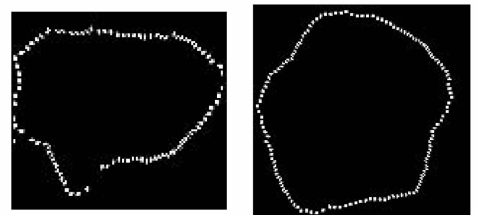


Fig. 9 Located contour points of the two cells in Fig. 6 by the contour searching method.

图 9 搜索算法处理图 6 中 2 幅细胞图像得到的轮廓点

出了该算法得到的细胞轮廓点(为了便于显示,图 9 中的轮廓点是放大显示的),由这些轮廓点组成的轮廓与真实细胞的轮廓已经非常接近.

3.3 GVF Snake 模型轮廓演化

为进一步得到目标的准确轮廓,本文将上述得到的轮廓点组成的轮廓作为 Snake 模型的初始轮廓,采用 GVF Snake 模型进行分割得到最后的细胞轮廓. 本文将边缘增强图  $f_{\text{enhance}}$  与原始边缘图采用式(14)进行组合得到用于计算 GVF 外力场的边缘图像  $f_{\text{fusion}}$ :

$$f_{\text{fusion}}(i,j) = \lambda \times |\mathbf{g}(i,j)| + (1-\lambda) \times f_{\text{enhance}}(i,j), \tag{14}$$

其中,  $\lambda$  为加权参数,在本文中取  $\lambda=0.35$ . 该参数控制了 GVF Snake 模型在进行轮廓逼近的过程中分别对原始细胞边缘图和边缘增强图的依赖权重. 本文所用到的 Snake 模型的定义及求解过程参见文献[11].

4 实验与讨论

实验部分主要验证算法对于单细胞、重叠细胞的分割. 而对于细胞核的分割,由于相关算法<sup>[2,4,7]</sup>得到的细胞核分割结果的准确度已经很高,所以细胞核的分割不是本文研究重点,本文没有给出细胞核的分割结果. 实验用的测试细胞图像样本包括 2 个部分:数据集 1 来自国际上公开的 Herlev 数据集<sup>[5]</sup>,包含 80 幅细胞图像;数据集 2 是由本文作者所在实验室研发的全自动显微镜系统<sup>[10]</sup>所采集的细胞图像数据集,共包含 200 幅单细胞、重叠细胞图像. 我们的采集设备的参数为:系统采用 Olympus 显微镜光路系统及大恒 CCD(charge-coupled device)相机采集图像,其中物镜放大率为 40 倍,细胞图像样本的尺寸约为  $270 \times 270$  像素点.

在细胞图像分割领域,RGVF Snake 算法<sup>[7]</sup>是国际上近年提出的细胞分割算法. 由于 RGVF Snake 算法的测试细胞图像同样来自 Herlev 数据集<sup>[5]</sup>,因此在实验中,我们将本文提出的算法与 RGVF Snake 算法<sup>[7]</sup>作对比,从而验证本文算法对于分割重叠细胞图像的优越性. RGVF Snake 算法的参数设置参照其原文. 本文算法的参数设置为:在 4.2 节动态轮廓搜索中,径向搜索区间的尺寸  $\tau=4$  个像素点宽度,最大步进次数  $\omega_{\text{small}}=3, \omega_{\text{large}}=7$ ;GVF Snake 参数设置参照文献[11].

为定量分析本文算法的分割精度,我们对图像

样本进行了统计实验,并采用 ZSI(Zijdenbos similarity index)相似指数<sup>[17]</sup>计算分割结果的精度:

$$Z_{\text{SI}} = 2 \times \frac{\#\{A_1 \cap A_2\}}{\#\{A_1\} + \#\{A_2\}}, \tag{15}$$

其中,  $A_1$  为真实细胞区域,  $A_2$  为分割结果得到的细胞区域,  $\#\{\}$  表示取图像区域像素点的数量. 真实细胞区域由专业的医师手动分割细胞图像得到. ZSI 指数反映了算法分割结果与真实细胞的相似度,其值越大(接近 1)表明相似度越高. 经过测试,在数据集 1 的实验结果中,有 51 幅图像的分割结果是本文算法与 RGVF Snake 算法的分割结果均较好,有 21 幅图像的分割结果是本文算法更好,有 4 幅图像的分割结果是 RGVF Snake 算法更好,而 2 个算法的分割结果都不好的图像有 4 幅;在数据集 2 的实验结果中,有 115 幅图像的分割结果是本文算法与 RGVF Snake 算法的分割结果均较好,有 61 幅图像的分割结果是本文算法更好,有 9 幅图像的分割结果是 RGVF Snake 算法更好,而 2 个算法的分割结果都不好的图像有 15 幅. RGVF Snake 算法分割效果不好的图像特点是图像光照过强或含有重叠细胞,本文算法分割效果不好的图像特点是图像中细胞的重叠程度比较高(细胞的大部分发生重叠). 另外,个别图像的背景区域染色偏红,对于这种图像 RGVF Snake 算法的分割效果更好. 为了定量分析本文算法与 RGVF Snake 算法的分割精度,本文给出了不同算法的平均 ZSI 指数与标准差对比,如表 1 所示:

Table 1 Average ZSI  $\mu_{\text{ZSI}}$  and Standard Deviation  $\sigma_{\text{ZSI}}$  of Different Methods

表 1 不同算法的平均 ZSI 指数  $\mu_{\text{ZSI}}$  与标准差  $\sigma_{\text{ZSI}}$  对比

| Dataset | $\mu_{\text{ZSI}} \pm \sigma_{\text{ZSI}}$ |                     |
|---------|--------------------------------------------|---------------------|
|         | RGVF Snake                                 | The Proposed Method |
| 1       | 0.937±0.051                                | 0.956±0.043         |
| 2       | 0.920±0.063                                | 0.951±0.059         |

图 10 中给出了一部分细胞图像的分割结果,其中图 10(a)的前 3 幅细胞图像来自数据集 1,后 3 幅细胞图像来自数据集 2,图中标记的轮廓线代表分割得到的细胞区域的边缘轮廓线. 在这些细胞图像中,有的是同种颜色重叠细胞图像,如图 10(a)的第 1 幅细胞图像;有的是不同颜色重叠细胞图像,如图 10(a)第 4 幅细胞图像;还有的既包含同种颜色,也包含不同颜色细胞重叠的复杂重叠细胞图像,如

图 10(a)的第 2 幅细胞图像. 可以看出, 本文算法得到的分割结果更准确, 可信度更高. 而 RGVF Snake 算法很容易受到重叠细胞区域的影响, 使得细胞分割的结果与真实细胞存在着较大的差异. 例如图 10(b)第 1 幅图像, RGVF Snake 算法的分割结果中漏掉了重叠区域的部分细胞, 第 2 幅图像既漏掉了部分亮度较高的细胞区域, 也将一部分属于其他细胞的区域误分割. 图 10(a)第 3 幅细胞图像是异常(可

疑癌变)细胞图像, 虽然该细胞没有重叠, 但是其细胞浆内部亮度差异大, 使得 RGVF Snake 算法出现了误分割. 而本文算法对这样的单个无重叠细胞图像也取得了好的分割结果. 与 RGVF Snake 算法相比, 本文算法的分割效果更好, 结果与真实细胞基本吻合. 从表 1 和图 10 可见, 对于单细胞图像和重叠细胞图像, 本文算法都能够取得较好的分割结果, 算法具有更高的准确性.

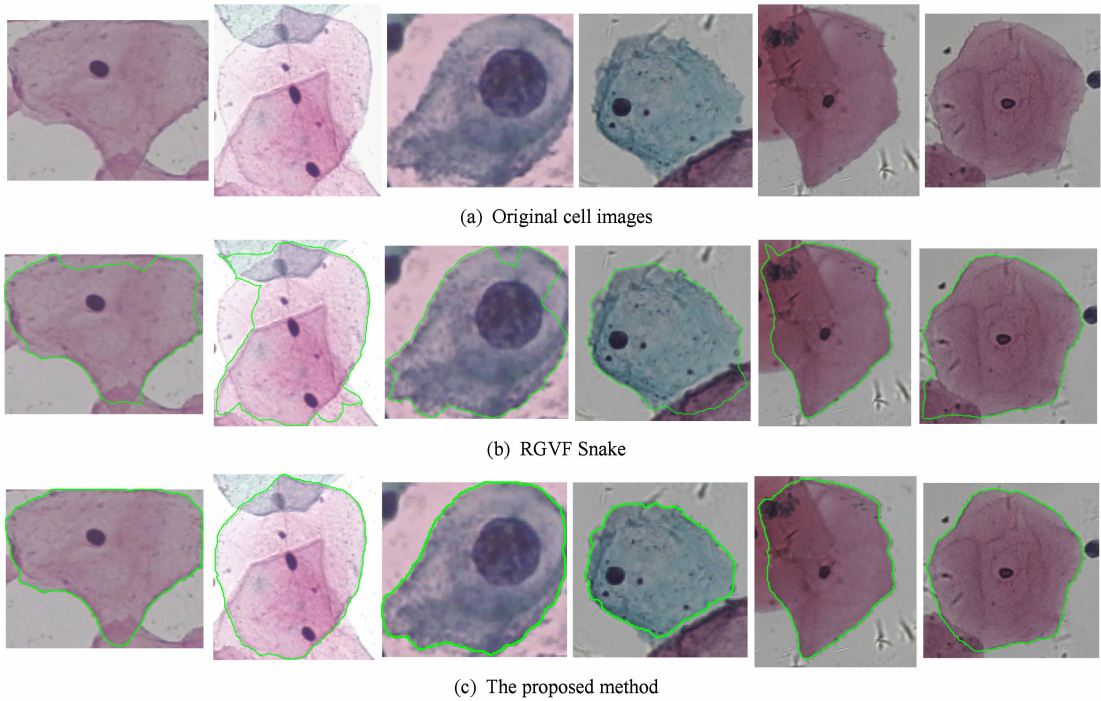


Fig. 10 Cell image segmentation results.  
图 10 细胞分割结果

在实际采集的宫颈细胞图像中, 细胞的形状变化多样, 细胞重叠的方式也是变化多样的. 因此单一的图像分割算法很难有效处理各种各样的重叠细胞的分割. 但是本文研究的是一类比较典型的细胞重叠情况, 即相邻细胞彼此的一部分细胞浆发生重叠. 算法对这一类细胞重叠分割能够取得比较好的分割结果, 而且对于一些更为复杂的情况, 也能得到较好的分割结果. 但是本文算法并不适合那一类极端复杂的细胞重叠问题的分割处理, 例如细胞的变形很严重(不再是凸状的), 或者一个细胞完全嵌入到另一个细胞中的情况.

5 结 论

本文研究了医学图像分割中重叠细胞图像的分割. 算法主要针对的是显微镜高倍物镜采集的单个

细胞图像和重叠细胞图像的分割研究, 在图像中只有一个主体细胞(文献[6-7]的研究同样是针对这类细胞图像). 有的细胞图像中只含有一个单独的细胞, 这种情况相对容易处理. 但是, 还有很多图像中的主体细胞与相邻细胞发生重叠, 致使细胞浆的大部分出现重叠, 这类图像的分割难度大、挑战性强. 本文在一定程度上解决了这种类型重叠细胞图像的分割问题.

通过对细胞轮廓的几何结构的分析, 本文提出了细胞的稀疏轮廓点模型, 并将该模型与 GVF Snake 模型结合提出了一种新颖的重叠细胞图像分割算法. 该算法的分割过程是自动不需人工干预的. 对 2 个细胞图像数据集进行的细胞分割实验结果证明, 本文算法能够有效实现单细胞图像、重叠细胞图像的分割. 目前我们正致力于研究针对显微镜低倍物镜采集的低分辨率多细胞图像的分割. 这类图像的

分割目的主要是定位图像中的细胞核与实体细胞区域,然后可以结合本文算法研究的高倍物镜下细胞图像的精细分割,使得算法更能够接近工程实际的应用。

## 参 考 文 献

- [1] Fan Jinping. The research on the method of cervical cell image segmentation and recognition [D]. Guangzhou: Jinan University, 2010 (in Chinese)  
(范金坪. 宫颈细胞图像的分割和识别方法研究[D]. 广州: 暨南大学, 2010)
- [2] Plissiti ME, Nikou C, Charchanti A. Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering [J]. IEEE Trans on Information Technology in Biomedicine, 2011, 15(2): 233-241
- [3] Bamford P, Lovell B. Unsupervised cell nucleus segmentation with active contours [J]. Signal Processing, 1998, 71(2): 203-213
- [4] Genctav A, Aksoy S, Onder S. Unsupervised segmentation and classification of cervical cell images [J]. Pattern Recognition, 2012, 45(12): 4151-4168
- [5] Martin E. Pap-smear classification [D]. Lyngby, Denmark: Technical University of Denmark, 2003
- [6] Yang-Mao SF, Chan YK, Chu YP. Edge enhancement nucleus and cytoplasm contour detector of cervical smear Images [J]. IEEE Trans on Systems, Man, and Cybernetics-Part B, 2008, 38(2): 353-366
- [7] Li K, Lu Z, Liu W, et al. Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake [J]. Pattern Recognition, 2012, 45(4): 1255-1264
- [8] Jung C, Kim C. Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization [J]. IEEE Trans on Biomedical Engineering, 2010, 57(10): 2600-2604
- [9] Park C, Huang JZ, Ji JX, et al. Segmentation, inference, and classification of partially overlapping nanoparticles [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35(3): 669-681
- [10] Guan Tao, Zhou Dongxiang, Liu Yunhui. Color optical microscopic cell image segmentation based on color difference vector field [J]. Acta Optica Sinica. 2014, 34(1): No. 0115001 (in Chinese)  
(关涛, 周东翔, 刘云辉. 基于色差向量场的彩色光学显微细胞图像分割[J]. 光学学报, 2014, 34(1): No. 0115001)
- [11] Xu C, Prince JL. Snakes, shapes, and gradient vector flow [J]. IEEE Trans on Image Processing, 1998, 7(3): 359-369
- [12] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models [J]. International Journal of Computer Vision, 1988, 1(4): 321-331
- [13] Soille P. Morphological Image Analysis: Principles and Applications [M]. New York: Springer, 1999: 100-150
- [14] Gonzalez RC, Woods RE. Digital Image Processing [M]. 2nd ed. New York: Prentice Hall, 2002: 120-180
- [15] Guan Tao, Zhou Dongxiang, Liu Yunhui, et al. Classification of cervical cell images based on adaptive thresholding segmentation [J]. Signal Processing, 2012, 28(9): 1262-1270 (in Chinese)  
(关涛, 周东翔, 刘云辉, 等. 基于自适应阈值分割的宫颈细胞图像分类算法[J]. 信号处理, 2012, 28(9): 1262-1270)
- [16] Harandi NM, Sadri S, Moghaddam NA, et al. An automated method for segmentation of epithelial cervical cells in images of ThinPrep [J]. Journal of Medical Systems, 2010, 34(6): 1043-1058
- [17] Zijdenbos A, Dawant B, Margolin R, et al. Morphometric analysis of white matter lesions in MR images: Method and validation [J]. IEEE Trans on Medical Imaging, 1994, 13(4): 716-724



**Guan Tao**, born in 1984. PhD candidate. His research interests include image processing, computer vision.



**Zhou Dongxiang**, born in 1968. PhD, professor. His major research interests include computer vision, image processing (dxzhou@nudt.edu.cn).



**Fan Weihong**, born in 1970. PhD, associate professor. Her major research interests include wireless communication, wireless sensor networks, and intelligent signal processing(whfan@nudt.edu.cn).



**Liu Yunhui**, born in 1965. PhD, professor. Fellow of IEEE. His research interests include artificial intelligence, robotics and automatic control(yunhui.liu@gmail.com).