

# 一种重叠胃腺癌细胞分离线估计算法

张云 刘建平

(国防科技大学机电工程及自动化学院, 长沙 410073)

**摘要** 针对胃腺癌细胞显微图像中经常出现重叠(或粘连)现象,对后续进一步的研究工作产生了很多不良影响,为了准确地分离重叠(或粘连)的胃腺癌细胞,采用了一种数学形态学与改进的B样条活动轮廓模型算法相结合的方法。该算法首先对图像进行二值化处理,然后利用变元腐蚀得到细胞的非线性距离图和背景的非线性距离图,接着对两者进行叠加,即得到一幅新的非线性距离图,最后使B样条曲线准确收敛到目标边界,以便自动分离细胞。实验结果表明,该算法不但能有效地构造分离线,并能估计出重叠(或粘连)胃腺癌细胞的原始形状。

**关键词** 数学形态学 B-Snake 分离线

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2007)11-2058-04

## A Estimating Separate Line Algorithm for Overlapping Gastric Adenocarcinoma Cell Image

ZHANG Yun, LIU Jian-ping

(College of Mechatronics Engineering & Automation of National University of Defense Technology, Changsha 410073)

**Abstract** Computer-aided medical image processing and analysis is a important application in the field of image processing. The introduction of computer-aided diagnosis system of cancerous tissues is highly helpful to clinical diagnosis and medical studies, especially in the case of lacking specialists. But the technology is difficult to put into practice. For example there are many overlapping cells in the gastric adenocarcinoma image because of the facture of the micrograph of gastric adenocarcinoma. If the cells are not separated accurately, the following research will not go on well. It is important to draw the separate-line and gain the borderline of each cell. So the paper presents a algorithm which integrated mathematical morphology with an improved B-spline active contour model (B-Snake model) in order to segment the overlapping parts in the gastric adenocarcinoma image. The algorithm adopted an improved nonlinear distance image to force B-spline curve close to the objective boundary automatically. The experimental result of shows that our algorithm can construct the separate-line exactly and draw the original figures of cells effectively.

**Keywords** mathematical morphology, B-Snake, separate-line

## 1 引言

借助计算机对肿瘤细胞的显微图像进行图像处理和分,以实现自动识别,在计算机辅助诊断中,尤其在缺少专家的情况下具有十分重要的意义。胃腺癌细胞自动识别系统的建立对于临床诊断、病理分析、治疗以及教学有着尤为重要的意

义。胃腺癌细胞区域的定位与分割是建立胃腺癌细胞自动识别系统的一项重要技术,其直接关系到癌病理形态学特征提取的可靠性。在实际的图像处理过程中,由于经常出现多个细胞重叠(或粘连)的现象,因此给图像后续的分析处理带了诸多困难。

文献[1]采用距离变换与分水岭算法相结合的方法,先把图像二值化,再通过距离变换得到灰度图像,然后用分水岭算法进行分割。文献[2]通过搜

索凹点,先构成分离点对,然后通过连接分离点来构成分离线。文献[3]基于数学形态学的方法对重叠细胞进行分割。由于前述这些方法都是简单地用分离线隔开重叠细胞,因此容易丢失细胞原来的一些形态信息,不利于图像后续的分析处理。

本文针对胃腺癌细胞图像中出现重叠(或粘连)现象,采用了一种数学形态学梯度与改进的 B 样条活动轮廓模型算法相结合的方法,有效地分离出了各细胞在没有发生重叠时的原始形状。该算法是先应用可变结构迭代腐蚀得到种子点和非线性距离图,并根据最大距离图的轮廓对 B 样条曲线进行初始化,然后利用原图像梯度和改进的非线性距离图,共同使 B 样条曲线准确收敛到目标边界。

## 2 传统的 Snake 模型

Kass 等人提出的 Snake 模型<sup>[4]</sup>,其基本思想是在图像中找出一条闭合的参数化曲线,并使该曲线处的总能量达到极小值。设  $s \in [0, 1]$  代表曲线归一化的长度,则可用  $v(s) = (x(s), y(s))$  表示曲线,而经典活动轮廓的能量则可以由下面的公式表示:

$$E_{Snake} = \int_0^1 E(v(s)) ds$$

$$= \int_0^1 \{ [E_{int}(v(s))] + [E_{ext}(v(s))] \} ds \quad (1)$$

其中,内部能量为

$$E_{int}(v(s)) = \alpha(s) \left| \frac{dv(s)}{ds} \right| + \beta(s) \left| \frac{d^2v(s)}{ds^2} \right| \quad (2)$$

式中,第 1 项为 Snake 曲线内部能量函数的一阶项,用于控制曲线轮廓的长度,并使其尽可能的短;第 2 项为 Snake 曲线内部能量函数的二阶项,用于控制轮廓线曲率,使轮廓线尽可能平滑,而 Snake 曲线外部能量  $E_{ext} = E_{image} + E_{cont}$  来自图像和外部的约束,由于  $E_{ext}(v(s))$  为外部能量函数,其多以图像的梯度充当<sup>[5]</sup>,它是活动轮廓线的收缩限制力,用于控制轮廓线,使其接近梯度变化大的区域,因此外部能量定义为

$$E_{ext}(v(s)) = -|\nabla I(v(s))|^2 \quad (3)$$

内部能量  $E_{int}(v(s))$  用于控制轮廓曲线的连续性和光滑性,其中  $\alpha(s)$  和  $\beta(s)$  为正的加权系数;外部能量  $E_{ext}(v(s))$  用于将曲线吸引到图像中的目标界。根据变分学中 Euler-Lagrange 方程,可得到以下一个随时间演变的方程:

$$\frac{\partial v(s,t)}{\partial t} - \frac{\partial}{\partial s} \left[ \alpha(s) \frac{\partial v(s,t)}{\partial s} \right] + \frac{\partial^2}{\partial s^2} \left[ \beta(s) \frac{\partial^2 v(s,t)}{\partial s^2} \right] + \nabla E_{ext}(v(s,t)) = 0 \quad (4)$$

如果  $\frac{\partial v(s,t)}{\partial t} = 0$ , 那么就得到了问题的解,即轮廓线收敛到目标边界。

## 3 改进的 B-Snake 模型

本文提出的改进的 B-Snake 模型是首先通过可变结构的迭代腐蚀求出种子点,并在距离图上找出各个最大的分离区域;然后在各分离区域边缘上,逆时针方向间隔一定角度  $\alpha$  取一系列点  $\{p_i\}$ , 作为 B 样条曲线的控制点,即完成了模型的初始化。

由于 B 样条曲线本身具有光滑性,因此可省略计算内能,使模型的计算简化,以提高计算效率。模型总能量为

$$E_{Snake} = E_{ext}(v(s)) \quad (5)$$

其中,  $E_{ext}(v(s))$  为模型外部能量,  $v(s)$  为归一化曲线长度。

外部能量  $E_{ext}(v(s))$  由图像能量项和非线性距离约束能量项两部分能量组成,这两部分能量分别由原图像和改进的距离图像所决定。其中非线性距离约束能量项在轮廓曲线距离细胞边界较远时发挥着主要作用,这时该能量项用于把曲线推动到细胞边界附近。在细胞边界附近时,非线性距离约束能量项能量较小,而图像能量项则起着主要作用,这时,轮廓曲线将终止到细胞边界上。外部能量  $E_{ext}(v(s))$  可以表示为

$$E_{ext}(v(s)) = E_{image}(v(s)) + E_{dist}(v(s)) \quad (6)$$

式(6)中  $E_{image}(v(s))$  为图像能量,  $E_{dist}(v(s))$  为非线性距离约束能量。图像能量  $E_{image}(v(s))$  可通过梯度图求得,由式(3)定义可知,该项能量与经典活动轮廓算法的外部能量相同。

由于改进的距离图像中,当轮廓线深入到细胞内部,或是远离细胞时,不但非线性距离值很大,变化迅速,并且如果轮廓曲线对细胞边界的偏离越大,则在距离图中对应位置的非线性距离值越大,变化越迅速;当轮廓线接近细胞实际边界时,距离图中对应像素的灰度值很小,而且灰度变化非常缓慢,因此,可以将非线性距离约束能量项  $E_{dist}(v(s))$  定义为

$$E_{dist}(v(s)) = \lambda D(v(s)) \quad (7)$$

其中,  $\lambda$  为一个加权系数, 其与产生距离图时的  $r_0$ ,  $r_{min}$ ,  $h$  有关。  $D(v(s))$  为改进后的非线性距离值。若  $-\lvert \nabla I(v(s)) \rvert^2 + \lambda D(v(s)) = 0$ , 即可使曲线收敛到目标轮廓。

#### 4 在重叠胃腺癌细胞图像分割中的应用

分离重叠细胞的工作是建立在细胞正确计数和搜索正确的细胞种子点位置的基础上。图 1 为需要处理的胃腺癌细胞显微图像, 图上已经圈注了本文算法所要处理的部分, 在对该图癌细胞进行分割前, 首先对胃腺癌细胞显微图像进行一系列的预处理, 得到二值化图像; 然后通过连续多次的腐蚀运算来求得二值图像的极限腐蚀<sup>[6,7]</sup>, 极限腐蚀是每个连通区域被连续腐蚀到完全消失前一次结果的并集(见图 2)。在此过程中, 每一步都要进行迭代, 且像素以迭代次数作为标记, 即得到距离变换图。实验中不难发现, 小尺寸腐蚀元会导致虚假像素聚集现象, 而大尺寸腐蚀元则会使检测区域丢失。为准确找到种子点, 本文是采用可变结构迭代腐蚀来得到非线性距离图。

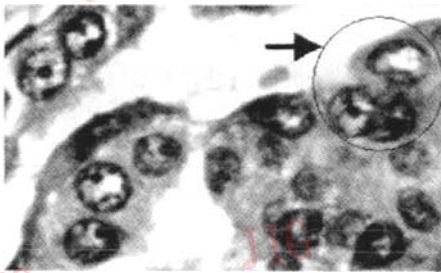


图 1 胃腺癌细胞显微图像  
Fig. 1 The micrograph of gastric adenocarcinoma

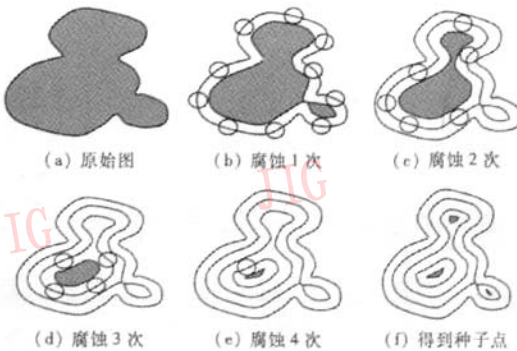


图 2 极限腐蚀  
Fig. 2 Utmost erosion

变元迭代腐蚀法中, 可设最初的腐蚀元半径为  $r_0$ , 最小腐蚀元半径为  $r_{min}$ , 定义第  $k$  次腐蚀的腐蚀元半径为

$$r_k = \text{round} \left( \max \left\{ \frac{r_0}{h^{k-1}}, r_{min} \right\} \right) \quad (8)$$

其中, 函数  $\text{round}(x)$  表示把  $x$  四舍五入到整数位,  $h$  为比例系数。一般情况下,  $r_{min}$  可以取 1, 此时的腐蚀元即为十字形结构。  $r_0$  应根据图像细胞类型、采集分辨率和定标系数等来确定, 在实际应用中,  $r_0$  一般在 3~8 之间取值。若细胞面积较大, 则  $r_0$  应取较大值; 若细胞面积较小或是细胞大小差异显著的图像, 则  $r_0$  应取值较小。  $h$  应大于 1, 一般取 1.2~1.5, 当取稍大值时, 则不仅可以增大腐蚀元尺寸减小的速度, 且容易出现连通区域尚未分离就被全部腐蚀消失的情况; 若取稍小值时, 则可以加快腐蚀速度。

在迭代过程中, 腐蚀元产生变化的腐蚀次数为

$$T = \left[ \frac{\ln r_0 - \ln r_{min}}{\ln h} \right] + 1 \quad (9)$$

其中,  $[\ ]$  表示取整,  $h$  为迭代腐蚀时的比例系数。

分割时, 首先对胃腺癌细胞显微图像(图 3(a))进行 RGB 通道分割, 即取 R 通道图像(图 3(b)), 经一系列预处理得到二值化图像(图 3(c)), 由非线性距离图(图 3(d))中可以看出, 细胞重叠部分的距离值都很小, 因此包含距离值作用的约束力就保证了 Snake 曲线不会向着与之重叠的细胞内部演化。本文将背景的非线性距离图和细胞的非线性距离图进行叠加, 得到一幅新的非线性距离图(图 3(e))。此图中, 由于距离极小值区域包含细胞与背景的分界区和细胞之间的重叠区, 因此可采用改进后的距离值作为控制力, 加上 Snake 曲线的光滑性, 在曲线远离细胞边界或是深入细胞内部时进行约束, 确保 Snake 曲线的演化被限制在图 3(f)中的黑色区域, 以便最终收敛到细胞边界。图 3(h)为两个重叠细胞的分离结果, 并与文献[8]提出的等值线跟踪算法的分割结果(图 3(g))进行了比较。

由图 3(g)可以看出, 文献[8]的方法虽然分离了重叠细胞, 但是因其仅仅依靠二值图像, 所以只能用简单的分隔线分离细胞, 而且无法准确定位各细胞重叠部分的边界。而本文则采用数学形态学与改进的 B 样条活动轮廓模型算法相结合的方法, 在进行变元腐蚀迭代时, 取  $r_0 = 8, r_{min} = 1, h = 1.6$ , 模型的加权系数  $\lambda = 1.0$ 。从图 3(h)可以看出, 本文算

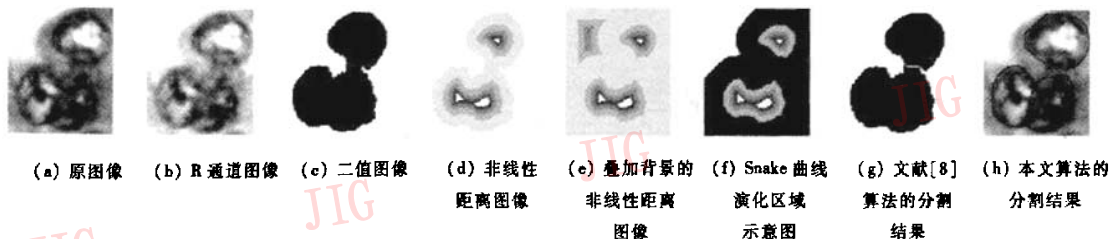


图 3 分离 3 个重叠胃腺癌细胞

Fig. 3 Separate three overlapping gastric adenocarcinoma cell

法不但分离了重叠细胞,而且构造了有效的分离线,并确定了重叠部分的完整边界。

## 5 结 论

实验表明,本文算法在运行过程中无需人工干预,利用背景的非线性距离图和细胞的非线性距离图进行叠加得到的新的非线性距离图,即可使 B 样条曲线准确收敛到目标边界,并能较准确地估计出重叠细胞分离线和再现细胞的原始形状,从而避免了因简单分离重叠细胞而无法准确定位各细胞重叠部分边界和导致丢失细胞原始形态信息的缺点。经本文算法处理后,可准确得到细胞面积、周长、形态因子等参数信息,从而为后续的胃腺癌细胞图像自动识别系统的建立奠定了良好的基础。由于胃腺癌细胞图像较为复杂,而模型参数和腐蚀参数的设定比较困难,因此今后应在参数设定的自动化方面对算法进行改进。

### 参考文献 (References)

- 1 Liu Sheng-hao, Zeng Li-bo, Liu Bin, *et al.* Separating algorithm for overlapping granule images [J]. *Computer Engineering*, 2002, 28(2): 198 ~ 210. [刘生浩,曾立波,刘斌等. 重叠颗粒图像的分离[J]. *计算机工程*, 2002, 28(2): 198 ~ 210.]
- 2 Lu Zong-qi, Tong Tao. The application of chain code sun in the edge form analysis [J]. *Journal of Image and Graphics*, 2002, 7(12): 1323 ~ 1328. [陆宗骥,童滔. 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.]
- 3 Ma Dong, Cao Pei-jie, Pan Kai-li, *et al.* Comparison of some methods for segmentation of overlapped nuclei [J]. *Beijing Biomedicine Engineering*, 1999, 18(3): 142 ~ 147. [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.]
- 4 Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models [J]. *International Journal of Computer Vision*, 1988, 1(4): 321 ~ 331.
- 5 Qiu Shu-bo, Wang Hua-xiang, Liang Zhi-wei. Tracking object contour using a novel B-Snake algorithm [J]. *Journal of Image and Graphics*, 2005, 10(5): 585 ~ 589. [邱书波,王化祥,梁志伟. 一种新的 B-Snake 算法在目标轮廓跟踪中的应用[J]. *中国图象图形学报*, 2005, 10(5): 585 ~ 589.]
- 6 Hu Jiong-jiong, Yu Hui-min, Fang Bo. Automatic cell image segmentation based on B-Snake model with constraint of morphology [J]. *Journal of Image and Graphics*, 2005, 10(1): 31 ~ 37. [胡炯炯,于慧敏,房波. 基于形态学约束的 B-Snake 模型的细胞图像自动分割算法[J]. *中国图象图形学报*, 2005, 10(1): 31 ~ 37.]
- 7 Pratt William K. *Digital Image Processing* [M]. New York: John Wiley&Sons, Inc. 2001.
- 8 Lu Zhong-qi. *C/C++ Image Pross* [M]. Beijing: Tsinghua University Press, 2005. [陆宗骥. *C/C++ 图像处理编程* [M]. 北京:清华大学出版社, 2005.]