

文章编号:1001-9081(2006)02-0400-03

基于综合技术的组织细胞图像分割

万卫兵,施鹏飞

(上海交通大学 图像处理与模式识别研究所, 上海 200030)

(wbwan@sjtu.edu.cn)

摘要:在组织工程培养细胞图像检测中,需要得到较完整的清晰轮廓线以获取生长信息。为了适应这种需求,提出了一种边缘检测综合算法。首先应用小波尺度独立边缘检测方法,得到简洁的边缘;然后采用图像融合的方法,结合 Canny 和动态聚类方法得到最终的细胞分割图。实验结果证明了该算法的可行性和有效性。

关键词:组织细胞; 小波变换; 边缘检测; 图像融合

中图分类号: TP391.41 **文献标识码:**A

Tissue cell segmentation based on integrate techniques

WAN Wei-bing, SHI Peng-fei

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: To investigate cultural cell growth information in tissue engineering, it is important to get clear contour line of cell. An edge detection algorithm was proposed. First it got rough edge by scale_independ algorithm of wavelet transform, then adopted algorithm of image fusion that combines Canny and clustering measures to get final segmenting result. It is proved validity and feasibility by experiment result. As an integrate technique for image segmentation can apply widely in image processing field.

Key words: tissue cell; wavelet transform; edge detection; image fusion

0 引言

目前,已有多种研究细胞的技术,从光镜到电子显微镜,从一般细胞化学法到免疫化学法。细胞在体外培养过程中需要间隔一段时间进行常规检查和显微镜观察,及时了解细胞生长状态、数量改变、细胞形态及细胞有无移动等等。细胞形态的观测处于动态中,因此它的细胞分割相对较难确定,但细胞大致的轮廓一定时期基本可以保持不变,所以得到基本轮廓后就可以精确确定它的其他参数及生长情况。

传统边缘提取算法以原始图像为基础,对噪声信号和边缘信号不加区分地使用边缘提取算子,如 Robert 算子、Laplacian 算子、Sobel 算子、Kirsch 算子等。以上算子都是使用差分算子,利用临近边缘地方的一阶和二阶方向导数变化来检测边缘。它们大多数可以分为两类:梯度方向法和零交叉点法。前者是通过寻找图像一阶导数的局部最大和最小值来进行边缘检测,后者通过寻找图像二阶导数引出的零交叉点去发现边缘。对图像而言,奇异点往往构成图像的边缘。

另一种有效的边缘提取方法是多尺度法,其思路是:在大尺度下抑制噪声,可靠地识别边缘;在小尺度下精确定位。一般的多尺度方法都是利用图象金字塔,多以减少计算量为主要目标;而文献[1]则是利用不同尺度的高斯函数的一次微分与图象卷积,以局部极大值点为边缘点,由粗到精确定图象的边缘,获得了较好的结果^[1]。但是,文献[1]由于采用高斯函数的一次微分作为卷积核,算法计算量大,且不能确定边缘的类型。

文献[2]采用二次样条二进小波对图象进行小波变换,

利用小波变换天生的多尺度特性,提取不同精度、不同奇异度的图象边缘,综合形成图象边缘,获得了良好效果。

根据数学分析上的 Lipschitz 知识,函数的奇异点主要集中在阶跃结构和端顶结构上,在图像处理中阶跃结构的边缘通常被认为是图像的轮廓。为提取图像轮廓,我们设想能不能构造这样一种小波变换,使它不仅能在不同的方向上进行边缘检测,而且能提取阶跃结构的边缘,还能消除噪声。多尺度小波就具有这样的功能。用于轮廓提取的尺度独立算法,就是基于多尺度小波变换理论,通过小波变换的振幅的局部最大值提取图像的边缘,并利用其与变换尺度无关性,有效地检测出阶跃结构的图像边缘,并消除噪声,从而提取出图像的轮廓。

1 边缘的奇异度和多尺度边缘的综合

图象的边缘对应于图象灰度的奇异点,不同类型的边缘有不同的奇异度。一个刻划奇异度的重要指标是 Lipschitz 指数。基于多尺度小波变换的局部极大值提取图象边缘方法的一个突出优点就是边缘点的振幅 $W_{2^j}f(x,y)$ 与 Lipschitz 指数 α 之间存在密切的联系,即:

$$|W_{2^j}f(x,y)| \leq k(2^j)^\alpha \quad (1)$$

式(1)可化为:

$$\log_2 |W_{2^j}f(x,y)| \leq \log_2 k + j\alpha \quad (2)$$

式(2)的求解等效于求下式极小值:

$$J = \sum_{j=1}^n (\log_2 |W_{2^j}f(x,y)| - \log_2 k - j\alpha)^2 \quad (3)$$

显然,已知边缘点 3 个以上尺度的振幅 $W_{2^j}f(x,y)$,由式

收稿日期:2005-08-19; 修订日期:2005-10-26

作者简介:万卫兵(1969-),男,江西南昌人,讲师,博士研究生,主要研究方向:生物医学图像处理与分析、图像压缩等; 施鹏飞(1939-),男,上海人,教授,博士生导师,主要研究方向:医学图像分析、生物特征识别。

(2) 可估计Lipschitz指数 α 。平滑的图象边缘可看成是原始边缘与高斯函数卷积的结果,有:

$$|W_{2^j}f(x,y)| \leq k2^{-j}(s)^{\alpha-1}, s = \sqrt{2^{2^j} + \sigma^2} \quad (4)$$

同样,可通过求下式的极小值估计 Lipschitz 指数 α 。

$$J = \sum_{j=1}^n (\log_2 |W_{2^j}f(x,y)| - \log 2^k - j - (\alpha - 1)s)^2 \quad (5)$$

一般说来边缘越平滑,奇异度越大,例如阴影类型边缘比较平滑,指数 α 较大,而几何类边缘的指数相对小些。在实用中则可通过设定阈值,输出一种类型的边缘而抑制其他类型的边缘。

2 尺度独立的算法

由于不同的 α 值对应于不同的几何边缘结构, α 越小, 光滑性越弱。根据图像处理知识,我们知道图像边缘的种类主要有两种:一种是阶跃结构的边缘,边缘点两边的像素的灰度值有着显著的不同;另一种是端顶结构的边缘,它位于灰度值从增加到减少的转折点。在实际的应用中,我们取两种特殊的 α 值,构成图像的基本边缘结构,就是 $\alpha = 0$, 它对应于阶跃结构的边缘; $\alpha = 1$, 它对应于端顶结构的边缘。对于这两种边缘结构取多尺度小波变换,可以证明阶跃结构阶跃结构边缘的小波变换的振幅是一个非零的常量,且与方向向量和变换的尺度无关。端顶结构边缘的小波变换的振幅随变换尺度变化而变化。同时可以证明噪声的多尺度小波变换是随着尺度的变换而变化的,尺度越大,变化越小^[3~4]。

在上述分析基础上可以得到尺度独立算法。该算法主要用于提取图像中的阶跃结构边缘,删除其他结构的边界,同时可以消除噪声。阶跃结构边缘的小波变换是一个与尺度无关的非零常数,即对不同的尺度的小波变换是相互等同的,在阶跃结构边缘点上,显然有:

$$\frac{W_{2^i}f(x,y)}{W_{2^j}f(x,y)} = 1, (i,j = 1,2,\dots,J) \quad (6)$$

虽然阶梯结构边缘的小波变换的振幅是非零的常量,但由于实际中的边缘总不是理想的,这样对阶跃结构边缘的小波变换来说,只能:

$$\frac{W_{2^i}f(x,y)}{W_{2^j}f(x,y)} \approx 1, (i,j = 1,2,\dots,J) \quad (7)$$

满足式(7)的点能被认为是阶跃结构边缘,但这个约等号提取的边缘点是不确定的,为了避免之,我们使用一个实数 R 去代替 1, 让 R 非常接近于 1, 但又比 1 大, 则有:

$$\frac{1}{R} \leq \frac{W_{2^i}f(x,y)}{W_{2^j}f(x,y)} \leq R, (i,j = 1,2,\dots,J) \quad (8)$$

这样,令 K, L 为有限的整数,结合上面的边缘检测方法,对于二维数字图像信号 $f(k,l)$ 我们得到下面的尺度独立算法^[5]:

1) 取不同的尺度 2^j , 计算出 $W_{2^j}f(k,l)$ 及 $A_{2^j}f(k,l)$, $(1 \leq j \leq J)$ 。

2) 确定阈值 T , 有:

$$|W_{2^j}f(k,l)| \geq T$$

且沿 $A_{2^j}f(k,l)$ 所代表的 4 个梯度方向达到局部最大值。

3) 确定实数 R , 有:

$$\frac{1}{R} \leq \frac{W_{2^i}f(k,l)}{W_{2^j}f(k,l)} \leq R, (i,j = 1,2,\dots,J)$$

那么满足上面条件的点 (k,l) 被认为是图像阶跃结构边

缘上的点,是图像的轮廓上的点。

3 图像融合的算法

图像融合就是将不同来源的同一对象的图像数据进行空间配准,然后采用一定的算法将各图像数据中所含的信息优势或互补性有机地结合起来产生新图像数据的技术。这种新数据具有描述所研究对象的较优化的信息表征,同单一信息源相比,能减少或抑制对被感知对象或环境解释中可能存在的多义性、不完全性、不确定性和误差,最大限度地利用各种信息源提供的信息。这里采用几种不同方法进行边缘检测,作为不同的信息源来对它们进行图像融合。

文献[1]的 Canny 算法已被证明是一非常有效的边缘检测方法,所以它的结果可以作为一种信息源进行比较参考。

另一种为基于动态聚类的边缘检测方法。在传统的阈值法图像分割中,往往将图像简单分成两类,这通常不符合实际情况,特别是对组织细胞。事实上,在图像分割中经常不能预先知道要分割成的类别数。动态聚类方法正好能解决这一难题,因为它可以将图像分割成任意设定的类别数。在此基础上利用边缘检测算子得到图像的边缘点。

加上前面介绍的小波尺度独立算法,这样对一幅细胞图像可以得到几种不同的边缘检测结果,它们对边缘的检测各有特点。这样可以利用它们的优点和互补,得到更为准确的边缘点。

由以上的分析,进一步可以得到基于图像融合的边缘分割算法,它的具体步骤如下:

1) 对原始图像分别采用尺度独立算法,Canny 算子和动态聚类算法进行边缘检测。

2) 将所得到的三副边缘图像对应点的像素值加权平均,其中尺度独立算法占得权值最重,把三副图像融合为一副新的图像。

3) 进行灰度阈值处理,去除不需要的多余信息。

4) 进行形态学腐蚀、膨胀、细化和边界闭合得到最终的边缘图像。

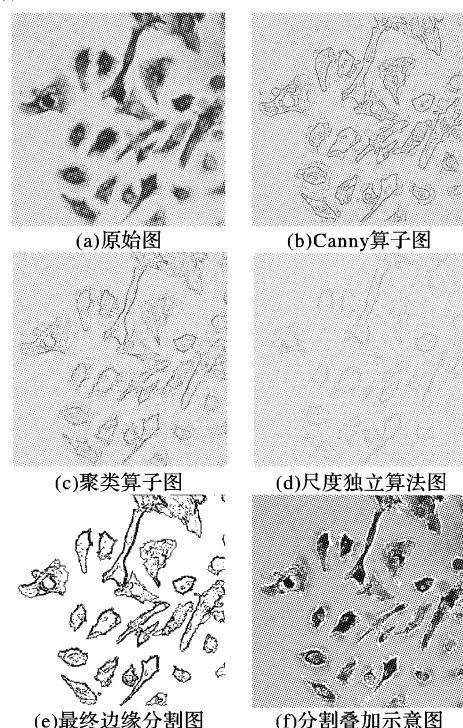


图 1 实验结果

4 实验结果与讨论

实验中选取一张培养细胞灰度图像,大小为 256×256 ,小波变换选取的是紧支集的二次 B 样条小波。

从图 1 中可以看出 canny 算法边缘太细,有些噪声;尺度独立算法连续性不太好,有边缘丢失;聚类算子介于两者之间,与分类数有很大关系;而采用本文算法能有效弥补三种方法的不足,又保持了它们的优点,取得了良好的折衷。边缘图像具有很好的清晰度和连续性,边缘信息比较完备,基本分割出了细胞的大致轮廓。

表 1 细胞边缘分割点的方法比较

方法	正确边缘点	误判点	正确率
Canny	14 534	2 346	83.8%
聚类	13 589	3 321	75.6%
尺度独立	13 986	2 312	83.4%
融合综合	15 123	2 190	85.5%

(上接第 392 页)

```

 $F_i^k = F_i^k \cup c;$ 
else
  delete  $c$  from  $L_i^k$ ; /* 根据定理 3 局部剪枝 */
}
}

3) 各节点都获取所有节点局部频繁项集的并集
for(  $i = 1; i <= n; i++$ )
   $P_i$  send  $F_i$  to  $P_a$ ;
  /*  $F_i$  表示  $P_i$  的局部频繁项集;  $P_a$  表示中心节点 */
   $P_a$  combine  $F_i$  equal  $F'$ ;
  /*  $F' = \bigcup_{i=1}^n F_i$ , 表示所有节点局部频繁项集的并集 */
   $P_a$  broadcast  $F'$ ;
4) 计算项集的全局频度
for(  $i = 1; i <= n; i++$ )
  { for each items  $d \in F'$ 
     $P_i$  send  $d.si$  to  $P_a$ ; /* 利用哈希树计算  $d.si$  */
  }
  for each items  $d \in F'$ 
     $d.s = \sum_{i=1}^n d.di$  /*  $d.s$  表示项集  $d$  的全局频度 */
5) 获得全局频繁项集  $F$ 
for each items  $d \in F'$ 
  if(  $d.s \geq min\_sup \times M$ )
    /*  $M$  表示全局数据库  $DB$  的元组数 */
     $F = F \cup d$ 

```

3 算法分析与性能测试

为了测试算法的性能,将 FPMFI 算法与经典的挖掘频繁项集的并行算法 CD、DD 和 FDM 算法进行性能比较。测试环境为 10M 局域网,采用 5 台联想 PC 机构成分布式计算机节点,其中 1 台作为中心节点,各 PC 机配置均为 P4 2.4G,内存 512M,Windows2000 professional 操作系统,SQL Server 2000 数据库管理系统。实验数据来自某大型商业连锁店的销售数据。在测试中,采用水平等间距投影方法分别在 5 台 PC 生成

从表 1 中可以看出,本文方法对细胞边缘点的检测上具有良好的正确率。本文提出的算法对于组织细胞的动态检测分割具有良好的效果,为下一步的参数测定和观察打下良好的基础。同时,该方法也可以应用到其他类型图像的检测分割。

参考文献:

- [1] CANNY J. A Computational Approach to Edge Detection. IEEE T-PAM II[J]. 1986, 8(6): 679 - 698.
- [2] MALLAT SG, ZHONG S. Charaterization of Signals from Multiscale Edges. IEEE T-PAM II[J]. 1992, 14(7): 710 - 732.
- [3] ER PIERRE K, MARC LJ, SAINT JP, et al. Wavelet based multi-fractal formalism to assist in diagnosis in digitized mammograms. Image Anal Stereo II[J]. 2001, 20 (3) : 169 - 174.
- [4] IEL FERNA'NDEZ G, BERGER TH. Waveletbased system for recognition and a beling of polyhedral junctions. Optical Engineering [J]. 1998, 37 (1) : 158 - 165.
- [5] 周敏, 龙昭华. 基于 MAS 小波变换的数字图像轮廓提取算法 [J]. 重庆邮电学院学报, 2004, 16(2): 44 - 48.

测试数据库,每台 PC 为 5000 条数据。测试程序的编程语言为 VC++ 6.0,消息传递库为标准 MPI,测试结果如图 1 所示。

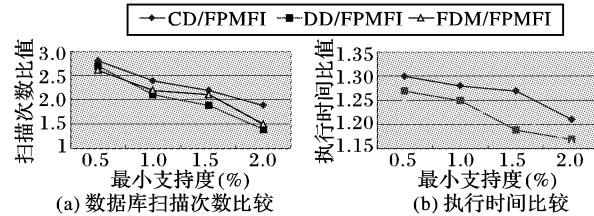


图 1 测试结果

由测试的结果可知,在相同支持度下,与 CD、DD 和 FDM 等并行挖掘算法相比,数据库扫描次数和执行时间都降低了,而且随着支持度的下降,FPMFI 算法性能优势更加明显。

4 结语

FPMFI 算法采用水平等间距投影方法进行数据分配,减少了数据偏移现象。算法让各计算机节点独立地计算局部频繁项集,再与中心节点交互实现数据交换及汇总,最终获得全局频繁项集,减少了同步次数、各节点间的通信量以及数据库的扫描次数。

参考文献:

- [1] 杨明, 孙志挥, 吉吉林. 快速挖掘全局频繁项目集[J]. 计算机研究与发展, 2003, 40(4): 620 - 626.
- [2] PARTHASARATHY S, ZAKI MJ, OGAWARA M. Parallel data mining for association rules on shared - memory systems[J]. Knowledge and Information Systems, 2001, 3(1): 1129.
- [3] ZAKI MJ. Parallel and distributed association mining: A survey[J]. IEEE Concurrency, Special Issue on Parallel Mechanisms for Data Mining, 1999, 7(4): 14 - 25.
- [4] HAN J, KAMBER M. Data Mining: Concepts and Techniques[M]. Beijing: High Education Press, 2001.
- [5] 李航, 刘宗田, 陈惠琼. 挖掘关联规则的并行算法[J]. 小型微型计算机系统, 2002, 23(10): 1231 - 1234.