

显微图象的数字图象处理及其在 血细胞形态分析中的应用

蒋勇^{1,2} 刘飞雷¹ 鲍红光³

¹(南京理工大学应用数学系信息与计算科学教研室, 南京 210094)

²(南京大学数学系, 南京 210093) ³(南京市第一医院, 南京 210006)

摘要 为了进行血细胞显微图象的快速分析, 将显微图象的数字图象处理技术引入到血细胞形态分析中. 在研究过程中, 首先通过高倍电子显微镜及数字图象采集设备对血细胞图象进行采集; 然后对血细胞的数字图象进行预处理, 其中包括对比度增强、图象的高通滤波、图象的二值化、图象的腐蚀和膨胀、边缘跟踪等等; 最后通过对血细胞的圆形度、矩形度和中心矩等等几何形态进行比较, 提取出能够较明显地辨析出它们的特征向量, 同时对分析结果的可信度进行了讨论.

关键词 显微图象 数字图象处理 血细胞 模式识别 形态分析 麻醉

中图法分类号: TP391.41 R446.113-39 **文献标识码:** A **文章编号:** 1006-8961(2001)11-1079-05

Microscope Digital Image Processing and Its Application in the Modality Analysis of Blood Cell

JIANG Yong^{1,2}, LIU Fei-lei¹, BAO Hong-guang³

^{1,2}(Dept. Math., Nanjing University of Science and Technology, Nanjing 210094)

²(Dept. Math., Nanjing University, Nanjing 210093)

³(The First Hospital of Nanjing, Nanjing 210006)

Abstract This paper introduced the application of the microscope digital image processing in the modality analysis of blood cell. We deal with the blood cell by the method of digital image processing, and have also thought over the processing step of concerned pattern recognition, so that, we can use the computer voluntarily to judging and distinguishes for the microscope of cell. The viewpoint of artificial intelligence has still been thought over in this paper, for that, the computer have the character of learning, this makes that its recognition result possesses reliability. After we pro-processed the digital image of blood cell, by comparing between the round degree, rectangle degree and center quadrature of the geometrical modality, we can get the eigenvectors of the different blood cell, we have discussed their reliability. We calculated the above geometry characteristic from the 56 examples cell samples (every sample includes three kinds of blood cells, and counts 168 images), and by standardization based the standard difference, we computed all values of characteristic weight, finally we got the characteristic vector of three kinds "classification", that are X_1, X_2 and X_3 . Using 2-norm $\|X - X_i\|$, we created the distinguish function. By train computer to remember the cell characteristic of new kind of cell, we have pledged to differentiate correctly the known classification, and have pledged to distinguishing correctly the new classification.

Keywords Microscope image, Digital image procestion, Blood cell, Pattern recognition, Modality analysis, Anesthesia

基金项目: 国家自然科学基金资助项目(1-99-71041)

收稿日期: 2000-08-24; 改回日期: 2000-12-22

0 前言

随着科学的发展,在对人类、动植物疾病的诊治和研究中,人们已不再只局限于对表面现象的处理,进而研究一些隐藏的深层次现象,这便导致了对人类、动植物细胞显微图象的研究,即通过对细胞状态的分析,来判断病因和研究相应的解决方案,以及观察疗效。

现有研究细胞的各种方法一般是把显微图象放大,尔后用人工进行辨识,这不仅耗费了许多宝贵的人力资源,也延误诊断时间,甚至使得临床中无法应用。本文利用数字图象处理的方法对其进行处理,合理设计了有关图象识别的处理步骤,实现了利用计算机来对细胞的显微图象进行自动判断和识别;本文还采用了人工智能的观点,使计算机有了“学习”的“品质”,这使得它的识别结果具有一定的可靠性。

1 实验条件及应用环境

1.1 硬件环境

本系统可在 IBM-PC 机及其兼容机上运行,并要求系统必须有分辨率至少为 640×480 pixel 的图形适配器,还要有至少 16M 的内存以及 540M 硬盘等硬件支持。

1.2 软件环境

本系统可在中文 Windows 9X 操作系统平台上编译及运行,源程序要求在 Visual C++ 5.0 以上的环境下编译。

2 软件系统结构及文档说明

2.1 数据的采集

本文所要用到的图象文件是利用图象采集卡辅助高倍电子显微镜实时采集到的数据(无条件的单位也可利用摄像仪将显微镜下的细胞拍摄成照片,再经过扫描仪录入计算机),它是以 .jpg 的文件格式或以灰度形式(Gray scale)存储的,raw 格式文件进行存放。

2.2 软件使用说明

系统使用时要注意以下两点:

(1) 待识别的图象文件和特征库文件 identity.

dat 要放在 C:\temp 子目录下,如果没有该子目录,系统则自动先行创建。

(2) 中间处理过程中形成的图象数据会以不同于被检测对象的文件名存放。

3 实验步骤

3.1 图象的获取

为了把工作重点放在图象的处理和识别上,本文数据采用 Photoshop 支持的,RAW 文件格式,由于该图象格式没有文件头,因此需要知道所使用文件的大小,本文所使用的图象大小是 450×450 pixel,以 1Byte 表示每个象素的灰度。

3.2 对比度增强

对读入计算机的图象,首先进行对比度增强处理,即对图象中的有用信息进行灰度变换,使图象更为清晰,以便于观察,本文使用线性灰度变换的方法来进行对比度增强处理,其变换公式为

$$g'(i, j) = 255 \times \frac{[g(i, j) - g_{\min.}]}{[g_{\max.} - g_{\min.}]}$$

3.3 图象的高通滤波

一般,图象的边缘轮廓信息都集中在频域中较高的频段,本文为了获取图象边缘轮廓等重要信息,因此选取了高通滤波方法。

3.4 图象的二值化

高通滤波后,就可对图象进行二值化处理,通过逐一实验,选用了阈值 220 来进行二值化,从系统测试结果来看,采用 220 作为阈值,处理的效果很好。

3.5 图象的腐蚀和膨胀

由于二值化后的图象上会仍存在一些零星的称之为孤立点和孤立块的小点和小块,且这种孤立点和小的孤立块的存在,对于以后进行的边缘跟踪和几何特征的计算会产生一些不良的影响,为了消除以上两种情况可能会带来的不良后果,因此要对二值化后的图象进行腐蚀和膨胀运算,本文是采取先腐蚀两次,再膨胀两次的方法,从系统测试的结果来看,也取得了比较理想的结果。

3.6 边缘跟踪与边界链码

为了取得图象的边界,本文又进行了边界跟踪处理,边界跟踪是采用链码技术,即先依某种规则找出一个边界点,尔后以该点为中心,再去找其他的边界点,因为此时链码规定的就是从当前边界点走到下一个边界点这一步骤必须采用的方向,本文所用

的链码结构为:首先以右上方为初始点,并定义为 1;然后链码值加 1,对应顺时针增加 45°,链码跟踪的方法是以逆时针进行,除第 1 次跟踪从正左方向开始外,其余各点都从以该点为中心的前一个点逆时针增加 45°(称为“前”面)的方向开始跟踪。

3.7 几何特征的选取

一般图象均是根据待识别对象的一些特征来进行识别的,因此几何特征选取是很重要的一步。在以前处理的基础上,即可决定选择哪些几何特征进行图象识别。这里考虑了如下几个比较重要的几何量度。

3.7.1 复杂度(圆形度)

面积一定的图形,一般周长越小,图形表面越光滑,越接近圆;反之,周长越大,则图形表面褶皱越多,形状越复杂,基于这个原因,人们采用复杂度来衡量图形的复杂程度,其最常用的复杂度可用如下公式表示

$$c = \frac{P^2}{4\pi A}$$

其中, P 为图形周长, A 为周长所围面积。

3.7.2 矩形度

另外一个反映物体形状的参数是它的矩形度。当图象给定后,即可求出物体的最小外接矩形(MER),而反映物体矩形度的一个参数是矩形拟合因子 R

$$R = A_0/A$$

其中, A_0 是物体的面积, A 是物体最小外接矩形的面积。

这里, R 反映了一个物体对其 MER 的充满程度。对于矩形, R 取得最大值 1.0;对于纤细的、弯曲的物体则其值变小,矩形拟合因子的值限定在 0 与 1 之间。

另一个矩形度参数是长宽比 K

$$K = W/L$$

其中, W 是物体外接矩形的宽, L 是物体外接矩形的长,这个特征可以把较纤细的物体与方形的或圆形的物体区分开来。

本文采用了圆形度指标,主要是因为计算 MER 比较繁琐,而且事实上只用圆形度就已经能够反映图象整体的几何特性了。

3.7.3 中心矩

为了描述物体内部形状细节,设 $f(x, y)$ 在物体内部为 1,而其余部分取 0, $f(x, y)$ 是描述物体形态分布是否均匀的函数。所谓重心矩(中心矩)就是指

以重心为原点进行计算的不变矩 M_{pq} 。

$$M_{pq} = \int \int (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

依此公式来看,重心矩具有位置无关性。本文采用的特征是如下二阶混合重心矩

$$M_{11} = \int \int (x - \bar{x})(y - \bar{y}) f(x, y) dx dy$$

把它离散化后,即为

$$\begin{aligned} \tilde{M}_{11} &= \sum_k (x - \bar{x})(y - \bar{y}) f(x, y) \\ &= \sum_{k=1}^M \sum_{l=1}^N (x - \bar{x})(y - \bar{y}) f(x, y) \end{aligned}$$

3.8 几何特征的计算

3.8.1 周长 P 的计算

设链码为偶的像素个数为 M ,为奇的像素个数为 O ,则

$$P = M + \sqrt{2}O$$

3.8.2 周长所围面积 A (包括图象白斑)的计算

$$A = S - [(M + O)/2 - 1]$$

式中, S 为所有像素个数。

3.8.3 重心和重心矩的计算

重心的行坐标和列坐标分别按下面公式进行计算

$$\bar{x} = \frac{\sum_R x f(x, y)}{\sum_R f(x, y)}, \bar{y} = \frac{\sum_R y f(x, y)}{\sum_R f(x, y)}$$

这样,通过前面的公式便可以求得二阶混合重心矩,为缩小该指标,本文采用标准化后的重心矩来代替它作为识别特征。

3.9 识别

本文对 36 例细胞样本(每个样本均包括 3 种血细胞,计 168 个图象)的上述几何特征进行了计算,即首先经过标准差标准化处理,然后计算出各个几何特征分量的均值,最后得到 3 种“分类”的特征向量 $X_1, X_2, X_3^{[5]}$,以此作为识别指标,并利用 2-范数 $\|X - X_i\|_2$ 来构造判别函数,以对细胞进行识别。同时通过“教授”计算机“记忆”新的种类的细胞特征,这样既保证了对于已知类别能进行正确判别,又保证了对新类别的正确辨识。

4 检测结果实例

下面就 3 种细胞图象的检测结果做一个简单分析。从待检测的 3 种细胞的图象(分别见图 1.2.3)来看,第 1 种细胞是正常细胞(其类型名是“good”)。

形状较为规则;第2种细胞是做了保护处理,并经过血液循环后的细胞(其类型名是“protected”),形状有些模糊,也不太规则;第3种细胞是经过血液循环后被破坏的细胞(其类型名是“bad”),形状紊乱,极不规则。

象,由图1~图7可以看出(处理最后的结果图象见图7),与原图象相比,虽形状和面积没有明显改变,但平滑了边界,为了形象地给出边界跟踪的效果,本文也给出了细胞1经边界提取处理后的边界图象(图8)。

图1至图7给出了细胞1经过每次处理后的图

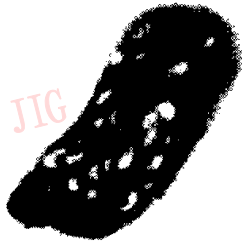


图1 细胞1原图象



图2 细胞2原图象



图3 细胞3原图象

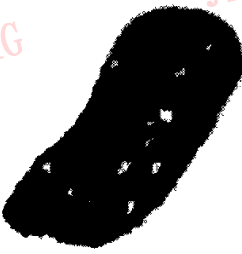


图4 经增强、高通滤波后的细胞1图象



图5 经二值化的细胞1图象



图6 腐蚀后的细胞1图象



图7 膨胀后的细胞1图象

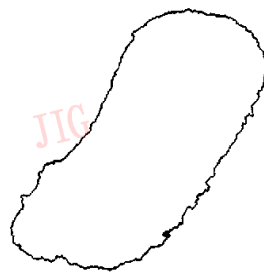


图8 提取的细胞1边界图象

将上述过程提取的特征参数,再通过对每种细胞图象的计算,其对应数据(此数据为未经过标准化等处理的数据)见表1。

由于检测已知样本较少(56例,168个样本),因而回顾性检验没有发生错判的现象,对15例待测样本的识别也未发生错判现象,但从理论上分析,如果

表1 用本文方法检测提取出的3种细胞特征参数

特征	细胞1	细胞2	细胞3
面积(μm^2)	94 261	37 447	109 232
周长(μm)	1 499.52	2 327.3	1 531.91
链码个数	2 306	7 067	15 791
重心距	0.539 473	0.017 413 8	0.113 771
复杂度	1.898 28	7.502 83	14.734 8

待检测样本数量增加,仍有发生错判的可能,其主要原因是显微镜下观测到的图象有时清晰,有时模糊,而图象采集是瞬时完成的,因而采集到的数字图象的清晰度可能会很差,这将导致在进行边界跟踪时,边界走形,复杂度增加^[5].有关这方面的工作,今后将继续进行研究.

参 考 文 献

- 1 鲍红光,杨海基,葛亚力. 围术期全程大剂量 6-氨基己酸对冠状动脉旁路术患者血液功能及手术后出血量的影响[J]. 临床麻醉学杂志,1999,11(5):54~55.
- 2 JIANG Yong, BAO Hongguang. Observation and analysis for clinical hemorheology on epidural block[A]. In: The Proceedings of Eighth International Congress of Bio-rheology[C], Beijing: Peking Univ. Press,1992.
- 3 国家自然科学基金委员会. 自然科学学科发展战略研究报告——生物医学工程[M]. 北京:科学出版社,1995.
- 4 BAO Hongguang, JIANG Yong, HUANG Jing. Image-sequence acquisition of liquid medicine diffusion in epidermal space[J]. J. Math. Medicine, 1998,11(1):70~72.
- 5 Castleman K R. Digital image processing [M]. New York: Prentice hall,1986.
- 6 鲍红光,蒋勇,刘飞雷. 血液麻醉的血细胞显微图象处理与形态分析[A]. 见:第八届国际生物学会会议论文集[C],北京,1992年8月.



蒋 勇 1959 年生,教授,中国工业与应用数学学会理事,中国数学学会 BM 学会副理事长,主要从事信息与计算科学和技术(包括生物信息技术等方面)的研究、教学与技术开发工作,连续主持多项国家自然科学基金等科研项目,已在国内外发表论文数十篇.



刘飞雷 1977 年生,工程师,1999 年毕业于南京理工大学计算数学与应用软件专业,现在华夏银行南京分行计算中心工作,主要从事信息与计算技术的应用开发.



鲍红光 1961 年生,主任医师,主要从事麻醉与复苏、生物医学工程(包括综合性血液保护等方面)的科研、临床与教学工作,主持、参加多项国家、省科技攻关项目,已在国内外发表论文数十篇,获得多项科技进步奖.

欢 迎 订 阅

地理信息系统与方法专辑

从传统的第三代 GIS 对图面信息的利用,向未来时空数据智能化共享的新一代 GIS 攀登,是当今数字信息时代必然发展的方向.我国“十五”规划已吹起向第四代 GIS 进军的号角.本刊为满足广大从事 GIS 应用和开发的科技人员对 GIS 领域研发现状及未来发展动态的渴求,特聘请著名 GIS 专家陈述彭院士、方裕教授等撰写了“第四代 GIS 软件研究”、“现代 IT 与第四代 GIS 软件”、“分布式地学虚拟环境研究”、“三维 GIS 的基本问题探讨”、“GIS 应用模型实现技术研究”、“万维网 GIS 发展及前景”、“可互操作之 GIS 研究”、“大型分布式 GIS 研究进展评述”、“GIS 空间分析能力探讨”等近 20 篇论文,集中刊登在《中国图象图形学报》2001 年第 9 期上,以飨读者,欢迎感兴趣的科技人员踊跃订阅.

为迎接 GIS 产业市场的兴旺与繁荣,《中国图象图形学报》竭诚为科技企业和研究机构认真做好广告宣传服务,同时欢迎从事相关行业研究和产品生产的企业、机构与个人踊跃订阅.

《中国图象图形学报》2001 年第 9 期“地理信息系统与方法专辑”定价:20 元

邮购地址:北京海淀区花园路 6 号《中国图象图形学报》编辑部 邮 编:100088

联系电话:(010)62378784,82611834 联系人:谷利英 王福美