

·成果简介·

计算机信号处理的细胞识别诊断系统

李根乾 余心宏 罗子健

(西北工业大学,西安 710072)

[关键词] 神经网络,细胞,特征库,分类器,专家系统

细胞的识别、分类和计数是医学研究和临床诊断的基本手段。采用计算机图形学原理完善、取代人工的识别、分类和计数早在 20 世纪 70 年代就有人提出^[1,2]。但是,由于当时计算机软硬件发展水平和识别算法的限制,上述方法未能实现,而基于红宝石毛细管的光电分类计数仪得到了广泛的临床应用。

血液在全身范围内不断流动,参与机体各种功能活动。除了造血系统本身的疾病可以直接引起血液发生相应病理变化外,人体全身或局部器质性、功能性疾病也会影响血液成分,因而血液检验可揭示其数量和质量上的变化。因此,血液细胞分类和计数是医学诊断的基本方法之一。它不仅是诊断血液疾病的主要依据,对其他疾病的诊断也有很大帮助。

目前,国际上血液细胞分类和计数所采用的方法主要有两种,即显微镜人工计数法和电子计数仪检测。前者设备简单,成本低。但是,劳动强度大,人为因素较多,检验结果与检验师的经验和个人情绪有关。后者虽然实现了自动化,但由于这种方法主要是根据细胞大小来判断,误差较大,而且只能进行常规分类,不能鉴别异常细胞。

随着电子技术和信息技术的发展,PC 机的速度、容量和软件性能有了很大提高,为采用计算机图形学原理对血液细胞分类和计数创造了条件。国内这方面的研究工作已初见成效。作者参考国内外研究成果,研制和开发了计算机信号处理的细胞诊断系统。该系统既可进行白细胞、红细胞、血小板的分类计数,也可对白细胞中的中性粒细胞、嗜酸性白细胞、嗜碱性白细胞、淋巴细胞、单核细胞进行分类计数。该系统具有自学习功能,通过不断积累样本,以吸收人类检验师的经验,使其决策逐步接近权威性。

该系统还可通过对病变异常细胞的命名和收集,及时发布异常细胞公告,以推进人类细胞学研究。通过判据库的建立和补充可不断扩充系统对各种细胞的识别功能。

1 系统的组成

系统由 Pentium II 计算机,高倍显微镜,彩色摄像机,彩色图像卡组成(图 1)。系统在 Windows 平台上运行,可进行真彩色细胞图像的采集、色彩分离、形态分析、分类识别。系统还有特征库的建立与编辑功能,以便不断扩大和完善系统功能。

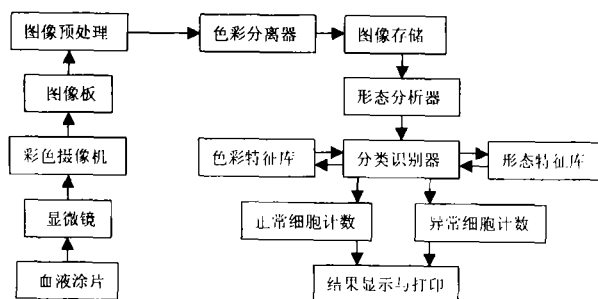


图 1 系统的组成

2 组成部分的功能和结构

2.1 图像预处理

图像预处理是对真彩色图像进行 R-G-B 三色分离和减色,以减少数据量,提高信噪比。图象处理的过程如下:采用金字塔压缩方法,将 R-G-B 分量的分辨率大幅度降低,变成 256 色,保留了识别的有用的信息,又减少了计算量,同时对噪声起到了抑制作用。由 R-G-B 分量值计算各像素饱和度,构成饱和度图像。对饱和度图像利用阈值分割技术对图像进行二值化。

本文于 2000 年 7 月 21 日收到。

2.2 细胞色彩、形态分离

在预处理的基础上,将各种对象(白细胞、红细胞、血小板)分离出来,形成可独立访问的实体。首先,对上述二值图进行形态运算,先进行N次(由本次要分离的对象大小决定)腐蚀,以消除噪声点干扰,再进行N次膨胀,使细胞边沿光滑连续,但对细胞的基本形态不产生影响。然后,对以上二值图像的目标区域进行标记。除白细胞核区域外这些标记区域中可能还有干扰,但其面积比较小且形状也不规则,可采用面积和圆弧度参数构成判别函数进行剔除。最后,对白细胞核标记区域确定质心位置,并根据白细胞的直径设置检出窗口,恢复原色彩分量图像,提取出视野中的多个白细胞。

2.3 白细胞核、浆分离

根据白细胞图像的特点和丰富的色彩信息,采用自动阈值分割技术完成白细胞内部的核、浆分离。由于背景区域的亮度分布均匀且远高于其他区域,从亮度直方图高端向下寻找,将第一个谷底作为阈值,扣除图像背景区域。对图像其余部分的饱和度进行统计结果表明,核的饱和度远大于其余区域,直方图呈双峰分布。因此,用最大方差作为阈值,可完成白细胞核区域的分离。图像其余部分的色调值计算结果表明,白细胞浆区域的颜色相似,色调分布集中,因而采用区域色调相似性增长准则进行白细胞浆区域的分割,并利用白细胞轮廓的凸性准则控制区域增长的迭代次数。

2.4 特征抽取

在医师的指导下,对白细胞、红细胞、血小板抽取色彩和形态特征。对每类白细胞的核、浆区域抽取形状、纹理、色彩和光密度等不同性质的特征。形状特征包括核叶数核浆面积比、形状因子、核/浆面积度等反映白细胞形态的参数。纹理特征反映白细胞的核、浆的粗糙程度,采用灰度游程长法和联合灰度统计法对核、浆进行纹理参数抽取。色彩特征包括核与浆区域的分量强度比、饱和度、色调、饱和度比、饱和度变化系数等。光密度特征包括:核/浆的平均光密度、方差、积分光密度、光密度变化系数、核/浆对比度、核浆积分光密度比等。

2.5 特征选择与分类器结构

系统采用树状判别机制实现血液细胞的三分类和白细胞的五分类。为提高系统的处理速度,首先采用将单个特征优劣排序法和序贯法相结合而形成的序贯增减特征筛选法对特征进行预选择^[3,4],以产生最有利于划分两-两类样本的优选特征子集。

最终选择的特征集根据要求的分类识别率、分类的速度和分类器的结构类决定。然后,在先验知识的指导下,进行树结构的设计,在决策树的各非终止节点所包含类别的次优特征中,集中精选出少数几个最优特征子集,用于分类判决。最后,非终止节点分别采用映射能力不同的线性判别、统计判别或BP神经网络,以发挥不同分类器的优点。各节点处的分类器经过训练后,便可用于白细胞的识别。

3 自学习、异常细胞处理和判据库的建立

系统建立后,首先要在医师指导下用人类已积累的样本进行训练。在使用过程中对异常细胞进行收集,及时发布异常细胞公告,由医师会诊确定其分类或重新命名,并将会诊结果加入样本库。当系统无检验任务或进入休息期间启动自学习功能,利用积累的新样本进行训练。这样不断积累、不断学习,逐步接近权威性。对新命名的细胞经过长期积累后可归纳其特征,为其编辑判据库,并经训练后可投入使用,从而达到对新细胞种类的识别。

通过对病变异常细胞的收集,及时发布异常细胞公告,可推进人类细胞学研究。通过判据库的建立和补充,可不断扩充系统对各种细胞的识别功能。

4 结论

所建立的系统利用了彩色细胞图像,使细胞的检出、分割过程有效而稳定。在特征提取方面,采用不同类型的参数,增强了互补性与容错性,使分类更为有利。在分类器机构中,采用了树状判别机制,并且在各非终止节点处针对不同情况采用不同复杂度和映射能力的分类器,使得系统总体正确识别率可达到94%。可以预见,通过对病变异常细胞的收集,及时发布异常细胞公告,将加快人类细胞学研究。通过判据库的建立和补充,可不断扩充系统对各种细胞的识别功能。随着分类器结构参数不断优化和特征集的优选,系统的性能将会不断提高。

参 考 文 献

- [1] Young I T. The classification of white blood cell. IEEE Trans on BMF, 1972, 19(1):291—298.
- [2] Levine M. Automated differentials; Geometric DATA'S HEMATRAK. Amer. J. UED. Tech., 1974, 40:462—468.
- [3] 扬晓敏. 血液白细胞特征抽取和分析. 北京生物医学工程, 1992, 11(4):123—129.
- [4] 罗立民. 血液白细胞计算机自动分类研究. 东南大学学报, 1993, 23(5):106—109.

IDENTIFICATION AND DIAGNOSIS SYSTEM FOR BLOOD CELL BASED ON COLOR CELL IMAGE AND COMPUTER INFORMATION PROCESSING

Li Genqian Yu Xinhong Luo Zijian

(Northwestern Polytechnical University, Xi'an 710072)

Key words NN, cell, characteristic lib, classify ware, expert system

·资料·信息·

1999 年 ISI 所收录中国科学期刊的引证数据 (1999 年度 JCR 发布)

| 期 刊 名 称 | 影响因子 | | 被引频次 | |
|--------------------------------------|--------|--------|--------|--------|
| | 1998 年 | 1999 年 | 1998 年 | 1999 年 |
| Acta Bioch Bioph Sinica(生物化学与生物物理学报) | 0.106 | 0.153 | 47 | 93 |
| Acta Chim Sinica(化学学报) | 0.178 | 0.317 | 566 | 730 |
| Acta Math Sci(数学学报) | 0.069 | 0.073 | 61 | 66 |
| Acta Mech Sinica(力学学报) | 0.506 | 0.346 | 164 | 157 |
| Acta Mech Solida Sinica(固体力学学报) | 0.041 | 0.200 | 32 | 65 |
| Acta Pharmacol Sinica(中国药理学报) | 0.224 | 0.196 | 439 | 488 |
| Acta Phys Sinica(物理学报) | 0.065 | 0.127 | 26 | 64 |
| Algebr Colloq(代数集刊) | 0 | 0.172 | 0 | 16 |
| Appl Math Mech(应用数学与力学) | 0 | 0.109 | 0 | 186 |
| Biomed Enviro Sci(生物医学与环境科学) | 0.430 | 0.839 | 212 | 318 |
| Chem J Chin U(高等学校化学学报) | 0.331 | 0.540 | 777 | 1203 |
| Chem Res Chin U(高等学校化学研究) | 0.208 | 0.275 | 83 | 124 |
| Chin Ann Math(中国数学年刊) | 0.173 | 0.193 | 92 | 95 |
| Chin Chem Lett(中国化学快报) | 0.195 | 0.224 | 357 | 500 |
| Chin J Chem(中国化学) | 0.229 | 0.346 | 152 | 241 |
| Chin J Chem Eng(中国化学工程快报) | 0.202 | 0.261 | 36 | 54 |
| Chin J Polym Sci(高分子学报) | 0.088 | 0.343 | 95 | 143 |
| Chin Med J(中华医学杂志) | 0.126 | 0.111 | 637 | 716 |
| Chin Phys Lett(中国物理快报) | 0.254 | 0.681 | 325 | 739 |
| Chin Sci Bull(科学通报) | 0.251 | 0.395 | 853 | 1198 |
| China Ocean Eng(中国海洋工程) | 0 | 0.089 | 0 | 31 |
| Comm Theor Phys(理论物理通讯) | 0.15 | 0.228 | 218 | 258 |
| High Energ Phy Nucl(高能物理核物理) | 0.818 | 0 | 167 | 0 |
| J Comput Math(计算数学) | 0.21 | 0.221 | 134 | 140 |
| J Iron Steel Res Int(高分子科学) | 0 | 0.209 | 0 | 10 |
| J Mater Sci Technol(材料科学技术学报) | 0.217 | 0.282 | 93 | 130 |
| J Rare Earth(中国稀土科学) | 0.068 | 0.127 | 37 | 45 |
| J Wuhan U Tec(武汉工业大学学报-材料科学版) | 0.135 | 0.034 | 16 | 13 |
| Prog Biochem Biophys(生物化学生物物理进展) | 0.041 | 0.070 | 32 | 63 |
| Prog Nat Sci(自然科学进展) | 0.195 | 0.269 | 116 | 170 |
| Rare Metal Mat Eng(稀土金属材料与工程) | 0 | 0.225 | 0 | 94 |
| Sci China-A(中国科学-A) | 0.273 | 0.297 | 506 | 516 |
| Sci China-B(中国科学-B) | 0.479 | 0.648 | 620 | 819 |
| Sci China-C(中国科学-C) | 0.21 | 0.361 | 36 | 105 |
| Sci China-D(中国科学-D) | 0.378 | 0.414 | 64 | 168 |
| Sci China-E(中国科学-E) | 0.116 | 0.333 | 25 | 76 |
| T Nonferr Metal Soc(中国有色金属学报) | 0.092 | 0.222 | 68 | 146 |