

基于嵌入式 Linux 的细胞特征提取算法设计

甘 岚, 龚礼武

(华东交通大学 信息工程学院, 江西 南昌 330013)

摘 要: 依据多连通域图像四邻域跟踪原理设计了嵌入式 Linux 下的细胞特征提取算法. 本算法解决了嵌入式下图像处理 API 封装不够完善以及硬件资源欠缺的问题. 实验结果表明, 该算法能够比较准确地提取细胞的周长、面积等特征参数. 该算法移植至移动嵌入式设备中解决了传统 PC 机携带不便的难题.

关键词: 嵌入式; 图像; 跟踪; MiniGUI

中图分类号: TP368.1

文献标识码: A

文章编号: 1000-7180(2012)08-0082-04

Cells Feature Extraction Algorithm Design Based on Embedded Linux

GAN Lan, GONG Li-wu

(School of Information Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: This paper designed a cell under the embedded Linux feature extraction algorithm based on multiply connected region image around domain tracking. The algorithm have resolved the the problems which embedded image processing API package is not perfect and the lack of hardware resources. Experimental results show that the algorithm can accurately extract the cell perimeter, area, and other characteristic parameters. The algorithm is portable to mobile embedded devices to carry a PC to solve the traditional problem of inconvenience.

Key words: embedded; image; track; MiniGU

1 引言

至今为止, 图像处理已经发展成为相对比较熟悉的研究领域, 并且在此领域也取得了巨大的研究成果, 但是, 绝大部分研究平台都是依赖于传统 PC 机或者研究成果应用于 PC 机, PC 机体积相对庞大, 携带不便, 这就凸显了其应用的局限性^[1]. 同时, 对于某些专业领域的应用还会造成大量资源的浪费.

为此本文提出了基于嵌入式 Linux 来实现细胞特征提取的方法. 本文算法首先对原始图像进行预处理来减小杂质、噪声等对细胞特征的影响; 其次, 对预处理过的细胞图像进行区域标号和邻域跟踪; 最后, 通过自己设计封装嵌入式 Linux 图像处理接口完成算法的封装实现. 实验结果表明, 本文提出的方法应用在嵌入式 Linux 系统下能够准确地提取出

细胞的周长及面积特征.

2 细胞特征提取原理

2.1 细胞图像预处理

从显微镜下采集细胞图片中, 会出现一些杂质、噪声以及其它各种细微组织的干扰. 原始细胞图像如图 1(a)所示, 为一幅胃粘膜上皮显微细胞图像. 为了在后文算法中精确地提取出细胞特征, 下面介绍图像几种预处理^[2]:

(1) 对图像进行平滑和滤波处理

对在显微镜下采集来的原始肿瘤细胞图像进行平滑和滤波操作, 可以改变图像周围像素灰度值相差较大的像素点的值, 从而抑制图像中的噪声, 处理效果如图 1(b)所示, 可知噪声点得到了很好的抑制.

(2) 对图像灰度化操作

收稿日期: 2011-12-10; **修回日期:** 2012-02-08

基金项目: 国家自然科学基金项目(61163040); 华东交通大学研究生创新基金项目(YC2011-X015)

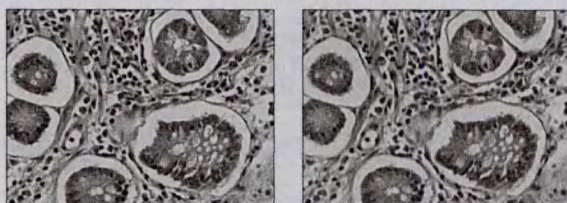
在表示颜色的方法中,除了RGB外,还有一种叫YUV的表示方法,在这种颜色表示方法中,Y分量的物理意义就是亮度,U和V分量代表了色差信号.Y分量包含了灰度图的所有信息,只用Y分量就能完全表示出一幅灰度图.YUV和RGB之间的对应关系如下:

$$[YUV] = [RGB] \begin{bmatrix} 0.299 & -0.148 & 0.615 \\ 0.587 & -0.289 & -0.515 \\ 0.114 & 0.437 & -0.100 \end{bmatrix} \quad (1)$$

利用上式,根据R、G、B的值求出Y值后,将R、G、B值都赋值成Y,就能表示出灰度图来。

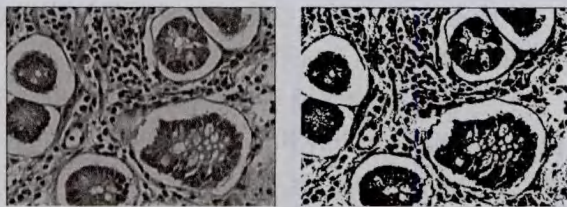
(3) 图像分割,对图像进行二值化处理

对经过上述两步预处理得到的灰度图像,首先遍历图像各像素点,按256个灰度级进行计算.本文的分割技术采用简单阈值分割,阈值设定为128,比它大就是白(用1表示),比它小就是黑(用0表示),经过阈值处理化处理后的图像变成了黑白二值图,如图1(d)所示.分割得到了明显的细胞目标区域和背景区域,为下文的细胞特征提取创造了条件。



(a) 原始图像

(b) 平滑、滤波后图像



(c) 灰度化图像

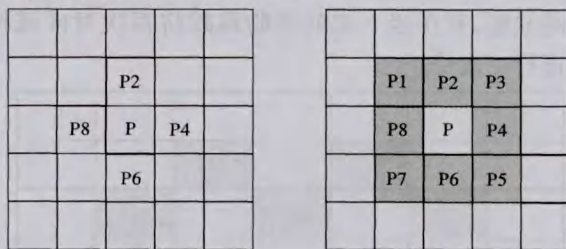
(d) 二值化图像

图1 原始图像及处理效果

2.2 特征提取原理

由于细胞图像的特征是由多个连通区域构成的,特征提取算法将对上文预处理后的细胞图像进行轮廓跟踪,最终得出细胞的周长与面积特征.对多连通域二值化(0表示背景像素,1表示目标像素)细胞图像的跟踪首先得对各个目标区域用不同的数字表示(同一个区域的所有像素点都用同一个数字)——细胞区域标号^[3].通常使用的区域标记算法有基于像素的标记算法和基于游程递归的标记算

法.完成区域标号后,紧接着就是对各个区域细胞进行轮廓跟踪^[4].后文的算法设计就是采用邻域跟踪的原理,邻域分为4邻域和8邻域.假设图像的当前目标点是 $f(x, y)$,则此点在图像中4个邻域像素点分别为 $f(x+1, y)$, $f(x, y+1)$, $f(x-1, y+1)$, $f(x-1, y)$;如果是8邻域跟踪,以 $p(x, y)$ 为当前目标像素点,那么其8邻域像素点分别是 $p(x, y-1)$, $p(x-1, y-1)$, $p(x-1, y)$, $p(x-1, y+1)$, $p(x, y+1)$, $p(x+1, y+1)$, $p(x+1, y)$, $p(x+1, y-1)$.4邻域像素标记如图2(a)所示,8邻域像素标记如图2(b)所示。



(a) 四邻域表示

(b) 八邻域表示

图2 四邻域和八邻域表示

3 算法设计

3.1 区域标号算法设计

经过细胞图像的阈值分割后,将细胞图分割为细胞目标区和背景区.从图像首行进行扫描,通过对图像灰度值进行比较,对每一个像素进行标号,将相同标号的像素归为同一个闭合区域。

综合考虑算法的精度与效率,笔者结合使用序贯算法和标签处理算法将细胞进行标号分区.算法描述如下:

(1) 如果当前点的左边点的像素值与上方点的像素值相等且都为目标点,并且左边点和上方点的标记号都等价,则当前点的标记号同左边点或者上方点。

(2) 如果当前点的左边点为目标点,而上方点为背景像素点且当前点为目标点,那么,当前点的标记号同左边点。

(3) 同样如果当前点的上方点是目标点,而左边点为背景点且当前点为目标点,那么,当前点的标记号同上方点。

(4) 如果当前点的上方点和左边点都为背景点且当前点为目标点,则当前点为新标号点并赋予新的标号。

(5) 如果当前点的左边点和上方点的像素值相

等且为目标点,但是左边点与上方点的标记号不相同,则选择左边点和上方点中标记号较小的值赋给当前点的标记号,并且修改左边点和上方点的标记号,将较大的标号值修正成与较小的标号值相同。

通过以上分析可知,标号过程中只要扫描当前点的左边和右边点,且不需要设计额外的数据结构来存储标号等价表,所以从算法效率上来考虑的话,该算法还是比较优越的。但是,对于比较特殊的不规则图像的标号却存在很大的偏差,如图 3 所示的锯齿型区域就会导致多标,本来五个目标像素点属于同一个连通区域,却标上了五个不同的标号。基于上述的缺陷,我在接下来的细胞跟踪将再次对区域标号进行一次修正。

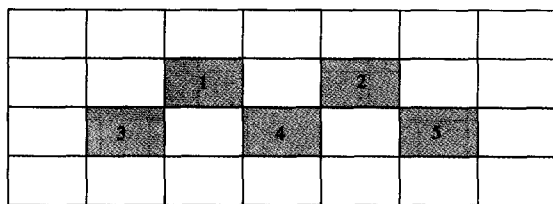


图 3 锯齿边界图

3.2 细胞特征提取算法设计

本算法从上往下、从左往右对图像进行扫描,采用 4 邻域连通,跟踪过程如图 4 所示。在 4 邻域跟踪后完整地跟踪出了细胞图像的边界,随着图像的跟踪结束,统计得到了每个细胞区域的边界像素点数及区域面积像素点,最后,综合所有的特征值得出细胞的平均周长和平均面积。综上所述,本算法实现的流程图如图 5 所示和具体步骤如下:

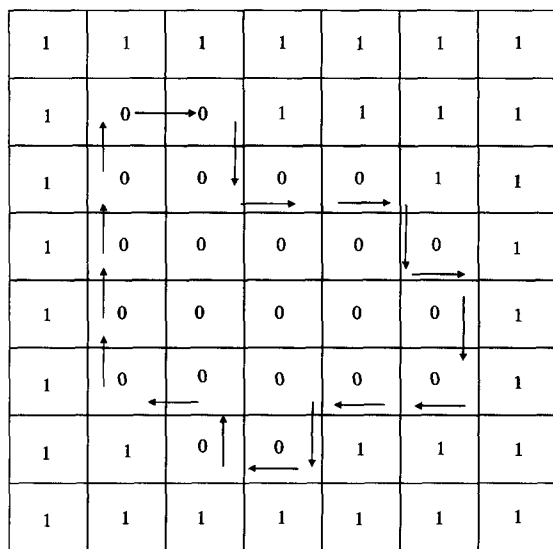


图 4 四邻域跟踪示例

(1) 从上往下,从左至右扫描图像,当 $f(x, y)$

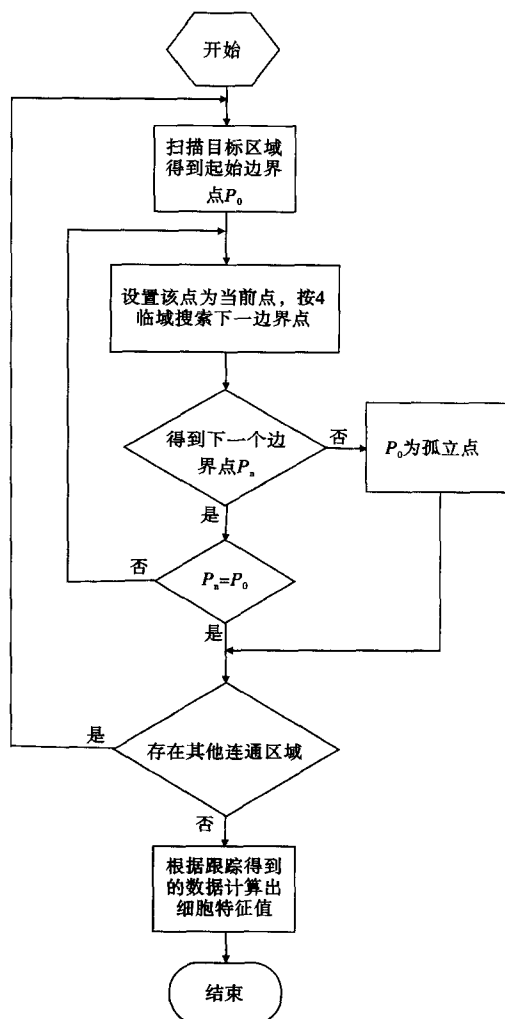


图 5 算法流程图

>0 时,确定该点一定是当前连通域的起始边界点 P_0 ,记录坐标 StartX, StartY.

(2) 设置此点为当前点,顺时针方向搜索 4 邻域像素点,当第一次出现其像素点的值为 0 时,这一像素点就是要求的新边界点 $P_n (n > 0)$,更新 P_n 点为当前点;如果没有发现像素值为 0 像素表示 P_0 为孤立点,则跳至(5).

(3) 若 $P_n = P_0$,则表示跟踪回到了起点,该连通区域跟踪完毕,统计此连通域的所有边界点总数和区域面积,然后跳转至步骤(5).

(4) 如果新 $P_n \neq P_0$,对当前点 P_n 转到步骤(2).

(5) 如果该图像还存在未跟踪的连通区域,选择下一个连通区域转至步骤(1),否则,算法结束,输出跟踪后的图像并且得出所有连通域的边界点总和与所有细胞面积总数。

(6) 根据上文得出的边界、面积等参数计算输

出细胞的平均周长和平均面积,实现细胞特征值的提取。

算法执行过程中,为了正确地选择下一个边界点,需要即时更新当前点与前一点坐标。同时,上文提到在跟踪过程中涉及到一次标号修正,实现方法如下:对比当前边界点和前一边界点的区域标号,如果不相等则需要记录这两个标号均标识相同区域,取两者中的最小者替代原标号。

通过上文的跟踪算法得到了真是的连通区域的个数以及整个图像边界点的个数,为了满足医疗应用的需求,算法通过设置两个参数 Min-Round(细胞周长的最小值)和 Max-Round(细胞周长最大值)来提取出正常细胞的特征,从而检测出图像中的肿瘤细胞。当单连通域边界周长值位于 Min-Round 至 Max-Round 这个范围中时,统计出符合要求的连通域细胞的个数。

算法中为了将每个区域的边界保存到相应的数据结构中并提高检索效率,笔者设计了二维链表结构,在此就不做详细介绍。

3.3 嵌入式 Linux 图像接口设计

由于上述算法最终应用于嵌入式 Linux 操作系统,鉴于嵌入式环境的特性(可使用的 API 较少):尤其是对于 Windows 下标准 BMP 图像没有相关数据结构定义以及对图像处理的支持,所以笔者在图像格式数据的转换环境做大量的工作,对于图像要求操作的结构、变量以及大部分算法都要由自己定义^[6]。算法中重新定义了三个主要的图像结构体:位图文件头 BITMAPFILEHEADER、位图信息头 BITMAPINFOHEADER 和调色板 Palette。

在 Windows 体系结构下,设计了 HLOCAL 和 HGLOBAL 等内存句柄,它们分别指向一个局部的和全局的内存块,然而,在 Linux 下并没有发现类似的内存句柄的定义,这给图像数据在内存中的操作造成了很大的不便,最后,笔者通过定义普通类型指针,然后动态分配内存的方式解决了上述问题。同理,在 Linux 下没有类似 Windows 下 VC 提供 Globallock()这样的 API 函数,因此使用 malloc()函数代替。而全局内存管理函数 Globallock()比其他内存管理函数慢,而且,考虑到兼容性,该函数仍然用于 Windows 下 DDE 和剪切板函数,并且不能提供更多的特征。这样,以 MiniGUI 为平台图像处理系统在速度上比 Windows 下要快^[7]。下面就是笔者设计的图像数据能够在内存操作的模块:

```
unsigned char * m_pImage; //图像数据指针
```

```
m_pImage=(unsignedchar *)malloc
(bmpInfoHead.biHeight * lLineBytes);
memcpy(m_pImage, imgData, bmpInfoHead.biHeight
* lLineBytes); //拷贝//图像数据至 m_pImage 中
```

m_pImage 为指向图像数据区的指针,通过 malloc()函数给此指针动态分配了 bmpInfoHead.biHeight * lLineBytes 字节大小的内存块, bmpInfoHead.biHeight 代表图像高度, lLineBytes 代表图像行字节数。最后,通过 memcpy()函数实现了图像数据在内存中拷贝,这样,就可以依赖指针方便地完成图像数据的处理。

考虑到工作效率和算法的可移植性,常用的接口函数和结构体都以 API 的形式封装起来,这样,在后续的开发中只要调用封装好的 API 即可。

4 实验结果及分析

在开发板上配置好 MiniGUI 环境,将上述算法移植至开发板,至此,该算法就能正常地运行在嵌入式 Linux 平台下了。算法在开发板上运行没到了满意的实验结果:为医生诊断细胞的肿瘤病变提供了真实可靠的数据^[8]。

5 结束语

嵌入式 Linux 下细胞特征提取算法初步完成之后主要对肿病理组织切片进行诊断和分析,它能够很好地完成细胞图像边界跟踪、细胞特征参数提取。此外,本算法是基于嵌入式技术开发,因此,它拥有实时性强、可靠性高、体积小、功耗低、成本低、较好的便携性等众多优点,方便了医生携带至边远落后地区使用,对提高人民的医疗水平有着很大帮助。

参考文献:

- [1] 孟建明,薛重德,李广涛.基于 ARM9 的嵌入式指纹图像处理算法研究[J].计算机应用与软件,2010,27(1):255.
- [2] 章毓晋.图像分割[M].北京:科学出版社,2001:101-102.
- [3] 刘相滨,向坚持,阳波.基于八邻域边界跟踪的标号算法[J].计算机工程与应用,2001(23):125-126.
- [4] 徐正光,鲍东来,张利欣.基于递归的二值图像连通域像素标记算法[J].计算机工程,2006,32(24):186-188.
- [5] 沈桥楠,安雪晖.基于游程递归的连通区域标;证算法[J].计算机应用,2012,30(6):1617-1618

(下转第 90 页)

nitor 传输过来的输入采样信息给 AES C 模型的驱动模块,该驱动模块会将输入采样信息解析成 AES C 模型能识别的数据类型,然后按照 AES C 模型的输入要求对其进行驱动,AES C 模型的监控模块会采样 AES C 模型的输出信息并将其打包传输给比较器,比较器接受 AES 的输出信息和 AES C 模型的输出信息,按照一个数据单元的信息长度对相关的命令与数据作比较,从而判断此次验证场景的正确性,若正确则打印 pass,否则打印 failed.

AES 模块的解密步骤是机密的一个逆过程,只需按照 AES 解密的协议标准和时序驱动、采样和比对即可.

4.3 验证结果分析

以 aes_multi_channel_enc_tc 测试用例为例分析仿真结果.此测试用例的验证场景是 AES 模块多个通道同时进行数据加密,密钥长度(128 比特或是 256 比特)、工作模式(xts 模式或 ecb 模式)、数据单元长度、待加密数据、密钥均为随机约束得到,AES_driver 使用这些数据项目驱动 AES DUT,AES_monitor 采样 AES DUT 的输入输出信息,并将其转化打包.ASE_monitor 采样得到的一个通道的数据项目可以看到该通道工作的信息,该通道加密密钥的长度为 256 比特,工作 ecb 模式,数据单元的大小为 4 kB,传输延时为 0,cts 功能禁止,还可以看到每笔待加密数据和其对应的加密后数据.

可以看到,此测试用例 aes_mukti_channel_enc_tc 比对成功.在此次仿真过程中一共打印了 1 040 条仿真信息,没有警告、一般错误或严重错误信息,在 rsnst, test_done, aes_driver, aes_monitor, aes_scoreboard 模块中分别打印的信息数量也在报告中给出了.若测试用例失败,这些打印信息便是至关重要的调试参考信息.

5 结束语

在验证领域,验证语言的整齐划一、仿真平台的统一、发展正规高效的验证方法学,一直是 IC 验证的发展趋势^[6].UVM 验证方法学吸收了当前诸多主流验证方法学的优点,并且获得了如 Cadence、Synopsys 和 Mentor 等 EDA 公司的仿真软件支持,它的高效、高可重用化、高度自动化、开源化已经得到了很多工程师们的肯定.

参考文献:

- [1] 李磊,罗胜钦.基于 VMM 方法的 SOC 集成验证[J].电子测量技术,2011(1):1
- [2] Announcing the Advanced Encrypyion Stadndard[S].美国空军用标准:FZPS-PVB-197 Federal Information Processing Standards Publication. 2001.
- [3] Accdlera. Universal verification methodology 1.1 user's guide[M]. Cadence Design Systems Inc., Mentor Graphics Corp., Synopsys Inc., 2011,5: 2
- [4] Accdlera. Universal verification methodology 1.1 class reference[M]. Cadence Design Systems Inc., Mentor Graphics Corp., Synopsys Inc., 2011,6: 342
- [5] Sharon Rosenberg. A practical guide to adopting the universal verification methodology [M]. Cadence Design Systems Inc., 2010,5: 9
- [6] Chris Spear. System Verilog for Verification[M]. Synopsys Inc., 2008: 5

作者简介:

王小力 男,(1956-),教授,博士生导师.研究方向为高性能 VLSI 与优化设计、集成化芯片系统设计与基础研究、光纤通讯与光波分复用(WDM)技术.

田 劲 男,(1985-),硕士研究生.研究方向为 SSD 固态硬盘主控芯片的设计与验证.

(上接第 85 页)

- [6] 杜华. Linux 编程技术详解 [M]. 北京:人民邮电出版社. 2007:280-281.
- [7] 周立功. ARM 嵌入式 MiniGUI 初步与应用开发范例 [M]. 北京:北京航空航天大学,2006:25-26.
- [8] 李广丽,刘遵雄,张宏斌. 肿瘤图像病理诊断系统的设计[J]. 华东交通大学学报,2007,24(5):99.

作者简介:

甘 岚 女(1964-),硕士,教授.研究方向为模式识别与图像处理.

龚礼武 男,(1985-),硕士研究生.研究方向为嵌入式系统与应用.