

基于图像处理的血液透析器纤维识别计数^①

曹婧蕾, 李军华

(南昌航空大学 信息工程学院, 南昌 330063)

摘要: 为获得血液透析膜纤维的高精度识别与计数, 通过对纤维图像颜色聚类特征的实验和分析, 得出在 YUV 空间中仅利用颜色分量 U 和 V 对图像进行颜色聚类. 利用 K-means 聚类算法对图像进行颜色聚类, 根据图像中各像素的聚类属性对图像进行分割. 在纤维的识别计数环节, 采用 Hough 变换的方式对分割出的纤维图像识别并计数. 实验结果表明, 该算法识别精度高, 鲁棒性好.

关键词: 颜色分析; 特征提取; 图像分割; Hough 变换; 计数

Recognition and Counting of Blood Dialyzer Fibers Based on Image Processing

CAO Jing-Lei, LI Jun-Hua

(School of Information and Engineering, Nanchang Hangkong University, Nanchang 330063, China)

Abstract: In order to obtain high-precision fiber identification and counting of fibers by color image analysis and clustering features of the experiment, obtained in the YUV space using only U, V components of the image color clustering. Using k-means clustering algorithm for color image clustering, each pixel according to the image clustering properties of the image segmentation. Identification of the fiber count links, using the segmented Hough transform the way out of the fiber image recognition and counting. Experiments show that the algorithm to identify high precision, can adapt to a range of optical changes in the environment, better robustness.

Key words: color analysis; feature extraction; image segmentation; hough transform; counting

1 引言

采用图像处理的方式, 对同类物体进行识别计数已有广泛的研究, 识别计数对象包括细胞、棒材、瓜果等. 该方法相对传统的方法具有人力少、处理速度快等优势. 本研究主要围绕血液透析器生产过程中透析膜纤维的计数展开研究, 这在透析器生产厂家中有实际的应用意义. 一次性空心纤维血液透析器血液透析是利用半透膜的原理, 将患者的血液与透析液同时引进透析器, 两者在透析膜的两侧呈反方向流动, 借助膜两侧的溶质梯度、渗透梯度和水压梯度, 以达到清除毒素和体内滞留过多的水分, 同时补充体内所需的物质, 并维持电解质和酸碱平衡的目的. 透析器主要由支撑结构和透析膜组成, 历史上先后出现过的透析器: 平板型、蟠管型和空心纤维型, 其中空心纤维型透析器是目前临床使用最多、效果最好的一类透析

器. 透析器最重要的部件是透析膜, 透析膜是由中空纤维组成的. 透析膜的主要评价标准是清除率和超滤系数, 透析膜的密度即单位面积纤维个数是透析器生产的考核指标.

以往是采用人工方式对透析器透析膜的纤维计数, 而人工计数劳动强度大, 耗时长, 还可能存在对纤维损坏等问题, 因此传统的计数方式难以满足实际需要, 这就需要找到一条快捷、无损的方式对透析器纤维数量进行统计. 本文引入图像处理的方式试图解决上述问题, 通过对图像中纤维颜色、形状等分析, 提取纤维的关键特征, 进行识别计数.

获取图像时, 外界光照的变化会直接影响图像明暗度及对比度, 这些都会最终影响图像的质量. 而引起光照变化的因素很多, 如拍摄时快门的开启与闪光灯到达最高亮度时不匹配, 提供的恒定光源明暗出现

^① 收稿时间:2012-03-05;收到修改稿时间:2012-04-16

变化等, 因外界环境一定范围变化的扰动而对识别效果造成较大的影响是不适于实际应用的. 因此, 针对上述问题, 本文将研究重点放在能否设计一种方法, 能够适应正常状况下光照在一定范围内的波动, 即该方法具备较好的稳定性.

由于彩色图像中含有丰富的目标信息, 因此, 本文在彩色图像中利用聚类方法直接进行识别, 避免二值化的信息过多损耗, 而对识别难度加大. 先将彩色纤维图像从 RGB 颜色空间转化到 YUV 颜色空间, 在对颜色进行聚类时, 将因不均匀而影响聚类效果的亮度分量 Y 去掉, 直接用色差信息 U 和 V 进行聚类^[1-2]. 利用获得的纤维区域的聚类, 对彩色纤维图像进行分割, 采用圆的 Hough 变换检测纤维圆, 并对检测的圆计数.

1 基于聚类分析的纤维颜色特征提取

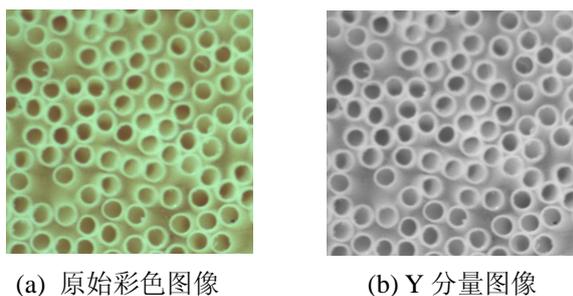
1.1 转换及直方图生成

本研究的纤维图像是 CCD 数码相机拍摄得到的, 图像格式采用 24 位的 BMP 位图. 彩色纤维图像在拍摄时由于光源的影响, 图像亮度分量是不均匀的. RGB 颜色空间是非均匀空间, 只要亮度改变, 3 个分量都会相应改变, 不具备视觉一致性. 因此 RGB 空间中由于亮度分量的影响, 不适宜于图像的分割和识别. YUV 颜色空间具有视觉一致性, 且亮度分量 Y 和色差分量 U 和 V 是分离的, 在亮度不均匀的图像中, 色差对区分不同颜色的目标非常有效. 因此, 本文直接在 YUV 颜色空间中构造纤维的聚类颜色模型, 在 YUV 颜色空间进行聚类分割和识别. 从 RGB 到 YUV 颜色空间转换公式:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

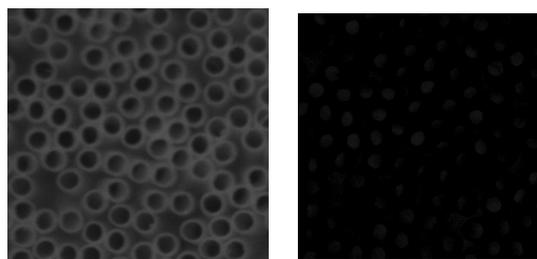
1.1.1 纤维图像颜色聚类特征

图 1 为 YUV 颜色空间中各分量图像.



(a) 原始彩色图像

(b) Y 分量图像



(c) U 分量图像

(d) V 分量图像

图 1 YUV 空间中各分量图像

从这些图像的比较中可以得出以下结论:

① Y 分量图像基本上保留了纤维区域的轮廓细节, U、V 分量图像主观感觉上十分模糊.

② 在观察合成的彩色纤维图像时没有觉得图像颜色的模糊.

③ 人眼对亮度信息敏感, 主要通过亮度差别来分辨物体细节, 而对彩色信息不敏感.

1.1.2 纤维图像颜色特征的实验结果分析

图 2 为 YUV 空间中三分量直方图分布, 从图中可以看出 Y 分量分布范围大, U、V 分布比较集中, 在 YUV 空间中 Y 分量和 U、V 分量是分离的.

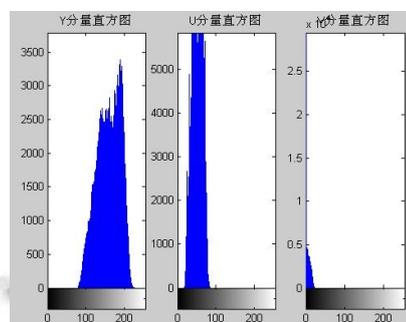


图 2 YUV 空间中各分量直方图

1.2 彩色纤维图像颜色算法研究

由于在 YUV 空间中 Y 分量和 U、V 分量是分离的, 只利用 U、V 分量进行颜色聚类. K-means 聚类算法最终的聚类依赖于初始聚类中心及聚类数目 K 的选择. 首先利用 K-means 算法对彩色图像进行初步颜色聚类, 然后用改进的 K-means 聚类算法对初步聚类的结果进行优化聚类, 最后对聚类结果结合空间因素进行处理.

1.3 彩色纤维图像颜色特征的聚类提取

1.3.1 K-means 聚类基本算法

K 均值算法^[3,4](K-means Algorithm)是基于准则函数最优的聚类算法, 它能够使各类样本到聚类中心的

聚类平方和取得极小值。

已知样本集合 $X=\{x_1, x_2, \dots, x_n\}$, x_j 是 d 维特征向量, $j=1, 2, \dots, n$; 已知类别数 K 和初始聚类中心 C_i ; 相似性测度可以采用欧氏距离; 聚类准则采用误差平方和准则,

其准则函数为 $J = \sum_{i=1}^k \sum_{x \in C_i} \|x - C_i\|^2$, C_i 是第 i 类的聚类中心。

K 均值算法就是通过不断调整聚类中心, 使得误差平方和准则函数 J 取得极小值。 K 均值算法流程如图 3 所示:

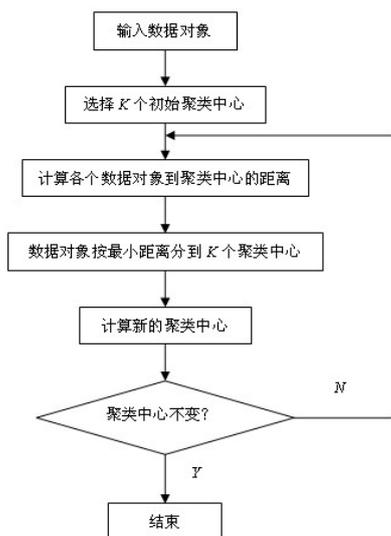


图 3 K 均值算法流程

1.3.2 YUV 颜色空间的 K-means 聚类

颜色特征向量 (U, V) , 其中 (\bar{U}, \bar{V}) 表示样本中心, m 为图像像素数, (U_i, V_i) 表示第 i 个像素点的颜色特征向量。

$$\begin{cases} \bar{U} = \frac{1}{m} \sum_{i=1}^m U_i \\ \bar{V} = \frac{1}{m} \sum_{i=1}^m V_i \end{cases} \quad (2)$$

1) 初始聚类中心设定方法

① 根据颜色向量 (U, V) 数据集的特征, 选取距离 (采用欧式距离) 所有样本中心 (\bar{U}, \bar{V}) 最近的一个样本作为第一个初始聚类中心 (\bar{U}_1, \bar{V}_1) 。

② 当聚类数为 2 时, 从待分类的样本中选出距离第一个初始聚类中心 (\bar{U}_1, \bar{V}_1) 最远的样本作为第二个初始聚类中心 (\bar{U}_2, \bar{V}_2) 。

③ 当聚类数为 3 时, 计算未被作为聚类中心的各样本与 (\bar{U}_1, \bar{V}_1) 、 (\bar{U}_2, \bar{V}_2) 之间的距离, 并求出它们之中的最小值 d_i , 根据 $D_t = \max\{d_i\}$, 选择第 t 个样本作为第三

个初始聚类中心 (\bar{U}_3, \bar{V}_3) 。

④ 当聚类数为 K 时, 根据已知的 $K-1$ 个初始聚类中心计算未被作为聚类中心的各样本到各聚类中心的距离 d_{ij} , 并计算出 $D_r = \max\{\min\{d_{i1}, d_{i2}, \dots, d_{i(K-1)}\}\}$, 选择第 r 个样本作为第 K 个初始聚类中心 (\bar{U}_k, \bar{V}_k) 。

2) 聚类有效性指标的选择:

设 $a(i)$ 为样本 i 与类内所有其他样本的平均距离, $b(i)$ 为样本 i 到其他每个类中样本平均距离的最小值。Silhouette 指标定义为:

$$sil(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (3)$$

Silhouette 指标的在 $[-1, 1]$ 范围内变动, 所有样本的平均 Silhouette 指标值越大表示聚类质量越好, 其最大值对应的类数为最佳聚类数。

2 基于聚类分析的彩色纤维图像的分割

2.1 基于聚类的纤维图像的粗分割

根据图像中各像素的聚类属性对图像进行分割。对图像中任一像素点 i , 若其聚类属性归类为背景类, 将该像素点置为黑色, 若其聚类属性归类为纤维类, 则该像素点的颜色保持不变。

2.2 基于聚类的纤维图像的细分割

在聚类粗分割基础上, 采用区域合并方式将过细的区域合并, 即每次都是合并图像中距离最相近的两个区域, 形成一个新的区域, 合并时的聚类必须满足以下两个条件:

- 1) 两个区域在颜色上距离相近。
- 2) 两个区域在空间距离上相近。

2.3 分割后图像噪声处理

彩色纤维图像经过聚类分割后, 纤维的横截面被分割成一个个独立的连通区域, 但是, 在区域合并过程中, 少量小的区域因为不符合合并的条件, 并没有合并到邻近的区域中去, 无规则地分布在细分割的图像中, 将这些小的区域视为噪声, 这些噪声会对后面的纤维识别计数造成干扰, 因此必须将这些噪声区域去除, 为识别计数创造条件。

3 基于Hough变换的纤维识别与计数

3.1 Hough 变换圆检测

圆的一般方程式: $(x - a)^2 + (y - b)^2 = r^2$, 圆心的计

算公式:
$$\begin{cases} a = x_i - r \cos\theta \\ b = y_i - r \sin\theta \end{cases}$$
 其中, $\theta = \arctg \frac{g_x}{g_y}$, g_x, g_y 分

别是两个方向的梯度. 实际计算中, 可以直接计算出:

$$\begin{cases} \cos\theta = \frac{g_x}{g} \\ \sin\theta = \frac{g_y}{g} \end{cases}, \text{ 其中 } g = \sqrt{g_x^2 + g_y^2}, \text{ 这种方法已成为圆检}$$

测的标准 Hough 变换^[5,6].

在上述圆的方程中, 圆有三个参数, 分别是圆心坐标与半径. 以图像上每一个像素为圆心, 以已知半径在参数平面上画圆, 将结果进行累加, 得出参数平面上的峰值点, 也即是对应原图像中的圆心. 图像的每一点对应到参数平面上的一个圆.

模糊理论中的 Hough 变换, 圆的方程可以表示为:

$$(x - a)^2 + (y - b)^2 = \rho^2 \quad (4)$$

Hough 变换后的参数空间为三维 (ρ, a, b) . 为了把随机 Hough 变换扩展到适用于模糊点, 对每一个参数 ρ , 隶属度都沿着曲线相加, 像随机 Hough 变换一样, 将一个点对映射到参数空间, 然后依照隶属函数进行卷积. Hough 变换圆检测流程图如图 4 所示:

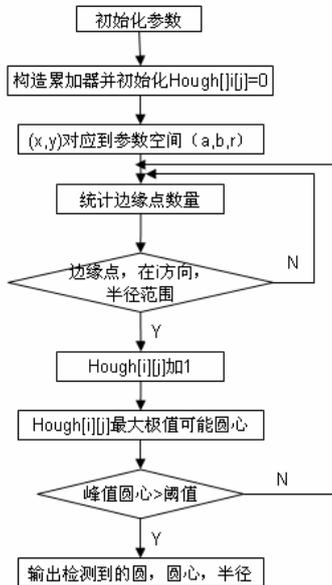


图 4 Hough 变换圆检测流程图

4 实验结果

由 Hough 变换圆检测流程图可以看出经过 Hough 变换识别出类圆纤维、圆心、半径, 然后对识别结果标识计数.

为验证算法的有效性, 采用不同光照下的图像进行实验, 实验结果如图 5 所示.

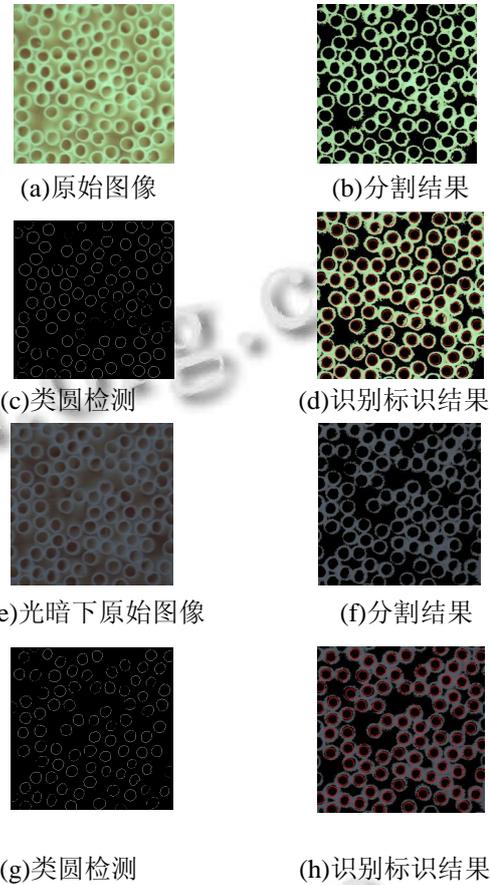


图 5 纤维的分割与识别图像

选择 100 幅纤维图像的识别结果进行统计如图 6 所示, 横坐标表示图片数据量(幅), 纵坐标表示识别结果的相对误差(%).

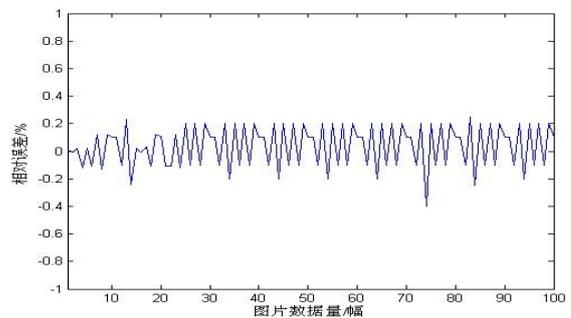


图 6 纤维识别结果及其分析

实验表明, 误差在 0.2% 以内的占样本总数的 86%, 其中实验样本中任一单幅图像的识别误差不超过 0.4%, 总体误差小于 0.1%.

5 结论

提出基本颜色聚类 and Hough 变换相结合的方法对纤维图像进行识别计数, 对不同光照情况下纤维识别的鲁棒性得到了增强. 下一步的工作主要考虑增强对少数变形纤维的识别能力.

参考文献

- Chen TW, Chen YL, Chien SY. Fast Image segmentation based on k-means clustering with histograms in HSV color space. *IEEE Trans. on Image Processing*, 2008,1(8): 322–325.
- Chen W, Shi YQ, Xuan GG. Identifying computer graphics using HSV color model and statistical moments of characteristic functions. *IEEE Proc. on Multimedia and Expo*, 2007,7(1):1123–1126.
- 孙可, 刘杰, 王学颖. K 均值聚类算法初始质心选择的改进. *沈阳师范大学学报(自然科学版)*, 2009,27(4):448–450.
- 丛培盛, 张洪江, 朱仲良. 用遗传算法 C-均值聚类分割医学彩色图像. *同济大学学报*, 2003,31(7):874–877.
- 袁理, 曹智睿. 改进的随机 Hough 变换圆检测算法. *计算机应用*, 2010,1(30):174–176.
- 刘念, 苏杭, 郭纯宏, 周静. 基于 Hough 变换圆检测的人眼定位方法改进. *急速阿奴及工程与设计*, 2011,32(4):1359–1362.
- Guo SY, Zhang XF, Zhang F. Adaptive randomized hough transform for circle detection using moving window. *International Conference on Machine Learning and Cybernetics*, 2006,10(13):3880–3885.
- Tang J. A color image segmentation algorithm based on region growing. *Computer Engineering and Technology*, 2010,6(2):634–637.
- Dragana C, Terry C. Region-Based coding of Color images Using Karhunen-Loeve Transform. *Department of Computer Science*, 1997,59(1):27–38.
- Shih FY, Cheng SX. Automatic seeded region growing for color image segmentation. *Image and Vision Computing*, 2005,23(10):877–886.
- Chien BC, Cheng MC. A color image segmentation approach based on fuzzy similarity measure. *IEEE International Conference on Fuzzy Systems*, 2002,1(2):449–454.
- Chitade AZ, et al. Color based Image segmentation using K-means clustering. *International Journal of Engineering Science and Technology*, 2010,2(10):5319–5325.

(上接第 173 页)

重新设置分割标志;

}
}

3 结语

本文利用 OpenGL 强大的图形功能在三维地形模拟方面进行了有益尝试, 在实践中取得了很好的效果, 通过键盘进行人机交互, 最终实现了三维地形的动态可视化.

参考文献

- 石教英. 虚拟现实基础及实用算法. 北京: 清华大学出版社, 1999.
- 和平鸽工作室. OpenGL 三维图形系统开发与实用技术. 重庆: 重庆大学出版社, 2003.
- 王玉琨, 王高峰, 刘启平. 基于四叉树的地形可视化研究. *地理与地理信息科学*, 2008,24(2):30–32.
- 王磊, 毛利民, 李骞. 基于 LOD 的三维地形可视化计算机与信息技术, 2007(7):39–41.
- 万定生, 龚汇丰. 大规模地形的 LOD 生成算法研究. *计算机工程与设计*, 2007,28(22):5345–5348.
- 殷宏, 许继恒, 周良伟, 等. 基于四叉树的大规模地形可视化及其实现. *计算机应用研究*, 2006,23(5):151–153.
- 张俊霞. 三维地形可视化及其实时显示方法概论. *北京测绘*, 2001(2):6–9.
- 廖中石, 杨化超. 基于 OpenGL 的三维地形实时交互显示现代测绘, 2004,27(3):43–44.
- 韩样. 基于 OpenGL 的三维地形可视化方法研究. *车辆与动力技术*, 2003,(2),11–15.
- 王玉琨, 姚华. 基于渐进网格的地形模型简化算法的研究. *微计算机信息*, 2010,8(3):203–205.
- 王臻, 胡敏, 李响. 基于四叉树的动态多分辨率 LOD 地形快速简化. *计算机应用*, 2007,(7):1642–1645.