

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2013 5
Vol.18 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013 年 5 月 第 18 卷 第 5 期(总第 205 期)

目 次

综述

- 中国图像工程:2012 章毓晋(483)
动态场景下的交通标识检测与识别研究进展 刘华平, 李建民, 胡晓林, 孙富春(493)

图像处理与编码

- 基于 PCA 硬阈值收缩的平滑投影 Landweber 图像压缩感知重构 李然, 干宗良, 朱秀昌(504)

图像分析和识别

- 利用视觉显著性和粒子滤波的运动目标跟踪 张巧荣, 冯新扬(515)
融合边缘测度的自然场景彩色图像区域分割 代沁伶, 王雷光, 洪亮(523)
场景语义树图像标注方法 刘咏梅, 杨帆, 于林森(529)
基于视锥细胞色适应生理机制的颜色恒常性计算 袁兴生, 王正志, 项凤涛(537)

医学图像处理

- 改进 FCM 和 LFP 相结合的白细胞图像分类 庞春颖, 刘记奎, 韩立喜(545)
局部高斯分布拟合的脑 MR 图像分割及有偏场校正 崔文超, 王毅, 樊养余, 冯燕, 郝重阳(552)

遥感图像处理

- 局部投影可分离的高光谱图像异常检测 刘明, 杜小平, 夏鲁瑞(558)
基于结构特征的遥感影像匹配 寇媛, 徐景中(565)

“第 9 届中国计算机图形学大会”专栏

- 结合浅景深与构图的图像质量评价 顾婷婷, 郭延文, 殷昆燕(574)
结合精确大气散射图计算的图像快速去雾 甘佳佳, 肖春霞(583)
多特征综合的视频拷贝检测 林莹, 杨扬, 凌康, 肖金伟, 武港山(591)
真实感雨线绘制 江隆盛, 吴恩华(600)
误差云映射的深度融合 3 维重建 陈亚当, 蔡仲谋, 吴恩华(607)

- “第八届中国系统建模与仿真技术高层论坛”征文通知 封 3

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 5 May 2013

Contents

Review

- Image engineering in China: 2012 Zhang Yujin(483)
- Recent progress in detection and recognition of the traffic signs in dynamic scenes
..... Liu Huaping, Li Jianmin, Hu Xiaolin, Sun Fuchun(493)

Image Processing and Coding

- Smoothed projected Landweber image compressed sensing reconstruction using hard thresholding based on principal components analysis
..... Li Ran, Gan Zongliang, Zhu Xiuchang(504)

Image Analysis and Recognition

- Object tracking based on visual saliency and particle filter Zhang Qiaorong, Feng Xinyang(515)
- Regional segmentation of natural scene color images by integrating edge cues Dai Qinling, Wang Leiguang, Hong Liang(523)
- Automatic image annotation based on scene semantic trees Liu Yongmei, Yang Fan, Yu Linsen(529)
- Color constancy algorithm based on biological mechanism; color adaptation of cone cells
..... Yuan Xingsheng, Wang Zhengzhi, Xiang Fengtao(537)

Medical Image Processing

- White blood cells image classification based on improving the connection of FCM and LFP
..... Pang Chunying, Liu Jikui, Han Lixi(545)
- Multiphase level set method for segmentation and bias correction of brain MR images
..... Cui Wenchao, Wang Yi, Fan Yangyu, Feng Yan, Hao Chongyang(552)

Remote Sensing Image Processing

- Anomaly detection method based on separable local projection in hyperspectral imagery ... Liu Ming, Du Xiaoping, Xia Lurui(558)
- Algorithm for remote sensing images matching based on the structure characteristics Kou Yuan, Xu Jingzhong(565)

Issum of the Chinagraph '2012

- Image quality assessment combining low DoF and composition Gu Tingting, Guo Yanwen, Yin Kunyan(574)
- Fast image dehazing based on accurate scattering map Gan Jiajia, Xiao Chunxia(583)
- Video copy detection based on multiple visual features synthesizing
..... Lin Ying, Yang Yang, Ling Kang, Xiao Jinwei, Wu Gangshan(591)
- Photorealistic rendering of rain streaks Jiang Longsheng, Wu Enhua(600)
- Three dimensional reconstructions based on error clouds mapping Chen Yadang, Cai Zhongmou, Wu Enhua(607)

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2013)05-0545-07

论文引用格式: 庞春颖, 刘记奎, 韩立喜. 改进 FCM 和 LFP 相结合的白细胞图像分类[J]. 中国图象图形学报, 2013, 18(5): 545-551.

改进 FCM 和 LFP 相结合的白细胞图像分类

庞春颖, 刘记奎, 韩立喜

长春理工大学生命科学技术学院, 长春 130022

摘要: 研究白细胞图像分类识别中有效的图像分割与特征提取方法, 以提高白细胞图像的正确识别率。由于某些白细胞(粒细胞)中颗粒的存在, 严重影响细胞核与细胞质区域的正确分割, 通过将空间信息与核函数融入模糊 C-均值聚类(FCM)算法, 提出一种改进的 FCM 算法。应用该算法对白细胞图像进行分割, 并采用数学形态学方法对分割后的图像进行处理, 获得了很好的分割效果, 解决了粒细胞的质核分割难题。对于细胞的纹理特征提取, 通过对局部二值模式(LBP)中阈值参数的模糊化, 建立了基于局部模糊模式(LFP)的纹理特征提取算法。运用本文方法进行图像分割和纹理提取, 以支持向量机作为分类器, 对 CellAtlas 的 100 幅白细胞图像进行了分类识别的实验, 结果表明白细胞的正确识别率达到 93%。

关键词: 白细胞分类; 图像分割; 模糊 C-均值聚类; 纹理特征提取; 局部模糊模式

White blood cells image classification based on improving the connection of FCM and LFP

Pang Chunying, Liu Jikui, Han Lixi

School of Life Science and Technology, Changchun University of Science and Technology, Changchun 130022, China

Abstract: To improve the correct recognition rate of white blood cells images, the effective methods of image segmentation and feature extraction are studied in this article. Because of the existence of grains in some type of white blood cells (granulocyte), the result of image segmentation is seriously affected. Integrating spatial information and kernel function into the fuzzy C-means clustering FCM algorithm, this paper proposes an improved FCM algorithm. Applying this new algorithm to image segmentation and taking the measure of mathematic morphology to process segmented image, the study gets a good segmentation effect and solves the problem of cytoplasm-nucleus of granulocyte segmentation. As for the feature extraction of cells, by fuzzification of the threshold parameter in local binary pattern (LBP), the texture feature extraction method based on local fuzzy pattern (LFP) is proposed. The employment of the methods above in image segmentation and texture extraction supports vector machine as the classifier and tests the classification of 100 CellAtlas's white blood cells images. The results indicate that the correct recognition rate is up to 93%.

Key words: white blood cells classification; image segmentation; fuzzy C-means clustering (FCM); texture feature extraction; local fuzzy pattern (LFP)

收稿日期: 2012-08-18; 修回日期: 2012-10-28

基金项目: 吉林省科技厅基金项目(2012006)

第一作者简介: 庞春颖(1971—), 女, 副教授, 2007 年于长春光机所获博士学位, 主要研究方向生物医学信号与图像处理。E-mail: pangchunying8888@163.com

通讯作者: 刘记奎, E-mail: liujikui2007@163.com

0 引言

白细胞的分类计数对疾病(如贫血、白血病和艾滋病等)的诊断以及血液成分的研究具有重要意义。细胞的分类识别技术也逐步由主观的人工分类发展到计算机软件的自动识别。

一幅通过流式细胞影像技术获得的白细胞图像主要包含 3 个区域:鞘流液区域、细胞核区域、细胞质区域。其中,鞘流液区为背景区域,细胞核与细胞质区为目标区域。白细胞图像的识别过程一般为:首先分割出细胞核与细胞质区域;然后分别提取细胞的几何特征以及细胞核与细胞质的纹理特征;最后通过神经网络或支持向量机的方法进行特征识别。细胞质和细胞核的分割是白细胞正确分类识别的前提。许多文献提出了采用诸如阈值分割^[1]、分水岭变换分割^[2]以及模糊 C-均值聚类(FCM)分割等算法^[3],但它们都不能很好地解决粒细胞图像错分率高的问题。

在细胞识别的过程中,细胞特征的选择和提取至关重要,是细胞识别的关键。细胞的图像特征主要包括几何特征、颜色特征与纹理特征。几何特征可以直观地反应白细胞的外在形态,是细胞识别的重要依据,它包括细胞与细胞核的周长与面积、细胞核圆度等。纹理特征描述的是物体表面局部的灰度变化,在模式识别领域得到广泛的应用。用于纹理分析的算法很多,如文献[1]采用统计学方法提取纹理特征,文献[4-6]提出了以数学形态学方法(模式谱或粒度测量)提取细胞的纹理特征,而更多的文献提出采用局部二值模式(LBP)与灰度共生矩阵(GLCM)的方法提取白细胞的纹理特征^[7]。但是这些方法因存在着对噪声敏感的缺点,正确识别率最高为 85%^[8]。

提出了将改进的 C-均值聚类 FCM 算法与数学形态学运算相结合的方法。通过将空间信息与核函数同时融入传统的 FCM 中,利用该算法对白细胞图像进行分割,并采用数学形态学的方法对分割后的二值图像进行修饰。很好地解决了粒细胞的分割难题。并采用局部模糊模式(LFP)提取白细胞图像的纹理特征。实验结果表明,LFP 比 LBP 在对含有噪声的白细胞图像分类识别中表现出优良的性能。这种图像分割方法和纹理特征提取方法使得细胞的正确识别率有了显著的提高。

1 白细胞分类识别方法

白细胞分类计数系统作为帮助医生或血液病学家对疾病做出快速与精确判断的智能医学系统,其核心是实现白细胞显微图像的自动分类。在临床上常依据白细胞的细胞质内有无颗粒分为粒细胞和无粒细胞两大类。其中,粒细胞根据其颗粒染色特征不同又分为嗜碱性粒细胞、嗜酸性粒细胞、嗜中性粒细胞,而无粒细胞又分为单核细胞、淋巴细胞。本文采用的白细胞分类过程如图 1 所示。首先对单个的白细胞图像进行细胞质与细胞核区域分割,再对其进行几何特征与纹理特征的提取,最后通过支持向量机(SVM)对其分类。

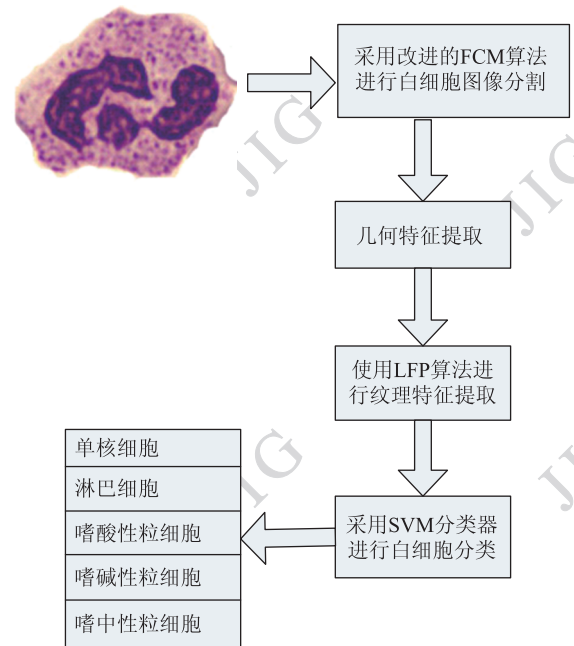


图 1 白细胞分类过程

Fig. 1 White blood cells classification process

1.1 白细胞图像分割

FCM 算法是一种特别适用于多元复杂图像的分割技术,采用该算法对白细胞图像进行质核分割。

1.1.1 传统 FCM 算法

假定有限数据集 $X = \{x_1, x_2, \dots, x_n\} \in \mathbf{R}^S$, 其中 x_n 为样本点向量, S 为 x_n 维数。 $P = \{v_1, v_2, \dots, v_c\}$ 为初始聚类中心集, v_c 为聚类中心点, c 为聚类数 ($1 < c < n$)。FCM 的聚集快慢取决于数据点集与聚类中心点间的平均距离,当聚类算法的目标函数 J_m 达到最小时获得最佳的分类效果, J_m 定义为

$$J_m = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \| \mathbf{x}_i - \mathbf{v}_j \|^2, 1 \leq m < \infty \quad (1)$$

式中, m 为加权系数; u_{ij} 为表示 $\mathbf{x}_i$ 与 j 类之间的隶属度; $\| \mathbf{x}_i - \mathbf{v}_j \|$ 表示样本 $\mathbf{x}_i$ 与 j 类间的欧氏距离。在每次迭代使目标函数最佳化时, FCM 算法也需要更新隶属度 u_{ij} 与聚类中心 $\mathbf{v}_j$, 即

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\| \mathbf{x}_i - \mathbf{v}_j \|}{\| \mathbf{x}_i - \mathbf{v}_k \|} \right)^{\frac{2}{m-1}}} \quad (2)$$

$$\mathbf{v}_j = \frac{\sum_{i=1}^n u_{ij}^m \cdot \mathbf{x}_i}{\sum_{i=1}^n u_{ij}^m} \quad (3)$$

对于隶属度 u_{ij} 有约束关系

$$\begin{cases} \sum_{j=1}^c u_{ij} = 1 \\ 0 \leq u_{ij} \leq 1 \\ 0 \leq \sum_{i=1}^n u_{ij} \leq 1 \end{cases} \quad (4)$$

根据以上原理, 设计 FCM 算法的迭代步骤如下:

- 1) 初始化聚类中心集 $\mathbf{P}$, 计算得到隶属度矩阵 $\mathbf{U} = [u_{ij}]$, 记为 $\mathbf{U}^{(0)}$;
- 2) 迭代步骤 d : 聚类中心向量 $\mathbf{P}^{(d)} = [\mathbf{v}_j]$, 隶属度矩阵为 $\mathbf{U}^{(d)}$;
- 3) 更新 $\mathbf{U}^{(d)}$ 、 $\mathbf{U}^{(d+1)}$;
- 4) 如果 $\| \mathbf{U}^{(d+1)} - \mathbf{U}^{(d)} \| < \varepsilon$, 迭代停止; 否则, 返回步骤 2)。

对白细胞图像使用上述 FCM 算法的分割结果如图 2 所示。由图 2 可见, 当所输入的细胞图像是较平滑的单核细胞或淋巴细胞图像时 (见图 2(a)), 可以得到较好的分割效果 (见图 2(b)); 但当输入的是较粗糙的粒细胞图像时 (见图 2(c)), 则有很大的误分率 (见图 2(d))。很显然, 传统的 FCM 算法对于细胞质内含有粗细不一颗粒的白细胞分割极不准确。

1.1.2 FCM 算法的改进

引起上述分割结果的主要原因是传统的 FCM 算法对噪声或局部斑点十分敏感, 以至于常常会把颗粒像素错误地划分到细胞核区域, 导致了粒细胞的细胞核和细胞质的无效分割。从 FCM 算法本质来看, 产生这种缺陷缘于两点: 1) 该算法忽视了图像的空间信息^[9-12]; 2) 该算法

所处理的数据集是多样性的, 造成在低维空间很难准确的对其分类^[9,13-14]。因此, 需要对 FCM 算法进行改进。

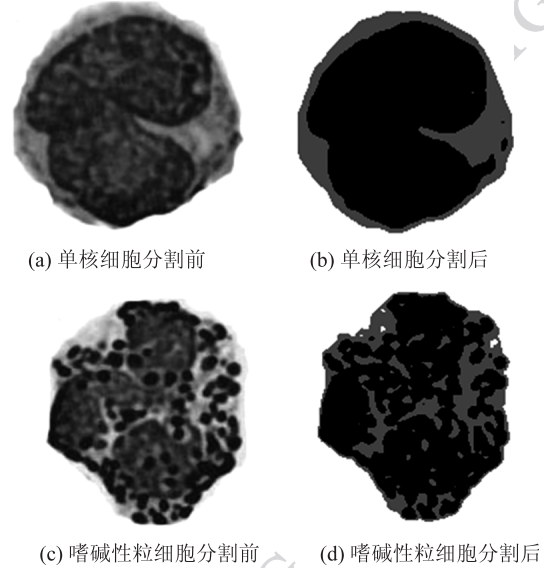


图 2 基于 FCM 算法的细胞图像分割

Fig. 2 Cell image segmentation based on FCM algorithm

如何将图像的空间信息融入传统 FCM 算法来克服上述缺点, 不少学者对此进行了研究。在文献 [15] 中, 通过将空间信息融入传统 FCM 算法, 并称之为 SFCM。空间函数定义为

$$h_{ij} = \sum_{k \in NB(\mathbf{x}_i)} u_{ik} \quad (5)$$

式中, $NB(\mathbf{x}_i)$ 表示以像素点 $\mathbf{x}_i$ 为中心的局部方形或圆形窗口; h_{ij} 表示像素 $\mathbf{x}_i$ 隶属于 j 聚类的几率, 它的作用是将中心像素归为与大多数邻域像素相同的簇。 h_{ij} 在数值上等于不包含中心像素 $\mathbf{x}_i$ 的所有邻域像素的隶属度之和。

改进的含有空间信息的模糊隶属度为

$$\mu_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{j=1}^c u_{ij}^p h_{ij}^q} \quad (6)$$

式中, 参数 p, q 分别表示中心像素与空间信息的权重。这样隶属度不仅取决于中心像素本身, 而且受邻域窗内所有像素的影响。由式 (6) 可知, 像素隶属度值随着空间函数值的增加而增加, 随之减小而减小。

实现 SFCM 算法分为两个步骤: 第 1 步如同传统的 FCM 算法, 计算出隶属度矩阵与聚类中心; 第 2 步将空间信息融入其中得到新的模糊隶属度。

解决第 2 步的关键是将处于低维空间的数据集映射至高维空间,即让 $\mathbf{x}_i \in \mathbf{x} \subseteq \mathbf{R}^s \rightarrow \phi(\mathbf{x}_i) \in \mathbf{F} \subseteq \mathbf{R}^p$, 在这里 $S \ll P$ 。为了避免映射为高维空间所引起的维度灾难,引入核函数。常见的核函数有高斯核、多项式核以及 S 形核。高斯核定义为

$$k(\mathbf{x}_i, \mathbf{y}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{y}_j\|^2}{\sigma^2}} \quad (7)$$

核函数的引入使得无需知道具体的映射函数 ϕ , 当满足 Mercer 条件时, $k(\mathbf{x}_i, \mathbf{y}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{y}_j)$ 。因此,在计算 $\mathbf{F}$ 空间内积时可以用 $k(\mathbf{x}_i, \mathbf{y}_j)$ 代替。这样,当在 SFCM 算法中嵌入核函数时,结合式(5)(6)可得目标函数为

$$\begin{aligned} J_m &= \sum_{i=1}^n \sum_{j=1}^c \mu_{ij}^m \|\phi(\mathbf{x}_i) - \phi(\mathbf{v}_j)\|^2 \\ \|\phi(\mathbf{x}_i) - \phi(\mathbf{v}_j)\|^2 &= \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_i) - \\ &2\phi(\mathbf{x}_i) \cdot \phi(\mathbf{v}_j) + \phi(\mathbf{v}_j) \cdot \phi(\mathbf{v}_j) = \\ &k(\mathbf{x}_i, \mathbf{x}_i) - 2k(\mathbf{x}_i, \mathbf{v}_j) + k(\mathbf{v}_j, \mathbf{v}_j) \end{aligned} \quad (8)$$

当在传统 FCM 算法中嵌入核函数时,模糊隶属度与聚类中心表示为

$$u_{ij} = \frac{(1 - k(\mathbf{x}_i, \mathbf{y}_j))^{-\frac{1}{(m-1)}}}{\sum_{k=1}^c (1 - k(\mathbf{x}_i, \mathbf{y}_k))^{-\frac{1}{(m-1)}}} \quad (9)$$

$$\mathbf{v}_j = \frac{\sum_{i=1}^n u_{ij}^m \cdot k(\mathbf{x}_i, \mathbf{y}_j) \cdot \mathbf{x}_i}{\sum_{i=1}^n u_{ij}^m \cdot k(\mathbf{x}_i, \mathbf{y}_j)} \quad (10)$$

由式(5)(6)(9)(10)可得出在 SFCM 中嵌入高斯核函数的模糊隶属度与聚类中心公式,这里不再赘述。

利用改进的 FCM 算法对粒细胞图像再次进行分割,如图 3(b)所示可得出较好的分割结果。但分割出的图像有时仍会存在孤岛或孔洞,这需要用数学形态学方法对分割后的二值图像进行修饰,最终结果如图 3(c)所示。

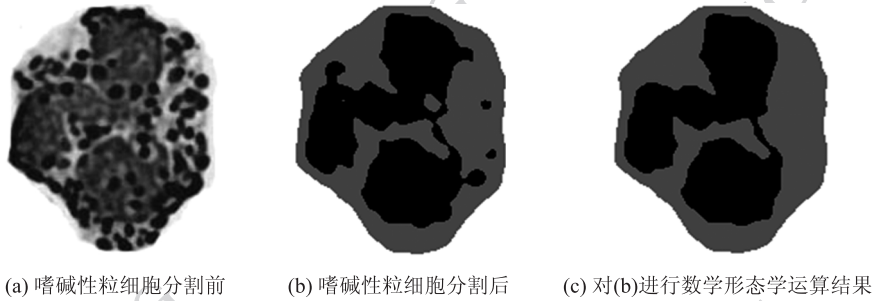


图 3 基于改进的 FCM 算法的细胞图像分割

Fig. 3 Cell image segmentation based on improved FCM algorithm

1.2 图像特征提取

特征向量是图像在某些方面的表征,是对白细胞图像正确识别的依据。使用几何与纹理两种特征。值得注意的是选用神经网络或支持向量机作为分类器时,特征向量维数不易过多,否则会引起维度灾难以及淹没主要的识别特征^[16-17]。

几何特征直观地反映了白细胞的外在形态,是细胞识别的重要依据。然而几何特征具有自身的缺陷,由于所处理图像反映的是白细胞的 2 维结构,也就是说同一种细胞不同的拍摄方位将会出现不同的外在形态,因此几何特征的选取不宜过多。另外所选择的图像特征需要尽可能具有旋转和平移不变性,又由于 5 类白细胞的细胞核大小与形状复杂度、细胞体大小以及细胞核体积占整个细胞体的比重等形态特征具有很明显的差异,而这些形态特征可以

用细胞核周长、细胞核面积、细胞核圆度、细胞面积以及细胞核与细胞面积比 5 种特征直接量化度量(细胞体或核的大小以及体积在 2 维图像中表现为面积尺寸)。例如,对于细胞核圆度整体上有:淋巴细胞 > 单核细胞 > 嗜酸性粒细胞 > 嗜碱性粒细胞 > 嗜中性粒细胞。细胞核圆度定义为 $\frac{C_{nucleus}^2}{A_{nucleus}}$, 其中, $C_{nucleus}$ 表示细胞核周长, $A_{nucleus}$ 表示细胞核面积。

在细胞识别领域,由于细胞内的染色质与颗粒的分布都有一定的随机性,这样图像无论取自哪个视角,都会得到比较相似的纹理分布^[17]。采用 LFP 提取纹理特征。LFP 算法的本质是将模糊逻辑融入 LBP 算法,通过模糊化方法将二值化按照相应的隶属度作“软”处理,很好地克服了 LBP 算法对斑点噪声敏感的缺陷^[17-18]。

具体来说,LBP 算法只是以中心像素灰度值为阈值对邻域像素作“硬”的二值化编码,原理上的缺陷造成两个问题:1)会将中心阈值附近相差甚微的两个灰度值作出“0”或“1”不同的编码,而人眼很难分辨这种差异;2)抗噪声性能差。LFP 算法模拟了人眼的含糊不清与数字量化不精确的特点,根据邻域内像素点与中心像素点的灰度差值大小,用模糊隶属度函数 m_0 和 m_1 来分别代替 LBP 算法中的“0”和“1”编码。两个隶属度函数 m_0 和 m_1 定义为

$$m_0(i) = \begin{cases} 0 & p_i \geq p_{\text{center}} + T \\ \frac{T - p_i + p_{\text{center}}}{2T} & p_i \geq p_{\text{center}} - T < p_i < p_{\text{center}} + T \\ 1 & p_i \leq p_{\text{center}} - T \end{cases} \quad (11)$$

$$m_1(i) = 1 - m_0(i) \quad (12)$$

式中, $p_i (0 \leq i \leq 7)$ 表示邻域像素值; p_{center} 表示中心点像素值; $T (>0)$ 为设定的阈值,控制 LFP 的模糊程度,不难看出 LBP 是 LFP 在“ $T=0$ ”时的一个特例。如图 4 所示,隶属度函数反映了相邻像素点的相似程度,当 m_0 和 m_1 均不为“0”或“1”时,表明该局部区域纹理是均匀的;反之,则不均匀。

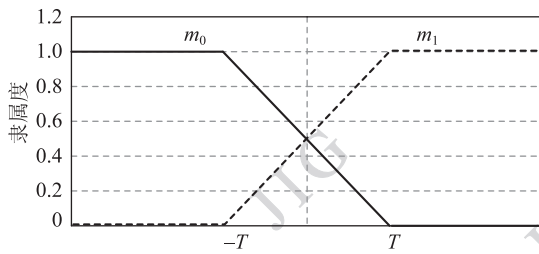


图4 隶属度函数

Fig. 4 Membership function

由于传统 LBP 算法不具备旋转不变性,以及所提取的特征数过多,十分不利于图像的识别,而 LFP 算法遗传了 LBP 算法上述缺点。为解决上述问题,首先引入了 Uniform 模式的 LBP 算子 $LBP_{P,R}$,所谓 Uniform 模式是指 LBP 算法中的局部二值序列中“0”与“1”之间的跳变次数小于等于 2 的模式。事实上,Uniform 模式的发生概率占有所有模式发生概率的 99% 以上^[19],而它却只有 58 种特征,远小于传统 LBP 的 256 种特征。为实现旋转不变性问题,对 LBP 算法设置规定,当一种局部二值序列通过旋转能够得到另一种二值序列时(二值序列是循环的),则这两种序列视为同一种。这样就得到了具有旋转不变性的 Uniform 模式的 LBP 算子 $LBP_{P,R}^{riu2}$,即

$$\begin{cases} LBP_{P,R}^{riu2} = \begin{cases} \sum_{i=0}^{p-1} u(P_i - P_{\text{center}}) & U \leq 2 \\ P + 1 & \text{其他} \end{cases} \\ u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \end{cases} \quad (13)$$

式中, U 表示二值序列中“0”到“1”或“1”到“0”的跳变次数。当 $P=8, R=1$ 时(8 邻域), $LBP_{P,R}^{riu2}$ 最大值为 9,这样特征值数量又进一步缩小为 10 个。

在 LBP 算法中,每个像素点具有单一的 LBP 值,而 LFP 中则具有多个 LBP 值,并且每个 LBP 值具有相应的隶属度权重 $M(LBP_i)$ 。LFP 算法的实现过程如下:

- 1) 设定阈值参数 T , 计算邻域内每个像素点的隶属度 $m_0(i)$ 和 $m_1(i)$;
- 2) 根据 LBP 算法中 LBP 值计算方法计算 LBP_i , 以及对应的 $M(LBP_i)$;
- 3) 计算 LFP 值, $C_{LFP} = \sum_i M(LBP_i) \times LBP_i$;
- 4) 对 C_{LFP} 作概率统计,得出 LFP 特征向量。

2 实验与结果分析

为了评估上述算法的性能,在实验中采用来自美国 CellAtlas 细胞图库的 100 幅白细胞图像作为样本。从样本中随机抽取 60 幅作为训练样本,剩下 40 幅作为测试样本,对于每类白细胞分别使用 12 幅图像进行训练。实验主要分为 3 个阶段:首先进行图像分割,采用改进的 FCM 聚类算法分割出细胞核与细胞质区域;其次进行特征提取,采用 5 个几何特征与 20 个纹理特征组成一组特征向量,其中 20 个纹理特征由 10 个细胞核纹理特征与 10 个细胞质纹理特征组成;最后进行图像分类识别,采用支持向量机作为分类器,它以结构风险最小化原理为基础,对小样本、非线性及高维模式识别问题表现出许多优势。

在图像分割阶段,如图 3 所示改进的 FCM 算法与传统的 FCM 算法相比展现出良好的分割效果。LFP 算法中参数 T 的选择尤为重要,在文献[20-21]中对其进行了探讨,指出当 $T=5$ 时为有最佳的分类效果。本文也对其进行了研究:当 $T=0$ 时,有与 LBP 相同的分类效果;当 $0 < T \leq 5$ 时,分类能力随着 T 的增加而增强;当 $T > 5$ 时,其分类能力不再增强,并且开始下降,表明 $T=5$ 是最佳分类参数。因此在

实验中最终将参数 T 设为 5。

实验中对使用粒度测量技术 (granulometry)、GLCM、LBP 以及 LFP 提取纹理特征的性能进行了对比研究,结果如图 5 所示。可以看出采用 LFP 算法提取纹理特征的精确度要明显高于 LBP 算法以及其他算法。

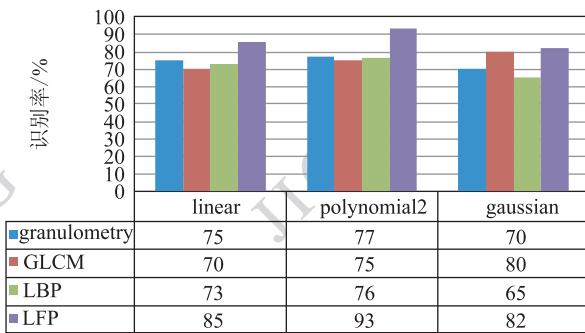


图 5 应用不同算法的白细胞图像分类效果对比图

Fig. 5 White blood cells image classification effect contrast figure for different algorithm

3 结 论

FCM 聚类算法是一种无监督聚类技术,其原理是基于图像像素点与聚类中心的欧氏距离进行聚类划分的,它适用于区域灰度较为平滑的图像分割。但是,它对含有噪声或者局部区域变化剧烈的图像分割效果较差,这主要由于一方面,它忽略了蕴含在图像中的像素与像素间的空间信息,而这些空间信息又是像素聚类的重要特征;另一方面,处理数据集的多样性造成在低维空间很难对其进行准确分类。

将空间信息与核函数同时嵌入传统 FCM 算法中,引入了改进的 FCM 算法,通过对比图 2 与图 3 中对嗜碱性粒细胞的分割结果可知,改进的 FCM 算法克服了传统 FCM 算法对噪声与局部斑点敏感的缺陷,对白细胞中粒细胞图像具有良好的分割效果。

在对白细胞图像进行特征提取的环节,由于几何特征具有相对的不确定性,在特征值提取中仅采用了 5 个关键的几何特征。相反,纹理特征具有相对的确定性,因此在特征提取中赋予其更高的权重,共采用了 20 个纹理特征。纹理特征提取采用的 LFP 算法通过将模糊逻辑引入 LBP 二值化处理中,充分考虑到了噪声和数字图像量化过程中引起的模糊性与不精确性,使之更符合人眼的视觉心理,所提取的特征也更能表征原始图像。

综上,采用的改进的 FCM 算法与数学形态学相结合的方法以及基于 LFP 的图像纹理特征提取方法使白细胞的正确识别率达到 93%,充分表明了这种图像分割算法与特征提取算法可以很好地应用于白细胞图像的分类计数系统。

参考文献 (References)

- [1] Zhu X F. Study of key techniques of imaging flow cytometry for bio—particle detection and recognition [D]. Tianjin: Tianjin University, 2010. [朱险峰. 生物颗粒流式成像关键技术的研究 [D]. 天津: 天津大学, 2010.]
- [2] Jiang K, Liao Q M, Dai S Y. A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering [J]. Proceedings of the 2nd International Conference on Machine Learning and Cybernetics, 2003, 5(5): 2820-2825.
- [3] Sudhavani G, Sathyaprasad K. Segmentation of lip images by modified fuzzy C-means clustering algorithm [J]. International Journal of Computer Science and Network Security, 2009, 9(4): 187-191.
- [4] Gómez-Gil P, Ramírez-Cortés M, González B J, et al. A feature extraction method based on morphological operators for automatic classification of leukocytes [C]//Proceedings of the 7th Mexican International Conference on Artificial Intelligence. USA: IEEE Computer Society, 2008, 10: 227-232.
- [5] Theera-Umporn N, Dhompangsa S. Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification [J]. IEEE Transactions on Information Technology in Biomedicine, 2007, 11(3): 353-359.
- [6] Angulo J, Klossa J, Flandarin G. Ontology-based lymphocyte population description using mathematical morphology on colour blood images [J]. Cellular and Molecular Biology TM, 2006, 52(6): 2-15.
- [7] Rezaatofghi S H, Khaksari K, Hamid S Z. Automatic recognition of five types of white blood cells in peripheral blood [C]//LNCS 6112. Berlin: Springer-Verlag, 2010: 161-172.
- [8] Fatichah C, Tangel M L, Widyanto M R, et al. Parameter Optimization of Local Fuzzy Patterns for Extracting White Blood Cell Texture Feature [C]//World Congress of Int. Indonesia: Fuzzy Systems Association, 2011: 153-159.
- [9] Zhang X Q, Li X F, Liu Y N. Gaussian kernelized fuzzy c-means with spatial information algorithm for image segmentation [J]. Journal of Computers, 2012, 7(6): 1511-1516.
- [10] Chen S C, Zhang D Q. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure [J]. IEEE Transactions on System, 2004, 134(4): 1907-1917.
- [11] Keh-Shih C, Tzeng H L, Chen S, et al. Fuzzy c-means clustering with spatial information for image segmentation [J]. Com-

- puterized Medical Imaging and Graphics, 2006,30(1):9-15.
- [12] Ge Q, Wei Z H, Zhang J W, et al. Multiphase CV model integrated with improved FCM algorithm[J]. Journal of Image and Graphics, 2011, 16(3):547-553. [葛琦, 韦志辉, 张建伟, 等. 结合改进 FCM 算法的多相位 CV 模型[J]. 中国图象图形学报, 2011, 16(3):547-553.]
- [13] Girolami M. Mercer kernel-based clustering in feature space[J]. IEEE Trans. on Neural Networks, 2002, 13(3):780-784.
- [14] Yang M S. A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction[J]. Pattern Recognition Letters, 2008, 29(1):1713-1725.
- [15] Ahmed M N, Yamany S M, Mohamed N. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data[J]. IEEE Transactions on Medical Imaging, 2002, 21(3):193-199.
- [16] Dong C, Yang S R, Li T F, et al. Modeling and verification for equivalent NN of general two-dimensional fuzzy controller[J]. Computer Engineering and Design, 2008, 29(11):2912-2917. [董淳, 杨绍荣, 李太傅, 等. 一般二维模糊控制器的等效神经网络建模与验证[J]. 计算机工程与设计, 2008, 29(11):2912-2917.]
- [17] Barros de Almeida M, De Pádua Braga A, Braga J P. SVM-KM: speeding SVMs Learning with a priori Cluster Selection and k-Means[C]//The 6th Brazilian Symposium on Neural Networks. USA: IEEE Xplore, 2000:162-167.
- [18] Caponetti L, Castiello C, Fanelli A M, et al. Texture segmentation with local fuzzy patterns and neuro-fuzzy decision support[C]//LNAI 4252. Berlin:Springer-Verlag, 2006:340-347.
- [19] Tajeripour F, Kabir E, Sheikhi A. Fabric defect detection using modified local binary patterns[J]. Eurasip Journal on Advances in Signal Processing, 2008:1-12.
- [20] Iakovidis K, Keramidas G, Dimitris M. Fuzzy local binary patterns for ultrasound texture characterization[C]//LNCS 5112. Berlin:Springer-Verlag, 2008:750-759.
- [21] Fatichah C, Tangel M L, Widyanto M R, et al. Parameter optimization of local fuzzy patterns based on fuzzy contrast measure for white blood cell texture feature extraction [J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2012, 16(3):412-419.