

骨髓细胞显微图像的分形特征分析

张小京 孙万蓉 钟政辉

(西安电子科技大学工程学院, 西安 710071)

摘要 为了更好地完成骨髓细胞图像的分类识别,在对临床血液学及细胞学图谱中的细胞图像进行切分的基础上,提出了一种基于分形维数和图像重心的算法,用来提取每一细胞图像的分维特征。由于图像的纹理不仅表现在图像结构上具有某种统计意义的相似性,还反映在彩色分布方面,因此,可从骨髓细胞的真彩色图像中提取一个新的彩色识别分量图像 C ,并与 HIS 彩色空间的饱和度分量图像 S 一起进行分维特征提取。实验数据表明,由于不同类型骨髓细胞图像的分形参数具有不同程度的差异,因此分形参数对于某些类别的骨髓细胞具有较好的区分能力。

关键词 骨髓细胞 分形维数 差分盒子法 彩色分量

中图分类号: R318.4 文献标识码: A 文章编号: 1006-8961(2006)05-0624-06

Fractal Feature of Marrow Cell Images

ZHANG Xiao-jing, SUN Wan-rong, ZHONG Zheng-hui

(School of Electronic Engineering, Xidian University, Xi'an 710071)

Abstract This paper proposes a new recognition algorithm based on fractal dimension and image center of gravity after the segmentation of marrow cell images in the clinical haematology and cytology atlas which can be used to extract. The fractal feature of every marrow cell image. Since the texture of a color image contains not only certain statistical similarity but also the color distributions, a new color parameter is extracted from the color images of marrow cells and discussed with parameter S of HIS for the fractal dimension. The experimental result shows that different marrow cell image has the certain discrimination in fractal dimension and some images can be recognized for classification.

Keywords marrow cell, fractal, differential box counting, color parameter

1 引言

由 Mandelbrot 等提出的分形几何认为复杂景象的维数可以是分数^[1],这样就为定量分析自然图像的复杂形貌提供了手段^[2]。本文将分形几何的这一特性应用于骨髓细胞显微图像的分类识别研究中,并在对《临床血液学及细胞学图谱》^[3]中的细胞图像进行切分的基础上,提出了一种基于分形维数和图像重心的算法,用来提取每一细胞图像的分维特征。由于不同纹理图像可能会有相同的分维值,而仅靠单一灰度图像的分维值又不能得到足够的信

息用来达到准确分类识别的目的,因此,本文从骨髓细胞的真彩色图像中提取了一个新的彩色识别分量 $C = B/(R + G + B)$,并将其与 HIS 彩色空间的饱和度和分量图像一起进行了分形讨论。

2 分形理论及其计算方法

2.1 分形的理论基础

众所周知,自然界中的大量形体和自然景物表面,例如曲折的海岸线、错综复杂的地貌等都是复杂不规则的,而且它们无法用简单的结构元素来描述。而在分形几何中,则可以用分形维数来度量事物的

基金项目:陕西省自然科学基金资助项目(2002F04)

收稿日期:2005-03-20;改回日期:2005-05-30

第一作者简介:张小京(1981 ~),女,2003年获西安电子科技大学工学学士学位,现为西安电子科技大学工程学院硕士研究生。主要研究方向为医学信号的图像处理。E-mail:flyandfly8@126.com

不规则程度,这种分形维数在直观上与物体表面的粗糙度相吻合。目前,定义分形维数的方法有多种,其中由于盒维数法便于实现计算机计算,因而成为应用最广泛的一种分形维数定义方法。

盒维数的基本定义为^[4]:设 F 是 \mathbf{R}^n 上任意非空的有界子集, $N_\sigma(F)$ 是最大直径为 σ 的盒子数,即可以覆盖 F 集的最少个数,则 F 的下、上盒维数分别定义为

$$D_F = \lim_{\sigma \rightarrow 0} \frac{\log N_\sigma(F)}{-\log \sigma} \quad (1)$$

$$\bar{D}_F = \lim_{\sigma \rightarrow 0} \frac{\log N_\sigma(F)}{-\log \sigma} \quad (2)$$

如果这两个值相等,则称这个值为非空的有界子集 F 的盒维数,并记为

$$D_F = \lim_{\sigma \rightarrow 0} \frac{\log N_\sigma(F)}{-\log \sigma} \quad (3)$$

2.2 分形维数的计算方法

关于图像分形维数的计算方法目前主要分为以下两类:一类是为了便于实现计算机计算而提出的方法,如 Sarkar 根据 Box-counting 维数的定义提出的差分盒子计数法^[5]等。另一类是基于分形模型的计算方法,如 Pentland 提出的基于分形布朗运动模型的傅里叶变换功率谱密度计算分形维数法^[6]。由于一般图像都是空间域的,因此如果用傅里叶域的功率谱来估计分形维数,则显然要增加大量的傅里叶变换计算,而基于空间域的差分盒子计数法则可使图像纹理的分形维数计算简洁方便。

本文将差分盒子计数法用于骨髓细胞图像分形维数的估计。对于给定面积为 $M \times M$ 的图像,假设它已被分解成 $s \times s$ 的小块, s 是介于 1 和 $M/2$ 之间的整数,此时尺度因子 $r = s/M$ 。通常可把图像看作一个 3 维空间 (x, y, z) , 其中 (x, y) 表示像素点的平面位置, z 表示像素点的灰度值。若把图像平面 (x, y) 划分为一个个 $s \times s$ 大小的网格,则可以认为在每一个网格的 z 方向上都有一系列 $s \times s \times \hat{s}$ 大小的长方体小盒子,其中 \hat{s} 为长方体的高。为保证研究对象在 x, y, z 3 个方向上有相同的压缩因子,可对盒子的尺寸作如下限制:

$$\frac{g}{\hat{s}} = \frac{M}{s} = \frac{1}{r} \quad (4)$$

其中, g 为图像的灰度级数。

假如在第 (i, j) 个网格内,灰度的最小值和最大值分别落在第 k 个和第 l 个小盒子内,那么第 (i, j) 个网格内的小盒子数目为 $n_r(i, j) = l - k + 1$, 而整

个图像的小盒子总数则为 $N_r = \sum n_r(i, j)$ 。对应于不同的 r 值,可以得到不同的 N_r 值;再由式(3),采用最小二乘法线性回归法拟合出 $(\log(N_r), \log(1/r))$ 的斜率,即为该图像的分形维数 D_F 。

令 $x_r = -\log r, y_r = \log N_r$, 则用最小二乘法估计的非空有限子集 F 的分形维数为

$$D_F = \frac{\sum_r (x_r - \bar{x}_r)(y_r - \bar{y}_r)}{\sum_r (x_r - \bar{x}_r)(x_r - \bar{x}_r)} \quad (5)$$

式中, \bar{x}_r, \bar{y}_r 是 x_r, y_r 的统计平均值。

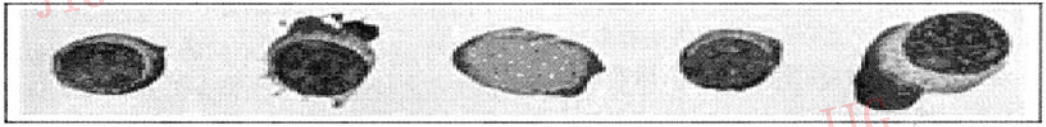
3 骨髓细胞显微图像的分形参数提取

本文主要针对《临床血液学及细胞学图谱》^[3] 中的红细胞、嗜碱粒细胞、嗜酸粒细胞、单核细胞、淋巴细胞、分叶核粒细胞、杆状核粒细胞及浆细胞进行分形特征的讨论,以便建立相应的查询列表来应用于临床分析。该《临床血液学及细胞学图谱》包含所有不同的细胞类型及各类型的不同演变阶段,而且它是采用瑞氏(Wright)染色,经 1000 倍油镜放大得到的。本文在对这些医学图像分形特征参数进行提取时,首先从医学图谱中切分出不同类别、不同演变阶段骨髓细胞图,其大小为 100×100 、存储格式为 BMP 格式,切分后得到了 22 幅不同演变阶段的红细胞图像、35 幅不同演变阶段的嗜碱粒细胞图、39 幅不同演变阶段的嗜酸粒细胞图像、43 幅不同演变阶段的单核细胞图像、38 幅不同演变阶段的淋巴细胞图像、25 幅不同演变阶段的嗜中性粒细胞(杆状核和分叶核中性粒细胞)图像及 30 幅不同演变阶段的浆细胞图像。图 1 分别列出了不同类型细胞的 5 个不同演变阶段的图像。

以骨髓细胞显微图像为研究对象,本文提出的分形维数分析算法如下:

(1) 首先从每幅切分图像的真彩色图像中提取饱和度分量图像 S ;

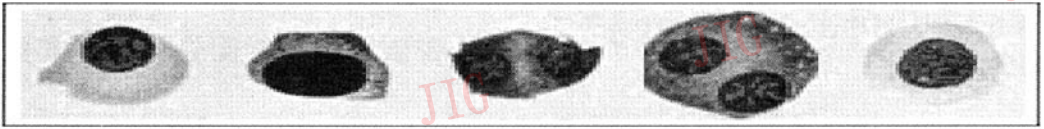
(2) 计算出每幅饱和度分量图像 S 中细胞的重心点,并以此重心点为参考点(即图像中心),再以细胞的长宽比为依据,提取出大小合适的细胞图像,由于细胞位置具有随机性,因此需先求出细胞的主轴方向(即有最大长度和最小宽度的方向),然后,相对于细胞主轴来测量细胞的长度及宽度,长度和宽度所在直线的交点即为图像的重心点,其中确定



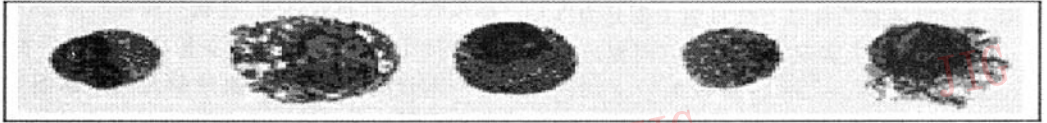
(a) 红细胞图像



(b) 淋巴细胞图像



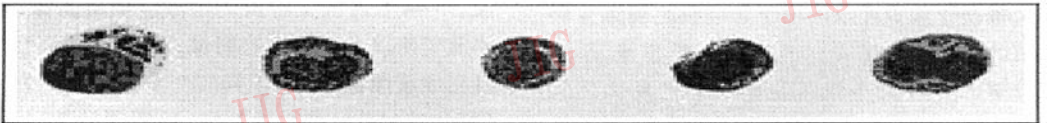
(c) 浆细胞图像



(d) 嗜酸粒细胞图像



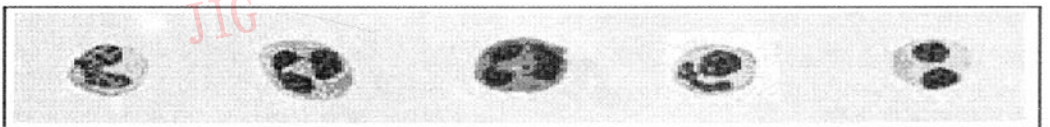
(e) 单核细胞图像



(f) 嗜碱粒细胞图像



(g) 杆状粒细胞图像



(h) 分叶核粒细胞图像

图1 不同类别、不同演变阶段骨髓细胞灰度图像
Fig.1 Eight types of marrow cells at different stages

主轴的方法可用目标的二阶中心矩获得;

(3)以细胞的长宽为标准,通过基于标准偏差的自适应无标度区求解方法^[7],对于长宽相近的细胞图像可采用相同的无标度区,而对于不同类别、不同演变阶段的细胞图像,则在相应的无标度区内采用最小二乘法进行线性回归,在本文选定的无标度区内,相关系数 R 的值介于 0.9994 ~ 1.0000 之间;

(4)利用差分盒子法求出不同阶段的红细胞、嗜碱粒细胞、嗜酸粒细胞、单核细胞、淋巴细胞、分叶核粒细胞、杆状核粒细胞及浆细胞图像的饱和度分量图像 S 的分形维数值(简称分维值) D_s 。

一方面由于考虑到图谱在录入的过程中不可避

免地带有一定的噪声;另一方面由于自然图像只是近似满足分形特性,因此,为了消除噪声对分形维数的影响及能较好地满足分形性质,可在算法的步骤(2)中通过旋转、平移细胞在图像中的位置来不断改变细胞重心的位置;然后,以细胞多次移动后重心的均值作为图像新的中心。这样在求细胞的分维值时,就不仅可以有效地满足算法对平移不变性的要求,而且通过两次均值化还可以削弱噪声的影响。经实验证明,增加 3 次或 4 次求重心的方法就可以满足这两方面的要求。表 1 列出了用该算法求出图 1 中 8 类细胞的部分演变阶段饱和度分量图像 S 的分维值 D_s 。

表 1 不同类型细胞的饱和度分量图像 S 的分维值
Tab. 1 The fractal dimension values of eight types of cell derived from S images

D_s							
嗜碱粒细胞	嗜酸粒细胞	淋巴细胞	单核细胞	分叶核粒细胞	杆状核粒细胞	浆细胞	红细胞
2.3945	2.3876	2.1178	2.0756	2.1941	2.1688	2.0811	2.0940
2.4334	2.4258	2.1022	2.0830	2.1930	2.1906	2.0817	2.0908
2.4599	2.4585	2.0872	2.1387	2.2358	2.2305	2.0481	2.0762
2.4996	2.4757	2.1151	2.0827	2.1789	2.2182	2.0935	2.0938
2.5065	2.5228	2.0823	2.1234	2.1844	2.1857	2.0585	2.1076
2.5732	2.5985	2.0761	2.0876	2.2462	2.2352	2.0541	2.0872

从表 1 的分维值 D_s 可以看出,嗜碱粒细胞与嗜酸粒细胞的 D_s 值较接近,它们的均值分别为 2.4779、2.4781,且总体上均值大于其他几类细胞的均值;分叶核粒细胞和杆状核粒细胞的 D_s 的均值分别为 2.2054、2.2048,总体上这类细胞的均值居中;单核细胞、淋巴细胞、浆细胞及红细胞的分维值接近,它们的均值小于其他几类细胞。这表明,嗜碱粒细胞和嗜酸粒细胞的灰度表面最为粗糙;分叶核粒细胞和杆状核粒细胞灰度表面的粗糙度较小,而单核细胞、淋巴细胞、浆细胞及红细胞则十分细腻。

由于不同纹理图像可能会有相同的分维值,仅靠单一灰度图像的分维值还不能得到足够的信息用来区分纹理图像;文献[8]中指出,细胞图像的纹理不仅表现在图像结构上具有某种统计意义的相似性和重复性,还反映在彩色分布方面,因此,为了提高细胞分类的准确性,本文一方面全面考虑了图谱中前述 8 类细胞所有不同演变阶段的图像,另一方面

则根据骨髓细胞真彩色图像的颜色信息来提取特征分量,并对提出的分量进行分形分析。

8 类骨髓细胞的颜色特征如表 2 所示,从中可以看出,由于不同类别真彩色骨髓细胞图像不仅存

表 2 8 类骨髓细胞的颜色特征(瑞氏染色)

Tab.2 The different color of eight types of marrow cells(wright)

骨髓细胞种类	颜色特征
嗜碱粒细胞	浆为粉红色带有蓝黑色颗粒,核为深紫色
嗜酸粒细胞	粉红色浆带有桔黄色及少量深蓝色颗粒,深紫色核
淋巴细胞	浆呈淡蓝色带有少量的紫红色颗粒,核为紫色
单核细胞	浆呈灰蓝色,核为紫红色
分叶核粒细胞	浆呈粉红色带有少量深蓝色颗粒,核为深紫色
杆状核粒细胞	浆呈粉红色带有少量深蓝色颗粒,核为深紫色
红细胞	浆呈深蓝色,核为紫红色
浆细胞	在不同的发展阶段,浆的颜色由淡粉色变为深蓝色

在不同的颜色信息,而且存在共有的颜色特征,因此,可以从其 RGB 图像中提取出每类细胞所特有的或所有细胞共有的颜色信息来进行分析。例如,由于蓝色(Blue)是这 8 类细胞所共有的颜色信息,所以对蓝分量 B 应该加倍重视;另外,本文提出了一

个新的彩色识别分量 $C = B/(R + G + B)$ 。显然,由于彩色分量 C 与蓝色分量 B 是密切相关的,并且是一个归一化值,因此说,彩色分量 C 的引入是恰当且有效的。表 3 列出了表 1 中所用到的细胞彩色分量图像 C 的分维值 D_c 。

表 3 不同类型细胞的彩色分量图像 C 的分维值

Tab. 3 The fractal dimension values of eight types of marrow cells derived from C images

D_c							
淋巴细胞	单核细胞	嗜碱粒细胞	红细胞	浆细胞	嗜酸粒细胞	分叶核粒细胞	杆状核粒细胞
2.026 1	2.007 2	2.053 0	2.000 4	2.010 5	2.153 5	2.032 4	2.026 3
2.029 6	2.017 9	2.054 8	2.000 6	2.019 2	2.160 1	2.035 7	2.030 8
2.029 3	2.018 9	2.062 3	2.001 5	2.025 5	2.171 3	2.040 0	2.036 7
2.020 7	2.022 6	2.065 8	2.004 6	2.023 0	2.073 2	2.042 5	2.042 1
2.024 0	2.026 3	2.074 1	2.011 2	2.014 1	2.180 9	2.048 3	2.044 9
2.015 5	2.039 6	2.082 4	2.015 4	2.035 1	2.212 5	2.051 7	2.048 9

对表 3 中的 D_c 值与表 2 中的 D_s 值进行比较发现,同一类型细胞的 D_c 值普遍比该类型细胞的 D_s 值偏小,表明这 8 类细胞的彩色分量图像 C 表面的粗糙度都有所降低。由于分形参数的差异,细胞本质上是由细胞所含染色颗粒的大小和分布特性决定的,所以从 D_c 值可以看出,在这 8 类细胞中,蓝色分量 B 占整个细胞的比例相对较均匀,而彩色分量图像 C 的细胞表面则较光滑。

4 细胞之间类别可分性实验分析

为考察 D_s 、 D_c 对每两类细胞之间类别的可分性,可用如下判据公式^[9]:

$$G_n = \frac{(m_{i,n} - m_{j,n})^2}{(\delta_{i,n}^2 + \delta_{j,n}^2)} \quad n = 1, 2, \dots \quad (6)$$

表 4 各类细胞的饱和度分量图像 S 和彩色分量图像 C 的分维值统计量

Tab. 4 The mean and standard deviation of D_s and D_c

统计量	嗜碱粒细胞	嗜酸粒细胞	淋巴细胞	单核细胞	分叶核粒细胞	杆状核粒细胞	红细胞	浆细胞
D_s 均值	2.3579	2.3681	2.0968	2.0985	2.2054	2.2048	2.0916	2.0695
D_s 标准差	0.067	0.063	0.080	0.063	0.087	0.093	0.066	0.078
D_c 均值	2.0560	2.1410	2.0242	2.0233	2.0454	2.0388	2.0089	2.0212
D_c 标准差	0.016	0.052	0.009	0.010	0.008	0.008	0.007	0.010

该判据值越大,表明类别可分性越好。其中,下标 n 表示参数类型,这里 n 最大为 2,只对两个参数类型 D_s 和 D_c 进行可分性分析,设 G_1 为 D_s 的判据值, G_2 为 D_c 的判据值;下标 i, j 表示模式类别; m, δ 分别表示某类模式对于某参数的均值和标准偏差。

实验中用到的细胞图像共 232 幅,其中,22 幅不同演变阶段的红细胞图像、35 幅不同演变阶段的嗜碱粒细胞图像、39 幅不同演变阶段的嗜酸粒细胞图像、43 幅不同演变阶段的单核细胞图像、38 幅不同演变阶段的淋巴细胞图像、25 幅不同演变阶段的嗜中性粒细胞图像及 30 幅不同演变阶段的浆细胞图像。表 4 分别统计出了图谱中这 8 类细胞所有演变阶段图像的 D_s 、 D_c 均值和标准差。由表 4 数据估算的这 8 类细胞之间的可分性判据值见表 5。

表 5 不同类别细胞之间可分性判据值
Tab. 5 The discrimination values between eight types of marrow cells

细胞类别	嗜碱粒细胞	嗜酸粒细胞	分叶核粒细胞	杆状核粒细胞	淋巴细胞	单核细胞	红细胞	浆细胞
嗜碱粒细胞	0	0.0122	1.9201	1.7768	6.2300	7.9053	7.9574	7.8269
嗜酸粒细胞	2.4409	0	2.2918	2.1113	7.0898	9.1420	9.1551	8.8581
分叶核粒细胞	0.3511	3.3018	0	0.0197	0.8443	0.9904	1.0848	1.3527
杆状核粒细胞	0.9245	3.7734	0.3403	0	0.7751	0.8955	0.9843	1.2425
淋巴细胞	3.0007	4.8985	0.1614	1.4701	0	0.0432	0.0325	0.0597
单核细胞	3.0036	4.9405	0.1742	1.4649	0.0045	0	0.0557	0.0837
原红细胞	7.5973	6.3688	13.3225	8.9401	2.0008	1.5427	0	0.0586
浆细胞	3.4018	5.1184	3.5710	1.8888	0.0497	0.0200	1.1124	0

注:表中,右上角为参数 D_s 的判据值,左下角为参数 D_c 的判据值。

由于自然景象只在一定程度上满足自相似性,因此,只要参数统计分布具有类别可分性,就可作为特征值用于分类识别。表 5 的实验数据 G_n 表明,综合考虑参数 D_s 、 D_c 就可对这 8 类细胞中的部分细胞进行分类识别。例如,划分嗜碱粒细胞与其他类别细胞时,只要某细胞参数 D_s 的判据值 G_1 小于 1 就可以排除该细胞属于分叶核粒细胞、杆状核粒细胞、淋巴细胞、单核细胞、红细胞和浆细胞类型,但是,却不能分辨出是嗜酸粒细胞,还是嗜碱粒细胞。如果结合 D_c 的判据值 G_2 ,则从表 5 中可以看出,只要 D_s 的判据值 G_1 小于 1, D_c 的判据值 G_2 大于 1,那么就可以将嗜碱粒细胞从嗜酸粒细胞中区分出来了。

5 结 论

本文针对骨髓细胞图像的特点,采用差分盒子计数法和求取图像重心的算法来提取《临床血液学及细胞学图谱》中不同类别细胞图像的分形特征。实验结果表明,由于不同类型骨髓细胞图像的分形参数具有不同程度的差异,通过对彩色图像中的饱和度分量图像 S 和彩色分量图像 C 的类别可分性判据值进行的分析可知,由于分形参数对于某些类别的骨髓细胞具有较好的区分能力,因此,在对骨髓细胞显微图像进行分类识别时,可将分形参数作为识别特征量。

参 考 文 献 (References)

1 Mandelbrot B B. The fractal geometry of nature[M]. San Francisco,

CA, USA: Freeman, 1982:1~83,109~279.

2 ZHANG Hou-qiang, WANG Fu-quan. Fractal geometry and application in molecular science[M]. Beijing: Science Press, 1993: 1~213. [李后强,汪富泉.分形理论及其在分子科学中的应用[M].北京:科学出版社,1993:1~213.]

3 The Cytology Group of Shangyang Medical College. Clinical hematology and cytology[M]. Beijing: People's Medical Publishing House, 1967. [沈阳医学院细胞图谱编绘小组.临床血液学及细胞学图谱[M].北京:人民卫生出版社,1967.]

4 ZHANG Ji-zhong. Fractal geometry [M]. Beijing: Tsinghua University Press, 1995. [张济忠.分形[M].北京:清华大学出版社,1995.]

5 Sarkar N, Chaudhuri B B. An efficient approach to estimate fractal dimension of textural images [J]. Pattern Recognition, 1992, 25(9):10~35.

6 Pentlend P. Fractal-based description of naural scenes [J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 1984, 6(6): 661~674.

7 WU Zhao-cong. Determination of fractal scaleless range [J]. Acta Geodaetica et Cartographica Sinica, 2002, 31(3):240~244. [巫兆聪.分形分析中的无标度区确定问题[J].测绘学报,2002, 31(3):240~244.]

8 Chang C C, Wang L L. Color texture segmentation for clothing in a computer-aided fashion design system [J]. Image Vision Computer, 1996, 14(9):685~702.

9 ZHANG Yong, BO Zi-you, ZHANG Qiang. Regional fractal property of leukocyte image [J]. Journal of Xi'an Jiaotong University, 1999, 33(10):60~63. [张勇,柏子游,张强.白细胞显微图像的区域分形特征分析[J].西安交通大学学报,1999, 33(10): 60~63.]