

红血球图像表面曲率的计算方法

王瑞胡¹, Brendan McCane², 尹章群³

(1. 重庆文理学院数学与计算机科学系, 重庆 402160; 2. 奥塔戈大学计算机科学系, 但尼丁, 新西兰; 3. 武汉大学医学院, 武汉 430060)

摘 要: 提出一种基于阴影恢复技术的细胞图像表面各点的曲率计算方法, 利用边界轮廓提取技术分离由扫描电镜 SEM 得到的红血球亮度图像, 对单个细胞从跟踪得到的中心点开始进行区域生长, 其每个像素点的亮度变化信息决定了与之相对应的高度场深度, 阴影恢复技术可以将这种变化的亮度转化为不同的深度, 从而得到像素的三维数据点。通过最小二乘法对三维点进行曲面拟合得到分段连续光滑的细胞表面, 其形状特征由平均曲率与高斯曲率来表示。结果表明, 该方法具有很强的实用性和应用价值。

关键词: 红血球; 阴影恢复; 轮廓跟踪; 区域生长; 曲率

Method for Curvature Calculation of Red Blood Cell Surface

WANG Rui-hu¹, Brendan McCane², YIN Zhang-qun³

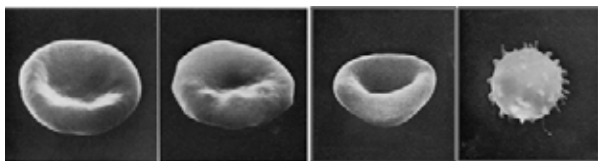
(1. Department of Mathematics and Computer Science, Chongqing University of Arts and Science, Chongqing 402160; 2. Department of Computer Science, University of Otago, Dunedin, New Zealand; 3. Division of Medical Science, Wuhan University, Wuhan 430060)

【Abstract】 This paper proposes an approach to compute curvature by means of Shape Form Shading(SFS) technique. Initially it traces the contour of SEM RBCs image, and grows regionally from center point of each cell. The height field depth is derived from the intensity variation with SFS, which is also the 3D points related to the corresponding pixel. As curvature is defined on smooth curved surface, the piece-wise smooth surface can be obtained using least squares method and surface fitting. The shape is featured by mean curvature and Gaussian curvature. Results show that it is promising and practicable.

【Key words】 red blood cell; Shape From Shading(SFS); contour tracing; region growing; curvature

1 概述

红血球对维持人体正常生理功能起着重要作用, 它向身体的各个组织细胞输送氧气, 同时将细胞代谢产生的二氧化碳运回肺部。这些功能能否正常运行很大程度上依赖于红血球的形变。临床诊断中大量实践经验表明很多疾病都与红血球的形变能力有关。图 1(a)、图 1(b)为 2 种正常的红血球形状, 呈圆盘状中间略有凹陷(双凹盘状); 图 1(c)中的红血球细胞形状常见于遗传性与后天溶血性疾病当中; 图 1(d)是红血球的另一种变形(棘红细胞)。



(a)正常红血球 1 (b)正常红血球 2 (c)疾病红血球 (d)变形红血球

图 1 4 种典型的红血球形状

分析红血球形变能力的一个主要途径是计算 SEM 图像中各类细胞的形状分布。由于红血球本身形状的多变性, 常用的分类方法并不适用于这类特殊对象, 文献[1]的红细胞形态特征统计分布算法^[1]将正常红细胞形态特征与病态相比, 能够在一定程度上用于红细胞形态变化相关疾病的辅助诊断, 但细胞和组织形态变化的复杂性, 仅用面积、周长、等效半径、形状因子等指标是不够的, 不足以反映细胞本身的形变特征。需要将红血球的形变特征与其形状结合起来考虑。文献[2]提出一种主要用于距离图像(range image)的基于曲率符

号的表面分割与重构算法, 直接利用距离图像像素点所表示的图像信息计算曲率, 再进行曲面拟合。本文是基于亮度图像的分割与重构, 其实现过程正好相反。

基于微分几何特征的细胞表面曲率计算按照先后关系分为如下 5 个步骤:

- (1)轮廓跟踪;
- (2)基于中心点与边界轮廓的区域生长;
- (3)利用阴影恢复重建细胞表面形状高度场;
- (4)利用最小二乘法拟合二次曲面;
- (5)计算高斯曲率与平均曲率。

2 细胞轮廓跟踪^[3]

通过扫描电镜对血涂样片进行 600 倍放大得到的图像如图 2 所示。可以发现, 图像中存在大量的相互重叠的细胞, 对于此类重叠物体分割的一种常见方法是分水岭分割^[4]。虽然分水岭方法能得到每个重叠细胞的基本轮廓, 但是其形状很不规则, 不利于后阶段的区域生长和形状重建。由于最终感兴趣的不是每一幅图像中所包含的细胞个数, 而是不同红血球的形状分布, 因此可作如下合理简化假设: 所有重叠细胞在整个分布中具有一致性, 即等概率分布。处理时只需检测和识别叠加细胞图像中处在最顶层的细胞。采用基于边界

作者简介: 王瑞胡(1975 -), 男, 讲师、博士研究生, 主研方向: 数字图像处理, 模式识别与信息智能; Brendan McCane, 高级讲师、博士; 尹章群, 护士

收稿日期: 2007-12-22

E-mail: wrh0319@163.com

的轮廓跟踪方法可以较好地得到边缘细节，如图 3 所示。

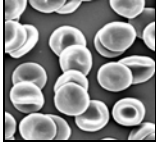


图 2 红血球数字图像



图 3 单个细胞

由图 2 可以发现细胞轮廓类似于椭圆，这一点可以作为先验知识运用到轮廓跟踪当中，通过最小二乘法对跟踪得到的所有像素点(型值点)进行二阶多项式参数拟合得到二次曲线方程

$$\begin{cases} x(t) = at^2 + bt + c \\ y(t) = et^2 + ft + g \end{cases} \quad (1)$$

为了增加精确度并减少计算量，仅仅考虑哪些与预测方向的夹角在一定范围之内的跟踪方向。用单位向量的集合 $u_i \in U$ 来表示：

$$U = \{(\cos(\alpha_i + \beta), \sin(\alpha_i + \beta)) : \beta_{\min} \leq \beta \leq \beta_{\max}\} \quad (2)$$

其中， α_i 为预测方向切向量角度； β_i 为在 $\beta_{\min} = -\frac{\pi}{2}$ 和 $\beta_{\max} = \frac{\pi}{2}$ 之间均匀取值。

在图 3 中总共有 7 个处于顶层的细胞的轮廓被检测出。所有的轮廓信息作为属性存储在 XML 文件中：

```
//每 1 条轮廓的中心点坐标及包含的像素点个数
<contour center_x="37.3884" center_y="175.931" size="182">
//轮廓上的每个像素点的 x, y 坐标
<point x="30.8831" y="203.469" />
...
```

3 区域生长

根据提取出的细胞轮廓信息，以各个轮廓的中心点为起始点采用区域生长算法可得到各个细胞的完整图像，其算法执行流程如下：(1)依次从 XML 文件中提取单个细胞轮廓，得到中心点和轮廓点位置信息；(2)以中心点为种子，按照 4 连通邻域向外生长直到边界轮廓点，并记录每个细胞像素点信息(位置和亮度)；(3)继续计算下一条轮廓所包含的细胞像素点，直到提取出所有的轮廓包含的各个细胞图像。

在算法具体实现过程中，利用队列来保存生长过程中的种子点，主要用到入队、判空和出队 3 种操作。另外还用到 2 幅工作图像，分别定义为布尔型和整型，前者用于存储是否入队标志，后者用于存储提取出的每一个像素点的灰度值。

经过区域生长后得到的整个细胞图像如图 4 所示。

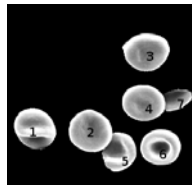


图 4 细胞区域生长

详细信息如表 1 所示。

表 1 通过区域生长分离出的细胞信息

No	中心点 x 坐标	中心点 y 坐标	所有像素点个数
1	37.388 4	175.931 0	2 697
2	114.131 0	179.526 0	2 927
3	191.346 0	65.334 8	2 777
4	187.606 0	138.718 0	2 516
5	154.363 0	209.667 0	2 027
6	211.253 0	202.634 0	2 359
7	233.308 0	134.118 0	953

4 阴影恢复

物体表面某一点的亮度与该点的梯度、光源位置和表面反射特性有关。图像中某一个像素点的灰度级别取决于光源方向和物体表面方向的法向量。这种关系可以用图像亮度方程 IRE 表示为

$$E(x, y) = R(p, q) \quad (3)$$

其中， E 表示二维图像中某一个像素点 (x, y) 处的灰度或亮度； R 为反射函数。

常用的 2 种反射模型为 Lambertian 模型和 Specular 模型，考虑到本文中涉及到的细胞表面特性，在此采用第 1 种模型，其反射函数^[5]为

$$E(x, y) = R(p, q) = \frac{1 + pp_s + qq_s}{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}} \quad (4)$$

通过扫描电镜得到的细胞图像非常适用于用 SFS 来解决。SEM 中物体表面发射的电子的数量是和光照方向向量与表面法向量夹角的 sec 函数成比例^[6]：

$$R(p, q) = \frac{1}{n \cdot l} = \frac{\sqrt{1 + p^2 + q^2} \sqrt{1 + p_s^2 + q_s^2}}{1 + pp_s + qq_s} \quad (5)$$

其中， $n = \frac{(-p_s, -q_s, 1)}{\sqrt{p_s^2 + q_s^2 + 1}}$ 为单位法向量； l 为光源方向的单位向量。

利用笛卡尔坐标系下的深度映射函数 $Z(x, y)$ 对梯度向量 (p, q) 进行线性逼近^[7]：

$$p = \frac{\partial Z}{\partial x} = Z(x, y) - Z(x-1, y), \quad q = \frac{\partial Z}{\partial y} = Z(x, y) - Z(x, y-1) \quad (6)$$

由式(5)、式(6)，将式(4)重写为

$$0 = f(E(x, y), Z(x, y), Z(x-1, y), Z(x, y-1)) = E(x, y) - R(Z(x, y) - Z(x-1, y), Z(x, y) - Z(x, y-1)) \quad (7)$$

将式(7) Taylor 展开，利用 Jacobi 迭代方法对其进行化简合并得到的第 n 次迭代结果 $Z^n(x, y) = Z^n(x, y)$ 的深度映射可以表示为

$$Z^n(x, y) = Z^{n-1}(x, y) + \frac{-f(Z^{n-1}(x, y))}{\frac{d}{dZ(x, y)} f(Z^{n-1}(x, y))} \quad (8)$$

由式(5)和式(8)有：

$$\begin{aligned} \frac{d}{dZ(x, y)} f(Z^{n-1}(x, y)) = & -\left(\frac{\partial R}{\partial p} + \frac{\partial R}{\partial q}\right) = \\ & -1 \cdot \frac{((1 + pp_s + qq_s)(p + q) - (1 + p^2 + q^2)(p_s + q_s)) \cdot \sqrt{1 + p_s^2 + q_s^2}}{(1 + pp_s + qq_s)^2 \cdot \sqrt{1 + p^2 + q^2}} \end{aligned} \quad (9)$$

重写式(9)得到：

$$Z^n(x, y) = Z^{n-1}(x, y) + \frac{f(Z^{n-1}(x, y)) \cdot (1 + pp_s + qq_s)^2 \cdot \sqrt{1 + p^2 + q^2}}{((1 + pp_s + qq_s)(p + q) - (1 + p^2 + q^2)(p_s + q_s)) \cdot \sqrt{1 + p_s^2 + q_s^2}} \quad (10)$$

利用线性逼近方法重建细胞高度场如图 5 所示。



图 5 细胞形状的三维重建

5 曲面拟合与曲率计算

由于曲率是定义在连续且二阶可导的曲面上，离散的三维数据点不能直接计算曲率。在利用阴影恢复技术得到的三维数据点基础上首先进行拟合，得到分段光滑的二次曲面，再利用二次曲面的系数计算高斯曲率和平均曲率。其详细计算过程如下：

- (1)从深度映射(depth map)数据 XML 文件中提取每一个细胞图像包含的像素点信息；
- (2)找与该点相对应的所有邻域点；
- (3)通过最小二乘法利用这些邻域点进行二次曲面拟合，解线性方程组得到的曲面方程的系数直接用于计算曲率；
- (4)分别计算高斯曲率和平均曲率的均值与方差。

主要代码如下：

```
depth_data dd(argv[1]); //argv[1]参数存储深度映射信息
for (int i = 0; i < dd.depths.size(); i++) {
    for (int j = 0; j < dd.depths.at(i).size(); j++) {
        vc.at(i).points.at(j).find_neighbors();
        //利用最小二乘法构造线性方程组的参数矩阵
        vc.at(i).points.at(j).setAb();
        //高斯消去法解线性方程组
        vc.at(i).points.at(j).gauss_elimination();
        //计算高斯曲率与平均曲率
        vc.at(i).points.at(j).compute_curvatures();
    }
    vc.at(i).get_statistics(); //计算统计量
}
```

本文选择基函数集合

$$\{\psi_0=1, \psi_1=x, \psi_2=y, \psi_3=xy, \psi_4=x^2, \psi_5=y^2\}$$

来拟合二次曲面，其方程可以表示为基函数的线性组合：

$$\hat{z} = \hat{f}(x, y) = \Phi_a(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2 \quad (11)$$

采用最小二乘法最小化逼近误差

$$E(a_0, a_1, \dots, a_n) = \sum_{i=0}^N (f(x_i, y_i) - \hat{f}(x_i, y_i))^2 = \sum_{i=0}^N (f(x_i, y_i) - (a_{00}\psi_0(x_i, y_i) + a_{10}\psi_1(x_i, y_i) + \dots + a_{02}\psi_5(x_i, y_i)))^2 \quad (12)$$

E 取最小值的必要条件是 $\frac{\partial E}{\partial a_r} = 0$ ，即

$$-2 \sum_{i=0}^N [f(x_i, y_i) - (a_{00}\psi_0(x_i, y_i) + a_{10}\psi_1(x_i, y_i) + \dots + a_{02}\psi_5(x_i, y_i))] \psi_r(x_i, y_i) = 0 \quad (13)$$

进一步得到：

$$a_{00} \sum_{i=0}^N \psi_0(x_i, y_i) \psi_r(x_i, y_i) + a_{10} \sum_{i=0}^N \psi_1(x_i, y_i) \psi_r(x_i, y_i) + \dots + a_{02} \sum_{i=0}^N \psi_5(x_i, y_i) \psi_r(x_i, y_i) = \sum_{i=0}^N f(x_i, y_i) \psi_r(x_i, y_i) \quad (14)$$

其中， N 为拟合的数据点的个数。

高斯消去法可以有效地解决由最小二乘法得到的系数矩阵构成的线性方程组。解上述线性方程组得到的二次曲面系数 $\{a_i |_{i=0}^5\}$ ，可以直接用来计算高斯曲率 K 和平均曲率 H 。

将式(11)表示为参数形式：

$$S = \{(x, y, \hat{z}) : x=u, y=v, \hat{z}=\hat{f}(u, v), (u, v) \in D \subset R^2\} \quad (15)$$

此时，有

$$K = \frac{LN - M^2}{EG - F^2}, \quad H = \frac{EN - 2FM + GL}{2(EG - F^2)} \quad (16)$$

其中， E, F, G 为曲面方程一次基本形变量(first fundamental form variables)：

$$E = S_u \cdot S_u, F = S_u \cdot S_v, G = S_v \cdot S_v \quad (17)$$

L, M, N 为二次基本形变量：

$$L = n \cdot S_{uu}, M = n \cdot S_{uv}, N = n \cdot S_{vv}, n = \frac{S_u \times S_v}{\|S_u \times S_v\|} \quad (18)$$

最后分别计算高斯曲率和平均曲率的均值与方差。

6 实验结果

利用基于边界的分割方法提取出由扫描电镜 SEM 得到的红血球细胞数字图像的轮廓信息，经实践证明能很好地应用于区域生长及三维高度场重建。三维数据点可进一步用于曲面拟合，得到的分段光滑曲面的系数直接用于计算曲率。本系统在 Ubuntu 7.04 操作系统下用 VXL 语言编程及调试通过，三维高度场数据利用 xd3d 软件包来显示。所有数据点的信息(包括轮廓点及三维高度点)用 XML 文件进行存储与交换。表 2 给出了平均曲率与高斯曲率的统计量信息。

表 2 平均曲率与高斯曲率的均值与方差

No	平均曲率均值	高斯曲率均值	平均曲率方差	高斯曲率方差
1	-0.003 582 35	-0.011 886 700	0.014 446 70	0.011 089 300
2	0.001 909 20	-0.003 196 100	0.000 746 70	0.001 353 110
3	-0.008 010 67	-0.000 136 584	0.007 765 77	8.059 16e-06
4	0.003 624 56	-0.004 921 720	0.001 742 06	0.002 800 740
5	4.1156 3e-05	-0.005 027 110	0.002 811 72	0.006 724 130
6	0.007 744 88	-0.002 439 660	0.007 153 60	0.000 373 712
7	0.002 352 31	-0.013 290 200	0.001 920 86	0.031 150 800

曲率计算结果的准确性主要是和曲面拟合的精确度有关，表 3 给出了 3 种不同窗口大小邻域点的曲面拟合均方无差(RMSE)：

$$\sigma_{S_N}^2(p) = \frac{1}{N} \sum_{(i,j) \in S_N} (z_{ij} - \Phi_a(x, y))^2 \quad (19)$$

其中， $a = (a_{00}, a_{10}, a_{01}, a_{11}, a_{20}, a_{02})$ 。

表 3 不同窗口大小邻域点参与曲面拟合的各个细胞图像均方无差

RMSE	Window Size=3	Window Size=5	Window Size=7
1	1.287 5	6.047 28	11.609 90
2	2.360 6	5.554 21	8.859 61
3	1.685 4	4.736 81	8.378 02
4	2.790 5	5.603 96	8.688 87
5	1.653 4	6.277 08	11.161 70
6	3.041 2	11.374 20	19.754 20
7	2.376 9	6.726 58	10.109 80

在实际算法运行过程中，参与拟合的邻域点的个数并不严格等于窗口尺寸的平方，因为有些像素的邻域点并没有落在当前所处理的细胞图像范围内，此时该邻域点有可能属于背景，也有可能属于其他的细胞图像。可以发现，随着窗口尺寸的增大，参与曲面拟合的数据点越多，均方无差也越大。

7 结束语

本文的工作是整个红血球分类识别系统的前部分，其表面形状的多形变特性需要借助于几何形状特征来描述，得到的高斯曲率和平均曲率统计信息将用来作为细胞特征，进而构成特征向量用以设计分类标准，后期将用模式分类技术，如支持向量机等方法实现多类分类。

参考文献

- [1] 郝葆青, 罗教明, 尹光福, 等. 红细胞形态特征统计分布及计算机图像分析方法研究[J]. 生物医学工程学杂志, 2000, 17(4): 429-432. (下转第 34 页)