

基于椭圆 Hough 变换的尿沉渣图像红细胞检测^①

Red Blood Cell Detection in Urinary Sediment Images Based on Ellipse Detecting Hough Transform

黄梓效 郭斯羽 廖建勇 (湖南大学 电气与信息工程学院 湖南 长沙 410082)

摘要: 传统的椭圆 Hough 变换存在运算速度慢, 计算量大, 占用储存空间大等缺点。针对尿沉渣镜检图像, 首先对利用圆环 Hough 变换对全图中的红细胞进行粗定位, 之后针对每个可能存在红细胞的小区域, 利用一个改进的随机椭圆 Hough 变换对图像中的红细胞进行定位识别。该方法由于采用了低维 Hough 参数空间且合理减少了随机椭圆 Hough 变换中的采样次数, 从而同时降低了算法的时间与空间开销, 同时还具有良好的抗噪能力, 提高了检测准确率。在真实图像上的实验结果证明了以上结论。

关键词: Hough 变换 圆环检测 椭圆检测 尿沉渣图像

1 引言

尿沉渣检查是医院中常见的检验项目之一, 对泌尿系统疾病、特别是肾病的诊断具有重要意义。基于显微图像的自动尿沉渣检验成本低, 能有效地减轻医务人员的劳动强度, 提高检验速度, 从而使医生能更专注于病情的分析与诊断。尿沉渣检查中, 红细胞检查是一个重要的指标。通过对大量显微图像的观察, 尿沉渣图像中的红细胞多为椭圆形, 因此, 我们通过 Hough 变换检测椭圆的方法来检测红细胞。

传统的 Hough 变换^[1]在椭圆检测上通常是利用 Hough 变换把图像空间转换到参数空间, 然后在椭圆的参数空间中进行峰值检测得到椭圆参数。由于椭圆有 5 个参数, 在 5 维参数空间进行累加以及峰值检测存在计算量大, 计算复杂, 需要储存空间大等缺点。针对传统 Hough 变换缺陷, 人们提出了很多 Hough 变换的改进方法, Xu 等人^[2]提出了随机 Hough 变换, 与传统的 Hough 相比, 由于采用多到一的映射, 随机 Hough 变换避免了传统 Hough 变换一到多映射的巨大计算量; Robert 等人^[3]提出的三点椭圆检测法, 通过降低随机采样的点数, 从而减少采样的概率, 从而快速检测; 还有利用椭圆的几何性质提出的改进方

法有: Yoo^[4]等人提出先利用几何性质确定椭圆圆心, 之后再利用 Hough 变换来确定其他的 3 个参数。

本文提出了一种红细胞检测的随机 Hough 变换。为提高随机采样的成功率, 结合红细胞大小的先验数据, 利用 Hough 变换在全图检测圆环, 以圆环所含的边缘点作为椭圆检测的候选边缘点, 然后在每一组这样的候选点上进行随机 Hough 椭圆检测。由于提高了采样成功率, 椭圆检测的速度和精度都得到明显提高。另外, 在随机 Hough 椭圆检测中, 使用 5 个一维累加器来记录 Hough 投票值, 以降低空间开销。

2 Hough 变换原理

Hough 变换是一种通过图像空间中的点向对偶的参数空间进行投票、并根据投票数的局部极大值所对应的参数来进行参数化曲线与形状检测的方法。我们以直线检测 Hough 变换为例来简要说明其原理。熟知直线可通过方程 $y=kx+b$ 来描述。对于一幅给定的图像(假定为二值图像, 其中白色像素表示特征点, 黑色像素表示背景), 我们希望找出其中的直线, 即找到这么一些直线的参数对 (k,b) , 它们所对应的直线上具有足够多的特征点。假设 (x_0,y_0) 为图像中某特征点,

^① 收稿时间:2009-01-02

则直线方程可重写为 $kx_0+b-y_0=0$, 如果考虑所有可能的直线参数 (k,b) 所构成的“参数空间”(即 R^2 平面), 则上方程描述了参数空间中的一条直线, 在该直线上的每个参数点 (k,b) 所对应的图像平面上的直线都经过点 (x_0,y_0) , 即它们至少得到了一个特征点的“支持”或“投票”。对参数空间进行量化后, 可以对每个参数点设置一个“累加器”用来记录该参数点得到的“投票数”, 即有多少个特征点处于该参数点所定义的直线上。然后遍历所有图像中的特征点, 每个特征点均进行这一投票过程。当投票结束后, 参数空间中的投票数将呈现一定的分布: 有的参数点获得的投票数很多, 而有的很少甚至没有。最后, 选取参数空间中具有足够多投票数的那些局部极大点, 它们对应的参数点就定义了若干经过了足够多特征点的图像中的直线, 从而完成了直线检测。

以上方法可推广到任意参数化的曲线或形状。假定某个曲线/区域对应的参数化方程/不等式为 $F(x,a)=0/F(x,a)>0$, 其中 x 为 n 维图像中的坐标, a 为参数矢量, 则对每个 n 维图像中的特征点 x_0 , 都可得到一个定义了参数空间中的曲线/区域的方程/不等式。按同样方式对这些参数空间中的点进行投票后, 就可通过参数空间中的局部极大点检测来得到图像中真实存在的曲线/区域所对应的参数, 从而完成检测。

3 检测方法

3.1 图像预处理

图像预处理过程的主要目的是获得边缘图像以进行后续的 Hough 变换。根据所获得的图像质量, 首先是对灰度图像进行中值滤波, 然后使用阈值化方法得到二值图像, 并填充孔洞, 最后由二值图像得到边缘。在其它应用中, 如果图像质量不适合直接使用阈值化处理^[5], 也可利用边缘检测算子如 Sobel 算子和 Canny 算子等, 进行二值边缘的提取。

3.2 随机 Hough 变换

二次曲线的一般性方程为:

$$x^2+Bxy+Cy^2+Dx+Ey+F=0 \quad (1)$$

其中, (x,y) 为图像空间坐标, 而参数组 (B,C,D,E,F) 则唯一地确定了一条特定的二次曲线。令

$$\Delta = \frac{1}{2} \begin{bmatrix} 2 & B & D \\ B & 2C & E \\ D & E & 2F \end{bmatrix}, J = \frac{1}{2} \begin{bmatrix} 2 & B \\ B & 2C \end{bmatrix}, I = 1+C$$

当 $\Delta \neq 0, J > 0$ 并且 $\Delta/I < 0$ 时, 二次曲线为椭圆^[6]。随机 Hough 变换的通常做法是在边缘图像上任取 5 点, 代入(1)式求出 5 个参数, 再判断是否是椭圆参数。若满足约束条件, 则这 5 个点就能确定一个椭圆, 所确定的椭圆的 5 个参数圆心坐标 (x_0, y_0) , 长轴 a , 短轴 b , 偏转角 ϕ 分别如下:

$$x_0 = \frac{2CD - BE}{B^2 - 4C}, y_0 = \frac{2E - BD}{B^2 - 4C} \quad (2)$$

$$\phi = \frac{1}{2} \arctan \frac{C-1}{B} \quad (3)$$

$$M = \sqrt{1 + \frac{B^2}{(1-C)^2}}, N = \frac{2(E^2 + CD^2 + FB^2 - BDF - 4CF)}{B^2 - 4C} \quad (4)$$

$$a = \sqrt{\frac{N}{(C-1)M - (C+1)}}, b = \sqrt{\frac{N}{(1-C)M - (C+1)}}$$

得到椭圆参数后, 需要在参数空间中进行投票累加。为了降低算法的复杂度, 我们采用类似于[7]中的方法, 使用 5 个一维数组对这 5 个参数分别进行累加。每次采样并成功得到一组椭圆参数后, 分别对这 5 个参数值所对应的累加器单元加 1, 即这些参数点多获得了 1 票。整个采样-投票操作进行预设的次数 T_s , 在投票完成后, 找到各个累加器中具有足够大(即获得票数超过一个给定阈值 T_v)的局部极值, 这些局部极值对应的参数值进行组合后可得到一组或多组椭圆参数, 这些参数组即确定了一个或多个可能的椭圆。

对这些检测得到的候选椭圆, 需要根据它们实际包含的边缘点数来确认是否是真实椭圆。对于实际的数字图像, 由于噪声的存在, 不能要求边缘点严格地落在椭圆上, 而只能给定一个误差限 T_e , 如果某边缘点到椭圆的距离不超过 T_e , 则认为该点落在了椭圆上。如果某候选椭圆上包含的边缘点数不小于一个给定的阈值 T_p , 则认为该候选椭圆为一个检测得到的真实椭圆。

3.3 性能分析

由于噪声点及其非细胞上的点的影响, 传统随机 Hough 变换在检测的过程中随机采样的成功率往往比降低, 假设一副图像中只有有一个红细胞, 红细胞的点数为 m , 还有 n 个不是在红细胞边缘上的点, 则随机采样 5 个点落在同一红细胞上的概率为:

$$P = \frac{C_m^5}{C_{m+n}^5} = \frac{m \times \dots \times (m-4)}{(m+n) \times \dots \times (m+n-4)} \quad (5)$$

当 $n \gg 1$ 时, $P \approx (m/n)^5$, 可见, 随着噪声点数量的增加, 采样成功率将迅速降低。因此, 要提高算法的性能, 关键在于在进行随机 Hough 椭圆检测时, 将采样边缘点集中的噪声点数量尽量减小。

Guo 等^[8]在 Hough 圆检测中采用了滑动窗口来减少噪声点数量, 提高了效率。本文也采用类似的思路来改善随机 Hough 椭圆检测。但是滑动窗口本身包含的范围较大, 因此仍包括了较多的噪声点。我们考虑通过圆环区域来进一步将参与随机 Hough 椭圆检测的噪声点排除。一个椭圆所占据的空间范围可通过其长、短轴长度 a 与 b 加以表示, 所有该椭圆上的点均处在一个圆心在椭圆圆心、内径为 b 而外径为 a 的圆环上, 因此, 如果有足够多的边缘点处于某个椭圆上, 则必然有至少同样多的点处于对应的圆环上。而对于成像条件确定的尿沉渣图像, 红细胞的长短轴长度落在一个确定的范围内, 则最大的长轴长 A 与最小的短轴长 B 可以先验地确定。由此, 我们首先利用传统 Hough 变换进行内径为 B 外径为 A 的圆环检测, 然后对每个圆环, 以落在该圆环内的边缘点作为采样点集, 进行随机椭圆检测。由于圆环的待确定参数只有其圆心位置, 对应了一个 2 维 Hough 变换, 因此时间与空间开销都较小; 而这个额外的开销带来的则是其后椭圆检测的过程由于采样成功率提高而带来的时间与精度总体性能的明显提高, 因此是值得的。

4 算法步骤

本文算法的具体步骤如下:

- ① 对原始图像进行预处理, 以获得二值边缘图像;
- ② 根据在训练图像集上得到的红细胞长短轴范围, 以最大长轴长度 A 和最小短轴长度 B 为外径和内径, 按照给定的边缘点数阈值 T_p , 在边缘图像全图利用传统 Hough 变换检测圆环, 找出各个圆环范围内的边缘点集。
- ③ 对每个圆环所包含的点集 Q , 如下进行随机 Hough 椭圆检测;
- ④ 令采样次数 $s=0$; 初始化 5 个 1 维椭圆参数投票累加器数组, 并对所有累加器单元清 0;

⑤ 从 Q 中随机取 5 个点, 代入①式后求解方程组, 如果方程组有解则求出参数, 并验证参数能构成椭圆, 至⑥; 否则 $s=s+1$, 至⑦;

⑥ 由解出的参数进一步求得椭圆的参数, 并将其离散化后, 对对应的投票累加器单元投票; $s=s+1$;

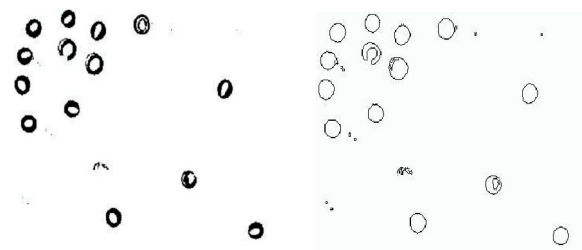
⑦ 如果采样次数 s 达到采样次数阈值 T_s , 至⑧; 否则至④;

⑧ 根据投票数阈值 T_v , 在 5 个 1 维累加器数组中找出所有值不小于该阈值的局部极值对应的参数值, 并根据这些参数值的所有可能组合, 得到一个或多个候选椭圆参数;

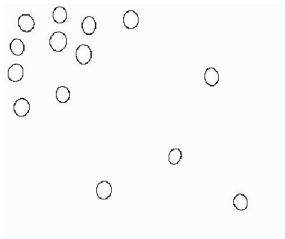
⑨ 对每个候选椭圆参数, 根据给定的距离误差阈值 T_e , 统计全图中落在该椭圆上的边缘点个数 n , 如果 n 不小于给定的点数阈值 T_p , 则认为该候选椭圆参数对应了一个真实椭圆; 如果存在若干类似的椭圆参数(即圆心位置接近, 且长短轴长度和朝向也近似相同), 则在这些椭圆中找到 n 值最大的一个作为检测结果; 如果 n 也相同, 则再考虑椭圆长短轴的差值, 在本文中认为是差值越小, 认为该椭圆就是所求的。

5 实验结果

为进行验证, 在 MATLAB6.5 中实现了算法, 并在一批实际的尿沉渣图像上进行了实验。图像大小均为 384×288 。作为对比, 还实现了直接在全图利用所使用的随机 Hough 椭圆检测方法类似于^[8]中的带滑动窗口扫描随机 Hough 椭圆检测。实验平台为 PC 机, CPU 为 AMD4300+, 内存 1G, 操作系统为 Windows XP。为节省版面, 这里给出其中一张图片上的结果。实验结果见图 1, 其中图 1(a)为原始图像, 1(b)为经预处理后得到边缘图像, 1(c)为检测得到的椭圆。



(a)原始图像的二值图像 (b)经预处理得到的边缘图像



(c)检测结果

图 1 实验结果

由于红细胞呈现出的是同心的椭圆，且大小比白细胞要大，在对图像预处理中对细胞填充后及其边缘轮廓的提取(如图 1 b)，更加能够方便准确的检测出红细胞的位置，之后运用本文改进后的方法进行检测得到结果(如图 1 c)。结果说明，该方法能准确有效的检测出红细胞，且对噪声不敏感。若直接在整幅图像做椭圆 Hough 变换，很多红细胞都检测不出来；在用 30×30 窗口扫描进行全图范围进行 5 个一维椭圆 Hough 变换检测，虽然能很好检测出红细胞，但是 Hough 采样次数过多，耗时太久，表 1 列举出 3 种方法的分别在不同 T_s 值(T_s 为某一个范围内进行的边缘采样次数)下对应的全图 Hough 采样平均次数 H (在每个固定的 T_s 值下重复 20 次后的均值)以及红细胞数量检出率(R_D)。

表 1 三种不同方法下的结果比较

T_s	直接随机 Hough 变换		带滑动窗口随机 Hough 变换		改进后的 Hough 变换	
	H	R_D	H	R_D	H	R_D
100	10,000	0%	130,123	96.28%	28,599	98.84%
200	15,000	0%	252,390	98.76%	42,813	99.31%
400	40,000	4%	481,752	100%	114,259	100%
800	60,000	6%	964,376	100%	228,512	100%

较之传统的 Hough 变换，本文最大的特点是降低了随机 Hough 变换采样次数以及计算量，从而在满足合理的检测率的前提下，提高了运算速度，这些优势使本算法在椭圆检测方面具有一定的使用价值。

6 结论

红细胞是尿沉渣的重要组成部分，对医生诊断具有及其重要的作用，在 Hough 变换广泛应用于计算机视觉领域的同时，由于传统的 Hough 变换检测椭圆计算量很大，而很难在图像中检测到椭圆，又由于医学图像的特殊性，要求必须对噪声不敏感，而且又要准确率高，通过实验可以证明，采用改进的椭圆 Hough 变换的算法，可以大大降低 Hough 变换检测椭圆的计算量，且对噪声的敏感度低，所以本算法在医学细胞图像检测还是一种有意义的方法。

参考文献

- Gonzalez RC, Woods RE. Digital Image Processing. 2nd ed. New Jersey: Prentice Hall, 2002.
- Xu L, Oja E. A New Curve Detection Method: Randomized Hough Transform(RHT). Pattern Recognition Letters, 1990,11(5):331-338.
- Robert AM. Randomized Hough transform: Improved ellipse detection with comparison. Pattern Recognition Letters, 1998,19:199-305.
- Yoo JH, Sethi IK. An ellipse detection method from the polar and pole definition of conics. Pattern Recognition, 1993,26(2):307-315.
- 艾红梅,王昌富.尿液分析的现状与展望.国外医学:临床生物化学与检验学分册, 2005,26(4):256.
- MathWorld W. Ellipse. 1999. <http://mathworld.wolfram.com/Ellipse.html>.
- Goneid A, El-Gindi S, Sewisy A. A method for the Hough transform detection of circle and ellipse using a 1-dimensional array. Proc. 1997 IEEE International Conference on Systems, Man and Cybernetics 4:3154-3157.
- Guo SY, Zhang XF, Zhang F. Adaptive randomized Hough transform for circle detection using moving window. Proc. Fifth International Conference on Machine Learning and Cybernetics. 2006:3880-3885.