

# 基于图像的酵母菌出芽率分析方法<sup>①</sup>

吕丹凤, 杨翠荣, 庞全

(杭州电子科技大学 自动化学院, 杭州 310018)

**摘要:** 出芽率是反映酵母菌生长的重要质量标准之一, 目前对于酵母菌的出芽率的统计分析主要通过技术人员人工显微镜观察和统计, 准确性和效率受到影响, 不能适应工业生产的需求。提出一种基于凹点检测的椭圆拟合方法, 根据酵母菌通常近似椭圆的特点, 通过对酵母菌进行凹点检测之后再行椭圆拟合的方法, 分别统计出图像中酵母菌总数和出芽菌体数, 进而判断出酵母菌的出芽率。实验表明, 该方法可以准确有效地检测出酵母菌的出芽率。

**关键词:** 椭圆检测; 酵母菌; 曲线拟合; 凹点检测;

## Analysis of Yeast Budding Rate Based on Images

LV Dan-Feng, YANG Cui-Rong, PANG Quan

(Institute of Automation, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** The budding rate was one of the standard to reflect the growth of yeast, but currently the analysis of the yeast growth were mainly through technical and statistical manual microscopy, cannot meet the needs of industrial production for accuracy and efficiency are affected. This paper presented a approach of ellipse fitting method based on concave points detection, according the characteristic that Yeast was usually approximate ellipsoid, by fitted ellipse after found concave points, counted the total number and the budding number of the yeast in the image respectively, then determined the rate of budding yeast. Experimental results show that the approach can detect budding yeast detection rate accurately and effectively.

**Key words:** elliptic detection; yeast; curve fitting; concave detection

在酒精生产中, 为满足发酵大生产的工艺要求, 应保证酵母细胞数目和质量。成熟酒母质量的重要标准之一就是酵母菌<sup>[1]</sup>的出芽率。出芽率过低时( $\leq 20\% \sim 30\%$ )应及时采取措施, 以提高出芽率。目前对酵母菌生长情况的分析主要是技术人员人工统计和识别, 通过人工在显微镜下的观察得出检测结果。但是长时间用肉眼观察易造成疲倦, 影响检测的准确性和效率; 而且由于检验检疫、质量检测等部门每天待检样品繁多, 靠人工肉眼检测, 效率和准确性都比较低, 不能满足社会实际需求。随着机器视觉技术的发展, 利用图像处理对酵母菌出芽率进行统计是一个新的可行方法。

分析酵母菌的形态特征可知, 酵母菌呈椭圆及近似椭圆状, 我们提出了以下方案进行酵母菌出芽率的统计。1) 采集酵母菌当前生长状况的图像信息, 2) 利用图像处理技术在图像中对椭圆进行拟合与统计, 3) 逐个分析椭圆信息判断是否为芽体, 4) 根据以下公式即可得到出芽率。

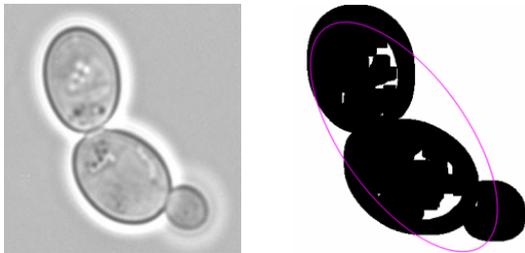
$$\text{出芽率} = \frac{\text{芽体数}}{\text{酵母菌总数}} \times 100\% \quad (1)$$

根据酵母菌的形状特点, 采用椭圆拟合的方法实现酵母菌出芽率的统计。椭圆拟合是指对平面上的一组样本点, 寻找一个尽可能靠近这些样本点的椭圆。

然而在实际的图像中, 由于拍摄角度以及酵母本

<sup>①</sup> 收稿时间:2011-04-19;收到修改稿时间:2011-06-27

身的问题，经常会出现多个菌体重叠在一起形成了一个较大的区域，如果直接采用椭圆拟合的方法，将会产生严重的误拟合（如图 1 所示），这将直接影响菌体的计数以及菌落信息的分析结果。



(a) 原图 (b) 椭圆拟合图

图 1 椭圆拟合效果图

针对这种情况，文献[1]中提出了几种算法，如基于分离点搜寻的分离算法、流域分割算法、水坝分割法等。由于其运算量较大，对计算机软硬件要求也比较高。本文提出了一种新的解决方案，即根据菌体轮廓的凹凸性，首先从重叠区域寻找到凹点，进而对聚合的菌体进行分离，最后再进行椭圆拟合分析。该算法具有实现简单，运算量小，对计算机软硬件的要求较低等特点。实验结果显示，对出芽率的分析结果相当准确，具有较高的使用价值。

### 1 凹区域轮廓的提取

凹点检测的基本思想是分析凹图形上点的几何关系，从轮廓上寻找凸集点，将这些点连接成凸闭包图形，并由凸闭包和凹图形的差值获得凹区，然后根据欧几里得距离公式求出凹点。

#### 1.1 重叠区域的判定

凹凸性<sup>[2]</sup>是图像的基本特征之一，当图像存在两个像素点连线穿过区域外部的图像时，称为凹图形；反之称为凸图形。在图像的重叠区域存在凹陷部分，使图像呈现凹凸的特征。由于凹点必然位于凹区域的轮廓上，因此提取出凹区域，就可以搜索出凹点。

区域重叠与否可根据形状因子来判断。形状因子公式定义如下：

$$PE = \frac{4\pi a^2}{c^2} \quad (2)$$

其中  $a$  是目标的面积，可通过扫描图像累加同一标记的区域中像素的个数来表示。 $c$  是目标的周长，

可用相邻边缘点的距离之和表示。形状因子的取值范围为  $0 < PE \leq 1$ 。细胞重叠时，其边界由于出现凹陷而变得复杂。在相同面积的情况下，存在凹陷情况的目标比没有凹陷的目标的周长更大，导致形状因子相应地变小。确定一个阈值  $P_0$ ，当  $PE \leq P_0$ ，认为目标存在细胞重叠。

#### 2.1 凸封闭包的计算

将所提取得到的重叠区域<sup>[3]</sup>边缘点的坐标序列<sup>[4]</sup>记为  $P_i = (x_i, y_i)$  ( $i=1,2,3,\dots,n$ ,  $n$  为边缘点的个数)。在  $P_i$  中求  $x, y$  的最大值和最小值，记

$$\begin{cases} P_{LU} = (x_{\min}, y_{\max}) \\ P_{LD} = (x_{\min}, y_{\min}) \\ P_{RU} = (x_{\max}, y_{\max}) \\ P_{RD} = (x_{\max}, y_{\min}) \end{cases} \sqrt{a^2 + b^2} \quad (3)$$

所讨论的凹图形必定落在以  $P_{LU}, P_{LD}, P_{RU}, P_{RD}$  为顶点的矩形之内，该矩形为凹图形的  $f(P_1, P_2, P_3) > 0$  矩形。下图为凹图形的最小矩形：

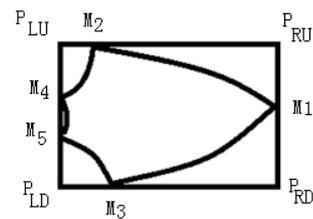


图 2 凹凸性边缘点的划分

凹图形边缘点的划分：

从最小矩形边上寻找一些特殊点，对凹图形边缘再进行划分，以便缩小凸集点的搜索范围：

$X$  坐标等于  $P_{RD}$  中  $y$  值最大的点记为  $M_1$ ， $y$  最小的点记为  $M_5$ ；其余点分别记为  $M_2, M_3, M_4$ 。如图 2 所示：

位于  $M_1M_5$  之间的点必为凸集包上的点，下面考虑其余三点之间的凸集包求法。这 3 个点列的拓扑结构是一样的，只需要讨论一个。

1) 若  $M_1M_2$  之间没有点，则线段  $M_1M_2$  为凸封闭包边界上的一条边，否则转入 2)

2) 假设  $M_1M_2$  之间有  $n$  个点，则这些点中能够成为凸封闭包边界上的点必落入由  $M_1, M_2$  及  $P_{LU}$  三点组成的三角形区域中。

3) 建一个栈存储凸集点。顺序提取  $n$  个点，判断为凸集点则压入堆栈。取一点  $q_1$ ，若  $y_{q_1} \leq y_{M_1}$  或

$x_{q_i} < x_{M_1}$ ，则  $q_i$  一定不会落入三角形区域中，故不可能为凸集点。

4) 顺序判断  $n$  个点，直至找到不满足上述条件的点  $q_i$ ，将  $q_i$  压入堆栈。取  $q_i$  的下一个点  $q_{i+1}$ ，分析  $q_i$  和  $q_{i+1}$  的位置关系：**a)** 当  $x_{q_i} = x_{q_{i+1}}$  时，若  $y_{q_i} > y_{q_{i+1}}$ ，则  $q_{i+1}$  不能成为凸集点，删除该点。若  $y_{q_i} < y_{q_{i+1}}$ ，则  $q_i$  不可能为凸集点，用  $q_{i+1}$  替代  $q_i$ 。**b)** 当  $x_{q_i} < x_{q_{i+1}} < x_{M_2}$  时，若  $x_{q_{i+1}}$  位于  $Mx_{q_i}$  左侧时， $q_i$  不可能为凸集点，以  $x_{q_{i+1}}$  代替  $x_{q_i}$ ，将  $x_{q_{i+1}}$  压入堆栈。若  $x_{q_{i+1}}$  位于  $Mx_{q_i}$  右侧时，只有当  $y_{q_i} < y_{q_{i+1}}$  时  $x_{q_{i+1}}$  为凸集点，压入堆栈。

判断一个点对于线段的位置可以采用如下的方法：设  $p_i = (x_i, y_i)$  ( $i=1,2,3$ )，定义函数

$$f(p_1, p_2, p_3) = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \quad (4)$$

该函数具有如下的性质：

$f(p_1, p_2, p_3) = (x_3 - x_2)(y_1 - y_2) - (x_1 - x_2)(y_3 - y_2)$ 。当  $f(p_1, p_2, p_3) = 0$  时，三点共线；当时， $p_2$  位于  $p_1, p_3$  右侧；当  $f(p_1, p_2, p_3) < 0$  时， $p_2$  位于  $p_1, p_3$  左侧。若  $f(M_1, q_i, q_{i+1}) > 0$  说明  $q_{i+1}$  位于线段  $Mx_{q_i}$  左侧或者之上；若  $f(M_1, q_i, q_{i+1}) < 0$  说明  $q_{i+1}$  位于线段  $Mx_{q_i}$  右侧。**c)** 当  $x_{q_i} > x_{q_{i+1}}$  时，若  $f(M_1, q_i, q_{i+1}) > 0$  说明  $q_{i+1}$  位于线段  $Mx_{q_i}$  左侧， $q_i$  不是凸集点，以  $q_{i+1}$  代替  $q_i$ ，压入堆栈。若  $f(M_1, q_i, q_{i+1}) < 0$  说明  $q_{i+1}$  位于线段  $Mx_{q_i}$  右侧，不可能为凸集点。

4. 对其余的子列按照上述处理，当所有点判断完毕，从栈内讲凸集点取出，依序链接成凸闭包。

### 1.3 凹点检测的步骤

凹点检测的步骤如下：

1) 采用八邻域的链码跟踪法提取图像中的重叠区域。

2) 根据凸封闭包的判定方法，计算重叠区域的凸封闭包。

3) 凸封闭包减去原重叠图，得到凹区域的轮廓。利用欧几里德距离公式：

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5)$$

找到凹区域中距离最小的两点即为凹点。由于细胞有串联和并联的情况，因此分别讨论：

细胞串联的情况：定  $(x_1, y_1)$  是第一个凹区域轮廓上的某点， $(x_2, y_2)$  是第二个凹区域轮廓上的

某点，直接计算它们之间的欧式距离，则  $D$  取值最小的情况下所对应的  $(x_1, y_1)$  和  $(x_2, y_2)$  为凹点。

2) 细胞并联的情况：将重叠区域的核心相加除以核心的个数，求出区域的中心。然后仿照步骤 1) 计算凹区域轮廓上每个点与中心间的欧几里德距离。其中使距离  $D$  最小的点对即为凹点。

检测出凹点之后，重叠区域的轮廓(C)可以分成  $N$  段<sup>[5]</sup>：

$$C = cp_1p_2 + cp_2p_3 \dots + cp_s p_{s+1} \quad (6)$$

其中  $cp_s p_{s+1}$  指凹点  $p_s, p_{s+1}$  之间的轮廓段。

## 2 椭圆拟合

### 2.1 椭圆拟合原理

椭圆的一般方程<sup>[6]</sup>为：

$$F(a, x) = ax^2 + bxy + xy^2 + dx + ey + f = 0 \quad (7)$$

并且：

$$4ac - b^2 < 0 \quad (8)$$

在理想的情况<sup>[7]</sup>下，将点  $(x_i, y_i)$  代入(7)， $F(a_i, x_i)$  的值为 0，但是由于误差的存在， $F(a_i, x_i)$  非零。将非零  $F(a_i, x_i)$  值称为点  $(x_i, y_i)$  到椭圆  $F(a, x) = 0$  的距离。(7)式等价于下面的向量形式：

$$F(a, x) = a \cdot x = 0 \quad (9)$$

其中  $a = [a, b, c, d, e, f]^T$ ，分别表示椭圆的参数向量和系数向量。参数向量可以通过使点到椭圆的代数距离的平方和<sup>[8]</sup>：

$$D_A(a) = \sum_{i=1}^N F(a, x_i)^2 \quad (10)$$

上述(10)式可以用向量的形式表示：

$$E = \|DK\|^2 \quad (11)$$

其中矩阵  $D$ ：

$$D = \begin{bmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i^2 & x_i y_i & y_i^2 & x_i & y_i & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_m y_m & y_m^2 & x_m & y_m & 1 \end{bmatrix}$$

由于椭圆参数不受定标因子影响，将不等式(8)简化为约束等式：

$$4ac - b^2 = 1 \quad (12)$$

这是一个二次约束式，可用矩阵形式表示：

$$a^T C a = 1 \quad (13)$$

其中矩阵 C 表示为:

$$C = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

参数向量 a 可以通过联立广义特征系统:

$$\begin{cases} D^T Da = \lambda Ca \\ a^T Ca = 1 \end{cases} \quad (14)$$

该广义特征系统只有一个正的特征值,该特征 对应的特征向量即所求椭圆的参数向量。

### 2.2 基于凹点的椭圆拟合

拟合步骤如下:

- 1) 遍历凹区域的轮廓,选择一对凹点( $p_s, p_{s+1}$ ),将两凹点之间的轮廓段  $p_s p_{s+1}$  保存并进行椭圆拟合。
- 2) 删除保存的轮廓段  $p_s p_{s+1}$ ,继续寻找下一对凹点,重复第一步,直到所有凹区域的轮廓都被拟合。
- 3) 求拟合椭圆的几何中心( $x_c, y_c$ ) [9]:

$$\begin{cases} x_c = (BE - 2CD) / (4AC - B^2) \\ y_c = (BD - 2AE) / (4AC - B^2) \end{cases} \quad (15)$$

求长轴 a 和短轴 b:

$$\begin{cases} a = \sqrt{\frac{2(Ax_c^2 + Cy_c^2 + Bx_c y_c - 1)}{A + C - \sqrt{(A-C)^2 + B^2}}} \\ b = \sqrt{\frac{2(Ax_c^2 + Cy_c^2 + Bx_c y_c - 1)}{A + C + \sqrt{(A-C)^2 + B^2}}} \end{cases} \quad (16)$$

通过以下步骤判断椭圆的出芽情况:

由公式:

$$\begin{cases} dx = dx_1 - dx_2 \\ dy = dy_1 - dy_2 \end{cases} \quad (17)$$

计算椭圆中心间的距离:  $d = (dx, dy)$ 。利用公式:

$$\begin{cases} a_{x_{ij}} = a_i \cos \theta_i + a_j \cos \theta_j \\ a_{y_{ij}} = a_i \sin \theta_i + a_j \sin \theta_j \end{cases} \quad (18)$$

计算长轴水平方向和垂直方向的分量( $a_x, a_y$ ), 将 ( $a_x, a_y$ ) 分别与椭圆中心间距离( $dx, dy$ ) 比较: 若  $a_x > dx$  或者  $a_y > dy$ , 可以断定两椭圆相交。

- 3) 利用椭圆的面积公式:

$$s = \pi * a * b \quad (19)$$

对于两相交的椭圆进行面积判定,若椭圆 ( $S_1$ ) 面积小于椭圆( $S_2$ )面积的 1/2, 可以判定  $S_1$  为  $S_2$  的芽体, 否则认为两椭圆粘连。

### 3 实验结果与分析

为了验证算法的有效性,进行了大量的实验。如图 4-1 所示的酵母菌存在出芽的情况并且粘连堆积较为明显。首先通过凹点检测,提取图像中的凹点,然后基于凹点进行椭圆拟合。将图片分为三部分以方便说明,图 3-2, 3-3, 3-4 展示了凹点检测和基于凹点检测的椭圆拟合的实验效果, 表格 1 记录了椭圆的中心及长短轴数据, 表格 2 记录了椭圆中心间距离及长轴

分量。分析图 4-2 可得到:  $\begin{cases} a_{y_{12}} < dx_{12} \\ a_{x_{33}} > dx_{23} \\ a_{y_{12}} > dy_{12} \\ a_{y_{23}} < dy_{23} \end{cases}$ , 所以三个椭圆两两相交, 又由于  $\begin{cases} S_1 = 22969 > 0.5S_2 \\ S_3 = 6908 < 0.5S_2 \end{cases}$ , 椭圆 1 与椭圆 2 是粘连的关系, 椭圆 3 是椭圆 2 的芽体。对于图 4-3: 由  $\begin{cases} ax_{12} > dx \\ ay_{12} > dy \end{cases}$  得出两个椭圆相交, 又, 所以椭圆 1 是椭圆 2 的芽体。分析图 4-4: 因为  $\begin{cases} a_{x_{33}} > dx_{13} \\ a_{x_{23}} > dx_{23} \\ a_{y_{33}} < dy_{13} \\ a_{y_{23}} < dy_{23} \end{cases}$ ,  $\begin{cases} a_{x_{34}} > dx_{34} \\ a_{y_{34}} > dy_{34} \end{cases}$ ,  $\begin{cases} a_{x_{34}} < dx_{24} \\ a_{y_{24}} > dy_{24} \end{cases}$  并且只有  $S_1 < 0.5S_3$ , 所以判定四个椭圆两两粘连, 椭圆 1 为椭圆 3 的芽体。

对图 4-1 经过上述分析后, 得到酵母菌的总数为 9, 而芽体数为 2, 根据公式(1)得到出芽率为: 22%。图 4-1 所示酵母菌具有良好的发酵能力, 属于成熟酒母。

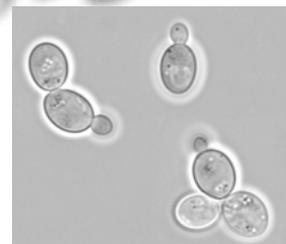
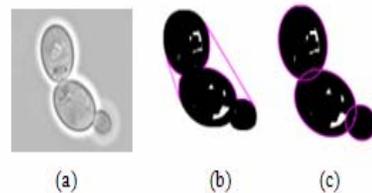
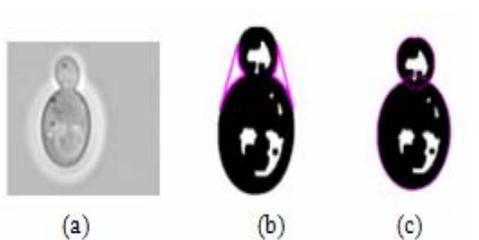


图 3-1 待检测图



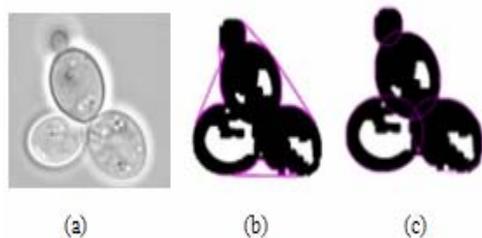
(a)原图 (b) 凹点检测图 (c)椭圆拟合图

图 3-2 实验效果图 1



(a) 原图 (b) 凹点检测图 (c) 椭圆拟合图

图 3-3 实验效果图 2



(a) 原图 (b) 凹点检测图 (c) 椭圆拟合图

图 3-4 实验效果图 3

#### 4 结论

针对酵母菌存在菌体粘连及芽体未与母体分离情况，本文提出了一种基于图像的酵母菌出芽率分析的新算法，利用菌体为椭圆形特征，以边缘和凹点为信息对酵母菌逐一识别。该方法根据粘连程度与芽体分离情况进行判定，能够准确有效的计算出酵母菌出芽率，相对于传统的人工识别方法有明显的改善。

表 1 拟合所得到椭圆的中心和长短轴

椭圆数据	倾斜角	中心	短轴	长轴	面积
4-2-1(上)	72°	(357,91)	77	95	22969
4-2-2(中)	62°	(243,140)	79	103	25550
4-2-3(下)	12°	(171,292)	44	50	6908

4-3-1(上)	356°	(126,125)	57	78	13960
4-3-2(下)	7°	(129,223)	30	36	3391
4-4-1(最小)	34°	(106,319)	32	36	3617
4-4-2(右下)	63°	(235,121)	64	84	16880
4-4-3(中间)	23°	(148,230)	70	82	18023
4-4-4(左下)	94°	(98,123)	70	85	18683

#### 参考文献

- 熊子书.中国酿酒酵母菌的研究.中国食品发酵工业研究所,2002,112(4):23-27.
- 傅蓉.基于凹点搜寻的重叠细胞图像自动分离的算法研究.南方医科大学,2007,43(17):21-28.
- 求是科技.Visual C++数字图像处理典型算法及实现.北京:人民邮电出版社,2006.
- 刘润涛.一种简单多边形凸包的新线性算法.工程图学学报,2002,2:120-125.
- Yan L, Park CW. New separation algorithm for touching grain kernels based on contour segments and ellipse fitting. Zhejiang Univ-Sci, 2011,12(1):54-61.
- Fitzgibbon A. Pilu M, Fisher RB. Direct least square fitting of ellipses. IEEE Trans. Pattern Anal. Mach. Intell., 1999,21(5): 476-480.
- Ge P. Research on Segmentation Algorithm of Adhesive Plant Grain Image. 8th Int. Conf. on Electronic Measurement and Instruents. 2007,2: 927-930.
- Carter RM. Digital imaging based classification and authentication of granular food products. Meas. Sci. Technol., 2006, 17(2):235-240.
- 彭青玉.木星土星边缘的椭圆拟合.云南天文台台刊,2003, (4):43-46.

(上接第 117 页)

算机工程,2006,(12):18-21.

- Ko SH. Prediction of Preferences Through Optimizing Users and Reducing Dimension in Collaborative Filtering System. Proceedings of the 17th International Conference on Innovations in Applied Artificial Intelligence. Ottawa, Canada, 2004
- 朱征宇,裴仰军,等.个性化服务中用户近期兴趣视图的生

成.计算机工程与设计,2005,(6):237-249.

- 张伟,廖晓峰,吴中福.一种基于遗传算法的聚类新方法.计算机科学,2002,(6):114-116.
- 陶俊,张宁.基于用户兴趣分类的协同过滤推荐算法.计算机系统应用,2011,20(5).
- 肖慧,王立.Web 日志挖掘中的用户识别算法.计算机系统应用,2011,20(5).