

文章编号 1004-924X(2005)02-0231-06

# 基于遗传算法的生物组织图像最佳挖取点寻优

王朝晖<sup>1</sup>, 李 莉<sup>2</sup>, 李引生<sup>1</sup>, 蒋庄德<sup>1</sup>

- (1. 西安交通大学 机械工程学院, 陕西 西安 710049;
2. 国家微检测系统工程技术研究中心, 陕西 西安 710069)

**摘要:**生物组织微阵列芯片的制备过程中,需要从包埋有生物组织的蜡块中挖取一块直径为0.5~1.5 mm的组织蜡柱。为了实现自动化作业,研究了基于数字图像处理的组织蜡柱最佳挖取点计算机寻优策略。确定了问题的数学描述为在选定组织图像区域中寻找最大内切圆的圆心,提出了基于遗传算法的寻优算法。重点探讨了坐标点编码,参数选取和遗传操作方法等。研究了个体适应度快速计算方法,初始种群选择等算法改进方案。经过实际编程反复数次实验,结果表明迭代60次算法完全收敛,证明其搜索速度和稳定性都比较好,完全满足自动化作业的在线计算要求。

**关键词:**组织微阵列;数字图像;遗传算法

**中图分类号:**TP391.41 **文献标识码:**A

## Seeking of tissue image digging point method based on genetic-arithmetic

WANG Chao-hui<sup>1</sup>, LI Li<sup>2</sup>, LI Yin-sheng<sup>1</sup>, JIANG Zhuang-de<sup>1</sup>

- (1. *Mechanical Engineering School, Xian Jiaotong University, Xian 710049, China;*
2. *National Engineering Research Center for Miniaturized Detection System, Xian 710069, China)*

**Abstract:** For tissue-micro-array manufacture automatization, the fitness point seeking in the wax where the tissue sample to be picked up was studied. Based on digital image of tissue, the mathematical model of seeking was described as seeking the biggest inscribed circle center in the tissue area. The seeking method based on genetic-arithmetic was researched, including the code principle, parameters selection and the implementing method. The arithmetic improvement consists of fitness calculation and initialization selection, etc. The results of theory analysis and experiment prove that this arithmetic is effective and useful to fulfill the on-line manipulation.

**Key words:** bio-tissue micro array; digital image; genetic-arithmetic

## 1 引言

生物芯片中的微阵列(micro array)技术是近年来新兴的分子生物学技术。在制备组织微阵列时,先要用空心针从初始包埋有生物组织的蜡块中,挖取一块直径为 0.5~1.5 mm 的组织蜡柱,然后将蜡柱转移到与之相配合的空白石蜡阵列孔中。逐次将不同病理组织的蜡柱填满空白阵列孔,就构成了组织微阵列<sup>[1-2]</sup>。在组织微阵列制备过程中,要求提取并转移到阵列孔中的组织蜡柱能够充分反应病理特征。因此,在初始蜡块样品中准确选择提取点,是保证最后制成的组织微阵列质量的关键。

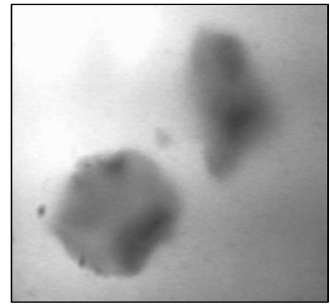
目前国内外组织芯片的制备大都采用手工操作方法,依靠操作人员的经验在初始组织蜡块中确定蜡柱提取点,不仅耗时费力,而且准确率较低,严重制约了组织芯片的发展。在研制自动化制备仪的过程中,为了使机器自动完成组织蜡块的识别定位打孔等工作,引入了计算机数字图像处理技术。这其中关键的一步是计算机如何在线判定最佳的打孔位置。为此将遗传算法与数字图像处理技术相结合,研究了新的搜索算法,实现了计算机自动搜索定位的功能。

## 2 最佳挖取点寻优的数学模型建立

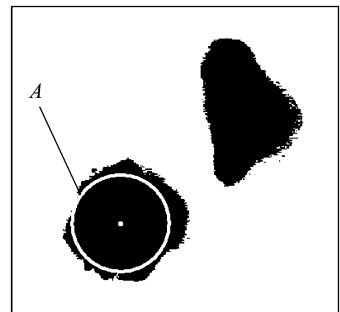
首先进行了问题的数学模型研究。根据 CCD 提取的组织蜡块图像(灰度图),研究其图像特征发现,在初始蜡块的图像中,特征元素可分为前景和背景两类。如图 1(a)是计算机提取的包埋有生物组织的蜡块图像,前景(偏黑色)是生物组织区域,其余的部分都是背景。因此可将图像平滑、降噪预处理后,转换成黑白二值函数图像来研究,如图 1(b)所示。由图中可以直观

地看出,所要选择的挖取点必然要在黑色区域内部。

按照操作要求,蜡柱的挖取点要尽量在组织区域的中心部位。因为中心位置的生物组织质量比较好,而且在深度方向能够保证有足够的挖取量。因为每块组织区域的图形是连通域,定义此区域为图形 A,并确定 A 中最大内切圆的圆心即为最佳挖取点,如图 1(b)所示。需要说明的是,图 1 中显示有两块组织区域,实际操作时可在任意一块上挖取组织蜡柱。



(a)包埋有两块生物组织的蜡块图像  
(a)Image of wax embed two tissue



(b)最佳挖取点示意图  
(b)Sketch map of the best digging point

图 1 生物组织蜡块图像

Fig. 1 The image of tissue wax

设数字图像矩阵中的像素坐标点为  $(x, y)$ ,  $R(x, y)$  表示以点  $(x, y)$  为圆心的区域 A 的内切圆半径,则最佳挖取点寻优的数学模型可以描述为:

$$\begin{cases} \max & R(x, y) \\ \text{s. t.} & (x, y) \in A, \end{cases} \quad (1)$$

至此,在初始组织蜡块中寻找最佳挖取点问题,转化成如何在图像区域  $A$  中寻找最大内切圆圆心的问题。

寻找  $A$  中的最大内切圆圆心,可用全局搜索的方法,比较  $A$  中所有的内切圆,找出最大的半径。也可采用数学形态学中的图形腐蚀方法,选定某一元素多次腐蚀,直至找到最内部的点。但这些搜索方法的计算量过大,影响了自动制备仪的工作效率。在研究中发现采用遗传算法不仅计算速度快,而且准确率很高,完全满足在线使用的要求。

### 3 基于遗传算法的寻优策略

遗传算法是一种借鉴生物界自然选择和进化机制发展起来的高度并行、随机、自适应搜索算法。它借用了群体搜索技术,将种群代表一组问题解,通过对当前种群施加选择、交叉和变异等一系列遗传操作,从而产生新一代的种群,并逐步使种群进化到包含近似最优解的状态<sup>[4-5]</sup>。在研究过程中发现,运用遗传算法来解决蜡柱最佳挖取点寻优问题,不仅可以找到比较满意的近似最优挖取点,而且收敛速度快。

#### 3.1 基本算法

关于遗传算法的原理不再赘述,应用于本文所研究问题的算法描述如下:

(1) 编码:数字图像的构成是二维矩阵,其每个像素点的位置表达是一对整数对。设图像区域  $A$  中的像素点的坐标为  $(x, y)$ ,将其编码为基因表达个体,可直接将其转换为二进制串,串的前后两部分分别代表  $x$  坐标和  $y$  坐标值。串的长度取决于图像矩阵的大小,例如对一个  $800 \times 600$  的图像矩阵来说,因为  $2^9 < 800 \times 600 < 2^{10}$ ,则串的长度可取为 20 位。

(2) 种群:种群的规模定为偶数,初始种群中的个体组成是在图像  $A$  中随机选取。

(3) 适应度:种群中各个体的适应度,取为组织区域中以像素点  $(x, y)$  为圆心的内切圆半径,半径越大,适应度越高。

(4) 遗传操作:遗传操作包括个体选择、交叉、变异。

(5) 终止条件:迭代的终止条件有 2 种。第 1,当种群中的每个个体适应度方差小于给定数值后,就终止新种群的繁殖计算。第 2,规定最大迭代次数,当迭代次数达到要求的最大的迭代次数时,终止迭代。

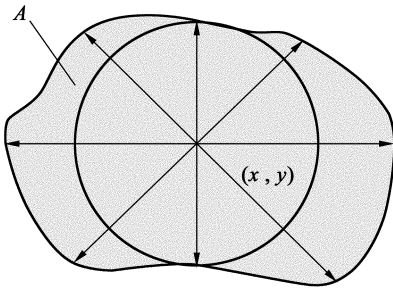
求解结束后采用保留最佳个体的方式确定最佳挖取点。即保留每一代种群适应度最高的个体,待繁殖计算终止后,比较这些局部最佳个体,选出全局最佳个体。

#### 3.2 算法的几点改进

在保证算法有效的前提下,根据生物组织图像的特点,作了如下几点改进<sup>[7-8]</sup>。

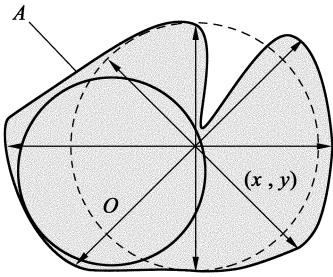
(1) 适应度的计算。组织图像区域  $A$  中以像素点  $(x, y)$  为圆心的内切圆半径的计算,采用计算如图 2(a)所示的,  $A$  中以点  $(x, y)$  为中心的 8 个矢量的长度,最小的长度即为内切圆半径。具体计算时,取二值图像中的组织区域灰度值为 1,背景区域灰度值为 0。在图像中沿着水平、垂直和  $45^\circ$  角斜方向这 8 个方向向外搜索,当灰度值由 1 变为 0 时,说明抵达了目标边缘。比较该点沿 8 个方向到边缘的直线距离,最小距离是该点的内切圆的半径,即是该点的适应度。这样做大大减少了计算量。当然这样有可能产生误判,如图 2(b)所示,虚线圆是误判,实际的最大内切圆应该是实线圆  $O$ 。但图形  $A$  出现这种形状的概率极小,因为图像中的生物组织区域一般呈现凸图形,其外轮廓不可能具如此陡峭的曲线形状。

(2) 初始种群的确定。先搜索出区域



(a) 8 个方向矢量确定内切圆

(a) Eight vector estimation of the inscribed circle



(b) 内切圆产生误判的示意图

(b) Mis-estimation of the inscribed circle

图 2 图形 A 的最大内切圆圆心确定示意图

Fig. 2 Sketch map of the inscribed circle center about area A

A 的 4 个顶点,在此范围内随机产生一定数量的 A 内部的点作为初始种群。这样可以最大程度避免近亲繁殖,不会导致算法早熟。

(3) 变异操作。在开始阶段采用较大的变异率进行,结束阶段减少变异率。这样有利于算法从最优点附近向最优点靠近。

## 4 算法的实现与实验结果

作者将此寻优算法编制了计算机程序并进行了实验验证。实验过程如下:

CCD 采集一幅初始组织蜡块图像如图 3 所示,数字图像尺度为  $800 \times 600$ ,实际测

量的组织区域跨度约 5 mm。图像经过平滑降噪处理后,转换为二值函数图像,确定出需要寻优计算的组织区域。将像素点坐标编码为 20 位的二进制序列。遗传操作中,初始种群是在组织区域内部随机取 20 个像素点坐标。先计算每个坐标点的适应度,用蒙特卡罗法确定个体的选择概率分配,采用轮盘赌选择出适应度高的点参加交叉和变异。一般交叉操作的交叉率取  $0.4 \sim 0.9$ ,变异操作的变异率取  $0.001 \sim 0.1$ 。现据多次实验选定交叉率为 0.8,变异率开始取为 0.05,40 代后取为 0.03。根据交叉率在已经选择的点中随机选取 2 个坐标点,按照线性重组的方法交叉。变异操作中,如果小概率事件发生,该点随机在 8 个自由方向上进行变异。检验遗传操作后产生新个体的坐标点是否在图像 A 的内部。如果不在内部要重新进行交叉和变异,直到产生新的满足条件的点。经过遗传操作后繁殖了新一代的种群,再次计算种群中各个体的适应度,然后进行新一轮的遗传迭代操作,直到满足终止条件,停止迭代运算。在确定算法的终止条件时,经过多次实验发现,种群迭代 30 代基本上就趋于收敛,到 50 代就完全收敛了。为此确定迭代 60 次后计算终止。然后比较 60 个局部最优点,确定全局的最佳挖取点坐标。

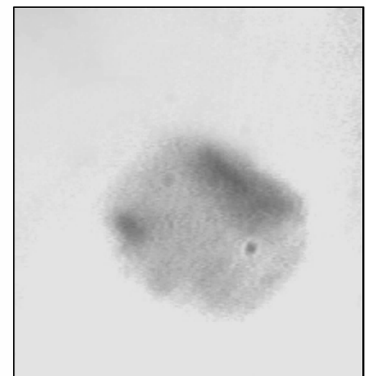


图 3 原始组织图像

Fig. 3 Initial tissue image

经遗传算法寻优后的图像标识如图 4 所示,图中的白色方点即为找到的最佳挖取点,与人工判断的一致。计算机从采集图像开始到标识出挖取点,整个计算过程耗时不到 10 ms。实验证明算法快速准确,达到了在线使用的要求。

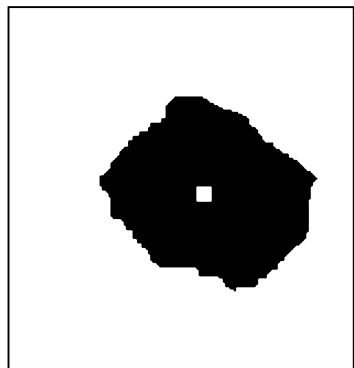


图 4 寻优结束后的图像

Fig. 4 Image after having sought the digging point

图 5 为计算过程中每代平均适应度曲线变化情况。图 6 是每代最大适应度曲线变化情况。从曲线可以看出,在第 20 代时适应度已经抵达最优解 38.2 附近,而后没有大的波动情况出现。

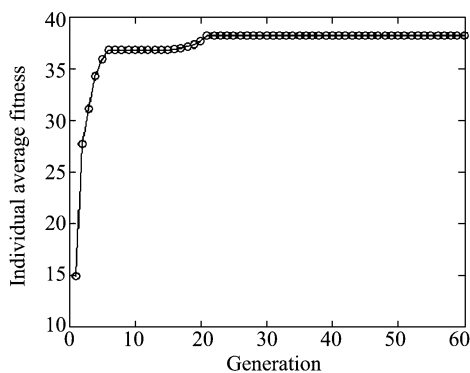


图 5 每代中个体平均适应度变化曲线

Fig. 5 Individual average fitness for each generation

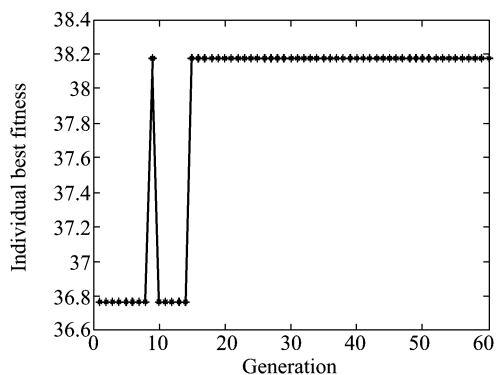


图 6 每代中最优个体适应度变化曲线

Fig. 6 Individual best fitness for each generation

对同一个组织蜡块图像反复数次实验,得到结果是一样的。而后,用同一程序计算不同的组织蜡块图像,效果同样令人满意。说明此算法的搜索速度和稳定性都比较好,完全满足自动制备仪在线计算要求。

## 5 结 论

针对生物组织微阵列自动化制备仪的研制需求,研究了在组织蜡块中确定最佳挖取点的计算机寻优算法。在初始组织蜡柱挖取操作中,引入计算机数字图像处理技术,将组织蜡块图像转换为二值图像,明确问题的数学模型为寻求图像中组织区域的最大内切圆圆心。采用遗传算法进行寻优计算,对图像坐标点采用二进制编码,随机选取初始种群中的个体组成,个体适应度为组织图像中的内切圆半径,遗传操作包括个体选择、交叉、变异等。多次实验结果表明,迭代计算 60 次算法完全收敛,计算耗时小于 10 ms,计算结果与人工判定的一致,证明此算法快速、有效,完全满足生物组织微阵列自动化制备的在线计算要求。

## 参考文献:

- [1] 杨琳,李锦毅,郑华川.组织微阵列/芯片技术及其在肿瘤病理研究中的应用[J].中国实验诊断学,2002,(4):267-269.  
YANG L, LI J Y, ZHENG H CH. Tissue microarray/chip technique and its application in pathology of tumors[J]. *Chinese Journal of Laboratory Diagnosis*, 2002, (4):267-269. (in Chinese)
- [2] 茹晓荣.组织微阵列技术及其应用[J].国外医学·生理、病理科学与临床分册,2004,24(1):104-107.  
RU X R. Tissue micro array technology and application[J]. *Foreign Medical Sciences · Section of Pathophysiology and Clinical Medicine*, 2004,24(1):104-107. (in Chinese)
- [3] 蒋庄德,王朝晖.生物组织微阵列自动制备仪及其控制方法[P].发明专利,专利申请号:200410073180.2.  
JIANG ZH D, WANG CH H. *Tissue microarray auto-manufacturer and control method*[P]. Invent Patent, 200410073180.2. (in Chinese)
- [4] 王小平,曹立明.遗传算法:理论、应用与软件实现[M].西安:西安交通大学出版社,2002.  
WANG X P, CAO L M. *Genetic-arithmetic: theory, application and software implement*[M]. Xi'an: Xi'an Jiaotong University Press, 2002. (in Chinese)
- [5] 李和平,李德华,朱洲,等.基于遗传算法的结构光条纹中心检测方法[J].光学精密工程,2004,12(1):82-87.  
LI H P, LI D H, ZHU ZH, *et al.* Detection of structured light strip center line based on genetic algorithm[J]. *Optics and Precision Engineering*, 2004, 12 (1):82-87. (in Chinese)
- [6] 邢婉丽,程京.生物芯片技术[M].北京:清华大学出版社,2004.  
XING W L, CHENG J. *Biochip technology*[M]. Beijing: Tsinghua University Press, 2004. (in Chinese)
- [7] 章毓晋.图像工程.图像理解与计算机视觉[M].北京:清华大学出版社,1995.  
ZHANG Y J. *Image engineering*[M]. Beijing: Tsinghua University Press,1995. (in Chinese)
- [8] 孙玉芹,车仁生.求解最大内切圆的一种新方法[J].光学精密工程,2003,11(2):181-185.  
SUN Y Q, CHE R SH. Novel method for solving maximum inscribed circle [J]. *Optics and Precision Engineering*, 2003, 11(2):181-185. (in Chinese)
- [9] 蔡戈,黄成伟,李庆祥,等.基于图像识别的动物纤维细度自动测量系统[J].光学精密工程,2002,10(3):239-242.  
CAI Y, HUANG CH W, LI Q X, *et al.* Automatic animal fibre measurement system based on image analysis[J]. *Optics and Precision Engineering*, 2002, 10(3):239-242. (in Chinese)

作者简介:王朝晖(1968—),男,博士,1991年在天津大学获机械制造专业和工业管理专业双学士学位,1998年和2003年在西安交通大学分别获得机电控制专业硕士学位和机械工程博士学位,研究方向为精密仪器和微型机械电子系统。