



(12) 发明专利

(10) 授权公告号 CN 102254109 B

(45) 授权公告日 2013.08.14

(21) 申请号 201110198107.8

(22) 申请日 2011.07.15

(73) 专利权人 西交利物浦大学

地址 215123 江苏省苏州市工业园区独墅湖
高等教育区仁爱路 111 号

(72) 发明人 史玉回

(74) 专利代理机构 苏州创元专利商标事务所有
限公司 32103

代理人 范晴

(51) Int. Cl.

G06F 19/10(2011.01)

G06N 3/00(2006.01)

审查员 金霞

权利要求书2页 说明书8页 附图2页

(54) 发明名称

采用演化算法进行基因芯片图像基因点自动定位方法

(57) 摘要

本发明公开了一种采用演化算法进行基因芯片图像基因点自动定位方法,其特征在于所述方法包括以下步骤:(1)获取基因芯片图像,然后在基因芯片图像中确定基因区网格定位时对整幅图像进行旋转操作的旋转点,采用演化算法找到旋转角度后,先对基因芯片图像进行旋转校正,然后通过图像投影法对基因芯片图像进行基因区网格定位;(2)基因芯片图像基因区网格定位后提取所有基因区,依次采用演化算法进行基因芯片图像的基因区内基因点网格定位并输出。该方法解决了对基因芯片图像基因点进行自动定位的难题。



1. 一种采用演化算法进行基因芯片图像基因点自动定位方法,其特征在于所述方法包括以下步骤:

(1) 获取基因芯片图像,然后在基因芯片图像中确定基因区网格定位时对整幅图像进行旋转操作的旋转点,采用粒子群优化算法找到旋转角度后,先对基因芯片图像进行旋转校正,然后通过图像投影法对基因芯片图像进行基因区网格定位;

(2) 基因芯片图像基因区网格定位后提取所有基因区,依次采用粒子群优化算法进行基因芯片图像的基因区内基因点网格定位并输出;所述粒子群优化算法包括以下步骤:

A) 先初始化一组解;

B) 评估这组解;

C) 对每一解,如解的评估值比记录的最好值好,则将当前的解的评估值记为最好值;如果它的评估值比记录的所有解的最好值还好,则将这一评估值记录为所有解的最好评估值;

D) 依据公式(1)、(2)对每一粒子进行迭代;

$$V_{id}=w*V_{id}+c_1*r_{1d}*(P_{id}-X_{id})+c_2*r_{2d}*(P_{gd}-X_{id}) \quad (1);$$

$$X_{id}=X_{id}+V_{id} \quad (2);$$

E) 如已达最大设置的迭代次数,则终止算法,输出找到的最好解;否则转去第B)步;

其中将第*i*个解记为 $X_i=(x_{i1}, x_{i2}, \dots, x_{in})$,解的飞行速度记为 $V_i=(v_{i1}, v_{i2}, \dots, v_{in})$,它的具有最好评估值的位置记为 $P_i=(p_{i1}, p_{i2}, \dots, p_{in})$,其中*n*为问题的维数,所有解中拥有最好评估值的粒子的下标记为*g*,每个解按照公式(1)和(2)进行飞行,一代一代迭代计算;*w*、*c*₁和*c*₂为常数,*r*_{1d}和*r*_{2d}为(0,1)之间的随机数;*V*_{id}为第*i*个粒子的第*d*个速度分量;*P*_{id}为第*i*个粒子所搜索到的当前最优解的第*d*个分量;*X*_{id}为第*i*个粒子的第*d*个位置分量;*P*_{gd}为整个群体所搜索到的当前最优解的第*d*个位置分量;步骤(1)中采用粒子群优化算法寻找旋转角度,步骤B)中解的评估按照如下步骤进行:

I) 对每一解,依据解所表示的旋转角度对基因芯片图像进行旋转校正;

II) 对旋转校正过的图像进行水平和垂直方向的投影;

III) 对能进行基因区网格定位的图像赋予大的评估值,并且网格间距越大,评估值越大;旋转后不能进行基因区网格定位的图像则赋予较小的评估值;

通过如上步骤评估每一解,采用粒子群优化算法找到旋转角度,然后对基因芯片图像进行旋转校正;步骤(2)中采用粒子群优化算法进行基因芯片图像的基因区内基因点网格定位,步骤B)中解的评估包括如下步骤:

(i) 对基因芯片图像的基因区图像依据解所表示的旋转角度进行旋转校正,然后将旋转校正过后的基因芯片图像的基因区依据已知行列数进行均匀行列分格,每一格的中心暂定为—基因点的中心;

(ii) 对每一基因点的中心依据解所表示的对应于每一基因点*x*和*y*方向值对每一基因点中心作*x*和*y*方向微调;

(iii) 在每一中心依据解所表示的半径作圆形成每一基因点的范围;

(iv) 在所有基因点内包含信息的像素越多和在所有基因点外包含信息的像素越少,则该解的评估值越大,否则越小;

通过上述步骤评估每一解,采用粒子群优化算法寻找到的最佳解将基因区内的基因点

网格定位。

采用演化算法进行基因芯片图像基因点自动定位方法

技术领域

[0001] 本发明属于生物医学图像处理及生物医学仪器技术领域,具体涉及一种采用演化算法进行基因芯片图像基因点自动定位方法。

背景技术

[0002] 基因芯片技术已被广泛地应用于疾病诊断,药物筛选等。基因芯片图像分析是应用基因芯片技术的一个重要环节。一幅基因图像一般有多个基因区子图像组成,每个基因区又有多个基因点组成。一般来说,基因芯片图像分析包含三部分:基因芯片图像预处理、基因芯片图像基因点分割和基因点数据提取与分析。

[0003] 基因芯片图像在获取和/或扫描过程中一般会受到一些噪声污染。基因芯片图像预处理就是采用滤波的方法将这些引入的噪声去掉。常用的滤波方法有中值滤波,均值滤波,和低通滤波等等。不同滤波方法各有其优缺点。

[0004] 基因芯片图像基因点定位一般需先对基因芯片图像进行网格定位。基因芯片图像网格定位是基因芯片图像基因点定位的关键一步。它将基因芯片图像中所有基因点进行网格定位从而可进一步将所有基因点数据提取与分析。网格定位一般分两步。第一步基因区网格定位。第二步对每一基因区中基因点进行网格定位。

[0005] 常用的基因芯片图像基因点网格定位方法可分为三类:手动、半自动、和全自动方法。对于一规则的基因芯片图像,它的基因区及基因点分布一般为已知值且规则分布,因而一般可手工选取基因区的区域,然后在该区域内按已知基因点分布规则地将该基因区内基因点进行网格定位。如有差错,再手工进行一些局部微调。由于基因芯片图像数量巨大并且每一基因芯片图像含有上千数量的基因点。从而手动非常耗时,不实用。

[0006] 常用的自动网格定位法有 Liu Yan, Zhang Yong-de, and Sha Xian-zheng, Automatic Recognition of Microarray Images Using Projection Algorithm, 4th International Conference on Bioinformatics and Biomedical Engineering, 2010, Page(s) :1-4 和胡园园,孙啸等. 基于图像投影的基因芯片图像网格定位. 生物医学工程杂志, 22(4), 2005年, 第668-671页中所述的图像投影法, 邵桂芳, 张倩, 许三耀, 罗林开, 基于启发式原理的基因芯片图像自动网格定位, 2010中国模式识别大会, 2010, 第1-5页提出的基于启发式的自动网格定位, 聂桂军, 王加俊等. 基于爬山法的cDNA基因芯片图像自动划格算法. 南京工业职业技术学院学报, 9(2), 2009年, 第42-46页和 Rueda L, Vidyadharan V. A Hill-Climbing Approach for Automatic Gridding of cDNA Microarray Images. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 3(1), 2006, Pages :72-83 提出的基于爬山法的自动网格定位, Roberth Irata Jr., Juniorb Arrera, Ronaldof Hashimoto, and Daniel Dantas, Microarray Gridding by Mathematical Morphology, 2001 Proceedings of XIV Brazilian Symposium on Computer Graphics and Image Processing, 2001, Page(s) :112-119 提出的基于遗传算法的网格定位算法等等。

[0007] 图像投影法通过基因芯片图像在水平和垂直方向的投影来对基因芯片图像进行网格定位。先针对基因区进行定位,再对每一基因区中的基因点进行网格定位。这种方法适用于规则分布的基因芯片图像。在实际中,基因芯片图像中并不都是规则的。操作及扫描过程形成的基因芯片图像可能不规则。如基因芯片图像中某一基因区可能被旋转了一定的角度。单纯地采用图像投影法不能很好地对基因芯片图像中旋转后的基因区中的基因点进行网格定位。需要对该基因区进行一定角度旋转校正,再采用图像投影法对该基因区中基因点进行网格定位。

[0008] 基于启发式的自动网格定位方法假定基因点行列间分布是等间隔的,没有变形的图像。在此前提下,基因芯片图像网格定位的关键是确定第一条线和线与线之间的间隔距离,然后采用了启发式自动搜索方法来对基因芯片图像进行自动网格定位。这种方法同样不适用于处理不规则的旋转过后的基因芯片图像。

[0009] 对于旋转了的不规则基因芯片图像可采用半自动和自动网格定位。半自动基因芯片图像网格定位一般采用在基因芯片图像上人工确定两点。依据这两点来计算旋转角度并将基因芯片图像进行旋转得到规则的图像。爬山法则是一种局部优化算法,从而基于爬山法的自动网格定位很容易陷入局部最优。基于遗传算法的自动网格定位算法通过对标准模板进行变形来与基因芯片图像进行匹配,并采用遗传算法来自动寻找最佳标准模板变形参数。李兴珊,孙啸,袁春伟,基于遗传算法的基因芯片图像网格定位,生物物理学报,第18卷,第2期,2006年6月,第223-226页中假定了基因芯片图像仅含有单一的基因区,并且假定基因区内所有基因点均匀规则地分布在基因区内。另外采用了二值遗传算法,既变形参数为二进制表示。如旋转角度由7位二进制数表示,位移值由4位二进制数表示,从而算法的精度有限。

发明内容

[0010] 本发明目的在于提供一种采用演化算法进行基因芯片图像基因点自动定位方法,解决了现有技术中难以对基因芯片图像基因点进行自动定位或者定位精度不佳等问题。

[0011] 为了解决现有技术中的这些问题,本发明提供的技术方案是:

[0012] 一种采用演化算法进行基因芯片图像基因点自动定位方法,其特征在于所述方法包括以下步骤:

[0013] (1) 获取基因芯片图像,然后在基因芯片图像中确定基因区网格定位时对整幅图像进行旋转操作的旋转点,采用演化算法找到旋转角度后,先对基因芯片图像进行旋转校正,然后通过图像投影法对基因芯片图像进行基因区网格定位;

[0014] (2) 基因芯片图像基因区网格定位后提取所有基因区,依次采用演化算法进行基因芯片图像的基因区内基因点网格定位并输出。

[0015] 优选的,所述演化算法选自粒子群优化算法、实数值遗传算法、演化规则算法。

[0016] 优选的,所述粒子群优化算法包括以下步骤:

[0017] A) 先初始化一组解;

[0018] B) 评估这组解;

[0019] C) 对每一解,如解的评估值比记录的最好值好,则将当前的解的评估值记为最好值;如果它的评估值比记录的所有解的最好值还好,则将这一评估值记录为所有解的最好

评估值；

[0020] D) 依据公式 (1)、(2) 对每一粒子进行迭代；

$$[0021] \quad v_{id} = w*v_{id} + c_1*r_{1d}(p_{id} - x_{id}) + c_2*r_{2d}(p_{gd} - x_{id}) \quad (1);$$

$$[0022] \quad x_{id} = x_{id} + v_{id} \quad (2);$$

[0023] E) 如已达最大设置的迭代次数, 则终止算法, 输出找到的最好解; 否则转去第 B) 步;

[0024] 其中将第 i 个解记为 $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$, 解的飞行速度记为 $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$, 它的具有最好评估值的位置记为 $P_i = (p_{i1}, p_{i2}, \dots, p_{in})$, 其中 n 为问题的维数, 所有解中拥有最好评估值的粒子的下标记为 g , 每个解按照公式 (1) 和 (2) 进行飞行, 一代一代迭代计算; w , c_1 , 和 c_2 为加速度常数, 如 $w = 0.729$, $c_1 = c_2 = 1.49445$; r_{1d} 和 r_{2d} 为 $(0, 1)$ 之间的随机数。

[0025] 其中 V_{id} 为第 i 个粒子的第 d 个速度分量; P_{id} 为第 i 个粒子所搜索到的当前最优解的第 d 个分量; X_{id} 为第 i 个粒子的第 d 个位置分量; P_{gd} 为整个群体所搜索到的当前最优解的第 d 个位置分量。

[0026] 优选的, 当使用粒子群优化算法进行基因芯片图像基因点自动定位时, 步骤 (1) 中采用粒子群优化算法寻找旋转角度, 其中步骤 B) 解的评估按照如下步骤进行:

[0027] I) 对每一解, 依据解所表示的旋转角度对基因芯片图像进行旋转校正;

[0028] II) 对旋转校正过的图像进行水平和垂直方向的投影;

[0029] III) 对能进行基因区网格定位的图像赋予大的评估值, 并且网格间距越大, 评估值越大; 旋转后不能进行基因区网格定位的图像则赋予较小的评估值;

[0030] 通过如上步骤评估每一解, 采用粒子群优化算法找到旋转角度, 然后对基因芯片图像进行旋转校正。

[0031] 优选的, 当使用粒子群优化算法进行基因芯片图像基因点自动定位时, 步骤 (2) 中采用粒子群优化算法进行基因芯片图像的基因区内基因点网格定位, 其中步骤 B) 解的评估包括如下步骤:

[0032] (i) 对基因芯片图像的基因区图像进行旋转校正, 然后将旋转校正过后的基因芯片图像的基因区依据已知行列数进行均匀行列分格, 每一格的中心暂定为每一基因点的中心;

[0033] (ii) 对每一基因点的中心依据解所表示的对应于每一基因点 x 和 y 方向值对每一基因点中心作 x 和 y 方向微调;

[0034] (iii) 在每一中心依据解所表示的半径作圆形形成每一基因点的范围;

[0035] (iv) 在所有基因点内包含信息的像素越多和在所有基因点外包含信息的像素越少, 则该解的评估值越大, 否则越小;

[0036] 通过上述步骤评估每一解, 采用粒子群优化算法寻找到的最佳解将基因区内的基因点网格定位。

[0037] 本发明优选采用粒子群优化算法能对这种具有不完全规则的基因芯片图像中的基因点进行自动定位, 通过粒子群优化算法在一维空间上获得基因芯片图像的旋转角度, 通过获得的旋转角度对基因芯片图像进行旋转校正; 然后在旋转后的基因芯片图像上进行基因点的网格定位。根据基因芯片图像每个基因区依据的已知行列数进行均匀行列分格,

将每一格的中心暂定为基因点的中心；然后采用演化算法确定基因点的圆半径和基因点的中心的偏移量；通过寻找最优解获得基因区内的基因点定位位置信息。

[0038] 由于在实际应用中，许多扫描获取的基因芯片图像存在一定程度的变形（如旋转，位移），基因区内基因点分布间距也不会完全等距，会有一些的偏差。本发明将基因点的定位问题分成两步：第一步对基因芯片图像中基因区进行定位；第二步分别对每一基因区中的基因点进行定位。每一步都将网格定位问题表示为一个优化问题，然后采用粒子群优化方法确定基因芯片图像中网格定位信息。

[0039] 一般来说，在确定的一段时间内要处理的一批基因芯片图像的基因区及基因点的排列是相同并已知的。基因芯片图像基因点定位方法要确定的是所有基因区及基因区基因点的位置信息。

[0040] 本发明采用的粒子群优化算法属于一种演化计算技术。它是通过模仿一群鸟寻找食物而产生的一种优化算法。它与其它演化计算算法一样，是一种基于一组解的迭代方法，而不像传统优化算法如爬山法那样是基于单一解的迭代算法。在粒子群算法中，每一解又称为一粒子。一组粒子在所要求解问题的解空间中飞行。每一粒子记录它飞行过程中至今为止所飞行轨迹中最好的位置（对优化问题则是至今为止所表示的最好的解）。每一粒子朝着自己最好的位置和所有同伴中最好的位置随机地飞行。统计意义上，这组粒子将会飞向越来越好的区域，从而能在有限的时间内寻找到足够好的解。

[0041] 采用粒子群优化寻找最优解的两个关键因素是：解的表征和解的评估。在基因芯片图像基因点定位问题中就是确定所有基因点的位置信息，所以它的一个解可用所有基因点的位置信息来表征。基因芯片图像中基因点可以通过直角坐标来确定，每个点都有 x 和 y 坐标值。所以粒子群优化得到的每个解由 2 倍基因点数量个实数值组成。粒子（及解）是随机地朝着自己最好的位置和所有同伴中最好的位置飞行，从而需要能评估每一解（粒子）的好坏。当某一解表征的基因点位置为基因芯片图像中真实的基因点位置信息时，该解具有最好的评估值。而某一解表征的基因点位置与基因芯片图像中真实的基因点位置信息完全不符合时，该解具有很差的评估值。粒子群算法的任务就是引导一组初始化时随机产生的解朝着越来越好的区域飞行，最终找到所需要的解，既基因芯片图像的所有基因点的真实位置信息。

[0042] 本发明中每一解（粒子）依据公式 (1) 和 (2) 飞行，既一代一代迭代。公式如下：

$$[0043] \quad v_{id} = w * v_{id} + c_1 * r_{1d} * (p_{id} - x_{id}) + c_2 * r_{2d} * (p_{gd} - x_{id}) \quad (1);$$

$$[0044] \quad x_{id} = x_{id} + v_{id} \quad (2);$$

[0045] 如已达最大设置的迭代次数，则终止算法，输出找到的最好解；否则继续寻找下一解。

[0046] 然而，当基因芯片图像的基因点定位问题可转化为在多维的解空间中寻找一最优解问题（既所有个基因点的定位信息）时，这会带来一高维空间优化问题不适合直接采用粒子群算法求解因为粒子群算法同其它演化算法及传统优化算法一样具有“维数的诅咒”的问题——即问题求解的难度随问题的维数指数增长。因而本发明将这一问题分成两步来解决：第一步对基因区进行优化网格定位，第二步再对每一基因区的基因点进行定位；这样可以减少求解的难度。

[0047] 第一步、基因区网格定位

[0048] 对基因芯片图像进行基因区网格定位可先对基因芯片图像进行旋转校正,再采用图像投影法。将旋转角度作为要寻找的解,采用粒子群算法求解这个一维问题。对每一解的评估采用如下步骤:

[0049] a) 依据解所表示的旋转角度对基因芯片图像进行旋转校正;

[0050] b) 对旋转校正过的图像进行水平和垂直方向的投影;

[0051] c) 对能进行基因区网格定位的图像赋予大的评估值,并且网格间距越大评估值相对来说越大;旋转后不能进行基因区网格定位的图像则赋予较小的评估值;

[0052] 通过粒子群算法找到旋转角度后,先对基因芯片图像进行旋转校正,然后通过图像投影法对图像进行基因区网格定位。

[0053] 第二步、依次对基因区内基因点定位

[0054] 对每一基因区基因点定位依据下列步骤:

[0055] a) 先对基因区图像进行旋转校正;

[0056] b) 再对旋转校正过后的基因区图像依据已知行列数进行均匀行列分格,每一格的中心暂定为每一基因点的中心;

[0057] c) 对每一基因点的中心作些微调,并以该中心作半径 r 的圆,该圆所对应区域则为该基因点的位置及大小。

[0058] 从以上步骤可看出对每一基因区基因点定位需找到以下信息:旋转校正角度大小、每一基因点的中心位置的 x 和 y 方向的微调值、每一基因点的半径。然后采用粒子群优化算法来寻找旋转校正角度大小、每一基因点的中心位置的 x 和 y 方向的微调值、每一基因点的半径的最佳解。每一解的评估依据以下步骤:

[0059] a) 依据每一解所表示的旋转角度对基因区图像进行旋转校正;

[0060] b) 对旋转校正过后的基因区图像依据已知行列数进行均匀行列分格,每一格的中心暂定为每一基因点的中心;

[0061] c) 对每一基因点的中心依据解所表示的对应于每一基因点 x 和 y 方向值对每一基因点中心作 x 和 y 方向微调;

[0062] d) 在每一中心依据解所表示的半径作圆形成每一基因点的范围;

[0063] e) 在所有基因点内包含信息的像素越多和在所有基因点外包含信息的像素越少,则该解的评估值越大,否则则越小。

[0064] 这样通过粒子群算法寻找到的最佳解就将基因区内的基因点定位。

[0065] 相对于现有技术中的方案,本发明的优点是:

[0066] 不同于已有方法,本发明所采用基于粒子群优化的基因芯片图像基因点自动定位方法能对变形的基因芯片图像的基因区及每一基因区内基因点进行自动定位。由于将定位问题分两步走,从而可将图一所示类型的高维(图一所以为 9600 维)优化问题分解为两类低维优化问题(1 维基因区网格定位问题和 48 个相同 202 维基因区内基因点网格定位问题)。因而采用的粒子群优化算法效果更好,解的精确度更高。本发明优化算法可采用其它演化计算算法如实数值遗传算法、演化规则算法等来替代粒子群优化算法。

附图说明

[0067] 下面结合附图及实施例对本发明作进一步描述:

[0068] 图 1 为本发明实施例采用的基因芯片图像；

[0069] 图 2 为本发明实施例采用粒子群算法进行基因芯片图像基因点自动定位方法的流程图。

具体实施方式

[0070] 以下结合具体实施例对上述方案做进一步说明。应理解,这些实施例是用于说明本发明而并不限于限制本发明的范围。实施例中采用的实施条件可以根据具体厂家的条件做进一步调整,未注明的实施条件通常为常规实验中的条件。

[0071] 实施例基因芯片图像的基因点网格定位

[0072] 本实施例采用的基因芯片图像如图 1 所示,它有 4×12 个基因区,每一基因区包含 10×10 个基因点的基因芯片图像。共含有 48 个基因区,每个基因区含有 100 个基因点。

[0073] 如图 2 所示,实施例按照图 2 的流程对基因芯片图像进行处理。首先对所获取的基因芯片图像首先进行预处理滤去噪声污染然后确定图像中一点作为基因区网格定位中对整幅图像进行旋转操作的旋转点。旋转点可选任意基因区内任何一点,这里取图像的中间点。一旦确定旋转点,则固定不变。然后采用粒子群算法对整幅图像进行基因区网格定位。

[0074] 整幅图像基因区网格定位完成后,则可提取出所有基因区。然后可以对每个基因区进行基因区内基因点的网格定位。如图 2 所示,先对第一块基因区采用粒子群算法进行基因点网格定位,然后对第二块,第三块等等,直至完成所有基因区内基因点网格定位,最后输出整幅基因芯片图像所有基因点定位信息。

[0075] 由于在确定的一段时间内要处理的一批基因芯片图像的基因区及基因点的排列是相同并已知的。如图 1 所示,不失一般性,可假定已知基因区排列是 4×12 ,每一基因区基因点排列是 10×10 。基因芯片图像基因点定位方法要确定的是所有基因区及基因区基因点的位置信息。

[0076] 在基因芯片图像基因点定位问题中就是确定所有基因点的位置信息,所以粒子群算法的一个解可用所有基因点的位置信息来表征。对图一所示基因芯片图像,有 4800 个基因点,每个点由 x 和 y 坐标值确定。所以每个解由 $4800 \times 2 = 9600$ 个实数值组成。粒子(及解)是随机地朝着自己最好的位置和所有同伴中最好的位置飞行,从而需要能评估每一解(粒子)的好坏。当某一解表征的基因点位置为基因芯片图像中真实的基因点位置信息时,该解具有最好的评估值。而某一解表征的基因点位置与基因芯片图像中真实的基因点位置信息完全不符合时,该解具有很差的评估值。粒子群算法的任务就是引导一组初始化时随机产生的解朝着越来越好的区域飞行,最终找到所需要的解,既基因芯片图像的所有基因点的真实位置信息。

[0077] 假设将第 i 个解记为 $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$,解的飞行速度记为 $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$,它的具有最好评估值的位置记为 $P_i = (p_{i1}, p_{i2}, \dots, p_{in})$,其中 n 为问题的维数,。所有解中拥有最好评估值的粒子的上标记为 g 。则每个解依据公式 (1) 和 (2) 飞行,既一代一代迭代。粒子群算法流程如下:

[0078] (a) 初始化一组解;

[0079] (b) 评估这组解;

[0080] (c) 对每一解,如解的评估值比记录的最好值好,则将当前的解的评估值记为最好值;如果它的评估值比记录的所有解的最好值还好,则将这一评估值记录为所有解的最好评估值;

[0081] (d) 依据公式 (1)、(2) 对每一粒子进行迭代;

$$[0082] \quad v_{id} = w*v_{id} + c_1*r_{1d}*(p_{id} - x_{id}) + c_2*r_{2d}*(p_{gd} - x_{id}) \quad (1);$$

$$[0083] \quad x_{id} = x_{id} + v_{id} \quad (2);$$

[0084] 如已达最大设置的迭代次数,则终止算法,输出找到的最好解;否则转去第 (b) 步;其中 $w, c_1,$ 和 c_2 一般为常数,如 $w = 0.729, c_1 = c_2 = 1.49445$; r_{1d} 和 r_{2d} 为 $(0, 1)$ 之间的随机数。

[0085] 对如图 1 所示基因芯片图像,它的基因点定位问题可转化为在一 9600 维的解空间中寻找一最优解(既 4800 个基因点的定位信息)。由于是高维解空间,求解复杂,本实施例将这一问题分成两步,第一步对基因区进行优化网格定位,第二步再对每一基因区的基因点进行定位。

[0086] 1) 基因区网格定位

[0087] 对基因芯片图像进行基因区网格定位可先对基因芯片图像进行旋转校正,再采用图像投影法。将旋转角度作为要寻找的解,采用粒子群算法求解这一一维问题。对每一解的评估采用如下步骤:

[0088] a) 依据解所表示的旋转角度对基因芯片图像进行旋转校正;

[0089] b) 对旋转校正过的图像进行水平和垂直方向的投影;

[0090] c) 对能进行基因区网格定位的图像赋予大的评估值,并且网格间距越大评估值相对来说越大;旋转后不能进行基因区网格定位的图像则赋予较小的评估值;

[0091] 通过粒子群算法找到旋转角度后,先对基因芯片图像进行旋转校正,图像投影法对图像进行基因区网格定位。

[0092] 2) 每一基因区内基因点定位

[0093] 对每一基因区基因点定位依据下列步骤:

[0094] a) 先对基因区图像进行旋转校正;

[0095] b) 再对旋转校正过后的基因区图像依据已知行列数进行均匀行列分格,每一格的中心暂定为每一基因点的中心;

[0096] c) 对每一基因点的中心作些微调,并以该中心作半径 r 的园,该园所对应区域则为该基因点的位置及大小。

[0097] 从以上步骤,可看出对每一基因区基因点定位需找到以下信息:旋转校正角度大小,每一基因点的中心位置的 x 和 y 方向的微调值,每一基因点的半径。一般来说,同一基因区内基因点的半径相同,因而对图一中的每一基因区需要寻找的信息为:1 个旋转校正角度,1 个基因点的半径,100 个 x 方向的微调值,和 100 个 y 方向的微调值,所以图一中每一基因区内基因点定位所需寻找的解的空间维数为 202 维。同样采用粒子群优化算法来寻找对应的 202 维最佳解。每一 202 维解的评估依据以下步骤:

[0098] a) 依据每一解所表示的旋转角度对基因区图像进行旋转校正;

[0099] b) 对旋转校正过后的基因区图像依据已知行列数进行均匀行列分格,每一格的中心暂定为每一基因点的中心;

[0100] c) 对每一基因点的中心依据解所表示的对应于每一基因点 x 和 y 方向值对每一基因点中心作 x 和 y 方向微调；

[0101] d) 在每一中心依据解所表示的半径作圆形成每一基因点的范围；

[0102] e) 在所有基因点内包含信息的像素越多和在所有基因点外包含信息的像素越少，则该解的评估值越大，否则则越小。

[0103] 这样通过粒子群算法寻找到的最佳解就将基因区内的基因点定位。

[0104] 上述实例只为说明本发明的技术构思及特点，其目的在于让熟悉此项技术的人是能够了解本发明的内容并据以实施，并不能以此限制本发明的保护范围。凡根据本发明精神实质所做的等效变换或修饰，都应涵盖在本发明的保护范围之内。

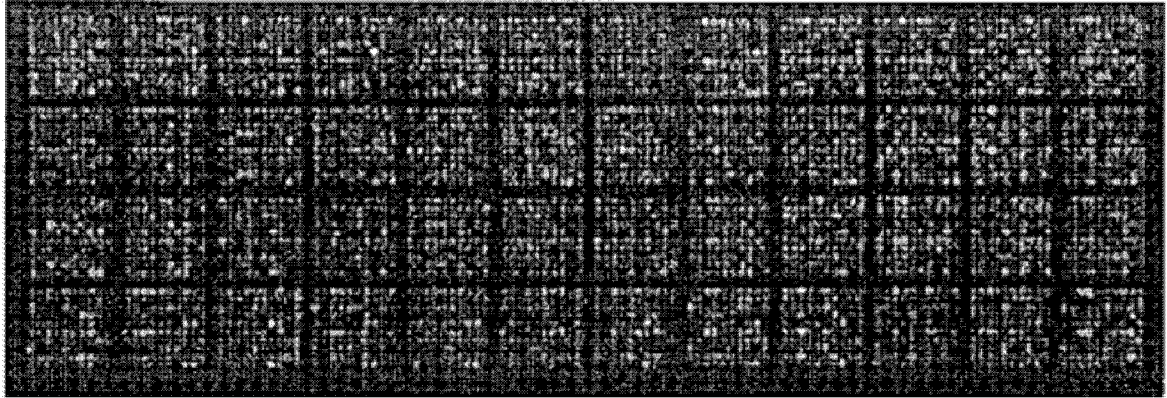


图 1

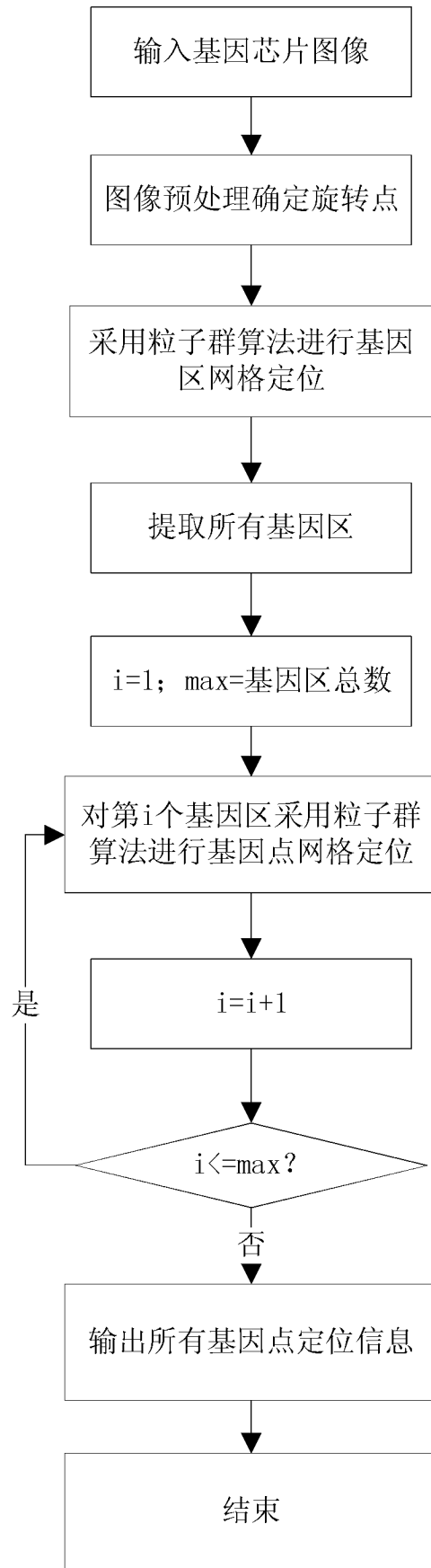


图 2