

1. 一种贝叶斯显微成像方法,其特征在于包括以下内容:
 - 1) 建立用于构建不同位置荧光分子出现概率分布图的初始模型;
 - 2) 采用聚类分析方法、限定的前向算法和共轭梯度法优化初始模型并进行模型选择;
 - 3) 重复步骤2)直到相邻两次迭代重构结果符合设定要求;
 - 4) 将每次迭代过程中得到的荧光分子位置信息进行重构,得到荧光分子的超分辨定位图。
2. 如权利要求1所述的一种贝叶斯显微成像方法,其特征在于,所述步骤1)初始模型建立的具体过程为:

连续采集一系列荧光图像作为原始图像序列;

利用荧光分子的转移概率,随机选择分析区域中的某些位置作为初始位置,并在初始位置处建立因子隐马尔可夫初始模型。
3. 如权利要求1所述的一种贝叶斯显微成像方法,其特征在于,所述步骤2)优化初始模型的具体过程为:

采用聚类分析方法对初始模型中荧光分子按照彼此相对距离的不同划分成N个不同的类别;以每一个类别中的初始位置出发,采用混合MCMC和限定的前向算法计算某一位置出现荧光分子的概率;

采用共轭梯度法计算出现荧光分子概率的极值点,选取概率最高的位置作为优化结果,合并N个类别,并在获得概率最高的位置建立新的隐马尔可夫初始模型。
4. 如权利要求3所述的一种贝叶斯显微成像方法,其特征在于,所述步骤2)模型选择的具体过程为:
 - ① 采用聚类分析方法对得到的新的初始模型中的荧光分子位置按照彼此相对距离的不同进行聚类分析,将其分为不同的类别;
 - ② 随机选择某个位置作为待新增的荧光分子位置或者在新的初始模型中随机选择待删除荧光分子;
 - ③ 判断步骤②新增或删除的荧光分子与步骤①聚类后哪一个类别最近,在其最近或所属的类别中,使用混合MCMC、限定的前向算法和共轭梯度法计算在此位置增加一个荧光分子或删除一个已有荧光分子后对模型概率的影响,如果概率增加则保留该位置信息,并将其加入到新的初始模型作为新的初始位置。
5. 一种贝叶斯显微成像方法,其特征在于包括以下内容:
 - 1) 提取单分子信息提取,并获取单分子候选点;
 - 2) 随机从单分子候选点中选择部分点作为初始点,建立用于构建不同位置荧光分子出现概率分布图的初始模型;
 - 3) 采用聚类分析方法将初始模型中的荧光分子按位置信息进行聚类,对每一类荧光分子进行优化处理得到扩展点;
 - 4) 从候选点中删除已经分析过的荧光分子,并从步骤2)中剩下的候选点随机选择建立初始模型,重复步骤3)得到新的扩展点,直到所有的候选点分析完成;
 - 5) 合并所有的扩展点作为新的候选点重复步骤3)和步骤4),直到相邻两次迭代重构结果满足设定要求,迭代停止;
 - 6) 将迭代过程中得到的荧光分子位置信息进行重构,得到超高分辨率荧光分子的定位

图。

6. 如权利要求5所述的一种基于单分子定位的贝叶斯显微方法,其特征在于,所述步骤1)提取单分子信息提取,并获取单分子候选点通过双通道荧光显微镜采集两通道采集单分子和高密度信息,或者通过单通道同时采集单分子和高密度信息。

7. 如权利要求6所述的一种贝叶斯显微方法,其特征在于,所述步骤1)提取单分子信息提取,并获取单分子候选点,具体过程为:

1.1)对于提供单分子信息的图像序列,利用单分子定位方法,精确得到荧光分子的位置信息;

1.2)当采用双通道进行数据采集时,在其中一通道图像序列中选择与待分析另一通道图像序列相同图像帧数所包含的单分子作为候选点;当采用单通道进行数据采集时,选择所有图像序列中的单分子作为候选点;

1.3)当采用双通道进行数据采集时需要进行漂移校准:分别校准两通道的图像数据,并将两通道之间建立联系,将其中一通道图像的单分子候选点校准到待分析另一通道对应区域中。

8. 如权利要求5所述的一种贝叶斯显微方法,其特征在于,所述步骤3)采用聚类分析方法将初始模型中的荧光分子按位置信息进行聚类,对每一类荧光分子进行优化处理得到扩展点,包括以下内容:

3.1)用混合前向算法和MCMC方法获得初始模型的采样图;

3.2)对类中的每一个荧光分子,用限定的前向算法和改进的共轭梯度方法进行优化获得新的荧光分子;

3.3)对每个新的荧光分子通过距离和强度阈值进行判断,确定是否将计算得到的新的荧光分子作为扩展点;

3.4)计算扩展点的采样图,并将这些采样图通过扩张采样方法添加到初始模型的采样图中。

9. 如权利要求1或5所述的一种贝叶斯显微成像方法,其特征在于,限定的前向算法是指当计算某个荧光分子的观测概率时,不需要计算整个图像区域,而仅仅计算以该荧光分子中心坐标为中心,3个标准差的范围内的值,当超过该范围时,认为不会影响周围的荧光分子。

10. 如权利要求1或5所述的一种贝叶斯显微成像方法,其特征在于,所述聚类分析方法采用k-means方法。

一种贝叶斯显微成像方法

技术领域

[0001] 本发明涉及一种贝叶斯显微成像方法,属于生物超高分辨显微成像技术领域。

背景技术

[0002] 超高分辨率荧光显微镜在生物学研究领域有着广泛的应用,使得人们能够更好、更深入地认识、理解细胞生物学。与电子显微镜一样,超高分辨率荧光显微镜可以让人们观察到以往无法观察到的细胞内部结构及动态变化过程。并且,超高分辨率荧光显微镜还具有连电子显微镜都无可比拟的优势—对活细胞成像,并能高特异性地标记多个分子。

[0003] 由于光学衍射的限制,传统的光学显微镜不能识别200nm以下的结构。超高分辨率荧光显微技术打破了这一限制,并在2014年获得诺贝尔化学奖,该技术主要可以分为两类,结构光照明技术(例如受激发射损耗显微技术STED和饱和结构照明显微技术SIM)和单分子荧光定位技术(例如随机光学重构显微技术STORM和光激活定位显微技术PALM)。其中,单分子定位技术通过大量具有稀疏荧光分子的图像拟合叠加得到超高分辨率图像,但是该方法要求图像中的荧光分子密度足够低以保证单个荧光分子之间不能重叠。为了获得超高分辨率图像就需要大量的图像,这样长时间拍照就会导致样品漂移和细胞损伤,时间分辨率低,影响该技术在活细胞成像中的应用。针对活细胞成像要求,当前有很多方法通过拟合多个荧光分子来得到更高时间分辨率的结果。但是,随着荧光密度的增加,很难确定真实的拟合模型,定位精度大大下降,这限制了时间分辨率的进一步提高。

[0004] 近年来,研究发现一种利用统计贝叶斯解决高密度荧光图像识别的方法(Bayesian analysis of the blinking and bleaching,3B),3B利用荧光分子漂白和闪烁的特性,通过分析整个图像序列的变化得到荧光分子概率图。作为细胞成像新的重要工具,尽管它可以用简单的光学设备实现活细胞动态结构的超高分辨率成像,但是仍然有三个关键的问题需要解决:1)在精度方面,存在严重的结构缺失,定位精度不高;2)在速度方面,该方法极其耗时,为了得到 $1.5\mu\text{m}\times 1.5\mu\text{m}$ 的超分辨率结构,大约需要6小时,并且随着图像尺寸的增加,计算时间急剧增长;3)在分析尺度方面,由于速度的限制,该方法很难获得整个活细胞大尺度长时间的动态变化。

发明内容

[0005] 针对上述问题,本发明的其中一个发明目的是提供一种能够加快高密度数据重构速度且能够提高重构精度的贝叶斯显微成像方法。

[0006] 本发明的另一发明目的是提供一种基于单分子定位的贝叶斯显微成像方法,能够确定细胞等生物样品中用来标记蛋白等生物物质的荧光分子在纳米尺度的精确定位,以及在活细胞中精确定位动态变化即活细胞中的超高分辨显微成像。

[0007] 为实现上述其中一发明目的,本发明采取以下技术方案:一种贝叶斯显微成像方法,其特征在于包括以下内容:1)建立用于构建不同位置荧光分子出现概率分布图的初始模型;2)采用聚类分析方法、限定的前向算法和共轭梯度法优化初始模型并进行模型选择;

3)重复步骤2)直到相邻两次迭代重构结果符合设定要求;4)将每次迭代过程中得到的荧光分子位置信息进行重构,得到荧光分子的超分辨定位图。

[0008] 进一步,所述步骤1)初始模型建立的具体过程为:连续采集一系列荧光图像作为原始图像序列;利用荧光分子的转移概率,随机选择分析区域中的某些位置作为初始位置,并在初始位置处建立因子隐马尔可夫初始模型。

[0009] 进一步,所述步骤2)优化初始模型的具体过程为:采用聚类分析方法对初始模型中荧光分子按照彼此相对距离的不同划分成N个不同的类别;以每一个类别中的初始位置出发,采用混合MCMC和限定的前向算法计算某一位置出现荧光分子的概率;采用共轭梯度法计算出现荧光分子概率的极值点,选取概率最高的位置作为优化结果,合并N个类别,并在获得概率最高的位置建立新的隐马尔可夫初始模型。

[0010] 进一步,所述步骤2)模型选择的具体过程为:①采用聚类分析方法对得到的新的初始模型中的荧光分子位置按照彼此相对距离的不同进行聚类分析,将其分为不同的类别;②随机选择某个位置作为待新增的荧光分子位置或者在新的初始模型中随机选择待删除荧光分子;③判断步骤②新增或删除的荧光分子与步骤①聚类后哪一个类别最近,在其最近或所属的类别中,使用混合MCMC、限定的前向算法和共轭梯度法计算在此位置增加一个荧光分子或删除一个已有荧光分子后对模型概率的影响,如果概率增加则保留该位置信息,并将其加入到新的初始模型作为新的初始位置。

[0011] 为实现上述另一发明目的,本发明采取以下技术方案:一种贝叶斯显微成像方法,其特征在于包括以下内容:1)提取单分子信息提取,并获取单分子候选点;2)随机从单分子候选点中选择部分点作为初始点,建立用于构建不同位置荧光分子出现概率分布图的初始模型;3)采用聚类分析方法将初始模型中的荧光分子按位置信息进行聚类,对每一类荧光分子进行优化处理得到扩展点;4)从候选点中删除已经分析过的荧光分子,并从步骤2)中剩下的候选点随机选择建立初始模型,重复步骤3)得到新的扩展点,直到所有的候选点分析完成;5)合并所有的扩展点作为新的候选点重复步骤3)和步骤4),直到相邻两次迭代重构结果满足设定要求,迭代停止;6)将迭代过程中得到的荧光分子位置信息进行重构,得到超高分辨率荧光分子的定位图。

[0012] 进一步,所述步骤1)提取单分子信息提取,并获取单分子候选点通过双通道荧光显微镜采集两通道采集单分子和高密度信息,或者通过单通道同时采集单分子和高密度信息。

[0013] 进一步,所述步骤1)提取单分子信息提取,并获取单分子候选点,具体过程为:1.1)对于提供单分子信息的图像序列,利用单分子定位方法,精确得到荧光分子的位置信息;1.2)当采用双通道进行数据采集时,在其中一通道图像序列中选择与待分析另一通道图像序列相同图像帧数所包含的单分子作为候选点;当采用单通道进行数据采集时,选择所有图像序列中的单分子作为候选点;1.3)当采用双通道进行数据采集时需要进行漂移校准:分别校准两通道的图像数据,并将两通道之间建立联系,将其中一通道图像的单分子候选点校准到待分析另一通道对应区域中。

[0014] 进一步,所述步骤3)采用聚类分析方法将初始模型中的荧光分子按位置信息进行聚类,对每一类荧光分子进行优化处理得到扩展点,包括以下内容:3.1)用混合前向算法和MCMC方法获得初始模型的采样图;3.2)对类中的每一个荧光分子,用限定的前向算法和改

进的共轭梯度方法进行优化获得新的荧光分子;3.3)对每个新的荧光分子通过距离和强度阈值进行判断,确定是否将计算得到的新的荧光分子作为扩展点;3.4)计算扩展点的采样图,并将这些采样图通过扩张采样方法添加到初始模型的采样图中。

[0015] 进一步,限定的前向算法是指当计算某个荧光分子的观测概率时,不需要计算整个图像区域,而仅仅计算以该荧光分子中心坐标为中心,3个标准差的范围内的值,当超过该范围时,认为不会影响周围的荧光分子。

[0016] 进一步,所述聚类分析方法采用k-means方法。

[0017] 由于本发明采取以上技术方案,其具有以下优点:1、在时间分辨率方面,本发明与单分子定位标准方法PALM相比,时间分辨率提高100~1000倍,与高密度定位显微3B相比,重构需要图像帧数从至少200帧降低到100帧,时间分辨率提高两倍;在空间分辨率方面,本发明与PALM相比,重构结构基本相同;与3B相比,结构连续性更好,不会出现结构的缺失,定位精度高;在运行速度方面,相比3B有几个数量级的提高,针对 50×50 的像素区域,加速比将近100倍,并且随着分析区域的增大,加速效果更加明显,可以在普通PC机对整个成像视野进行超高分辨显微成像;在分析尺度方面,由于图像连续性改善和运行速度大幅提高,使得计算全细胞大尺度动态变化成为可能。2、本发明基于限定的前向算法和聚类分析方法实现了贝叶斯显微成像方法,大大加快了高密度超分辨率成像技术的运行速度和定位精度。3、本发明提出基于单分子定位的贝叶斯显微成像方法是结合单分子定位显微技术(SIMBA),利用双通道荧光显微镜或单通道荧光显微镜进行实现,进一步提高了运行速度和时空分辨率,使得全细胞大尺度长时间动态变化分析成为可能。本发明广泛适用于大视野的活细胞及固定细胞的荧光显微成像。

附图说明

[0018] 图1是不同图像尺寸下,本发明执行一次前向算法的运行时间示意图,其中,虚线为3B前向算法运行时间,实线为采用限定的前向算法后的运行时间;

[0019] 图2是本发明的贝叶斯显微成像方法流程示意图;

[0020] 图3是本发明荧光分子的状态转移示意图;

[0021] 图4是本发明的3B、quick-3B和SIMBA运行时间和重构结果对比效果示意图,图(a)~(d)分析区域均为 38×50 的像素区域,图(a)为200帧原始图像叠加,图(b)为3B重构结果,图(c)quick-3B重构结果,图(d)SIMBA重构结果,图(e)为3B、quick-3B和SIMBA运行时间对比示意图,运行时间分别为19.02小时、1.13小时和11分钟,quick-3B加速比为16.8,SIMBA加速比约为100;

[0022] 图5是本发明的荧光蛋白模型示意图,采用因子的隐马尔可夫模型(FHMM)表示多个荧光蛋白共同作用;

[0023] 图6是本发明的SIMBA方法流程示意图;

[0024] 图7是本发明肌动蛋白全细胞超分辨率重构结果,共包含红绿两个通道数据集,其中2万帧红通道数据用于PALM成像,200帧绿通道数据用于3B和SIMBA成像,其中,图(a)是绿通道SIMBA数据集200帧原始图像的叠加,显示了衍射现象的影响,图(b1)~(b3)是(a)方框区域PALM(b1)、3B(b2)和SIMBA(b3)重构结果比较,图(b4)是用NCC(归一化相关系数)比较了3B与PALM、SIMBA与PALM的相似性,NCC值分别为0.509和0.744,图(c1)是不同图像尺寸下

3B和SIMBA运行时间对比,对于图像尺寸从10X10像素区域到50X50像素区域的数据集,3B的运行时间分别为0.81小时、3.17小时、8.05小时、13.87小时和23.4小时,如虚线所示;SIMBA的运行时间分别为3分钟、7分钟、10分钟、13分钟和16分钟,如实线所示,图(c2)是相对3B, SIMBA加速比曲线,不同尺寸下加速比分别为16.2、27.2、48.3、64和87.7,图(d1)、(d2)分别PALM分析、SIMBA分析全细胞重构结果比较区域放大图,图(d3)是PALM分析、SIMBA分析全细胞重构结果局部区域放大图,图(e1)、(e2)是图(d1)和图(d2)线条的属性图,显示其分辨率曲线长虚线是PALM结果,实线是SIMBA结果,图(e3)是图(d3)线条分辨率曲线,分辨率为49nm;

[0025] 图8(a)~(c)是本发明实施例肌动蛋白全细胞大尺度长时间动态变化。

具体实施方式

[0026] 以下结合附图来对本发明进行详细的描绘。然而应当理解,附图的提供仅为了更好地理解本发明,它们不应该理解成对本发明的限制。

[0027] 为了加快高密度数据重构的速度,提高重构精度,本发明提出了一种快速的贝叶斯显微成像方法(quick-3B),包括两方面内容:1)限定的前向算法用于快速的计算荧光分子的观测概率,加快分析速度;2)聚类分析方法用于减少不相关的荧光分子影响,提高重构精度,下面对上述两方面内容的原理进行详细阐述。

[0028] 1)限定的前向算法

[0029] 当采用3B方法分析荧光数据时,需要选择荧光分子的初始数目,但是随着图像分析区域的增大,即使选择相同初始数目的荧光分子,计算时间也会急剧增长,甚至到无法计算的地步。为了高效地确定荧光分子的参数,3B使用了一种混合前向算法和MCMC(马尔科夫蒙特卡洛)的方法。在这个过程中,3B利用前向算法计算了某一位置是否是荧光分子的概率。如图1所示,通过虚线可以看出,在选择相同初始数目的荧光分子情况下,前向算法的运行时间和图像的尺寸成正比。进一步分析发现,前向算法在计算过程中需要确定某一数据区域的观测概率,即求某一帧图像某一个区域包含荧光分子或者不含荧光分子的概率,在这个过程中3B需要计算整个图像区域。因此,3B的时间复杂度与荧光分子数目和图像像素数目的乘积成正比。然而,图像系统是受点扩散函数(PSF)的影响,PSF可以假设是服从正态分布, $N(u, \sigma^2)$ (均值是 u ,标准差是 σ)。荧光分子产生的高斯斑是以坐标原点为中心,呈指数下降的趋势,并且99.7%的置信区间是以荧光分子中心坐标为中心3个标准差范围以内,因此本发明提出了一种限定的前向算法,对荧光分子的作用范围加以限制。当计算某个荧光分子的观测概率时,不需要计算整个图像区域,而仅仅计算以该荧光分子中心坐标为中心,3个标准差的范围内的值,当超过该范围时,认为不会影响周围的荧光分子。如图1所示的实线表示在使用限定的前向算法时不同图像尺寸的运行时间,可以看出,选择相同初始数目的荧光分子,随着图像尺寸的增大,计算时间基本不变。在优化初始模型中使用限定的前向算法,使得时间复杂度从荧光分子数目与图像像素数目乘积成正比,变成荧光分子数目与常数的乘积成正比。

[0030] 2)聚类分析方法

[0031] 3B通过因子的隐马尔科夫模型(FHMM)对模型中所有荧光分子进行建模,在该方法中,某个荧光分子位置的确定需要考虑其他所有的荧光分子的影响。然而,由于PSF作用范

围的影响,如果一个荧光分子远离其他的荧光分子一定距离以上,那么该荧光分子几乎不会对其他荧光分子产生影响,而将所有荧光分子引入会导致计算误差的增大。从上述分析可以看出,为了消除不相关荧光分子的影响,把相邻的、可能会相互作用的荧光分子聚在一起,然后分别进行计算。这样,自然的引入了聚类的概念,聚类是将一系列对象组成的集合进行划分,具有相似属性的对象划分成一类。在同一类别中,对象之间彼此相似,但与其他类别中的对象相异。聚类分析将最有可能相互作用的对象聚合在一起,考虑了对象之间的固有属性,正好可以解决当前的问题。本发明采用k-means算法对荧光分子进行聚类,对于包含M个荧光分子的模型,将其划分成k个类别,并用荧光分子之间的位置关系来衡量是否分在同一类别中,计算如公式(1)所示:

$$[0032] \quad \underset{S}{\operatorname{arg\,min}} \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - u_i\|^2 \quad (1)$$

[0033] 式中, x_j 为荧光分子j的位置信息, u_i 为类别i的中心位置,k-means算法使得荧光分子都被分配到最近的类别 S_i 中, $\|x_j - u_i\|^2$ 表示荧光分子 x_j 和类别中心 u_i 的距离。k-means方法目的是要最小化每个荧光分子到它对应类别中心的距离之和。

[0034] 在每轮新的计算过程中都采用k-means算法将荧光分子重新划分成不同的类别,然后这些荧光分子分别在各自的类中优化,最后合并结果,进一步降低了计算量。

[0035] 综上所述,如图2所示,基于限定的前向算法和聚类分析方法,本发明提出了一种快速的贝叶斯显微成像方法(quick-3B),通过对连续多张采样图像序列进行分析,在纳米尺度确定每个发光的荧光分子在细胞中的精确定位,包括以下步骤:

[0036] 1、初始模型建立,具体过程为:

[0037] 1.1)采集原始图像序列,即连续采集一系列一定曝光时间的荧光图像作为原始图像序列,本发明实施例中获取200帧曝光时间为10毫秒的荧光图像,但是以此为例,不限于此。

[0038] 1.2)利用荧光分子的转移概率(如图3所示),随机选择分析区域中的某些位置作为初始位置,并在初始位置处建立因子隐马尔可夫(factorial hidden Markov models)初始模型,用来最终构建不同位置荧光分子出现的概率分布图。

[0039] 2、优化初始模型,即以初始模型中的初始位置出发,计算初始位置周围荧光分子可能出现的概率,选择概率最高的位置重新建立新的初始模型,具体过程为:采用k-means算法对初始模型中的荧光分子按照彼此相对距离的不同划分成N个不同的类别,以每一个类别中的初始位置出发,采用混合MCMC和限定的前向算法计算某一位置出现荧光分子的概率,并用共轭梯度法计算出现荧光分子概率的极值点(梯度的步长是0.5),选取概率最高的位置作为优化结果,合并N个类别,并在获得概率最高的位置建立新的隐马尔可夫初始模型。

[0040] 3、模型选择,具体过程为:

[0041] 3.1)采用k-means算法,对步骤2)得到的新的初始模型中的荧光分子位置按照彼此相对距离的不同进行聚类分析,将其分为不同的类别;

[0042] 3.2)随机选择某个位置作为待新增的荧光分子位置或者在新的初始模型中随机选择待删除荧光分子;

[0043] 3.3)判断步骤3.2)新增或删除的荧光分子与步骤3.1)聚类后哪一个类别最近。在

其最近或所属的类别中,使用混合MCMC、限定的前向算法和共轭梯度计算在此位置增加一个荧光分子或删除一个已有荧光分子后对模型概率的影响,如果概率增加则保留该位置信息,并将其加入到新的初始模型作为新的初始位置。

[0044] 4、重复步骤2和步骤3,直到相邻两次迭代重构结果没有明显的变化,可以用迭代过程中新增荧光点的数目占总荧光点数目的百分比作为截止条件,例如1%。本发明实施例中迭代次数约200次,但是仅以此为例,不限于此。

[0045] 5、将每次迭代过程中得到的位置信息重构为一张概率分布图,得到荧光分子的超分辨率定位图。

[0046] 图4对比了3B和quick-3B运行时间和超分辨率重构结果,在运行速度方面,对于200帧 38×50 像素分析区域的数据集,运行时间从19.02小时缩短到了1.13小时,加速比为16.8;在重构效果方面,quick-3B结构连续性更好,与原始叠加图更加一致。

[0047] 为了进一步提高时空分辨率和运行速度,使得全细胞大尺度长时间动态变化分析成为可能,结合单分子定位显微技术,利用单通道荧光显微镜或双通道荧光显微镜(本发明实施例是基于双通道荧光显微镜进行说明,但是不限于此),本发明提出了一种基于单分子定位的贝叶斯显微成像方法(SIMBA),本发明使用一种光转换荧光蛋白mEos3.2,该蛋白在不同波长的激光照射下会发出红绿两种不同荧光,其中,发出红色荧光的通道称之为红通道,该通道荧光分子相互不重叠,发出绿色荧光的通道称之为绿通道,该通道荧光分子信号相互之间是重叠的,通过二分镜进行分光,这样针对同一细胞结构就可以在不同通道观测到不同的荧光状态,这是SIMBA方法的基础。首先通过单分子定位方法从红通道提取出单分子,这些单分子是用于指导对应绿通道荧光分子位置优化的,利用双通道图像校准技术将红通道的单分子数据准确的校准到对应的绿通道区域。然后,在绿通道中基于这些单分子组成的候选点,利用贝叶斯分析方法得到新的荧光分子位置,称之为扩展点,这些扩展点继续扩张,得到更多新的荧光分子。在这个扩张的过程中,随机选择已经得到的部分荧光分子作为初始模型,然后优化这个初始模型得到新的荧光分子,不断重复模型优化和扩张,最终得到超高分辨率的重构结果。

[0048] SIMBA方法主要包括四个方面:1)红通道单分子定位用于得到红通道真实的荧光分子定位信息;2)双通道图像校准用于将红通道的单分子数据准确的校准到对应的绿通道中;3)扩张采样和改进的共轭梯度用于结合单分子来计算未获得的荧光分子,减少结构信息的缺失;4)距离和密度阈值用于去除错误的荧光分子,下面对上述四方面的原理进行详细阐述。

[0049] 1)红通道单分子定位

[0050] 为了精确的获得荧光分子的位置信息,SIMBA使用了一种光转换荧光蛋白mEos3.2来标记细胞结构,该荧光蛋白在波长561nm的激光照射下发出红色的荧光,呈现的是单分子的形态,荧光分子之间互不重叠,每一帧图像的荧光密度低,在波长488nm的激光照射下发出绿色的荧光,荧光蛋白之间互相重叠,每一帧图像的荧光密度很高,通过二分镜进行分光,同时采集两个通道的数据。对于红通道的数据,利用单分子定位的方法精确的得到荧光蛋白的位置信息,这些单分子用于指导绿通道荧光蛋白位置的优化。

[0051] 2)双通道图像校准

[0052] 由于拍摄过程中,实验设备会存在一定漂移,红绿两个通道是通过用二分镜进行

分光,其位置和形变程度也不同。因此,需要将每个通道产生的图像序列和红绿灯通道之间进行校准。对于单个通道的图像序列校准,通过追踪序列里某个荧光珠的轨迹,定位该荧光珠,然后计算图像序列和第一帧图像之间的位移差就可以校准序列间的位置漂移。对于红绿灯通道之间的校准,将整个样品区域洒满了双色荧光珠(beads),此荧光可同时被波长为488nm和561nm的激光激发,在两个通道中就显示了相同的beads信息,采用仿射变换(Affine Transformation)对两个通道进行校准。仿射变换是一种二维坐标到二维坐标之间的线性变换,公式如(2)所示,它能实现图像的平移(Translation)、缩放(Scale)、旋转(Rotation)、翻转(Flip)和错切(Shear)。首先利用双色荧光珠计算出仿射变换的变换参数,然后用这一参数将红通道得到的单分子数据准确的校准到对应的绿通道区域。

$$[0053] \quad \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_2 & a_1 & a_0 \\ b_2 & b_1 & b_0 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

[0054] 式中,(x,y)为变换前的坐标,(u,v)为变换后的坐标,(a₂,a₁,a₀,b₂,b₁,b₀)为变换参数。

[0055] 3)扩张采样和改进的共轭梯度

[0056] 3B方法利用因子的马尔可夫模型(FHMM)对整个数据集进行建模,模型如图5所示。FHMM的求解过程是期望最大化(EM)的过程,E过程为采样,即已知模型中荧光分子的参数信息,得到符合原图的采样图,M过程为最大化的过程,通过共轭梯度来优化模型,搜寻更可能的荧光分子参数信息,使得模型中新的荧光分子采样图更符合原图,不断重复这个过程,最终得到更精确的FHMM模型参数,如公式3所示:

$$[0057] \quad Q(\phi^{new} | \phi) = E\{\log P(\{F_t, D_t\} | \phi^{new}) | \phi, \{D_t\}\} \quad (3)$$

[0058] 式中,{D_t}表示荧光分子的观测值,即我们可以看到的荧光图像,t=1,⋯,T,共T帧图像,{F_t}表示荧光分子的隐状态。

[0059] 荧光分子独立的在不同状态进行转换:发光、不发光、漂白,如图3所示。Q为模型参数的优化函数,通过给定当前荧光分子参数φ和观测序列{D_t}来逐步优化得到模型中荧光分子新的参数φ^{new}。由于SIMBA已经知道了部分真实的荧光分子定位信息,改变了E过程和M过程来进一步得到周围可能的荧光分子,减少了结构缺失,主要包括两个方面:1、扩张采样用于计算未获得的荧光分子,减少结构信息的缺失;2、改进的共轭梯度用于防止距离相近的荧光分子计算到相同的局部极值。

[0060] 扩张采样:首先用混合前向算法和MCMC方法得到初始模型的采样图,然后在该采样图的基础上得到周围的荧光分子,当得到一个新的荧光分子时,先计算这个荧光分子的采样图,再将该荧光分子的采样图添加到整个模型的采样图中,不断扩张这个模型的采样。这个过程相当于“拼图游戏”,先把部分轮廓填充,然后在子图中优化模型,不断执行该操作,得到了周围的荧光分子信息,最终得到整张图像的超分辨率结构。这样就避免了亮区域荧光分子的重复计算,使得更多的荧光分子计算到未获得的荧光分子区域,使重构结构更加平滑,扩充了结构信息。而且,在这个过程中仅仅计算新获得荧光分子的采样图,标记其变化,然后将该荧光分子的采样图添加到整体模型的采样图中。这样就避免了每次需要确定一个荧光分子时,都要重新计算模型中所有荧光分子的采样图,降低了计算量。

[0061] 改进的共轭梯度:之前提到的扩张采样方法,当得到一个新的荧光分子时,首先计

算这个荧光分子的采样图,然后将该荧光分子的采样图添加到整个模型的采样图中。但是在计算模型周围的荧光分子时,如果还是以模型中的荧光分子的位置信息作为初始位置来优化,那么不仅计算耗时而且邻近的荧光分子还容易跳进相同的局部极值点。因此,为了更准确的优化模型周围的荧光分子,改变了优化的初始位置。假设荧光蛋服从正态分布, $N(u, \sigma^2)$ (均值是 u ,标准差是 σ)。为了保证原来模型中的荧光分子会对新优化的荧光分子产生影响,不能将新优化的初始位置移动过远,经过实验,以模型中待优化的荧光分子坐标原点 u 为中心, 2σ 半径内,随机选择一个坐标作为新优化的初始位置。

[0062] 4)距离和强度阈值,使用距离和强度阈值来去除计算过程中错误的荧光分子,主要包含两个方面:(1)距离阈值来避免相同位置荧光分子的重复计算;(2)强度阈值来平衡荧光分子的强度分布。

[0063] 距离阈值:SIMBA使用共轭梯度得到新的荧光分子,如果优化前后两个荧光分子位置几乎相同,那么在迭代计算过程中,相同的位置就会重复计算多次,这会导致局部亮点的出现。特别的,当需要计算的荧光分子落入到背景区域中,由于周围临近区域没有其他荧光分子,这就使得周围是荧光分子的概率并没有比该位置更高。按照共轭梯度计算极值的方式,就有可能一直计算到相同的区域。因此,使用距离阈值来进行筛选,当优化前后两点位置相差不超过0.01像素,认为是同一个荧光分子,舍弃该荧光分子。

[0064] 强度阈值:使用扩张采样的方式得到新的荧光分子,如果某个区域有大量的荧光分子,那么这个区域将会扩张出更多新的荧光分子,由于某个图像区域的荧光强度是固定的,所以每个荧光分子所具有的强度值就会越来越小,但是荧光分子的强度值不会一直小下去,其强度值服从对数正态分布, $\text{Log-N}(u, \sigma^2)$,这里采用作为强度阈值的下界来平衡荧光分子的亮度。

[0065] 综上所述,如图6所示,结合单分子定位显微技术,本发明提出了一种基于单分子定位的贝叶斯显微成像方法(SIMBA),可以在活细胞进行超高时空分辨率的显微成像,具体步骤为:

[0066] 1、单分子信息提取。

[0067] 单分子信息提取可以包括两种情况,取决于荧光分子的性质,一种是单分子信息和高密度信息不在同一通道,基于双通道荧光显微镜采集两通道原始图像序列,如光转化荧光分子mEos3.2、mEos2以及其他相同功能的荧光染料或蛋白。另一种情况是单通道同时采集单分子和高密度信息,使用光开关荧光分子,例如Skylan-S、mGeos、Dronpa以及其他相同功能的荧光染料或蛋白,一个通道就能同时提供单分子和高密度信息,本发明实施例中使用光转化荧光蛋白mEos3.2利用双通道成像,并对获取的两通道图像序列数据进行预处理,包括:

[0068] 1.1)对于提供单分子信息的图像序列(对于双通道情况,此通道为红通道),利用PALM\STORM\dSTORM\压缩感知等单分子定位方法,精确得到荧光分子的位置信息;

[0069] 1.2)对于双通道情况,在红通道图像序列中选择与待分析绿通道图像序列相同图像帧数所包含的单分子作为候选点;对于单通道情况,选择所有图像序列中的单分子作为候选点;

[0070] 1.3)对双通道图像进行漂移校准:分别校准红通道和绿通道的图像数据,并将红绿通道之间建立联系,将红通道图像的单分子候选点校准到绿通道图像对应区域中。单通

道情况此步省略。

[0071] 2、随机从绿通道图像的候选点中选择部分点作为初始点,建立初始模型。

[0072] 3、优化初始模型,即:采用聚类分析方法(本发明中使用的是k-means)将初始模型中的荧光分子按位置信息进行聚类,对每一类荧光分子,进行如下操作:

[0073] 3.1)用混合前向算法和MCMC方法获得初始模型的采样图;

[0074] 3.2)对类中的每一个荧光分子,用限定的前向算法和改进的共轭梯度方法沿着四个方向优化得到4个新的荧光分子;

[0075] 3.3)对每个新的荧光分子通过距离和强度阈值进行判断,确定是否将计算得到的新的荧光分子作为扩展点;

[0076] 3.4)计算扩展点的采样图,并将这些采样图通过扩张采样方法添加到初始模型的采样图中。

[0077] 4、从候选点中删除已经分析过的荧光分子,并从步骤2中剩下的候选点随机选择建立初始模型,重复步骤3,得到新的扩展点,直到所有的候选点分析完成。

[0078] 5、合并所有的扩展点作为新的候选点重复步骤3和步骤4,直到相邻两次迭代重构结果中增加的新的荧光分子数量与总荧光分子数量的比值小于1%则迭代停止,本发明实施例中迭代次数约120次,但是仅以此为例,不限于此。

[0079] 6、将绿通道图像迭代过程中得到的位置信息重构为一张概率分布图,即为最终重构的超高分辨率荧光分子的定位图。

[0080] 如图4所示,对比了3B、quick-3B和SIMBA运行时间和超分辨率重构结果,在运行速度方面,SIMBA性能继续大幅提升,对于同样200帧 38×50 像素分析区域的数据集,与原始3B方法相比,运行时间从19.02小时缩短到了11分钟,加速比约为100倍;在重构效果方面,SIMBA的连续性进一步改善,基本没有间断结构出现,与原始图像叠加图更加一致。

[0081] 如图7所示,显示了肌动蛋白全细胞超分辨率重构结果,该数据集包含红绿两个通道数据,红通道为2万帧具有单分子性质的图像,用于PALM分析,绿通道为200帧荧光分子高度重叠的图像,用于3B和SIMBA分析。图(a)显示了200帧绿通道原始图像的叠加图,由于3B不能计算大尺寸的图像,选择了图(a)中方框的区域进行分析,分别比较了该区域2万帧PALM重构结果、200帧3B和200帧SIMBA重构结果,可以看出3B重构结果有大量的结构缺失,200帧的SIMBA重构结果与2万帧的PALM重构结果基本一致,时间分辨率提高100倍。接下来,比较了不同尺寸下3B和SIMBA的运行时间,即使是 50×50 的像素区域,加速比也近似90倍,而且随着图像尺寸的增大,加速效果更加明显。

[0082] 由于速度和精度的提升,进一步计算了全细胞区域的结构,SIMBA重构结果可以看到肌动蛋白的丝、束和褶皱边等结构,与PALM重构结果一致。进一步测量空间分辨率,如图7(e3)所示为49nm。

[0083] 如图8所示,相机拍摄的帧率为18毫秒每帧,用100帧图像进行SIMBA重构,时间分辨率为1.8秒。重构图像之间间隔为20秒,共得到30幅SIMBA重构图,选择其中的部分时间点组成动态图进行展示。重构结果同样可以看到肌动蛋白的丝状和束状结构,进一步从子结构动态图可以看出,肌动蛋白的束状结构比丝状结构相对稳定,如图8(c)所示。可以看出,SIMBA可以实现全细胞大尺度长时间动态变化分析。

[0084] 上述各实施例仅用于说明本发明,其中方法的各实施步骤等都是可以有所变化

的,凡是在本发明技术方案的基础上进行的等同变换和改进,均不应排除在本发明的保护范围之外。

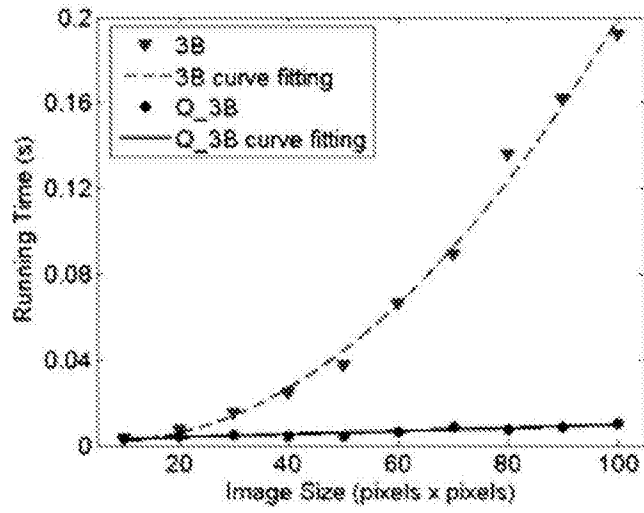


图1

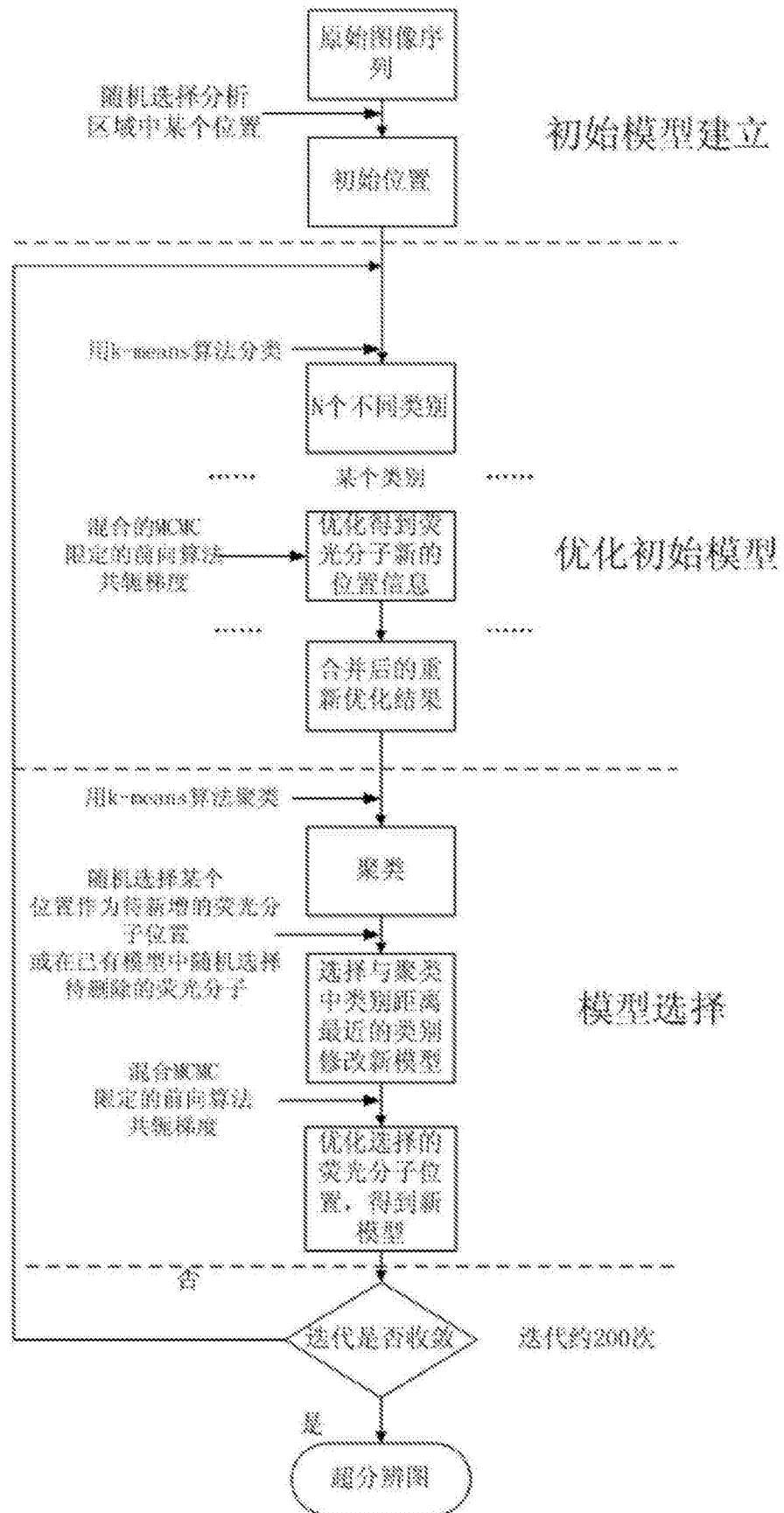


图2

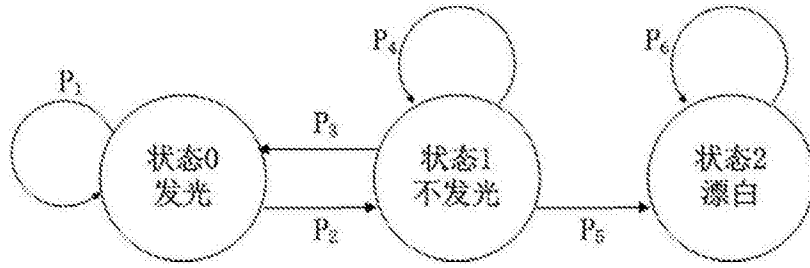
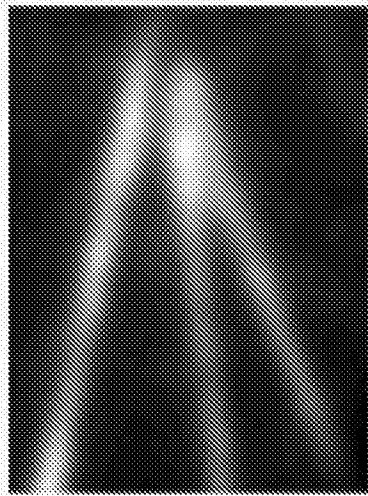
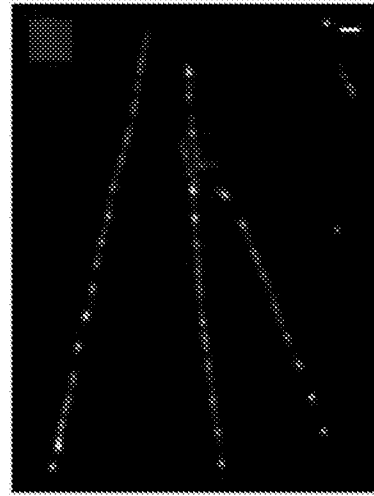


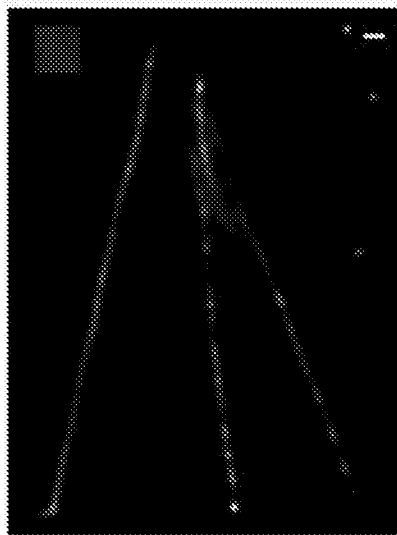
图3



(a)



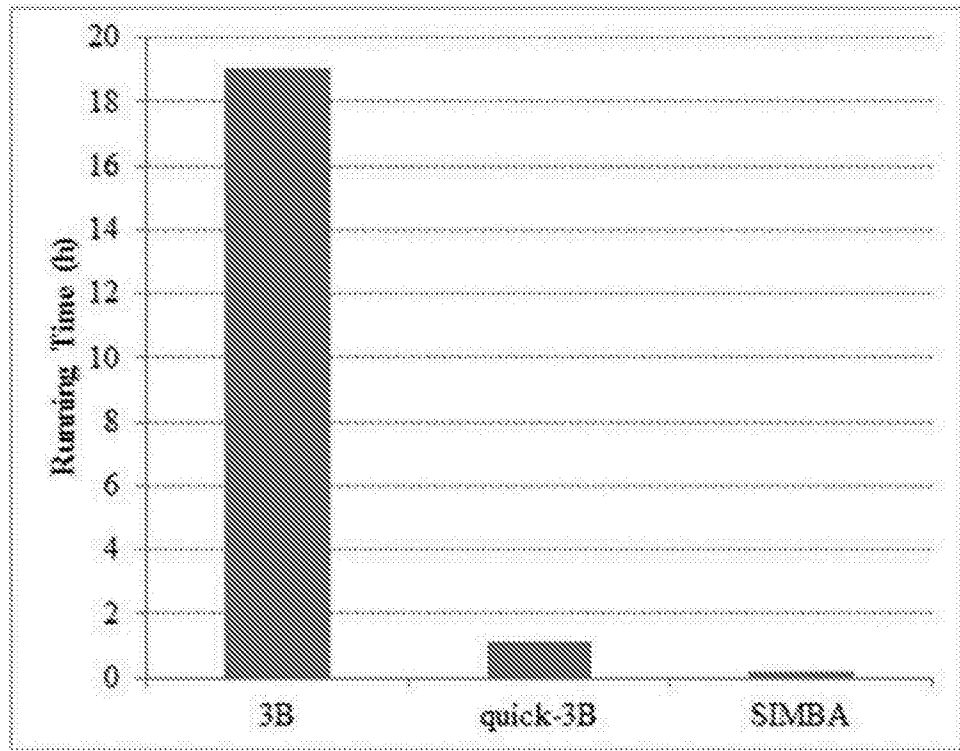
(b)



(c)



(d)



(e)

图4

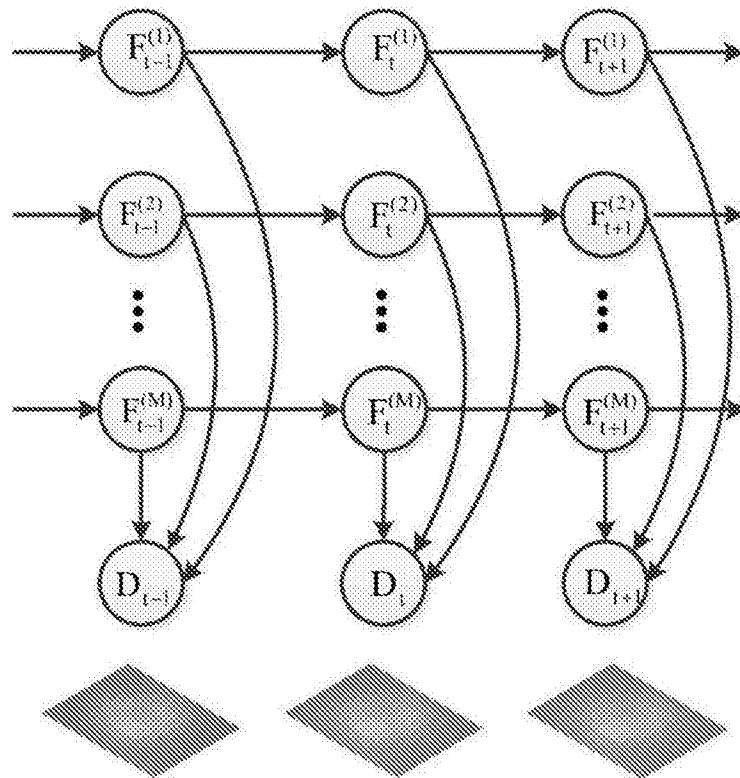


图5

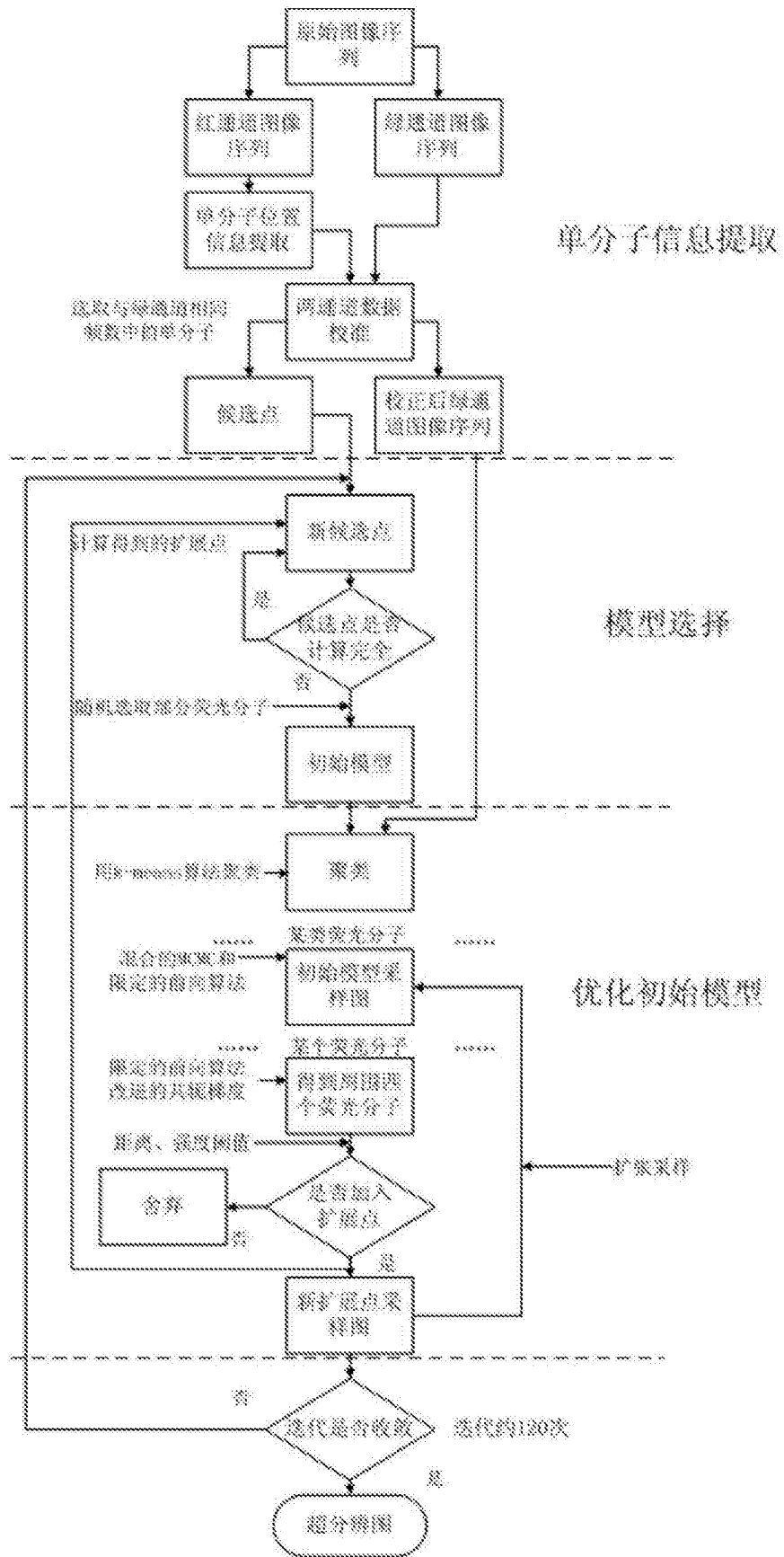
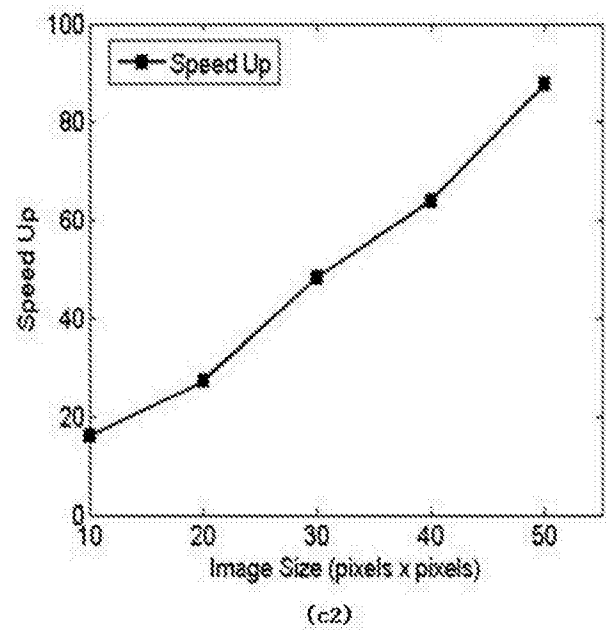
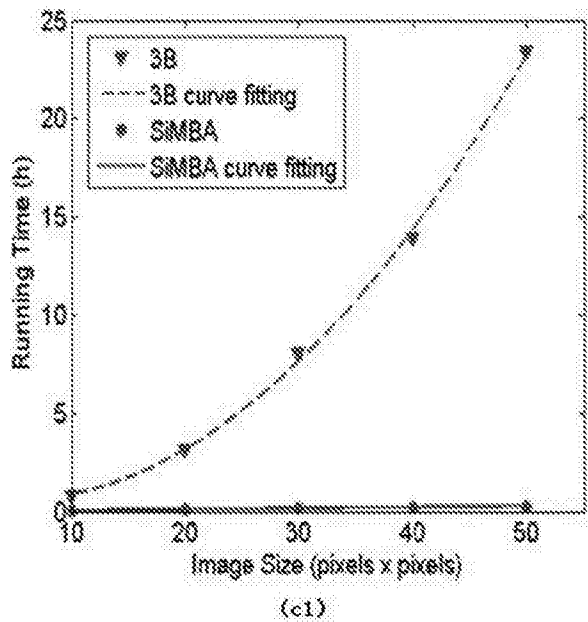
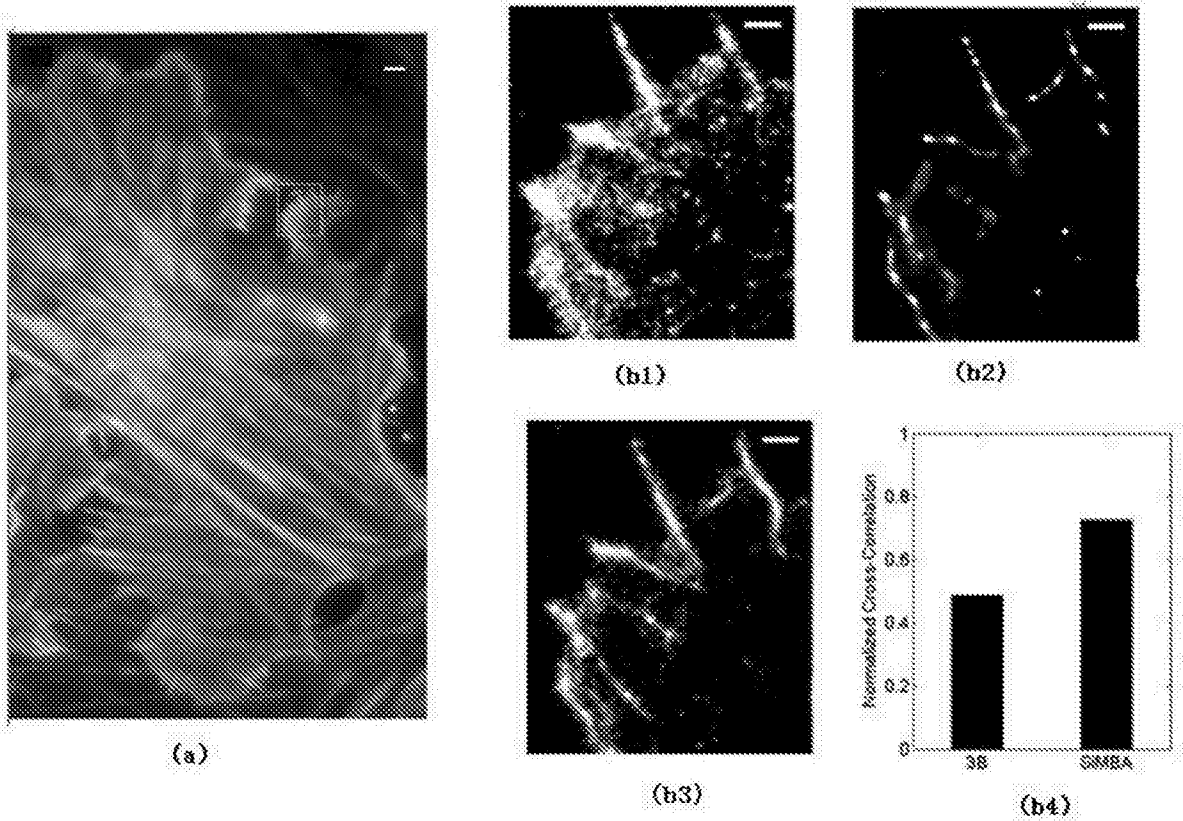
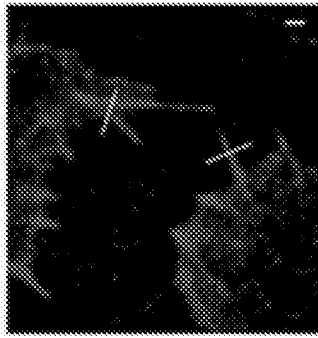
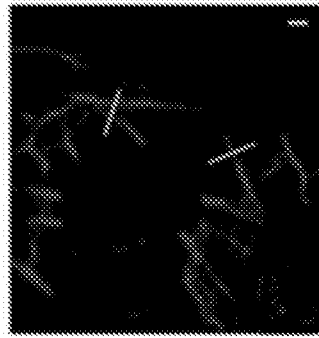


图6





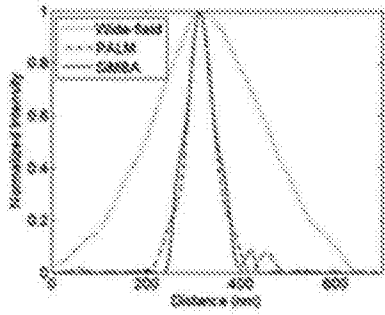
(d1)



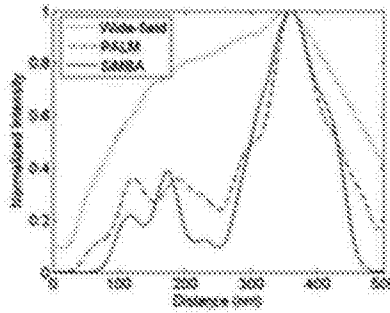
(d2)



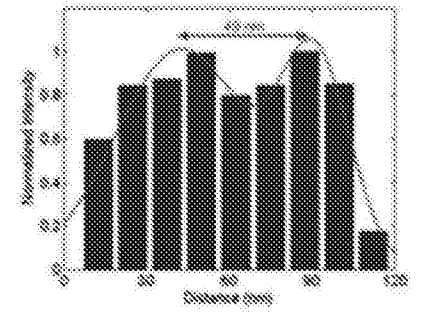
(d3)



(e1)



(e2)

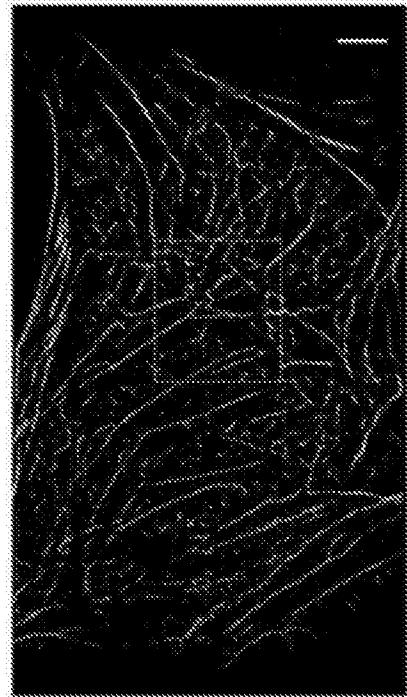


(e3)

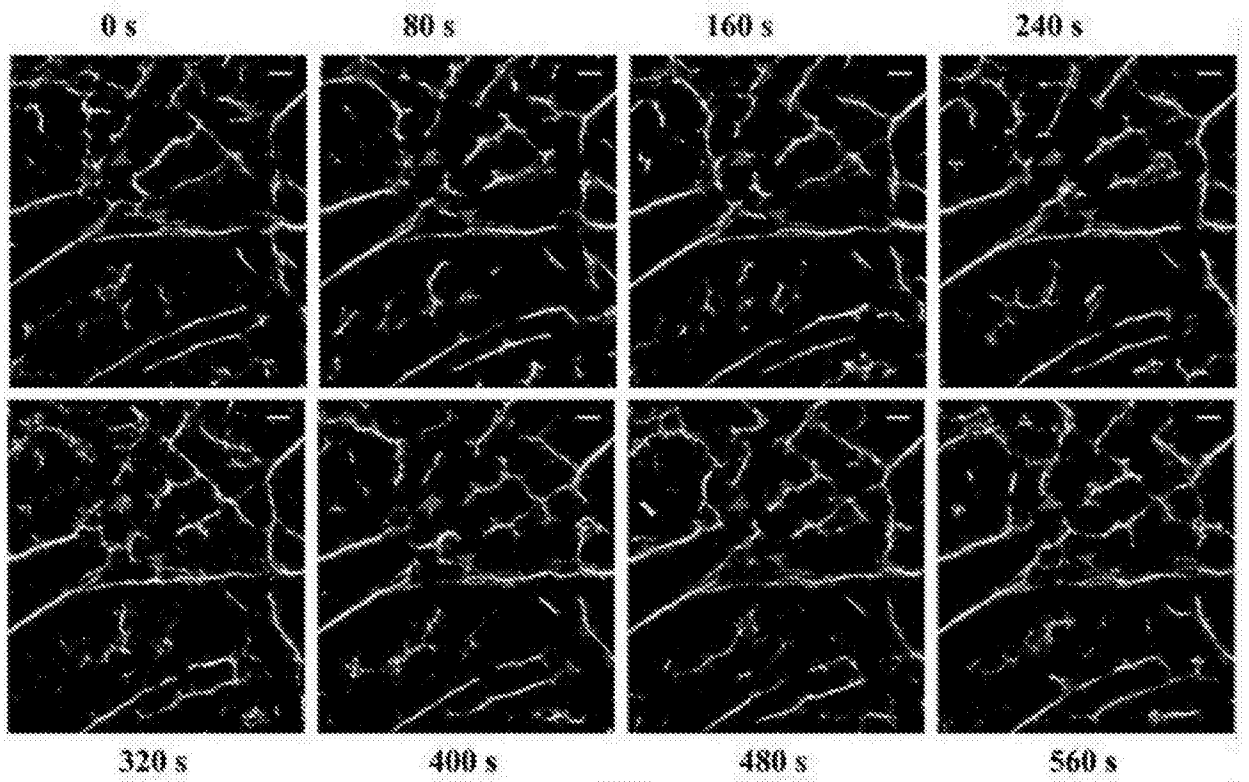
图7



(a)



(b)



(c)

图8