

基于压缩感知的高密度分子定位算法比较

张赛文^{1,2,3}, 于斌^{1,2,3*}, 陈丹妮^{1,2,3}, 吴晶晶^{1,2,3}, 李四维^{1,2,3}, 屈军乐^{1,2,3}

¹深圳大学光电子器件与系统教育部/广东省重点实验室, 广东 深圳 518060;

²深圳大学光电工程学院, 广东 深圳 518060;

³深圳生物医学工程重点实验室, 广东 深圳 518060

摘要 为了提高荧光超分辨显微技术的时间分辨率, 提出了各种高密度荧光分子定位算法。对基于压缩感知模型的凸优化(CVX)工具箱的内点算法、同伦算法以及正交匹配追踪(OMP)算法的重构密度、定位精度、定位时间进行比较。模拟结果表明, CVX方法和同伦算法能够在高密度情况下实现精确定位, OMP算法与同伦算法运行时间比 CVX 算法短, 但 OMP 算法在高密度的情况下定位精度相对其他 2 种算法明显下降。实验结果表明, 3 种算法都能实现高密度的荧光分子定位, CVX 方法和同伦算法具有较好的重构效果; 在 500 幅图像重构中, 同伦算法与 OMP 算法的速度相比于 CVX 算法分别提高了 14.9 倍和 21.2 倍, 大幅度缩短了重构时间。

关键词 生物医学光子学; 荧光显微镜; 超分辨成像; 压缩感知; 荧光分子定位; 定位算法

中图分类号 O439

文献标识码 A

doi: 10.3788/CJL201845.0307014

Comparison of Algorithms of High-Density Molecule Localization Based on Compressed Sensing

Zhang Saiwen^{1,2,3}, Yu Bin^{1,2,3}, Chen Danni^{1,2,3}, Wu Jingjing^{1,2,3}, Li Siwei^{1,2,3}, Qu Junle^{1,2,3}

¹Key Laboratory of Optoelectronic Devices and Systems of the Ministry of Education and Guangdong Province, Shenzhen University, Shenzhen, Guangdong 518060, China;

²College of Optoelectronic Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China;

³Shenzhen Key Laboratory of Biomedicine Engineering, Shenzhen, Guangdong 518060, China

Abstract In order to improve the time resolution of super-resolution fluorescent microscopy, the methods of high-density molecule localization have been proposed. Three algorithms based on compressed sensing models, including the interior-point method in the CVX toolbox, the homotopy method, and the orthogonal matching pursuit (OMP) algorithm, are investigated. We compare the identified density, localization precision, and execution time by using these algorithms in the simulations and experiments. Simulation results show that the CVX and homotopy methods can accurately locate in the high molecule density, but the CVX method has the longest running time among these methods. The OMP method has low localization precision in the high density. The experimental results show that these algorithms can realize the localization of high molecule density. The CVX and homotopy methods get better results than OMP method in the localization precision. For the localization of 500 images, the homotopy and OMP methods are 14.9-fold and 21.2-fold faster than CVX method, which can greatly shorten the reconstruction time.

Key words biomedical photonics; fluorescent microscopy; super-resolution imaging; compressed sensing; localization for single fluorescent molecule; localization algorithm

OCIS codes 170.2520; 180.2520; 100.3010; 100.6640

收稿日期: 2017-08-07; 收到修改稿日期: 2017-09-14

基金项目: 国家 973 计划(2015CB352005, 2012CB825802)、国家自然科学基金(61335001, 61178080, 61235012, 11004136)、国家重大科学仪器设备开发专项(2012YQ15009203)、广东省自然科学基金(2014A030312008)、深圳市科技计划项目(JCYJ20150324141711698)、国家留学基金

作者简介: 张赛文(1988—), 男, 博士研究生, 主要从事生物医学光子学和压缩感知方面的研究。

E-mail: zhangsaiwen2012@163.com

导师简介: 于斌(1976—), 男, 博士, 教授, 主要从事超分辨成像方面的研究。E-mail: yubin@szu.edu.cn

1 引 言

近年来,为了实现超衍射极限分辨的荧光显微成像,提出了一系列方法,如受激辐射损耗显微(STED)^[1]、(饱和)结构光照明显微[(S)SIM]^[2]、通过荧光分子的开关效应和单分子定位的随机光学重建显微术(STORM)^[3]、光敏定位显微术(PALM)^[3-4]等。STORM利用不同波长的激活光和激发光控制化学荧光分子Cy5在荧光态与暗态之间的转换以实现超分辨成像。将Cy3与Cy5分子对交联到特异的蛋白质抗体上,就可以用抗体来标记细胞的内源蛋白。用特定波长的激光随机地激活荧光探针,然后用另一个波长的激发光来观察,精确定位以及漂白的荧光分子,此过程循环上百次后就可以得到内源蛋白的高分辨率影像。发展快速、高精度的单分子定位算法,是STORM实现超分辨成像的关键技术之一。

为了实现单个荧光分子的纳米尺度定位,研究人员提出了一系列定位算法,包括质心法^[5]、高斯拟合法^[5]、极大似然法^[6]、基于频率域内快速定位算法^[7-8]等。但上述方法都只能实现不存在重叠图像的分子定位,而较低的激发密度导致成像时间过长,只能以牺牲时间分辨率来达到所需的分辨率,这非常不利于活细胞这类运动样品的成像。因此,可通过缩短单幅曝光时间或者减少重构所需画幅数量这两种方法来获得所需的时空分辨率。在一定分辨率的前提下,若采取第1种措施提高时间分辨率,则需要提高荧光分子的光子产生速率;若采取第2种措施则要求每幅图像能够重构出更多的分子数,即提高单幅图像的分子密度。

近两年提出了很多高密度目标分子的定位方法,如DAOSTORM算法^[9]、贝叶斯统计方法^[10]和压缩感知算法^[11-12]。其中,对压缩感知算法的研究最活跃,该算法广泛应用于高密度分子图像的二维和三维定位^[13-16]。压缩感知的重构算法非常多,为了分析不同重构算法的重构精度以及重构时间,本文比较了STORM高密度图像定位领域最常见的3种压缩感知重构算法:基于凸优化(CVX)工具箱的内点算法、基于贪婪思想的正交匹配追踪(OMP)算法和同伦(homotopy)算法。从定位精度、重构密度和重构时间3个方面进行仿真比较,同时对一个序列的高密度实验图像进行重构对比。

2 基本原理

根据荧光显微成像模型,可以模拟产生不同密度分布的荧光分子图像,然后分别用3种算法对产生的图像进行重构定位计算。

2.1 荧光显微成像模型

在荧光显微成像中,单个荧光分子图像的信号分布,即系统的点扩展函数,可以近似用二维高斯函数进行描述^[17]:

$$f(x, y) = \frac{p}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (1)$$

式中, p 为单个分子辐射的光子数, σ 为高斯函数的标准差,其大小与参考显微系统的关系为

$$\sigma = \frac{0.21 \times \lambda}{NA}, \quad (2)$$

其中, λ 为荧光分子辐射波长, NA 为光学系统的数值孔径。

在STORM成像中,显微系统有线性移不变性,因此, K 个荧光分子的探测图像可以表示为

$$I(x, y) = \sum_{i=1}^K f(x - x_i, y - y_i), \quad (3)$$

式中, (x_i, y_i) 为第 i 个分子所在的位置。在成像过程中,散射体、溶液、盖玻片等产生均匀背景噪声,光信号在探测器上引入散粒噪声,所以任意像素处接收到 Q 个光子的概率为

$$P = \frac{\exp[-(I + b_g)](I + b_g)^Q}{Q!}, \quad (4)$$

式中 b_g 为背景噪声。由分子和背景得到的光子数服从泊松分布。

2.2 压缩感知模型的3种重构算法

如果一个信号 $\mathbf{X} \in \mathbf{R}^N$ 是稀疏的或者通过某种变换可以稀疏表示,通过一个矩阵 $\mathbf{A} \in \mathbf{R}^{M \times N}$ 对其进行线性测量得到观测向量 $\mathbf{Y} \in \mathbf{R}^M$,其中 $M \ll N$ 。在矩阵 \mathbf{A} 和向量 \mathbf{Y} 已知的情况下,求解的最小化模型为^[11]

$$\min \|\mathbf{X}\|_1, \text{ s.t. } \mathbf{Y} = \mathbf{A}\mathbf{X}, \quad (5)$$

式中s. t.代表约束条件。

在荧光显微成像中, \mathbf{Y} 为由探测图像 \mathbf{I} 排成的列向量, \mathbf{A} 为由成像系统的点扩展函数构造成的矩阵,向量 \mathbf{X} 为定位的超分辨图像排成的列向量,且 \mathbf{X} 可以通过向量转换成超分辨图像矩阵。由于在实际成像过程中,探测器的像素尺寸不可能无限小,并且分子经过光学系统后成一个高斯斑,因此需要定位的超分辨图像尺寸要比探测器的小,这就相当

于在探测图像中依次进行插值和定位。被定位分子稀疏分布在样品中,满足信号稀疏性的要求,可以进行定位。

2.2.1 基于 CVX 工具箱的内点法

在实际成像过程中,光子散粒噪声服从泊松分布,可以将(5)式的约束条件模型修改为

$$\min \| \mathbf{X} \|_1, \text{ s.t. } \| \mathbf{A}\mathbf{X} - \mathbf{Y} \|_2 \leq \chi, \quad (6)$$

式中, $\chi = \epsilon \left(\sum Y_j \right)^{1/2}$, ϵ 为根据实际图像进行设置的一个常数。在单分子定位显微中, \mathbf{X} 是非负的,因此,(6)式可改写为

$$\min(\mathbf{c}^T \mathbf{X}), \text{ s.t. } X_i \geq 0 \text{ and } \| \mathbf{A}\mathbf{X} - \mathbf{Y} \|_2 \leq \chi, \quad (7)$$

最小化目标函数值 $\mathbf{c}^T \mathbf{X}$ 相当于一个加权的 L_1 范数,其中的 c_i 是矩阵 \mathbf{A} 的第 i 列的和,相当于在 i 的位置分子对图像产生的总强度,可通过在 MATLAB 软件里使用 CVX 工具箱求解以上最优化问题。

2.2.2 同伦算法

相对于(6)式,通常可以转换为无约束最优化问题:

$$\min(\| \mathbf{A}\mathbf{X} - \mathbf{Y} \|_2 + \lambda \| \mathbf{X} \|_1), \quad (8)$$

其中 λ 的值控制着稀疏性与求解误差之间的平衡。有许多算法可以求解(8)式中的最优化模型,如梯度投影法(GP)^[18-19]、近端梯度法^[20]、增广拉格朗日乘法^[21]、快速软阈值迭代算法(FISTA)^[22]、同伦算法^[23]等。在 STORM 数据分析中,可从图像数据直接计算出(6)式中约束条件的误差限,但对于参数 χ 在(6)式中产生的解,(8)式却很难得到同样解的参数 λ 值,这是因为 λ 与 χ 对一般问题不存在解析关系。因此,要用于 STORM 数据分析的算法必须首先确定合适的 λ 值。在所列举的算法中,同伦算法对于找到合适的 λ 值特别有效。因为同伦算法中 λ 的初始值很大,随着 λ 逐渐减小, $\| \mathbf{A}\mathbf{X} - \mathbf{Y} \|_2$ 方差越来越小,当方差值小于误差限 χ 时,取这 2 个 λ 的中值作为最终值,(8)式的解即为(6)式的近似解。

2.2.3 OMP 算法

匹配追踪类算法是将观测矩阵中部分列向量线性组合得到观测信号或者图像的算法。匹配追踪(MP)算法是依次选择与残差最匹配的原子(即列向量)来得到稀疏解向量的方法。但是由于观测残差仅与当前选择的原子正交,所以其收敛速度慢。OMP 算法可以保证每次迭代后的残差与所选择的原子都正交,并且可以保证每次都不重复选择原子,收敛速度快。

OMP 算法的步骤修改为如下^[24]。

- 1) 初始化参数为:残差 $\mathbf{r}_0 = \mathbf{Y}$, 被选原子索引集 $s = \emptyset$ 即为空集,迭代次数 $i = 0$,解初始化为 $\mathbf{X} = \mathbf{0}$;
- 2) 当 $\| \mathbf{r}_i \|_2 \geq \chi$ 并且 $\max(\mathbf{A}^T \mathbf{r}_i) > 0$ 时,开始迭代;
- 3) 找到 $\max(\mathbf{A}^T \mathbf{r}_i)$ 的索引并添加集合 s ;
- 4) 通过 $\arg \min_{\mathbf{X}_s > 0} \| \mathbf{Y} - \mathbf{A}_s \mathbf{X}_s \|_2$ 最小化目标函数求得解 \mathbf{X}_s ;
- 5) 通过 $\mathbf{Y} - \mathbf{A}_s \mathbf{X}_s$ 求得残差 \mathbf{r}_{i+1} , 并且 $i = i + 1$, 返回第 2) 步;
- 6) 当不满足第 2) 步条件时,停止迭代, $\mathbf{X} |_s = \mathbf{X}_s$ 即为最后所求的解。

3 模拟仿真及实验结果

为了比较上述 3 种算法,分别进行了模拟仿真和真实实验图像的重构。模拟仿真分别从重构密度、定位时间及定位精度 3 个方面进行对比。对 500 幅实验图像进行了重构,分析其定位精度与重构时间。

3.1 单分子定位算法模拟仿真

为了与实验设备相一致,模拟参数设置为:物镜的数值孔径 NA 为 1.3,发射波长为 690 nm,像元尺寸为 $100 \text{ nm} \times 100 \text{ nm}$,像元数为 $32 \text{ pixel} \times 32 \text{ pixel}$ 。为了避免边缘影响,让 n 个分子随机分布在中间 $24 \text{ pixel} \times 24 \text{ pixel}$ 区域[图 1(a)]。设置每个分子发射的光子数服从对数正态分布,其峰值为 3000 光子数,标准差为 1700 光子数。根据(1)~(3)式生成模拟图像。对生成图像的每个像素添加 70 个光子数的均匀背景噪声,相当于目前使用的电子倍增电荷耦合器件(EMCCD)参数,并对图像添加泊松化噪声,如图 1(b)所示。为了减少测量矩阵过大而导致重构时间过长和存储内存过大,对图像进行分块处理,分块方法参考文献[11]。采用不同算法对图像进行重构,图 1(c)所示为 CVX 方法定位的结果,然后利用定位分子与原始分子位置计算分子重构密度与定位精度。根据文献[25]的补充材料计算分子重构密度,其中误差圆盘取为 10 pixel (125 nm);而定位精度采用的方法如下:在不同分子密度情况下,一共统计 2000 个分子,统计每一个原始分子位置与定位分子位置的偏差,然后用高斯函数拟合偏差曲线,得到如图 2 所示的高斯函数的标准差,即定位精度。图 3(a)所示为不同算法的识别密度随分子密度的变化曲线,可以看出,随着分子密度的增加,3 种算法识别密度均增大,但是恢复率变差,同伦算法与 OMP 算法的增加的速度较慢。

图 3(b)所示为 3 种算法的定位精度随分子密度的变化曲线,其中定位精度用高斯拟合函数的标准差来表示,可以看出,随着分子密度的增加,3 种算法的定位精度都越来越低,OMP 算法的定位精度随着分子密度的增加显著降低,即其他 2 种算法的定

位精度相对较好。图 3(c)所示为 3 种算法的运行时间随分子密度的变化曲线,同伦算法和 OMP 算法的重构时间很短, CVX 内点算法对一幅 $32 \text{ pixel} \times 32 \text{ pixel}$ 图像的重构时间比其他 2 种算法高出一个量级。

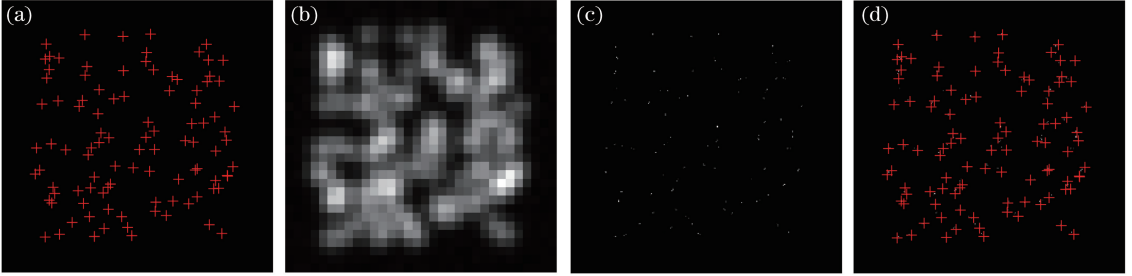


图 1 分子图像定位重构。(a)原始分子位置;(b)系统成像;(c)定位重构图像;(d)原始分子与定位重构分子的位置对比图像
Fig. 1 Reconstructed images of simulated samples. (a) Distribution of original molecules; (b) system imaging; (c) reconstructed image; (d) position comparison between original and reconstructed molecules

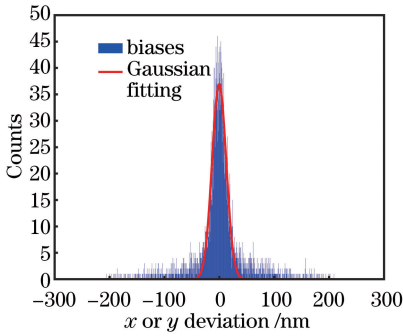


图 2 重构分子的定位精度

Fig. 2 Localization precision of reconstructed molecules

3.2 高密度序列实验图像重构

为了验证算法的有效性,从国外公开的网站^[26]上获取 500 幅高密度实验微管图像数据,其光学系统参数与 3.1 节的模拟参数相一致。3 种算法对 500 幅图像重构结果如图 4 所示,图 4(a)是由 500 幅图像直接累加而得到,图 4(b)~(d)分别由 CVX

内点法、OMP 算法、同伦算法重构之后累加得到。从重构的结果可以看出, CVX 内点算法重构的图像更加均匀、无断线,其重构效果最佳,表明其重构的分子密度和定位精度更高。从 OMP 算法重构结果可以看出,微管旁边增加了一些分子,说明 OMP 算法的定位精度比其他 2 种算法差,定位分子偏离了真实分子的位置。

相比直接累加的原始图像,3 种算法对高密度定位重构图像都具有很好的效果。图 5 是在 3 种算法重构的结果中取如图 4 所示的白色直线,从对比的结果可以看出, OMP 算法会多出一些不正确的点,如图 5 中椭圆所示,重构结果相对其他 2 种算法稍差。对 500 幅实验图像进行重构,3 种算法运行时间的对比如下: CVX 内点法的运行时间为 131930.45 s(约 36.6 h); OMP 算法的运行时间为 8857.50 s(约 2.4 h);同伦算法的运行时间为 6229.35 s(约 1.7 h)。从 3 种算法的运行时间可以

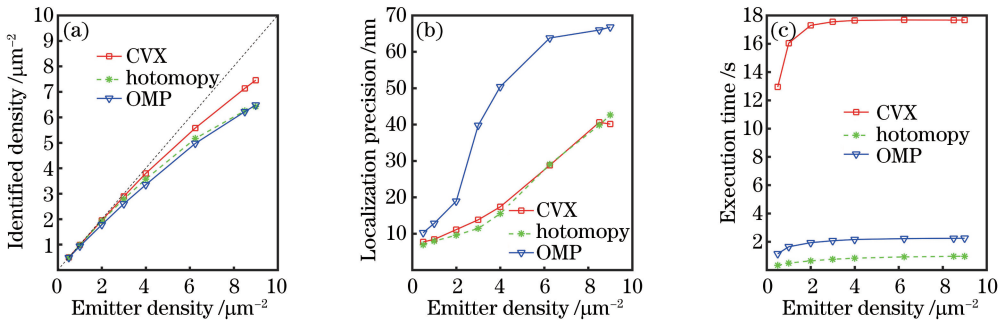


图 3 三种算法的(a)识别密度,(b)定位精度以及(c)执行时间随分子密度的变化关系

Fig. 3 (a) Identified density, (b) localization precision and (c) execution time versus emitter density by using three different algorithms

看出,对 500 幅图像进行重构,CVX 算法所需要的时间为 OMP 算法的 14.9 倍,为同伦算法的 21.2

倍。随着图像的增大和序列的增加,同伦算法和 OMP 算法的时间优势更明显。

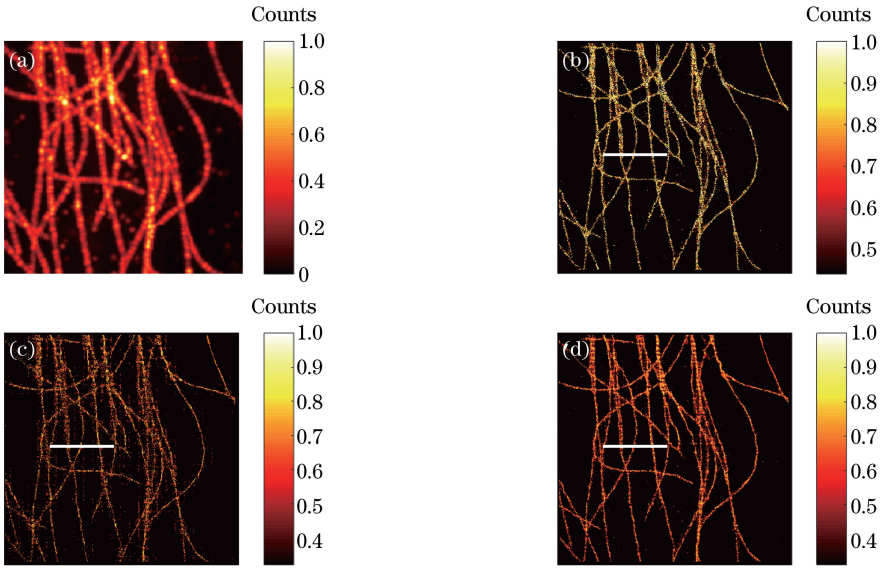


图 4 3 种算法的重构图像。(a)直接累加的 500 幅图像;(b) CVX 内点法;(c) OMP 算法;(d) Homotopy 算法
Fig. 4 Reconstructed images of three algorithms. (a) Accumulative image from 500 original images; (b) CVX interior-point method; (c) OMP algorithm; (d) homotopy method

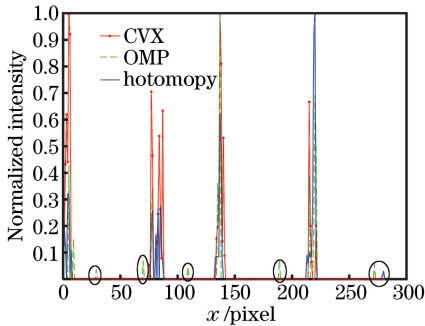


图 5 图 4 白线的定位结果

Fig. 5 Reconstructed results along the white lines in Fig. 4

4 结 论

在高密度单分子定位超分辨成像中,对基于压缩感知模型中 3 种算法进行了模拟与实验的比较。基于 CVX 工具箱的内点算法的定位精度最高,但是其运行时间太长;OMP 算法与同伦算法能够快速实现高密度的单分子定位,但其重构密度相对较差。上述 3 种算法都能够实现高密度的单分子定位,在图像尺寸较小、图像帧数不多的情况下,若想获得高精度的重构结果,可以使用 CVX 内点法。对于上千幅图像的重构,使用同伦算法可以节省运行时间,也可以获得较好的重构效果。为了进一步提高算法的重构精度,可以对观测矩阵进行优化,提高观测矩阵的非相关性。

参 考 文 献

- [1] Hell S W, Wichmann J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy [J]. *Optics Letters*, 1994, 19(11): 780-782.
- [2] Gustafsson M G. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution [J]. *Proceedings of National Academy Sciences of the United States of America*, 2005, 102(37): 13081-13086.
- [3] Rust M J, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM) [J]. *Nature Methods*, 2006, 3(10): 793-796.
- [4] Hess S T, Girirajan T P K, Mason M D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy [J]. *Biophysical Journal*, 2006, 91(11): 4258-4272.
- [5] Thompson R E, Larson D R, Webb W W. Precise nanometer localization analysis for individual fluorescent probes [J]. *Biophysical Journal*, 2002, 82(5): 2775-2783.
- [6] Aguet F, van de Ville D, Unser M. A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles [J]. *Optics Express*, 2005, 13(26): 10503-10522.
- [7] Yu B, Chen D N, Liu L, *et al.* Fast Fourier-domain

- localization algorithm of single molecule with nanometer resolution for super-resolution fluorescence imaging[J]. *Acta Optica Sinica*, 2012, 32(2): 0218001.
- 于斌, 陈丹妮, 刘磊, 等. 荧光单分子的频率域纳米级快速定位算法及其在超分辨荧光成像中的应用[J]. *光学学报*, 2012, 32(2): 0218001.
- [8] Yu B, Chen D N, Qu J L, *et al.* Fast Fourier domain localization algorithm of a single molecule with nanometer precision[J]. *Optics Letters*, 2011, 36(22): 4317-4319.
- [9] Holden S J, Uphoff S, Kapanidis A N. DAOSTORM: an algorithm for high-density super-resolution microscopy[J]. *Nature Methods*, 2011, 8(4): 279-280.
- [10] Quan T W, Zhu H Y, Liu X M, *et al.* High-density localization of active molecules using structured sparse model and bayesian information criterion[J]. *Optics Express*, 2011, 19(18): 16963-16974.
- [11] Zhu L, Zhang W, Elnatan D, *et al.* Faster STORM using compressed sensing [J]. *Nature Methods*, 2012, 9(7): 721-723.
- [12] Babcock H P, Moffitt J R, Cao Y L, *et al.* Fast compressed sensing analysis for super-resolution imaging using L1-homotopy [J]. *Optics Express*, 2013, 21(23): 28583-28596.
- [13] Cheng T, Chen D N, Yu B, *et al.* Reconstruction of super-resolution STORM images using compressed sensing based on low-resolution raw images and interpolation[J]. *Biomedical Optics Express*, 2017, 8(5): 2445-2457.
- [14] Zhang S W, Chen D N, Niu H B. 3D localization of high particle density images using sparse recovery [J]. *Applied Optics*, 2015, 54(26): 7859-7864.
- [15] Du Y J, Zhang H, Zhao M Y, *et al.* Faster super-resolution imaging of high density molecules via a cascading algorithm based on compressed sensing[J]. *Optics Express*, 2015, 23(14): 18563-18576.
- [16] Gu L S, Sheng Y, Chen Y, *et al.* High-density 3D single molecular analysis based on compressed sensing [J]. *Biophysical Journal*, 2014, 106(11): 2443-2449.
- [17] Huang J Q, Gumper K, Chi Y J, *et al.* Fast two-dimensional super-resolution image reconstruction algorithm for ultra-high emitter density [J]. *Optics Letters*, 2015, 40(13): 2989-2992.
- [18] Daubechies I, Fornasier M, Loris I. Accelerated projected gradient method for linear inverse problems with sparsity constraints [J]. *Journal of Fourier Analysis and Applications*, 2008, 14(5/6): 764-792.
- [19] Figueiredo M A, Nowak R D, Wright S J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(4): 586-597.
- [20] Parikh N, Boyd S. *Proximal algorithms* [M]. Boston: Now Publishers, 2014, 1(3): 127-239.
- [21] Boyd S, Parikh N, Chu E, *et al.* *Distributed Optimization and Statistical Learning via the alternating direction method of multipliers* [J]. *Foundations and Trends © in Machine Learning*, 2010, 3(1): 1-122.
- [22] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems [J]. *SIAM Journal on Imaging Sciences*, 2009, 2(1): 183-202.
- [23] Donoho D L, Tsaig Y. Fast solution of L1-norm minimization problems when the solution may be sparse [J]. *IEEE Transactions on Information Theory*, 2008, 54(11): 4789-4812.
- [24] Cai T T, Wang L. Orthogonal matching pursuit for sparse signal recovery with noise [J]. *IEEE Transactions on Information Theory*, 2011, 57(7): 4680-4688.
- [25] Hugelier S, de Rooi J J, Bernex R, *et al.* Sparse deconvolution of high-density super-resolution images [J]. *Scientific Reports*, 2016, 6: 21413.
- [26] <http://bigwww.epfl.ch/smlm/>.