High-density localization of active molecules using Structured Sparse Model and Bayesian Information Criterion

Tingwei Quan,[1,2,3] Hongyu Zhu,[1,2] Xiaomao Liu,[4] Yongfeng Liu,[5] Jiuping Ding,[5] Shaoqun Zeng,[1,2] and Zhen-Li Huang[1,2,*]

[1]*Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China*
[2]*Key Laboratory of Biomedical Photonics of Ministry of Education, Huazhong University of Science and Technology, Wuhan 430074, China*
[3]*School of Mathematics and Economics, Hubei University of Education, Wuhan 430205, China*
[4]*School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China*
[5]*Key Laboratory of Molecular Biophysics of Ministry of Education, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*
**leo@mail.hust.edu.cn*

Abstract: Localization-based super-resolution microscopy (or called localization microscopy) rely on repeated imaging and localization of active molecules, and the spatial resolution enhancement of localization microscopy is built upon the sacrifice of its temporal resolution. Developing algorithms for high-density localization of active molecules is a promising approach to increase the speed of localization microscopy. Here we present a new algorithm called SSM_BIC for such purpose. The SSM_BIC combines the advantages of the Structured Sparse Model (SSM) and the Bayesian Information Criterion (BIC). Through simulation and experimental studies, we evaluate systematically the performance between the SSM_BIC and the conventional Sparse algorithm in high-density localization of active molecules. We show that the SSM_BIC is superior in processing single molecule images with weak signal embedded in strong background.

OCIS codes: (180.2520) Fluorescence microscopy; (100.6640) Superresolution; (110.2960) Image analysis.

References and links

1. S. W. Hell, "Far-field optical nanoscopy," Science **316**(5828), 1153–1158 (2007).
2. B. Huang, H. Babcock, and X. Zhuang, "Breaking the diffraction barrier: super-resolution imaging of cells," Cell **143**(7), 1047–1058 (2010).
3. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," Science **313**(5793), 1642–1645 (2006).
4. S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," Biophys. J. **91**(11), 4258–4272 (2006).
5. M. J. Rust, M. Bates, and X. W. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," Nat. Methods **3**(10), 793–796 (2006).
6. G. Patterson, M. Davidson, S. Manley, and J. Lippincott-Schwartz, "Superresolution imaging using single-molecule localization," Annu. Rev. Phys. Chem. **61**(1), 345–367 (2010).
7. S. A. Jones, S. H. Shim, J. He, and X. Zhuang, "Fast, three-dimensional super-resolution imaging of live cells," Nat. Methods **8**(6), 499–505 (2011).
8. T. J. Gould, V. V. Verkhusha, and S. T. Hess, "Imaging biological structures with fluorescence photoactivation localization microscopy," Nat. Protoc. **4**(3), 291–308 (2009).
9. S. J. Holden, S. Uphoff, and A. N. Kapanidis, "DAOSTORM: an algorithm for high- density super-resolution microscopy," Nat. Methods **8**(4), 279–280 (2011).
10. F. Huang, S. L. Schwartz, J. M. Byars, and K. A. Lidke, "Simultaneous multiple-emitter fitting for single molecule super-resolution imaging," Biomed. Opt. Express **2**(5), 1377–1393 (2011).

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

11. C. Hegde, M. F. Duarte, and V. Cevher, "Compressive sensing recovery of spike trains using a structured sparsity model," in Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS), (Saint-Malo, France) (2009), pp. 13–16.
12. K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* 2nd ed. (Springer, 2002).
13. M. Gu, *Advanced Optical Imaging Theory* (Springer, 2000).
14. B. Zhang, J. Zerubia, and J. C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," Appl. Opt. **46**(10), 1819–1829 (2007).
15. M. Elbaum and P. Diament, "SNR in photocounting images of rough objects in partially coherent light," Appl. Opt. **15**(9), 2268–2275 (1976).
16. Z. Gao, Y. Lai, and Z. Hu, "A generalized gradient projection method for optimization problems with equality and inequality constraints about arbitrary initial point," Acta Appl. Math. Sin. **12**(1), 40–49 (1996).
17. D. Mayne and E. Polak, "Feasible directions algorithms for optimization problems with equality and inequality constraints," Math. Program. **11**(1), 67–80 (1976).
18. R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single-molecule microscopy," Biophys. J. **86**(2), 1185–1200 (2004).
19. T. W. Quan, P. C. Li, F. Long, S. Q. Zeng, Q. M. Luo, P. N. Hedde, G. U. Nienhaus, and Z. L. Huang, "Ultra-fast, high-precision image analysis for localization-based super resolution microscopy," Opt. Express **18**(11), 11867–11876 (2010).
20. P. N. Hedde, J. Fuchs, F. Oswald, J. Wiedenmann, and G. U. Nienhaus, "Online image analysis software for photoactivation localization microscopy," Nat. Methods **6**(10), 689–690 (2009).
21. R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," Biophys. J. **82**(5), 2775–2783 (2002).
22. Y. Cheng, "Mean shift, mode seeking, and clustering," IEEE Trans. Pattern Anal. Mach. Intell. **17**(8), 790–799 (1995).
23. S. Wolter, M. Schüttpelz, M. Tscherepanow, S. Van De Linde, M. Heilemann, and M. Sauer, "Real-time computation of subdiffraction-resolution fluorescence images," J. Microsc. **237**(1), 12–22 (2010).
24. A. Sergé, N. Bertaux, H. Rigneault, and D. Marguet, "Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes," Nat. Methods **5**(8), 687–694 (2008).
25. M. P. Gordon, T. Ha, and P. R. Selvin, "Single-molecule high-resolution imaging with photobleaching," Proc. Natl. Acad. Sci. U.S.A. **101**(17), 6462–6465 (2004).
26. X. H. Qu, D. Wu, L. Mets, and N. F. Scherer, "Nanometer-localized multiple single-molecule fluorescence microscopy," Proc. Natl. Acad. Sci. U.S.A. **101**(31), 11298–11303 (2004).
27. S. H. DeCenzo, M. C. DeSantis, and Y. M. Wang, "Single-image separation measurements of two unresolved fluorophores," Opt. Express **18**(16), 16628–16639 (2010).
28. X. Q. Li, G. H. Shi, and Y. D. Zhang, "Time-domain interpolation on graphics processing unit," J. Innovative Opt. Health Sci. **4**(1), 89–95 (2011).
29. C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," Nat. Methods **7**(5), 373–375 (2010).

1. Introduction

Recently, a series of breakthroughs in far-field optical microscopy have been made, which dramatically improve the spatial resolution of conventional fluorescence microscopy by over an order of magnitude in both lateral and axial directions [1,2]. Among them, a wide-field approach called localization microscopy, including photo-activation localization microscopy (PALM) [3], fluorescence photoactivation localization microscopy (fPALM) [4] and stochastic optical reconstruction microscopy (STORM) [5], holds superior capability in live cell imaging.

The operating principle of localization microscopy is based on repeated cycles of imaging and localization of active molecules. Within one cycle, a small fraction of densely distributed molecules are activated and make capable of fluorescence (that is, become active molecules), followed by conventional single molecule imaging steps including excitation, emission and bleaching, finalized with molecule localization procedures which are used to find the localizations of active molecules in single molecule image frames. A final super-resolution fluorescence image could be obtained through repeating this cycle with typically many thousands of times, which depends mainly on the total number of active molecules necessary to be found within a diffraction limited area (or called Airy disk) for a desired spatial resolution [6].

It is well-known that the high spatial resolution of localization microscopy is built upon the sacrifice of its temporal resolution. Therefore, ever since its birth there is a strong need to improve the imaging speed of localization microscopy, while maintaining a desired spatial resolution. Recently, with a tight control on the spectroscopic characteristics of bright

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

fluorescence probes and the use of fast low light detector, Zhuang et al [7] achieved 2D super-resolution imaging at spatial resolution of ~25 nm and temporal resolution as fast as 0.5 s in living cells. This achievement almost reaches the upper speed limit of localization microscopy, which was estimated to be on the order of sub-second [8]. Note that this estimation was based on imaging and localization of sparse distributed active molecules, that is, one or zero molecule within an Airy disk.

If it is possible to localize more than one active molecules inside an Airy disk in each image frame, the total number of repeated imaging cycles could thus be significantly reduced to obtain the same number of localized molecules inside that Airy disk, which is surely effective for enhancing the imaging speed without sacrificing the spatial resolution. In fact, a number of algorithms for high-density localization of active molecules have been developed in the past several years for such purpose. For example, based on multiple Gaussian kernel fitting and residual image analysis, Holden et al designed an algorithm that extremely boosts the density of localized active molecules [9]. While at the same time, Huang et al applied multiple-emitter fitting and graphics processing unit (GPU) computation for fast and high precision localization of active molecules with high density [10]. However, these algorithms were developed for single molecule images with very weak fluorescence background, which was calculated to have signal-background-ratio (SBR) of 3-10 and is not easily obtained from localization microscopy experiments (especially for those without TIRF illumination). Note that the SBR is defined as the ratio of the peak value of signal and the intensity of background.

Here we present an algorithm for processing single molecule images with weak signal embedded in strong background (SBR < 1), which are typical from localization microscopy experiments. This algorithm, called SSM_BIC for the combination use of the Structured Sparse Model [11] and the Bayesian Information Criterion [12], exploits the fact that the number of active molecules in an image with low SBR could be reasonably estimated with high robustness by the Structured Sparse Model (SSM), which has become a major research field in signal processing for sparse image representation, and that a high probability for obtaining an optimal model for describing a low SBR image can be achieved with the Bayesian Information Criterion (BIC), a well-known criterion for model selection with the power of balancing training error and model complexity.

2 Simulation and experimental methods

2.1 The imaging model of multiple active molecules

Based on optical imaging theory, point spread function (PSF) can well describe the imaging model of individual molecules at the focal plane of an optical system [13]. Assuming that multiple active molecules lie in the focal plane, the imaging model of these molecules is the convolution of PSF and the function with respect to the positions of the molecules, which can be represented as

$$L(o_1, o_2, \cdots, o_n) \otimes PSF = F \tag{1}$$

Here, $L(O_1,O_2,...,\ O_n)$ denotes the function that returns the fixed values in the positions $(O_1,O_2,...,\ O_n)$ of n active molecules and zeros otherwise. Considered that PSF can be approached by two dimensional Gaussian function, the theoretical signal $(F_{i,j})$ of pixel (i,j) in detector plane can be written as [14]

$$F_{i,j} = \sum_{l=1}^{n} A_l \exp\left(-\frac{(i-o_{lx})^2 + (j-o_{ly})^2}{2\omega^2} \right) \tag{2}$$

where O_{lx} and O_{ly} represent the x- and y- coordinates of the molecule in position O, A_l is the amplitude, and ω is the width of Gaussian kernel. Taking the influence of noise (including mainly shot noise and background noise) into account, the finally observed signal $(S_{i,j})$ in pixel (i,j) is [15]

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

$$S_{i,j} = Pois(F_{i,j} + N_b) \tag{3}$$

where N_b is the intensity of background noise and $Pois(x)$ is a Poisson random number with mean value of x. The task of high-density localization of active molecules is thus to find the positions of multiple active molecules ($O_1, O_2, ..., O_n$) from the observed signal $S_{i,j}$ (i,j = 1,2,…,m).

2.2 Pre-estimation of the positions of active molecules using SSM

Pre-estimation of the positions of active molecules in an extracted sub-region (m × m pixels) is a key step in SSM_BIC. Assuming that if an active molecule lies in pixel (i,j), the corresponding observed signal $S_{i,j}$ should be larger than a threshold value, and that the distribution of active molecules is still sparse in sub-pixel scale. Then an SSM [11] can be obtained

$$\begin{aligned} & e^* = -\arg\min_{e} \sum_{i,j=1}^{m} (L_{i,j})^2 e_{i,j} \\ & subject\ to: \sum_{i,j=1}^{m} e_{i,j} \le n \\ & \qquad e_{i-1,j} + e_{i-1,j+1} + e_{i,j-1} + e_{i,j} + e_{i,j+1} \le 1 \quad (i, j = 2,3,\cdots,m-1) \end{aligned} \tag{4}$$

Here, $e_{i,j}$ is an indicator function whose value is 1 if there is one molecule in pixel (i,j) and 0 if there is no molecule. L_{ij} is determined by the observed signal $S_{i,j}$ with the following procedures. $S_{i,j}$ is firstly subtracted by the associated background (the average intensity in the edge of the extracted sub-region), then normalized by dividing by its maximum value (after background subtraction) to give $L_{i,j}$. If $L_{i,j}$ is larger than a given threshold value, it keeps unchanged and zeros otherwise. In optimization problem (4), the inequality constraints describe the sparse distribution characteristic of active molecules. Specifically, the first inequality constraint indicates that the total number of active molecules does not exceed n, while the following inequality constraints show that there is only one active molecule at most in a pixel and its neighborhood. Furthermore, the inequality constraints in optimization problem (4) guarantee that, in most cases, the optimal solver e^* could be obtained using common optimization algorithm [16,17]. For some extreme cases, an auxiliary branch and bound procedure should be also used. According to the definition of $e_{i,j}$, pre-estimated positions of active molecules are given by the optimal solver e^*.

Note that the threshold value and the maximum number of particles n can effectively control the sparsity of pre-estimated molecules in a sub-region, and thus a reasonable pre-estimation of the positions of active molecules is beneficial for the procedures in the following sections. Considering the size of Airy disk, the density of molecules, and pixel size, the threshold value and n were set to be 0.4 and 6, respectively, unless otherwise specified.

2.3 The selection of optimal model using BIC

Since the pre-estimation in Section 2.2 is rough, the pre-estimated positions of active molecules usually contain some false-positive positions and may have pixel-level errors compared to the corresponding real positions. To eliminate the false-positive positions and improve localization precision, BIC is introduced for the selection of an optimal model that describes the observed signal $S_{i,j}$ (i,j = 1,2,…,m). A detailed description is presented as follows.

If the number of pre-estimated positions is assumed to be s, we can obtain these sub-pixel positions using the maximum likelihood estimation [18]

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

$$
\begin{aligned}
&(\hat{A}_1, \hat{o}_{1x}, \hat{o}_{1y}; \hat{A}_2, \hat{o}_{2x}, \hat{o}_{2y};, \cdots, \hat{N}_b) \\
&= \arg\min\left(\sum_{1 \le i,j \le m} F_{i,j} + N_b \right) - \left(\sum_{1 \le i,j \le m} S_{i,j} \log(F_{i,j} + N_b) \right) \qquad (5) \\
&subject\ \ to: \ A_i > 0, o_{ix} > 0, o_{iy} > 0, (i = 1, 2, \cdots, s); N_b > 0
\end{aligned}
$$

Here the Airy disk is fixed, i.e., the width of Gaussian kernel ω in Eq. (2) is unchanged. The symbols in Eq. (5) have the same meanings as those in Eqs. (2) and (3). Equation (5) can be solved using the optimization algorithm reported in [16,17]. The amplitude and the posizition of each molecule and the intensity of background noise were ascertained and then substituted into the function $F_{i,j}$ + N_b. Then, we can approximately calculate the corresponding BIC statistics [12]

$$
BIC(s) = \sum_{i,j=1} (S_{i,j} - \hat{F}_{i,j} - \hat{N}_b)^2 / S_{i,j} + (\log m^2)(3s+1) \qquad (6)
$$

In a total number of s molecules, the molecule with minimum fitting amplitude A_i is removed and a new model describing the remaining s-1 molecules is generated. This new model is used to fit the observed signal $S_{i,j}$ (i,j = 1,2,…,m), and the statistics BIC(s-1) is calculated again. In this way, a series of statistics BIC(s-2),…, BIC(1) can be obtained by repeating the above procedures. The model with minimum value of BIC*(s*)* is defined to be the optimal model.

Using the optimal model, the observed signal $S_{i,j}$ is fitted to give positions of the active molecules, which are regarded as their real positions. Because the value of amplitude A_i is typically one or two orders of magnitude larger than the position parameters of active molecules in a sub-region, it is difficult to search for the optimal solver through iteration method. Therefore in the fitting the position parameters of active molecules were enlarged ten times and then recovered afterwards.

2.4 The extraction of sub-region

This step includes de-noising and sub-region extraction. For de-noising, an original image is convoluted three times with an averaging template containing 3 × 3 pixels, and then convoluted with an annular averaging template (5 × 5 pixels) in which the values in inner part (3 × 3 pixels) are set to be zero and the values in marginal part to be −1/16, respectively. Here the triple application of the 3x3 averaging template is used to effectively collect the signal energy of active molecules which are distributed within a distance of 1-2 Airy disk radius, and the annular averaging process is used to reduce the influence of background intensity difference in the entire sub-region extraction process [19,20]. In the de-noised image, pixels with local maximum values are detected by threshold method, where the threshold value is set to be more than 5 times the standard deviation of background. Taking the coordinates of the detected pixels as the centers, sub-regions with a size of 9 × 9 pixels are extracted. Note that the sizes of the convolution template and the sub-region are selected according to the fact that Airy disk typically occupies 5 × 5 pixels for a balance on localization precision and signal intensity [21].

After the sub-region extraction, active molecules are localized simultaneously through the procedures described in Sections 2.2 and 2.3. However, the size of sub-region may be too large for one active molecule and may be too small for multiple active molecules. Therefore, the positions of the active molecules in each image, which were estimated by Eqs. (5) and (6), are automatically divided into multiple groups by a cluster procedure basing on a mean-shift algorithm [22], In the process of clustering, the kernel width should be set to ensure that any molecule pairs within a distance of less than 4 × Airy disk radius can be attributed to the same group. Then another round of sub-region subtraction is performed to guarantee that the Airy disk of each group of active molecules can be covered completely. Finally, the positions of active molecules can be ascertained using the procedures described in Section 2.3.

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

The complete procedures for the SSM_BIC algorithm are summarized as follows.

Step 1. De-noising and sub-region extraction.

Step 2. Pre-estimation of the positions of active molecules using SSM.

Step 3. Select the optimal model by BIC and search for the positions of active molecules.

Step 4. Cluster the positions of active molecules with mean-shift algorithm for another round of sub-region extraction.

Step 5. Repeat *Step 3* with the refined sub-regions found in *Step 4*.

2.5 Localization microscopy imaging

2.5.1 Cell culture and plasmid transfection

HEK293T cells were cultured on 25-mm-diameter No. 1.5 cover-glass-bottom dishes (Electron Microscopy Sciences, Cat. # 72225-01) at 37 °C with 5% CO_2 and grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen). Transfection of the eukaryotic expression vector for actin bundles labeling, pcDNA3-lifeact-d2EosFP, was carried out using Lipofectamine LTX (Invitrogen) according to the manufacturer's instructions. Cells were maintained for 24 h after transfection in culture medium and then cleaned and fixed for further single molecule imaging.

2.5.2 Single molecule imaging

Single molecule imaging experiments were performed with a home-built TIRF microscope setup consisting of an Olympus IX 71 inverted microscope, a 100x/NA1.49 oil immersion TIRF objective (UAPON 100XOTIRF, Olympus), a 405 nm laser diode and a 561 nm diode-pumped solid-state laser (both from CNILaser, China), and an Andor iXon 897 EMCCD camera. A dichroic mirror (Di01-R488/561, Semrock) and a longpass filter (BLP01-561R-25, Semrock) were used to separate the collected fluorescence from scattering laser and impurity fluorescence. Data were acquired by the software provided by the camera manufacturer. Signal in single molecule images were converted from gray level into photons with the camera conversion factor provided by the camera manufacture. Sample drift was not corrected. The pixel size at sample plane is 160 nm and an Airy disk covers 3 × 3 pixels.

3. Results

Simulation and experimental data sets were used to evaluate the performance of SSM_BIC in high-density localization. For all simulations, the Airy disk radius is 250 nm, the pixel size at sample plane is 100 nm and the background intensity is 100 photons per pixel, and we assume that all emitted fluorescence photons from individual molecules can be detected completely by detector. The performance of SSM_BIC was measured by false-positive rate, detection rate and localization precision. The false-positive rate is defined as the ratio of false-positive positions to suspected positions found by the localization algorithm. The detection rate is defined as the ratio of real positive positions from suspected positions to true positions of the simulated molecules. The localization precision (σ) represents the average of the absolute error between true positions and real positive positions. Suspected position is recognized as real positive position providing that the distance between the suspected and true positions is less than 50 nm, and false-positive positions otherwise. Comparison on the localization performance between SSM_BIC and Sparse Algorithm (SA) was also performed. Note that SA is based on the assumption that only one active molecule lies inside an Airy disk [19,20,23], and the SA used in this paper was described previously in detail [19].

3.1 Analysis of simulated images

To evaluate the applicability of SSM_BIC algorithm in high-density localization, 10 simulated image frames (100 × 100 pixels) with 20-180 molecules in each frame were

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

generated according to the description in Section 2.1. The molecules were immobilized and placed across the image uniformly and randomly. Key steps of SSM_BIC are shown in Fig. 1a. Specifically, SSM is used to calculate the pre-estimated positions (denoted by ' × ') from the observed signal (real positions, denoted by 'o') of three representative active molecules. With the pre-estimated positions, four models are produced and the optimal model is determined by selecting the minimum of BICs (marked in red in Fig. 1a). The fitting positions in the optimal model are seen to be fairly close to the real positions.

The simulated images in Fig. 1b show the overall effects of the number of active molecules, signal, and background to image quality. Overlaps of Airy disks hardly occur in the simulated images with 20 active molecules and are obvious for those with 100 or more active molecules. Quantitative image analysis (Fig. 1c) shows that, for a false-positive rate better than 10% and a detection rate higher than 85%, SA can only process images with 20-40 active molecules, while SSM_BIC can process those with much higher molecule density (100-180 active molecules). Clearly, SSM_BIC shows 3-6 folds improvement on the localization density of active molecules than SA and without loss of localization precision.

3.2 Resolving active molecule pairs

To further evaluate the performance of SSM_BIC in high-density localization, the efficiency of SSM_BIC in resolving molecule pairs with various distances was also quantified. A total number of 100 simulated images (9 × 9 pixels) with one molecule pair inside were generated for each computation.

Sample images are shown in Fig. 2a, where each molecule in the pair emits 350 photons. For this kind of images where weak signals are embedded in strong background, it was found that SSM_BIC can resolve molecule pair with a distance of 250 nm, evidencing by a false-positive rate of 8% and a detection rate of 93% (shown by the red line in Fig. 2b). If the signals increase to 500 photons for each molecule, molecule pair with a distance of 200 nm are still resolvable by SSM_BIC with a false-positive rate of 7% and a detection rate of 94% (shown by the black line in Fig. 2b). In contrast, SA is only suitable for the cases where the distance is larger than 350 nm and the signal is much stronger (700 photons), as shown by the dashed black line in Fig. 2b. These results verify the applicability of SSM_BIC in resolving molecule pairs with weak signals and short separate distance (200-250 nm).

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

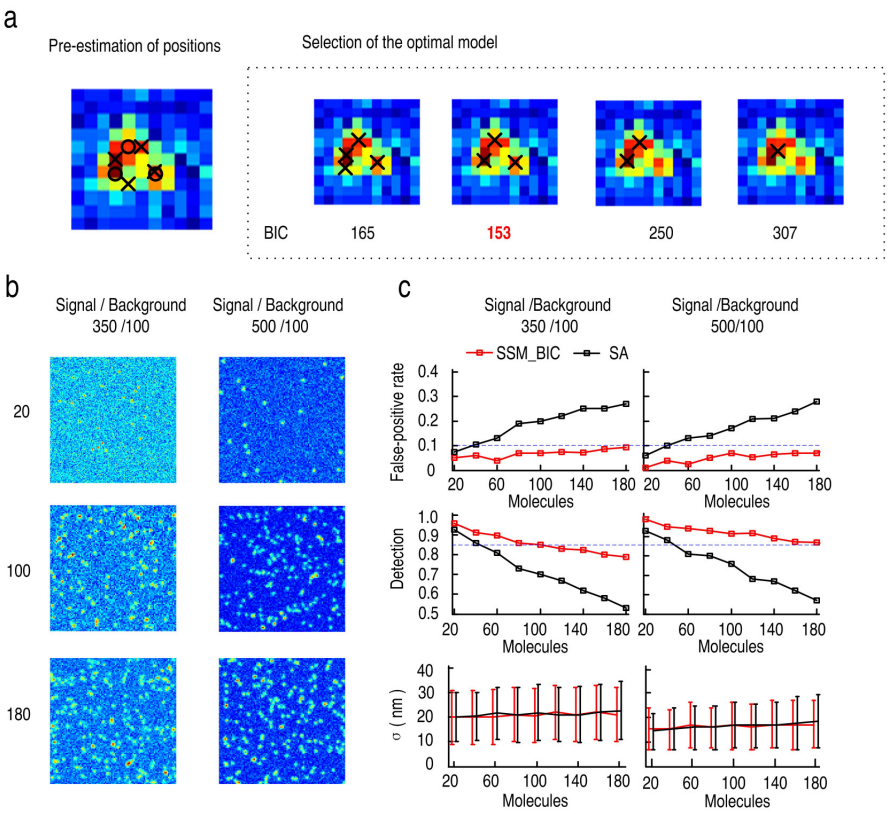

Fig. 1. Localizing simulated images with SSM_BIC and SA. (a) A brief description of the SSM_BIC. Circles and crosses represent the real and pre-estimated positions of active molecules, respectively. (b) Simulated images (100 × 100 pixels) containing active molecules with different densities shown in the left corner of the corresponding images. (c) The performance of SSM_BIC and SA in analyzing images with different number of active molecules.

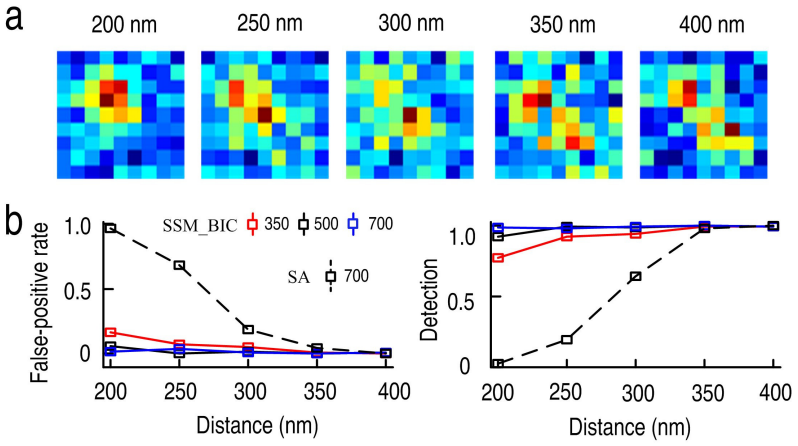

Fig. 2. Capability of SSM_BIC in resolving molecule pairs. (a) Sample images for molecule pairs with different distances, where each molecule emits 350 photons. (b) The performance of SSM_BIC and SA in resolving molecule pairs with different signal intensities and separate distances. Note that the background is 100 photons per pixel for all simulations.

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

3.3 Analysis of experimental data

The applicability of SSM_BIC in processing experimental data was evaluated by analyzing single molecule images of actin bundles in fixed HEK293T cells. A stack of 600 experimental image frames was overlaid to build a TIRF image, where some structures are seen to be labeled densely (as shown by the green arrows in Fig. 3a). Two image processing algorithms, SSM_BIC and SA, were used to process the TIRF image and reconstruct final super-resolution images (Figs. 3b and 3c). It was found that the number of detected molecules by SSM_BIC (~62000) is 2.4 times that by SA (~25000), showing the advantage of SSM_BIC in detecting more active molecules in a fixed number of images. This result indicates that SSM_BIC can be used to reduce the number of image frames for a given spatial resolution, which is surely beneficial for enhancing the image speed of localization microscopy.

Interestingly, some missing structures marked by the green arrows in Fig. 3b are clearly visualized in Fig. 3c, evidencing the power of SSM_BIC in high-density localization. For better understanding on the quality of the experimental single molecule images, Fig. 3b was further analyzed to give histograms in Figs. 3d and 3e. Note that only those molecules with fitting width of Gaussian kernel of less than 140 nm (corresponding to Airy disk radius of 350 nm) were included in the statistics, and that the peak value of signal is equal to ~44% of the total number of detected photons from single molecule when the Airy disk is covered by 3×3 pixels. Thus the average peak value of signal was calculated to be 380 and the average SBR was estimated to be 0.72, showing the weak signal characteristics of the experimental images.

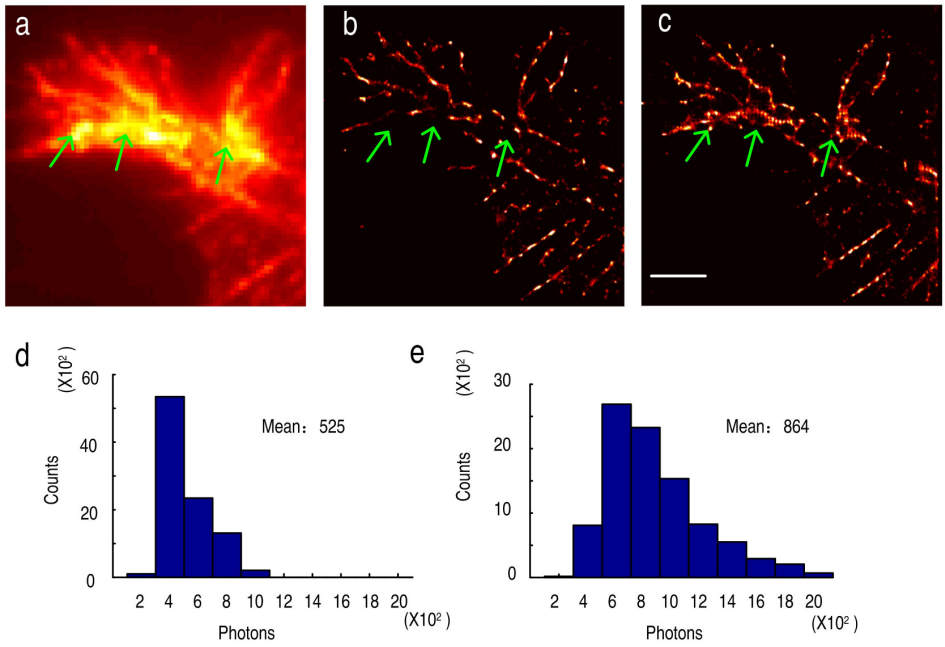

Fig. 3. TIRF image from a stack of 600 image frames (a) and super-resolution images reconstructed by SA (b) and SSM_BIC (c). The green arrows show some densely labeled structures. The data in (b) was further analyzed to give the histograms of background (d) and the total number of detected photons from single molecules (e). The mean values of individual histograms are shown in the right corner of the corresponding figures. Scale bar: 2 μm.

4 Discussions

4.1 The capability of detecting weak signals with SSM_BIC

Pre-estimation of the positions of active molecules and selection of an optimal model are two key components in SSM_BIC. The use of SSM ensures a reasonable estimation to the real

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

number and positions of active molecules (Fig. 1a). Further investigation shows that, if a given threshold is used in the selection of optimal model (see Section 2.2), the detection rate in SSM_BIC is insensitive to the signal intensity, while for Gaussian deflation [9,24] this rate depends strongly on the signal level (Fig. 4a). In fact, it was reported that the threshold in Gaussian deflation should be modified according to signal intensity to ensure detection efficiency [9]. Furthermore, it was found that changes on the threshold plays negligible influence on the detection rate in SSM_BIC (Fig. 4b).

It is well-known that BIC is a powerful tool for the selection of optimal model, since it keeps a balance on training error and model complexity. Although BIC originates from the identification of optimal linear model, it is still effective for optimal nonlinear model [12], which is also supported by our simulation results shown in Fig. 4c. For simulated images with real molecule number of 2, if the pre-estimated molecule number is set to be 4, the identification accuracy using the criterion of minimal training error increases slowly with the signal intensity and approaches the real value at signal levels higher than 2000 photons, while for BIC a satisfactory identification accuracy could be obtained at signal levels as low as 350 photons (Fig. 4c).

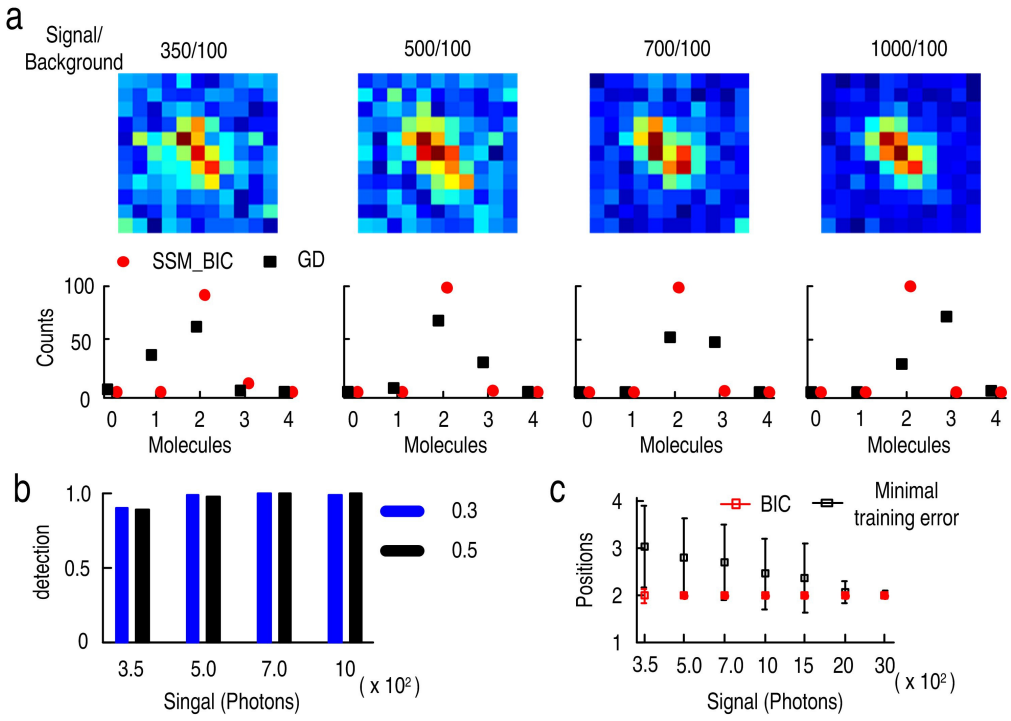

Fig. 4. Dependence of detection rate on threshold and optimal model selection in SSM_BIC. (a, upper) Simulated images containing active molecule pairs with a distance of 250 nm and different signal levels. (a, lower) Detection efficiency of SSM_BIC and Gaussian deflation (GD) using a given threshold of 0.4. (b) Influence of threshold on detection rate in SSM_BIC. (c) Comparison on optimal model selection between BIC and minimal training error.

4.2 Identification of the molecule pairs with strong signal

In the previous simulations with weak signals, the maximum molecule density is limited to 1.8 molecules/μm^2. Processing images with higher molecule density is possible with SSM_BIC algorithm, although in this case it is necessary to increase signal intensity and make some modifications to the optical system. For example, for a molecule pair with 2000 detected photons from each molecule, the resolved distance reaches 125 nm for a pixel size of 50 nm

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

(the blue lines in Fig. 5), which is much better than those with 500 detected photons and 100 nm pixel size (the black lines in Fig. 2b).

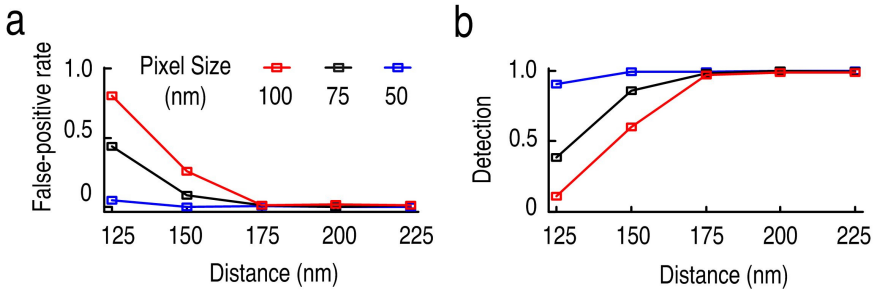

Fig. 5. Performance of SSM_BIC in the identification of molecule pairs with strong signal. Note that the background is 100 photons per pixel for all simulations.

4.3 Identification of the molecule pair with non-equal emission

The capability of SSM_BIC in identifying the molecule pairs with different detected photons was also investigated. In experimental single molecule images, the detected photons from individual molecule in a molecule pair may be different and the ratio could be as big as 2 [25,26]. Simulation results show that the localization performance of SSM_BIC is insensitive to the ratio changes (Fig. 6). This is from the fact that BIC is suitable of identifying multi-model composed with different weights. However, variance analysis [27] can only resolve the molecule pair with equal emission.

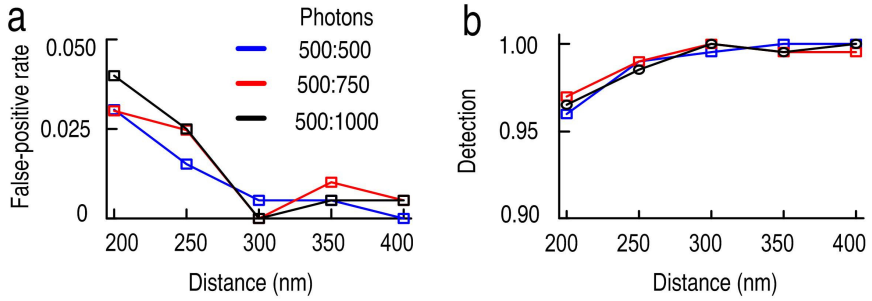

Fig. 6. The performance of SSM_BIC in the identification of the molecule pair with non-equal emission. Note that the background is 100 photons per pixel for all simulations.

4.4 Computation load of SSM_BIC

It was found that SSM_BIC can only process ~2 positions per second when the code is run in Matlab using an Intel i5-750, 2.68 GHz personal computer with 4.0 GB memory. The reasons for this slow speed are from the selection of optimal model and the solving of optimization problem (5) the iteration number of which is as big as 50. A combination of GPU and SSM_BIC can be expected to boost the computation speed with up to several hundred folds [19,28,29].

5. Conclusion

Through both simulation and experimental studies, we demonstrated the applicability of the new SSM_BIC algorithm for high-density localization of active molecules. We found that SSM_BIC has 3-6 folds improvement on the localization density than SA, and that SSM_BIC is capable of resolving the molecule pairs with weak signal (i.e. a total number of 350-500

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011

signal photons from single molecules embedded in a background intensity of 100 photons per pixel) and short separate distance (200-250 nm). We also verified the superior performance of SSM_BIC over SA in processing experimental single molecule images with densely labeled structures. We further discussed the reason on why SSM_BIC could detect weak signal, proved that SSM_BIC can identify the molecule pairs with strong signal or non-equal emission. Although currently our SSM_BIC algorithm runs much slower than SA, we expect significant speed enhancement with future combination use of GPU computation with SSM_BIC, which would surely extend the potentials of SSM_BIC in the field of super-resolution microscopy.

Acknowledgments

This work was supported by National Basic Research Program of China (Grant No. 2011CB910401), the National Natural Science Foundation of China (Grant Nos. 30970691 and 30925013), the Program for New Century Excellent Talents in University of China (Grant No. NCET-10-0407). We thank Prof. Gerd Ulrich Nienhaus and Mr. Per Niklas Hedde from Karlsruhe Institute of Technology for the d2EosFP plasmid.

#149878 - $15.00 USD Received 24 Jun 2011; revised 4 Aug 2011; accepted 11 Aug 2011; published 15 Aug 2011