Fast Maximum Likelihood High-density Low-SNR Super-resolution Localization Microscopy

Kyungsang Kim[1], Junhong Min[1], Lina Carlini[2], Michael Unser[3], Suliana Manley[2], Daejong Jeon[1] and Jong Chul Ye[1]
Bio-Imaging & Signal Processing Lab[1], Laboratory of Experimental Biophysics[2] and Biomedical Imaging Group[3]
KAIST, Republic of Korea[1] and EPFL, Switzerland[2,3]
Email:[kssigari, minimok]@kaist.ac.kr, [lina.carlini, michael.unser, suliana.manley]@epfl.ch and jong.ye@kaist.ac.kr

***Abstract*—Localization microscopy such as STORM/PALM achieves the super-resolution by sparsely activating photo-switchable probes. However, to make the activation sparse enough to obtain reconstruction images using conventional algorithms, only small set of probes need to be activated simultaneously, which limits the temporal resolution. Hence, to improve temporal resolution up to a level of live cell imaging, high-density imaging algorithms that can resolve several overlapping PSFs are required. In this paper, we propose a maximum likelihood algorithm under Poisson noise model for the high-density low-SNR STORM/PALM imaging. Using a sparsity promoting prior with concave-convex procedure (CCCP) optimization algorithm, we achieved high performance reconstructions with fast reconstruction speed of 5 second per frame under high density low SNR imaging conditions. Experimental results using simulated and real live-cell imaging data demonstrate that proposed algorithm is more robust than previous methods in terms of both localization accuracy and molecular recall rate.**

I. Introduction

For the past decades, several innovative methods for surpassing the diffraction limit in far-field optical microscopy have been proposed. It is now well known that their significant resolution improvement was originated from exploiting the optical non-linearity. For example, STED and SSIM can achieve the super-resolution by exploiting the non-linearity of high power illumination, whereas the stochastic optical reconstruction microscopy (STORM) [1] and photo-activated localization microscopy (FALM) [2] exploit non-linearity of photoswitchable fluorescence dyes. Specifically, STORM/FALM rely on sparse fluorophore activations such that fluorophores are sparsely activated in both spatial and temporal domain. When the point spread functions (PSFs) of the activated fluorophore are usually not overlapped, these fluorophores can be localized individually based on the least-squares [1], [2] or the maximum-likelihood [3] PSF fitting. To achieve sparse activation, an accumulation rate of localized fluorophores is, however, limited; so that typically several thousands frames are required to reconstruct a single super-resolution image. In other words, its temporal resolution is on the order of minutes, which allows only limited live-cell imaging.

In order to improve the temporal resolution, one of the possible approaches is high-density imaging. However, in the high-density imaging, many fluorophores are activated at the same time so that there are many overlapping PSFs at each snapshot. There have been several approaches to resolve the overlapping PSFs. For example, DAOSTORM algorithm [4] iteratively fits overlapping spots in a greedy manner. CSSTORM (compressed sensing STORM) [5] and DeconSTROM [6] solve this problem as a sparse recovery among which the latter approach has been demonstrated to be more efficient for high-density imaging in terms of localization accuracy as well as molecular recall rates. For example, in CSSTORM, Gaussian noise model with sparsity constraint is assumed, which is solved by linear programming. As linear programming is computationally expensive, it adopts the local approach in which a reconstructed image is divided into several small-sized blocks processed individually, which potentially degrades the localization accuracy. In DeconSTORM, they use a modified Lucy-Richardson deconvolution in order to utilize Poisson statistics and temporal correlation of activated fluorophores.

In this paper, we present a new localization algorithm for high-density imaging by using a maximum likelihood estimation with a sparsity constraint, which is extremely fast compared to the existing approaches due to perfectly parallelizable structure. Using both simulation and real experiment, we confirmed that the proposed algorithm is especially robust in high-density live-cell imaging at low SNR by low emitted photons from activated fluorophores and high background level.

II. CCCP Framework using Generalized Huber Penalty

A. Problem Formulation

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{r} \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{R}^m$ denote the unknown fluorophore distribution, background fluorescence signals, and detector measurements, respectively; and $A = [a_{ij}]_{i,j=1}^{m,n}$ denotes the probability matrix that a emission photon from a voxel is detected at a detector position. Then, the negative loglikelihood function from Poisson intensity measurement is given by:

$$L(\mathbf{x}) \quad = \quad \mathbf{1}^T(A\mathbf{x}+\mathbf{r}) - \mathbf{y}^T \log(A\mathbf{x}+\mathbf{r}) \qquad (1)$$

where $\mathbf{1}$ denotes a vector with elements of ones with an appropriate size and $\log(A\mathbf{x}+\mathbf{r})$ is treated as element by element operation. Then, our superresolution imaging problem can be formulated as the following minimization problem:

$$\min_{\mathbf{x}} J(\mathbf{x}) \qquad \text{where} \quad J(\mathbf{x}) = L(\mathbf{x}) + pen(\mathbf{x}) \ , \qquad (2)$$

where the penalty function $pen(\mathbf{x})$ imposes a penalty to guide the reconstruction. Note that the optimization problem is not trivial since 1) the gradient of $L(\mathbf{x})$ is non-Lipschitz, and 2) each element of $\mathbf{x}$ should be nonnegative. Another technical difficulties in minimizing $L(\mathbf{x})$ in Eq. (1) is the existence of non-separable term in the likelihood, *i.e.* $\log(A\mathbf{x})$. Quadratic approximation [7] or Anscombe transform [8] was used to make this separable. Recently, for the case of Poisson image deconvolution using total variation (TV) or frame based analysis/synthesis penalty, Figueiredo and Bioucas-Dia [9] proposed so-called PIDAL algorithm using alternating direction of method of multiplier (ADMM) without approximating the Poisson loglikelihood. However, these ADMM algorithm requires huge additional memory to store the Lagrangian parameters to deal with non-separability of loglikelihood and the non-negativeness of $\mathbf{x}$.

To overcome these issues, this paper proposes a new optimization algorithm using the concave-convex procedure [10], which does not need any approximation of the cost function, or to store additional Lagrangian parameters. CCCP is a special case of majorized-minimzation algorithm, which utilizes the Legendre-Fenchel transform as a majorization function.

Legendre-Fenchel Transform of the Penalty: More specifically, as a sparsity inducing penalty, consider the following:

$$\|\mathbf{x}\|_{\mu,p} = \sum_{j=1}^{n} h_{\mu,p}(x_j), \quad 0 < p \leq 1. \tag{3}$$

where the generalized p-Huber function $h_{\mu,p}(t)$ is defined as

$$h_{\mu,p}(t) = \begin{cases} |t|^2/2\mu, & \text{if } |t| < \mu^{1/(2-p)} \\ |t|^p/p - \delta & \text{if } |t| \geq \mu^{1/(2-p)} \end{cases} \tag{4}$$

and where $\delta = (1/p - 1/2)\mu^{p/(2-p)}$ to make the function continuous and differentiable [11]. Note that for $p < 1$ the prior is non-convex. For the generalized p-Huber function in Eq. (4), it is easy to show that $|t|^2/\mu - h_{\mu,p}(t)$ is strictly convex [12]. Therefore, the Legendre-Fenchel transform tells us that there exist $g_{\mu,p}$ such that

$$h_{\mu,p}(t) = \min_{s}\{|s-t|^2/\mu + g_{\mu,p}(s)\} \ . \tag{5}$$

Chartrand [13], [11] showed that $g_{\mu,p}(s)$ is convex when $p = 1$, but in general it is not convex. However, even when $g_{\mu,p}(s)$ is non-convex, $|s|^2/\mu + g_{\mu,p}(s)$ is convex and Eq. (5) becomes a convex minimization problem with respect to s that has a closed form expression for the minimizer given as

$$\begin{aligned} \text{shrink}_p(t,\mu) &:= \arg\min_{s}\{|s-t|^2/\mu + g_{\mu,p}(s)\} \\ &= \max\{0, |t| - \mu|t|^{p-1}\}t/|t| \ . \end{aligned} \tag{6}$$

Here, $p \in [0,1]$ in which $p = 0$ is similar to hard thresholding and $p = 1$ is the same as soft thresholding [11].

Legendre-Fenchel Transform of the Negative Loglikelihood: Note that the negative loglikelihood term for the Poisson noise in Eq. (1) is convex. However, to deal with the existence of non-separable term in the likelihood, we utilize the CCCP with the help of a concave coordinate transform. More specifically, using an appropriate coordinate transform and application of Legendre-Fenchel transform, we can show that

$$\begin{aligned} L(\mathbf{x}) &= \min_{\mathbf{v}} L_c(\mathbf{x}, \mathbf{c}) \\ L_c(\mathbf{x}, \mathbf{c}) &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j + c_{ij}\log\frac{c_{ij}}{a_{ij}x_j} - y_i\log(y_i)\right) \\ &\quad + \sum_{i=1}^{m}\left(r_i + c_i\log\frac{c_i}{r_i}\right) \end{aligned} \tag{7}$$

and $c_i + \sum_{j=1}^{m} c_{ij} = y_i$.

B. Optimization Framework

Now, we have the following minimization problem:

$$\min_{\mathbf{x},\mathbf{c},\mathbf{w}} L_c(\mathbf{x},\mathbf{c}) + \lambda\sum_{j=1}^{n}\left(\frac{1}{\mu}\|x_j - w_j\|^2 + g_{\mu,p}(w_j)\right) \ , \tag{8}$$

1) Minimization with respect to $\mathbf{w}$*:* Using the shrinkage relationship in Eq. (6), the close form solution is given by

$$w_j^{(k+1)} = \text{shrink}_p(x_j^{(k)}, \mu).$$

2) Minimization with respect to $\mathbf{c}$*:* The minimization problem has been studied by Hsiao *et al* [14] using Lagrangian for the constraint $c_i + \sum_{j=1}^{n} c_{ij} = y_i$ and it has been shown that we have the following closed form solution for the constrained optimization problem [15]:

$$c_{ij}^{(k+1)} = \frac{y_i a_{ij} x_j^{(k)}}{\sum_{j'=1}^{n} a_{ij'}x_{j'}^{(k)} + r_i}, \quad c_i^{(k+1)} = \frac{y_i r_i}{\sum_{j'=1}^{n} a_{ij'}x_{j'}^{(k)} + r_i}, \tag{9}$$

3) Minimization with respect to $\mathbf{x}$*:* Finally, for given $\mathbf{c}^{(k+1)}$ and $\mathbf{w}^{(k+1)}$, we can obtain a closed form solution for the update of $\mathbf{x}^{(k+1)}$. More specifically, a fixed point equation of the gradient of the cost function with respect to x_j satisfies the following second order polynomial:

$$0 = \sum_i a_{ij} - \frac{\sum_i c_{ij}^{(k+1)}}{x_j} + \frac{\lambda}{\mu}(x_j - w_j^{(k+1)}) \ , \tag{10}$$

Define

$$d := \frac{\lambda}{\mu}, \quad b_j^{(k+1)} = \left(\sum_i a_{ij}\right) - dw_j^{(k+1)}.$$

Then, the closed form solution is given by

$$x_j^{(k+1)} = \frac{-b_j^{(k+1)} + \sqrt{(b_j^{(k+1)})^2 + 4dx_j^{EM(k+1)}\sum_{i=1}^{m} a_{ij}}}{2d} \tag{11}$$

where $x_{ns}^{EM(k+1)}$ is similar to an ML-EM update given by

$$x_j^{EM(k+1)} = \frac{x_j^{(k)}}{\sum_{i=1}^{m} a_{ij}}\sum_{i=1}^{m}\frac{a_{ij}y_i}{\sum_{j'=1}^{n} a_{ij'}x_{j'}^{(k)} + r_i}. \tag{12}$$

Note that the solution is always non-negative, satisfying the positivity constraint. Moreover, our update equation is a pixel-by-pixel update similar to ML-EM algorithm or Lucy-Richardson method.

4) Advantages of the Proposed Method: Compared to a PI-DAL type algorithm, the proposed method has a unique advantage well-matched to super-resolution localization microscopy. The additional memory requirement for a Lagrangian approach is eliminated. Indeed, the computational complexity and memory usage during the calculation of x_j is similar to the standard Lucy-Richardson deconvolution method. Hence, the algorithm lends itself to a fast GPU implementation thanks to the efficient memory utilization and pixel-by-pixel update. On a intel i7 920 (CPU) and a Tesla C1060 (GPU), our GPU implementation of the proposed method takes only 5 seconds in reconstructing a five over-sampled image of a 128×128 CCD image with 1500 iterations.

III. Experimental Results

We performed experiments using simulated data and real high-density live-imaging PALM data. We compared the following algorithms: the least-square Gaussian fitting[1], CSSTORM[5], FISTA[16] using l1 norm, and the proposed algorithm. In CSSTORM, FISTA, and the proposed algorithm, uniform background is assumed and estimated. In the least-squares method, an elliptical Gaussian PSF to local maxima of the image is fitted.

A. Simluation

In the simulated data, each nanoscale molecule provides a Gaussian PSF of 340nm full width half maximum (FWHM). Emitted photons of the molecules follow the log-normal distribution with mean of 500 and standard deviation of 100. In order to generate low-SNR data, 70 background photons are added to every CCD pixel of 100 nm. In addition, we introduced Poisson shot noise and Gaussian readout noise with unit variance. We generated a data set of a wide range of imaging densities, from $0.2\mu m^{-2}$ to $3.4\mu m^{-2}$. At each density level, flurophores are generated at random locations within 40×40 pixel image, and the total of 30 realizations were used. To quantify the error, each true molecular positions are matched to the closest localized fluorophore within 200nm radius. Then, we calculated the standard deviation of the localization errors and the molecular recall rates.

In all ranges of the density level, the proposed algorithm demonstrated better recall rates than the others. Specifically, the proposed algorithm can identify 10 times more fluorophore molecules than the least-squares method, and improve about 10-30% compared to CSSTROM and FISTA. Moreover, the proposed method is much accurate than CSSTROM & FISTA in terms of localization accuracy. While the least squares method have the smallest localization error, these errors are only for the corrected identified fluoresphores, whose number is significantly smaller compared to others. Therefore, this confirmed that the proposed algorithm is more effective in low-SNR & high-density imaging data than the conventional approaches.

B. Live-cell Super-resolution Imaging

U2OS cells were maintained in Dulbecco's Modified Eagles's Medium (DMEM) (Gibco) supplemented with 10% Fetal Bovine Serum (FBS) (Gibco) in an atmosphere containing 5% CO2 at 37 °C. Cells were cultured and maintained in T-25 flasks and grown to about 70% confluency (corresponding to 2 days) before they were passaged. Prior to staining, cells were washed once with PBS (Sigma). A 200 nM dilution of Mitotracker Red CMH2XRos (Invitrogen) was made in Leibovitz (Invitrogen) and labelled in an inner membrane of the mitochondria. Cells were incubated with the dye for 1520 min at 37 °C in a CO2 atmosphere.

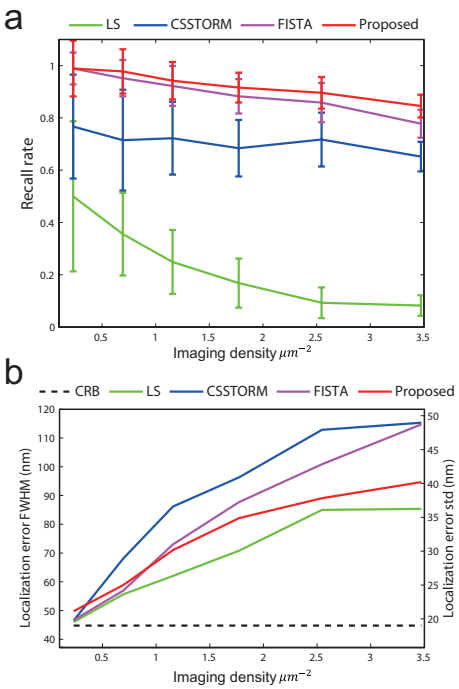

Fig. 1. Simulation analysis on low-SNR STORM/PALM data. (a) Localization error. (b) Identified molecular density. Cramer-Rao bound(CRB) is theoretical minimum accuracy of single molecule localization.

Imaging was performed on an inverted microscope (IX71, Olympus), equipped with an oil-immersion objective (UPlanSAPO 100 x, NA=1.40, Olympus). A 561 nm (Sapphire 561, Coherent) was used to excite Mitotracker Red CMH2XRos and fluorescence was directed onto an electron multiplying CCD camera (iXon+, Andor) with a resulting pixel size of 100 nm. The laser intensity was approximately 3 kW cm $^{-2}$ and an ET605/70 (Chroma) emission filter was used. 4000 frames were collected with a 20 ms exposure time per frame. Using the experimental data, we compared the reconstruction results using the three algorithms (Least-squares, FISTA, and the proposed). The proposed algorithm localized 30 % more molecules than FISTA and 8 times more than the least-squares. In figure 2(b-d), the proposed results show better internal matrix structure and have much clear boundaries of mitochondria than the others. Moreover, the size of the reconstructed mitochondria using FISTA (c) seems to be reduced compared to that of the proposed method (d). In order to observe the dynamic of mitochondria, we created time-lapse images (e,f). Every images in (e,f) were generated from the 1000 consecutive CCD frames for 20 sec and the time-gap

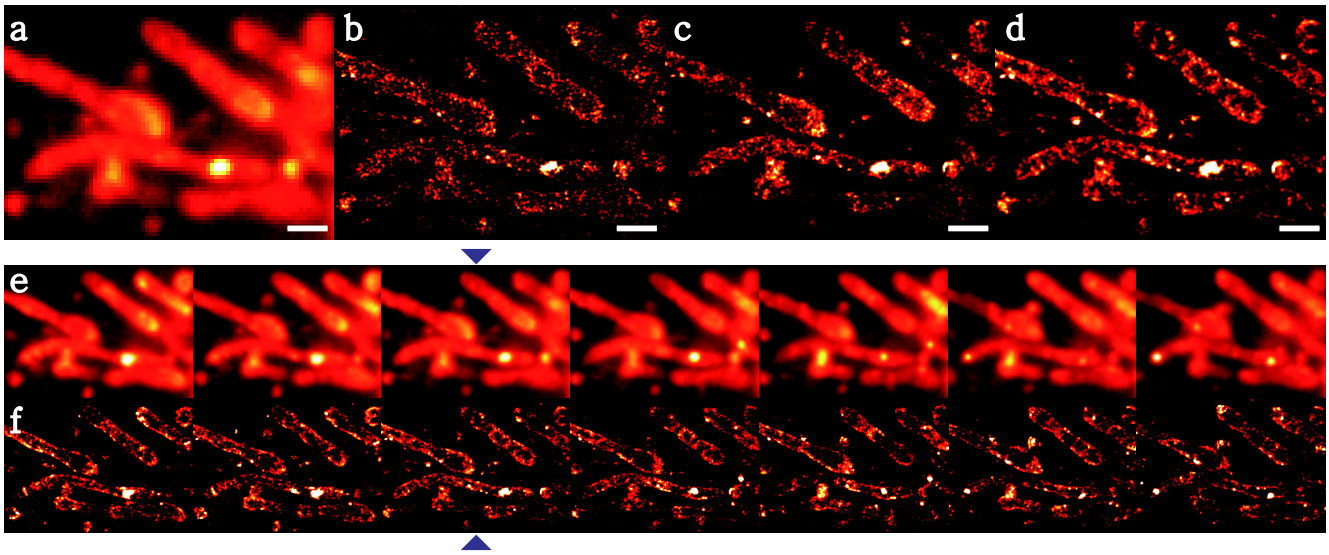

Fig. 2. Live-cell imaging of Mitochondria. Inner membrane of the Mitochondria was labbeled by Mitotracker. (a) Conventional image. (b)Least-square fitting. (c) FISTA decon. (d)Proposed. (e,f) Conventional and Proposed time-lapse images. Every image is generated from consecutive 1000 CCD frames(20sec) and time-gap is 10sec. (a-d) images correspond to blue maker in (e,f). Scale-bar in (a-d) is $1\mu m$.

between successive acquisitions was 10 sec. In the time-lapse images (e-f), we observed slow movements of mitochondria.

IV. Conclusion

We present a new localization algorithm for high-density super-resolution microscopy using the maximum-likelihood estimation of the Poisson noise model with sparsity promoting prior. Using concave-convex procedure, a highly parallelizable algorithm has been derived, which results in a fast GPU implementation with speed of 5 sec per frame. We demonstrated that our algorithm is much effective in low-SNR PALM data over wide range of imaging density in terms of recall rate and localization accuracy. Therefore, we expect that the proposed approach can significantly reduce the number of required CCD frames for super-resolution imaging, which can improve the temporal resolution significantly. Thus, our approach is appropriate for live-cell imaging to investigate biological interactions at the nanometer scale.

Acknowledgment

This research was supported by the National Research Foundation of Korea(NRF) grant funded in part by the Korean government (MEST) (No.2011-0030933) and (No.2012-0000173)

References

[1] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods*, vol. 3, no. 10, pp. 793–796, 2006.

[2] 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*, vol. 313, no. 5793, pp. 1642, 2006.

[3] Carlas S. Smith., Nikolai Joseph., Bernd Rieger., and Keith A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nature Methods*, vol. 7, no. 5, pp. 373–375, 2010.

[4] Seamus J Holden., Stephan Uphoff., and Achillefs N Kapanidis, "DAOSTORM: an algorithm for high-density super-resolution microscopy," *Nature Methods*, vol. 8, no. 4, pp. 279–280, 2011.

[5] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nature Methods*, vol. 9, no. 7, pp. 721–723, 2012.

[6] Eran A Mukamel, Hazen Babcock, and Xiaowei Zhuang, "Statistical deconvolution for superresolution fluorescence microscopy," *Biophysical Journal*, vol. 102, no. 10, pp. 2391–2400, 2012.

[7] Z. Harmany, R. Marcia, and R. Willett, "This is SPIRAL-TAP: sparse Poisson intensity reconstruction algorithms - theory and practice," *IEEE Transactions on Image Processing*, , no. 99, pp. 1084–1096, 2010.

[8] F.X. Dupé, J.M. Fadili, and J.L. Starck, "A proximal iteration for deconvolving Poisson noisy images using sparse representations," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 310–321, 2009.

[9] M.A.T. Figueiredo and J.M. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.

[10] A.L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[11] R. Chartrand, "Nonconvex Splitting for Regularized Low-Rank Sparse Decomposition," *Los Alamos National Laboratory Report: LA-UR-11-11298*, 2012.

[12] S.P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[13] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, pp. 035020.1–035020.14, 2008.

[14] I.T. Hsiao, A. Rangarajan, P. Khurd, and G. Gindi, "An accelerated convergent ordered subsets algorithm for emission tomography," *Physics in Medicine and Biology*, vol. 49, pp. 2145–2156, 2004.

[15] D.P. Bertsekas, "Constrained optimization and Lagrange multiplier methods," *Computer Science and Applied Mathematics*, vol. 1, 1982.

[16] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2419–2434, 2009.