Biomedical Signal Processing and Control 10 (2014) 11–20

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Statistical denoising scheme for single molecule fluorescence microscopic images

Ji Won Yoon*

Korea University, Republic of Korea

ARTICLE INFO

Article history:
Received 13 July 2013
Received in revised form 9 December 2013
Accepted 10 December 2013
Available online 14 January 2014

Keywords:
De-noising
Bayesian
Adaptive prior

ABSTRACT

Single molecule fluorescence microscopy is a powerful technique for uncovering detailed information about biological systems, both *in vitro* and *in vivo*. In such experiments, the inherently low signal to noise ratios mean that accurate algorithms to separate true signal and background noise are essential to generate meaningful results. To this end, we have developed a new and robust method to reduce noise in single molecule fluorescence images by using a Gaussian Markov random field (GMRF) prior in a Bayesian framework. Two different strategies are proposed to build the prior—an intrinsic GMRF, with a stationary relationship between pixels and a heterogeneous intrinsic GMRF, with a differently weighted relationship between pixels classified as molecules and background. Testing with synthetic and real experimental fluorescence images demonstrates that the heterogeneous intrinsic GMRF is superior to other conventional de-noising approaches.

1. Introduction

Using wide-field fluorescence microscopy with single-molecule sensitivity, it is now possible to track the movement of individual fluorophore-tagged molecules such as proteins and lipids in the cell membrane with nanometer precision. This ability has been used to characterize the diffusional properties of molecules [1], as well as monitor their organization relative to other molecules [2] and cellular structures [3], with the potential to work below the diffraction limit of light.

The design of automatic, efficient computer algorithms for data analysis is an extremely important facet of such experiments. Ideally, the program should be able to tolerate the low signal to noise ratios inherent to single-molecule data, must work at high object density and be able to cope with large data volumes at reasonable speed. Suitable methods can usually be decomposed into several sequential steps: filtering of the data, identification of fluorescent objects and their precise positions, and tracking of these objects by linking together their positions over an image sequence. For precise tracking care must be taken to optimize each of these stages.

Previous work in single molecule fluorescence image processing has focused on the tracking step of the problem, with several suggested approaches resulting in excellent performance, including a grid algorithm [4], and several Monte Carlo-based algorithms [5,2]. However, with one exception [5], the initial de-noising step utilized conventional de-noising algorithms, such as Gaussian filtering. Afterwards, the detection algorithms are applied to the de-noised images in order to separate signals and background noise. As a result, the tracking algorithms had to assume a cluttered environment with false alarms and missed detections, impacting their performance. Thus in this paper we concentrate on the first step in the data analysis procedure: data filtering or de-noising.

De-noising is commonly carried out using one of two approaches: Gaussian smoothing or wavelet de-noising. In the first, the image is smoothed through convolution with a Gaussian or average kernel, which is related to matched filtering for additive uncorrelated noise [6]. Use of a Gaussian filter is justified since the fluorescence profile of an individual molecule is well approximated by (and hence frequently fitted to) a Gaussian function [7]. In cases where the noise addition is linear or stationary, a Wiener filter is also used to reduce the noise of the image [8]. For the non-stationary problem, wavelet based de-noising algorithms are also well-known for image restoration [9,10]. In computer vision and image processing society, Markov random field is also popularly used to reduce noise from images [11–14].

To date, there have been several studies about denoising noisy single molecule images [15–17]. Lee et al. uses higher order singular value decomposition (HOSVD) method to extract spatial and temporal features [15]. Noise reduction in single-molecule experiments was improved by using molecular constructs with short handles [17]. There are also several approaches to de-noise the trajectories of the single molecules rather than image denoising [18–20]. A generalized nonlinear filtration technique introduced by Haran reduces the noise of single molecule fluorescence trajectories [18]. Taylor et al. introduced wavelets and Bayesian inference to

* Tel.: +82232904886.

1746-8094/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bspc.2013.12.005

de-noise single molecule fluorescence resonance energy (smFRET) trajectories [19].

In this work we develop and describe a new de-noising algorithm based on a Gaussian Markov random field (GMRF) model. Our proposed algorithm is a fully Bayesian approach with few tuned parameters: we need to set only a small relative threshold parameter and specify hyper-parameters to build prior distributions. The performance of this method relative to previous methods is evaluated using both synthetic and real single molecule data and found to display significant advantages.

The paper is structured as follows. We first introduce the Gaussian Markov random field in a latent Gaussian model (Section 2) which is then used to propose mathematical models for de-noising images (Sections 3 and 4). Finally, we compare the results of the various algorithms when applied to synthetic and real data (Section 5).

2. Intrinsic Gaussian models

Gaussian Markov random fields (GMRFs) are Gaussian fields defined on a discrete grid with a Markov property of conditional independence of a component with all others given its neighbours [21,22]. They have seen widespread application in statistical modelling, for example in spatio-temporal models [23] and dynamic linear models [24].

In this paper, a GMRF $\mathbf{f}$ is an one dimensional vector which corresponds to a two dimensional image lattice $\{(i,j)|i=1,\ldots,n_1\,;j=1,\ldots,n_2\}$ [22]. Let $\Delta_{i,j}$ be differenced values of $\mathbf{f}$ at a site (i,j) on the lattice with neighbours. There are many possible ways to define a GMRF for $\mathbf{f}$ through the $\Delta_{i,j}$. For example, a first order model can be defined through assuming the differences with the 4 nearest neighbours $\eta_{i,j}=\{(i-1,j),(i+1,j),(i,j-1),(i,j+1)\}$,

$$\Delta_{i,j}=\sum_{(k,l)\in\eta_{i,j}}(f_{k,l}-f_{i,j}),$$

to be Gaussian with mean 0 and precision (inverse of variance) κ_f. The distribution of $\mathbf{f}$ is of multivariate Gaussian form:

$$p(\mathbf{f}|\kappa_f)\propto\exp\left\{-\frac{\kappa_f}{2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\Delta_{i,j}^2\right\}=\exp\left\{-0.5\kappa_f(\mathbf{Df})^T(\mathbf{Df})\right\}$$
$$=\exp\left\{-0.5\kappa_f\mathbf{f}^T\mathbf{Q}_f\mathbf{f}\right\},\qquad(1)$$

where $\mathbf{Q}_f=\mathbf{D}^T\mathbf{D}$, and $\mathbf{D}$ is a $n_1n_2\times n_1n_2$ matrix with elements $D_{m,m}=-\sum_{d=1,d\neq m}^{n_1n_2}D_{m,d}$, $D_{m_1,m_2}=-1$ if the pixels represented by the m_1th and m_2th components are neighbours, and 0 otherwise. The resulting precision matrix $\mathbf{Q}_f$ is not of full rank, in which case it is known as an intrinsic GMRF (IGMRF). This form is used often as a prior in Bayesian inference. For de-noising, it has the desirable properties of placing more prior weight on smooth images while avoiding the specification of a mean value of $\mathbf{f}$. Under very general conditions, the posterior of $\mathbf{f}$ will become a proper distribution [22].

3. The de-noising model

Note that we focus on de-noising a single image in this paper although we have a sequence of images for processing. For simplicity, we assume images in the sequence are independent so that the de-noising algorithm can be applied separately. The dependency of spots across the sequence of images is accounted for in the tracking process.

3.1. Linear model for an intrinsic GMRF

We assume a Gaussian model for an observed image $\mathbf{y}$ in terms of the underlying signal $\mathbf{f}$ and some regression coefficients $\gamma=(\gamma_1,\gamma_2,\gamma_3)$:

$$\mathbf{y}=\mathbf{z}^T\gamma+\mathbf{f}+\epsilon\qquad(2)$$

where $\epsilon\sim\mathcal{N}(\cdot\,;\mathbf{0},\kappa_l^{-1}\mathbf{I})$ for an unstructured term and $\mathcal{N}$ denotes the Gaussian distribution. In this model, $\mathbf{z}$ are lattice location indices so that $\mathbf{z}^T\gamma$ models any global background noise trend. One of the simplest and the most practical ways to design $\mathbf{z}$ is to use low-pass filters like average and Gaussian filter to model the background noise.

3.2. Priors

According to Eqs. (1) and (2), we find that there are unknown regression coefficients γ and a set of two hidden parameters denoted by $\theta=\{\kappa_l,\kappa_f\}$. Given this model we assume conjugate prior forms of the regression coefficients

$$\gamma\sim p(\gamma)=\mathcal{N}(\gamma;\mathbf{0},\mathbf{Q}_\gamma^{-1}),\qquad(3)$$

where $\mathbf{Q}_\gamma=(1/10^3)\mathbf{I}_{3\times3}$, giving a weak prior, and $\kappa_f\sim\mathcal{G}(\cdot\,;\alpha_f,\beta_f)$ and $\kappa_l\sim\mathcal{G}(\cdot\,;\alpha_l,\beta_l)$, where $\mathcal{G}(\cdot\,;\alpha,\beta)$ denotes the Gamma distribution with hyper-parameters α and β. Since κ_f and κ_l are non-negative real variables, we model their priors by using Gamma distribution. In machine learning and data mining literatures, the Gamma distribution is commonly used for designing inverse of variance since the Gaussian distribution of a random variable $\mathbf{x}$ with mean μ and covariance $(\kappa\mathbf{Q})^{-1}$ can be transformed to Gamma distribution with respect to κ.[1]. In addition, Gamma distribution is also a member of an exponential family so that we can use it as conjugate prior to obtain a closed form for the posterior distribution.

According to the linearity of the hidden parameters in Eq. (2), we can apply Rao-Blackwellization technique [25] to reduce the dimension of the hidden variables and variances. Now, we have marginalized likelihood defined by

$$p(\mathbf{y}|\mathbf{f},\theta)=\int_\gamma p(\mathbf{y}|\mathbf{f},\gamma,\theta)p(\gamma|\theta)d\gamma$$
$$=\int_\gamma\mathcal{N}(\mathbf{y};\mathbf{z}^T\gamma+\mathbf{f},\kappa_l^{-1}\mathbf{I})\mathcal{N}(\gamma;\mathbf{0},\mathbf{Q}_\gamma^{-1})d\gamma=\mathcal{N}(\mathbf{y};\mathbf{f},\Phi)\qquad(4)$$

where $\Phi=\kappa_l^{-1}\mathbf{z}^T(\mathbf{z}\mathbf{z}^T)^{-1}\mathbf{z}+\mathbf{z}^T\mathbf{Q}_\gamma^{-1}\mathbf{z}$.

3.3. Problems with the IGMRF prior

Ref. [22] describes several limitations of the intrinsic GMRF. The limitations will bring following problems in de-noising images:

1 Unwanted blurring effects can be introduced because the IGMRF penalises discrete boundaries or sharp gradients between neighbours;
2 Weak signals can be ignored by smoothing with the background;
3 Stationary (homogeneous) fields across an image rarely exist in practice.

[1] Given a d dimensional random variable $\mathbf{x}$ with μ and variance $(\kappa\mathbf{Q})^{-1}$, we have $p(\mathbf{x}|\mu,\kappa,\mathbf{Q})=\mathcal{N}(\mathbf{x};\mu,(\kappa\mathbf{Q})^{-1})=|2\pi\kappa\mathbf{Q}|^{(1/2)}\exp\left\{-\frac{\kappa}{2}(\mathbf{x}-\mu)^T\mathbf{Q}(\mathbf{x}-\mu)\right\}=\kappa^{(d/2)}|2\pi\mathbf{Q}|^{(1/2)}\exp\left\{-\frac{(\mathbf{x}-\mu)^T\mathbf{Q}(\mathbf{x}-\mu)}{2}\kappa\right\}=\mathcal{G}\left(\kappa;\frac{d}{2}+1,\frac{(\mathbf{x}-\mu)^T(\mathbf{x}-\mu)}{2}\right)$ where $\mathcal{G}(\cdot\,;a,b)$ is the Gamma distribution with a shape parameter a and a rate parameter b. That is, normal distribution can be described in the Gaussian distribution with respect to κ.