Published in final edited form as:

Proc IEEE Int Symp Biomed Imaging. 2009 June 28; : 803–806. doi:10.1109/ISBI.2009.5193171.

IMAGE RECONSTRUCTION OF MULTIPHOTON MICROSCOPY DATA

Jared M. Doot,

University of Wisconsin-Madison, Electrical and Computer Engineering, Madison, WI 53706

Kevin W. Eliceiri^{*},

University of Wisconsin-Madison, Electrical and Computer Engineering, Madison, WI 53706

Robert D. Nowak[†], and

University of Wisconsin-Madison, Electrical and Computer Engineering, Madison, WI 53706

Rebecca Willett

Duke University, Electrical and Computer Engineering Durham, NC 27708

Abstract

The techniques introduced in this paper allow for accurate multi-scale image reconstruction of multi-photon microscopy data. Multi-photon microscopy (MPM) is a tool for the non-invasive imaging of living organisms and tissue. The data acquired using this technique can contain information about the position, excited state lifetime, and spectra of the observed photons. The small number of photons collected, however, limits the quality of the reconstruction. The multiscale framework in this paper results in an accurate representation of both the intensity and excited state lifetime information. Using a multiscale reconstruction approach based on a penalized likelihood function, the underlying image is more accurately represented as compared to a naive aggregate binning approach.

Index Terms

Biomedical microscopy; Poisson processes; Multidimensional signal processing; Maximum likelihood estimation

1. MULTIPHOTON MICROSCOPY

Optical sectioning fluorescence microscopy has become the method of choice for imaging living specimens as it offers high signal-to-background ratio and the ability to spectrally discriminate between multiple fluorophores. Recently developed techniques such as confocal [1] or multi-photon [2] imaging allow optical sections to be made of intact specimens. These may be collected as stacks of images at different focal depths to obtain three-dimensional data. Stacks of images may be collected at regular time intervals in order to reveal the dynamics of three-dimensional structures in living tissue [3, 4].

Multiphoton microscopy (MPM) is a non-invasive imaging modality that can provide high resolution images of fluorescence with low toxicity, high signal to noise ratio and optical sectioning deeper than confocal microscopy [2, 5, 6].

Formally speaking, photon data are distributed according to an inhomogeneous Poisson process, where the inhomogeneity refers to spatio-temporal variations in the underlying photon emission process. Wavelet denoising is now a classical multiresolution technique for spatially adaptive noise reduction, however, most wavelet-based approaches are inadequate for photon data denoising since they are usually designed with Gaussian noise in mind. Gaussian approximations to Poisson data statistics are undesirable because the approximations do not support non-negativity constraints (photon intensities are non-negative by nature) and since they are usually only reasonable if the number of photon events occurring is sufficiently large (so that the data, possibly after a suitable transformation, are roughly Gaussian distributed). To obtain high count levels in practice the detections must be artificially binned or aggregated over sufficiently large regions. Thus, one must immediately sacrifice resolution in order to accommodate the approximations. As a result, standard wavelet techniques are inappropriate for MPM.

When combined with complimentary detection methods, MPM can give rise to a series of detected photons with spatial and temporal resolution, its characteristic emission spectrum [7], and its characteristic excited state lifetime [7]. These characteristics are referred to as "marks". One challenge of this work is to perform accurate inference on the underlying phenomenon using small numbers of random, discrete, marked events.

1.1. Problem Statement

In this paper we focus on one example of marked Poisson processes in MPM: photon limited spatio-lifetime analysis. Binning the observed photons into small spatial and lifetime bins results in a data cube with a small number of observations in each bin. Specifically, assume that we measure Poisson observations of a spatially-varying intensity over a fixed lifetime window, such that the intensity is the sum of the events over the lifetime window. Assume that the number of events in each bin is small. Let f denote the $M \times N \times T$ true array of photon counts, where the first two dimensions represent the spatial locations and the third dimension corresponds to the excited state lifetimes. Our observations are of the form:

$$y \sim \text{Poisson}(f)$$

where y has the same dimensions as f . Our goal is then to infer f from the measurements y as accurately as possible. We wish to denoise the observations in such a manner that the lifetime information is retained.

1.2. Current Analysis Approaches

Current approaches estimate the lifetime distribution by fitting a multi-exponential model to the lifetime data at each pixel. Since each bin may have very few or even no photons, it is often impossible to obtain robust estimates of multiple parameters due to the limited photon counts per bin. Thus an aggregate of bins is needed for the calculation of parameters of interest. Current approaches address this by using a moving average approach to aggregate neighboring bins together [8]. This approach works well when the lifetime information is similar for surrounding pixels, but in regions of varying lifetime values, the data are erroneously aggregated together, resulting in an inaccurate representation. We seek to improve on these results by adaptively selecting the optimal partition of the observations.

Multiscale image estimation methods based on translation-invariant (TI) Haar wavelets, wedgelets, and platelets are near minimax optimal reconstruction techniques for photon limited images [9, 10]. These methods calculate image estimates by determining the ideal partition of the domain of observations (assumed to be $[0, 1]^2$) and using maximum

likelihood estimation to fit a model to each square of the optimal partition. The space of all possible partitions is a nested hierarchy defined through a recursive dyadic partition (RDP) of $[0, 1]^2$, and the optimal partition is selected by pruning a quad-tree representation of the observed data. An example of such a partition is shown in Figure 1. This gives the estimators the capacity of spatially varying the resolution to automatically increase the smoothing in very regular regions of the image and to preserve detailed structure in less regular regions. The accuracy of these estimates can be augmented by a process called cycle-spinning, or averaging over shifts, resulting in translation-invariant estimates [11]. Pruning decisions are made using a penalized likelihood criterion. [12]

1.3. Extention to lifetime data

We wish to use an RDP to optimally partition our spatio-lifetime data. The optimal partition should distinguish between regions of differing lifetime. Each square at the terminal end of the optimal RDP contains the exited state lifetime information and should be spatially separated from regions of differing lifetimes.

2. MULTISCALE LIKELIHOOD ESTIMATION

We first introduce a model for the spatio-lifetime distribution, then factor this model into temporal and spatial components. These distributions are used in a penalized likelihood setting for the calculation of the optimal partition.

2.1. Data

We define X to be an $M \times N \times T$ matrix of photon counts. $X_{i,j,k} \in \mathbb{N}$ is the photon count at spatial location i, j and time bin k .

Assume that there is a photon intensity function Λ over time and space. Without loss of generality, assume that the spatial field of view is the unit square $[0, 1]^2$ and that temporal variable ranges in the unit interval $[0, 1]$. Each count $X_{i,j,k}$ is assumed to be independent and Poisson distributed with intensity

$$\lambda_{i,j,k} := \int_{(x,y) \in S_{i,j}} \int_{t \in T_k} \Lambda(x, y, t) dx dy dt,$$

where $S_{i,j} = \left[\frac{i-1}{M}, \frac{i}{M} \right) \times \left[\frac{j-1}{N}, \frac{j}{N} \right)$ denotes the i, j -th pixel region in the unit square and $T_k = \left[\frac{k-1}{T}, \frac{k}{T} \right)$ is the k -th time bin. That is, $\lambda_{i,j,k}$ is the total photon intensity in pixel i, j and time bin k .

2.2. Semiparametric Model

Define $X_{i,j} := [X_{i,j,1}, \dots, X_{i,j,T}]$ and let $\lambda_{i,j}$ denote the corresponding vector of intensities.

Define $Y_{i,j} := \sum_{k=1}^T X_{i,j,k}$ and let $\mu_{i,j} = \sum_{k=1}^T \lambda_{i,j,k}$. Then we can factorize the distribution of the vector $X_{i,j}$ as follows.

$$P(X_{i,j} | \lambda_{i,j}) = P(X_{i,j} | Y_{i,j}, \alpha_{i,j}) P(Y_{i,j} | \mu_{i,j}).$$

The distribution $P(Y_{i,j}|\mu_{i,j})$ is Poisson with parameter $\mu_{i,j}$. This follows from the basic fact that the sum of Poisson random variables is Poisson distributed. The distribution $P(X_{i,j}|Y_{i,j}, \alpha_{i,j})$ is multinomial, and $\alpha_{i,j} = [\lambda_{i,j,1}/\mu_{i,j}, \dots, \lambda_{i,j,T}/\mu_{i,j}]$ are the parameters of this distribution. This, too, follows from standard properties of the Poisson distribution. Based on the physics of the multiphoton process, it is known that $P(X_{i,j}|Y_{i,j}, \alpha_{i,j})$ is well modeled as a mixture of exponential decaying distributions [7]. That is, the parameter $\alpha_{i,j}$ is a mixture of $L < T$ probability mass functions of the form

$$\alpha_{i,j} = \sum_{\ell=1}^L p_{i,j,\ell} \beta_{i,j,\ell},$$

where each component probability mass function is parameterized by a scalar parameter $\gamma_{i,j,\ell}$ and has the form

$$\beta_{i,j,\ell} = [\gamma_{i,j,\ell}, \gamma_{i,j,\ell}^2, \dots, \gamma_{i,j,\ell}^T].$$

This arises as the natural discretization of a continuous exponential density function of the form $e^{-\gamma t}$ on $[0, 1]$. In summary, we have specified a semiparametric form for λ . The spatial intensity function μ is nonparametric, but the temporal intensity function is a mixture of $L < T$ parametric probability mass functions.

2.3. Multiresolution Inference

Let $\mathcal{A}$ denote a recursive partition of $[0, 1]^2$ that determinates with the pixels $S_{i,j}$. Let $A \in \mathcal{A}$ be any cell in the partition. We say that the underlying photon process is *spatially homogeneous* on A if for each value of k , $\lambda_{i,j,k}$ is a constant for all $S_{i,j} \in A$. In that case, following the derivations above, we have the following factorization of the distribution of each $X_{i,j}$ such that $S_{i,j} \in A$:

$$P(X_{i,j}|\lambda_{i,j}) = P(X_{i,j}|Y_{i,j}, \alpha_A) P(Y_{i,j}|\mu_A).$$

In other words, the parameters of the distribution are constant within each cell of the partition. While our process may not be strictly homogeneous, it is a very good approximation.

Suppose A is the union of four disjoint cells $A_1, \dots, A_4 \in \mathcal{A}$. Then the data $X_A := \{X_{i,j} : S_{i,j} \in A\}$ could be modeled in two ways: homogeneously on A or homogeneously on $A_1, \dots, A_4$, separately. The model consisting of four separate parts is more flexible and will always fit the data more closely. Let $L(X_A; A)$ denote the likelihood of the data using the single model on A and let $L(X_{A_r}; A_r)$, $r = 1, \dots, 4$, denote the four likelihoods of the separate models.

Then $L(X_A; A) \leq \sum_{r=1}^4 L(X_{A_r}; A_r)$. We can adaptively prune a partition by introducing a penalty parameter $\xi > 0$ and accept the single model whenever

$$-L(X_A; A) \leq -\sum_{r=1}^4 L(X_{A_r}; A_r) + \xi$$

Where ζ is of the form $c \cdot \log(\sum_{i,j,k} X_{i,j,k})$ and c is a user defined parameter. [13]

An EM algorithm is used to calculate the mixture parameters of the lifetime distribution. Using these parameters, a likelihood can be calculated and used in the decision to split or merge based on the likelihood and penalty. The decision to split or merge is used to find the optimal partition of the data to denoise and appropriately bin lifetime information.

3. EXPERIMENTAL RESULTS

The above method exploits spatial homogeneities as well as lifetime information to find the optimal partition. This reduces over-smoothing across bins with differing lifetime values. To test the effectiveness of the proposed method, we develop a phantom image as follows. The phantom contains 4 regions. The background is a fixed intensity with no associated lifetime value. There are 3 overlapping regions of interest. As seen in Figure 1, the largest has an associated intensity value and lifetime value, λ_1, τ_1 . The two smaller regions have values λ_2, τ_1 and λ_1, τ_2 . The lifetime component is modeled as mixture of a Gaussian and an geometric distribution. The Gaussian component of the mixture distribution models the system response function of the detector in the imaging equipment.

The intensity image for the phantom is displayed in Figure 3(a). A lifetime map is created by generating a lifetime value for each pixel in the image. This lifetime image is shown in Figure 3(e). Photon-noisy observations of the intensity at each pixel are displayed in Figure 3(b). Using the Poisson observations and the EM algorithm to calculate the lifetime value at each pixel, Figure 3(f) shows the Poisson observations of the lifetimes. The data cube is $128 \times 128 \times 64$, and the total number of photons is 737,800. This is an average of .7036 photon events per bin. In this simulation, c was chosen to improve empirical performance. The weight was not selected to minimize a particular metric, but instead to yield more accurate reconstructions. The error metric used for comparison of the different results is as follows: $\|f - \hat{f}\|^2 / \|f\|^2$.

If we were to reconstruct the intensity image by using the moving average approach, as in other work, we achieve $\hat{f}_{I_a}$ seen in Figure 3(c) with an average MSE of 0.0283. The radius of pixels in the average was chosen to minimize the MSE. Using our penalized likelihood decision model, we achieve $\hat{f}_{I_{ple}}$ in Figure 3(d) with an average MSE of 0.0011, a significant improvement. After applying a moving average filter and a penalized likelihood approach to the data cube, a lifetime map is estimated. This lifetime map is shown in Figure 3(g) and (h). Using this information an average MSE for the lifetime can be calculated for the two approaches. The average MSE for the moving average estimate, $\hat{f}_{LT_a}$ is 0.0765, while the average MSE for the penalized likelihood estimate, $\hat{f}_{LT_{ple}}$ is 0.0214. This improvement can be attributed to the ability of our algorithm to better estimate the edges in the image in addition to better estimations over large regular regions.

We have also started to apply these methods on biological multiphoton data. Our study examined data gathered from a histology slide at various collection times. We compare the lifetime values at a collection time of 60 seconds with values calculated at 1 second. The two above methods are tested and the results are displayed in Figure 4(a)–(c). The poor performance of the fixed binning approach may be due to the fact that the collected data has a short lifetime. It seems as if the fixed binning approach has a more difficult time fitting fast lifetime values on photon limited data, while our method is able to reveal the underlying structure of the data. If a more intermediate collection time was used, the fixed binning would have performed better. But in practice, it is undesirable to collect data for longer than necessary due to possible tissue damage or to study changes over time.

4. DISCUSSION

In our proposed method, we build a quad tree decomposition of the data cube so that each leaf node of the tree corresponds to a lifetime distribution. This tree is then pruned using the penalized likelihood criterion. Instead of only reconstructing the intensity, our method also can reconstruct the lifetime. This allows for the image to be separated into distinct lifetime regions as well as denoised. We would like to move forward with this work by using additional marks to determine as much information from the data as possible. This includes using the spectral information to further understand MPM images.

References

1. White JG, Amos WB, Fordham M. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *The Journal of Cell Biology*. July.1987 105:41–48. [PubMed: 3112165]
2. Denk W, Strickler JH, Webb WW. Two-photon laser scanning fluorescence microscopy. *Science*. April.1990 248:73–76. [PubMed: 2321027]
3. White JG, Squirrell JM, Eliceiri KW. Applying multi-photon imaging to the study of membrane dynamics in living cells. *Traffic*. November.2001 2:775–780. [PubMed: 11733043]
4. White JG, Thomas CF. Four-dimensional imaging: the exploration of space and time. *Trends*. April. 1998 16:175–182.
5. White JG, Centonze VE. Multiphoton excitation provides optical sections from deeper within scattering specimens than confocal imaging. *Biophysical Journal*. October.1998 75:2015–2024. [PubMed: 9746543]
6. Squirrell JM, Wokosin DL, White JG, Bavister BD. Long-term two-photon fluorescence imaging of mammalian embryos without compromising visibility. *Nature Biotechnology*. August.1999 17:763–767.
7. Lakowicz, JR. *Principles of Fluorescence Spectroscopy*. Kluwer Academic/Plenum Publishers; 1999.
8. Provenzano PP, Rueden CT, Trier SM, Yan L, Ponik SM, Inman DR, Keely PJ, Eliceiri KW. Nonlinear optical imaging and spectral-lifetime computational analysis of endogenous and exogenous fluorophores in breast cancer. *Journal of Biomedical Optics*. 2008; 13(3)
9. Willett R, Nowak R. Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging. *IEEE Transactions on Medical Imaging*. 2003; 22:232–250.
10. Willett, R.; Nowak, R. Technical Report TREE0303. May. 2003 Multiscale density estimation.
11. Willett R, Nowak R. Fast, near-optimal, multiresolution estimation of poisson signals and images. *EUSIPCO*. 2004
12. Kolaczyk, E.; Nowak, R. Multiscale likelihood analysis and complexity penalized estimation. <http://www.ece.wisc.edu/nowak/pubs.html> Submitted to *Annals of Statistics* Aug. 2001
13. Willett R, Nowak R. Multiscale poisson intensity and density estimation. *IEEE Transactions on Information Theory*. Sept.2007 53:3171–3187.

Fig. 1.
Example RDP.

Fig. 2. Spatio-lifetime intensity. (a) Phantom with labeled intensity and lifetime regions (b) Lifetime values used in the phantom

Fig. 3. Spatio-lifetime analysis Results. (a) Original intensity image. (b) Pixel-based intensity estimate based on Poisson observations. (c) Reconstructed intensity image using moving average approach. (d) Reconstructed intensity image using penalized likelihood approach. (e) Original lifetime map. (f) Pixel-based lifetime map estimate using Poisson observations. (g) Reconstructed lifetime map using moving average reconstruction. (h) Lifetime map using penalized likelihood reconstruction.

Fig. 4.
Dog Bladder Histology Slide Results at 890nm with a 10 $\times$ lens. (a) Lifetime at 60 seconds.
(b) Fixed binning calculated lifetime at one second. (c) Adaptive binning calculated lifetime
at one second.