A Wavelet-Based Approach to
3-D Confocal Microscopy Image Reconstruction

by

Ben David Graf

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Science

APPROVED, THESIS COMMITTEE:

Robert Nowak, Associate Professor, Chair
Electrical and Computer Engineering

Don H. Johnson, Professor
Electrical and Computer Engineering

Richard Baraniuk, Professor
Electrical and Computer Engineering

HOUSTON, TEXAS

MAY 2004
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI

UMI Microform 1419079
Copyright 2004 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346
Abstract

A Wavelet-Based Approach to

3-D Confocal Microscopy Image Reconstruction

by

Ben David Graf

An algorithm based on the Haar wavelet basis and implementing an expectation maximization-maximum penalized likelihood estimator in 3-D is shown to provide dramatic improvement over traditional stopped-EM algorithms in terms of mean-squared error on simulated data for confocal microscopy systems. Confocal microscopy is one of many modern medical imaging systems changing the landscape of medical research and practice, and the blurred and grainy images produced are much more useful when suitable, accurate reconstruction algorithms are applied. The industry standard, the stopped expectation-maximization algorithm proves unreliable and inadequate when compared to penalized likelihood estimators based on spatially adaptive bases such as wavelets. In addition, processing confocal microscopy images in 3-D, rather than slice-wise in 2-D, takes into account the blurring that occurs between slices as a result of the microscope’s point spread function.
Acknowledgements

My advisor, Dr. Robert Nowak, was an invaluable source of information, inspiration, and motivation. Without his guidance, this research would have had far less direction and purpose. His knowledge and creativity made my task much easier, and his willingness to work around the situations that arose along the way was unbounded.

This thesis benefited as well from the wisdom and expertise of my committee members, Dr. Don Johnson and Dr. Rich Baraniuk, whose comments and advice are much appreciated.

An unexpected but much appreciated contribution came from fellow student Rebecca Willett. This research in some ways branched from her previous work, so I expected some collaboration, but her assistance exceeded my expectations, and I hate to think how I would have managed without it.

For sharing their in-depth knowledge of medical imaging, I have to add a special thanks to the confocal group at the Baylor College of Medicine. Specifically, Vivek Bansal provided a wealth of documentation and answers to my myriad questions regarding confocal microscopy and the nature of dendritic imaging. Tycho Hoogland provided real data on request, taking time out of his own work to do it.

Elizabeth Hickman greatly facilitated the processing of this thesis and all of the related forms and applications and always kept me aware of the numerous approaching deadlines.

And of course, my family provided inestimable love, encouragement, and understanding. No amount of gratitude would be enough to repay such support.
Contents

Abstract ii
Acknowledgements iii
List of Figures v

I. Introduction................................................................. 1

II. Confocal Microscopy and Neuronal Imaging......................... 2

III. Expectation-Maximization versus Penalized Approaches.......... 10

IV. Multiscale Analysis and Wavelets................................... 14

V. The 3-D EM-MPLE Algorithm with Haar Thresholding............... 18

VI. Simulation Results....................................................... 21

VII. Conclusions and Future Work......................................... 25

References 28
List of Figures

1 Scanning Microscopy ......................................................... 3
2 Reduction of Out-of-Focus Light Using a Confocal Pinhole .......... 4
3 Galvanometer-based Descan .................................................. 5
4 Complete Galvanometer-based Confocal Microscopy System ....... 5
5 Optical Transfer Function True Plot ....................................... 7
6 Optical Transfer Function Log Plot ....................................... 7
7 Hippocampal Neuron ............................................................ 9
8 Poisson System from Original Intensity to Observed Intensity ...... 11
9 Haar Scaling Function and Wavelet Function ........................... 17
10 Synthetic Data – 3-D Cylinder .............................................. 22
11 Original Data, λ ................................................................. 23
12 Blurred Data, μ ................................................................. 23
13 Received Poisson Data, y ....................................................... 23
14 Best EM Result ................................................................. 24
15 Best EM-MPLE Result .......................................................... 24
16 MSE Comparison ............................................................... 25
I. INTRODUCTION

Medical imaging systems such as confocal microscopy are becoming more prevalent and are increasingly relied upon by medical professionals in diagnosis and treatment decisions. Unfortunately, such systems do not deliver crystal clear images. In fact, image quality can be quite poor, due to the nature of the equipment and human body tissues. Such a situation begs for good image reconstruction algorithms.

The standard reconstruction method implemented is the expectation-maximization algorithm, based on the maximum likelihood estimator. It is preferred because of its relative simplicity but has a number of drawbacks to be outlined in Section III. Alternative methods utilizing a penalized estimator have proven superior in accuracy and dependability and hold the most promise for ongoing improvement.

One such method, the maximum penalized likelihood estimator, can make use of the wavelet basis to improve reconstruction. Wavelets have dominated more than a decade of signal research, and for good reason. They have proven both powerful and versatile in a wide range of applications of which reconstruction is only one. Wavelets effectively represent smooth data (which represents a large portion of real life data sets, including images) sparsely. In the case of images, most of the wavelet coefficients are used on the most interesting part of the image, the edges (discontinuities), and the smooth areas require few coefficients. The Haar wavelet basis proves especially useful.

This paper examines a wavelet-based penalized estimator that is performed in 3-D. Confocal microscopy, the imaging system analyzed here, produces a 3-D image of a volume of tissue. Existing reconstruction methods are typically performed in a slice-wise 2-D
fashion. As Section II will demonstrate, for confocal microscopy at least, this idea is fundamentally flawed.

As for the layout of this paper, Section II discusses confocal microscopy and the imaging of neurons, the specific application. Section III covers the expectation-maximization algorithm and penalized alternatives. Section IV outlines the basics of wavelets, specifically the Haar wavelet basis. The description of the algorithm devised for this paper is completed in Section V, while Section VI covers simulation results. Finally, Section VII ties everything together and looks at potential avenues for further development.

II. CONFOCAL MICROSCOPY AND NEURONAL IMAGING

In confocal microscopy, as in many imaging systems used in modern medicine, images are formed by photons emitted from a specimen hitting a detector. Specifically, fluorescent dyes applied to the specimen attach themselves to key structures and emit photons from primarily these locations when excited by a laser or other light source [1]. These emissions are typically modeled as Poisson noise, variation of the intensity about the Poisson mean. Because the number of photons detected is small in comparison to the number emitted, the signal-to-noise ratio can be undesirably low, and the techniques are known as photon-limited. A critical result of this photon-limited nature is that such methods are poorly approximated as Gaussian [2]; reconstruction algorithms should ideally be based around Poisson assumptions.

Wide-field microscopes attempt to illuminate an entire specimen at once and view it using a photodiode array. Unfortunately, such arrays are saddled with poor resolution, making imaging of features as small as those in neurons, the specimens used in Section VI,
virtually impossible. In addition, illuminating the entire specimen at one time introduces a lot of avoidable noise due to scattering of photons from one region to another [1].

These problems can be downplayed through the use of scanning microscopes, which focus excitation illumination on a tight spot in the 3-D specimen and then scan over the field of view to create a complete image. This tight focus limits the scattering to areas local to the region of interest [1]. This method can be shown theoretically and empirically to improve resolution significantly. The resultant loss in field of view is countered by the scanning nature of the setup [3]. Still, out-of-focus light arises from a cone-shaped region above and below the focal plane and is unavoidable in the basic scanning microscope setup [1]. Figure 1 diagrams the basic scanning microscopy setup.

This is where confocal microscopy comes into play. A spatial filter such as a simple pinhole is placed between the detector and the specimen to almost entirely eliminate out-of-focus light from outside the focal plane; scattered photons are largely eliminated as an additional bonus [1]. Confocal microscopy allows for thick, translucent objects to be imaged in 3-D through optical sectioning. A single level of the imaged 3-D object can be viewed because light from the material is focused through a small aperture [3]. Figure 2 diagrams confocal microscopy’s ability to eliminate of out-of-focus light. Image A shows the light from the focal point passing through the filter. Image B demonstrates how regions in the focal plane, but away from the actual focal point, are completely blocked by the filter.
Figure 2: Reduction of Out-of-Focus Light Using a Confocal Pinhole
A) Light from the focal point passes to the detector. B) Light from other areas in the focal plane will be completely blocked. C) Light from above the focal plane and D) Light from below the focal plane will be mostly blocked. [1]

Images C and D show how light from regions above and below the focal plane is mostly blocked as it attempts to focus behind or in front of the filter, respectively [1].

One specific confocal microscope setup, and the one used to generate the real life data examined in Section VI, is the galvanometer-based confocal microscope. This setup is the most widely used confocal setup and derives its name from the galvanometer-driven mirrors used to direct the excitation laser beam as well as the return fluorescence beam from the specimen. By positioning two separate mirrors, one for the x-direction and one for the y-direction, any location in the specimen can be focused upon, and a serial scan can be performed. The beam of fluorescence returning from the specimen along the same optical path (the “descan”) has a longer wavelength than the laser excitation beam, and so can be redirected using a dichroic mirror, one that passes certain wavelengths and deflects others. A
spatial filter can be inserted after this dichroic mirror to complete the confocal setup [1].

Figure 3 demonstrates the use of the galvanometer-driven mirrors in the descan process, and Figure 4 shows a typical galvanometer-based confocal setup.

The primary disadvantage of the galvanometer-based confocal microscope is in temporal resolution. While spatial resolution is superior to many opposing methods, the

![Diagram of confocal microscopy system](image)

**Figure 3:** Galvanometer-based Descan [1]

**Figure 4:** Complete Galvanometer-based Confocal Microscopy System [1]
scanning method using moving mirrors results in slow (relatively) scans. The mirrors’ mass and inertia dictate physical limits in their movement capabilities. For this reason, studies attempting to focus on rapidly occurring changes in a specimen often utilize line scans or faster confocal setups, such as the Nipkow disk-based confocal microscope. However, for three-dimensional structural imaging, galvanometer-based confocal microscopes are superb [1].

Confocal microscopes, like all such imaging systems, introduce blurring or other such image corruptions. An ideal reconstruction algorithm would not only denoise, but deblur as well. Proper approximation of the transfer function for the imaging process needs to be established to allow for improved reconstruction algorithms. The blur induced by confocal microscopy systems can be modeled based on the microscope physics. Two types of transfer functions can be described for scanning optical microscopes. Coherent systems appropriately use the coherent transfer function (CTF), while incoherent systems use the optical transfer function (OTF), which can be defined as the Fourier transform of the point spread function (PSF) [4]. The confocal setup falls under the latter heading. The 3-D OTF for one-photon fluorescence microscopy is

\[ C(t, s) = \frac{1}{l} \Re \left[ 1 - \left( \frac{|s|}{l} + \frac{l}{2} \right)^2 \right]^{1/2} \]

Figures 5 and 6 illustrate the OTF. Both plot C over s, the axial direction, and l, the radial direction. Figure 5 displays a standard plot of the OTF, while Figure 6 provides a log plot of C that sheds additional light on the function’s behavior. In both images, note that pixel 25 on each axis corresponds with a normalized frequency of one. This means that frequencies outside of a normalized axial frequency range of 0.5 (12.5 pixels in Figures 5 and 6) or
outside of a normalized radial frequency range of 2.0 (50 pixels) will have zero value in the OTF.

As discussed above, this paper examines confocal one-photon excitation microscopy, whose OTF is the convolution of the above equation with itself [5]. As for the two variables in the OTF, s is the axial frequency, normalized to the cutoff spatial frequency, given by $4n_i \sin^2 (\alpha) / \lambda$. The radial (or lateral) frequency is l, which comes from the transverse
coordinates as expected: \( l = \sqrt{m^2 + n^2} \). The transverse frequencies \((m, n)\) are normalized to \( (n, \sin \alpha)/\lambda \). The refractive index is \(n_r\), the wavelength is \(\lambda\), and the aperture angle is \(\alpha\) [4]. Note that the transfer function is radially symmetric.

As stated, the system produces 3-D images of a specimen, but they are generally reconstructed in 2-D, slice-by-slice [2]. However, the PSF for the confocal microscope has a focal spot that is longer between slices (in the axial direction) than within slices (in the radial direction). This leads to an OTF that is less extended in the axial direction, resulting in radial resolutions that are at least three to four times better than their axial counterparts [6], signifying that such a slice-wise procedure is fundamentally flawed. Another glance at Figures 5 and 6 confirms this idea. The frequency support of the OTF is much wider in the radial direction, indicating more of an input image will be maintained. The axial direction’s low frequency support acts as a low-pass filter. Most blurring will therefore occur between slices; reconstruction of a single slice will ignore this entirely. As a result, improved reconstruction is possible with a coherent method that processes the 3-D image as a whole.

The real (not synthetic) data used in the testing of the algorithm described in this paper are images of neuron structure, especially the dendrites. Recent neuronal research has focused on dendritic structure and function more heavily as evidence has grown to indicate their importance in neural function [7]. While neuron size and shape can vary greatly, neurons from the hippocampus are often the focus of research due to their association with learning. Figure 7 shows a hippocampal neuron. The cell body is known as the soma and is generally 15-20 \(\mu\)m in diameter. The dendrites, which range in diameter from 0.6 to 3 \(\mu\)m, extend out from the cell body in a branching network up to 450 \(\mu\)m in length. The output
tract of the neuron, the axon, extends out from the cell body about 400 μm and is generally 1 μm in diameter [1].

Dendrites themselves possess an additional group of structures, known as dendritic spines. Neurons receive synaptic inputs on these spines, which are mushroom-shaped, with a 1-μm-diameter head and a 110-nm-long neck [1]. Because of their suspected importance in neural function, imaging techniques (along with their respective reconstruction techniques) used for research on neurons must be able to distinguish and adequately display the dendritic spines. In addition, because the preferred study environment for neurons is in living brain tissue, imaging systems must be able deal with a 3-D specimen, ideally by creating a 3-D image [1].

Optical measurements of dendritic activity can be achieved through the use of fluorescent dyes that are sensitive to electrical and chemical changes in the neuron. Calcium is a popular dye trigger, due to its status as the most important ion studied in neurophysiology. The most common of the calcium-sensitive dyes is the class based on the calcium chelator BAPTA (1,2-bis(o-aminophenoxy)ethane tetraacetate) [1].

Other than the Poisson noise that comes from the randomness of photon emission (discussed earlier in this section), two other primary noise sources exist. Light sources, such as lasers, that are used to excite fluorescent dyes can fluctuate from the exact wavelengths required by the dyes. This is referred to as source noise and is typically only a problem at
high intensities. The other noise type is known as dark noise and is the result of thermal photons being detected in addition to fluorescence photons [1]. Dark noise is proportional to the number of pixels in the detector and can be ignored if the number of detected fluorescence photons is not extremely small.

III. EXPECTATION-MAXIMIZATION VERSUS PENALIZED APPROACHES

The maximum likelihood estimator (MLE), as implemented in the expectation-maximization (EM) algorithm is the most common tool for denoising and reconstructing confocal images [8-10]. Specifically, the standard technique in confocal image reconstruction is the stopped expectation-maximization (EM) algorithm [11]. As the traditional EM algorithm iterates, its error typically diverges. Currently, researchers attempt to estimate when the errors start to diverge and stop the EM algorithm at that point. However, this stopping point is difficult to determine in practice and difficult to analyze theoretically.

We define $y$ as the Poisson data, generated from the original (true) data, $\lambda$, by

$$y(d) \sim \text{Poisson} \left( \sum_{b=1}^{B} \lambda(b) p(b, d) \right)$$

The variable $p$ is the probability that a photon emitted from $b$ (in $\lambda$) will be detected at receiver location $d$ (in $y$). The sum in parentheses is often referred to by the variable $\mu$.

Figure 8 diagrams this Poisson system.

We also define $z_{ij}$, related but not identical to $p$, as the number of photons emitted from $j$ in the original data space and received at $i$ in the observed data space; $z$ is of course unknown in a real situation. Now, we have the following results:
Figure 8: Poisson System from Original Intensity toObserved Intensity

\[ y_i = \sum_j z_{ij} \quad \text{and} \quad x_j = \sum_i z_{ij} \quad \text{with} \quad \lambda_j = x_j \]

If \( \bar{z} \) were known, then the estimate of the true data would be the summation of \( \bar{z} \) over \( i \).

Without knowing \( \bar{z} \) exactly, we are forced to estimate it: the EM algorithm does the trick.

The basic algorithm has two iterated steps. The first is the expectation step, expressed mathematically as

\[ \bar{z}^{est} = E[\bar{z} | \bar{y}, \bar{\lambda}^{old}] \]

Here, \( \bar{\lambda}^{old} \) and \( \bar{z}^{est} \) refer to the previous iteration. The second step is the maximization and is expressed as

\[ \bar{\lambda}^{new} = \sum_i z_{ij}^{est} \]

In this case, \( \bar{\lambda}^{new} \) is the new maximized estimate of the original \( \lambda \). These steps can be combined into a single recursive equation as described in [9]. That equation in 1-D is

\[ \lambda^{new}(b) = \lambda^{old}(b) \sum_{d=1}^{B} \frac{Y(d)p(b,d)}{\sum_{b'} \lambda^{old}(b')p(b',d)} , \quad b = 1,\ldots,B \]

In a real life problem, \( \lambda \) is completely unknown; \( y \) and \( p \) (or an estimate of it) are the only tools to approximate \( \lambda \) [9].
A special case involves a parameter where each row is a shifted version of the same thing; in other words, $p$ is Toeplitz. This causes the two sums in the EM equation to become a pair of convolutions. Toeplitz matrices can be approximated as circulant, and that allows the data to be converted to the Fourier domain using the fast Fourier transform (FFT). The convolutions can now be performed as multiplications in frequency. It is widely known, of course, that the FFT and multiplication can be executed in $O(n \ln n)$, while convolution in time takes $O(n^2)$. Performing the calculations in the Fourier domain, especially when the algorithm is moved to higher dimensions (i.e., three) thereby allows for a dramatic decrease in runtime and even makes the variables far more manageable. The assumption that the transition matrix is Toeplitz (and approximately circulant) allows it to be represented in half the number of dimensions. Note in the original EM equation that $p$ has two dimensions. In 3-D, this would skyrocket to six dimensions (each location has three coordinates and can transition to any other location in 3-D, adding another three coordinates) unless Toeplitz structure is assumed (allowing a three-dimensional $p$). Therefore, for the purposes of the simulations that follow, the OTF outlined in Section II will be assumed to apply to each 3-D location in the image (a circulant assumption).

By approximating Toeplitz structure as circulant, a minor inaccuracy is created. Described another way, multiplication in the Fourier domain parallels circular convolution in the time domain. Circular convolution is not what we have in the EM equation, hence the inaccuracy. This can be mostly circumvented, however, through the use of zero-padding. By increasing the volume of the received image by tacking on zeros at both ends in each direction, artifacts created by this circular convolution issue are almost entirely eliminated.
Such artifacts will tend to appear in the padded regions, which are then eliminated once the image has been fully reconstructed.

An alternative to the stopped-EM algorithm is the maximum penalized likelihood estimator (MPLE). This method balances the log-likelihood of the image estimate against some sort of penalty, often based on the complexity of the estimate’s representation [2,12,13]. It can be expressed mathematically as follows:

\[
\hat{\theta} = \arg \max_{\theta} \{- \log \Pr[y | \theta; \rho] + \text{pen}(\theta)\}
\]

where \(\theta\) is the estimate, potentially equivalent to \(\lambda^{\text{new}}\) in the preceding discussion.

As mentioned above, the penalty is often a measure of the representation’s complexity. Complexity can be measured any number of ways but often is simply proportional to the number of data elements required to represent an estimate. This definition of complexity works well with the Haar wavelet transform, the method utilized in the simulations to follow. Section IV provides an in-depth description of the Haar wavelet basis and the application of complexity regularization to it.

Suffice to say at this point that methods using multiscale representations of the data in a complexity-regularized maximum penalized likelihood estimation (MPLE) setting have been shown in 2-D to offer significant improvements in reconstruction results. In addition, convergence is far easier to accomplish than in stopped-EM methods [2], and near optimal error bounds are achievable [14]. The results of the research outlined here demonstrate that similar gains are realizable for 3-D reconstruction. By employing the MPLE in the maximization step of each iteration of the EM algorithm, a so-called EM-MPLE method can be implemented.

IV. MULTISCALE ANALYSIS AND WAVELETS
Although physiology and anatomy dictate that much of medical images should be expected to be smooth, the interesting places are the discontinuities and edges. The wavelet transform, a multiscale approach, is adept at edge detection while sparsely representing smooth regions, leading to a highly accurate representation of an original image with few elements. Multiscale, or multiresolution, analysis refers to the study of signal behavior at a variety of scales or resolutions. Recursive algorithms or basis functions such as wavelets are designed with the inherent ability to take advantage of changes in scale.

A complete multiresolution formulation of wavelets can be found in [15]. The first step is to define a set of scaling functions,

$$\varphi_k(t) = \varphi(t - k) \text{ where } k \text{ is an integer.}$$

The span of the scaling functions is $\nu_0$, and every other function in $\nu_0$ is a weighted sum of the scaling functions. A change in resolution can be represented by scaling and translating the basic scaling function as follows:

$$\varphi_{j,k}(t) = 2^{j/2} \varphi(2^j t - k)$$

Here, the span of $\varphi_{j,k}(t)$ is $\nu_j$. These spanning spaces can be nested to set up basic multiscale analysis:

$$\nu_j \subset \nu_{j+1} \text{ where } j \text{ is any integer.}$$

This allows for the basic scaling function to be represented as a weighted sum of higher resolution scaling functions as follows:

$$\varphi(t) = \sum_n h(n) \sqrt{2} \varphi(2t - n)$$
The Haar wavelet basis will be the focus of the algorithm devised in this paper, for reasons to be discussed shortly. Its scaling function is a pulse of unit width and height. This makes the following true [15]:

$$\varphi(t) = \varphi(2t) + \varphi(2t - 1)$$

The coefficients that apply are $h(0) = h(1) = 1/\sqrt{2}$.

The other component of a wavelet basis is the wavelet function. Wavelet functions are typically orthogonal to the scaling function to simplify coefficient calculation. The span $\nu_{j+1}$ is therefore made up of $\nu_j$ and $w_j$, the wavelet function's span. Because $w_0 \subset \nu_1$, the wavelet function can also be defined as a weighted sum of higher resolution scaling functions [15].

$$\psi(t) = \sum_n h_1(n) \sqrt{2} \varphi(2t - n)$$

This leads to a pair of wavelet function equations that parallel those for the scaling function.

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$$

$$\psi(t) = \varphi(2t) - \varphi(2t - 1)$$

The second equation is again Haar-specific, and has corresponding coefficients $h_1(0) = 1/\sqrt{2}$ and $h_1(1) = -1/\sqrt{2}$. Of note is the property, derived in [15], that $h_1(n) = (-1)^n h(1 - n)$.

With the formulas established at this point, the Discrete Wavelet Transform (DWT) can now be detailed. Any function can be written as the following weighted sum of scaling and wavelet coefficients:

$$g(t) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(t) + \sum_{j=j_0}^{\infty} \sum_{k} d_j(k) \psi_{j,k}(t)$$
The coarsest scale is set by $j_0$. The DWT is made up of the coefficients of this expansion [15], and are found by

$$c_j(k) = \langle g(t), \varphi_{j,k}(t) \rangle \quad \text{and} \quad d_j(k) = \langle g(t), \psi_{j,k}(t) \rangle$$

These straightforward results are a direct consequence of choosing wavelet functions that form an orthonormal basis, hence the near universality of the choice. Another key advantage of orthonormality is that Parseval’s Theorem can equate signal energy to wavelet coefficient energy [15].

$$\int |g(t)|^2 \, dt = \sum_{l=-\infty}^{\infty} |c_l|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_j(k)|^2$$

In 2-D (i.e., images), the time domain provides detailed information about image space, whereas the Fourier domain provides detailed frequency information. The wavelet domain has the invaluable property that it simultaneously provides detailed information about spatial and frequency relationships within an image.

Wavelets are commonly used in image reconstruction [2,16-19], but are generally difficult to apply to non-Gaussian scenarios. Usually, Gaussian approximations of Poisson data may only be made if spatial resolution is sacrificed to achieve higher counts in larger bins, an undesirable solution. However, the Haar wavelet basis has the rare ability to represent Poisson-distributed data well. The distribution reproduces under unweighted summation, and so Haar scaling coefficients at any scale are Poisson, whereas other wavelets result only in arbitrary linear combinations of Poisson random variables and are difficult to analyze [20]. In addition, it has been demonstrated that complexity penalized Haar wavelet estimation methods exhibit near minimax optimality [14].
To clarify the 1-D Haar description touched upon earlier, let us assume compact support over \(0 \leq t \leq 1\). The Haar scaling function \(\varphi(t)\) is then 1 over the support and 0 otherwise. The Haar wavelet function \(\psi(t)\) is 1 in the region \(0 < t < 0.5\), -1 in the region \(0.5 < t < 1\), and 0 otherwise [15]. Figure 9 depicts the 1-D Haar scaling and wavelet functions.

![Figure 9: Haar Scaling Function and Wavelet Function](image)

In 2-D, the Haar wavelet system is analogous to dividing an image in a dyadic fashion and performing piecewise constant estimation [2]. The Haar system functions similarly in 3-D to the way it does in 2-D. The image that results from one pass of a Haar transform is split into eight dyadic cubes, rather than four dyadic squares. These eight children can potentially be recursively split as well, though in practice only one of them is. (It will be identified momentarily.) The number or depth of these recursive splits relates directly to the multiscale nature of wavelets. The eight children, each containing the original data decimated by 2 in each dimension (for a volume decimation of eight), can be thought of as being derived in the following manner. One of them has undergone a low-pass filter in the \(x\)-direction, a low-pass filter in the \(y\)-direction, and a low-pass filter in the \(z\)-direction. Any or all of these low-pass filters may be replaced with high-pass ones to achieve the other seven children. There are eight permutations and therefore eight children. The only one that is recursively acted upon by the Haar system (or any wavelet system in general) is the one spelled out above: low, low, low.
Complexity regularization, as mentioned above, can play the penalization role in the MPLE. The complexity of an image's Haar wavelet transform is simply the number of non-zero wavelet coefficients. Noise reduction can be quickly and easily achieved by zeroing small coefficient values. The sparseness of the wavelet transform allows this procedure to effectively kill noise while maintaining the desired signal. The MPLE selects the Haar wavelet coefficients that lie below a chosen threshold and are to be eliminated. The choice of this threshold brings the penalization aspect to the estimator. A large threshold keeps fewer elements and produces smooth estimators while small thresholds produce complicated estimators [2]. Section V includes a discussion of the threshold chosen for use in these simulations. An additional requirement of proper Haar wavelet thresholding is that a child coefficient value can be kept only if its parent values are all kept as well, but this requirement is achieved automatically through the recursive multiscale definition of the Haar wavelet basis. By defining the Haar wavelet transform recursively, the elimination of a parent coefficient automatically eliminates all of its children.

V. THE 3-D EM-MPLE ALGORITHM WITH HAAR THRESHOLDING

Using the various ideas and algorithms discussed in the previous sections, a complete complexity-regularized EM-MPLE reconstruction using Haar wavelet thresholding can be devised in three-dimensional space. The algorithm strongly resembles the basic EM algorithm discussed in-depth in Section III. Added at the maximization step of each iteration of the basic algorithm is the Haar thresholding, the preferred complexity-regularized MPLE.

A threshold must be chosen for the wavelet-based penalization scheme. The standard deviation of the noise seems an obvious selection. By eliminating wavelet coefficients below
the noise's standard deviation, a large portion of the noise can be eliminated altogether while maintaining much of the target original image (assuming, of course, that the signal-to-noise ratio (SNR) is reasonable). The variance of the Poisson data is of course the Poisson mean, which is data-dependent. However, taking the square root of the Poisson data has the interesting result of making the data behave as Gaussian with variance of 0.25, regardless of the Poisson mean. This is a variant of the Anscombe transform [23], and this data-independent result provides a convenient threshold for the Haar wavelet transform.

Removing (by setting to zero) all coefficients less than 0.5 (the standard deviation) should therefore eliminate the majority of noise-induced corruptions of the data. So the process for Haar thresholding is to take an estimate, calculate its square root, pass it through a wavelet transform, eliminate values less than 0.5, pass it back through the inverse wavelet transform, and square the data.

As discussed in Section III, zero-padding of the data is employed to eliminate artifacts caused by the approximation of the OTF, a Toeplitz function, as circulant. For the simulation data, which will be described later as 48-by-48-by-48, this means 8 voxels (simply 3-D pixels, in this usage) are added to each end of each of the three dimensions, resulting in a 64-by-64-by-64 image for processing. The estimator is initialized such that every voxel contains the mean value of the received data.

The OTF discussed in Section II for confocal one-photon excitation microscopy requires three constants to be calculated in full, all relating to the specific microscope setup, and all located within the frequency normalizations. For the synthetic data in the simulations that follow, the constant values used in the OTF to blur the true data before it is Poisson randomized were as follows:
refractive index \( n_i = 1.0 \)

wavelength \( \lambda = 488 \text{ nm} \)

aperture angle \( \alpha = 50^\circ \)

For the real data, the refractive index becomes 1.3, and the wavelength of light used is 525 nm.

Real data, however, is a bit more complicated, as the tissue sample has its own transfer function in addition to that of the microscope. For this reason, simply plugging the parameters into the OTF may not give the best possible transfer function estimate. If supplied with test images with known structure, an alternate approximate transfer function can be calculated through deconvolution, and this can either be used in this form or enhanced in a number of ways. A useful confocal test image involves the placement of a fluorescent bead in the sample. This bead has known spherical dimensions. A pseudo-original image can be synthesized with a sphere of the appropriate size located in the approximate position of the bead in the test image. Let us refer to the FFT of the test image as \( X \) and the FFT of the pseudo-original as \( Y \). The deconvolution can be performed in the frequency domain as a division. The transfer function estimate is therefore

\[
H = \frac{X}{Y}
\]

As mentioned, this estimate can be enhanced. First, because \( X \) is noisy and \( Y \) is not, \( H \) will be a noisy approximate of the true transfer function. We expect \( H \) to be low-pass, based on the general form of the OTF outlined in Section II. For this reason, it is sensible to zero \( H \) for frequencies above a value chosen carefully based on careful examination of the estimate. This will eliminate high-frequency noise with little or no effect on the low-frequency regions where the true transfer function resides.
Another option, potentially performed in addition to (following) the previous enhancement, is to assume the OTF formula is correct and estimate the parameters that give the closest approximation of H. This can be done in an MSE sense as follows:

\[ \| H(a) - X/Y \|^2 \]

The variable (potentially a vector) \( a \) represents the various parameters of the OTF, either individually or combined in some reasonable fashion. The correct parameters could be calculated if \( H(a) \) were linear or searched for exhaustively otherwise. The latter is the case with the confocal microscopy OTF.

For each data set (a single 3-D image), the stopped-EM is compared against the EM-AMPLE algorithm whose description culminates in this section. For simulated data, the best result for each case will be compared even though what is termed “best” is unknown (or at least poorly defined) with real data, as the original, target image is unknown. For test data, visual inspection is the only realistic measure of success.

VI. Simulation Results

The real-life confocal microscopy specimens utilized in these simulations are neuronal structures such as dendrites, which are long, thin, tubular structures passing through several axial layers. Simulation data using thin cylinders are therefore appropriate, and they should exist at least partially in the axial direction, passing between many slices. For our experiments, we constructed a phantom consisting of a thin, angled cylinder as approximated in Figure 10. The cylinder is rotated 45° from the positive x-axis in the x-y plane and then 45° axially above that same plane.
Figures 10 through 16 demonstrate one example. Note that all of the images are 48 voxels in each dimension and are displayed in a tiled fashion. This results in an overall image that is 8 tiles by 6 tiles, with each tile 48-by-48. The upper left 48-by-48 section of the images below is therefore the topmost slice of the 3-D image. The 48-by-48 section immediately below this is one slice below the topmost, et cetera. Recall that, although all of the images shown of the synthetic data are shown as 48-by-48-by-48 (tiled), during the processing they are actually 64-by-64-by-64 as a result of zero-padding (see Section III).

The first image, Figure 11, shows the synthetic specimen, $\lambda$ in the equations in Section III. The second image, Figure 12, shows the simulated data blurred using the OTF
from Section II and corresponds with $\mu$
above. Note the nature of the blurring; it is
almost entirely axial in nature. Figure 13, $y$
above, shows a simulated confocal image
taken of this specimen, with the blurred
image received as Poisson noise. Figure 14
shows the best result achievable in terms of
mean-squared error (MSE) by a stopped-EM
method, while Figure 15 shows the result of
the multiscale EM-MPLE method. Finally,
the sixth image, Figure 16, compares the
relative MSE for the two methods, with the

Figure 11: Original Data, $\lambda$

Figure 12: Blurred Data, $\mu$

Figure 13: Received Poisson Data, $y$
Figure 14: Best EM Result

basic EM algorithm on the upper line and EM-MPLE the lower line. In this image, the $x$-axis shows the iteration count, while the $y$-axis shows MSE relative to the received data (third image). So an estimate with an MSE of 1 would be no better (in MSE terms) than the received data.

The multiscale method achieves a more accurate reconstruction both visually and in terms of mean squared error. Visually, more of the blurring has been eliminated and the estimate is almost everywhere more accurate. In MSE terms, the EM-MPLE method is again superior over the entire range of iterations. Additionally, the divergence issues with the basic EM algorithm become clear through examination of these error curves, revealing the EM-MPLE method as the more reliable algorithm. Before one hundred iterations have been run, the basic EM algorithm ceases its improvement and begins to generate worse estimates with each successive attempt. The image shown in Figure 14 is from around iteration seventy.
Figure 16: MSE Comparison
The upper line is the EM result; the lower is the EM-MPLE result.

Ideally, a stopped-EM algorithm would successfully cut off around this point, but as discussed in Section III, this is not an easy task to accomplish. Because the original specimen is unknown in a real situation, the stopped-EM will tend to perform even more poorly than a convergent algorithm.

VII. CONCLUSIONS AND FUTURE WORK

Medical imaging systems can prove invaluable aids in modern diagnosis and treatment methods, but procedures for repairing noise and blurring in the created images, when done right, can further elevate the utility of such systems. For a setup like that found in confocal microscopy, in which image corruption occurs primarily in the axial direction, slice-
wise 2-D reconstruction algorithms miss the boat entirely. A 3-D reconstruction scheme is critical for successful recovery of the true image.

The MLE, in the form of the EM algorithm, is the most commonly selected method for medical imaging reconstruction due to its simplicity and ease of use. However, its unreliability and poor convergence behavior, even in its stopped-EM variation, beg for an alternative solution. A penalized estimator, using multiscale complexity regularization in the form of wavelets, can provide improved convergence as well as historically (in other problems) greater success. In particular, the Haar wavelet basis provides the proper setting for the Poisson data involved in medical imaging.

The EM-MPLE algorithm devised in this paper is performed in 3-D to make the most of the knowledge of the confocal microscope's transfer function. It has been shown to perform better, both visually and in an MSE sense, on simulated data.

The algorithm described in this treatise takes advantage of many facets of the specific problem, but the Haar thresholding solution for the MPLE may be improved upon. Ongoing research in 2-D image reconstruction attempts to move beyond wavelets. It seems likely that such innovative methods as wedgelets [21], platelets [2], or beamlets [22], among others, may be able to act as replacements for the Haar wavelet thresholding. These groundbreaking algorithms have been shown to drastically improve on results obtained using more traditional wavelets when applied to the right problems. The key to successful implementation of alternative algorithms will be their ability to translate to 3-D.

Multiscale methods have the tremendous advantage of allowing resolution to be variable across an image. Highly detailed regions can receive correspondingly high resolutions, while smoothly varying regions can be allotted low resolutions to conserve
complexity. Unfortunately, most existing imaging systems undercut this fantastic advantage by binning photon events. This limits the achievable resolution unnecessarily. A preferable setup would be for the imaging system to record the exact location of each photon event rather than binning them. This would allow the reconstruction algorithm to determine the most appropriate resolution for each portion of the image [2].

In conclusion, through proper analysis of a problem, even relatively simple algorithms can become more effective, as insight gained can be applied to the solution. In medical imaging, simplicity and ease of use become far more important factors in the balance with reliability and accuracy. Any useful, successful algorithm must take all of these factors into account. And finally, to create the best possible medical images, the designers of the imaging systems must work together with the designers of the reconstruction algorithms to take full advantage of what each component can offer.
REFERENCES


