Model-based image reconstruction from time-resolved diffusion data

Suhail S. Saquib[a], K. M. Hanson[b], and G. S. Cunningham[b]

[a]School of Electrical & Computer Engineering, Purdue University
West Lafayette, IN 47907 USA

[b]Los Alamos National Laboratory, MS P940
Los Alamos, New Mexico 87545 USA

ABSTRACT

This paper addresses the issue of reconstructing the unknown field of absorption and scattering coefficients from time-resolved measurements of diffused light in a computationally efficient manner. The intended application is optical tomography, which has generated considerable interest in recent times. The inverse problem is posed in the Bayesian framework. The maximum *a posteriori* (MAP) estimate is used to compute the reconstruction. We use an edge-preserving generalized Gaussian Markov random field to model the unknown image. The diffusion model used for the measurements is solved forward in time using a finite-difference approach known as the alternating-directions implicit method. This method requires the inversion of a tridiagonal matrix at each time step and is therefore of $O(N)$ complexity, where N is the dimensionality of the image. Adjoint differentiation is used to compute the sensitivity of the measurements with respect to the unknown image. The novelty of our method lies in the computation of the sensitivity since we can achieve it in $O(N)$ time as opposed to $O(N^2)$ time required by the perturbation approach. We present results using simulated data to show that the proposed method yields superior quality reconstructions with substantial savings in computation.

Keywords: optical tomography, diffusion, Bayesian estimation, MAP, Markov random field, adjoint differentiation, finite-difference.

1. INTRODUCTION

Medical optical tomography has generated considerable interest in recent times.[1] The advantage of using infrared light as an imaging modality stems from the fact that it is non-ionizing and hence can be used for continuous monitoring. More importantly, it is capable of revealing the functioning of the tissues as opposed to other standard imaging modalities that reflect only the static physical structure. A final advantage is that optical imaging systems can be made portable, making them useful in clinical situations such as surgery, trauma, and intensive care.

An accurate model for the propagation of photons through tissue can be obtained from transport theory. There are two basic approaches using this theory: an essentially discrete model of individual photon interactions, such as Monte-Carlo,[2] or a continuous model based on a differential equation approximation, such as the diffusion equation. While the Monte-Carlo method is more generally applicable, it is computationally expensive to implement. On the other hand, the diffusion approximation is accurate for highly scattering media (which is the case for tissues) while being computationally tractable. Therefore we will use the diffusion equation as our data model.

The inverse problem of reconstructing the absorption and scattering coefficients from diffuse measurements of light is highly nonlinear. To facilitate the computation of the unknown coefficients, several approaches attempt to locally linearize the original inverse problem. The linear perturbation model[3–5] is one such method that employs a Taylor series approximation about a reference distribution for the unknown coefficients to obtain a set of linear equations. The associated weight matrix is computed using Monte-Carlo simulations.[6] However, the utility of this method is limited if we do not have a reference distribution that is close to the actual distribution of the unknown coefficients. Alternatively, the Newton-Raphson (NR) method employs a Taylor series expansion about the current estimate of the unknown coefficients to obtain a more refined estimate. This procedure is then iterated until convergence. The NR method has also been used with the Levenberg-Marquardt procedure to control the nonlinearity of the underlying problem.[7] Other numerical optimization methods that have been used include projection onto convex

sets (POCS),[3] simultaneous algebraic reconstruction technique (SART),[5] SART-type algorithm,[8] and conjugate gradient descent (CGD).[6] We will be using a modified version of CGD for our optimization because of its superior convergence properties. The modification involves replacing simple line searching with bent-line searching to enforce the positivity constraint on the unknown coefficients.

The inverse problem is known to be ill-posed and some form of regularization is necessary to make the solution space well-behaved. Toward this end, we formulate the inverse problem in the Bayesian framework and use the maximum *a posteriori* (MAP) estimation criterion to compute the reconstruction. This approach enables us to incorporate *a priori* knowledge of the unknown field through an image model to regularize the solution. The regularization methods that have previously been employed include truncated singular-value decomposition, Tikhonov and constraint-based regularization that impose a penalty on the norm of the solution vector.[7] However, none of these methods properly model edges that are normally present in the unknown field for real objects, resulting in smooth reconstructions with blurred edges. An important contribution of the present work is the use of an edge-preserving generalized Gaussian Markov random field (GGMRF) model[9] for the unknown field.

To compute the likelihood of the data, we need to solve the diffusion equation forward in time. The finite element method (FEM) has been widely used for this purpose.[7,10,11] However, we propose to do this by discretizing the diffusion equation using a finite-difference approach.[12] The discretization can be accomplished in a number of ways since the spatial derivatives can be evaluated at the present (implicit) or the past (explicit) time instance. In particular, we use an alternating-directions method[13] that computes the spatial derivative implicitly for one spatial direction and explicitly for the other spatial direction in the first half of the time step. In the next half time step, the implicit and explicit directions are switched. This method is known to be stable even for large time steps. Also, by virtue of the alternating-directions, the resulting matrix that needs to be inverted in the forward computation is tridiagonal. The inversion can therefore easily be done in $O(N)$ time, where N is the number of pixels in the unknown image.

Since we will use CGD to compute the MAP estimate, we require the gradient of data likelihood with respect to the unknown field. The novelty of our approach lies in the proposed method of this computation. We show that by working backward in time, and using the discretized equations that are employed to compute the forward solution, the gradient computation parallels the forward computation in complexity and can be accomplished in $O(N)$ time. This is in contrast to the computationally intensive perturbation approach, which is widely used to compute the gradient, but requires N forward computations and is therefore $O(N^2)$. The method we propose is known as adjoint differentiation,[14] which has been used to solve oceanographic and other computationally intensive inverse problems.

We present experimental results using simulated data to show that the proposed method results in superior quality reconstructions with substantial savings in computation.

2. DIFFUSION DATA MODEL

Let $U(x,y,t)$ be the intensity of light and $R(x,y,t)$ be the source strength at position (x,y) and time t. Let $\mu_a(x,y)$ and $\mu_s(x,y)$ denote the space-varying absorption and scattering coefficients. Let $D(x,y)$ denote the diffusion coefficient, given as

$$D(x,y) = \frac{c}{3[\mu_a(x,y) + (1-g)\mu_s(x,y)]} \quad , \tag{1}$$

where c is the speed of the light in the medium and g is the scattering anisotropy parameter, equal to the average cosine of the scattering angle distribution*. Then the diffusion equation is given as

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x} D \frac{\partial U}{\partial x} + \frac{\partial}{\partial y} D \frac{\partial U}{\partial y} - c\mu_a U + R \quad , \tag{2}$$

where the spatial and temporal dependence of the parameters has been suppressed. In this paper, we parametrize the inverse problem in terms of D and μ_a. This is equivalent to recovering μ_a and μ_s due to Eq. (1).

*$\mu_s' = \mu_s(1-g)$ is known as the effective scattering coefficient

2.1. Notation

Let S denote the set of discrete lattice points and let $s \in S$ denote the spatial position of a particular lattice point. In some instances we need to distinguish between the two spatial directions. In this case we use the subscript (i, j) to denote the spatial position. When the spatial position subscript is present, the resulting quantity is a scalar with corresponding value at that spatial position (e.g. U_s or $U_{i,j}$). When the spatial position subscript is dropped, the resulting quantity is a column vector obtained by either row-ordering or column-ordering the corresponding two-dimensional field (e.g. U). We will use the superscript n to denote the discretized time index.

2.2. Measurement Model

Let $\mathcal{M}$ denote the set of detector positions and $\mathcal{T}$ denote the set of time indices when the measurements are recorded. Let Y denote the measurements of the diffuse intensity U for all $s \in \mathcal{M}$ and $n \in \mathcal{T}$. In the interest of simplicity, we assume that the measurements are corrupted by uncorrelated Gaussian noise. However, the method we propose is not restricted to this choice. The log-likelihood of the observations Y given D and μ_a is

$$\log P(Y|D, \mu_a) = -\sum_{s \in \mathcal{M}} \frac{1}{2\sigma_s^2} \sum_{n \in \mathcal{T}} (Y_s^n - U_s^n)^2 \quad , \tag{3}$$

where σ_s^2 is the noise variance at spatial position s.

2.3. Computation Of $\log P(Y|D, \mu_a)$

To compute the log-likelihood of the measurements Y given D and μ_a, we need to solve the diffusion equation (2) forward in time to obtain the diffuse intensity U_s^n for all time $n \in \mathcal{T}$ and spatial positions $s \in \mathcal{M}$. We propose to do this using a finite-difference approach where the spatial and temporal derivatives in Eq. (2) are replaced by their finite-difference approximations as follows

$$\frac{\partial}{\partial x} D \frac{\partial U}{\partial x} \approx \frac{D_{i+1/2,j}(U_{i+1,j} - U_{i,j}) - D_{i-1/2,j}(U_{i,j} - U_{i-1,j})}{\Delta^2} \triangleq \delta_x(U_{i,j}) \tag{4}$$

$$\frac{\partial}{\partial y} D \frac{\partial U}{\partial y} \approx \frac{D_{i,j+1/2}(U_{i,j+1} - U_{i,j}) - D_{i,j-1/2}(U_{i,j} - U_{i,j-1})}{\Delta^2} \triangleq \delta_y(U_{i,j}) \tag{5}$$

$$\frac{\partial U}{\partial t} \approx \frac{U^{n+1} - U^n}{\Delta t} \quad , \tag{6}$$

where $\Delta = \Delta x = \Delta y$ is assumed for simplicity. The approximations are obtained by simply differencing the second partial derivatives and centering each term appropriately. In doing so, we require the interpolated value of D halfway between the grid points. Linear interpolation is used to achieve this, i.e. $D_{i+1/2,j} = (D_{i,j} + D_{i+1,j})/2$.

By substituting the finite-difference approximations (4-6) in the diffusion equation (2), we obtain a difference equation that needs to be solved forward in time. When solving the difference equation for U^{n+1}, the finite-difference approximations to the spatial derivatives (4,5) can be evaluated at time index $n+1$ or n. The three methods we discuss in the following sections differ in this choice of the time index.

2.3.1. Explicit method

In this method, the spatial derivatives (4,5) are evaluated at the past time instance n when computing the diffuse intensity U^{n+1}. Substituting Eqs. (4-6) in Eq. (2) and using time index n for the spatial derivatives, we obtain

$$U_{i,j}^{n+1} = (1 - c\mu_a \Delta t) U_{i,j}^n + \delta_x(U_{i,j}^n)\Delta t + \delta_y(U_{i,j}^n)\Delta t + (R_{i,j}^{n+1} + R_{i,j}^n)\frac{\Delta t}{2} \quad . \tag{7}$$

Rewriting the above equation in vector-matrix notation, we obtain

$$U^{n+1} = B\, U^n + \bar{R}^{n+1/2} \quad , \tag{8}$$

where U^{n+1} and U^n are row-ordered column vectors and B is a sparse matrix with four off-diagonal elements (corresponding to the vertical and horizontal neighbors of a pixel). $\bar{R}^{n+1/2}$ denotes the integrated source strength

between time instances n and $n+1$. Using Eq. (8), we can compute U^n for any n by starting at $n=0$ and moving forward in time. Furthermore, since B is a sparse matrix, the forward computation is $O(N)$, where N is the total number of discretized spatial positions. However, the disadvantage of this simple method is that it becomes unstable when[13]

$$\Delta t > \frac{\Delta^2}{4(\max_{s \in S} D_s)} \quad .$$

This method is not very useful because it can dictate very small time steps in the forward simulation.

2.3.2. Implicit method

In this method, the spatial derivatives (4,5) are evaluated at the present time instance $(n+1)$ when computing the diffuse intensity U^{n+1}. Substituting Eqs. (4-6) in Eq. (2) and using time index $n+1$ for the spatial derivatives, we obtain

$$(1 + c\mu_a \Delta t) U_{i,j}^{n+1} - \delta_x(U_{i,j}^{n+1})\Delta t - \delta_y(U_{i,j}^{n+1})\Delta t = U_{i,j}^n + (R_{i,j}^{n+1} + R_{i,j}^n)\frac{\Delta t}{2} \quad . \tag{9}$$

In vector-matrix notation the above equation is given as

$$A\, U^{n+1} = U^n + \bar{R}^{n+1/2} \quad , \tag{10}$$

where A is a sparse matrix having exactly the same structure as the B matrix in Sect. 2.3.1. The advantage of this method is that it is unconditionally stable for any value of Δt. However, the computation of U^{n+1} from U^n now requires the inversion of matrix A, which is not trivial anymore in terms of computation. Multigrid relaxation methods[15] can be employed to efficiently solve Eq. (10) but we prefer to use the method described in the next section.

2.3.3. Alternating-directions implicit method (ADI)

In this method, the computation of U^{n+1} from U^n is broken up in two time steps.[13] In the first half time step, only the spatial derivative in one direction is evaluated at the present time instance (implicit) and the other spatial derivative is evaluated at the previous time instance (explicit). In the next half time step, the implicit and explicit directions are switched. The difference equations for the two half time steps are given as

$$(1 + c\mu_a \frac{\Delta t}{2}) U_{i,j}^{n+1/2} - \delta_y(U_{i,j}^{n+1/2})\frac{\Delta t}{2} = U_{i,j}^n + \delta_x(U_{i,j}^n)\frac{\Delta t}{2} + (R_{i,j}^{n+1} + R_{i,j}^n)\frac{\Delta t}{4} \tag{11}$$

$$(1 + c\mu_a \frac{\Delta t}{2}) U_{i,j}^{n+1} - \delta_x(U_{i,j}^{n+1})\frac{\Delta t}{2} = U_{i,j}^{n+1/2} + \delta_y(U_{i,j}^{n+1/2})\frac{\Delta t}{2} + (R_{i,j}^{n+1} + R_{i,j}^n)\frac{\Delta t}{4} \quad . \tag{12}$$

The difference Eqs. (11,12) for both halves of the time step can be compactly represented in vector-matrix notation by the single equation

$$A\, U^{n+1/2} = B\, U^n + \bar{R}^{n+1/4} \tag{13}$$

if we use row-ordering for U in Eq. (11), column-ordering for U in Eq. (12) and let n take on fractional values. By switching the ordering, we force the structure of matrices A and B to remain the same while the absolute values differ for both halves of the time step. However, for the sake of notational simplicity, we will not distinguish between the two halves of each time step and use Eq. (13) to represent both of them.

To compute $U^{n+1/2}$ from U^n, we need to invert A. However, in this case A is always tridiagonal due the spatial derivative being implicit only in one direction. The inversion can therefore be done in $O(N)$ computation. Furthermore, the method is also unconditionally stable for any value of Δt. In the rest of the paper, we assume that the ADI method will be used to do the forward computation.

2.4. Sensitivity Computation

Let $\theta = [D\ \mu_a]^T$ be the column vector of the unknown parameters D and μ_a. Define

$$\phi(\theta) \triangleq \log P(Y|\theta) \quad .$$

To facilitate the solution of the inverse problem, we require the derivative or sensitivity of $\phi(\theta)$ with respect to θ. The technique that we propose for this computation is known as adjoint differentiation.[14] The method requires us to work backwards in time using the same discretized equations that were used to compute the forward solution in Sect. 2.3.3. The sensitivity of ϕ with respect to θ is obtained by computing the intermediate sensitivity of ϕ with respect to the diffuse intensity U.

2.4.1. Sensitivity computation with respect to U

The sensitivity of ϕ with respect to U^n is obtained recursively by using the sensitivity of ϕ with respect to $U^{n+1/2}$. Application of the chain rule yields

$$\begin{aligned}\frac{d\phi}{dU_s^n} &= \sum_{r\in S}\frac{d\phi}{dU_r^{n+1/2}}\frac{dU_r^{n+1/2}}{dU_s^n}+\frac{\partial\phi}{\partial U_s^n}\\ &= \left[\frac{dU^{n+1/2}}{dU_s^n}\right]^T\frac{d\phi}{dU^{n+1/2}}+\frac{\partial\phi}{\partial U_s^n} \quad \text{for all } s\in S, \end{aligned} \tag{14}$$

where $\frac{\partial\phi}{\partial U_s^n}$ denotes the change in ϕ when only U_s^n is varied keeping all other variables constant while $\frac{d\phi}{dU_s^n}$ denotes the total change in ϕ when U_s^n is varied along with all variables that depend on U_s^n. Differentiating Eq. (3) with respect to U_s^n, we obtain

$$\frac{\partial\phi}{\partial U_s^n}=\begin{cases}\frac{1}{\sigma_s^2}(Y_s^n-U_s^n) & s\in\mathcal{M},\ n\in\mathcal{T}\\ 0 & \text{otherwise}\end{cases}.$$

Rewriting Eq. (14) in vector-matrix notation, we obtain

$$\frac{d\phi}{dU^n}=\left[\frac{dU^{n+1/2}}{dU^n}\right]^T\frac{d\phi}{dU^{n+1/2}}+\frac{\partial\phi}{\partial U^n}\ , \tag{15}$$

where the matrix $\frac{dU^{n+1/2}}{dU^n}$ is given as

$$\frac{dU^{n+1/2}}{dU^n}=\left[\frac{dU^{n+1/2}}{dU_{s_1}^n}\ \frac{dU^{n+1/2}}{dU_{s_2}^n}\ \cdots\ \frac{dU^{n+1/2}}{dU_{s_N}^n}\right] \quad \{s_1,\ldots,s_N\}\in S. \tag{16}$$

Differentiating Eq. (13) with respect to U_s^n, we obtain

$$\frac{dU^{n+1/2}}{dU_s^n}=A^{-1}B\,I_s \quad \text{for all } s\in S\ , \tag{17}$$

where I_s is column vector that is zero everywhere except at spatial point s where it is unity. Using Eqs. (17), (16) and (15), we obtain the sensitivity of ϕ with respect to U^n as

$$\frac{d\phi}{dU^n}=B^T(A^{-1})^T\frac{d\phi}{dU^{n+1/2}}+\frac{\partial\phi}{\partial U^n}\ . \tag{18}$$

Since A is tridiagonal and B is a sparse matrix with only two off-diagonal elements, the above computation can be done in $O(N)$ time.

2.4.2. Sensitivity computation with respect to θ

The chain rule is applied again to obtain the sensitivity of ϕ with respect to θ_r as

$$\frac{d\phi}{d\theta_r}=\sum_n\sum_{s\in S}\frac{d\phi}{dU_s^{n+1/2}}\frac{\partial U_s^{n+1/2}}{\partial\theta_r} \quad r=1,\ldots,2N.$$

Rewriting the above equation in vector-matrix notation, we obtain

$$\frac{d\phi}{d\theta}=\sum_n\left[\frac{\partial U^{n+1/2}}{\partial\theta}\right]^T\frac{d\phi}{dU^{n+1/2}}\ , \tag{19}$$

where

$$\frac{\partial U^{n+1/2}}{\partial\theta}=\left[\frac{\partial U^{n+1/2}}{\partial\theta_1}\ \frac{\partial U^{n+1/2}}{\partial\theta_2}\ \cdots\ \frac{\partial U^{n+1/2}}{\partial\theta_{2N}}\right]\ . \tag{20}$$

Taking the partial derivative of Eq. (13) with respect to θ_r, we obtain

$$\frac{dA}{d\theta_r}U^{n+1/2} + A\frac{\partial U^{n+1/2}}{\partial \theta_r} = \frac{dB}{d\theta_r}U^n$$

$$\frac{\partial U^{n+1/2}}{\partial \theta_r} = A^{-1}\overbrace{\left(\frac{dB}{d\theta_r}U^n - \frac{dA}{d\theta_r}U^{n+1/2}\right)}^{X_r} \quad . \tag{21}$$

Using Eqs. (20) and (21), we obtain

$$\frac{\partial U^{n+1/2}}{\partial \theta} = A^{-1}[X_1 \; X_2 \; \ldots \; X_{2N}] = A^{-1}X \quad . \tag{22}$$

Substituting Eq. (22) in Eq. (19), we obtain

$$\frac{d\phi}{d\theta} = \sum_n X^T (A^{-1})^T \frac{d\phi}{dU^{n+1/2}} \quad . \tag{23}$$

This computation can again be done in $O(N)$ time since A is tridiagonal and X is sparse. By starting at the last time instance and using the recursion in Eq. (18), the contribution at each time step can be accumulated using Eq. (23) to obtain the complete sensitivity with respect to θ in $O(N)$ time.

3. IMAGE MODEL

The problem of reconstructing the unknown parameters D and μ_a from the measurements Y is an ill-posed inverse problem and some form of regularization is necessary to make the solution well behaved. This is accomplished by incorporating an image model in the reconstruction process, which models our *a priori* knowledge regarding the unknown fields D and μ_a. The following development is for D; however, the development for μ_a follows the same pattern.

Markov random fields (MRF) have been extensively used in image processing applications since they yield a rich class of models with relatively simple parameterization. With this in view, we model D as a MRF or, equivalently, a Gibbs distribution with the following form

$$P(D) = \begin{cases} \frac{1}{z(\sigma_D)} \exp\{-u(D/\sigma_D)\} & \text{if } D \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad , \tag{24}$$

where σ_D is the scale parameter of the model and $z(\cdot)$ is the normalizing constant of the distribution known as the partition function. We impose a nonnegativity constraint on D because negative values of D are physically meaningless. We restrict the energy function $u(\cdot)$ to be of the form

$$u(\frac{D}{\sigma_D}) = \sum_{\{s,r\}\in\mathcal{N}} b_{s-r}\rho(\frac{D_s - D_r}{\sigma_D}) \quad , \tag{25}$$

where $\mathcal{N}$ is the set of all neighboring pixel pairs, and $\rho(\cdot)$ is the potential function that assigns a cost to differences between neighboring pixel values.

A popular choice for $\rho(\cdot)$ in the signal-processing literature has been the quadratic function

$$\rho(\frac{\Delta}{\sigma}) = \frac{1}{2}\left|\frac{\Delta}{\sigma}\right|^2 \quad .$$

This choice is similar to regularization methods that have been used earlier in optical tomography that impose constraints on the derivatives of the solution to enforce smoothness.[7] However, the quadratic cost function tends to excessively penalize large pixel differences resulting in blurred edges. This will be demonstrated in the results section.

Alternatively, a host of edge-preserving potential functions have been suggested in the literature.[16–18,9] In particular, the generalized Gaussian MRF (GGMRF)[9] uses the following potential function

$$\rho(\Delta, \sigma) = \frac{1}{p} \left| \frac{\Delta}{\sigma} \right|^p \quad . \tag{26}$$

Note that when $p = 2$, the potential function is quadratic and the model reduces to a Gaussian MRF (GMRF). The advantage of using the GGMRF as the image model is two fold: first, it has been shown to provide good edge-preservation in the reconstructed image[19] for $p \approx 1$; second, the form of the model facilitates the estimation of the scale parameter directly from the data.[19]

Using Eqs. (24), (25), and (26), the log-likelihood of the image D is given as

$$\log P(D) = -\frac{1}{p\sigma_D^p} \sum_{\{s,r\} \in \mathcal{N}} b_{s-r} \left| D_s - D_r \right|^p \quad . \tag{27}$$

The derivative of the log-likelihood with respect to D is given as

$$\frac{d \log P(D)}{dD_s} = \frac{1}{\sigma^p} \sum_{r \in \mathcal{N}_s} \text{sign}(D_r - D_s) b_{s-r} \left| D_s - D_r \right|^{p-1} \text{ for all } s \in S, \tag{28}$$

where $\mathcal{N}_s$ is the neighborhood of pixel s. The model for the absorption coefficient can be obtained by substituting σ_{μ_a} for σ_D and μ_a for D in Eqs. (27) and (28).

4. RECONSTRUCTION CRITERION

We use the maximum *a posteriori* (MAP) estimation criterion to compute the reconstructions of D and μ_a. The MAP estimate is defined as

$$[\hat{D}, \hat{\mu}_a] = \arg \max_{[D, \mu_a] \geq 0} \log P(D, \mu_a | Y) \quad .$$

Using Bayes rule in the above equation, we have

$$[\hat{D}, \hat{\mu}_a] = \arg \max_{[D, \mu_a] \geq 0} \{\log P(Y | D, \mu_a) + \log P(D) + \log P(\mu_a)\} \quad .$$

Since we can compute the derivative of each of the terms in the cost function (as shown in Sects. 2.4. and 3.), we can use gradient-based methods to do the above optimization. In particular, we use the conjugate gradient algorithm because of its superior convergence properties. However, conventional CGD cannot be used directly since it requires line searching that can violate the positivity constraint. We modify the search procedure so that the estimate is projected back on the convex set $[D\ \mu_a] \geq 0$ when searching for a minimum in a particular direction. This procedure is referred to as bent-line searching.

5. RESULTS

In this section, we study the performance of the proposed algorithm using simulated data. Although the method developed in this paper can be used to estimate D and μ_a simultaneously, we will restrict ourselves to the simple case of just estimating D and assume μ_a is known.

We will use a GGMRF with $p = 1.1$ as the image model since it results in good edge-preservation. We will also show the reconstruction corresponding to $p = 2.0$ (GMRF) to compare the quality of our reconstruction to previously used constraint-based regularization methods that use a quadratic penalty.[7] We use an 8 point neighborhood for the MRF with $b_{s-r} = (2\sqrt{2} + 4)^{-1}$ for nearest neighbors and $b_{s-r} = (4\sqrt{2} + 4)^{-1}$ for diagonal neighbors. Ideally, the scale parameter σ_D needs to be estimated directly from the measurements Y before computing the unknown field D. This unsupervised estimation of σ_D is a challenging problem in itself and is not within the scope of this research. Therefore, for the purpose of this research, we will fix the value of σ_D to the ML estimate[19] obtained from the original D, which is known in our simulation examples.

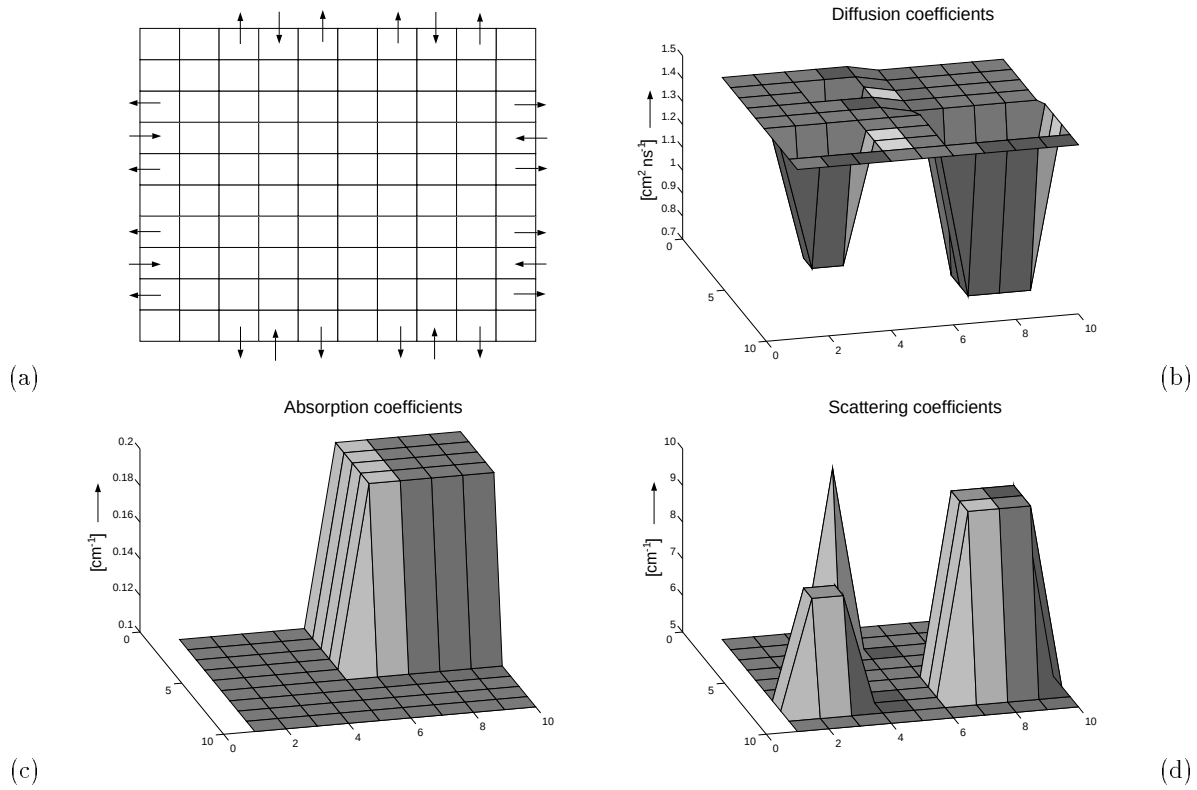
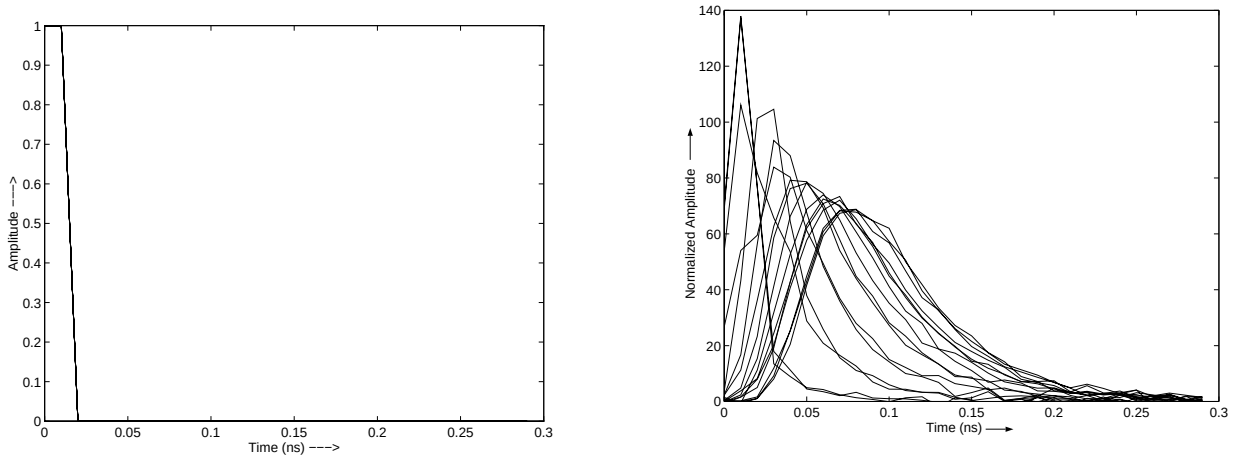


Figure 1. (a) Shows the set-up for a 10x10 grid with 0.1 cm separation between grid points. The arrows pointing inward show the location of the sources and the arrows pointing outward show the detector positions. (b), (c), and (d) show the diffusion, absorption, and scattering coefficients respectively.

The values of the absorption and scattering coefficients used in our examples have been chosen to reflect those of real tissues. The physical dimensions have also been chosen to reflect a real medical imaging set-up. Figure 1 shows the set-up for the first example, which is a 10x10 grid with 0.1 cm spatial separation between grid points. The placement of the sources and detectors are shown in Fig. 1(a). The data are collected by turning on one source at a time and making measurements at all the detector locations for all $n \in \mathcal{T}$. The data collected for each active source constitutes a single view of the object. Let the data collected in view v be denoted as $Y^{(v)}$. Let $\phi^{(v)} \stackrel{\Delta}{=} \log P(Y^{(v)}|\theta)$. Then the total log-likelihood of the data for multiple views is given as $\phi \stackrel{\Delta}{=} \sum_v \phi^{(v)}$, where each $\phi^{(v)}$ is computed using the procedure described in Sect. 2.3.3. Similarly, the total sensitivity of ϕ with respect to θ is given as $\frac{d\phi}{d\theta} = \sum_v \frac{d\phi^{(v)}}{d\theta}$, where each $\frac{d\phi^{(v)}}{d\theta}$ is computed using the procedure in Sect. 2.4.

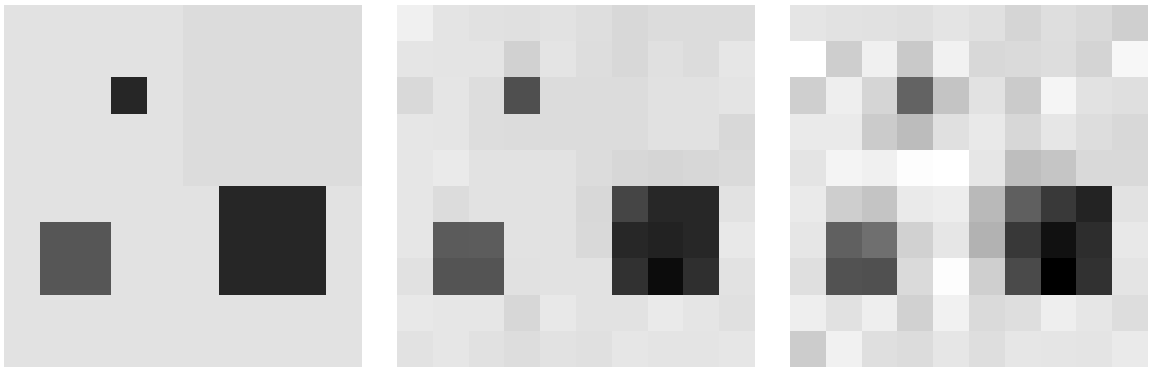
Figures 1(c) and 1(d) show the numerical values for the absorption and scattering coefficients respectively chosen for the first example. Figure 1(b) shows the diffusion coefficients computed from the absorption and scattering coefficients using Eq. (1) and $c = 22$ cm/ns. Figure 2 shows the source distribution, which is a single pulse, and all the detector responses for a single view for the 10x10 grid. Gaussian noise is added to the simulated signals with an rms noise value that is 3% of the rms signal value over the 0.3 ns observation time. This corresponds to a signal-to-noise ratio (SNR) of 30 dB. The simulation is done using $\Delta t = 0.01$ ns for a total time of 0.3 ns and the detector resolution is set to 0.01 ns.

Figure 3 shows the reconstructions of the diffusion coefficients for the 10x10 grid. The conjugate gradient algorithm initialized with a constant diffusion coefficient field of value 3.0 $\text{cm}^2\text{ns}^{-1}$ was run to convergence to obtain the MAP estimate. Figures 3(b) and 3(c) show the reconstructions corresponding to a GGMRF ($p = 1.1$) and a GMRF ($p = 2.0$) respectively. Notice that the GGMRF reconstruction has sharp edges and is a very good reproduction of



(a) (b)

Figure 2. (a) Source distribution used for the 10x10 grid. (b) Normalized measurements obtained at all the detector positions with 30 dB SNR when one of the sources was active. The normalizations was done by the standard deviation of the noise present in each measurement.



(a) (b) (c)

Figure 3. (a) Original diffusion coefficients. Reconstructed diffusion coefficients for the 10x10 grid using the GGMRF image model with (b) $p = 1.1$, and (c) $p = 2.0$ (Quadratic regularization). Note that white in the images corresponds to 1.54 $\text{cm}^2\text{ns}^{-1}$ and black corresponds to 0.58 $\text{cm}^2\text{ns}^{-1}$.

the original image. On the other hand, the GMRF reconstruction has blurred edges and excessive residual noise.

The purpose of the first example was to show how accurately we can reconstruct the original image when there are a large number of views and relatively few unknown coefficients to estimate. The second example we consider uses a much larger grid and only a limited number of views. Figure 4 shows the set-up for the second example, which is a 64x64 grid with 0.1 cm spatial separation between grid points. The placement of the 4 sources and 52 detectors are shown in Fig. 4(a). This case is under-determined because only four views of the object are taken. Figures 4(c) and 4(d) show the numerical values for the absorption and scattering coefficients respectively chosen for this example. Figure 4(b) shows the diffusion coefficients computed from the absorption and scattering coefficients. Figure 5 shows the source distribution, which is a single pulse, and all the detector responses for a single view for the 64x64 grid. The SNR of the detectors is set at 30 dB. The simulation is done using $\Delta t = 0.005$ ns for a total time of 1 ns while the detector resolution is kept at 0.02 ns.

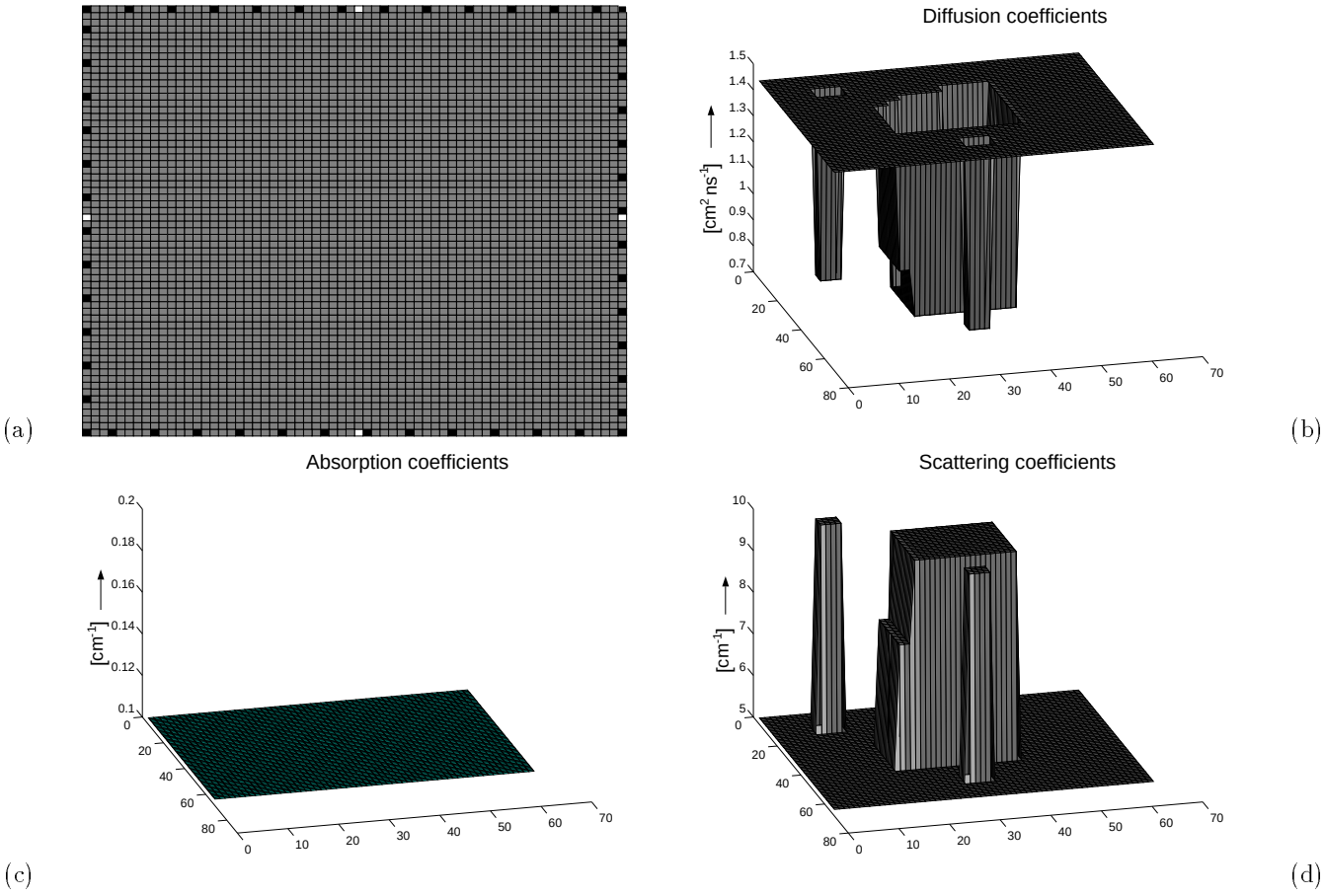
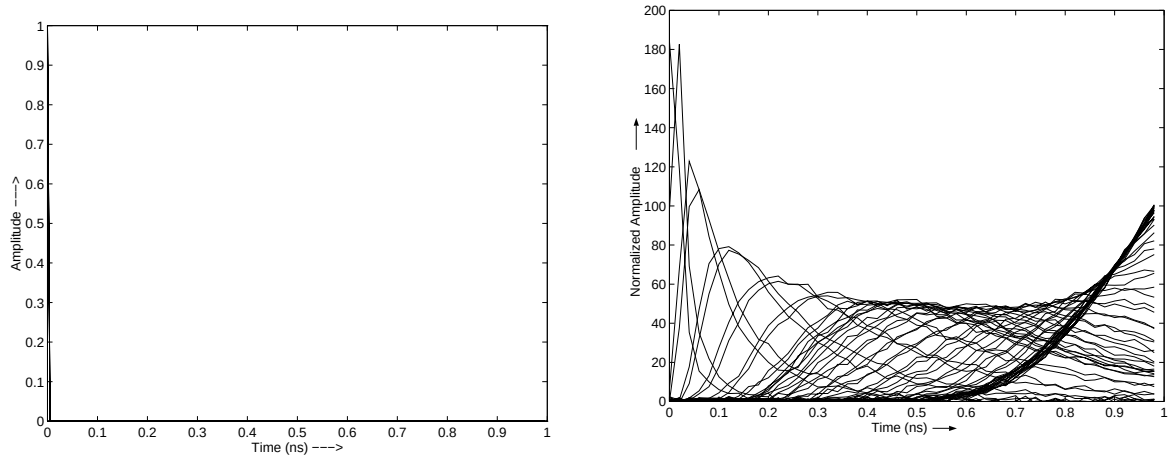


Figure 4. (a) Shows the set-up for a 64x64 grid with 0.1 cm separation between grid points. The black dots indicate measurements positions and the white dots indicate source positions. (b), (c), and (d) show the original diffusion, absorption, and scattering coefficients respectively.

Figure 6 shows the MAP reconstructions of the diffusion coefficients for the 64x64 grid using the conjugate gradient algorithm. The algorithm was initialized with a constant diffusion coefficient field with value 1.0 $\text{cm}^2\text{ns}^{-1}$. Figures 6(b) and 6(c) show the reconstructions corresponding to a GGMRF ($p = 1.1$) and a GMRF ($p = 2.0$) respectively. Notice that although the GGMRF estimate is much better compared to the GMRF estimate, both reconstructions are not very good reproductions of the original image. This can be attributed to the extremely small number of views that were used and the much higher resolution of the reconstruction. An improved estimate can be obtained if we use a lower resolution for the reconstruction as seen from the results of the first example. This observation suggests the use of a multiresolution strategy where the reconstructions are performed at progressively higher resolutions using the coarser resolution reconstruction as the initial condition. Such a strategy is left for future research.

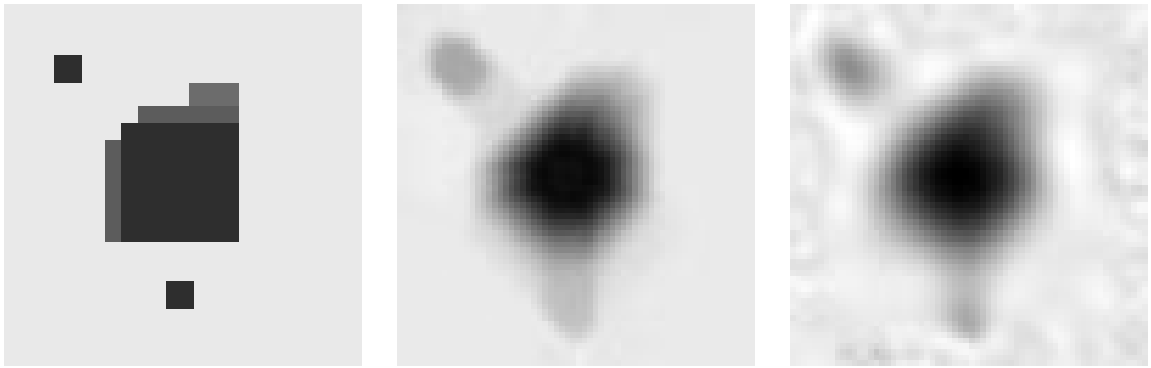
6. CONCLUSION

The contribution of this research has been threefold. First, the use of a new alternating-directions implicit method to solve the forward diffusion problem in O(N) time. Second, the use of adjoint differentiation to compute the sensitivity of the measurements with respect to the unknown parameters. By working backwards in time, and using the discretized equations that are employed to compute the forward solution, we have shown that this can be done in $O(N)$ time as opposed to $O(N^2)$ time required by the perturbation approach. Third, an edge-preserving GGMRF



(a) (b)

Figure 5. (a) Source distribution used for the 64x64 grid. (b) Normalized measurements obtained at all the detector positions with 30 dB SNR when one of the sources was active. The normalizations was done by the standard deviation of the noise present in each measurement.



(a) (b) (c)

Figure 6. (a) Original diffusion coefficients. Reconstructed diffusion coefficients for the 64x64 grid using the GGMRF image model with (b) $p = 1.1$, and (c) $p = 2.0$ (Quadratic regularization). Note that white in the images corresponds to 1.52 $\text{cm}^2\text{ns}^{-1}$ and black corresponds to 0.55 $\text{cm}^2\text{ns}^{-1}$.

model has been employed and shown to be superior to the standard quadratic regularization method, which results in blurred edges and excess residual noise.

Moreover, the proposed method of solving the inverse problem is very general and can be applied to complex problems such as geophysical structure estimation from seismic data and ocean surface reconstruction using sonar. The only requirement for the applicability of the proposed method is that the measurements for the process in question be adequately described by a predictive forward computational model. It can also be used to optimize engineering designs in complex situations such as streamlining of airplane foils and automobile bodies to reduce drag.

Future research should focus on multiresolution strategies for computing the MAP estimate, employing multiscale image models[20] as opposed to fixed-resolution image models to describe the unknown image, and unsupervised methods for optimal estimation of image and data model parameters.

ACKNOWLEDGEMENT

The authors would like to thank Andreas Hielscher for help in formulating the simulation examples in this paper and for useful discussions on this topic.

REFERENCES

1. G. Muller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, eds., *Medical optical tomography: Functional imaging and monitoring*, vol. IS 11, SPIE press, Bellingham, WA, 1993.
2. S. Flock, M. Patterson, B. Wilson, and D. Wyman, "Monte Carlo modeling of light propagation in highly scattering tissues - I: Model predictions and comparison with diffusion theory," *IEEE Trans. Biomed. Eng.* **36**(12), pp. 1162–1168, 1989.
3. Y. Wang, J. Chang, R. Aronson, R. Barbour, H. Graber, and J. Lubowsky, "Imaging of scattering media by diffusion tomography: an iterative perturbative approach," in *Physiological monitoring and early detection diagnostic methods*, *Proc. SPIE* **1641**, pp. 58–71, 1992.
4. R. Barbour, H. Graber, Y. Wang, J. Chang, and R. Aronson, "A perturbation approach for optical diffusion tomography using continuous-wave and time-resolved data," in *Medical optical tomography: Functional imaging and monitoring*, vol. IS 11, pp. 87–120, SPIE Press, 1993.
5. J. Chang, H. Graber, and R. Barbour, "Image reconstruction of targets in random media from continuous wave laser measurements and simulated data," in *Advances in optical imaging and photon migration*, R. R. Alfano, ed., *OSA Proc.* **21**, pp. 193–201, 1994.
6. B. B. Das, J. Dolne, R. L. Barbour, H. L. Graber, J. Chang, M. Zevallos, F. Liu, and R. R. Alfano, "Analysis of time-resolved data for tomographical image reconstruction of opaque phantoms and finite absorbers in diffusive media," *Proc. SPIE* **2389**, pp. 16–28, 1995.
7. S. R. Arridge, "Forward and inverse problems in time-resolved infrared imaging," in *Medical optical tomography: Functional imaging and monitoring*, vol. IS 11, pp. 35–64, SPIE Press, 1993.
8. R. Barbour, H. Graber, R. Aronson, and J. Lubowsky, "Imaging of subsurface regions of random media by remote sensing," *Proc. SPIE* **1431**, pp. 192–203, 1991.
9. C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. on Image Processing* **2**(3), pp. 296–310, 1993.
10. R. Model, R. Hunlich, M. Orlt, and M. Walzel, "Image reconstruction for random media by diffusion tomography," *Proc. SPIE* **2389**, pp. 400–410, 1995.
11. D. S. Burnett, *Finite element analysis: From concepts to applications*, Addison-Wesley, New York, 1987.
12. W. F. Ames, *Numerical methods for partial differential equations*, Academic Press, New York, 1977.
13. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, second ed., 1992.
14. K. M. Hanson and G. S. Cunningham, "A computational approach to Bayesian inference," in *Computing Science and Statistics,* **27**, M. M. Meyer and J. L. Rosenberger, eds., pp. 202–211, Interface Foundation, Fairfax Station, VA 22039-7460, 1996.
15. W. Briggs, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics, Philadelphia, 1987.
16. P. J. Green, "Bayesian reconstruction from emission tomography data using a modified EM algorithm," *IEEE Trans. on Medical Imaging* **9**(1), pp. 84–93, 1990.
17. R. Stevenson and E. Delp, "Fitting curves with discontinuities," *Proc. of the First International Workshop on Robust Computer Vision* , pp. 127–136, (Seattle, WA), 1990.
18. K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Trans. on Medical Imaging* **9**(4), pp. 439–446, 1990.
19. S. S. Saquib, C. A. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields, with applications to Bayesian tomography," Tech. Rep. TR-ECE 95-24, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, 1995.
20. S. S. Saquib, C. A. Bouman, and K. Sauer, "A non-homogeneous MRF model for multiresolution Bayesian estimation," *Proc. of IEEE Int'l Conf. on Image Proc.* **2**, pp. 445–448, (Lausanne, Switzerland), 1996.