

Bayesian Optical Diffusion Imaging

J. C. Ye^a, C. A. Bouman^a, K. J. Webb^a and R. P. Millane^b

^aSchool of Electrical and Computer Engineering, Purdue University, W. Lafayette, IN 47907

^cComputational Science and Engineering Program, Purdue University, W. Lafayette, IN 47907

ABSTRACT

Frequency-domain diffusion imaging is a new imaging modality which uses the magnitude and phase of modulated light propagating through a highly scattering medium to reconstruct an image of the scattering and/or the absorption coefficient in the medium. In this paper, the inversion algorithm is formulated in a Bayesian framework and an efficient optimization technique is presented for calculating the maximum a posteriori image. Numerical results show that the Bayesian framework with the new optimization scheme out-performs conventional approaches in both speed and reconstruction quality.

Keywords: optical diffusion tomography, Bayesian image reconstruction, shot noise statistics, generalized Gaussian Markov random field, iterative coordinate descent algorithm

1. INTRODUCTION

Optical diffusion tomography has generated considerable recent interest, and its potential for imaging in highly scattering media such as tissue, as an alternative to x-ray or ultrasonic tomography, has been demonstrated.^{1,2} In this technique, with a red or near infra-red light source, the detected transmitted light is used to reconstruct the absorption and/or the scattering properties of the medium as a function of position. The low energy optical radiation presents significantly lower health risks than x-ray radiation. Also, suitable sources and detectors are relatively inexpensive, making such an instrument considerably less expensive than computed tomography (CT) and magnetic resonance imaging (MRI) systems.

An accurate model for the propagation of photons through tissue can be obtained from transport theory.³ This model ignores the optical phase, treating the photons as particles. A solution can be obtained by means of Monte-Carlo methods,⁴ which describes individual photon paths, or by means of the diffusion approximation.¹ While the Monte-Carlo method can model the photon path more accurately, the diffusion approximation is sufficiently accurate in highly scattering media such as tissue, and provides a computationally tractable forward model. Therefore, we use the diffusion equation as our forward model.

The inverse problem of reconstructing the absorption and/or scattering coefficients from measurements of scattered light is highly nonlinear because of the nonlinear coupling between the coefficients and the photon flux in the diffusion equation. To facilitate the computation of the unknown coefficients, several approaches attempt to locally linearize the inverse problem. For this class of problems, the Newton-Raphson method has been commonly used with the Levenberg-Marquardt procedure based on a Taylor series expansion.^{5,6,2,7}

However, the Levenberg-Marquardt type algorithm which imposes a penalty on the L_2 norm of the update at each iteration, tends to over-smooth edges in the image or produce excessively noisy-images, depending on the control parameter value. This is because the L_2 penalty term for the new update is not a regularization in the Tikhonov sense,⁸ but a “trust region” constraint for a nonlinear least squares problem.⁹

The artifacts due to poor regularization can be reduced by incorporation of a priori information using a Bayesian framework. In this framework, the maximum *a posteriori* (MAP) reconstruction is often computed by maximizing the posterior distribution. This Bayesian approach has been applied in many image restoration and reconstruction problems.^{10,11} More recently, Bayesian (or other regularization) methods have been applied to nonlinear inverse problems such as microwave imaging, impedance tomography, and optical imaging.^{12,6,13}

The individual approaches have differed both in terms of the prior model (or stabilizing functional) used and the optimization algorithms employed to compute the MAP reconstruction. For example, Paulsen and Jiang¹² added a quadratic regularization term to their previous formulation⁶ to stabilize the reconstruction. Each iteration of the optimization performed a linearization (similar to the Born approximation), followed by a full matrix inversion to solve the linearized problem. Saquib, Hanson and Cunningham proposed a more computationally efficient algorithm for the time domain diffusion problem in which each iteration alternates a linearization step with a single step of a conjugate gradient algorithm.¹³ Arridge and Schweiger adapted this method for the frequency domain diffusion tomography problem.¹⁴ However, their method is computationally expensive because the line search used for each conjugate gradient update requires repeated evaluations of the forward model. Perhaps the research of Carfantan, Mohammad-Djafari and Idier is most closely related to ours, in that they used exact single-site updates for each pixel.¹⁵ They observed that the single-site optimization had rapid convergence in terms of number of iterations. However, each iteration of this method is computationally expensive, requiring $O(N^2)$ complex multiplications. Finally, we note that previous Bayesian approaches have not incorporated the physics of the measurement noise into the Bayesian framework.

In this paper, we formulate the frequency domain optical diffusion inverse problem in a Bayesian framework and derive the maximum a posteriori (MAP) estimate for the reconstruction. Although the methodology we describe can in principle be applied in the general case of unknown absorption and scattering coefficients, for simplicity we focus on the estimation of the absorption coefficient under the assumption that the scattering coefficient is known.

Similar to the previous approach of Saquib, Hanson and Cunningham,¹³ we use the generalized Gaussian Markov random field (GGMRF) as a prior model of the unknown parameters. This results in stable and edge preserving regularization for the optical diffusion imaging problem. In addition, we incorporate a model for the detection statistics derived from the physics of the measurement system. This “shot noise” model provides a natural scaling for the data, which is based on the square root of the time average of the measurements. We believe that our approach is superior because it is based on the accuracy of the actual measurements and is extendible to a wide variety of physical measurement systems.

Another contribution of this work is the introduction of a new optimization technique that we call the iterative coordinate descent Born (ICD/Born) method. Each iteration of the ICD/Born method consists of a linearization step using the Born approximation, followed by a single pass of the ICD algorithm.^{16,17} Since the computational complexity of ICD/Born is $O(N)$, the method requires much less computation per iteration than the exact single pixel update algorithm,¹⁵ the Gaussian elimination technique for total variation minimization,¹² and the conventional distorted Born iterative method (DBIM).¹⁸ Our numerical results indicate that the ICD/Born method together with the Bayesian framework yields accurate and fast reconstructions from synthetic data.

2. THE OPTICAL DIFFUSION TOMOGRAPHY PROBLEM

In a highly scattering medium with low absorption, such as soft tissue in the 650-1300 nm wavelength range, the photon flux density is accurately modeled by the diffusion equation.^{1,19,20} More specifically, let

$\psi_k(r, t)$ be the photon rate per unit area generated at time t and position $r \in \Omega$ due to a modulated point source of light at position $s_k \in \Omega$. Then, $\psi_k(r, t)$ is given by the time domain diffusion equation as

$$\frac{1}{c} \frac{\partial}{\partial t} \psi_k(r, t) - \nabla \cdot D(r) \nabla \psi_k(r, t) + \mu_a(r) \psi_k(r, t) = S(t) \delta(r - s_k), \quad (1)$$

where c is the speed of light in the medium, $S(t)$ is the time varying photon source density, and $D(r)$ is the diffusion constant given by $D(r) = 1/3(\mu_a(r) + \mu'_s(r))$, where $\mu_a(r)$ is the absorption coefficient, and $\mu'_s(r)$ is the reduced scattering coefficient.

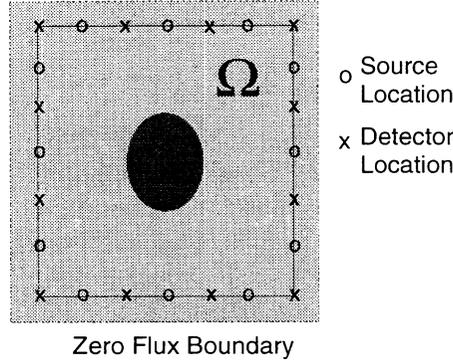


Figure 1. Simulation geometry with the locations of sources and detectors for inversion of synthetic data. The sources and detectors are uniformly spaced along the edges.

Practical systems based on time domain measurements have been implemented,^{21,1} but these systems tend to be expensive and noise sensitive. In order to circumvent these problems, we adopt a frequency domain approach to the optical diffusion problem.^{22,6} To do this, we assume that the light source is amplitude modulated at a fixed angular frequency $\omega (\neq 0)$, so that $S(t) = Re[1 + \beta \exp(-j\omega t)]$, where β is the modulation depth. At the detector, the complex modulation envelope is then measured by demodulating the in-phase and quadrature components of the measured sinusoidal signal $\psi_k(r, t)$. This technique allows low noise narrow-band heterodyne detection.²³ By taking the Fourier transform of (1), the partial differential equation that governs the complex modulation envelope $\phi_k(r, \omega)$ becomes

$$\nabla \cdot D(r) \nabla \phi_k(r, \omega) + (-\mu_a(r) + j\omega/c) \phi_k(r, \omega) = -\delta(r - s_k). \quad (2)$$

The region to be imaged is denoted by Ω and is surrounded by K point sources uniformly distributed around the 2-D boundary at positions $s_k \in \Omega$, and M detectors interspersed between the sources at positions $d_m \in \Omega$ (Figure 1). The reduced scattering coefficient $\mu'_s(r)$ is assumed to be known for all points $r \in \Omega$, but the absorption coefficient $\mu_a(r)$ in Ω is unknown. The domain Ω is discretized into N pixels, where the position of the i -th pixel is denoted by r_i for $1 \leq i \leq N$. The set of unknown absorption coefficients is then denoted by the vector \mathbf{x} where $\mathbf{x} = [\mu_a(r_1), \dots, \mu_a(r_N)]^T$. Using this notation, we may express the forward model as a vector valued function $\mathbf{f}(\mathbf{x})$. The function $\mathbf{f}(\mathbf{x})$ then takes on the values given by

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_P(\mathbf{x})]^T \\ &= [\phi_1(d_1, \omega), \phi_1(d_2, \omega), \dots, \phi_1(d_M, \omega), \phi_2(d_1, \omega), \dots, \phi_K(d_M, \omega)]^T. \end{aligned} \quad (3)$$

The measurements of the complex envelope $\phi_k(d_m, \omega)$ for source k and detector m are denoted by y_{km} . We also organize the measurements as a single column vector of length $P = KM$,

$$\mathbf{y} = [y_{11}, y_{12}, \dots, y_{KM}]^T. \quad (4)$$

Then, the MAP estimate, $\hat{\mathbf{x}}_{MAP}$, is given by

$$\hat{\mathbf{x}}_{MAP} = \arg \max_{\mathbf{x}} \{ \log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x}) \}, \quad (5)$$

where $p(\mathbf{x}|\mathbf{y})$ is the posterior density of \mathbf{x} given \mathbf{y} , $p(\mathbf{y}|\mathbf{x})$ is the conditional probability density of \mathbf{y} given \mathbf{x} , and $p(\mathbf{x})$ is the prior density for the image.

3. MEASUREMENT AND PRIOR MODEL

In this section, we derive an expression for the distribution $p(\mathbf{y}|\mathbf{x})$ in terms of the photon flux density $\phi_k(d_m, \omega)$. The details of the model are derived in a recent publication²⁴ and are based on a shot noise model for the detected signal. In this model, the measurements are normally distributed with a mean equal to the exact (noiseless) measurement and a variance proportional to the exact measurement at a modulation frequency of zero (DC). The density function for a single datum is given by²⁴

$$p(y_{km}|\mathbf{x}) = \frac{1}{2\pi\alpha|\phi_k(d_m, \omega)|} \exp \left[-\frac{|y_{km} - \phi_k(d_m, \omega)|^2}{2\alpha|\phi_k(d_m, \omega)|} \right], \quad (6)$$

where α is a constant determined by the modulation depth and the physical characteristics of the detector. We assume that the noise signals are independent for all source-detector pairs, so that the covariance matrix C for the data vector \mathbf{y} is diagonal and given by

$$C_{ii} = \alpha|\phi_k(d_m, \omega)| \simeq \alpha|y_{km}|, \quad \text{where } i = M(k-1) + m. \quad (7)$$

To simplify notation, we define the diagonal matrix Λ by

$$\Lambda = \frac{1}{2}C^{-1}. \quad (8)$$

The data likelihood is then given by

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^N |\Lambda|^{-1}} \exp \left[-\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_{\Lambda}^2 \right] \quad (9)$$

where $\|\mathbf{z}\|_{\Lambda}^2 = \mathbf{z}^H \Lambda \mathbf{z}$, and H denotes the Hermitian transpose.

In many image reconstruction problems, the Markov random field (MRF) model has proved useful in describing spatial correlations between neighboring pixels. MRFs have the property that the conditional distribution of a pixel given all other pixels is only a function of the pixel's neighbors. A wide variety of MRF models have been used.^{25,26,10,11} However, we use here the generalized Gaussian MRF (GGMRF) model because it is both convex and scale invariant.²⁶ The convexity of the potential function of the GGMRF leads to continuous or stable MAP estimates, and the scale invariant property of the GGMRF potential functions eliminates the necessity of choosing an edge threshold which is often required for non-Gaussian potential functions.²⁶ For the GGMRF model, the density function for \mathbf{x} is given by

$$p(\mathbf{x}) = \frac{1}{\sigma^N z(p)} \exp \left[-\frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p \right], \quad 1 \leq p \leq 2. \quad (10)$$

Furthermore, since the absorption must be positive we also impose the constraint $x_i \geq 0, i = 1, \dots, N$.

4. ICD/BORN OPTIMIZATION TECHNIQUE

Referring to (5), (9), (10), and the positivity constraint, the MAP estimate for \mathbf{x} is given by

$$\hat{\mathbf{x}}_{MAP} = \arg \min_{\mathbf{x} \geq 0} \left[\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_{\Lambda}^2 + \frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p \right]. \quad (11)$$

To compute the MAP reconstruction, we must perform the optimization (11). We choose to use the ICD algorithm¹⁷ for a number of reasons. First, it has been shown that ICD updates work well with non-Gaussian prior models.¹⁷ Second, the ICD algorithm is easily implemented with a positivity constraint. On the other hand, a drawback of the conjugate gradient method is the difficulty of incorporating positivity constraints.¹⁷

The ICD algorithm is implemented by sequentially updating each pixel of the image. After every pixel has been updated, the procedure is repeated starting from the first pixel again. We refer to a single update of every pixel in the image as a “scan”. The ICD algorithm therefore consists of a number of scans until some convergence criterion is satisfied. Each scan consists of N pixel updates. Each pixel update is chosen to minimize the MAP cost function, so that the update \hat{x}_i of the absorption value of the i -th pixel is given by

$$\hat{x}_i = \arg \min_{\tilde{x}_i \geq 0} \left[\|\mathbf{y} - \mathbf{f}(\tilde{\mathbf{x}}_i)\|_{\Lambda}^2 + \frac{1}{p\sigma^p} \sum_{j \in \mathcal{N}_i} b_{i-j} |\tilde{x}_i - x_j|^p \right], \quad (12)$$

where $\tilde{\mathbf{x}}_i = [x_1, x_2, \dots, x_{i-1}, \tilde{x}_i, x_{i+1}, \dots, x_N]^T$ and \mathcal{N}_i is the set of pixels neighboring pixel i . Note that x_i is replaced by \tilde{x}_i before the next pixel update. However, a direct approach¹⁵ to the update equation of (12) is very computationally expensive due to the highly nonlinear nature of the forward model $\mathbf{f}(\mathbf{x})$. Furthermore, each evaluation of the function $\mathbf{f}(\mathbf{x})$ requires the solution of the full partial differential equation (2) for each source.

The computational inefficiency is overcome by using the Born approximation at the beginning of each scan, and we call this approach the ICD/Born algorithm. We use the integer n to index the scans of the algorithm, and \mathbf{x}^n denotes the image after the n -th scan. At the beginning of the $(n+1)$ -th scan, the approximation

$$\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_{\Lambda}^2 \simeq \|\mathbf{y} - \mathbf{f}(\mathbf{x}^n) - \mathbf{f}'(\mathbf{x}^n)\Delta\mathbf{x}\|_{\Lambda}^2 \quad (13)$$

is used, where $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^n$ and $\mathbf{f}'(\mathbf{x})$ represents the Fréchet derivative of $\mathbf{f}(\mathbf{x})$, which for the discretized problem is the $P \times N$ complex matrix

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_P(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_P(\mathbf{x})}{\partial x_N} \end{bmatrix} \quad (14)$$

and $\mathbf{f}'(\mathbf{x}^n)$ denotes the Fréchet derivative computed for the absorption parameter estimate \mathbf{x}^n , and $g(d_m, r, \omega)$ is the Green's function for the frequency domain diffusion equation (2). For the discretized domain, $\Delta x_i = \mu_a(r_i) - \mu_a^n(r_i)$ and the elements of the matrix $\mathbf{f}'(\mathbf{x})$ in (14) are approximated as⁷

$$\frac{\partial \mathbf{f}_i(d_m, \omega)}{\partial x_i} = g(d_m, r_i, \omega) \phi_k(r_i, \omega) \left\{ -1 + \frac{\mu_a^n(r_i) + j\omega/c}{\mu_a^n(r_i) + \mu_s(r_i)} \right\} A. \quad (15)$$

where A is the pixel area, $l = M(k - 1) + m$, and $g(d_m, r_i, \omega)$ is a Green's function for frequency domain diffusion equation (2). After the n -th scan, $\mathbf{f}(\mathbf{x}^n)$ and $\mathbf{f}'(\mathbf{x}^n)$ are calculated by computing $g(d_m, r_i, \omega)$ and $\phi_k(r_i, \omega)$ of (15) using a linear PDE solver for the diffusion equation (2) with the n -th estimate of absorption coefficient, \mathbf{x}^n .

In an ICD scan, each pixel is updated in turn, and the new value \hat{x}_i is given by

$$\hat{x}_i = \arg \min_{\tilde{x}_i \geq 0} \left[\|\mathbf{y} - \mathbf{f}(\mathbf{x}^n) - [\mathbf{f}'(\mathbf{x}^n)]_{*i}(\tilde{x}_i - x_i^n)\|_{\Lambda}^2 + \frac{1}{p\sigma^p} \sum_{j \in \mathcal{N}_i} b_{i-j} |\tilde{x}_i - x_j|^p \right], \quad (16)$$

where $[\mathbf{f}'(\mathbf{x}^n)]_{*i}$ is the i -th column of the Fréchet derivative matrix. To compute the solution to (16), we express the first term as a quadratic function of \tilde{x}_i to obtain the expression

$$\hat{x}_i = \arg \min_{\tilde{x}_i \geq 0} \left[\theta_1(\tilde{x}_i - x_i^n) + \frac{\theta_2}{2}(\tilde{x}_i - x_i^n)^2 + \frac{1}{p\sigma^p} \sum_{j \in \mathcal{N}_i} b_{i-j} |\tilde{x}_i - x_j|^p \right], \quad (17)$$

where θ_1 and θ_2 are given by

$$\theta_1 = -2Re \left[[\mathbf{f}'(\mathbf{x}^n)_{*i}]^H \Lambda \mathbf{e}(i) \right], \quad \theta_2 = 2[\mathbf{f}'(\mathbf{x}^n)_{*i}]^H \Lambda \mathbf{f}'(\mathbf{x}^n)_{*i}, \quad (18)$$

and the error vector $\mathbf{e}(\cdot)$ for the first pixel is

$$\mathbf{e}(1) = \mathbf{y} - \mathbf{f}(\mathbf{x}^n), \quad (19)$$

and for subsequent pixels is updated as

$$\mathbf{e}(i + 1) = \mathbf{e}(i) - [\mathbf{f}'(\mathbf{x}^n)]_{*i}(\hat{x}_i - x_i^n). \quad (20)$$

Solution of (17) requires minimization of a one-dimensional function. We achieve this by solving for the root of the derivative of the expression in the square brackets in (17), i.e.,

$$\theta_1 + \theta_2(\tilde{x}_i - x_i^n) + \frac{1}{\sigma^p} \sum_{j \in \mathcal{N}_i} b_{i-j} |\tilde{x}_i - x_j|^{p-1} \text{sign}(\tilde{x}_i - x_j) = 0. \quad (21)$$

This root-finding procedure is done using a half-interval search because the function in (21) is monotone decreasing.¹⁷

5. COMPUTATIONAL COMPLEXITY

In this section, we compare the computational cost of our ICD/Born method with the conventional DBIM. This is done by counting the number of complex multiplications (referred to here as cflops) required for one update of the whole image. Recall that one complete iteration or scan of the ICD/Born method implies a single update of each pixel in the image formed by the unknown absorption coefficients. Table 1 summarizes the computational complexity analysis, and Table 2 gives comparisons in the number of complex multiplications for two typical cases

We first analyze the ICD/Born algorithm. The computational cost for evaluating an element of the Fréchet derivative $\mathbf{f}'(\mathbf{x}^n)$ in (14) consists of calculating the Green's function of (2), the flux $\phi_k(r_i, \omega)$, and evaluating (15). Evaluation of $\phi_k(r_i, \omega)$ requires K forward solutions of (2), one solution for each source location. The evaluation of the Green's function in (15) implicitly involves placing a source at each grid

	DBIM (cflops)	ICD/Born (cflops)
Green's function and ϕ_k update	$5(M + K)LN$	$5(M + K)LN$
Fréchet derivative	$(2MK + 2)N$	$(2MK + 2)N$
Pixel update	$2(MK)^2N - (MK)^3/3$	$5MKN$
Total order of computation	$2(MK)^2N + 5(M + K)LN$	$7MKN + 5(M + K)LN$

Table 1. Computational complexity of the DBIM and the ICD/Born method in terms of the number of complex multiplications per iteration.

Parameters	DBIM (cflops)	ICD/Born (cflops)	Speed up
$L_{DBIM} = 10, L_{ICD} = 10, K = 12, M = 12, N = 33^2$	45,790,290	2,406,690	19 to 1
$L_{DBIM} = 30, L_{ICD} = 20, K = 12, M = 12, N = 33^2$	48,403,890	3,713,490	13 to 1
$L_{DBIM} = 30, L_{ICD} = 30, K = 54, M = 54, N = 33^3$	6.034×10^{11}	1.316×10^9	459 to 1

Table 2. Comparison of the computation required (complex multiplications) for one iteration of the DBIM and the ICD/Born method.

point r_i and computing the flux at each detector point d_m , which requires N forward solutions. This computational cost can be dramatically reduced by using reciprocity²⁷ of $\phi_k(r_i, \omega)$ in (2) so that

$$g(d_m, r_i, \omega) = g(r_i, d_m, \omega). \quad (22)$$

Hence, we can place a source at each detector point, thereby requiring only $M \ll N$ forward solutions. Furthermore, to compute the Green's function and flux for the new parameter \mathbf{x}^{n+1} , we can use the corresponding values from the previous iteration as initial guesses in an iterative solver. Since the ICD/Born algorithm typically produces a small change ($\mathbf{x}^{n+1} - \mathbf{x}^n$), this approach gives rapid convergence by reducing the number of iterations required for the forward solver. Using a standard five point Laplacian discrete approximation, each forward iteration of (2) by an iterative linear solver (e.g. SOR²⁸) requires $5N$ complex multiplications. Therefore, the total multiplications required for the Green's function and flux update is $5(K + M)LN$, where L is the number of iterations required for the linear solver. In addition, $(2MK + 2)N$ multiplications are required to fill the Fréchet derivative matrix in (14).

The solution of (21) is usually computationally inexpensive since the neighborhood \mathcal{N}_i typically contains only a few pixels. Therefore, the computation is dominated by the calculation of θ_1 and θ_2 in (18). Since the number of columns of $\mathbf{f}'(\mathbf{x}^n)$ is MK and Λ is diagonal, the number of multiplications required to compute an individual value of θ_1 is $2MK$. Similarly, the update of θ_2 is also $2MK$. In addition, the update of the projected error vector $\mathbf{e}(i + 1)$ requires MK multiplications. Therefore, the total number of multiplications required for the pixel update is $5MKN$. This results in a total computational cost for the ICD/Born method of $5(M + K)LN + (7MK + 2)N$ complex multiplications per iteration.

To compare this result with the computational cost for DBIM, let us first briefly explain the DBIM algorithm. In the DBIM, a new parameter estimate is computed from the perturbation equation using a trust region constraint, where the Fréchet derivative $\mathbf{f}'(\mathbf{x}^n)$ is again defined as in (14) and (15). Each iteration of the DBIM also requires the computation of the Fréchet derivative and a regularized inverse. If we use the same linear solver for the computation of the Green's function and flux, and if we also use the reciprocity relation (22), the number of multiplications required for the Fréchet derivative update is again $5(M + K)LN$, where the iteration number L must typically be chosen to be larger due to the greater change

in \mathbf{x} for each DBIM iteration. To fill in the Fréchet derivative (14) a total of $(2MK + 2)N$ multiplications is again required. The computational cost for the regularized inverse using QR-decomposition is at least $2(MK)^2N - (MK)^3/3$.²⁸ Therefore, the total number of multiplications for a complete update of the DBIM is $5(M + K)LN + 2((MK)^2 + MK + 1)N - (MK)^3/3$.

Table 1 summarizes the computational complexity results. The bottom row of the table lists the dominant (i.e. highest order) terms for each method. Notice that the essential difference is that the DBIM contains a $(MK)^2N$ term, whereas ICD only contains a MKN term. This difference is particularly significant as the number of sources and detectors grows.

The computational advantage of ICD/Born algorithm over the DBIM becomes clear when we use actual numbers. Table 2 shows the number of complex multiplications required when we use $M = 12$ detectors, $K = 12$ sources, and $N = 1089$ pixels (e.g. a 33×33 discretization domain). As discussed earlier, the number of the forward linear solver iterations for the two algorithms will vary. We used the MUDPACK (multigrid software for elliptic partial differential equations) libraries²⁹ as our forward solver, and controlled the number of iterations by setting the relative error tolerance in MUDPACK.²⁹ Therefore, the number of forward solver iterations varies dynamically with respect to \mathbf{x} . For the results shown in Table 2, we picked typical iteration numbers L_{DBIM} and L_{ICD} that we believe to be reasonable. These choices are justified later by Table 3 which shows the actual CPU time per iteration.

Two cases are considered, one which uses the same number of forward iterations for both methods, and one which uses different numbers of forward iterations. The results listed in Table 2 show that for these two cases the computation using the ICD/Born method is reduced by a factor of 13 and 19 as compared to the DBIM. The bottom row of the table lists the estimated complex multiplications for a three-dimensional reconstruction scenario with similar geometry. We can see that for this case the computational speedup of the ICD/Born method is dramatic.

	DBIM	ICD/Born ($p = 2.0$)	ICD/Born ($p = 1.1$)
CPU time per iteration	4 sec	0.228 sec	0.318 sec

Table 3. Average CPU time per iteration for the first simulation (Figure 2).

6. NUMERICAL RESULTS

Simulation results are presented here to assess the performance of the new algorithm. The entire region, Ω (including the homogeneous background) is considered unknown. A total of 12 sources and 12 detectors are located uniformly over the boundary of a $8cm \times 8cm$ domain, as shown in Figure 1.

To illustrate the effect of noise on the stability of the algorithm, random noise with a complex Gaussian distribution is added to the measured flux data. For the simulations, the image is discretized on a 33×33 grid. In all the simulations, the scattering coefficient is set to a constant value $\mu_s = 10.0cm^{-1}$. The simulation was performed on a Sun Ultra Sparc 30.

Figure 2(a) shows the original μ_a image used for the first numerical experiment. Figure 2(b) shows the DBIM reconstruction result.⁷ The reconstruction shows a high noise level and incorrect peak heights. This is due to the L_2 norm used in each update as a trust region constraint, which does not have a noise smoothing effect. For the GGMRF prior model, we used an eight-point neighborhood system, with $b_{i-j} = (2\sqrt{2} + 4)^{-1}$ for nearest neighbors and $b_{i-j} = (4\sqrt{2} + 4)^{-1}$ for diagonal neighbors. The MAP reconstructions were computed by running the ICD/Born algorithm with a positivity constraint. The

stopping criterion used was a fixed CPU time of 60 seconds (in order to allow comparison of the two algorithms). Figure 2(c) shows the MAP reconstruction for a Gaussian prior ($p = 2, \sigma = 1.0 \times 10^{-3}$), while Figure 2(d) gives the reconstruction for $p = 1.1$ and $\sigma = 2.3 \times 10^{-4}$. The reconstruction using the Gaussian prior ($p = 2$) reduces the background noise compared to the reconstruction by the DBIM, but suffers from some smoothing of the edges as the price for noise suppression. The boundaries can be sharpened by using a larger value of σ , but at the expense of larger noise artifacts. Sharper edges and good noise suppression are obtained for $p = 1.1$ (Figure 2(d)). Here, due to the edge preserving nature of the GGMRF prior, the edges are more noticeable in the reconstruction while suppressing the noise.

Figure 3(a) shows a plot of the normalized root mean square error (NRMSE) in μ_a versus CPU time for this simulation, which illustrates that the DBIM eventually diverges. This is again due to the insufficient regularization of the DBIM.¹⁸ Figure 3(b) shows a plot of cost function in (11) versus iteration for the ICD/Born method with $p = 2.0$ and $p = 1.1$ for the above simulation. Note that the ICD/Born is very stable and has reasonably fast convergence. The convergence behavior of ICD/Born is also consistent with a previously described single site update algorithm.¹⁵

Figure 4 shows a variety of more complicated absorption images. Reconstructions using the ICD/Born method for these images are shown in Figure 5. The parameters used for the reconstructions and the final NRMSE are given in the caption to Figure 5. In each case, the reconstructions are accurate both quantitatively and qualitatively, demonstrating that our new algorithm performs well on more complex images.

7. CONCLUSION

We have formulated the optical diffusion inverse problem in a Bayesian framework and implemented a maximum *a posteriori* (MAP) reconstruction of the absorption coefficient. The Bayesian framework enables incorporation of prior knowledge of the unknown parameters as well as detection statistics. As an optimization technique for the Bayesian framework, we developed a new method which combines the iterative coordinate descent (ICD) and the Born approximation which we call the ICD/Born method. This method significantly reduces the computational complexity as compared to methods such as the conventional distorted Born iterative method (DBIM). Numerical simulations show that the Bayesian framework together with the ICD/Born method significantly improves the quality of reconstructions.

REFERENCES

1. M. S. Patterson, B. Chance, and B. Wilson, "Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties," *Appl. Opt.* **28**, pp. 2331–2336, June 1989.
2. Y. Yao, Y. Wang, Y. Pei, W. Zhu, and R. L. Barbour, "Frequency domain optical imaging of absorption and scattering distributions by a Born iterative method," *J. Opt. Soc. Am. A* **14**, pp. 325–342, January 1997.
3. J. J. Duderstadt and L. J. Hamilton, *Nuclear Reactor Analysis*, Wiley, New York, 1976.
4. 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.
5. S. R. Arridge, M. Schweiger, M. Hiraoka, and D. T. Delpy, "Performance of an iterative reconstruction algorithm for near infrared absorption and scattering imaging," in *Proc. of SPIE Conf. on Photon Migration and Imaging in Random Media and Tissues*, G. Muller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, eds., vol. 1888, pp. 360–371, (Los Angeles, CA.), May 1993.
6. H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency-domain data: simulation and experiment," *J. Opt. Soc. Am. A* **13**, pp. 253–266, February 1996.

7. J. C. Ye, K. J. Webb, R. P. Millane, and T. J. Downar, "Modified distorted Born iterative method algorithm with an approximate Fréchet derivative for optical diffusion tomography," *J. Opt. Soc. Am. A* **16**, July 1999.
8. A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, Winston and Sons, New York, 1977.
9. R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons Ltd., England, 2 ed., 1987.
10. T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. on Medical Imaging* **8**, pp. 194–202, June 1989.
11. D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **14**(3), pp. 367–383, 1992.
12. K. D. Paulsen and H. Jiang, "Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization," *Appl. Opt* **35**, pp. 3447–3458, July 1996.
13. S. S. Saquib, K. M. Hanson, and G. S. Cunningham, "Model-based image reconstruction from time-resolved diffusion data," in *Proc. of SPIE Conf. on Medical Imaging: Image Processing*, vol. 3034, pp. 369–380, (Newport Beach, CA), February 25-28 1997.
14. S. R. Arridge and M. Schweiger, "A gradient-based optimisation scheme for optical tomography," *Optics Express* **2**, pp. 213–226, March 1998.
15. H. Carfantan, A. Mohammad-Djafari, and J. Idier, "A single site update algorithm for nonlinear diffraction tomography," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, vol. 4, pp. 2837–2840, (Munich, Germany), April 21-24 1997.
16. K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Trans. on Signal Processing* **41**, February 1993.
17. C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Processing* **5**, pp. 480–492, March 1996.
18. W. C. Chew, *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, New York, 1990.
19. M. S. Patterson, J. D. Moulton, B. C. Wilson, K. W. Berndtand, and J. R. Lakowicz, "Frequency-domain reflectance for the determination of the scattering and absorption properties of tissue," *Appl. Opt* **30**(31), pp. 4474–4476, 1991.
20. M. O'Leary, D. Boas, B. Chance, and A. Yodh, "Refraction of diffuse photon density waves," *Physical Review Letters* **69**(18), pp. 2658–2661, 1992.
21. S. Jacques, "Time resolved propagation of ultrashort laser pulsed within turbid tissues," *Appl. Opt* **28**, pp. 2223–2229, June 1989.
22. J. S. Reynolds, S. Yeung, A. Przada, and K. J. Webb, "Optical diffusion imaging: a comparative numerical and experimental study," *Appl. Opt* **35**, pp. 3671–3679, July 1996.
23. R. H. Kingston, *Detection of Optical and Infrared Radiation*, Springer-Verlag, New York, 1978.
24. J. C. Ye, K. J. Webb, C. A. Bouman, and R. P. Millane, "Optical diffusion tomography using iterative coordinate descent optimization in a Bayesian framework," *Accepted for publication, J. Opt. Soc. Am. A* , 1999.
25. J. Besag, "Towards Bayesian image analysis," *Journal of Applied Statistics* **16**(3), pp. 395–407, 1989.
26. C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving map estimation," *IEEE Trans. on Image Processing* **2**, pp. 296–310, July 1993.
27. S. R. Arridge, "Photon-measurement density functions. Part 1: Analytical forms," *Appl. Opt* **34**, pp. 7395–7409, November 1995.
28. G. H. Golub and C. F. V. Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore, 2 ed., 1989.
29. J. C. Adams, "MUDPACK: Multigrid portable FORTRAN software for the efficient solution of linear elliptic partial differential equations," *Appl. Math. Comput.* **34**, pp. 113–146, 1989.

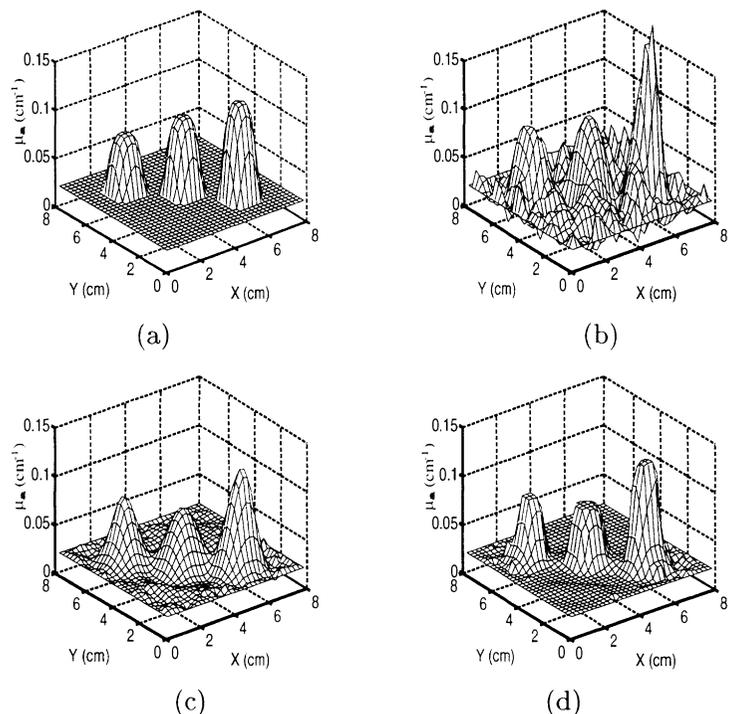


Figure 2. Reconstruction results for μ_a : (a) original absorption image; (b) reconstruction by the DBIM; (c) reconstruction by the new algorithm with a Gaussian prior ($p = 2.0, \sigma = 1.00 \times 10^{-3}$); (d) reconstruction by the new algorithm with a GGMRF prior ($p = 1.1, \sigma = 2.31 \times 10^{-4}$).

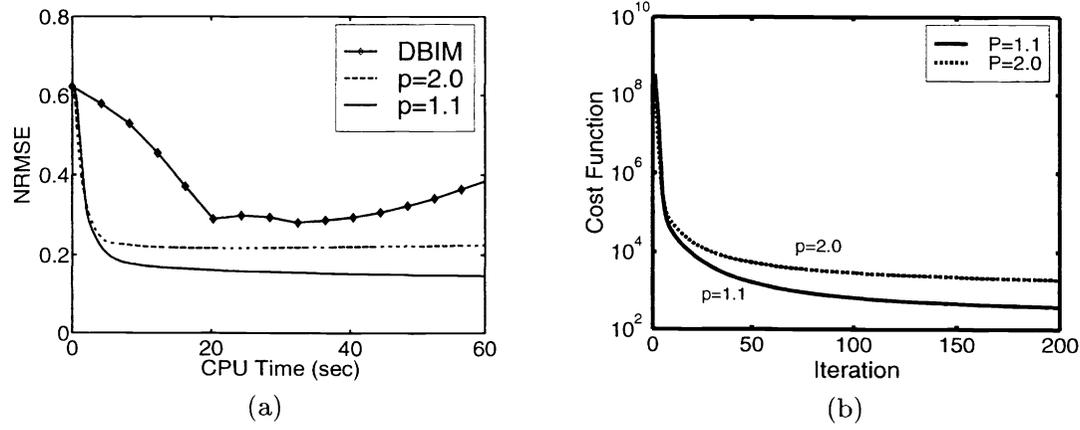


Figure 3. (a) Normalized root mean square error (NRMSE) versus CPU time for the ICD/Born method (with $p = 1.1$ and $p = 2$) and the DBIM. (b) Cost function (log scale) versus iteration for the ICD/Born algorithm with $p = 1.1$ and $p = 2$.

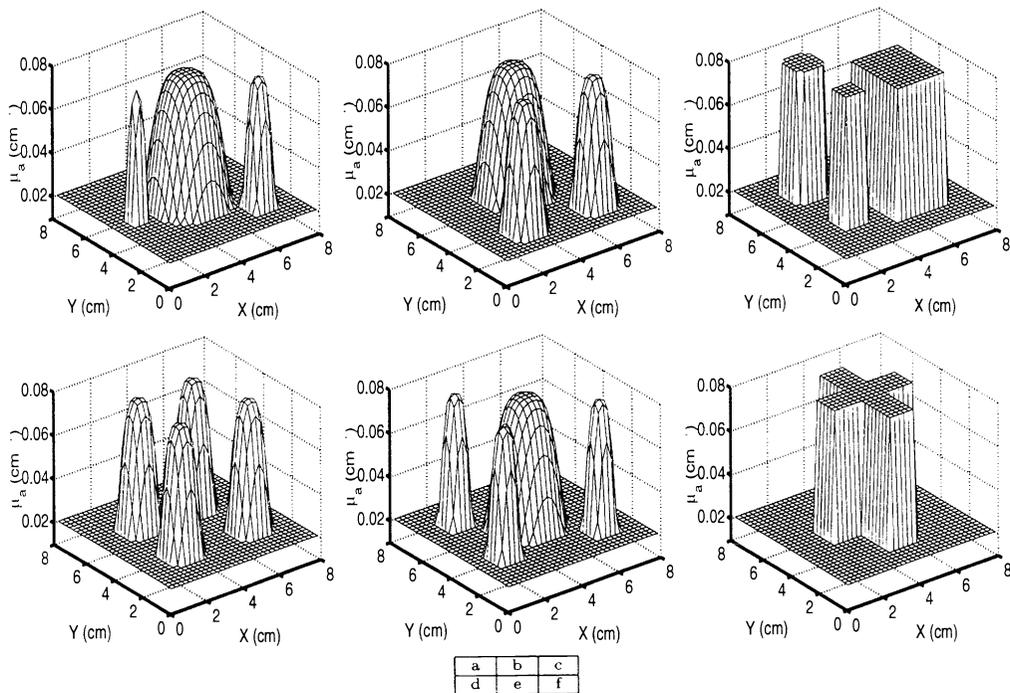


Figure 4. A variety of true absorption images used for simulations.

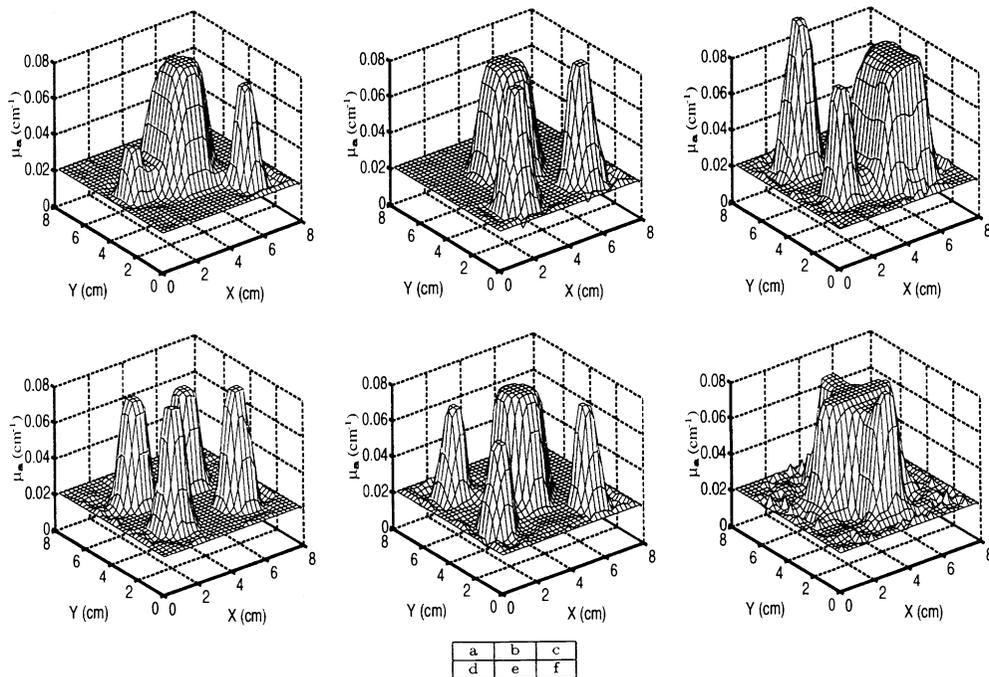


Figure 5. Reconstructions of the absorption images shown in Figure 4 using the ICD/Born algorithm with the GGMRF prior with $p = 1.1$ and $\sigma = 2.31 \times 10^{-4}$. NRMSE values for the final reconstructions are: (a) 1.25×10^{-1} , (b) 7.70×10^{-2} , (c) 2.18×10^{-1} , (d) 8.34×10^{-2} , (e) 1.26×10^{-1} , (f) 2.08×10^{-1} .