Optical diffusion tomography by iterative-coordinate-descent optimization in a Bayesian framework

Jong Chul Ye,* Kevin J. Webb, and Charles A. Bouman

School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907

R. P. Millane

Computational Science and Engineering Program, Purdue University, West Lafayette, Indiana 47907

Received January 4, 1999; accepted March 31, 1999

Frequency-domain diffusion imaging uses the magnitude and phase of modulated light propagating through a highly scattering medium to reconstruct an image of the spatially dependent scattering or absorption coefficients in the medium. An inversion algorithm is formulated in a Bayesian framework and an efficient optimization technique is presented for calculating the maximum *a posteriori* image. In this framework the data are modeled as a complex Gaussian random vector with shot-noise statistics, and the unknown image is modeled as a generalized Gaussian Markov random field. The shot-noise statistics provide correct weighting for the measurement, and the generalized Gaussian Markov random field prior enhances the reconstruction quality and retains edges in the reconstruction. A localized relaxation algorithm, the iterative-coordinate-descent algorithm, is employed as a computationally efficient optimization technique. Numerical results for two-dimensional images show that the Bayesian framework with the new optimization scheme outperforms conventional approaches in both speed and reconstruction quality. © 1999 Optical Society of America [S0740-3232(99)01410-6]

OCIS codes: 100.3010, 100.3190, 100.6950, 170.3010, 290.3200.

1. INTRODUCTION

Optical diffusion tomography has generated considerable recent interest, and its potential for imaging, as an alternative to x-ray or ultrasonic tomography, in highly scattering media such as tissue has been demonstrated.[1–3] In this technique, with a red or near-infrared light source, the detected transmitted light is used to reconstruct the absorption 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. Furthermore, in an optical imaging application a host of spectroscopic techniques can be applied. Given these desirable features, optical imaging has become a candidate for the screening of soft-tissue tumors.

Two common approaches for optical diffusion tomography are a frequency-domain method with an amplitude-modulated optical source, in which a coherent measurement is performed at the modulation frequency, and a time-domain method with short optical pulses, in which temporal gating methods are employed. We focus in this paper on the frequency-domain method.

An accurate model for the propagation of photons through tissue can be obtained from transport theory.[4] This model ignores the optical phase, treating the photons as particles. A solution can be obtained by means of Monte Carlo methods,[5] which would describe individual photon paths, or by means of the diffusion approximation.[1] Although 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 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 have attempted 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. A Levenberg–Marquardt method based on a variational formulation of the time-domain diffusion equation has been applied to time-domain data.[6] In this technique the moments of the photon current at the detector locations are used as data in the inversion algorithm. In frequency-domain methods, with sinusoidally modulated light, a similar algorithm has been developed with the magnitude and phase of the modulation envelope used as data.[7] The forward frequency-domain diffusion equation has been further approximated and formulated as an integral equation, and the Born approximation has been used to derive a linear integral equation.[3] In this approach the equivalent wave number, which is a nonlinear function of absorption and

0740-3232/99/102400-13$15.00 © 1999 Optical Society of America

scattering coefficients, is computed, and the unknown absorption and the scattering coefficients are recovered from the reconstructed equivalent wave number. These approaches use heuristic linearization to obtain the gradient necessary for the iterative solutions rather than rigorous optimization. Recently we provided a detailed analysis of previous approaches (especially the integral equation approach) from a standard nonlinear optimization point of view and showed that the conventional integral equation approach[3] does not use the correct Fréchet differential for the absorption and the scattering parameters.[8,9] A new integral equation approach was derived and shown to have a performance superior to previous methods.[8] Furthermore, the iterative method used in conventional approaches,[3,6,7] which imposes a penalty on the L_2 norm of the update at each iteration, tends to oversmooth 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,[10] but a trust-region constraint for a nonlinear least-squares problem.[11,12]

The artifacts that are due to poor regularization can be reduced by incorporation of *a priori* information by use of a Bayesian framework. In this framework the maximum *a posteriori* (MAP) reconstruction is often computed by maximization of the posterior distribution. This Bayesian approach has been applied in many image-restoration and image-reconstruction problems.[13,14] More recently, Bayesian (or other regularization) methods have been applied to nonlinear inverse problems such as microwave imaging, impedance tomography, and optical imaging.[7,15,16] 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 Jiang[15] added a quadratic regularization term to their previous formulation[7] 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. The computational complexity of this method is very high because $O(N^3)$ complex multiplications are required at each iteration, where N is the number of image pixels. Saquib *et al.*[16] 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. Arridge and Schweiger[17] adapted this method for the frequency-domain diffusion tomography problem. 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 *et al.*[18] is most closely related to ours, in that they used exact single-site updates for each pixel. They observed that the single-site optimization had rapid convergence in terms of the 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 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.

As in the previous approach of Saquib *et al.*,[16] we use the generalized Gaussian Markov random field (GGMRF) as the 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. Since the dynamic range of the data is usually very large owing to the source–detector geometry and strong attenuation in the medium, large intensity measurements may be overweighted in the inversion procedure. Previously, heuristic scaling based on the time average of the measurements[17] was used to ameliorate this problem. We address the scaling problem by deriving a model for the noise based on shot-noise detection statistics.[19] This 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.[20,21] Since the computational complexity of the ICD–Born method is $O(N)$, it requires much less computation per iteration than the exact single-pixel update algorithm,[18] the Gaussian elimination technique for total variation minimization,[15] and the conventional distorted Born iterative method[22] (DBIM).

Our numerical results for a two-dimensional geometry indicate that the ICD–Born method together with the Bayesian framework yields accurate and fast reconstructions from synthetic data. The implications of the new algorithm will be more dramatic for three-dimensional imaging.

Notation: The following notation is used in this paper:

K,	number of sources;
M,	number of detectors;
$P(=KM)$,	number of measurements;
N,	number of image pixels;
Ω; $\partial\Omega$,	image domain; boundary of Ω;
d_m,	position vector of the mth detector;
s_k,	position vector of the kth point source;
r; r_i,	position vector in Ω; position of the ith voxel or pixel in Ω;
y_{km},	complex measurement at the mth detector due to the kth source;
$\mathbf{y}$,	measurement vector, $\mathbf{y} = [y_{11}, y_{12}, \ldots, y_{1M}, y_{21}, \ldots, y_{KM}]^T$;

$\mathbf{x}$, vector of unknown absorption coefficients, $\mathbf{x} = [\mu_a(r_1),\ldots,\mu_a(r_N)]^T$;
$\psi_k(r, t)$, time-domain solution of diffusion equation (photon flux) at r due to the kth source;
$\phi_k(r, \omega)$, frequency domain solution of diffusion equation at r due to the kth source.

2. 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.[1,23,24] More specifically, let $\psi_k(r, t)$ be the photon rate per unit area generated at time t and position $r \in \Omega$ originating from 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

$$D(r) = \frac{1}{3[\mu_a(r) + \mu_s'(r)]}; \quad (2)$$

where $\mu_a(r)$ is the absorption coefficient, and $\mu_s'(r)$ is the reduced scattering coefficient. The reduced scattering coefficient is defined by $\mu_s'(r) = (1 - g)\mu_s(r)$, where $\mu_s(r)$ is the scattering coefficient and g is the mean cosine of the scattering angle. Although alternative definitions for D have been proposed,[25] we use the the classical definition in Eq. (2). Note that $\psi_k(r, t)$ takes on real positive values, since it corresponds to the number of photons that pass through a unit surface area per unit time.

Practical systems based on time-domain measurements have been implemented,[1,26] but these systems tend to be expensive and noise sensitive. To circumvent these problems, we adopt a frequency-domain approach to the optical diffusion problem.[7,27] To do this, we assume that the light source is amplitude modulated at a fixed angular frequency ω ($\neq 0$), so that $S(t) = \mathrm{Re}[1 + \beta\exp(-j\omega t)]$, 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.[19] When the Fourier transform of Eq. (1) is taken, the partial differential equation that governs the complex modulation envelope $\phi_k(r, \omega)$ becomes

$$\nabla \cdot D(r)\nabla\phi_k(r, \omega) + [-\mu_a(r) + j\omega/c]\phi_k(r, \omega) = -\beta\delta(r - s_k). \quad (3)$$

In the frequency-domain imaging approach, Eq. (3) is used as a forward model.

Throughout this paper a two-dimensional domain is considered, but the approach can be generalized to three dimensions. Figure 1 illustrates the typical experimental scheme that we consider. The region to be imaged is denoted by Ω and is surrounded by K point sources, distributed around the two-dimensional boundary at positions $s_k \in \Omega$, and M detectors, at positions $d_m \in \Omega$. 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. Our objective is then to determine the values of $\mu_a(r)$ from the measured values of $\phi_k(d_m, \omega)$. Note that the complex measurements of $\phi_k(d_m, \omega)$ must be made for each source and detector combination. The specific relationship between $\phi_k(r, \omega)$ and the physical measurements is described in Appendix A.

The domain Ω is discretized into N pixels, where the position of the ith pixel is denoted r_i for $1 \leq i \leq N$. The set of unknown absorption coefficients is denoted by the vector $\mathbf{x}$, where

$$\mathbf{x} = [\mu_a(r_1),\ldots,\mu_a(r_N)]^T. \quad (4)$$

Using this notation, we may express the forward model as a vector-valued function $\mathbf{f}(\mathbf{x})$. The function $\mathbf{f}(\mathbf{x})$ takes on the values of a $P = KM$ dimensional column vector with elements given by

$$\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots, f_P(\mathbf{x})]^T = [\phi_1(d_1, \omega), \phi_1(d_2, \omega),\ldots, \phi_1(d_M, \omega), \phi_2(d_1, \omega),\ldots, \phi_K(d_M, \omega)]^T. \quad (5)$$

The elements of $\mathbf{f}(\mathbf{x})$ then represent the exact values of the flux density for the assumed values of the absorption coefficient $\mathbf{x}$.

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

$$\mathbf{y} = [y_{11}, y_{12},\ldots, y_{1M}, y_{21},\ldots, y_{KM}]^T. \quad (6)$$

Note that there is a measurement corresponding to each source–detector pair.

The estimation of $\mathbf{x}$ from the measurement vector $\mathbf{y}$ is a classic example of an ill-posed inverse problem in which the solution is often underdetermined, nonunique, and noise sensitive. To address this problem, we formulate the solution in a Bayesian framework by computing the MAP estimate for $\mathbf{x}$ given $\mathbf{y}$. The MAP estimate, $\hat{\mathbf{x}}_{\mathrm{MAP}}$, is given by

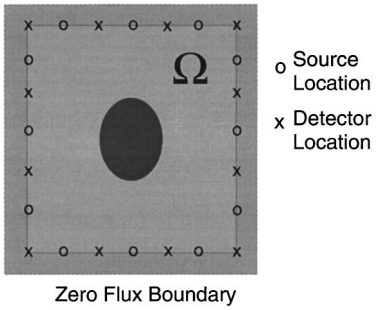

Fig. 1. Simulation geometry with the locations of sources and detectors for inversion of synthetic data. The sources and detectors are uniformly spaced along the edges.

$$\hat{\mathbf{x}}_{\mathrm{MAP}} = \arg\max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y}) \tag{7}$$

$$= \arg\max_{\mathbf{x}} [\log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x})], \tag{8}$$

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. The density $p(\mathbf{y}|\mathbf{x})$ models the physical properties of the measurement system, while the prior density $p(\mathbf{x})$ models image characteristics such as smoothness that one would expect in the solution. In the following sections we derive an expression $p(\mathbf{y}|\mathbf{x})$ based on our modeling assumptions, and we adopt a prior model $p(\mathbf{x})$ that enforces smoothness while allowing for abrupt changes in the estimate of $\mathbf{x}$.

3. MEASUREMENT 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 Appendix A 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 by (see Appendix A)

$$p(y_{km}|\mathbf{x}) = \frac{1}{2\pi\alpha|\phi_k(d_m, \omega)|} \exp\left[-\frac{|y_{km} - \phi_k(d_m, \omega)|^2}{2\alpha|\phi_k(d_m, \omega)|} \right], \tag{9}$$

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

$$C_{ii} = \alpha|\phi_k(d_m, \omega)| \simeq \alpha|y_{km}|, \quad \text{where } i = M(k-1) + m. \tag{10}$$

To simplify notation, we define the diagonal matrix Λ as

$$\Lambda = {}^1\!/_2 C^{-1}. \tag{11}$$

The data likelihood is then given by

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^N|\Lambda|^{-1}} \exp[-\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_\Lambda{}^2], \tag{12}$$

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

4. GENERALIZED GAUSSIAN MARKOV RANDOM-FIELD PRIOR MODEL

In this section, we describe the prior model $p(\mathbf{x})$ that we use for the absorption image $\mathbf{x}$. In many image-reconstruction problems, the Markov random field model[28] has proved useful in describing spatial correlations between neighboring pixels. Markov random fields have the property that the conditional distribution of a pixel, given all other pixels, is a function only of the pixel's neighbors, i.e.,

$$p(x_i|x_j,\ i \neq j) = p(x_i|x_{\partial i}), \tag{13}$$

where ∂i denotes the pixel's neighboring pixel i. If the density function is constrained to be strictly positive, then a random field is a Markov random field if and only if its density function has the form of a Gibbs distribution (Hammersley–Clifford theorem[28]). A Gibbs distribution is any distribution with a density function that can be put in the form

$$p(\mathbf{x}) = \frac{1}{\sigma^N z(p)} \exp\left[-\frac{1}{p} u\left(\frac{\mathbf{x}}{\sigma}, p \right) \right], \tag{14}$$

where σ and p are constants representing scale and shape parameters for the distribution and $z(p)$ is a normalizing constant.

We further assume that the function $u(\mathbf{x}/\sigma, p)$ has the form[29]

$$u(\mathbf{x}/\sigma, p) = \sum_{\{i,j\}\in\mathcal{N}} b_{i-j}\rho\left(\frac{x_i - x_j}{\sigma}, p \right), \tag{15}$$

where $\mathcal{N}$ is the set of all neighboring pixel pairs and $\rho(\cdot,\cdot)$ is a potential function that assigns a cost to differences between neighboring pixel values. A wide variety of functions for $\rho(\cdot,\cdot)$ have been used.[13,14,30,31] However, here we use the GGMRF model because it is both convex and scale invariant.[31] The convexity of the potential function of the GGMRF model 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.[31] For the GGMRF model the density function for $\mathbf{x}$ is given by

$$p(\mathbf{x}) = \frac{1}{\sigma^N z(p)} \exp\left(-\frac{1}{p\sigma^p} \sum_{\{i,j\}\in\mathcal{N}} b_{i-j}|x_i - x_j|^p \right), \quad 1 \leqslant p \leqslant 2. \tag{16}$$

Furthermore, since the absorption must be positive, we also impose the constraint

$$x_i \geqslant 0, \qquad i = 1, \ldots, N. \tag{17}$$

5. ITERATIVE-COORDINATE-DESCENT–BORN OPTIMIZATION TECHNIQUE

Referring to Eqs. (8), (12), and (16) and conditions (17), the MAP estimate for $\mathbf{x}$ is given by

$$\hat{\mathbf{x}}_{\mathrm{MAP}} = \arg\min_{\mathbf{x}\geqslant 0} \left[\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_\Lambda{}^2 + \frac{1}{p\sigma^p} \sum_{\{i,j\}\in\mathcal{N}} b_{i-j}|x_i - x_j|^p \right]. \tag{18}$$

To compute the MAP reconstruction, we must perform optimization (18). We choose to use the ICD algorithm[21] for a number of reasons. First, it has been shown that ICD updates work well with non-Gaussian prior models.[21]

Second, the ICD algorithm is easily implemented with a positivity constraint. In contrast, a drawback of the conjugate gradient method is the difficulty of incorporating positivity constraints.[21]

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 performed 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 ith pixel is given by

$$\hat{x}_i = \arg\min_{\tilde{x}_i \geq 0}\left[\|\mathbf{y} - \mathbf{f}(\tilde{\mathbf{x}}_i)\|_\Lambda{}^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i} b_{i-j}|\tilde{x}_i - x_j|^p\right], \tag{19}$$

where $\tilde{\mathbf{x}}_i = [x_1, x_2, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots, x_N]^T$ and $\mathcal{N}_i$ is the set of pixels neighboring pixel i. Note that x_i is replaced by $\hat{x}_i$ before the next pixel update. However, a direct approach[18] to the update equation (19) is computationally very expensive owing 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 (3) for each source.

The computational inefficiency is overcome by use of 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 nth scan. At the beginning of the $(n+1)$th scan, the approximation

$$\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_\Lambda{}^2 \simeq \|\mathbf{y} - \mathbf{f}(\mathbf{x}^n) - \mathbf{f}'(\mathbf{x}^n)\Delta\mathbf{x}\|_\Lambda{}^2 \tag{20}$$

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

$$f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_P(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_P(\mathbf{x})}{\partial x_N} \end{bmatrix} = \begin{bmatrix} \frac{\partial\phi_1(d_1,\omega)}{\partial x_1} & \frac{\partial\phi_1(d_1,\omega)}{\partial x_2} & \cdots & \frac{\partial\phi_1(d_1,\omega)}{\partial x_{N-1}} & \frac{\partial\phi_1(d_1,\omega)}{\partial x_N} \\ \frac{\partial\phi_1(d_2,\omega)}{\partial x_1} & \frac{\partial\phi_1(d_2,\omega)}{\partial x_2} & \cdots & \frac{\partial\phi_1(d_2,\omega)}{\partial x_{N-1}} & \frac{\partial\phi_1(d_2,\omega)}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial\phi_1(d_M,\omega)}{\partial x_1} & \frac{\partial\phi_1(d_M,\omega)}{\partial x_2} & \cdots & \frac{\partial\phi_1(d_M,\omega)}{\partial x_{N-1}} & \frac{\partial\phi_1(d_M,\omega)}{\partial x_N} \\ \frac{\partial\phi_2(d_1,\omega)}{\partial x_1} & \frac{\partial\phi_2(d_1,\omega)}{\partial x_2} & \cdots & \frac{\partial\phi_2(d_1,\omega)}{\partial x_{N-1}} & \frac{\partial\phi_2(d_1,\omega)}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial\phi_K(d_M,\omega)}{\partial x_1} & \frac{\partial\phi_K(d_M,\omega)}{\partial x_2} & \cdots & \frac{\partial\phi_K(d_M,\omega)}{\partial x_{N-1}} & \frac{\partial\phi_K(d_M,\omega)}{\partial x_N} \end{bmatrix}, \tag{21}$$

and $\mathbf{f}'(\mathbf{x}^n)$ denotes the Fréchet derivative computed for the absorption parameter estimate $\mathbf{x}^n$. We showed previously[8] that the Fréchet differential for the mth detector associated with the kth source can be approximated by

$$\lim_{N\to\infty}\sum_{i=1}^{N}\frac{\partial\phi_k(d_m,\omega)}{\partial x_i}\Delta x_i = \int_\Omega \mathrm{d}r\, g(d_m, r, \omega)\phi_k(r,\omega) \times\left[-1 + \frac{\mu_a{}^n(r) + j\omega/c}{\mu_a{}^n(r) + \mu_s(r)}\right]\Delta\mu_a(r), \tag{22}$$

where $\Delta\mu_a(r) = \mu_a(r) - \mu_a{}^n(r)$ denotes the change in μ_a, and $g(d_m, r, \omega)$ is the Green's function for the frequency domain diffusion equation (3). 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 Eq. (21) are given by

$$\frac{\partial\phi_k(d_m,\omega)}{\partial x_i} = g(d_m, r_i, \omega)\phi_k(r_i,\omega) \times\left[-1 + \frac{\mu_a{}^n(r_i) + j\omega/c}{\mu_a{}^n(r_i) + \mu_s(r_i)}\right]A, \tag{23}$$

where A is the pixel area. After the nth scan, $\mathbf{f}(\mathbf{x}^n)$ and $\mathbf{f}'(\mathbf{x}^n)$ are calculated by computation of $g(d_m, r_i, \omega)$ and $\phi_k(r_i, \omega)$ of Eq. (23) with a linear partial-differential equation solver for the diffusion equation (3) with the nth estimate of the 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

$$\hat{x}_i = \arg\min_{\tilde{x}_i\geq 0}\left\{\|\mathbf{y} - \mathbf{f}(\mathbf{x}^n) - [\mathbf{f}'(\mathbf{x}^n)]_{*i}(\tilde{x}_i - x_i{}^n)\|_\Lambda{}^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i} b_{i-j}|\tilde{x}_i - x_j|^p\right\}, \tag{24}$$

where $[\mathbf{f}'(\mathbf{x}^n)]_{*i}$ is the ith column of the Fréchet derivative matrix. To compute the solution to Eq. (24), we express the first term as a quadratic function of $\tilde{x}_i$ to obtain the expression

For n:=1 **until** stopping criterion; /* for each full iteration through $\mathbf{x}$ */

1. Solve the frequency domain diffusion equation for each source and detector combination using an absorption coefficient estimate $\mathbf{x}^n = \{\mu_a^n(r_i)\}_{i=1}^N$.
2. Fill in the Fréchet derivative matrix entries (21) by computing (23).
3. Compute the initial error vector, $\mathbf{e}(1) = \mathbf{y} - \mathbf{f}(\mathbf{x}^n)$.
4. **For** i:=1 **until** N; /* for each pixel in $\mathbf{x}$ */
   (a) $\theta_1 = -2Re\left[\,[\mathbf{f}'(\mathbf{x}^n)_{*i}]^H \Lambda \mathbf{e}(i)\right]$.
   (b) $\theta_2 = 2[\mathbf{f}'(\mathbf{x}^n)]_{*i}^H \Lambda \mathbf{f}'(\mathbf{x}^n)_{*i}$.
   (c) $\hat{x}_i = \arg\min_{\tilde{x}_i \geq 0}\left[\theta_1(\tilde{x}_i - x_i^n) + \frac{\theta_2}{2}(\tilde{x}_i - x_i^n)^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i} b_{i-j}|\tilde{x}_i - x_j|^p\right]$.
   (d) $\mathbf{e}(i+1) = \mathbf{e}(i) - \mathbf{f}'(\mathbf{x}^n)_{*i}(\hat{x}_i - x_i^n)$;
   (e) $x_i^n := \hat{x}_i$.
   end

end;

Fig. 2. Pseudo-code specification of the ICD–Born algorithm.

$$\hat{x}_i = \arg\min_{\tilde{x}_i \geq 0}\left[\theta_1(\tilde{x}_i - x_i{}^n) + \frac{\theta_2}{2}(\tilde{x}_i - x_i{}^n)^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i} b_{i-j}|\tilde{x}_i - x_j|^p\right], \tag{25}$$

where θ_1 and θ_2 are given by

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

$$\theta_2 = 2[\mathbf{f}'(\mathbf{x}^n)_{*i}]^H \Lambda \mathbf{f}'(\mathbf{x}^n)_{*i}, \tag{26}$$

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

$$\mathbf{e}(1) = \mathbf{y} - \mathbf{f}(\mathbf{x}^n), \tag{27}$$

and for subsequent pixels is updated as

$$\mathbf{e}(i+1) = \mathbf{e}(i) - [\mathbf{f}'(\mathbf{x}^n)]_{*i}(\hat{x}_i - x_i{}^n). \tag{28}$$

Solution of Eq. (25) 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 Eq. (25), i.e.,

$$\theta_1 + \theta_2(\tilde{x}_i - x_i{}^n) + \frac{1}{\sigma^p}\sum_{j\in\mathcal{N}_i} b_{i-j}|\tilde{x}_i - x_j|^{p-1}\,\mathrm{sgn}(\tilde{x}_i - x_j) = 0. \tag{29}$$

This root-finding procedure is performed with a half-interval search[32] because the function in Eq. (29) is monotonically decreasing.[21] The lower and the upper bounds of update $\hat{x}_i$ in Eq. (25) are first computed from the observation that[21]

$$\min\left(x_i{}^n - \frac{\theta_1}{\theta_2},\ x_j \in \mathcal{N}_i\right) \leq \hat{x}_i \leq \max\left(x_i{}^n - \frac{\theta_1}{\theta_2},\ x_j \in \mathcal{N}_i\right). \tag{30}$$

Then these bounds are used as end points for initiating a half-interval search. Since the half-interval search has guaranteed exponential convergence, one can choose to terminate either after a fixed number of iterations or after a fixed tolerance is reached.[32] Figure 2 summarizes the complete procedure used to implement the ICD–Born algorithm.

6. 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 as cflops in the tables) 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 Eq. (21) consists of calculating the Green's function of Eq. (3) and the flux $\phi_k(r_i, \omega)$ and evaluating Eq. (23). Evaluation of $\phi_k(r_i, \omega)$ requires K forward solutions of Eq. (3), one solution for each source location. The evaluation of the Green's function in Eq. (23) implicitly involves placing a source at each grid 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 use of the reciprocity[33] of $\phi_k(r_i, \omega)$ in Eq. (3) so that

$$g(d_m, r_i, \omega) = g(r_i, d_m, \omega). \tag{31}$$

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 Eq. (3) by an iterative linear solver (e.g., SOR[34]) 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 necessary to fill the Fréchet derivative matrix in Eq. (21).

The solution of Eq. (29) 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 Eqs. (26). Since the number of columns of $\mathbf{f}'(\mathbf{x}^n)$ is MK and Λ is diagonal, the number of multiplications necessary 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 the DBIM, let us first briefly explain the DBIM algorithm. In the DBIM, a new parameter estimate is computed from the perturbation equation with a trust-region constraint, where the Fréchet derivative $\mathbf{f}'(\mathbf{x}^n)$ is again defined as in Eqs. (21) and (23). 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 (31), 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 owing to the greater change in $\mathbf{x}$ for each DBIM iteration. To fill in the Fréchet derivative (21), a total of $(2MK + 2)N$ multiplications is again required. The computational cost for the regularized inverse by QR decomposition is at least $2(MK)^2N - (MK)^3/3$ (Ref. 34). 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 contains only a MKN term. This difference is particularly significant as the number of sources and detectors grows.

The computational advantage of the 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 above, the number of forward linear solver iterations for the two algorithms will vary. We used the MUDPACK (multigrid software for elliptic partial differential equations) libraries[35] as our forward solver and controlled the number of iterations by setting the relative error tolerance in MUDPACK.[35] 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

Table 1. Computational Complexity of the DBIM and the ICD–Born Method: Number of Complex Multiplications (cflops) per Iteration[a]

Calculation	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	$2(MK)^2N + 5(M + K)LN$	$7MKN + 5(M + K)LN$

[a] M, number of detectors; K, number of sources; L, number of relaxations for the forward partial-differential equation solver; N, number of image pixels.

Table 2. Comparison of the Computation Required (Complex Multiplications, cflops) for One Iteration of the DBIM and the ICD–Born Method[a]

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

[a] M, number of detectors; K, number of sources; L, number of relaxations for the forward partial-differential equation solver; N, number of image pixels.

Table 3. Average CPU Time (s) per Interation for the First Simulation (Fig. 3)

	ICD–Born	
DBIM	$p = 2.0$	$p = 1.1$
4.0	0.23	0.32

L_{ICD} that we believe to be reasonable. These choices are justified by the results presented in Table 3, which shows the actual CPU times per iteration.

Two cases are considered, one that uses the same number of forward iterations for both methods and one that uses different numbers of forward iterations. The results listed in Table 2 show that for these two cases the computation by the ICD–Born method is reduced by factors of 19 and 13, respectively, as compared with that by the DBIM. Although most investigations of optical diffusion imaging consider two-dimensional geometries, the real problem is three dimensional. Therefore in the bottom row of Table 2 we show the estimated number of complex multiplications for a three-dimensional reconstruction scheme with a similar geometry. The computational speedup of the ICD–Born method is even more dramatic in this case.

7. 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 8 cm $\times$ 8 cm domain, as shown in Fig. 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. Referring to Eq. (9), the signal-to-noise ratio (SNR) of the current measurement for the mth detector, and the kth source can be represented by the magnitude of the flux and is given by

$$\mathrm{SNR}_{mk} = \frac{|E[i]|^2}{{\sigma_0}^2} \simeq \frac{1}{2\alpha}|\phi_k(d_m, \omega)|, \tag{32}$$

where the constant α and noise current variance ${\sigma_0}^2$ are defined in Appendix A. (Note that the signal-to-noise ratio increases proportionally to the magnitude of the flux at the detector.) To determine the value of α in relation (32), we consider a 1-MHz bandwidth detection system ($B = 1$ MHz) and a 200-MHz modulation frequency with a modulation depth of $\beta = 1.0$. For detector parameters, a common laboratory photomultiplier tube, the R928 from Hamamatsu, is considered. Here the cathode responsivity κ is 68 mA/W (Ref. 36), and γ is set to 1 mm^2. This gives $\alpha = 4.7059 \times 10^{-6}$. The noise is generated by a Gaussian random-number generator and is added to the signal as

$$\mathrm{Re}[y_{mk}] = \mathrm{Re}[\phi_k(d_m, \omega)] + [\alpha|\phi_k(d_m, \omega)|]^{1/2} \times N(0,1),$$

$$\mathrm{Im}[y_{mk}] = \mathrm{Im}[\phi_k(d_m, \omega)] + [\alpha|\phi_k(d_m, \omega)|]^{1/2} \times N(0,1), \tag{33}$$

where $N(0,1)$ is a zero mean Gaussian random variable with unit variance. The average signal-to-noise ratio over the detectors was approximately 30 dB for each simulation. To compare the overall accuracy of the reconstructions, we introduce the normalized root-mean-square error (NRMSE) of the reconstructed image as a function of iteration, defined as

$$\mathrm{NRMSE}_n = \left\{ \frac{\sum_{i=1}^{N} [\mu_a{}^n(r_i) - \mu_a(r_i)]^2}{\sum_{i=1}^{N} [\mu_a(r_i)]^2} \right\}^{1/2}, \tag{34}$$

where $\mu_a{}^n(r_i)$ is the reconstructed value of the absorption coefficient at mesh location r_i at the nth iteration and $\mu_a(r_i)$ is the correct value. For the simulations the image is discretized on a 33 $\times$ 33 grid. In all the simulations the scattering coefficient is set to a constant value $\mu_s = 10.0\,\mathrm{cm}^{-1}$. The simulation was performed on a Sun Ultra Sparc 30.

Figure 3(a) shows the original μ_a image used for the first numerical experiment. Figure 3(b) shows the DBIM reconstruction result.[8] 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

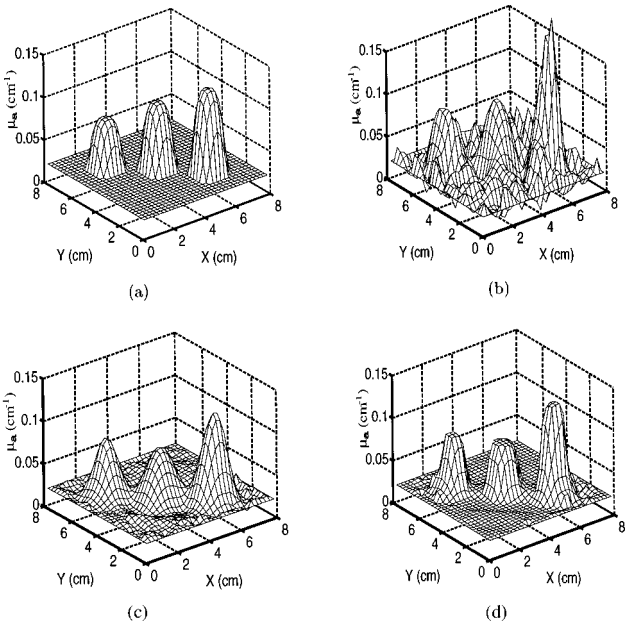

Fig. 3. Reconstruction results for μ_a: (a) original absorption image; (b) reconstruction by the DBIM; (c) reconstruction by the new algorithm with a Gaussian prior ($p = 2.0$, $\sigma = 1.00 \times 10^{-3}$); (d) reconstruction by the new algorithm with a GGMRF prior ($p = 1.1$, $\sigma = 2.31 \times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

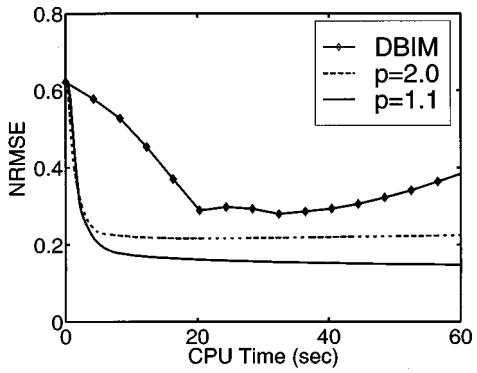

Fig. 4. NRMSE versus CPU time for the ICD–Born method (with $p = 1.1$ and $p = 2$) and the DBIM.

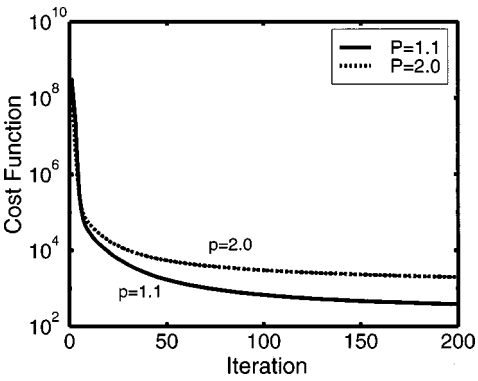

Fig. 5. Cost function (log scale) versus iteration for the ICD–Born algorithm with $p = 1.1$ and $p = 2$.

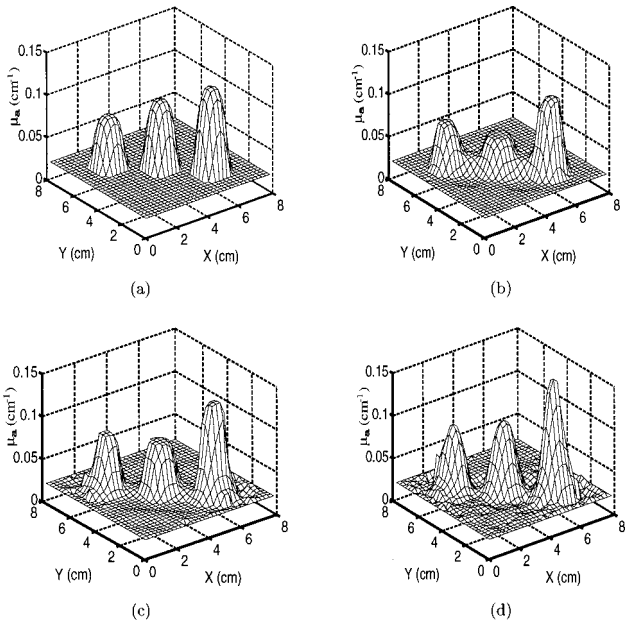

Fig. 6. Reconstructions of μ_a showing the effect of σ for a GGMRF prior model and with $p = 1.1$: (a) original absorption image, and reconstructions with (b) $\sigma = 2.85 \times 10^{-5}$, (c) $\sigma = 2.31 \times 10^{-4}$, (d) $\sigma = 1.52 \times 10^{-2}$.

running the ICD–Born algorithm with a positivity constraint. The stopping criterion used was a fixed CPU time of 60 s (to allow comparison of the two algorithms). Figure 3(c) shows the MAP reconstruction for a Gaussian prior model ($p = 2$, $\sigma = 1.0 \times 10^{-3}$), and Fig. 3(d) gives the reconstruction for $p = 1.1$ and $\sigma = 2.31 \times 10^{-4}$. The reconstruction for the Gaussian prior ($p = 2$) reduces the background noise compared with the reconstruction by the DBIM but suffers from some smoothing of the edges as the price for noise suppression. The boundaries can be sharpened when a larger value of σ is used, but at the expense of larger noise artifacts. Sharper edges and good noise suppression are obtained for $p = 1.1$ [Fig. 3(d)]. Here, because of the edge-preserving nature of the GGMRF prior, the edges are more noticeable in the reconstruction while the noise is suppressed.

Figure 4 shows a plot of the NRMSE versus CPU time for this simulation. With a single-user environment, the CPU times were ~80% of the wall clock time. The average CPU time for a single iteration of each algorithm is also given in Table 3. The relative computational speeds for the DBIM and the ICD–Born method are consistent with the results shown in the first two rows of Tables 1 and 2. Note that more CPU time is required for the ICD–Born method with $p = 1.1$ than for $p = 2.0$. This is due to the increased number of the half-interval searches for $p = 1.1$. Figure 4 shows that the DBIM eventually diverges. This is again due to the insufficient regularization of the DBIM.[22]

Figure 5 shows a plot of the cost function in Eq. (18) versus the iteration number for the ICD–Born method with $p = 2.0$ and $p = 1.1$ for the above simulation. Note that the ICD–Born method is very stable and has reasonably fast convergence. The convergence behavior of the ICD–Born method is also consistent with a previously described single-site update algorithm.[18]

The problem of choosing p and σ is complex, and a maximum-likelihood estimation technique can be used.[29] However, in this paper we empirically choose values that give the best results. For σ, the range of $0 < \sigma < 0.02\ \mathrm{cm}^{-1}$ was considered (note the units of σ). In Fig. 6 the reconstructions obtained with the GGMRF prior with $p = 1.1$ for different values of σ are shown. A small value of $\sigma = 2.85 \times 10^{-5}\ \mathrm{cm}^{-1}$ results in an oversmoothed reconstruction, as is observed in Fig. 6(b). Note that in this case the amplitude of the inhomogeneity at the center is underestimated. A larger value of $\sigma = 1.52 \times 10^{-2}\ \mathrm{cm}^{-1}$ produces background noise in the reconstruction, even though the edges are improved and the value of the inhomogeneity at the center is improved [Fig. 6(d)]. Figure 6(c), with $\sigma = 2.31 \times 10^{-4}\ \mathrm{cm}^{-1}$, shows the best trade-off between the smoothness of the image and the edge improvement.

Figure 7 shows a variety of more complicated absorption images. Reconstructions using the ICD–Born method for these images are shown in Fig. 8. The parameters used for the reconstructions and the final NRMSE are given in the caption of Fig. 8. In each case the reconstructions are accurate both quantitatively and qualitatively, demonstrating that our new algorithm performs well on more complex images.

8. 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 permits the incorporation of prior knowledge of the unknown parameters as well as detection statistics. The shot-noise detection statistics we describe here provide a natural weighting for the measurement data. As a prior model of the unknown parameter, we use the generalized Gaussian random field (GGMRF). This results in stable and edge-preserving regularization for this inverse problem. As an optimization technique for the Bayesian framework, we use a new method that combines iterative coordinate descent (ICD) and the Born approximation, which we call the ICD–Born method. This method significantly reduces the computational complexity compared with methods such as the conventional distorted Born iterative method (DBIM). Numerical simulations for a two-dimensional geometry show that the Bayesian framework together with the ICD–Born method significantly improves the quality of reconstructions. Although practically all simulations of optical diffusion have considered a two-dimensional geometry, the problem is fundamentally three dimensional (unlike x-ray tomography), since the photon paths are not restricted to a plane. Efficient and effective inversion algorithms for optical diffusion tomography in three dimensions is currently an open problem that needs investigation. The overall properties of the inversion algorithm described here are expected to be similar for the three-dimensional problem. However, the improvement in computational efficiency over that of existing algorithms will be even more significant in three dimensions than in two dimensions because of the large increase in the number of unknowns in three dimensions. The algorithm, therefore, has significant potential in practical optical diffusion imaging.

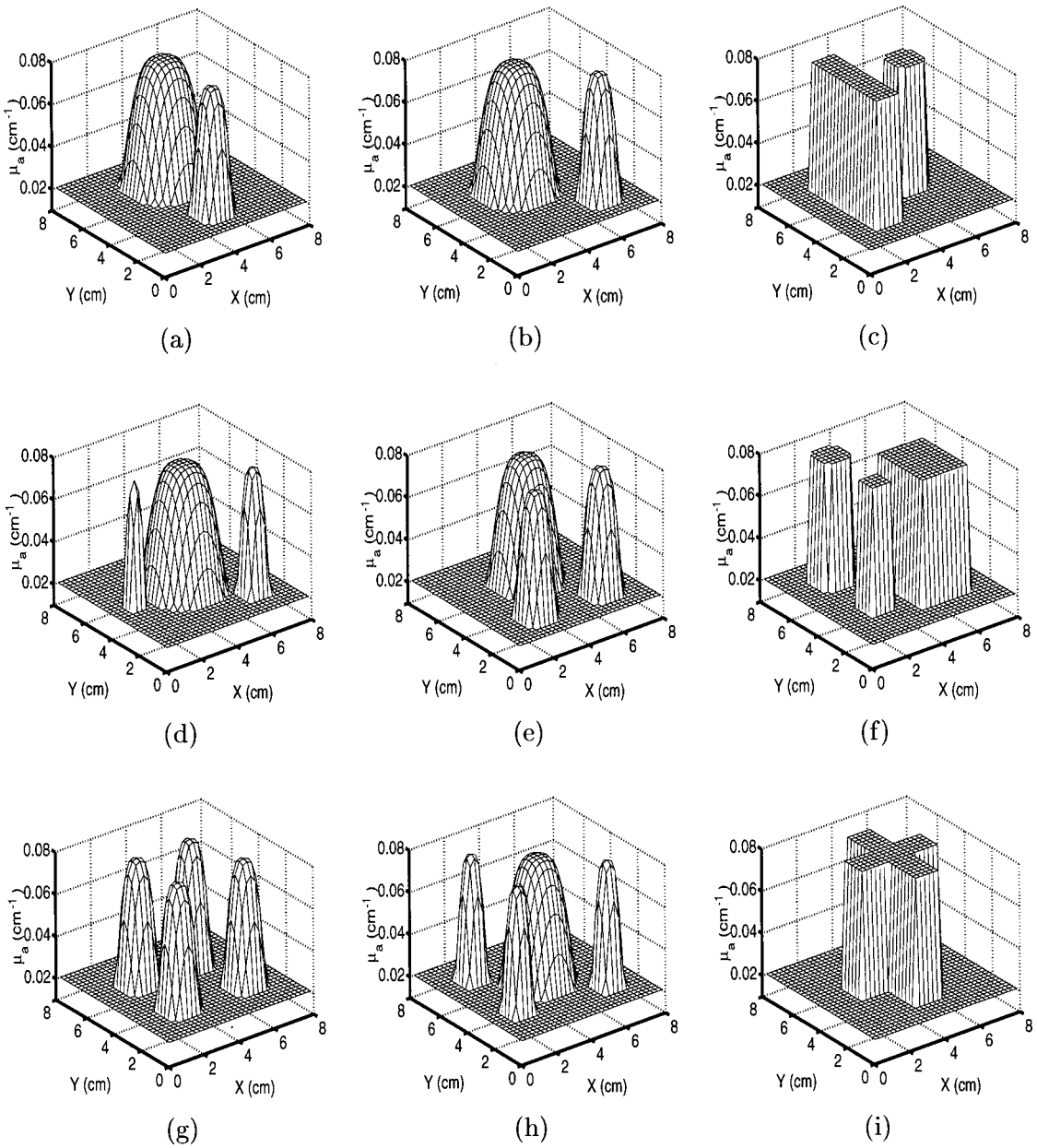

Fig. 7. Variety of true absorption images used for simulations.

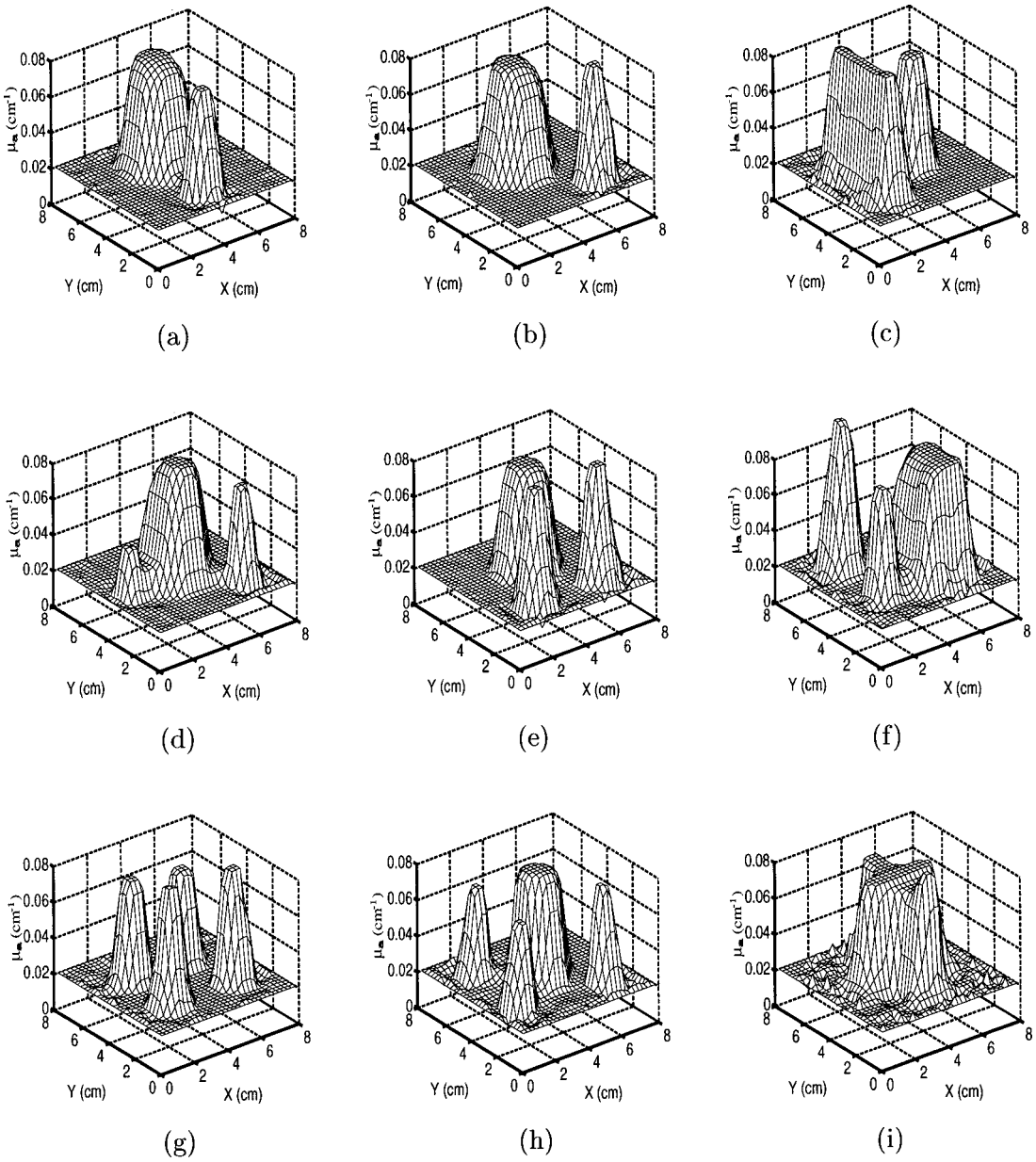

Fig. 8. Reconstructions of the absorption images shown in Fig. 7 by the ICD–Born algorithm with the GGMRF prior with $p = 1.1$ and $\sigma = 2.31 \times 10^{-4}$. NRMSE values for the final reconstructions are (a) 5.83×10^{-2}, (b) 5.56×10^{-2}, (c) 1.92×10^{-1}, (d) 1.25×10^{-1}, (e) 7.70×10^{-2}, (f) 2.18×10^{-1}, (g) 8.34×10^{-2}, (h) 1.26×10^{-1}, (i) 2.08×10^{-1}.

APPENDIX A: DETECTOR NOISE MODEL

We develop in this appendix a noise model for the measured data y_{km}. A measurement will sense the optical power leaving the scattering domain, the photon current $\mathbf{J}$, which is given by Fick's law[4] as $\mathbf{J} = -D\nabla\phi$. The average (dc) optical power P at a detector is therefore

$$P = |\mathbf{J}^+(d_m)|A_e = D(d_m)|\nabla\phi_k(d_m, \omega)|A_e, \quad \text{(A1)}$$

where $\mathbf{J}^+(d_m)$ is the outward photon current from the domain at the detector, A_e is the effective detector aperture, and $D(d_m)$ is the diffusion constant at the receiver location d_m.

We consider an absorbing boundary condition as follows[4] (Fig. 9). The detectors are located on a boundary, denoted $\partial\Omega_d$, inside the computational boundary $\partial\Omega$, such that all the incident photons are absorbed at the detector boundary. This means that the inward photon current at the detectors $\mathbf{J}^-(d_m) = \mathbf{0}$. This boundary condition is implemented by setting the flux $\phi = 0$ on the computational boundary $\partial\Omega$ at a distance $2.131D(d_m)$ from $\partial\Omega_d$ (Ref. 4). Referring to Fig. 9, we can therefore approximate $|\nabla\phi_k(d_m, 0)|$ by $\phi_k(d_m, 0)/[2.131D(d_m)]$ and the optical power by

$$P = \gamma|\phi_k(d_m, 0)|, \quad \text{(A2)}$$

where $\gamma = A_e/2.131$.

We assume an ideal photodetector in which thermal noise is negligible so that the noise is dominated by shot noise. Shot noise for the photodetector current is described by Gaussian statistics in which the variance is

proportional to the mean.[19] The probability density function for the complex envelope i of the detector current is then

$$p(i) = \frac{1}{2\pi\sigma_0^2}\exp\left(-\frac{|i - E[i]|^2}{2\sigma_0^2}\right), \tag{A3}$$

where the noise current variance for shot noise is given by

$$\sigma_0^2 = 2eBi_0, \tag{A4}$$

with e the electron charge, B the bandwidth and $i_0 = E[i]$ the dc current.[19] The current i_0 is proportional to the detector dc optical power, i.e, $i_0 = \kappa P$, where κ is the detector responsivity,[36] so that using Eqs. (A2)–(A4) allows the probability density function for the data $y_{km} = \phi_k(d_m, \omega)$ to be derived as

$$p(y_{km}|\mathbf{x}) = \frac{1}{2\pi\eta\phi_k(d_m, 0)}\exp\left[-\frac{|y_{km} - \phi_k(d_m, \omega)|^2}{2\eta\phi_k(d_m, 0)}\right], \tag{A5}$$

where $\eta = 2eB/(\kappa\gamma)$ and $p(y_{km}|\mathbf{x})$ is the probability density function for each measured datum given a particular image $\mathbf{x}$. Furthermore, we show in Appendix B that, for typical tissue optical parameters and typical (low) modulation frequencies, $\phi_k(d_m, 0) \simeq |\phi_k(d_m, \omega)|/\beta$, so that (A5) can be replaced by

$$p(y_{km}|\mathbf{x}) = \frac{1}{2\pi\alpha|\phi_k(d_m, \omega)|}\exp\left[-\frac{|y_{km} - \phi_k(d_m, \omega)|^2}{2\alpha|\phi_k(d_m, \omega)|}\right], \tag{A6}$$

where $\alpha = \eta/\beta$.

APPENDIX B: APPROXIMATION OF DC PHOTON FLUX DENSITY

Consider the time-domain diffusion equation (1) in an infinite homogeneous region with $S(t) = 1 + \beta\exp(-j\omega t)$ located at s_k, which yields the closed-form solution[37]

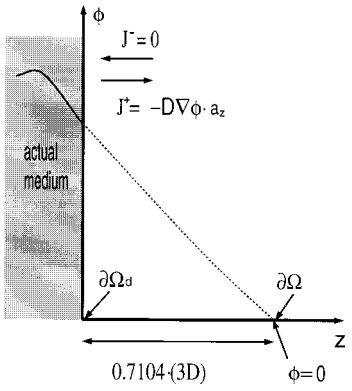

Fig. 9. Illustration of the zero-input photon current or absorbing boundary condition for the diffusion equation, where all incident light from within the scattering boundary is lost to free space. Setting $\phi = 0$ on an extrapolated boundary at $0.7104(3D)$, where $3D$ is the mean free path, is equivalent to the zero-input current condition on the physical boundary. In this figure z is the distance variable perpendicular to the interface and $\mathbf{a}_z$ is the unit vector.

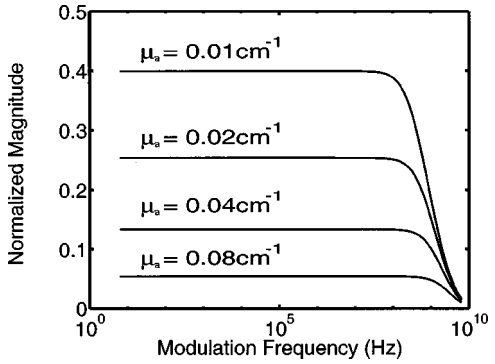

Fig. 10. Normalized (with respect to the source strength) magnitude of the complex optical signal versus modulation frequency for several absorption coefficients. The scattering coefficient is set to be $\mu_s = 10.0\,\mathrm{cm}^{-1}$ and the distance is 4 cm. The plots are computed for the analytic form, Eq. (B2).

$$\psi_k(r, t) = \mathrm{Re}\left(\frac{1}{4\pi Dz}\exp\left[-z\left(\frac{\mu_a}{D}\right)^{1/2}\right] + \frac{\beta}{4\pi Dz} \times \exp\left\{-z\left(\frac{c^2\mu_a^2 + \omega^2}{c^2D^2}\right)^{1/4} \times \cos\left[\frac{1}{2}\arctan\left(\frac{\omega}{c\mu_a}\right)\right]\right\} \times \exp\left\{-jz\left(\frac{c^2\mu_a^2 + \omega^2}{c^2D^2}\right)^{1/4} \times \sin\left[\frac{1}{2}\arctan\left(\frac{\omega}{c\mu_a}\right)\right] + j\omega t\right\}\right), \tag{B1}$$

where $z = |s_k - d_m|$. When reference signals $2\cos(\omega t)$ and $2\sin(\omega t)$ are used for the in-phase and quadrature phase of the measured current, respectively, the magnitude of the modulation envelope of the current becomes $|i| = \kappa\gamma|\phi_k(d_m, \omega)|$, where, for $\omega \neq 0$,

$$|\phi_k(d_m, \omega)| = \frac{\beta}{4\pi Dz}\exp\left\{-z\left(\frac{c^2\mu_a^2 + \omega^2}{c^2D^2}\right)^{1/4} \times \cos\left[\frac{1}{2}\arctan\left(\frac{\omega}{c\mu_a}\right)\right]\right\}. \tag{B2}$$

For typical optical parameters in tissue and modulation frequencies ($\omega/2\pi \leqslant 2 \times 10^8$ Hz), simulations (Fig. 10) show that Eq. (B2) is nearly independent of ω and is equal to the dc value times β. This is because $[(c^2\mu_a^2 + \omega^2)/(c^2D^2)]^{1/4}$ is an increasing function of ω, but $\cos\{1/2\arctan[\omega/(c\mu_a)]\}$ is a decreasing function, for small ω. Therefore $\phi_k(d_m, 0) \simeq 1/\beta|\phi_k(d_m, \omega)|$. We do not know the exact value of $\phi_k(d_m, \omega)$ because of measurement noise, so we use the estimate $\phi_k(d_m, \omega) \simeq y_{mk}$.

ACKNOWLEDGMENTS

The authors thank Thomas J. Downar of the School of Nuclear Engineering at Purdue University for discussions

concerning solution to the diffusion equation. This work was supported by a Purdue Research Foundation Fellowship and National Science Foundation grant MIP 97-07763.

Address correspondence to Kevin Webb, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907-1285; telephone, 765-494-3373; fax, 765-494-2706; e-mail, webb@ecn.purdue.edu.

*Present address: Coordinate Science Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1308 W. Main Street, Urbana, Illinois 61801.

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**, 2331–2336 (1989).
2. E. M. Sevick, J. K. Frisoli, C. L. Burch, and J. R. Lakowicz, "Localization of absorbers in scattering media by use of frequency-domain measurements of time-dependent photon migration," Appl. Opt. **33**, 3562–3570 (1994).
3. 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**, 325–342 (1997).
4. J. J. Duderstadt and L. J. Hamilton, *Nuclear Reactor Analysis* (Wiley, New York, 1976).
5. 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**, 1162–1168 (1989).
6. 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 *Photon Migration and Imaging in Random Media and Tissues*, B. Chance and R. R. Alfano, eds., Proc. SPIE **1888**, 360–371 (1993).
7. 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**, 253–266 (1996).
8. J. C. Ye, K. J. Webb, R. P. Millane, and T. J. Downar, "Modified distorted Born iterative method with an approximate Fréchet derivative for optical diffusion tomography," J. Opt. Soc. Am. A **16**, 1814–1826 (1999).
9. J. C. Ye, R. P. Millane, K. J. Webb, and T. J. Downar, "Importance of the ∇D term in frequency-resolved optical diffusion imaging," Opt. Lett. **23**, 1423–1425 (1998).
10. A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, F. John, transl. ed. (Winston, New York, 1977).
11. J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, N.J., 1983).
12. R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (Wiley, Chichester, UK, 1987).
13. T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," IEEE Trans. Med. Imaging **8**, 194–202 (1989).
14. D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," IEEE Trans. Pattern. Anal. Mach. Intell. **14**, 367–383 (1992).
15. K. D. Paulsen and H. Jiang, "Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization," Appl. Opt. **35**, 3447–3458 (1996).
16. S. S. Saquib, K. M. Hanson, and G. S. Cunningham, "Model-based image reconstruction from time-resolved diffusion data," in *Medical Imaging 1997: Image Processing*, K. M. Hanson, ed., Proc. SPIE **3034**, 369–380 (1997).
17. S. R. Arridge and M. Schweiger, "A gradient-based optimisation scheme for optical tomography," Opt. Express **2**, 213–226 (1998), http://epubs.osa.org/opticsexpress.
18. H. Carfantan, A. Mohammad-Djafari, and J. Idier, "A single site update algorithm for nonlinear diffraction tomography," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing IV* (Institute of Electrical and Electronics Engineers, New York, 1997), pp. 2837–2840.
19. R. H. Kingston, *Detection of Optical and Infrared Radiation* (Springer-Verlag, New York, 1978).
20. K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," IEEE Trans. Signal Process. **41**, 534–548 (1993).
21. C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," IEEE Trans. Image Process. **5**, 480–492 (1996).
22. W. C. Chew, *Waves and Fields in Inhomogeneous Media* (Van Nostrand Reinhold, New York, 1990).
23. M. S. Patterson, J. D. Moultan, 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**, 4474–4476 (1991).
24. M. O'Leary, D. Boas, B. Chance, and A. Yodh, "Refraction of diffuse photon density waves," Phys. Rev. Lett. **69**, 2658–2661 (1992).
25. R. Aronson and N. Corngold, "The photon diffusion coefficient in an absorbing medium," J. Opt. Soc. Am. A **16**, 1066–1071 (1999).
26. S. Jacques, "Time resolved propagation of ultrashort laser pulses within turbid tissues," Appl. Opt. **28**, 2223–2229 (1989).
27. J. S. Reynolds, S. Yeung, A. Przadka, and K. J. Webb, "Optical diffusion imaging: a comparative numerical and experimental study," Appl. Opt. **35**, 3671–3679 (1996).
28. J. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. R. Stat. Soc. B **36**, 192–236 (1974).
29. S. S. Saquib, C. A. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," IEEE Trans. Image Process. **7**, 1029–1044 (1998).
30. J. Besag, "Towards Bayesian image analysis," J. Appl. Stat. **16**, 395–407 (1989).
31. C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving map estimation," IEEE Trans. Image Process. **2**, 296–310 (1993).
32. D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. (Addison-Wesley, Reading, Mass., 1989).
33. S. R. Arridge, "Photon-measurement density functions. Part 1: Analytical forms," Appl. Opt. **34**, 7395–7409 (1995).
34. G. H. Golub and C. F. V. Loan, *Matrix Computations*, 2nd ed. (The Johns Hopkins U. Press, Baltimore, 1989).
35. J. C. Adams, "MUDPACK: Multigrid portable FORTRAN software for the efficient solution of linear elliptic partial differential equations," Appl. Math. Comput. **34**, 113–146 (1989).
36. M. G. Erickson, J. S. Reynolds, and K. J. Webb, "Comparison of sensitivity for single-source and dual-interfering-source configurations in optical diffusion imaging," J. Opt. Soc. Am. A **14**, 3083–3092 (1997).
37. J. B. Fishkin and E. Gratton, "Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge," J. Opt. Soc. Am. A **10**, 127–140 (1993).