Gradient-Based Iterative Image Reconstruction Scheme for Time-Resolved Optical Tomography

Andreas H. Hielscher,* Alexander D. Klose, and Kenneth M. Hanson

***Abstract*—Currently available tomographic image reconstruction schemes for optical tomography (OT) are mostly based on the limiting assumptions of small perturbations and *a priori* knowledge of the optical properties of a reference medium. Furthermore, these algorithms usually require the inversion of large, full, ill-conditioned Jacobian matrixes. In this work a gradient-based iterative image reconstruction (GIIR) method is presented that promises to overcome current limitations. The code consists of three major parts: 1) A finite-difference, time-resolved, diffusion forward model is used to predict detector readings based on the spatial distribution of optical properties; 2) An objective function that describes the difference between predicted and measured data; 3) An updating method that uses the gradient of the objective function in a line minimization scheme to provide subsequent guesses of the spatial distribution of the optical properties for the forward model. The reconstruction of these properties is completed, once a minimum of this objective function is found. After a presentation of the mathematical background, two- and three-dimensional reconstruction of simple heterogeneous media as well as the clinically relevant example of ventricular bleeding in the brain are discussed. Numerical studies suggest that intraventricular hemorrhages can be detected using the GIIR technique, even in the presence of a heterogeneous background.**

***Index Terms*— Infrared imaging, optical tomography, time-resolved imaging, tomographic reconstruction, turbid media.**

I. INTRODUCTION

In recent years researchers have invested considerable efforts toward tomographic imaging systems that use near-infrared (NIR) light [1]–[5]. In this novel medical imaging technique, commonly referred to as optical tomography (OT), one attempts to reconstruct the spatial distribution of optical properties (absorption and transport scattering coefficients, μ_a and μ_s') within the body from measurements of transmitted near-infrared light intensities. The source is typically a laser whose light is delivered through optical fibers to several locations around or inside the organ under investigation. The technology for making such light-transmission measurements on human subjects is nowadays readily available [6]–[10] and has been applied in a variety of pilot studies concerned with monitoring of blood oxygenation [11]–[18], hemorrhage detection [8], [19]–[21], functional imaging of brain activities [22]–[26], Alzheimer diagnosis [12], [27], [28], early diagnosis of rheumatic disease in joints [29]–[31], and breast cancer detection [7], [32]–[37]. However, a major challenge remains the development of algorithms that efficiently transform these measurements into accurate cross-section images of various body parts.

Currently available reconstruction algorithms for OT have several limitations that need to be overcome before they can be routinely applied in a clinical setting. In contrast to X rays, near-infrared photons used in OT do not cross the medium on a straight line from the source to the detector. The photons are strongly scattered throughout the tissue. Hence, standard backprojection methods have only limited success [38]–[41].

A majority of available reconstruction algorithms are based on perturbation methods [42]–[59]. These algorithms have limited practical application because of their inherent assumption that the variations in the optical properties within the medium are small, or that the properties of a reference medium similar to the unknown medium are available. Furthermore, they are computationally very expensive since they require the inversion of large full ill-conditioned Jacobian matrixes.

In this paper we report on a new approach, which we refer to as gradient-based iterative image reconstruction (GIIR), that overcomes problems encountered with the perturbation method. First we will review the basic features of perturbation schemes to contrast them to the GIIR technique. This review will be followed by a detailed account of the particular structure and mathematical background of a time-resolved image reconstruction algorithm. To illustrate the performance of that code we will present results for two-dimensional (2-D) and three-dimensional (3-D) reconstruction of simple heterogeneous media as well as simulated reconstructions of a slice through an infant's brain.

II. BACKGROUND

A. Perturbation Methods with Full Matrix Inversion

As pointed out in a recent review by Arridge and Hebden [60] the majority of the current research in OT is based on a perturbation approach involving the inversion of large Jacobian matrixes. (Arridge and Schweiger [42]–[45], Barbour *et al.* [46]–[49], Paulsen and Jiang [50]–[54], Schottland *et al.* [55], O'Leary *et al.* [56], Paithankar *et al.* [57], and Klibanov and Lucas [58], [59]). All these methods have

---

Manuscript received July 21, 1998; revised February 16, 1999. This work was supported in part by a fellowship from the Los Alamos National Laboratory and by the Dean's Office of the College of Medicine, State University of New York Health Science Center. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was R. Barbour. *Asterisk indicates corresponding author.*

*A. H. Hielscher and A. D. Klose are with the State University of New York, Downstate Medical Center, Brooklyn Department of Pathology, Box 25, 450 Clarkson Avenue, Brooklyn, NY 11203 USA.

K. M. Hanson is with the Los Alamos National Laboratory, DX-3: Hydrodynamic Applications, MS P940, Los Alamos, NM 87545 USA.

Publisher Item Identifier S 0278-0062(99)04008-2.

in common the assumption that the unknown distribution of optical properties, $\zeta = [\mu_a(\boldsymbol{r}), \mu'_s(\boldsymbol{r})]$ is a small perturbation to an estimated distribution ζ_e. From the assumed distribution of optical properties, it is possible to predict measurements values $\boldsymbol{M}_p$ given a forward model or theory $\boldsymbol{F}, \boldsymbol{M}_p = \boldsymbol{F}[\zeta_e]$. The forward model is typically a diffusion equation derived as an approximation to the equation of radiative transfer. Experimental measurement values $\boldsymbol{M}$ are commonly considered to be taken on the boundary $\partial\Omega$ of the system under investigation. A variety of parameters have been considered for $\boldsymbol{M}$, e.g., fluence rates, either obtained by time- or frequency-dependent measurements, maximum of fluence in time-resolved measurements, phase shifts of photon density waves, etc.

Assuming that the estimated spatial distribution ζ_e is close to the actual distribution ζ, one can perform a Taylor expansion as follows:

$$M = F[\zeta_e] + F'[\zeta_e](\zeta-\zeta_e) + (\zeta-\zeta_e)^T F''[\zeta_e](\zeta-\zeta_e) + \cdots \qquad (1)$$

where $\boldsymbol{F}'$ and $\boldsymbol{F}''$ are the first- and second-order derivatives of the forward model with respect to the optical properties ζ. The derivatives $\boldsymbol{F}'$ and $\boldsymbol{F}''$ can be represented as matrixes and are commonly referred to as the Jacobian or the weightfunction $(\boldsymbol{F}' \rightarrow \boldsymbol{J})$ and the Hessian $(\boldsymbol{F}'' \rightarrow \boldsymbol{H})$ of the problem. If we define the difference between experimentally obtained and predicted measurement values as $\Delta\boldsymbol{M} = \boldsymbol{M} - \boldsymbol{M}_p$ and the difference between actual and estimated distribution of optical properties as $\Delta\zeta = \zeta - \zeta_e$, (1) becomes

$$\Delta M = J[\zeta_e]\Delta\zeta + \Delta\zeta^T H[\zeta_e]\Delta\zeta + \cdots. \qquad (2)$$

By neglecting second-order terms on the right-hand side, (2) becomes linear. The reconstruction problem is then reduced to solving a set of linear equations for $\Delta\zeta$

$$\Delta\zeta = J^{-1}[\zeta_e]\Delta M. \qquad (3)$$

Knowing $\Delta\zeta$ and the reference medium ζ_e allows to calculate the distribution $\zeta = \Delta\zeta + \zeta_e$, which is the desired image. This approach may be generalized to iterative reconstruction schemes, which successively update ζ_e so that $\zeta = \zeta_e + \Delta\zeta_1 + \Delta\zeta_2 + \cdots$. In Fig. 1 these successive updates are described as outer iteration.

The major computational effort of this approach lies in the inversion of the large full ill-conditioned Jacobian matrix. Usually this inversion is not attempted directly, but is considered as an optimization problem in which the functional

$$||J[p_e]\Delta\zeta - \Delta M|| = \Phi(\Delta\zeta) \qquad (4)$$

is minimized. Commonly applied techniques are conjugated gradient descent (CGD), singular value decomposition (SVG), projection onto convex sets (POCS), algebraic or simultaneous algebraic reconstruction techniques (ART or SART) [48], or the recently developed elliptic system method (ESM) [58]. In Fig. 1, this step is called inner iteration.

Since the problem of solving (3) is ill posed, regularization techniques are often employed, which put additional constraints on the solution vector $\Delta\zeta$. Regularization of these problems is obtained by making the highly ill-conditioned Jacobian matrix more diagonally dominant [51]. For more details on the perturbation techniques and other reconstruction methods currently applied in OT, see the review by Arridge and Hebden [60].

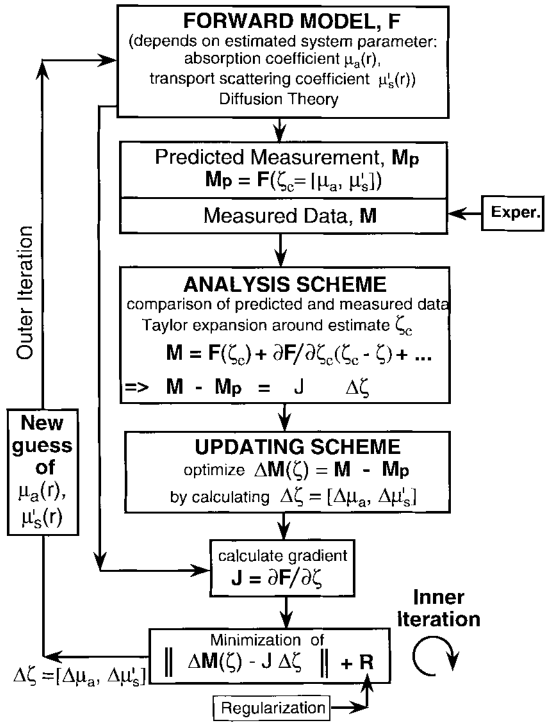


Fig. 1. Basic flow chart for reconstruction algorithms based on perturbation methods that employ full Jacobian matrix inversion.

B. GIIR Schemes

GIIR schemes [61]–[63] differ fundamentally from the perturbation approach employed by the majority of researchers in the field of optical tomography. This statement is true even though many common features exist. As in all other methods, the goal of the GIIR scheme is to reconstruct the distribution of the optical properties inside a medium from a given set $\boldsymbol{M}$ of measurements on the circumference, $\partial\Omega$ of the medium. Analogous to the perturbation approach depicted in Fig. 1, we can divide the GIIR scheme in three different major components (see Fig. 2).

1) *Forward Model:* As in the perturbation approach, this model is a theory or algorithm that predicts a set of measured signals $\boldsymbol{M_p}$ based on the position of the light source and the spatial distribution of optical properties $\zeta = [\mu_a(\boldsymbol{r}), \mu'_s(\boldsymbol{r})]$.
2) *Analysis Scheme:* Here, an objective function ϕ is defined, which describes the difference between the measured, $\boldsymbol{M}$ and predicted data, $\boldsymbol{M_p}$. A simple example is the least square error norm $\phi(\zeta) \sim (\boldsymbol{M} - \boldsymbol{M_p}(\zeta))^2$. Since the problem is highly ill posed, a penalty or regularization term R is usually added to the objective function. Note that by defining the objective function in this way, **no** linearization of the problem is performed.
3) *Updating Scheme:* Once the objective function is defined, the task becomes to minimize ϕ. This is accom-

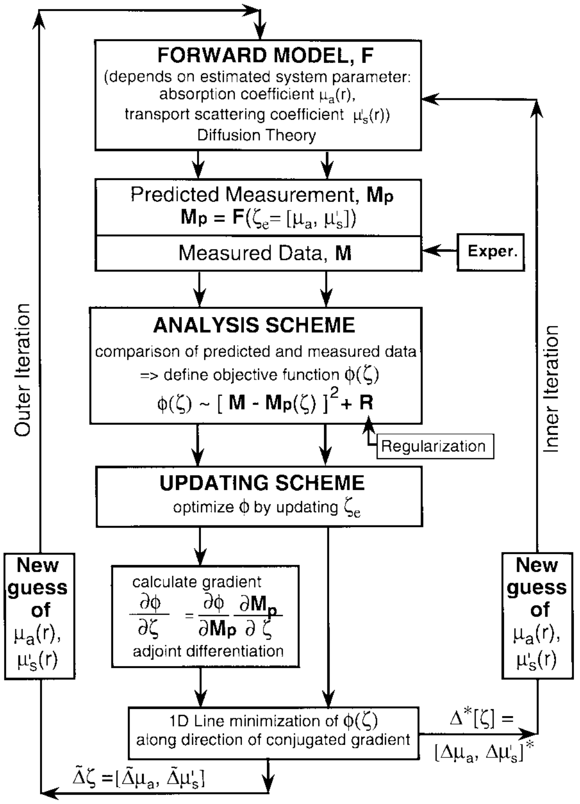


Fig. 2. Flow diagram of GIIR scheme used in this work.

plished in two substeps.

a) First, the gradient of the objective function $d\phi(\zeta)/d\zeta$ is calculated by means of so-called adjoint differentiation. Note that this gradient is not identical to the weight function or Jacobian used in the perturbation approach. The effective calculation of this gradient poses a major challenge in itself and is described in detail later.
b) Second, given the gradient an iterative line minimization in the direction of the gradient is performed. This step is labeled inner iteration in Fig. 2 and consists of several forward calculations in which the optical parameters ζ are varied. Once the minimum along the line is found, a new gradient is calculated at this minimum (outer iteration) and another line minimization is performed, now along a different direction in the ζ space. These steps are repeated until a distribution ζ is found for which $\phi(\zeta)$ is smallest.

The inner iteration in the perturbation approach (Fig. 1) consists in solving (4) which contains the large full ill-conditioned Jacobian matrix and is usually very time consuming. During the inner iteration of the perturbation approach the parameters ζ are fixed, unlike in the line-minimization scheme (Fig. 2). While the outcome of the inner iteration in Fig. 1 is one update $\zeta + \Delta\zeta$, the inner iteration in Fig. 2 provides several updates $\zeta + \Delta\zeta_1 + \Delta\zeta_2 + \cdots$ before a new gradient needs to be calculated. The superiority of this method was most recently acknowledged by Arridge and Schweiger [64], who currently work on similar reconstruction schemes as presented in this paper.

In summary, the major difference between the perturbation approaches described in Section II-A and the proposed GIIR scheme is that the inverse problem is to be solved as an optimization problem, rather than a perturbation problem. While in both methods Jacobian matrixes are calculated, they are used in different ways in the reconstruction process. In the perturbation approach, the Jacobian matrix needs to be inverted to obtain one update $\Delta\zeta$ of the optical properties in the medium. This inversion, in itself, is an extremely time-consuming iterative process, which sometimes also is referred to as optimization. In the GIIR scheme, the Jacobian is part of the gradient calculation of the objective function. Once this gradient is found, a line minimization of the objective function along the direction of the gradient is performed to find the update for the optical properties. This line minimization does not require the computationally expensive inversion of the Jacobian matrix.

In the remainder of this paper we will discuss in detail the mathematical background and actual implementation of the time-resolved image reconstruction scheme, which is based on a finite-difference forward model of the diffusion equation. To illustrate the performance of that code, we will show simulated reconstruction examples of well-defined objects in a homogeneous background and a more clinically relevant reconstruction of brain tissue.

III. THEORY

A. Forward Model

As a forward model that describes the photon propagation in the turbid medium and predicts the measurements at the detector position $\boldsymbol{r_d}$ we use the 2- or 3-D time-dependent diffusion equation with zero-boundary condition. For notational simplicity we will first develop all concepts in two dimensions. The extensions to three dimensions is straightforward and will be outlined whenever appropriate. The time-dependent diffusion equation is given as

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial U}{\partial x}\right) + \frac{\partial}{\partial y}\left(D\frac{\partial U}{\partial y}\right) - c\mu_a U + S. \quad (5)$$

Here, $U = U(x, y, t)$ is the fluence rate [Wcm^{-2}] and $S = S(x, y, t)$ is the source strength at position (x, y) and time t. The position-dependent absorption and diffusion coefficients are denoted by $\mu_a = \mu_a(x, y)$ and $D = D(x, y)$, respectively. The speed of light in the medium is represented by c. The diffusion coefficient is defined as $D = c \cdot [3\mu_a + 3(1-g)\mu_s]^{-1}$ where g is the scattering anisotropy value equal to the average value of the cosine of the angle through which photons are scattered. The coefficient $\mu'_s = (1-g)\mu_s$ is called the reduced or transport scattering coefficient.

Equation (5) is solved by replacing the temporal and spatial derivatives by their finite-difference approximations as follows:

$$\frac{\partial U}{\partial t} \approx \frac{U^n - U^{n-1}}{\Delta t} \quad (6)$$

$$\frac{\partial}{\partial x}\left(D\frac{\partial U}{\partial x}\right) \approx \delta_x(U_{i,j}D_{i,j}) \tag{7a}$$

$$\frac{\partial}{\partial y}\left(D\frac{\partial U}{\partial y}\right) \approx \delta_y(U_{i,j}, D_{i,j}) \tag{7b}$$

where

$$\begin{aligned}\delta_x(U_{i,j}) \equiv {} & [D_{i+1/2,j}(U_{i+1,j} - U_{i,j}) \\ & - D_{i-1/2,j}(U_{i,j} - U_{i-1,j})]/\Delta x^2 \end{aligned} \tag{8a}$$

$$\begin{aligned}\delta_y(U_{i,j}) \equiv {} & [D_{i,j+1/2}(U_{i,j+1} - U_{i,j}0 \\ & - D_{i,j-1/2}(U_{i,j} - U_{i,j-1})]/\Delta y^2. \end{aligned} \tag{8b}$$

The distances between spatial and temporal mesh points are given by $\Delta x, \Delta y$, and Δt, respectively. The approximations are obtained by simply differencing the partial derivatives as it stands and centering everything appropriately. The interpolated value of D halfway between the grid points is gotten by linear interpolation $D_{i+1/2,j} = (D_{i,j} + D_{i+1,j})/2$.

By substituting the finite-difference approximations (7) and (8) in the diffusion equation (5), we obtain a difference equation that needs to be solved forward in time. The finite-difference approximations to the spatial derivatives (7) can be evaluated at time index $n+1$ or n when we are solving the difference equation for U^{n+1}. Methods that evaluate the spatial derivatives at the past time instance (n) are called explicit, while methods that evaluate the spatial derivatives at the present time instance $(n+1)$ are called implicit. Explicit methods lead to simple and fast solution of the finite-difference equation, however, they are only conditionally stable. Implicit methods prove to be unconditionally stable, but lead to more complex solution schemes.

To combine the simplicity of explicit methods with the stability of fully implicit methods, we use in this work a technique known as alternating-directions implicit (ADI) method [65]. In this method, the computation of U^{n+1} from U^n is broken up into two time steps. In the first half-time step, the spatial derivative in only one direction is evaluated at the present time instance (implicit) and the other spatial derivatives are evaluated at the previous time instance (explicit). In the next half-time step, the implicit and explicit directions are switched, etc. The difference equations for the two half-time steps are given as

$$\begin{aligned}&\left(1 + c\mu_a\frac{\Delta t}{2}\right)U_{i,j}^{n+1/2} - \frac{\Delta t}{2}\delta_x(U_{i,j}^{n+1/2}) \\ &\qquad = \frac{\Delta t}{2}\delta_y(U_{i,j}^n) + U_{i,j}^n + \frac{\Delta t}{4}(S_{i,j}^{n+1} + S_{i,j}^n)\end{aligned} \tag{9a}$$

$$\begin{aligned}&\left(1 + c\mu_c\frac{\Delta t}{2}\right)U_{i,j}^{n+1} - \frac{\Delta t}{2}\delta_y(U_{i,j}^{n+1}) \\ &\qquad = \frac{\Delta t}{2}\delta_x(U_{i,j}^{n+1/2}) + U_{i,j}^{n+1/2} + \frac{\Delta t}{4}(S_{i,j}^{n+1} + S_{i,j}^n).\end{aligned} \tag{9b}$$

We call the vector U row ordered if the elements of the vector U are ordered such that we first enumerate the x values before increasing the y value $([U_{x1,y1}, U_{x2,y1}, \cdots, U_{xn,y1}, U_{x1,y2}, U_{x2,y2}, \cdots])$, and column ordered if we first enumerate the y values before increasing the x values $([U_{x1,y1}, U_{x1,y2}, \cdots, U_{x,yn}, U_{x2,y1}, U_{x2,y2}, \cdots])$. If we perform row ordering in the first half step [(9a)], column-ordering in the second half step [(9b)], and letting n take on fractional values, we can compactly represent (9a) and (9b) as

$$AU^{n+1/2} = BU^n + S^{n+1/4} \tag{10a}$$

or

$$U^{n+1/2} = A^{-1}BU^n + A^{-1}S^{n+1.4}. \tag{10b}$$

The structures of the matrixes A and B remain the same in both half-time steps, but the absolute values of the matrix elements differ. The advantage of the ADI method is that by correct ordering the matrix, A can always be made tridiagonal, e.g., for the first half step

$$A_{ii} = 1 + c\mu_a\frac{\Delta t}{2} + (D_{i+1/2,j} + D_{i-1/2,j})\frac{\Delta t}{2\Delta x} \tag{11a}$$

$$A_{i,i\pm1} = -D_{i\pm1/2,j}\frac{\Delta t}{2\Delta x}. \tag{11b}$$

The matrix A needs to be inverted to compute $U^{n+1/2}$ from U^n (10b). Since A is always tridiagonal, the inversion can easily be done in $O(N)$ computations [66]. Furthermore, the ADI method is unconditionally stable for any value of Δt and is acccurate to $O[(\Delta t)^2] + O[(\Delta x)^2]$ [65]. From now on we do not distinguish between the two half-time steps for notational simplicity. For further information on explicit, implicit, and ADI methods see [63], [65], [67], and [68].

The generalization of the ADI method to three dimensions is straightforward. Equations (9a) and (9b) have to be replaced by three equations containing, respectively, only U^n and $U^{n+1/3}, U^{n+1/3}$ and $U^{n+2/3}$ and $U^{n+2/3}$ and U^{n+1}. These three equations are identical, except for a cyclic shifting of the implicit term among the x, y, and z derivatives. The major difference between the 2- and 3-D ADI method is that, in three dimensions, the unconditional stability is lost and is replaced by the stability condition $\Delta t < [1.5 \cdot \Delta x / D_{i,j}]$. Furthermore, the accuracy drops to $O[(\Delta t)] + O[(\Delta x)^2]$ [65].

B. Analysis Scheme

The forward model is used to calculate detector responses at a set of detector position M. Let Y denote the measurements of the diffuse intensity U for all $s \in M$. We assume, for simplicity, that the measurements are corrupted by uncorrelated Gaussian noise. However, note that the method we propose is not limited to this choice and could just as well have chosen a more complex model such as Poisson noise. In this work the objective function is defined as

$$\Phi(\zeta) = \sum_{s \in M}\sum_{n}\frac{(Y_s^n - U_s^n(\zeta))^2}{2\sigma(s,n)^2} + R(\zeta). \tag{12}$$

The first sum over s indicates a summation over all source-detector pairs, while the second sum over n represents the summation over all time steps. The parameter $\sigma(s,n)$ is the time and spatially varying noise variance. The function $R(\zeta)$ describes any additional penalty or regularization term, which generally depends on the spatial distribution of optical properties $\zeta = [c\mu_a(r), D(r)]$. An example would be a Markov random field regularizer [69], defined as

$$R(\zeta) = \frac{\lambda}{p} \sum_{(m,r)\in N} \left| \frac{\zeta_m - \zeta_r}{\sigma_\zeta(m)} \right|^p. \tag{13}$$

The parameter λ is a weighting factor for the regularization term, often also referred to as hyperparameter [70]. The regularization term compares ζ at position r with ζ at all neighboring positions m. Large variations between neighboring ζ are more or less penalized, depending on the choice of p and the scaling parameter σ_ζ. In this work we used $\lambda = 1, p = 1.1$ and $\sigma_\zeta = 1$. This regularization term imposes a certain smoothness on the solution. A more detailed discussion on Markov random fields, as well as the interpretaion of (12) and (13) in a Bayesian framework, can be found elsewhere [63], [69]. The goal of the reconstruction algorithm is now to minimize this objective function with respect to the system parameters $D(\boldsymbol{r})$ and $c\mu_a(\boldsymbol{r})$.

C. Gradient Calculation

The effective solution to optimization problems involving many variables (here the spatial distribution of optical properties) relies on knowing the gradient of the objective function with respect to the variables $d\phi/d\zeta$. In this work, we employ the method of adjoint differentiation [71]–[73], sometimes also called reverse differentiation [74]. To outline this method we assume, for simplicity, that no regularization term is used, i.e., $R(\zeta) = 0$. Furthermore, we choose the following notation: $p, q, r, d \in \Omega$ are grid points; $d\partial \in \Omega \subset \Omega$ are detector positions on the boundary $\partial\Omega$; and M denote measurements on the boundary $\partial\Omega$. The derivative of the objective function (8) with respect to the optical properties ζ is given by

$$\frac{d\Phi}{d\zeta_r} = \sum_{n\in T} \sum_{p\in\Omega} \frac{d\Phi}{dU_p^n} \frac{\partial U_p^n}{\partial \zeta_r}. \tag{14}$$

Here, we simply applied the chain rule and the first term in the sum is the outer derivative, while the second term is the inner derivative of (12). The second term is easily obtained by differentiating (10a) at time $[n]$ with respect to the optical parameters, which yields

$$\frac{dA}{d\zeta_r} U^n + A \frac{\partial U^n}{\partial \zeta_r} = \frac{dB}{d\zeta_r} U^{n-1/2} \tag{15a}$$

$$\Rightarrow \frac{\partial U^n}{\partial \zeta_r} = A^{-1} \left(\frac{dB}{d\zeta_r} U^{n-1/2} - \frac{dA}{d\zeta_r} U^n \right). \tag{15b}$$

(Here we consider (10a) at times $[n - 1/2]$ and $[n]$ rather than $[n]$ and $[n + 1/2]$). Equation (15b) can be easily calculated by means of the known intensity U and the matrixes A and B of the forward model.

More involved is the calculation of the derivative of the objective function Φ with respect to the intensities U^n, which is the first term in (14). We obtain this term recursively by applying the chain rule and stepping backward (reverse) in time from $n + 1/2$ to n

$$\frac{d\Phi}{dU_p^n} = \sum_{q\in\Omega} \frac{d\Phi}{dU_q^{n+1/2}} \frac{dU_q^{n+1/2}}{dU_p^n} + \frac{\partial\Phi}{\partial U_p^n} \tag{16a}$$

with

$$\frac{d\Phi}{dU_q^N} = \frac{\partial\Phi}{\partial U_q^N} \quad \text{for } N = \max(n \in T). \tag{16b}$$

Here, $\partial\phi/U_p^n$ denotes the change in ϕ when only U_p^n is varied, keeping all other variables constant, while $d\phi/dU_p^n$ denotes the total change in ϕ when U_p^n is varied, as well as all variables that depend on U_p^n. Partially differentiating (12) with respect to U_p^n, we obtain

$$\frac{\partial\Phi}{\partial U_p^n} = \begin{cases} \frac{1}{\sigma}(M_p^n - U_p^n) & p = d \in \Omega_d \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

The term $dU_p^{n+1/2}/dU_p^n$ can again be calculated from the forward finite difference (6b) and we obtain

$$\frac{dU_q^{n+1/2}}{dU_p^n} = A^{-1} B I_p \quad \text{with} \quad I_p = \begin{cases} 1 & q = p \\ 0 & q \neq p \end{cases} \tag{18}$$

where I_p is a column vector that is zero everywhere except at the spatial point p, where it is unity.

To further enhance the efficient calculation of the gradient (14) we make use if the following matrix-multiplication properties. Consider the multiplication $(M_1 \cdot M_2 \cdot V)$ where M_1 and M_2 are $N \times N$ matrixes and V is a vector of length N. This multiplication can be done in two ways. Either a matrix-matrix multiplication is performed before a matrix-vector multiplication $(M_1 \cdot M_2) \cdot V$ or two matrix-vector multiplications are performed first, $M_1 \cdot (M_2 \cdot V)$. The first approach, $(M_1 \cdot M_2) \cdot V$, requires $N^3 + N^2$ scalar multiplications, while the second approach requires $2N^2$ scalar multiplications. Therefore, the second approach is $(N/2 + 1)$ times faster than the first approach.

How this method can be applied in our gradient calculation can be seen by inserting (16a) into (14)

$$\frac{d\Phi}{d\zeta_r} = \sum_{n\in T} \sum_{p\in\Omega} \left(\sum_{q\in\Omega} \frac{d\Phi}{dU_q^{n+1/2}} \frac{dU_q^{n+1/2}}{dU_p^n} + \frac{\partial\Phi}{\partial U_p^n} \right) \frac{\partial U_p^n}{\partial \zeta_r} \tag{19a}$$

pulling the last term into the bracket

$$\frac{d\Phi}{d\zeta_r} = \sum_{n\in T} \sum_{p\in\Omega} \sum_{q\in\Omega} \left(\frac{d\Phi}{dU_q^{n+1/2}} \frac{dU_q^{n+1/2}}{dU_p^n} \frac{\partial U_p^n}{\partial \zeta_r} + \frac{\partial\Phi}{\partial U_p^n} \frac{\partial U_p^n}{\partial \zeta_r} \right) \tag{19b}$$

and reordering the sums yields

$$\frac{d\Phi}{d\zeta_r} = \sum_{n\in T} \left(\sum_{p\in\Omega} \sum_{q\in\Omega} \frac{d\Phi}{dU_q^{n+1/2}} \frac{dU_q^{n+1/2}}{dU_p^n} \frac{\partial U_p^n}{\partial \zeta_r} \right) + \sum_{n\in T} \left(\sum_{p\in\Omega} \sum_{r\in\Omega} \frac{\partial\Phi}{\partial U_p^n} \frac{\partial U_p^n}{\partial \zeta_r} \right). \tag{19c}$$

The double sum in the first bracket of (19c) is just a multiplication of two matrixes ($[dU_q^{n+1}/dU_p^n]$ and $[dU_p^n; d\zeta_r]$), with a vector $(d\phi/dU_q^{n+1})$. The order of the summation can be done in two ways. If the sum over p is done before the sum over q, we effectively perform a matrix–matrix multiplication before the matrix-vector multiplication. Summing first over q and then over p corresponds to the case of two matrix-vector multiplication and, thus, can be performed $(N+1)/2$ faster than the first approach. This amounts to the adjoint differentiation scheme applied in this work.

The method of adjoint differentiation provides an efficient means to calculate the needed gradients for a complex sequence of calculations, such as the forward simulation of the migration of photons described above. Its power lies in its ability to calculate the derivatives with respect to all the variables in a CPU time that is comparable to one forward calculation.

D. Line Minimization

Once the gradient $d\phi/d\zeta$ for a point ζ_0 is obtained, a one-dimensional (1-D) line minimization along the given gradient direction is performed. In this method one moves along a direction given by the gradient until a minimum $\Phi(\zeta_m)$ is found. Various techniques can be applied to perform such 1-D minimization [75]. In this work we apply an iterative strategy by starting with a small step $\zeta_0 + \Delta\zeta = \zeta_1$ in the direction of the gradient, calculating the function values $\Phi(\zeta_1)$, taking a bigger step $\zeta_1 + \alpha\Delta\zeta = \zeta_2$, with $\alpha > 1$, in the gradient direction, calculating the function value $\Phi(\zeta_2)$, etc. until $\Phi(\zeta_v) < \Phi(\zeta_{v-1})$ and $\Phi(\zeta_v) < \Phi(\zeta_{v+1})$. At this point, a parabola is fitted to the three points and the minimum of the parabola is assumed to be a minimum along the direction of the gradient.

We typically chose $\Delta\zeta$ in a way that ensures that in each component $\Delta\mu_s' < 1$ and $\Delta\mu_a < 0.01$. Therefore, the changes of optical properties in each pixel are initially smaller than approximately 10% of typical values for biomedical tissues. The parameter α is set to 1.618, which is the ratio of the golden sections. Furthermore, we impose positivity constrains that limited the search to values $\zeta > 0$. For more detailed discussion of line-minimization schemes see, for example [75].

Once the minimum along the direction of the gradient is found, a new gradient calculation is performed and a new direction is chosen in a conjugate-gradient framework. That is, the new direction is determined by a weighted sum of the previous gradients. In this study we employ a Polak–Riberie conjugated-gradient scheme [75], [76].

IV. RESULTS

A. Problem Setup

The GIIR algorithm is tested by simulating time-resolved measurements of infrared light passing through a heterogeneous medium and using these simulated measurements as input to the reconstruction code. The measurements are simulated with the finite-difference, time-resolved diffusion code. To mimic real-world measurements, Gaussian noise is added to each measurement at a signal-to-noise (SNR) level of 30 db. This value is typical for time-resolved single-photon counting measurements [77], [78]. When single-photon counting is used, relatively noise free data can be obtained over 2–3 orders of magnitude. Signals that are approximately 1000 times smaller than the maximum are increasingly affected by statistical noise.

Once the detector readings are simulated, they are used as the actual measurement data Y in the reconstruction code. To start the reconstruction program an initial guess of the optical properties is necessary. If not stated otherwise, a medium with constant absorption and diffusion coefficients is used as a first guess. Based on this guess, the forward code calculates the detector responses U and compares these predicted measurements with the measurement data Y by calculating the objective function ϕ (12). The derivatives of ϕ with respect to ζ at all grid points are calculated and a new spatial distribution of D and $c\mu_a$ is put in place for the next forward calculation. The speed of light in the medium is set to $c = 22$ cm ns^{-1} in all cases. All calculations are done on an HP 9000/770 workstation.

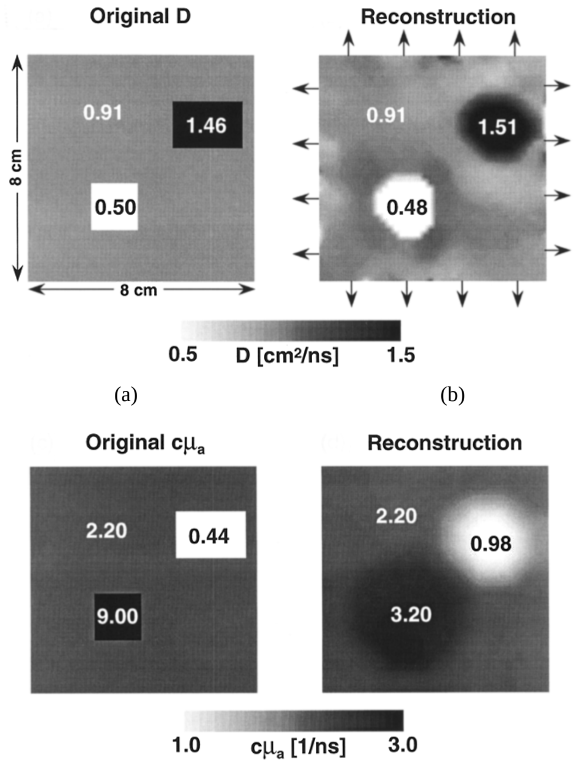
B. Simple Systems

First we consider a simple heterogeneous system [Fig. 3(a)] which consists of a background medium with ($D = 0.9$ cm^2 ns$^{-1}, c\mu_a = 2.2$ ns^{-1}) and two objects with ($D = 0.5$ cm^2 ns$^{-1}, c\mu_a = 9.0$ ns^{-1}) and ($D = 1.46$ cm^2 ns$^{-1}, c\mu_a = 0.44$ ns^{-1}), respectively. A 40 × 40 $x - y$ grid with a spatial resolution of 0.2 cm is used to simulate the 8 × 8-cm medium. The medium is surrounded by 16 source detector positions. Moving the source around the medium results in $16 \times 15 = 240$ detector readings. Each detector reading consists of 50 time points, which are equally spaced by $\Delta t = 0.1$ ns. In the reconstruction code, the derivatives of the objective function (12) with respect to $c\mu_a$ and D at each grid point are calculated, which results in $2 \times 40 \times 40 = 3200$ derivative calculations in each iteration step. Fig. 3(b) shows the reconstructed D image, which was obtained after 15 iterations (20 min). The initial guess is a homogeneous medium with $D = 1.0$ cm^2 ns^{-1} and $c\mu_a = 2.0$ ns^{-1}. The locations of the two objects and absolute values of D are reconstructed with high accuracy.

However the absolute values of $c\mu_a$ are not as accurate. For example, the small object with $c\mu_a = 9.0$ ns^{-1} appears as a larger object with $c\mu_a = 3.2$ ns^{-1}. Similar results have been observed by other researchers using different diffusion-based image algorithms [64], [79], [80]. Further studies are necessary to fully explore the reasons for this behavior.

The extension of the code to three dimensions is straight forward. An example of a 3-D reconstruction of a 8 × 8 × 8 cm heterogeneous medium is shown in Fig. 4. Sources and detectors were arranged in ten layers. Layers two and seven are shown in Fig. 4(b). Each layer has four sources, one centered on each side of the cube, and 20 detectors, four on each side and one on each corner. A medium with constant absorption ($\mu_a = 0.12$ cm^{-1}) and diffusion coefficients ($D = 1.0$ cm^{-1}) was used as a first guess.

We find that the time it takes for the GIIR code to complete a reconstruction on a given problem is proportional to the number of grid points $N_{\rm grid}$ and number of sources $N_{\rm source}$.



(a) (b)

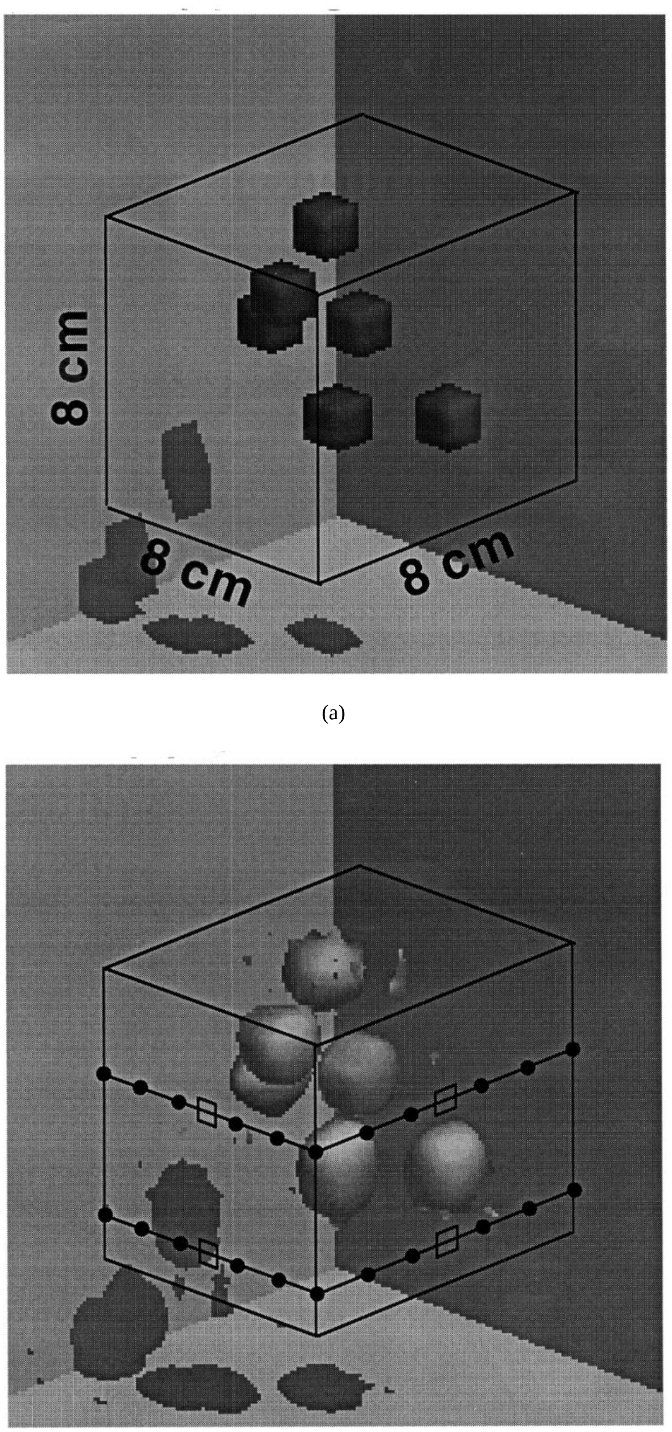
(c) (d)

Fig. 3. (a) and (c) The original medium that contains two inhomogeneities, which differ in the diffusion and absorption coefficient. For the reconstruction the system is discretized into $40 \times 40 = 1600$ voxels. (b) and (d) Here the reconstructed medium after 15 iterations (20 min) is shown for both optical parameters. The arrows indicate the 16 source/detector positions surrounding the medium.

Doubling N_{grid} or N_{source} leads to a doubling of the reconstruction time. Doubling the number of detectors increases the computation time only by a factor of 1.1. Furthermore, we observe that keeping $N_{\text{source}}, N_{\text{grid}}$, and N_{det}, constant, but increasing the dimension of the problem from 2-D → 3-D, leads to an increase of the computation time by a factor of 1.5, which can be explained by the additional overhead in the implementation of the 3-D code.

C. Brain Imaging

To test the algorithms in a situation closer to a clinical problem, MRI density maps are used to generate optical property maps of a brain. MRI imaging techniques allow one to distinguish between skull, white matter, gray matter and cerebrospinal-fluid-filled spaces in the head. These different tissues appear in MRI scans with different densities. From this an optical property map $(c\mu_a(\boldsymbol{r}), D(\boldsymbol{r}))$ is obtained by assigning different optical properties to different density values [81]. Fig. 5(a) shows such a segmented scan for a slice through the head. Clearly identifiable features include the cerebrospinal-fluid-filled ventricle in the center of the brain and a hematoma near the forehead. Furthermore, the brain is surrounded by cerebrospinal fluid, as can be seen in the light areas in Fig. 5(a). Fig. 5(b) shows the reconstruction, based on simulated time-resolved measurements, for 12 source/detector positions. In this case, 60 iterations are performed, which take about 70 min. While fine structures are not resolved, clearly visible in the reconstruction are the ventricles and other areas filled with cerebrospinal fluid. In addition, the hematoma at the forehead can be seen.



(a)

(b)

Fig. 4. (a) and (b) Original and reconstruction of six 1.6 × 1.6 × 1.6 cm^3 cubes ($D = 0.45$ cm^2 ns^{-1}) in 8 × 8 × 8 cm^3 background medium ($D = 0.9$ cm^2 ns^{-1}). The absorption is constant $\mu_a = 0.1$ cm^{-1}. The system is discretized into $20 \times 20 \times 20 = 8000$ voxels. The six heterogeneities are 4 × 4 × 4 voxels in size. For two sides of the cube the source detector locations are shown for two out of ten measurement layers. Each layer has four sources (open squares) and 20 detectors (full circles). Displayed are volumes for which $D < 0.76$ cm^2 ns^{-1}.

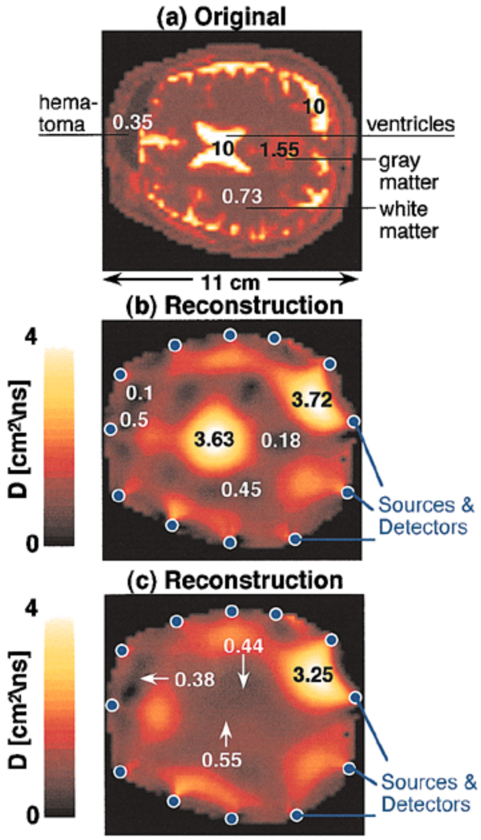


Fig. 5. (a)–(c) Reconstruction of 2-D slice through the brain based on simulated time-resolved measurements (75 time steps with $\Delta t = 0.1$ ns). The reconstruction was started with a homogenous initial guess ($D = 1.0$ cm^2 ns^{-1}). The numbers indicate the reconstructed diffusion coefficient D [cm^2 ns^{-1}] at various positions. The spatial resolution is 0.2 cm. The circles in the bottom image indicate source and detector positions.

Fig. 5(c) shows the same reconstruction, only this time the ventricles in the center of the brain are filled with blood. Ventricular bleeding is a major problem in preterm neonatal care. The incident of intraventricular hemorrhage among preterm infants ranges from 15–30% and often leads to severe brain damage or death of the infant [82], [83]. The reconstructions shown in Fig. 5(b) and (c) differ significantly and suggest that optical imaging can be used to detect ventricular bleeding. This would allow physicians to intervene promptly and effectively treat the problem.

V. CONCLUSION

We have developed a GIIR algorithm for time-resolved photon migration tomography. This algorithm does not require small perturbation or advance knowledge of a reference medium. The algorithm consists of three components.

1) A forward model that is used to predict the measurements assuming a certain distribution of optical properties. The forward model is currently based on a finite-difference formulation of the time-resolved diffusion equation.
2) An analysis scheme in which an objective function is defined which needs to be minimized.
3) An updating scheme, which allows to subsequent guess the optical properties of the medium based on the gradient of the objective function. The gradient is calculated with an adjoint differentiation technique.

The code was successfully used to reconstruct simple 2-D and 3-D heterogeneous media, which contain well-defined objects in a homogeneous background. The code was also tested on MRI-generated optical property maps of the brain.

Further studies will be necessary to fully explore all possibilities and problems connected with GIIR-type algorithm for optical tomography. Open questions are, for example, what type of regularization terms and hyperparameters are optimal, how can cross talk between absorption and diffusion reconstruction be addressed, how sensitive is the algorithm to the choice of the initial guess, and how can global maxima be found most efficiently.

Furthermore, experimental validation of the findings will be necessary. Since only numerical simulations were performed, the reconstructions shown do not provide an estimate of the model error. In other words, the results obtained in this study are based on the assumption that the diffusion model correctly describes the light propagation in the media. However, our group and others have shown that the diffusion model fails to accurately describe the propagation of light in highly absorbing regions, such as hematoma and void-like low scattering and low absorbing regions, such as the ventricles in the brain [81], [84]. How model errors effect the reconstruction results is currently under investigation and first results have been reported elsewhere [85].

ACKNOWLEDGMENT

The authors would like to thank Dr. S. S. Saquib, of Purdue University, for providing many ideas in implementing a prototype algorithm and frequent fruitful discussions. Furthermore, they would like to thank Dr. J. George, of Los Alamos National Laboratory, for his generous supply of computational time on HP workstations.

REFERENCES

[1] B. Chance and R. R. Alfano, Eds., Optical tomography and spectroscopy of tissue: Theory, instrumentation, model, and human studies. *Proc. SPIE-Int. Soc. Opt. Eng.*, no. 2979, 1997.
[2] R. R. Alfano and J. G. Fujimoto, Eds., *Advances in Optical Imaging and Photon Migrations*, *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 2, 1996.
[3] B. Chance and R. R. Alfano, Eds., Optical tomography, photon migration and spectroscopy of tissue and model media: theory, human studies, and instrumentation, *Proc. SPIE-Int. Soc. Opt. Eng.*, no. 2389, pts. 1 and 2, 1995.
[4] A. Yodh and B. Chance, "Spectroscopy and Imaging with diffusing light," *Physics Today*, vol. 48, pp. 34–40, 1995.
[5] G. Müller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, *Medical Optical Tomography: Functional Imaging and Monitoring* (SPIE Institutes for Advanced Optical Technologies Series). Bellingham, WA: SPIE, 1993, vol. IS11.
[6] K. Wells, J. C. Hebden. F. E. W. Schmidt, and D. T. Delpy, "The UCL multichannel time-resolved system for optical tomography," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 2979, pp. 599–607, 1997.
[7] H. Jess, H. Erdl, K. T. Moesta, S. Fantini, M. A. Franceschini, E. Gratton, and M. Kaschke, "Intensity modulated breast imaging: Technology and clinical pilot study results," in *Advances in Optical Imaging and Photon Migration* (OSA Trends in Optics and Phonics

Series), R. R. Alfano and J. G. Fujimoto, Eds. Washington, DC, Optical Society of America, 1996, vol. II, pp. 126–129.

[8] J. P. Vanhouten, D. A. Benaron, S. Spilman, and D. K. Stevenson, "Imaging brain injury using time-resolved near-infrared light scanning," *Pediatric Research*, vol. 39, pp. 470–476, 1996.

[9] M. Miwa and Y. Ueda, "Development of time-resolved spectroscopy system for quantitative noninvasive tissue measurement,"*Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 2389, pp. 142–149, 1995.

[10] S. Fantini, M. A. Franceschini, J. S. Maier, S. A. Walker, B. Barbieri, and E. Gratton, "Frequency domain multichannel optical detector for noninvasive tissue spectroscopy and oximetry," *Opt. Eng.*, vol. 34, pp. 32–42, 1995.

[11] P. J. Kirkpatrick, "Use of near-infrared spectroscopy in the adult," *Philos. Trans. R. Soc. London B, Biologi. Sci.*, vol. 352, pp. 701–705, 1997.

[12] C. Hock, K. Villringer, F. Muller-Spahn, R. Wenzel, H. Heekeren, S. Schuh-Hofer, M. Hofmann, S. Minoshima, M. Schwaiger, U. Dirnagl, and A. Villringer, "Decrease in parietal cerebral hemoglobin oxygenation during performance of a verbal fluency task in patients with Alzheimer's disease monitored by means of near-infrared spectroscopy (NIRS)-correlation with simultaneous CBF-PET measurements," *Brain Res.*, vol. 755, pp. 293–303, 1997.

[13] G. Buunk, J. G. van der Hoeven, and A. E. Meinders, "A comparison of near-infrared spectroscopy and jugular bulb oximetry in comatose patients resuscitated from a cardiac arrest," *Anaesthesia*, vol. 53, pp. 13–19, 1998.

[14] L. C. Henson, C. Calalang, J. A. Temp, and D. S. Ward, "Accuracy of a cerebral oximeter in healthy volunteers under conditions of isocapnic hypoxia," *Anesthesiology*, vol. 88, pp. 58–65, 1998.

[15] R. A. De Blasi, S. Fantini, M. A. Franceschini, M. Ferrari, and E. Gratton, "Cerebral and muscle oxygen saturation measurement by frequency-domain near-infrared spectrometer," *Med. Biol. Eng. Comput.*, vol. 33, pp. 228–230, 1995.

[16] W. J. Levin, S. Levin, and B. Chance, "Near-infrared measurement of cerebral oxygenation: Correlation with electroencephalographic ischemia during ventricular fibrillation," *Anesthesiology*, vol. 83, pp. 738–746, 1995.

[17] T. Noriyuki, H. Ohdan, S. Yoshioka, Y. Miyata, T. Asahara, and K. Dohi, "Near-infrared spectroscopic method for assessing the tissue oxygenation state of living lung," *Amer. J. Respir. Critical Care Med.*, vol. 156, pp. 1656–1661, 1997.

[18] K. J. Sapire, S. P. Gopinath, G. Farhat, D. R. Thakar, A. Gabrielli, J. W. Jones, C. S. Robertson, and B. Chance, "Cerebral oxygenation during warming after cardiopulmonary bypass," *Critical Care Med.*, vol. 25, pp. 1655–1662, 1997.

[19] D. A. Benaron, J. P. Vanhouten, W. Cheong, E. L. Kermit, and R. A. King, "Early clinical results of time-of-flight optical tomography in a neonatal intensive care unit," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 2389, pp. 582–596, 1995.

[20] W. F. Cheong, J. P. Vanhouten, E. L. Kermit, T. R. Machold, D. K. Stevenson, and D. A. Benaron, "Pilot comparison of light-based optical tomography versus ultrasound for real-time imaging of neonatal intraventricular hemorrhage," *Pediatric Res.*, vol. 39, pt. 2, p. 201A, 1996.

[21] S. P. Gopinath, C. S. Robertson, C. F. Contant, R. K. Narayan, R. G. Grossman, and B. Chance, "Early detection of delayed traumatic intracranial hematomas using near-infrared spectroscopy," *J. Neurosurgery*, vol. 83, pp. 438–444, 1995.

[22] G. Gratton, M. Fabiani, P. M. Corballis, D. C. Hood, M. R. Goodman-Wood, J. Hirsch, K. Kim, D. Friedman, and E. Gratton, "Fast and localized event-related optical signals (EROS) in the human occipital cortex: comparisons with the visual evoked potential and fMRI," *Neuroimage*, vol. 6, pp. 168–180, 1997.

[23] A. Maki, Y. Yamashita, E. Watanabe, and H. Koizumi, "Visualizing human motor activity by using noninvasive optical topography," *Frontiers Med. Biol. Eng.*, vol. 7, pp. 285–297, 1996.

[24] C. Hirth, H. Obrig, J. Valdueza, U. Dirnagl, and A. Villringer, "Simultaneous assessment of cerebral oxygenation and hemodynamics during a motor task. A combined near-infrared and transcranial Doppler sonography study," *Adv. Experimental Med. Biol.*, vol. 411, pp. 461–469, 1997.

[25] A. Maki, Y. Yamashita, Y. Ito, E. Watanabe, Y. Mayanagi, and H. Koizumi, "Spatial and temporal analysis of human motor activity using noninvasive NIR topography," *Med. Phys.*, vol. 22, pp. 1997–2005, 1995.

[26] R. Wenzel, H. Obrig, J. Ruben, K. Villringer, A. Thiel, J. Bernardinger, U. Dirnagl, And A. Villringer, "Cerebral blood oxygenation changes induced by visual stimulation in humans," *J. Biomed. Opt.*, vol. 4, pp. 399–404, 1996.

[27] C. Hock, K. Villringer, F. Muller-Spahn, M. Hofmann, S. Schuh-Hofer, H. Heekeren, R. Wenzel, U. Dirnagl, and A. Villringer, "Near infrared spectroscopy in the diagnosis of Alzheimer's disease," *Ann. New York Acad. Sci.*, vol. 777, pp. 22–29, 1996.

[28] A. J. Fallgatter, M. Roesler, L. Sitzmann, A. Heidrich, T. J. Mueller, and W. K. Strik, "Loss of functional hemispheric asymmetry in Alzheimer's dementia assessed with near-infrared spectroscopy," *Brain Res., Cognitive Brain Res.*, vol. 6, pp. 67–72, 1997.

[29] B. Devaraj, Y. Watanabe, M. Takeda, M. Usa, T. Yuasa, T. Akatsuka, and H. Inaba, "A novel application of coherent detection imaging methods for diagnosis of rheumatoid arthritis in the near-infrared region," in *Advances in Optical Imaging and Photon Migration* (OSA Trends in Optics and Photonics Series), J. G. Fujimoto and M. S. Patterson, Eds. Washington, DC: Optical Society of America, 1998, vol. 21, pp. 338–340.

[30] V. Prapavat, W. Runge, J. Mans, A. Krause, J. Beuthan, and G. Müller, "The development of a finger joint phantom for the optical simulation of early inflammatory rheumatic changes," *Biomedizinische Technik*, vol. 42, pp. 319–326, 1997.

[31] A. Klose, V. Prapavat, O. Minet, J. Beuthan, and G. Müller, "RA diagnostics applying optical tomography in frequency-domain," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 3196, pp. 194–204, 1997.

[32] O. Jarlman, R. Berg, S. Andersson-Engels, S. Svanberg, and H. Pettersson, "Time-resolved white light transillumination for optical imaging," *Acta Radiologica*, vol. 38, pp. 185–189, 1997.

[33] S. Nioka, Y. Yung, M. Shnall, S. Zhao, S. Orel, C. Xie, B. Chance, and L. Solin, "Optical imaging of breast tumor by means of continuous waves," *Adv. Experimental Med. Biol.*, vol. 411, pp. 227–232, 1997.

[34] B. J. Tromberg, O. Coquoz, J. B. Fishkin, T. Pham, E. R. Anderson, J. Butler, M. Cahn, J. D. Gross, V. Venugopalan, and D. Pham, "Noninvasive measurements of breast tissue optical properties using frequency-domain photon migration," *Philos. Trans. R. Soc. London B, Biol. Sci.*, vol. 352, pp. 661–668, 1997.

[35] R. R. Alfano, S. G. Demos, and S. K. Gayen, "Advances in optical imaging of biomedical media," *Ann. New York Acad. Sci.*, vol. 820, pp. 248–270, 1997.

[36] M. A. Franceschini, K. T. Moesta, S. Fantini, G. Gaida, E. Gratton, H. Jess, W. W. Mantulin, M. Seeber, P. M. Schlag, and M. Kaschke, "Frequency-domain techniques enhance optical mammography: initial clinical results," *Proc. Nat. Acad. Sci. U. S. A.*, vol. 94, pp. 6468–6473, 1997.

[37] S. Fantini, M. A. Franceschini, G. Gaida, E. Gratton, H. Jess, W. W. Mantulin, K. T. Moesta, P. M. Schlag, and M. Kaschke, "Frequency-domain optical mammography: Edge effect corrections," *Medical Physics*, vol. 23, pp. 149–157, 1996.

[38] J. A. Parker, *Image Reconstruction in Radiology*. Boca Raton, FL: CRC, 1990.

[39] D. A. Benaron, D. C. Ho, S. Spilman, J. P. van Houten, and D. K. Stevenson, "Non-recursive linear algorithms for optical imaging in diffusive media," in *Advances in Experimental Medicine & Biology: Oxygen Transport to Tissue XVI*. New York: Plenum, 1994, pp. 215–222.

[40] S. A. Walker, S. Fantini, and E. Gratton, "Image reconstruction by backprojection from frequency domain optical measurements in highly scattering media," *Appl. Opt.*, vol. 36, pp. 170–179, 1997.

[41] S. B. Colak, D. G. Papaioannou, G. W. 't Hooft, M. B. van der Mark, H. Schomberg, J. C. J. Paasschens, J. B. M. Melissen, and N. A. A. J. van Asten, "Tomographic image reconstruction from optical projections in light-diffusing media," *App. Opt.*, vol. 36, pp. 180–213, 1997.

[42] S. R. Arridge, "The Forward and inverse problems in time-resolved infrared imaging," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. IS11, pp. 35–64, 1993.

[43] ——, "Photon-measurement density functions. Part I: Analytical forms," *Appl. Opt.*, vol. 34, pp. 7395–7409, 1995.

[44] S. R. Arridge and M. Schweiger, "Photon measurement density functions. Part 2: Finite-element-method calculation," *Appl. Opt.*, vol. 34, pp. 8026–8037, 1995.

[45] M. Schweiger and S. R. Arridge, " A system for solving the forward and inverse problems in optical spectroscopy and imaging," in *Advances in Optical Imaging and Photon Migrations* (OSA Trends in Optics and Photonics Series), R. R. Alfano and J. G. Fujimoto, Eds. Washington, DC, Optical Society of America, 1996, vol. 2, pp. 263–268.

[46] R. L. Barbour, H. L. Graber, Y. Wang, J. H. Chang, and R. Aronson, "A perturbation approach for optical diffusion tomography using continous-wave and time-resolved data," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. IS11, pp. 87–120, 1993.

[47] H .L. Graber, J. Chang, R. Aronson, and R. L. Barbour, "A perturbation model for imaging in dense scattering media: Derivation and evaluation

of imaging operators," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. IS11, pp. 121–143, 1993.

[48] R. L. Barbour, H. L. Graber, J. W. Chang, S. L. S. Barbour, P. C. Koo, and R. Aronson, "MRI-guided optical tomography: Prospects and computation for a new imaging method," *IEEE Trans. Comput. Sci. Eng.*, vol. 2, pp. 63–77, 1995.

[49] Y. Q. Yao, Y. Wang, Y. L. Pei, W. W. Zhu, and R. L. Barbour, "Frequency-domain optical imaging of absorption and scattering distrinbutions by Born iterative method," *J. Opt. Soc. Amer. A*, vol. 14, pp. 325–342, 1997.

[50] K. D. Paulsen and H. Jiang, "Spatially varying optical property reconstruction using a finite element diffusion equation approximation," *Med. Phys.*, vol. 22, pp. 691–701, 1995.

[51] ———, "Enhanced frequency domain optical image reconstruction in tissues through total variation minimization," *Appl. Opt.*, vol. 35, pp. 3447–3458, 1996.

[52] H. Jiang, K. D. Paulsen, and U. L. Österberg, "Optical image reconstruction using DC data: Simulations and experiments," *Phys. Med. Biol.*, vol. 41, pp. 1483–1498, 1996.

[53] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Simultaneous reconstruction of optical absorption and scattering maps in turbid media from near-infrared frequency-domain data," *Opt. Lett.*, vol. 20, pp. 2128–2130, 1995.

[54] ———, "Optical image reconstruction using frequency domain data: Simulations and experiments," *J. Opt. Soc. Amer. A*, vol. 13, pp. 253–266, 1996.

[55] J. C. Schottland, J. C. Haselgrove, and J. S. Leigh, "Photon hitting density," *Appl. Opt.*, vol. 32, pp. 448–453, 1993.

[56] M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Experimental images of heterogeneous turbid media by frequency-domain diffusion-photon tomography," *Opt. Lett.*, vol. 20, pp. 426–428, 1995.

[57] D. Y. Paithankar, A. U. Chen, B. W. Pogue, M. S. Patterson, and E. M. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," *Appl. Opt.*, vol. 36, pp. 2260–2272, 1997.

[58] M. V. Klibanov, T. R. Lucas, and R. M. Frank, "A fast and accurate imaging algorithm in optical diffusion tomography," *Inverse Problems*, vol. 13, pp. 1341–1361, 1997.

[59] T. R. Lucas, M. V. Klibanov, and R. M. Frank, "Imaging experimental data from optical tomography by the elliptic system method," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 3171, pp. 22–33, 1997.

[60] S. R. Arridge and J. C. Hebden, "Optical imaging in medicine: II. Modeling and reconstruction," *Phys. Med. Biol.*, vol. 42, pp. 841–853, 1997.

[61] A. H. Hielscher, "Model-based iterative image reconstruction for photon migration tomography," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 3171, pp. 106–117, 1997.

[62] A. H. Hielscher, A. Klose, D. M. Catarious, Jr., and K. M. Hanson, "Tomographic imaging of biological tissue by time-resolved, model-based, iterative, image reconstruction," in *Advances in Optical Imaging and Photon Migration II* (OSA Trends in Optics and Photonics Series), R. R. Alfano and J. G. Fujimoto, Eds. Washington, DC, Optical Society of America, 1998, vol. 21, pp. 156–161.

[63] S. S. Saquib, K. M. Hanson, and G. S. Cunningham, "Model-based image reconstruction from time-resolved diffusion data," *Proc. SPIE-Int. Soc. Opt. Eng.*, vol. 3034, pp. 369–380, 1997.

[64] S. R. Arridge and M. Schweiger, "A gradient-based optimization scheme for optical tomography," *Optics Express*, vol. 2, no. 6, pp. 213–226, 1998.

[65] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*. New York: Wiley, 1967, pp. 211–215.

[66] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. New York: Cambridge Univ. Press, 1992, pp. 50–55.

[67] G. Birkhoff and R. E. Lynch, *Numerical Solution of Elliptical Problems*. Philadelphia, PA: SIAM, 1984, pp. 173–178.

[68] W. F. Ames, *Numerical Methods for Partial Differential Equations*. New York: Academic, 1977, pp. 148–157.

[69] S. S. Saquib, C. A. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," *IEEE Trans. Image Processing*, vol. 7, pp. 1029–1044, 1998.

[70] Z. Zhou, R. Leahy, and J. Qi, "Approximate maximum likelihood hyperparameter estimation for Gibbs priors," *IEEE Trans. Image Processing*, vol. 6, pp. 844–861, 1997.

[71] K. M. Hanson and G. S. Cunningham, "The Bayes inference engine," *Maximum Entropy and Bayesian Methods*, K. M. Hanson and R. N. Silver, Eds. Dordrecht, The Netherlands: Kluwer, 1996, pp. 125–134.

[72] A. Griewank and G. F. Corliss, Eds., *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*. Philadelphia, PA: SIAM, 1991.

[73] O. M. Alifanov, E. A. Artiukhin, and S. V. Rumiantsev, *Extreme Methods for Solving Ill-Posed Problems With Applications to Inverse Heat Transfer Problems* Begell House, 1995.

[74] A. J. Davies, D. B. Christianson, L. C. W. Dixon, R. Roy, and P. van der Zee, "Reverse differentiation and the inverse diffusion problem," *Adv. Eng. Software*, vol. 28, pp. 217–221, 1997.

[75] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. New York: Cambridge Univ. Press, 1992, ch. 10, pp. 394–455.

[76] D. A. H. Jacobs, Ed., *The State of the Art in Numerical Analysis*. London, U.K.: Academic, ch. III.

[77] D. V. O'Connor and D. Phillips, *Time-Correlated Single Photon Counting*. Orlando, FL: Academic, 1984.

[78] H. Liu, M. Miwa, B. Beauvoit, N. G. Wang, and B. Chance, "Characterization of absorption and scatetring properties of small volume biological samples using time-resolved spectroscopy," *Anal. Biochem.*, vol. 213, pp. 378–385, 1993.

[79] M. Schweiger and S. R. Arridge, "Direct calculation with a finite-element method of the Laplace transform of the distribution of photon time of flight in tissue," *Appl. Opt.*, vol. 36, pp. 9042–9049, 1997.

[80] M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Simultaneous scattering and absorption images of heterogeneous media using diffusive waves within the Rytov approximation," *Proc. SPIE-Int. Soc. Opt. Eng.*, pt. 1, pp. 320–327, 1995.

[81] A. H. Hielscher, R. E. Alcouffe, and R. L. Barbour, "Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues," *Phys. Med. Biol.*, vol. 43, pp. 1285–1302, 1998.

[82] L. R. Ment, M. Westerveld, R. Makuch, B. Vohr, and W. C. Allan, "Cognitive outcome at 4 1/2 years of very low birth weight infants enrolled in the multicenter indomethacin intraventricular hemorrhage prevention trial," *Pediatrics*, vol. 102, pp. 159–160, 1998.

[83] L. R. Ment, B. Vohr, W. Oh, D. T. Scott, W. C. Allan, M. D. Westerveld, C. Charles, R. A. Ehrenkranz, K. H. Katz, K. C. Schneider, and R. W. Makuch, "Neurodevelopmental outcome at 36 months' corrected age of preterm infants in the multicenter indomethacin intraventricular hemorrhage prevention trial," *Pediatrics*, vol. 98, no. 4, pp. 714–718, 1996.

[84] M. Firbank, S. R. Arridge, M. Schweiger, and D. T. Delpy, "An investigation of light transport through scattering bodies with nonscattering regions," *Phys. Med. Biol.*, vol. 41, pp. 767–783, 1996.

[85] A. D. Klose and A. H. Hielscher, "A transport-theory based reconstruction algorithm for optical tomography" *Proc. SPIE-Int. Soc. Opt. Eng.*, to be published.