Three-dimensional optical tomography with the equation of radiative transfer

Gassan S. Abdoulaev
Andreas H. Hielscher
Columbia University
Departments of Biomedical Engineering & Radiology
500 West 120th Street
ET351 Mudd Building, MC 8904
New York, New York 10027
E-mail: ahh2004@columbia.edu

Abstract. We report on the derivation and implementation of the first three-dimensional optical tomographic image reconstruction scheme that is based on the time-independent equation of radiative transfer (ERT) and allows for arbitrarily shaped medium boundaries and arbitrary spatial material distributions. The scheme builds on the concept of model-based iterative image reconstruction, in which a forward model provides prediction of detector readings, and a gradient-based updating scheme minimizes an appropriately defined objective function. The forward model is solved by using an even-parity formulation of the ERT, which lends itself to a finite-element discretization method. The finite-element technique provides the suitable framework for predicting light propagation in arbitrarily shaped three-dimensional media. For an efficient way of calculating the gradient of the objective function we have implemented an adjoint differentiation scheme. Initial reconstruction results using synthetic data from simple media and a three-dimensional mesh of the human forehead illustrate the performance of the code. © 2003 SPIE and IS&T. [DOI: 10.1117/1.1587730]

1 Introduction

Most of the currently available codes for optical tomography (OT) are based on the diffusion model of light propagation in tissue. While the diffusion model is a good model for many cases in biomedical imaging, several works have clearly established the shortcomings of the diffusion approximation. For example, the diffusion model is incapable of accurately describing light propagation in media that contain void-like regions with very low scattering and absorption coefficients. These types of media play an important role in several biomedical imaging applications. For example, in optical tomographic brain imaging one has to take into account the effects of the low-scattering and low-absorbing cerebrospinal fluid that surrounds the brain. The affect of this layer on light propagation has been the subject of many studies and discussions. Another example is the almost clear synovial fluid in joints, which plays an important role in the optical detection of rheumatoid arthritis. Other cases in which the diffusion model has potentially limited applicability are strongly absorbing media where the absorption coefficient $\mu_a$ is not much smaller than the scattering coefficient $\mu_s$, and optically thin media, in which the photon escapes the medium before it has been scattered sufficiently often.

Several groups have tried to develop image reconstruction codes that go beyond the diffusion approximation. These works typically employ the equation of radiative transfer (ERT) from which the diffusion equation can be derived assuming various approximations (see, e.g., Duderstadt and Martin or Case and Zweifel). Aronson et al. were the first to consider the full use of the equation of radiative transfer in OT. They used a Monte-Carlo solution of the adjoint transport equation and applied a back-projection method in order to obtain an image of the optical properties of a highly scattering medium. Dorn reported on a transport-backtransport method applied to the two-dimensional time-dependent equation of radiative transfer. Similarly, Klose et al. reported on a radiative-transfer-based reconstruction technique, which employed an upwind-difference, discrete-ordinates method applied to the two-dimensional time-dependent equation of radiative transfer.

Additional studies that consider the inverse problem with the ERT can be found outside the field of diffuse optical tomography. A majority of these investigations deal with plane-parallel media. For this geometry time-dependent and time-independent solutions have been provided for spatially homogeneous or nonhomogeneous media that have source and detectors either inside or outside the medium. Non-plane-parallel configurations were considered only by a few scientists. Larsen reports on a method to solve the time-independent inverse transport problem for optically thin media. Elliott found a solution for a weakly absorbing medium with a point source, but did not present actual reconstructions. Wang and Ueno establish a theoretical procedure for estimating the two-dimensional distribution of the ground albedo of the
earth by measurements made outside the atmosphere. A detailed discussion of all these methods has been given by McCormick. What is still missing is a three-dimensional image reconstruction scheme based on the ERT, which provides the spatial distribution of tissue optical properties for a generally heterogeneous medium of arbitrary shape, given external sources and detector readings obtained on the surface of the medium. In this work we present our first approach toward such an algorithm. The algorithm consists of a forward and inverse model that are used within a model-based iterative image reconstruction scheme (MOBIIR). First we will describe how the forward model is solved by using an even-parity approach, which reduces the time-independent radiative transfer equation to a system of partial differential equations of second order. These equations are discretized using a finite element method. Unlike finite-difference schemes this allows for more flexibility when arbitrary boundaries are considered. We then present in detail how the method of adjoint differentiation is used to solve the inverse problem. Examples using synthetic data illustrate the potential for three-dimensional optical tomographic imaging. A discussion of the advantages and limitations of the current code and an outlook concerning directions of future developments conclude this report.

2 Numerical Methods

2.1 Forward Problem

Within MOBIIR codes the forward model is used to predict the detector readings on the surface of the domain of interest, given the source location and a guess of the distribution of optical properties \( \mu_a(x) \) and \( \mu_s(x) \) inside the medium. In this work we consider the steady state ERT:

\[
\Omega \cdot \nabla \Psi(x, \Omega) + (\mu_a + \mu_s) \Psi(x, \Omega) = S(x, \Omega, \Omega') \Psi(x, \Omega') \mathrm{d}\Omega', \tag{1}
\]

as such a model. Here \( \Psi(x, \Omega) \) is the energy radiance in units of \( \text{W cm}^{-2} \text{s}^{-1} \), the source term \( S(x, \Omega, \Omega') \) is the density of the energy injected in the direction \( \Omega = (\omega_x, \omega_y, \omega_z) \), and the location \( x \in G \), and \( p(\Omega, \Omega') \) is the phase function equal to the probability that a photon with the direction \( \Omega' \) is scattered into the direction \( \Omega \). We assume the vacuum condition on the boundary of the domain:

\[
\Psi(x, \Omega) = 0, \quad \forall x \in \{ y \in \partial G : \Omega \cdot n_y < 0 \}, \tag{2}
\]

where \( n_y \) is an outward normal at the boundary point \( y \).

To bring this equation into a suitable form for a finite element scheme and to be able to later develop an efficient adjoint differentiation algorithm for the inverse problem, we employ an even-parity approach to solve Eq. (1). We start by noticing that since Eq. (1) is valid for all \( \Omega \), it also holds for \( -\Omega \):

\[
-\Omega \cdot \nabla \Psi(x, -\Omega) + (\mu_a + \mu_s) \Psi(x, -\Omega) = S(x, -\Omega, \Omega') \Psi(x, \Omega') \mathrm{d}\Omega'. \tag{3}
\]

Adding and subtracting Eqs. (1) and (3), we get a system of two coupled first-order integro-differential equations:

\[
\begin{align*}
\Omega \cdot \nabla \Psi^+(x, \Omega) + \sigma \Psi^+(x, \Omega) &= S^+(x, \Omega, \Omega') \Psi^+(x, \Omega') \mathrm{d}\Omega', \\
\Omega \cdot \nabla \Psi^-(x, \Omega) + \sigma \Psi^-(x, \Omega) &= S^-(x, \Omega, \Omega') \Psi^-(x, \Omega') \mathrm{d}\Omega',
\end{align*}
\]

where the superscripts ‘+’ and ‘−’ denote even and odd components of a function with respect to variable \( \Omega \), i.e.,

\[
F^\pm(x, \Omega) = \frac{F(x, \Omega) \pm F(x, -\Omega)}{2}, \tag{5}
\]

and \( \sigma = \mu_a + \mu_s \) is a total cross section. Note that \( p^+(\Omega, \Omega') \) and \( p^-(\Omega, \Omega') \) are actually functions of one independent variable \( y = (\Omega - \Omega')/2 \).

We can rewrite Eq. (4) in the following operator form:

\[
\begin{align*}
\Omega \cdot \nabla \Psi^+(x, \Omega) + H_+ \Psi^+(x, \Omega) &= S_+(x, \Omega), \\
\Omega \cdot \nabla \Psi^-(x, \Omega) + H_- \Psi^-(x, \Omega) &= S_-(x, \Omega),
\end{align*}
\]

where operators \( H_+ \) and \( H_- \) are defined as

\[
H_f(x, \Omega) = \sigma f(x, \Omega) - \mu_s \int_{4\pi} p^+(\Omega', \Omega') f(x, \Omega') \mathrm{d}\Omega'. \tag{7}
\]

One can easily exclude \( \Psi^- \) from Eq. (6), arriving at an equation for \( \Psi^+ \):

\[
\begin{align*}
-\Omega \cdot \nabla \{ H^{-1}_- [\Omega \cdot \nabla \Psi^+(x, \Omega)] \} + H_+ \Psi^+(x, \Omega) &= S^+(x, \Omega), \\
\Omega \cdot \nabla \Psi^+(x, \Omega) + H_- \Psi^+(x, \Omega) &= S^+(x, \Omega).
\end{align*}
\]

Using Eqs. (2), (5), and (6) we can show that the following condition for \( \Psi^+ \) holds on the boundary of the domain \( G \):

\[
H_- \Psi^+(x, \Omega) + \Omega \cdot \nabla \Psi^+(x, \Omega) = \frac{\partial}{\partial n_y} S^+(x, \Omega, \Omega), \quad \mbox{if} \quad \Omega \cdot n_y \leq 0. \tag{9}
\]

To solve Eq. (8) as is one has to invert the operator \( H_- \) explicitly. For instance, it is possible to take advantage of the expansion of the phase function \( p^-(\Omega, \Omega') \) in Legendre polynomials \( P_l(s) \):

\[
\begin{align*}
p^-(\Omega, \Omega') &= \sum_{l=0}^{\infty} \frac{(2l+1)}{4\pi} \alpha_l P_l(\Omega \cdot \Omega'), \tag{10}
\end{align*}
\]

where \( \alpha_l \) are the expansion coefficients. Then \( H^{-1}_- \) can be represented as a sum...
\[
H^{-1}f(x, \Omega) = \sum_{l=0}^{\infty} \left( \frac{2l+1}{4\pi} \right) \beta_l \int_{4\pi} P_l(\Omega \cdot \Omega') f(x, \Omega') d\Omega',
\]

with some coefficients \(\beta_l\).\(^{32,33}\) The representation above allows us to introduce a variational principle\(^{32-34}\) and use the finite element method to solve the problem numerically. This approach had been implemented in the EVENT radiation transport code.\(^{35}\)

In the case of isotropic scattering \([\rho = (4\pi)^{-1}]\) the inversion of \(H\) can be simplified: \(H^{-1}f = \sigma^{-1}f\). If the source is also isotropic, then Eq. (8) takes the following form:

\[
-\Omega \cdot \nabla (\sigma^{-1} \Omega \cdot \nabla \Psi^+) + \sigma \Psi^+ = S(x) + \frac{\mu_s}{4\pi} \int_{4\pi} \Psi^+(x, \Omega') d\Omega',
\]

where \(\Psi^+(x, \Omega) = [\Psi(x, \Omega) + \Psi(x, -\Omega)]/2\).

The vacuum boundary condition in Eq. (9) can now be rewritten as:\(^{36}\)

\[
\Psi^+(x, \Omega) = -\text{sign}(\Omega \cdot n) \sigma^{-1} \Omega \cdot \nabla \Psi^+(x, \Omega).
\]

The last integral in Eq. (12) can be computed approximately using a quadrature formula:

\[
\int_{4\pi} \Psi^+(x, \Omega) d\Omega = \sum_{l=1}^{L} W_l \Psi^+_l(x),
\]

where \(\Psi^+_l(x) = \Psi^+(x, \Omega_l)\). Here \(\{\Omega_l\}_{l=1, \ldots, L}\) is a (finite) set of three-dimensional vectors, such that \(|\Omega_l| = 1\), which are usually referred to as discrete ordinates, and parameters \(W_l\) are the weights of the quadrature formula. If we consider Eq. (12) for \(\Omega = \Omega_k\), \(k = 1, \ldots, L\), we arrive at a finite system of linear partial differential equations:

\[
-\Omega_k \cdot \nabla (\sigma^{-1} \Omega_k \cdot \nabla \Psi_k) + \sigma \Psi_k = S(x) + \frac{\mu_s}{4\pi} \sum_{l=1}^{L} W_l \Psi^+_l,
\]

where \(\Psi^+_k(x) = \Psi^+(x, \Omega_k)\).

The operator on the left-hand side of Eq. (14) is a self-adjoint for any fixed \(\Omega_k\). Numerical methods for solving this type of equation are well developed.\(^{37,38}\) We take advantage of the Galerkin finite-element method because of its flexibility in solving problems with complex geometries, as they are typically encountered in clinically relevant applications (e.g., brain, breast, or joint imaging).

As a result of the finite-element discretization of Eq. (14) we get a linear algebraic system

\[
A_k \psi_k = \hat{S} + B \sum_{l=1}^{L} W_l \psi_l, \quad k = 1, \ldots, L,
\]

where matrices \(A_k\) and \(B\) are symmetric positive-definite. The entries of \(A_k\), \(B\), and \(\hat{S}\) are defined as follows:

\[
(A_k)_{ij} = \left[ \left( \sigma^{-1} (\Omega_k \cdot \nabla \varphi_j)(\Omega_k \cdot \nabla \varphi_i) + \sigma \varphi_i \varphi_j \right) \right] dx,
\]

\[
B_{ij} = \left[ \frac{\mu_s}{4\pi} \varphi_j \varphi_i \right] dx, \quad \hat{S}_i = \left[ S(x) \varphi_i \right] dx,
\]

where \(\varphi_i\) is a basis function in the finite element method, corresponding to the mesh node \(i\). This system can be represented in a block-wise form:

\[
A \begin{pmatrix} \psi_0 \\ \cdots \\ \psi_L \end{pmatrix} = \begin{pmatrix} \hat{S} \\ \cdots \\ \hat{S} \end{pmatrix},
\]

\[
A = \begin{pmatrix} A_1 & W_2 B & \cdots & W_L B \\ -W_1 B & A_2 - W_2 B & \cdots & -W_L B \\ \vdots & \vdots & \ddots & \vdots \\ -W_1 B & -W_2 B & \cdots & A_L - W_L B \end{pmatrix}.
\]

The vectors \(\psi_k\), solving Eq. (15), have as their entries the nodal values of the approximate finite-element solutions of Eq. (14).

### 2.2 Inverse Problem

Our goal is to solve the inverse problem for Eq. (14), i.e., to find such optical properties \(\mu_a\) and \(\mu_s\) that would provide the best correlation between the measured and predicted data. Assume that we have \(n_s\) light sources and for each source there is a set of \(n_{s,d}\) detectors. Then we have \(\Sigma \mu_s\) source-detector pairs. Let \(M_{s,d}\) be a detector reading, corresponding to a source \(s\) and a detector \(d\). In this paper it is assumed that \(M_{s,d}\) is an outgoing fluence on the boundary of the domain. We denote \(\Psi^+_d\) as a solution of Eq. (14) for a source \(s\), and \(x_d\) as a location of a detector \(d\). Then an objective function can be defined as

\[
\Phi(M, \Psi) = \sum_{s,d} \left( \frac{\Sigma \left| W_s \Psi^+_s(x_d) - M_{s,d} \right|^2}{M_{s,d}} \right).
\]

Note that \(\Psi^+_d\) is summed over outward pointing ordinates and is considered as a function of optical properties, for instance, the absorption coefficient \(\mu_a\). Other more complex objective functions can and have been defined and studied.\(^{39}\) However, the influence of different objective functions on the reconstruction process is not the focus of this report, and we limit ourselves here to the least-squares error case.

In this work, we employ a gradient-based image reconstruction scheme as previously developed in our group.\(^{19,31}\) To find the gradient of the objective function we adapted an adjoint differentiation method presented by Davies and Christianson \textit{et al.} for the diffusion equation.\(^{40,41}\) This approach can be best understood as follows. Assume that we have a linear algebraic problem
\( Au = f, \)  \hspace{1cm} (18) 

where the entries of matrix A depend on the entries of some vector \( \sigma = \{ \sigma_k \}_{k=1, N} \): 

\[ A_{ij} = A_{ij}(\sigma). \]

Let \( \Phi(u) \) be a known function of \( u = \{ u_k \}_{k=1, N} \), and we want to find a gradient of \( \Phi \) with respect to \( \sigma \). We find that 

\[ \frac{\partial \Phi}{\partial \sigma_i} = \sum_{k=1}^{N} \frac{\partial \Phi}{\partial u_k} \frac{\partial u_k}{\partial \sigma_i} = \nabla_u \Phi \cdot \frac{\partial u}{\partial \sigma_i}. \]

Taking the partial derivative with respect to \( \sigma_i \) of Eq. (18) yields: 

\[ 0 = \frac{\partial (Au)}{\partial \sigma_i} = \frac{\partial A}{\partial \sigma_i} u + A \frac{\partial u}{\partial \sigma_i}. \]  \hspace{1cm} (19) 

Hence, 

\[ \frac{\partial \Phi}{\partial \sigma_i} = -\nabla_u \Phi \cdot A^{-1} \frac{\partial A}{\partial \sigma_i} u. \]

If Eq. (18) is generated approximating, for instance, a differential equation with a finite element or finite difference method, then \( \sigma \) could be a vector, representing a grid function, i.e., a function defined by its values in the grid nodes. In that case the matrix \( \partial A/\partial \sigma_i \) has very few nonzero entries, and computation of \( \partial \Phi/\partial \sigma_i \) can be significantly simplified.

Applying this approach to our inverse problem we proceed as follows. The goal is to find the gradient of the objective function as defined in Eq. (17), where \( \{ \Psi^l \}_{l=1, L} \) solves Eq. (14). The gradient of \( \Phi(M, \Psi) \) with respect to \( \mu_a \) reads as 

\[ \nabla_{\mu_a} \Phi(M, \Psi) = \sum_{s, d} \frac{2[\sum_l W_l \Psi^l_s(x_d) - M_s]}{M_{s,d}^2} \sum_l W_l \nabla_{\mu_a} \Psi^l_s(x_d). \]  \hspace{1cm} (20) 

We want to find a partial derivative

\[ \frac{\partial \Psi^l_s(x_d)}{\partial \mu_a(p)}, \]

where \( p \) is the point of the mesh \( G_h \), where \( \mu_a(p) \) is the value of the absorption coefficient at that point. Similarly to Eq. (19) we get:

\[ A_k \frac{\partial \psi^l_s}{\partial \mu_a(p)} - B \sum_{l=1}^{L} W_l \frac{\partial \psi^l_s}{\partial \mu_a(p)} = -\frac{\partial A_k}{\partial \mu_a(p)} \psi^l_s + \frac{\partial B}{\partial \mu_a(p)} \sum_{l=1}^{L} W_l \psi^l_s. \]  \hspace{1cm} (21)

Here \( (\partial A_k)/[\partial \mu_a(p)] \) denotes the matrix with the entries \( [\partial A_k]/[\partial \mu_a(p)] \). Obviously, \( (\partial B)/[\partial \mu_a(p)] = 0 \). Using Eq. (16) we deduce from Eq. (21) that 

\[ \frac{\partial \psi^l_s}{\partial \mu_a(p)} = A^{-1} \cdot \begin{pmatrix} -\frac{\partial A_k}{\partial \mu_a(p)} \psi^l_s \\ \vdots \\ -\frac{\partial A_l}{\partial \mu_a(p)} \psi^l_s \end{pmatrix}. \]  \hspace{1cm} (22) 

We define a projection vector \( P_d \) as a vector of size equal to the number of mesh nodes. The entry of \( P_d \) corresponding to the node \( x_d \) is equal to 1, and all other entries are zeros. Then

\[ \frac{\partial \Psi^l_s(x_d)}{\partial \mu_a(p)} = P^T_d \cdot \frac{\partial \psi^l_s}{\partial \mu_a(p)}. \]

From Eq. (20) we have:

\[ \frac{\partial \Phi(M, \Psi)}{\partial \mu_a(p)} = \sum_s \left( \sum_d \frac{2[\sum_l W_l \Psi^l_s(x_d) - M_s]}{M_{s,d}^2} P^T_d \right) \times \sum_l W_l \frac{\partial \Psi^l_s}{\partial \mu_a(p)}. \]  \hspace{1cm} (23) 

Let vector \( [(\psi^1_s)^T, \ldots, (\psi^L_s)^T]^T \) be a solution of the following system:

\[ A \begin{pmatrix} \psi^1_s \\ \vdots \\ \psi^L_s \end{pmatrix} = \begin{pmatrix} -W_1 \sum_d \frac{2[\sum_l W_l \Psi^l_s(x_d) - M_s]}{M_{s,d}^2} P_d \\ \vdots \\ -W_L \sum_d \frac{2[\sum_l W_l \Psi^l_s(x_d) - M_s]}{M_{s,d}^2} P_d \end{pmatrix}. \]  \hspace{1cm} (24) 

Equation (24) is usually referred as an adjoint system. Then from Eqs. (22) and (23) it follows that

\[ \frac{\partial \Phi(M, \Psi)}{\partial \mu_a(p)} = \sum_s \sum_l (\psi^*_l)^T \left( -\frac{\partial A_l}{\partial \mu_a(p)} \right) \psi^l_s. \]  \hspace{1cm} (25) 

The matrix \((\partial A_l)/[\partial \mu_a(p)]\) is sparse and only entries corresponding to the node \( p \) and to adjacent nodes are non-zero. Taking advantage of the sparsity pattern of this matrix, the matrix-vector multiplication procedure in Eq. (25) can be implemented very efficiently.

In much the same way as in Eq. (25), differentiating the objective function [Eq. (17)] with respect to \( \mu_s(p) \) yields

\[ \frac{\partial \Phi(M, \Psi)}{\partial \mu_s(p)} = \sum_s \sum_l (\psi^*_l)^T \left( -\frac{\partial A_l}{\partial \mu_s(p)} \right) \psi^l_s + \sum_s \sum_{l=1}^{L} (\psi^*_l)^T \left( -\frac{\partial B}{\partial \mu_s(p)} \right) \sum_k W_k \psi^l_k. \]  \hspace{1cm} (26)
Once the gradient of the objective function [Eq. (17)] is computed for an initial guess of optical properties \( \mu_a^k \) and \( \mu_s^k \), we can find an update of these parameters using a gradient minimization scheme. These schemes are of the general form

\[
\begin{align*}
\mu_a^{k+1} &= \mu_a^k - \alpha \nabla_{\mu_a} \Phi[M, \Psi(\mu_a^k, \mu_s^k)], \\
\mu_s^{k+1} &= \mu_s^k - \alpha \nabla_{\mu_s} \Phi[M, \Psi(\mu_a^k, \mu_s^k)],
\end{align*}
\]

where \( \alpha \) is some positive parameter. For instance, \( \alpha \) can be chosen to minimize \( \Phi(M, \Psi) \) along the gradient direction or the conjugate gradient method can be employed.\(^{42}\) Most recently our group has implemented a quasi-Newton method, where

\[
\begin{align*}
\mu_a^{k+1} &= \mu_a^k - H_k \nabla_{\mu_a} \Phi[M, \Psi(\mu_a^k, \mu_s^k)], \\
\mu_s^{k+1} &= \mu_s^k - H_k \nabla_{\mu_s} \Phi[M, \Psi(\mu_a^k, \mu_s^k)],
\end{align*}
\]

Here \( H_k \) is an approximate inverse of the Hessian matrix. Different strategies for computing \( H_k \) can be used. We employ in this work the limited-memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) method.\(^{43}\)

3 Results

3.1 Hexahedral Mesh

To test the code we initially consider the reconstruction of a medium with two spherical objects located in a cubic domain with the edge size 3 cm (Fig. 1). We have 8 sources and 36 detectors, which are distributed in three layers, producing 280 source-detector pairs. As measurements we use synthetic data obtained solving the forward problem with "exact" optical properties. Therefore the forward solver is used with the two heterogeneities present to generate the data that is input for the reconstruction algorithm. No noise is added to the data. For the calculations we discretize the medium using a hexagonal mesh and trilinear finite elements. A total of 33\(^3\) mesh nodes with 4 discrete ordinates per hemisphere is used. This results in 143,748 (number of mesh nodes times number of discrete ordinates) unknowns for the forward and 35,937 (number of mesh nodes) unknowns for the inverse problem. One iteration in the gradient-based scheme takes approximately 30 min on a workstation with a Pentium III XEON 700 MHz processor.

In the first example we reconstruct the absorption coefficient \( \mu_a \). The background value of \( \mu_a \) is equal to 0.1 cm\(^{-1}\) and in the heterogeneities \( \mu_a = 0.5 \) cm\(^{-1}\). The scattering coefficient is \( \mu_s = 5 \) cm\(^{-1}\) throughout the medium. The results of this reconstruction after 15 iterations of the limited-memory BFGS method are presented in Fig. 2. The iterations started with a homogeneous medium (\( \mu_a = 0.1 \) cm\(^{-1}\) and \( \mu_s = 5 \) cm\(^{-1}\) ) as an initial guess. The problem of reconstruction of the scattering coefficient is considered in the second example (Fig. 3), where the background scattering is set to \( \mu_s = 1 \) cm\(^{-1}\) and the scattering coefficient inside the heterogeneities is \( \mu_s = 3 \) cm\(^{-1}\). In this example the absorption coefficient was constant (\( \mu_a = 0.1 \) cm\(^{-1}\) ), and as an initial guess a homogeneous me-
dium with $\mu_a = 0.1 \text{ cm}^{-1}$ and $\mu_s = 1 \text{ cm}^{-1}$ was used. In both examples the MOBIIR algorithm is able to locate the objects, although for the absorption coefficient reconstruction the value of $\mu_a$ in the absorbers was underestimated, and for the scattering reconstruction some artifacts are present. Improved performances can be expected by using, for example, additional regularization mechanisms.

### 3.2 Tetrahedral Mesh

Going beyond the simple hexahedral mesh we considered a tetrahedral mesh of the human forehead (Fig. 4). In previous works we had used this mesh to reconstructed three-dimensional images of hemodynamic changes in the brain during a Valsalva maneuver. The size of the domain is approximately $11 \times 6 \times 5 \text{ cm}$. The surface of this mesh was obtained with photogrammetric methods applied to three photographs of a head. For more details concerning this mesh and the experimental procedures see Ref. 44.

Here we attempt to reconstruct a spherical absorbing object with a radius of 0.5 cm, which is embedded close to the frontal surface of the forehead-shaped three-dimensional domain. The background absorption coefficient is equal to $1 \text{ cm}^{-1}$, and the heterogeneity has an absorption coefficient of $\mu_a = 2 \text{ cm}^{-1}$. The scattering coefficient $\mu_s = 5 \text{ cm}^{-1}$ is constant throughout the mesh. Note that in this example we intentionally chose a rather high $\mu_a$. In this way we mimic a situation in which the commonly used diffusion approximation is not valid. For the reconstruction we use synthetic data from 4 sources and 15 detectors, which are arranged in a grid on the surface of the domain [see Fig. 4(b)] with a total of 56 source-detector pairs. Linear finite elements on a tetrahedral mesh are employed to approximate the even-parity equation [Eq. (12)].

A total of 11,244 mesh nodes with 4 discrete ordinates per hemisphere are used. One iteration of the limited-memory BFGS method takes on average 25 to 35 min on a LINUX workstation with a Pentium III XEON 700 MHz processor.

Results obtained using the MOBIIR scheme are displayed in Fig. 5. We started the iterations with a homogeneous medium ($\mu_a = 1 \text{ cm}^{-1}$ and $\mu_s = 5 \text{ cm}^{-1}$) as an initial guess. Figures 5(a) and 5(c) show the exact distributions of the absorption coefficient $\mu_a$ in a cross-sectional plane cutting through the center of the absorber for different depths of the absorber from the front surface. Figures 5(b) and 5(d) display corresponding reconstructions of $\mu_a$ after 15 iterations of the limited-memory BFGS method. The reconstruction of the object located closer to the surface is more accurate than the object that is deeper situated inside the head. This is expected, because of the particular source-detector configuration that only covers an area on the surface of the forehead and does not surround the inhomogeneities. As reported in earlier studies, in this case reconstructed objects tend to be pulled toward the surface.

### 3.3 Discussion

The presented code is the first three-dimensional image reconstruction scheme for arbitrarily shaped media solely based on the ERT. The initial results are promising, but further improvements are necessary. Currently the code is limited to isotropic scattering. However, the anisotropic scattering case does not require conceptually new insights. Using the even-parity approach, the forward problem has already been solved for anisotropic scattering in one and two dimensions by Kaplan and Davis, Lillie and Robinson, and Oliveira. The fundamental structure of the equations for the inverse problem [Eqs. (18) to (24)]
does not change. However, the matrices involved become much more complex, and efficient numerical schemes to solve the ensuing system of equations for the inverse problem still need to be developed.

Nevertheless we feel that the presented methods and results are an important first step in the right direction. Especially given the fact that a transport code with isotropic scattering is already an advance over diffusion-based methods as illustrated by the example in Fig. 6 (adapted from Hielser et al.47). This figure shows the calculated fluences on a line through an anatomically correct, hematoma-containing human head for three different cases. Using the DANTSYS (Diffusion Accelerated Neutal Particle Transport Code System48,49) software package fluences were calculated for an isotropic scattering phase function (anisotropy factor \(g=0\), which is the expectation value of the scattering angle’s cosine50), a highly forward peaked scattering phase function (\(g=0.99\)), and with the diffusion approximation. As can be seen diffusion and transport calculations differ substantially (up to 3 orders of magnitude right behind the hematoma, in which \(\mu_s'=(1-g)\mu_s\) is not much larger than \(\mu_s\)), while differences between transport calculation for \(g=0\) and \(g=0.99\) are comparatively small. For more details concerning this study see Ref. 47.

Furthermore, we have currently not used regularization terms. However, as Hielser et al. showed, other objective functions that include penalty terms may lead to improved performance39 and future work will consider these regularization schemes. In addition a thorough analysis concerning the differences between diffusion-theory-based and transport-theory-based reconstructions is highly desirable. To date, most comparisons of this kind have been limited to the forward model only. How errors in the forward model affect the reconstruction especially in real world applications such as brain and joint imaging is an important question and should be pursued in future studies. In our particular simulations we did not notice a strong dependence of the accuracy of the reconstruction on the number of ordinates, however more detailed investigations are required to fully explore this issue.

4 Summary

We have implemented a three-dimensional image reconstruction code that is based on the equation of radiative transfer (ERT). Unlike previously available schemes our approach allows one to consider arbitrary shaped media with arbitrary distribution of optical properties. The newly developed algorithm employs an even-parity formulation of the ERT, which reduces the ERT to a system of second-order equations. These equations can be conveniently discretized using a finite element method. The arising linear algebraic system is used as the forward model in a model-based iterative image reconstruction scheme (MOBIIR), in which the image reconstruction problem can be interpreted as an optimization problem. The gradient of the objective function is found by employing an adjoint differentiation method to the forward solver. Initial tests using synthetic data show that heterogeneities in scattering media can be accurately located. While the implemented code only allows for isotropic scattering, it nevertheless constitutes an advance over currently available diffusion-based algorithms. Future improvements of the code will focus on the numerical implementation of schemes that allow consideration of anisotropic scattering, the inclusion of regularization methods, and a detailed comparison of diffusion- and transport-based image reconstruction codes for important practical problems, such as brain or joint imaging.

Acknowledgments

We would like to thank Avraham Bluestone and Dr. Alexander Klose, both with the Dept. of Biomedical Engineering, Columbia University, New York, NY, for many fruitful discussions concerning inverse problems and adjoint differentiation. This work was supported in part by the National Institute of Arthritis and Musculoskeletal Diseases, a division of the National Institutes of Health (ROI-AR46255), the Whitaker Foundation (98-0244), and the New York Council Speaker’s Fund for Biomedical Research: Towards the Science of Patient Care.

References

Three-dimensional optical tomography


