REVIEW

Numerical modelling and image reconstruction in diffuse optical tomography

BY HAMID DEHGHANI1,* , SUBHADRA SRINIVASAN2, BRIAN W. POGUE2 AND ADAM GIBSON3

1 School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
2 Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA
3 Department of Medical Physics and Bioengineering, University College London, Gower Street, London WC1E 6BT, UK

The development of diffuse optical tomography as a functional imaging modality has relied largely on the use of model-based image reconstruction. The recovery of optical parameters from boundary measurements of light propagation within tissue is inherently a difficult one, because the problem is nonlinear, ill-posed and ill-conditioned. Additionally, although the measured near-infrared signals of light transmission through tissue provide high imaging contrast, the reconstructed images suffer from poor spatial resolution due to the diffuse propagation of light in biological tissue. The application of model-based image reconstruction is reviewed in this paper, together with a numerical modelling approach to light propagation in tissue as well as generalized image reconstruction using boundary data. A comprehensive review and details of the basis for using spatial and structural prior information are also discussed, whereby the use of spectral and dual-modality systems can improve contrast and spatial resolution.

Keywords: near-infrared imaging; image reconstruction; optical tomography

1. Introduction

Diffuse optical tomography (DOT) is a non-invasive imaging technique that has been under development since the early 1990s for the detection and characterization of functional changes within biological tissue (Boas et al. 2001; Gibson et al. 2005). Owing to the relatively low absorption of haemoglobin, water and lipid at wavelengths of 650–1000 nm, near-infrared (NIR) light can transmit through several centimetres of tissue with adequate signal-to-noise ratio to allow tomographic detection of light transmission. Using the measured NIR signal, together with the known spectral absorption and scattering coefficients of tissue, it is possible to extract functional information about the tissue being imaged.

* Author for correspondence (h.dehghani@cs.bham.ac.uk).

One contribution of 7 to a Theme Issue ‘New and emerging tomographic imaging’.

This journal is © 2009 The Royal Society
The development of the technique, from spectroscopic point measurements to full tomographic imaging systems, has already been outlined in the review by Gibson & Dehghani (2009).

Several imaging modalities have arisen with specific application to the pathophysiological imaging of biological tissue, including applications in breast cancer detection and characterization (Tromberg et al. 1997; Fantini et al. 1998; Liu et al. 2000; Zhu et al. 2000; Chance 2001; Jiang et al. 2001; Hebden et al. 2001; Ntziachristos & Chance 2001; Pogue et al. 2001; Shah et al. 2001; Dehghani et al. 2003c; Srinivasan et al. 2003b, 2005a; Choe et al. 2005; Yates et al. 2005; Boverman et al. 2007; Enfield et al. 2007), brain functional imaging (Chance et al. 1988a,b; Cooper et al. 1996; Siegel et al. 1999; Bluestone et al. 2001; Zaramella et al. 2001; Hebden et al. 2002; Franceschini et al. 2003; Obrig & Villringer 2003; Boas et al. 2004; Selb et al. 2005; Austin et al. 2006; Zeff et al. 2007), imaging of the finger joint for detection and quantification of arthritis (Hielscher et al. 2004) as well as muscle function imaging (Chance et al. 1988a,b; Wilson et al. 1989; Mancini et al. 1994; Breit et al. 1997; Hillman et al. 2001). Typically in DOT, a number of optical connectors are placed on the periphery of the domain being imaged and NIR light at specific wavelengths is injected into the domain using one source connector at a time, while all other remaining connectors are used to detect the transmitted light. The collected data, called ‘boundary data’, are then used for image reconstruction. Typically, images of optical absorption and reduced scatter are reconstructed, one wavelength at a time, from which maps of internal chromophore concentrations and scattering properties can be calculated using Beer’s law and approximations to Mie scattering (Dehghani et al. 2003a). However, more recently, the use of spectral image reconstruction has become widespread, whereby the incorporation of multi-wavelength data can allow for the direct reconstruction of tissue chromophore and scattering properties simultaneously (Li et al. 2004; Corlu et al. 2005; Srinivasan et al. 2005b; Wang et al. 2006).

In most DOT imaging studies, the majority of the work has relied on the use of model-based image reconstruction. A large number of different models can be used to predict light propagation within tissue, including stochastic, analytical and numerical. The definition and description of all such modelling techniques is beyond the scope of this review, and readers should refer to previously published work (Arridge & Hebden 1997; Arridge & Schweiger 1997; Arridge 1999).

Stochastic models involve predicting individual photon interactions using either explicit or implicit methods. Two of the most common methods include Monte Carlo methods (Wilson & Adam 1983; Wang et al. 1995) and random walk theory (Gandjbakhche & Weiss 1995). Analytical models have the advantage of being computationally fast but suffer from the disadvantage of being limited to simple geometries with nearly homogeneous interior values. Analytical solutions for various geometries using Green’s function have been derived and discussed by Arridge et al. (1992).

Conversely, numerical models have the potential of being able to model complex geometries as well as complex heterogeneous media, but have historically required longer computation times. But, perhaps the most promising reason for adoption of numerical approaches is to facilitate the combination of NIR tomography with standard clinical imaging systems, using predefined tissue geometries as the input domain. A number of different numerical models have
been developed and used with specific application in DOT, including finite elements (Arridge et al. 1993; Jiang & Paulsen 1995; Schweiger et al. 1995; Gao et al. 1998; Jiang 1998; Dehghani et al. 2003), finite difference (Hielscher et al. 1998; Klose & Hielscher 1999), finite volume (Ren et al. 2004) and boundary elements (Zacharopoulos et al. 2006; Srinivasan et al. 2007).

2. The forward model

Several models have been developed to predict the propagation of NIR light within biological tissue. For a good description of various models, readers are referred to the review paper by Arridge (1999). Largely, these models have relied on the use of either the radiative transport equation (RTE), or the simpler diffusion approximation. In this section we give a general account of these two main models being adapted to specific application of DOT.

(a) Radiative transport model

By treating light as photons, and thus ignoring any wave effects, the propagation of light in tissue can be well described by the RTE

\[
\frac{i\omega}{c} I(r, \omega, \hat{s}) + \hat{s} \cdot \nabla I(r, \omega, \hat{s}) + (\mu_a + \mu_s) I(r, \omega, \hat{s}) = \mu_s \int_{4\pi} f(\hat{s}, \hat{s}') I(r, \omega, \hat{s}') d^2\hat{s}' + q(r, \omega, \hat{s}),
\]

where \( I(r, \omega, \hat{s}) \) is the radiance at point \( r \), modulation frequency \( \omega \) and in the direction \( \hat{s} \); \( \mu_a \) and \( \mu_s \) are the absorption and scattering coefficients, respectively; and \( c \) is the speed of light in the medium. The \( f(\hat{s}, \hat{s}') \) term is the scattering phase function, which characterizes the intensity of a beam that is scattered from direction \( \hat{s}' \) into direction \( \hat{s} \) (Arridge 1999). The scattering phase function most typically employed is the commonly used Henyey–Greenstein scattering function (Welch & van Gemert 1995; Klose et al. 2002; Tarvainen et al. 2005)

\[
f(\cos \theta) = \frac{1 - \eta^2}{2(1 + \eta^2 - 2\eta \cos \theta)^{3/2}},
\]

where \( \theta \) is the angle between the two directions \( \hat{s} \) and \( \hat{s}' \); and \( \eta \) is the anisotropy factor, which is used to characterize the angular distribution of tissue scattering.

Several groups have developed and used a number of forward models based on the RTE for DOT (Hielscher et al. 1998, 2004; Klose & Hielscher 1999; Klose et al. 2002; Tarvainen et al. 2005), with most clinical applications relying on the accuracy of the RTE for problems where the propagation of photons may not be assumed as diffuse, for example small animal imaging and problems where low scattering due to clear layers may be present.

A major drawback of the use of RTE is the complex implementation within a numerical setting. One specific challenge is the use of an appropriate method for the incorporation of the angular dependence of the problem. The discrete-ordinates method is widely used with several different finite-difference approximations (Lathrop 1972), such as the diamond difference scheme, the weighted diamond
difference scheme, the centred difference scheme (Reed 1971) or the upwind-difference scheme (Klose et al. 2002). Another common approach is the use of a spherical harmonics expansion, whereby the angular dependence can be described by a set of spherical harmonics (Boas et al. 1995; Aydin et al. 2002), and more recently by the use of simplified spherical harmonics (Jiang 1999; Klose & Larsen 2006).

(b) Diffusion approximation

It is generally accepted that, if the magnitude of the isotropic fluence within tissue is significantly larger than the directional flux magnitude, the light field is ‘diffuse’, which occurs when the scattering interactions dominate over absorption and the region of interest is far from sources and boundaries, provided the light fluence is not rapidly changing with time (i.e. such as in the sub-picosecond time frame). This assumption allows a transition from the RTE, which is used to describe an anisotropic light field, to the diffusion equation approximation, which is used for isotropic fluence fields (Arridge 1999). The diffusion approximation in the frequency domain is given by

\[-\nabla \cdot \kappa(r) \nabla \Phi(r, \omega) + \left( \mu_a(r) + \frac{i\omega}{c(r)} \right) \Phi(r, \omega) = q_0(r, \omega),\]

where \(\mu_a\) and \(\mu_s'\) are absorption and reduced scattering (or transport scattering) coefficients, respectively; \(q_0(r, \omega)\) is an isotropic source; \(\Phi(r, \omega)\) is the photon fluence rate at position \(r\) and modulation frequency \(\omega\); \(\kappa = 1/(3(\mu_a + \mu_s'))\) is the diffusion coefficient; and \(c(r)\) is the speed of light in the medium at any point, defined by \(c_0/n(r)\), where \(n(r)\) is the index of refraction at the same point and \(c_0\) is the speed of light in vacuum.

The air–tissue boundary is represented by an index-mismatched type III condition (also known as Robin or mixed boundary condition), in which the fluence at the edge of the tissue exits but does not return (Schweiger et al. 1995; Dehghani et al. 2003b). The flux leaving the external boundary is equal to the fluence rate at the boundary weighted by a factor that accounts for the internal reflection of light back into the tissue. This relationship is described as

\[\Phi(\xi, \omega) + 2A\hat{n} \cdot \kappa(\xi) \nabla \Phi(\xi, \omega) = 0,\]

where \(\xi\) is a point on the external boundary, and \(A\) depends upon the relative refractive index (RI) mismatch between the tissue domain \(\omega\) and air. Here, \(A\) can be derived from Fresnel’s law

\[A = \frac{2/(1 - R_0) - 1 + |\cos \theta_c|^3}{1 - |\cos \theta_c|^2},\]

where \(\theta_c = \arcsin(n_{\text{air}}/n_1)\), the angle at which total internal reflection occurs for photons moving from region \(\omega\) with RI \(n_1\) to air with RI \(n_{\text{air}}\), and \(R_0 = (n_1/n_{\text{air}} - 1)^2/(n_1/n_{\text{air}} + 1)^2\). At the external boundaries, RI is generally assumed to be equal to that of free space, so that \(n_{\text{air}} = 1\).
(c) Data types

From either equation (2.1) or (2.3), two sets of boundary data can be extracted from models. These boundary data correspond to the fluence measured at the external boundary at points where the detector fibres are present. In the case of a frequency-domain problem, the measured boundary data include the intensity (or amplitude) of the measured signal and the corresponding phase. By setting the modulation frequency $\omega = 0$ MHz, this will lead to the continuous wave (CW) system, whereby only the amplitude of the measured signal is available.

Although the models defined in equations (2.1) and (2.3) are frequency-domain models, they can easily be adapted for time-resolved cases. The development of time-resolved models has been discussed extensively elsewhere (Patterson et al. 1989; Arridge et al. 1993), whereby the propagation of photons throughout the imaging model can be simulated as a function of pulsed (delta function at time = 0) sources. Using such an approach, it is possible to calculate the temporal point spread function (TPSF) of the measured signal at each detector position. Using the TPSF, it is possible to extract a number of data types from the measurement, including total intensity (analogous to intensity using CW or frequency-domain models), mean time of flight (analogous to phase measurement using frequency-domain models), variance, skew and other higher-order moments (Arridge & Schweiger 1995a, b).

Although some work has been done to use the full time-resolved data for image reconstruction (Gao et al. 2002), it is generally accepted that the use of at least the two data types of intensity and mean time of flight from the measured TPSF is sufficient to provide the full amount of information required for the imaging of both absorption and scattering coefficients in DOT (Gruenbaum 2001). A common misconception is, however, that, in order to model and calculate the higher moments (and hence data types other than intensity), it is necessary to calculate the full TPSF for a given model, which can be computationally intensive and time consuming. However, Arridge & Schweiger (1995a, b) demonstrated that ‘direct’ calculation of the moments of the distribution of photon time of flight in tissue can be achieved without the requirement for the calculation of the full TPSF.

3. The inverse model

The goal of the inverse problem is the recovery of the unknown optical properties (either absorption and scattering at each wavelength, or chromophore concentration and scattering properties from multi-wavelength data) using the measured boundary data. There are generally two approaches taken when attempting to reconstruct the unknown distribution of optical properties: (a) linear single-step and (b) nonlinear iterative reconstruction schemes.

(a) Linear single-step reconstruction

The goal of linear image reconstruction is to obtain images of temporal (time-dependent) changes of optical properties. Without exception, this method requires, either explicitly or implicitly, a difference experiment that measures the boundary data as the difference between two states, $\delta \phi = \phi_{\text{anom}} - \phi_{\text{ref}}$, where
\( \phi_{\text{anom}} \) and \( \phi_{\text{ref}} \) correspond to data acquired with and without a change in optical properties, respectively. This approach provides a means of imaging that is sensitive to changes in optical properties, which may be particularly useful for functional imaging of the brain, for example (Zeff et al. 2007). However, given that the problem is inherently nonlinear, care must be taken such that the imaged changes are relatively small. Nonetheless, this method is only suitable for providing qualitative images of measured changes, rather than absolute quantitative changes.

(b) Nonlinear iterative reconstruction

The aim of nonlinear image reconstruction is to calculate optical properties \( \mu = (\mu_a, \kappa) \) at each point within the model using measurements of light fluence from the tissue surface. There are two distinct approaches that can be used for such ‘optimization’ technique, namely those that use gradient-based reconstruction techniques (Arridge & Schwieger 1998; Hielscher et al. 1999; Hielscher & Bartel 2001), or those that require the direct calculation and inversion of the Jacobian (also known as the sensitivity matrix), which can be classed as Newton-like methods (Jiang et al. 1996; Dehghani et al. in press). Using the gradient-based optimization technique is accepted as being less computationally intensive, since rather than calculating the objective function directly, only the gradient need be calculated, therefore reducing the total amount of computational resources needed. However, one drawback of such a scheme is that the total number of ‘iterations’ required is substantially more than the alternative, whereby the Jacobian must be calculated and inverted. In the following section, we limit our discussion to the Newton-like optimization techniques, but the interested reader can refer to the works by Arridge & Schwieger (1998) and Hielscher et al. (1999) and Hielscher & Bartel (2001) for more information on gradient-based techniques.

If the measured fluence at the surface of the domain being imaged is represented by \( \Phi^M \) and the calculated model-based data using the forward solver by \( \Phi^C \), then the standard Tikhonov minimization function is

\[
\Omega = \min_\mu \left\{ \sum_{i=1}^{NM} (\Phi^M_i - \Phi^C_i)^2 + \lambda \sum_{j=1}^{NN} (\mu_j - \mu_0)^2 \right\}. \tag{3.1}
\]

Here \( NM \) is the total number of boundary measurements obtained from the imaging device; \( NN \) is the number of unknown parameters; \( \lambda \) is the Tikhonov regularization parameter, which is defined as the ratio of the variances of the measurement data and optical properties \( (\lambda = \sigma_{\phi M}^2 / \sigma_{\mu}^2) \) (Yalavarthy et al. 2007); and \( \mu_0 \) is either the initial estimate of the optical properties, which is generally obtained by a data-calibration procedure (McBride et al. 2003), or an a priori optical property distribution, which may be available from either other imaging modalities or published literature values. It has been found that, if the initial estimate, \( \mu_0 \), is not too far from the actual parameter distribution, this term can be ignored (Dehghani et al. 2003c; Brooksby 2005).

The minimization with respect to \( \mu \) in equation (3.1) involves setting the first-order derivative equal to zero, \( (\partial \Omega / \partial \mu) = 0 \), and ignoring higher-order terms. The first-order condition is therefore given by

\[\text{[Equation]}\]
The derivative matrix \( \left( \frac{\partial \Phi^C}{\partial \mu} \right)^T \) is known as the Jacobian matrix, \( J \), and is also referred to as the weight or sensitivity matrix. Using this linear approximation of the problem, and solving it as an iterative scheme, we get
\[
(J^T J + \lambda I) \delta \mu = J^T \delta \Phi - \lambda (\mu_i - \mu_0),
\]
where \( \delta \mu \) is the update for the optical properties; \( \delta \Phi \) is the data–model misfit at the current iteration; and \( I \) is the identity matrix.

(c) The Jacobian

The Jacobian defines the relationship between changes in boundary data \( \Phi^C \), resulting from small changes in optical properties \( \mu = (\mu_a, \kappa) \). Since both amplitude and phase data types are available from a frequency-domain system, and since the problem considers the effects of absorption and diffusion, the structure of the Jacobian is given, for example, by the following equation:

\[
J = \begin{bmatrix}
\frac{\delta \ln I_1}{\delta \kappa_1} & \frac{\delta \ln I_1}{\delta \kappa_2} & \cdots & \frac{\delta \ln I_1}{\delta \kappa_{NN}} & \frac{\delta \ln I_1}{\delta \mu_{a1}} & \frac{\delta \ln I_1}{\delta \mu_{a2}} & \cdots & \frac{\delta \ln I_1}{\delta \mu_{aNN}} \\
\frac{\delta \theta_1}{\delta \kappa_1} & \frac{\delta \theta_1}{\delta \kappa_2} & \cdots & \frac{\delta \theta_1}{\delta \kappa_{NN}} & \frac{\delta \theta_1}{\delta \mu_{a1}} & \frac{\delta \theta_1}{\delta \mu_{a2}} & \cdots & \frac{\delta \theta_1}{\delta \mu_{aNN}} \\
\frac{\delta \ln I_2}{\delta \kappa_1} & \frac{\delta \ln I_2}{\delta \kappa_2} & \cdots & \frac{\delta \ln I_2}{\delta \kappa_{NN}} & \frac{\delta \ln I_2}{\delta \mu_{a1}} & \frac{\delta \ln I_2}{\delta \mu_{a2}} & \cdots & \frac{\delta \ln I_2}{\delta \mu_{aNN}} \\
\frac{\delta \theta_2}{\delta \kappa_1} & \frac{\delta \theta_2}{\delta \kappa_2} & \cdots & \frac{\delta \theta_2}{\delta \kappa_{NN}} & \frac{\delta \theta_2}{\delta \mu_{a1}} & \frac{\delta \theta_2}{\delta \mu_{a2}} & \cdots & \frac{\delta \theta_2}{\delta \mu_{aNN}} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\delta \ln I_{NM}}{\delta \kappa_1} & \frac{\delta \ln I_{NM}}{\delta \kappa_2} & \cdots & \frac{\delta \ln I_{NM}}{\delta \kappa_{NN}} & \frac{\delta \ln I_{NM}}{\delta \mu_{a1}} & \frac{\delta \ln I_{NM}}{\delta \mu_{a2}} & \cdots & \frac{\delta \ln I_{NM}}{\delta \mu_{aNN}} \\
\frac{\delta \theta_{NM}}{\delta \kappa_1} & \frac{\delta \theta_{NM}}{\delta \kappa_2} & \cdots & \frac{\delta \theta_{NM}}{\delta \kappa_{NN}} & \frac{\delta \theta_{NM}}{\delta \mu_{a1}} & \frac{\delta \theta_{NM}}{\delta \mu_{a2}} & \cdots & \frac{\delta \theta_{NM}}{\delta \mu_{aNN}}
\end{bmatrix},
\]

where \( \delta \ln I_i/\delta \kappa_j \) and \( \delta \ln I_i/\delta \mu_{aj} \) are sub-matrices that define the change in the log of the amplitude of the \( i \)th measurement arising from a small change in \( \kappa \) and \( \mu_a \) at the \( j \)th reconstructed node, respectively; and \( \delta \theta_i/\delta \kappa_j \) and \( \delta \theta_i/\delta \mu_{aj} \) are sub-matrices that give the change in phase of the \( i \)th measurement arising from a change in \( \kappa \) and \( \mu_a \) at the \( j \)th node, respectively.

The calculation of the Jacobian matrix can take one of three possible forms: (i) the perturbation method, (ii) the direct method, and (iii) using the adjoint theorem. The perturbation method requires the calculation of the boundary data, before and after each point within the model has been perturbed by a small
amount in either of the unknown parameters, and will therefore require a complete set of \((NN \times 2) + 2\) forward calculations; that is, two sets of boundary data for the unperturbed model and NN sets of boundary data for absorption and diffusion coefficients. The use of the direct method involves the differentiation of equation (2.3) with respect to either the absorption or diffusion coefficient and then solution to obtain the individual values of the Jacobian matrix, therefore requiring \((NN \times 2)\) forward calculations. The adjoint theorem, however, makes use of the reciprocity theorem, which simply states that the measurement of flux at detector \(j\) that is due to a source at node \(i\) is equal to the measurement of the photon density at node \(i\) that is attributed to a source at detector \(j\). Using this method, it can be easily shown that, in order to calculate the Jacobian, only \(NS + ND\) forward calculations are needed, where \(NS\) is the total number of sources and \(ND\) is the total number of detectors (Arridge & Schweiger 1995a,b).

As indicated in equation (3.4), the Jacobian involves both diffusion coefficient \((\kappa)\) and absorption coefficient \((\mu_a)\) derivatives, so the Jacobian in the update equation (equation (3.3)) is normalized by a diagonal matrix \((G)\) consisting of the initial estimate of the optical properties \((\mu_0)\), such that

\[
\tilde{J} = JG,
\]

where \(G = \text{diag}(\kappa; \mu_a)\); moreover, \(\lambda\) is implemented in a modified Levenberg–Marquardt algorithm (Levenberg 1944), where it is initialized as the variances ratio and is systematically reduced at each iteration.

Once the optical properties are obtained at each wavelength, the chromophore concentrations are calculated using a constrained least-squares fit to the Beer’s law relationship

\[
\mu_a = [\varepsilon]c,
\]

where \(\varepsilon\) is the molar absorption spectra of the tissue’s absorbing chromophores and \(c\) is the concentration of these chromophores. Oxyhaemoglobin (HbO\(_2\)), deoxyhaemoglobin (Hb) and water are assumed to be the main absorbers and their molar absorption spectra have been obtained experimentally (Srinivasan et al. 2003a). By fitting for the concentrations, total haemoglobin is calculated as \(Hb_T = HbO_2 + Hb\) (in \(\mu\)M), and oxygen saturation as \(S_o_{2} = HbO_2/Hb_T \times 100\) (in %).

Similarly, the \(\mu'_s\) spectrum of tissue has been shown to fit well to an empirical approximation to Mie scattering theory (van Staveren et al. 1991; Mourant et al. 1997) given by

\[
\mu'_s = a\lambda^{-b}.
\]

Equation (3.7) is used to estimate the model parameters scattering amplitude \((a)\) and scattering power \((b)\) with wavelength in \(\mu\)m (van Staveren et al. 1991). The coefficient \(\mu'_s\) has units of \(\text{mm}^{-1}\). Both the scattering power and amplitude depend on the scattering centre size and number density and may reflect variations in tissue composition due to different cellular, organelle and structural sizes/densities (Wang et al. 2006). Typically, large scatterers have lower \(b\) and \(a\) values, whereas small scatterers have higher \(b\) and \(a\) values.
As an example of the application of this method in a clinical setting, NIR imaging of a volunteer subject is presented. The patient presented for standard screening mammography, which revealed a subtle nodular density and associated architectural distortion in the lateral aspect of the right breast. Pathology showed an invasive carcinoma of 20 mm size. The patient was presented for NIR measurement soon after biopsy.

Three-dimensional images of internal absorption and reduced scattering were reconstructed simultaneously from NIR data collected at each wavelength. The mesh used for the calculation of the Jacobian contained 8334 nodes corresponding to 41 623 linear tetrahedral elements (figure 1a). For the reconstruction basis, a second mesh of the same geometry was used, but with 2521 nodes, corresponding to 11 575 linear tetrahedral elements (figure 1b). Images were reconstructed at four wavelengths of 761, 785, 808 and 826 nm, and they are shown in figure 2. For image reconstruction, the regularization parameter (\( \lambda \)) used was initially set to 10, and was allowed to decrease by a factor of \( 10^{1/4} \) if the projection error (equation (3.1)) had decreased with respect to the previous iteration. The images shown are those at the 10th iteration. Here the images are true three-dimensional reconstructions, and coronal slices at \( z = -60, -45, -30, -15 \) and 0 mm only are shown. From the reconstructed images, it can be seen that an anomaly is found within the mid-plane at approximately the 9 o’clock position. The anomaly shows an absorption variation with wavelength, whereas the reduced scattering is almost constant over all the reconstructed wavelengths. The absorption images were used together with published values of extinction coefficients for oxy- and deoxyhaemoglobin (HbO\(_2\) and Hb, respectively) for the calculation of Hb, HbO\(_2\), total haemoglobin (Hb\(_T\)) and oxygen saturation (\( S_tO_2 \)). The calculated three-dimensional maps of Hb, HbO\(_2\), total Hb\(_T\) and \( S_tO_2 \) are also shown in figure 2.

From the calculated values of Hb, it is seen that the anomaly shows a peak value of 47.7 \( \mu \)mol, compared with a background of 32.87 \( \mu \)mol, whereas the HbO\(_2\) image shows a peak value of 24.73 \( \mu \)mol at a location on the periphery of the skin. The total haemoglobin value also shows a peak at the location of the anomaly, with a value of 65.28 \( \mu \)mol. \( S_tO_2 \) value, calculated by taking the ratio of oxygenated blood and total blood content, shows a marked decrease at the location of the anomaly, with a value of 25.9 per cent, as compared with a background value of 38.11 per cent.
Instead of reconstructing for optical properties at each wavelength and then applying equations (3.6) and (3.7) in a post-processing step, these constraints can be incorporated into the reconstruction directly to estimate the chromophore and scattering properties, thus reducing the parameter space (Srinivasan et al. 2003a, 2005b; Corlu et al. 2005). In this approach, the measurements at all measured wavelengths are coupled together and the relationships in equations (3.6) and (3.7) are combined to create a new set of equations. The Jacobian in equation (3.3), instead of relating the changes in optical properties to measured amplitude and phase, now relates the changes in chromophore concentrations and scattering parameters directly to changes in the log of the amplitude and phase data. The new Jacobian becomes

\[
J_{c,\lambda} = \frac{\partial \Phi}{\partial \mu} \bigg\rvert_{\lambda} = \frac{\partial \Phi}{\partial \mu} \frac{\partial \mu}{\partial c} \bigg\rvert_{\lambda}.
\]  

(3.8)

From equation (3.6), \( \partial \mu / \partial c \) is used, so that substituting for \( \partial \mu / \partial c \) yields

\[
J_{c,\lambda} = \frac{\partial \Phi}{\partial c} \bigg\rvert_{\lambda} = \frac{\partial \Phi}{\partial \mu} \epsilon \bigg\rvert_{\lambda} = \left( \frac{\partial \Phi}{\partial \mu} \bigg\rvert_{\lambda} \right) \otimes \left( \epsilon^{c1,c2,c3} \right) = J_{\mu,a,\lambda} \otimes \left( \epsilon^{c1,c2,c3} \right),
\]

(3.9)

Figure 2. Reconstructed images of (a) absorption and (b) reduced scattering at each wavelength from measured volunteer data: (i) 761 nm, (ii) 785 nm, (iii) 808 nm, (iv) 826 nm. (c) Calculated maps of blood content: (i) Hb, (ii) HbO2, (iii) HbT, (iv) \( S_{O2} \). Each slice represents a plane through the mesh, from the bottom near the nipple to the top near the chest. The images are coronal views of the cross section through the breast at the 10th iteration at the wavelengths indicated.

\( (a) \quad (i) \quad 10^{-3} \quad (ii) \quad 10^{-3} \quad (iii) \quad 10^{-3} \quad (iv) \quad 10^{-3} \)

\( (b) \quad (i) \quad 10^{-3} \quad (ii) \quad 10^{-3} \quad (iii) \quad 10^{-3} \quad (iv) \quad 10^{-3} \)

\( (c) \quad (i) \quad (ii) \quad (iii) \quad (iv) \)

(d) Spectral constraint

Instead of reconstructing for optical properties at each wavelength and then applying equations (3.6) and (3.7) in a post-processing step, these constraints can be incorporated into the reconstruction directly to estimate the chromophore and scattering properties, thus reducing the parameter space (Srinivasan et al. 2003a, 2005b; Corlu et al. 2005). In this approach, the measurements at all measured wavelengths are coupled together and the relationships in equations (3.6) and (3.7) are combined to create a new set of equations. The Jacobian in equation (3.3), instead of relating the changes in optical properties to measured amplitude and phase, now relates the changes in chromophore concentrations and scattering parameters directly to changes in the log of the amplitude and phase data. The new Jacobian becomes
where $\otimes$ refers to the Kronecker tensor product. Similarly,

$$J_{a,\lambda} = \frac{\partial \Phi}{\partial a} \bigg|_\lambda = \frac{\partial \Phi}{\partial \kappa} \frac{\partial \kappa}{\partial a} \bigg|_\lambda. \quad (3.10)$$

Rewriting $\frac{\partial \kappa}{\partial a} = \frac{\partial \kappa}{\partial \mu_s'}(\frac{\partial \mu_s'}{\partial a})$ results in

$$\left( \frac{\partial \kappa}{\partial \mu_s'} \right) = \frac{1}{3} \left( \frac{-1}{(\mu_a + \mu_s')^2} \right) = \frac{1}{3} (-9\kappa^2) = -3\kappa^2$$

and $\frac{\partial \mu_s'}{\partial a} = \lambda^{-b}$. Substituting these expressions in equation (3.10) leads to

$$J_{a,\lambda} = \frac{\partial \Phi}{\partial a} \bigg|_\lambda = \frac{\partial \Phi}{\partial \kappa} \frac{\partial \kappa}{\partial a} \bigg|_\lambda = J_k(-3\kappa^2)(\lambda^{-b})|_{\lambda}. \quad (3.11)$$

Similarly, expressions can be derived relating $J_{b,\lambda}$ and $J_{\kappa,\lambda}$ for scattering power $b$ (Srinivasan et al. 2005a). The overall system of equations is then given by these relationships:

$$\begin{pmatrix}
\frac{\partial \Phi_{11}}{\partial a} & \frac{\partial \Phi_{12}}{\partial a} & \cdots & \frac{\partial \Phi_{1n}}{\partial a} \\
\frac{\partial \Phi_{21}}{\partial a} & \frac{\partial \Phi_{22}}{\partial a} & \cdots & \frac{\partial \Phi_{2n}}{\partial a} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial \Phi_{n1}}{\partial a} & \frac{\partial \Phi_{n2}}{\partial a} & \cdots & \frac{\partial \Phi_{nn}}{\partial a}
\end{pmatrix}
\begin{pmatrix}
\frac{\partial c_1}{\partial a} \\
\frac{\partial c_2}{\partial a} \\
\vdots \\
\frac{\partial c_n}{\partial a}
\end{pmatrix}.$$

The Jacobian in equation (3.12) has dimensional size equal to the number of wavelengths times the number of measurements per wavelength times the number of nodes times the number of wavelength-independent parameters. The same Levenberg–Marquardt regularization scheme can be applied as in the conventional approach. However, since the size of the Jacobian is now dominated by the larger number of unknowns, the Moore–Penrose generalized inverse is used, which is more suitable for underdetermined problems (Penrose 1955). This gives rise to an update equation

$$J^T(JJ^T + \lambda I)^{-1}\partial \Phi = \partial \mu, \quad (3.13)$$

where $\mu$ is now the update vector consisting of the chromophore concentration and the scattering parameter. As in the single-wavelength case, since the Jacobian involves derivatives of different chromophores as well as scattering parameters, the Jacobian in the update equation (equation (3.13)) is normalized by a diagonal matrix $(G)$ consisting of the initial estimates of the unknown parameters.

As an example to demonstrate the capability of the spectrally constrained image reconstruction approach, a phantom using gelatin with whole blood added for absorption and titanium dioxide for scattering was made with a 25 mm hole drilled 10 mm from the boundary. This hole was filled with a saline solution containing 4 per cent pig blood (the haematocrit level of the blood was measured by a clinical co-oximeter so that 4 per cent blood $= 43.2 \mu M$ total haemoglobin)
with 0.75 per cent Intralipid for scattering. The background chromophore concentrations and scattering for the gel were determined by imaging the phantom in its homogeneous state and using the mean from the reconstructed NIR images, ignoring contributions close to the boundary. The inclusion had a contrast of nearly 2:1 in total haemoglobin, with respect to the background, and was expected to have 100 per cent oxygen saturation and water. The scattering images were expected to be almost homogeneous since 0.75 per cent Intralipid was measured to be similar in scattering quantitatively to the background gelatin in the phantom. Amplitude and phase data were collected from this heterogeneous phantom and image reconstruction was carried out using both the conventional technique of separate wavelength reconstruction as well as the spectrally constrained procedure. Figure 3a shows expected images for the five NIR parameter images, followed by the images recovered from the conventional technique (figure 3b) and the spectrally constrained reconstruction (figure 3c).

The images obtained by the spectrally constrained reconstruction are qualitatively much smoother and more accurate than those from the conventional technique. The simultaneous usage of six wavelengths (649–850 nm) of measurements along with the spectral priors makes the inverse problem better posed, and this, along with the parameter reduction procedure, provides the smoothness in the spectral images. The conventional technique images have more spatial artefacts, including higher crosstalk between oxyhaemoglobin and water, which have similar spectral behaviour, resulting in underestimation of total haemoglobin in the anomaly and saturation of water.

Figure 3. Comparison of images obtained by applying the spectrally constrained direct chromophore reconstruction and the conventional technique of separate wavelength optical properties recovery on measurements from a gelatin phantom with a 25 mm inclusion. The gelatin phantom contained whole blood and TiO₂ for scattering and the inclusion was filled with 4% pig blood and 0.75% Intralipid in buffered saline. (a) The expected images are shown for (i) [HbT] (µM), (ii) oxygen saturation S\textsubscript{T}O₂(%), (iii) water H\textsubscript{2}O(%), (iv) scattering amplitude \(a\) and (v) scattering power \(b\), along with (b) conventional technique images and (c) spectral method images.
A set of reasonable physiological constraints can also be added by fixing the water to be less than 100 per cent and the scattering parameters to be non-negative and to have an upper limit of $6.2 \, \text{mm}^{-1}$. These constraints are applied only at the boundary of their respective ranges, which occurs only in cases of very noisy data. In the majority of reconstructions, the constraints do not come into play, as the updated values at each iteration lie well within the acceptable range. The approach is easily extendable to additional wavelengths and chromophores, as has recently been shown (Wang et al. 2006; Davis et al. 2007). The use of a large number of wavelengths in NIR imaging can help to optimize the inverse problem and are especially useful for systems where the magnitude of measurement noise is large. However, use of a large number of wavelengths also increases the size of the computational problem and thus the memory requirement, which becomes more important for large, complicated models. It has been demonstrated that using an optimized set of a small number of wavelengths based on the residual and conditioning of the problem is computationally more efficient and can provide unique images using continuous-wave measurements (Eames et al. 2008).

(e) Use of prior information

(i) Hard prior

It is possible to segment the forward model (mesh) by appropriately labelling all the nodes within each region. In general, given structural segmentation into $n$ regions, we reconstruct for single values of $\mu_0$ within each region. We apply a matrix transformation to $J$, such that

$$\tilde{J} = JK,$$

where the dimensions of $\tilde{J}$ are the number of measurements times the number of regions ($NM \times NR$, where NR is the total number of regions). We call $K$ the a priori matrix

$$K = \begin{bmatrix}
R_1 & R_2 & \cdots & R_n \\
k_{1,1} & k_{1,2} & \cdots & k_{1,n} \\
k_{2,1} & k_{2,2} & \cdots & k_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
k_{j,1} & k_{j,2} & \cdots & k_{j,n}
\end{bmatrix}, \quad \text{where } k_{\xi,\eta} = \begin{cases}
1, & \xi \in R_\eta \\
0, & \xi \not\in R_\eta
\end{cases}.$$  

(3.15)

In effect, we produce a new Jacobian matrix, $\tilde{J}$, where we have added together all elements from the columns corresponding to similar regions. We then solve

$$\tilde{a} = (\tilde{J}^T \tilde{J})^{-1} \tilde{J}^T b,$$

where the dimensions of the solution update vector, $\tilde{a}$, are $2 \times NR$. In the update process, we apply

$$a = K \tilde{a}.$$  

(3.17)
Note that regularization is not typically required in solving matrix equation (3.16) because \( NR \ll NM \), and therefore the Hessian is a small well-conditioned matrix. However, in the presence of noise, it may be desirable to add some form of regularization into equation (3.16) to achieve a stable solution.

The advantage of using a hard prior in image reconstruction is that the total number of unknowns is reduced dramatically, making the problem better posed. However, one drawback of such a scheme is that spatial resolution is now limited by the size of the defined regions and its stability is highly dependent on the accuracy of the \( a priori \) information. Additionally, using this scheme, only bulk homogeneous values for each region may be reconstructed, with the loss of spatial information within each specified region.

(ii) \textit{Soft prior}

In order to add the spatial constraint, the minimization functional (equation (3.1)) is modified to include a penalty term for \( a priori \) information of tissue structure, given by

\[
(\hat{\mu}_a, \hat{\kappa}) = \arg \min_{\mu_a, \kappa} ||(y^* - F(\mu_a, \kappa))|| + \beta ||L(\mu - \mu_0)||. \quad (3.18)
\]

where \( \beta \) is the regularizing term for spatial prior and \( L \) is a matrix generated using spatial information, acting on the solution \( \mu \). Typically \( L \) is derived from anatomical imaging modalities such as magnetic resonance imaging (MRI) and here it is given by

\[
L_{i,j} = \begin{cases} 
-1/n & \text{region}_i = \text{region}_j \\
1 & i = j \\
0 & \text{region}_i \neq \text{region}_j \\
& i \neq j 
\end{cases} \quad (3.19)
\]

where \( i \) and \( j \) are points within a region and \( n \) is the total number of unknowns in a given region.

The \( L \) matrix links all the nodes in a particular tissue type (glandular or fatty) so that a second differential operator is approximated within each region. This is similar to the total variation minimization approach, which allows sharp boundaries to exist while providing flexibility to encode these boundaries from MRI information. Each node in the model (mesh) is labelled according to the region, or tissue type, with which it is associated (in the MR image). For the \( i \)th node in region \( R \), \( L_{i,i} = 1 \). When nodes \( i \) and \( j \) are in the same region, \( L_{i,j} = -1/n \), where \( n \) is the total number of nodes within region \( R \), otherwise \( L_{i,j} = 0 \). Applied with the spectral prior, the final matrix equation is

\[
a = (J^T J + \beta L^T L)^{-1} J^T b. \quad (3.20)
\]

Soft prior information has been extensively used in a number of experimental and clinical settings and has been found to be much more robust in the presence of uncertainty in prior information (Brooksby et al. 2003, 2005a, b).
As an example of the use of both spatial and spectral \textit{a priori} information, a case study using a combined NIR–MRI imaging system is presented to estimate the properties of normal breast tissue. An FEM mesh generated from the MRI in figure 4 \textit{a} was used in the reconstruction. For each mesh point, figure 4 \textit{b} was associated with a greyscale intensity in the co-registered MRI and was classified as representing either glandular or adipose tissue. Figure 4 \textit{c} shows the tissue properties estimated with the four procedures of (i) no prior information except outer boundary, (ii) spatial soft prior, (iii) spectral constraints and (iv) spatial and spectral constraints. The images obtained using an unconstrained reconstruction are noisy and exhibit boundary artefacts. The spatial priors act on these images, making them smoother, but preserve the trends in chromophore and scattering quantification. For example, the scattering power shows a decrease in the glandular tissue (row ii) similar in value to that obtained without priors (row i). Previous studies suggest that glandular tissue has a higher number density of scatterers, and may therefore have a greater scattering power than fat. Hence, the results from the spatially constrained reconstruction, while appearing smoother, may be misleading. The scattering power image obtained by the application of the spectrally constrained method (row iii) is more quantitatively acceptable. Including the spatial priors within this spectral method (row iv) produces the most intuitively appealing image for this parameter by also showing
the layered structure of the breast. We observed elevated \([\text{Hb}_T]\) (25 : 13 \(\mu\text{M}\)), water (91 : 49\%) and scattering power (1.0 : 0.5) in glandular relative to adipose tissue using the combined priors, which matches the higher degree of vascularization expected.

4. Conclusions

The development of model-based image reconstruction for DOT has been extensive, and has provided the means for developing and testing the clinical application of this non-invasive technique. The overall theory of model-based image reconstruction, as applied to DOT, has been presented with special attention given to the use of the diffusion approximation. Image reconstruction in DOT is a nonlinear problem, and although temporal imaging techniques can be employed to look at changes of optical parameters as a function of time, most imaging techniques have relied on the use of absolute, ‘static’ nonlinear image reconstruction algorithms. The theory behind Newton-like image reconstruction, together with the details of the calculation of the sensitivity functions, has been given. The concept of spectral image reconstruction, whereby functional information can be derived ‘directly’, has also been outlined. Using spectral image reconstruction has been shown to be less prone to noise within the measured NIR signal as well as having improved quantitative accuracy in the recovered contrast (Srinivasan et al. 2005b). There is however some debate on the number of wavelengths required to provide unique solutions to the chromophore and scattering properties to be recovered (Eames et al. 2008). Finally, the use of prior information in DOT has been outlined, demonstrating that it is possible to use such information either to reduce the number of imaging parameters, or to regularize the optimization problem appropriately to enhance spatial resolution.

Two modelling and image reconstruction packages are available for the interested reader to download and use:

TOAST is a software suite for image reconstruction in optical diffusion tomography. It contains binary command line tools for numerical modelling of light transport and parameter recovery.

http://web4.cs.ucl.ac.uk/research/vis/toast/

NIRFAST is a Matlab-based toolbox that also includes capabilities for multimodal NIR imaging, fluorescence and bioluminescence imaging.

http://www.cs.bham.ac.uk/~dehghan/research/NIRFAST.php

References


*Phil. Trans. R. Soc. A* (2009)


Phil. Trans. R. Soc. A (2009)


