

Quantitative spectroscopic diffuse optical tomography of the breast guided by imperfect *a priori* structural information

Gregory Boverman¹, Eric L Miller¹, Ang Li², Quan Zhang³,
Tina Chaves³, Dana H Brooks¹ and David A Boas³

¹ Department of Electrical and Computer Engineering, Northeastern University, 302 Stearns Hall, Boston, MA 02115, USA

² Beckman Laser Institute, University of California, Irvine, 1002 Health Sciences Road, East Irvine, CA 92612, USA

³ Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 Thirteenth Street, Suite 2301, Charlestown, MA 02129, USA

E-mail: gboverma@ece.neu.edu

Received 13 April 2005, in final form 5 July 2005

Published 11 August 2005

Online at stacks.iop.org/PMB/50/3941

Abstract

Spectroscopic diffuse optical tomography (DOT) can directly image the concentrations of physiologically significant chromophores in the body. This information may be of importance in characterizing breast tumours and distinguishing them from benign structures. This paper studies the accuracy with which lesions can be characterized given a physiologically realistic situation in which the background architecture of the breast is heterogeneous yet highly structured. Specifically, in simulation studies, we assume that the breast is segmented into distinct glandular and adipose regions. Imaging with a high-resolution imaging modality, such as magnetic resonance imaging, in conjunction with a segmentation by a clinical expert, allows the glandular/adipose boundary to be determined. We then apply a two-step approach in which the background chromophore concentrations of each region are estimated in a nonlinear fashion, and a more localized lesion is subsequently estimated using a linear perturbational approach. In addition, we examine the consequences which errors in the breast segmentation have on estimating both the background and inhomogeneity chromophore concentrations.

1. Introduction

Near-infrared (NIR) light is beginning to show great promise as a means of noninvasively probing tissue. NIR light is able to penetrate several centimetres through the body, and the differences between the spectra of oxy-haemoglobin (HbO) and deoxy-haemoglobin (HbR)

make spatially localized functional imaging of the body's haemodynamics possible. In addition, H₂O and lipids have spectral peaks in this region, in principle making possible the imaging of these chromophores as well (Cubbedu *et al* 1999). In contrast to imaging modalities such as x-rays and MRI, NIR imaging can be accomplished using non-ionizing radiation and low-cost, potentially portable, electronic components. However, imaging the body's chromophore distributions, which we will refer to here as diffuse optical tomography (DOT), is made quite difficult by the highly turbid nature of NIR light propagation within tissue, with photons generally experiencing many scattering events in their paths.

Near-infrared spectroscopy and imaging has been applied to a number of biomedical applications (Boas *et al* 2001a), including functional brain imaging (Chance *et al* 1998, Franceschini *et al* 2000), monitoring of strokes (Vernieri *et al* 1999), neonatal haemodynamics (Hebden *et al* 2002) and the imaging of breast tumours (Dehghani *et al* 2003, Franceschini *et al* 1997, Grosenick *et al* 1999). In the domain of breast imaging, DOT, by virtue of its ability to directly measure physiologically significant parameters, may prove to be a valuable adjunct to x-ray mammography, a technology that is currently prone to a high rate of false-positive tumour detections (Elmore *et al* 1998, Banks *et al* 2004, Fletcher and Elmore 2003).

In DOT, tissue is illuminated by NIR light sequentially at a number of source locations, generally coupled to the body by means of fibre-optic components. In frequency-domain DOT, the light is RF modulated, producing diffuse photon density waves (DPDWs) within the body. The resulting attenuation and phase shifts are measured at a number of detector positions, with photomultiplier tubes (PMTs), avalanche photodiodes (APDs) or charged-coupled devices (CCDs) used for light detection. The significant problems of sensor calibration are described elsewhere (Boas *et al* 2001b, Oh *et al* 2002). The measured amplitude and phase shifts (with respect to a theoretically known background) are used to reconstruct the optical properties, specifically the absorption and reduced scattering coefficient, $\mu_a(\lambda)$ and $\mu'_s(\lambda)$, at a given wavelength of light, λ . This reconstruction is typically accomplished using either linearized methods, which make a number of assumptions with respect to inhomogeneity size and/or contrast, or through the use of nonlinear optimization approaches (i.e. gradient-descent or Newton-based), which can be very computationally costly.

Although DOT holds a great deal of promise for breast-cancer screening, it presents very difficult technical problems. The most significant limitations of DOT imaging are the smoothing and nonlinear properties of the light propagation model, the consequences of which are the ill-posedness and nonlinearity of the inverse problem, respectively. This ill-posedness can be ameliorated by means of regularization, but this introduces the further difficulty of regularization operator and regularization parameter selection.

Early work in this field concentrated on reconstruction of the medium's optical properties for a given wavelength of light (Jiang *et al* 1996, 1997, Milstein *et al* 2002). Recently, researchers have begun to combine measurements taken at a number of wavelengths, reconstructing an image at each wavelength and using a least-squares fit in order to estimate the distribution of chromophores within the body (McBride *et al* 1999, Pogue *et al* 2001).

Most recently, a method has been proposed to directly image the chromophore concentrations within the body (Gaudette *et al* 1999, Hillman 2002, Li *et al* 2004, 2005) without first estimating the optical properties. Simulation results have shown a reduction in cross-talk and improvement in resolution made possible by this 'direct' form of spectroscopic imaging as compared to the 'indirect' approach previously reported in the literature.

The issue of the heterogeneity of breast tissue, with respect to tissue optical properties, has only recently begun to be examined. One study (Shah *et al* 2004) directly examined this question, making use of NIR frequency-domain spectroscopy at seven wavelengths for 31 patients. Significant, reproducible, spatial heterogeneity was observed for haemoglobin

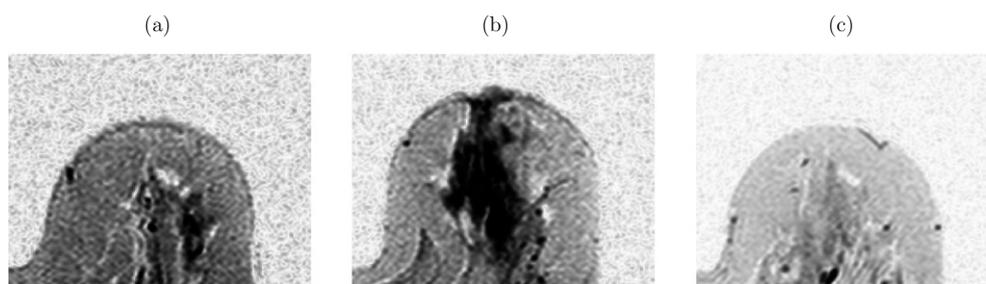


Figure 1. Slices of a fat-suppressed MRI scan at several depths. (a) $z = 1.2$ cm, (b) $z = 2.4$ cm and (c) $z = 4.0$ cm.

concentration, H_2O concentration and lipid concentration, with the degree of heterogeneity being somewhat age dependent. As this research made use of spectroscopy rather than diffuse optical imaging, the question of how the heterogeneity within the breast is spatially distributed has not yet been clearly answered. A study (Cubbedu *et al* 2000), making use of time-domain data, also found significant differences between measurements of total haemoglobin, oxygen saturation, and amounts of H_2O and lipids made at different points on the surface of the breast, as well as differences between reflectance and transmission measurements. Another study (Shah *et al* 2001), making use of frequency-domain data for four wavelengths over a wide range of modulation frequencies, found variations in breast haemoglobin concentration, μ'_s , and breast water content to be correlated with age. These variations were postulated to be a result of the breast's changing composition, and specifically with the atrophy of the glandular tissue as a woman ages. A larger study (Cerussi *et al* 2001), using measurements at seven wavelengths, confirmed these results and found breast lipid content to be correlated to age as well. Two studies (Srinivasan *et al* 2003, Durduran *et al* 2002) investigating the bulk properties of the breast found a correlation between body mass index (BMI) and total haemoglobin concentration. It therefore seems reasonable to postulate that there are significant optical differences between the glandular and adipose tissues present in the female breast, although this assumption has not been definitively validated by empirical investigations.

In this paper, we assume that the breast is composed of distinct glandular and adipose regions with piecewise constant optical properties, and that the boundary between them can be determined by concurrent imaging with a higher spatial resolution imaging modality, such as x-ray imaging, ultrasound or magnetic resonance imaging (MRI). The existence of these regions, whose size and properties vary as a woman ages and, most likely, at different points in the menstrual cycle, is clear from MRI scans, as shown in figure 1. In our simulations, we examine the effect on reconstruction quality of significant differences between glandular and adipose tissues in blood volume and in lipid concentration.

A number of papers have addressed the use of prior information in the solution of the inverse problem. Researchers (Pogue and Paulsen 1998) showed that haemodynamic imaging of a rat cranium could be improved by including anatomical information from a high-resolution MRI scan. Li *et al* (2003) examines the linear inverse problem with a spatial constraint, incorporating the spatial prior by means of a second regularization parameter in the region in which the anomaly is believed to reside. Simulation results showed that the suggested multiple regularization parameter selection algorithm is robust to incorrect prior information. Others (Hero *et al* 1999) introduced an algorithm to make use of noisy information about the shape of a region's boundary in estimating the radionuclide uptake within the region. A number of studies (Huang *et al* 2003, Brooksby *et al* 2003) have combined NIR imaging

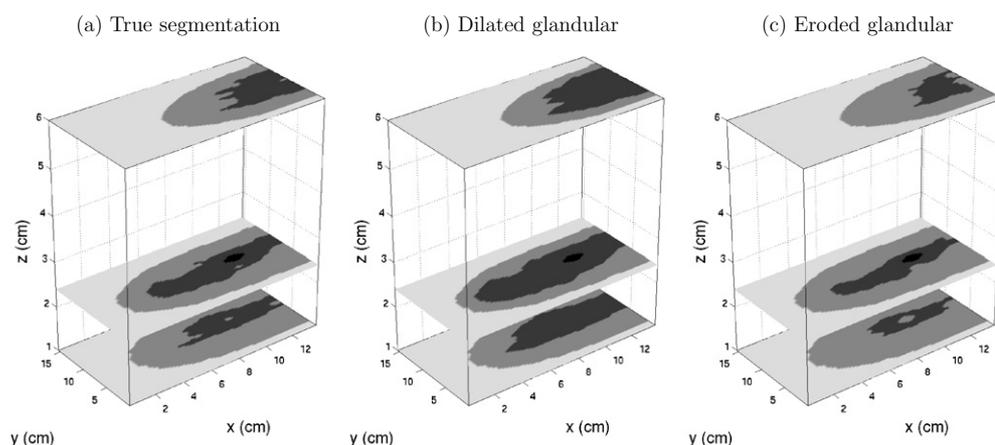


Figure 2. Breast geometries used in the simulation. (a) Geometry derived from manually segmented MRI image, (b) geometry with a dilated glandular region and (c) geometry with an eroded glandular region.

with other modalities and have quantified the improvement in anomaly characterization, using measurements at a single wavelength, achieved by means of incorporating prior spatial information from coregistered MRI and ultrasound images. The use of prior spatial and spectroscopic prior information has been reported (Intes *et al* 2004), as well as a Bayesian approach to the inclusion of *a priori* anatomical information (Guyen *et al* 2005).

Here, we study the importance of accurately modelling the background structure to localizing a tumour and characterizing its spectroscopic properties. To this end, information from other imaging modalities can be used to demarcate the boundary between the two tissues, and the estimation of the spectral properties of each tissue type is reduced to a low-dimensional nonlinear optimization problem. In a realistic imaging situation, our information about the breast structure is likely to be imperfect, due to a number of factors: noise in the underlying x-ray or MRI measurements and image artefacts in the reconstructions for these imaging modalities, segmentation errors and errors in the simultaneous co-registration of several three-dimensional images.

In this work, we examine, both qualitatively and quantitatively, the sensitivity of the tumour spectroscopic reconstruction with respect to errors in our prior knowledge of the breast background structure. Four cases are explored: assuming that the breast is homogeneous, assuming a dilated glandular region, assuming an eroded glandular region and knowledge of the true background structure. In the dilated and eroded cases, the estimated glandular segmentations are greater and smaller in volume, respectively, than the true glandular region, as shown in figure 2. Given our segmentation, we first estimate the overall chromophore concentrations of the two regions directly, using a Gauss–Newton optimization approach, and then estimate the inhomogeneity using a linear, perturbational algorithm. Thus, errors in estimation of the tumour’s spectroscopic properties will occur in two places: error in computing the mismatch between the expected and actually measured data and error in the sensitivity matrices used in the linear perturbational model.

Our results qualitatively show that even incorrect information about the background structure is useful in quantifying tumour chromophore concentrations and in reducing imaging artefacts, particularly for perturbations in oxy-haemoglobin and deoxy-haemoglobin (HbR and HbO), which are assumed to be much greater than the differences in blood volume and oxygen

saturation between the two tissue types. As we have assumed that glandular and adipose tissues differ more sharply in their lipid content, reconstruction of perturbations of the same order is more problematic. We also quantify the reconstruction error as a function of regularization parameter, both within and outside the tumour, for all four cases, assuming perturbations in individual chromophores within the tumour. We find that large errors in the estimation of one chromophore's perturbation tend to be relatively isolated within that chromophore's reconstruction.

2. Methods

2.1. Forward modelling

In the simulations that follow, we make use of the diffusion approximation (Arridge 1999, Boas 1996) to model the steady-state intensity and phase distribution due to a modulated source, making use of a zero partial flux boundary condition (Aronson 1995, Haskell *et al* 1994) to model the air-tissue interface:

$$\begin{cases} -\nabla \cdot D(\mathbf{r})\nabla\phi(\mathbf{r}, \omega) + \left(\mu_a(\mathbf{r}) + \frac{j\omega}{v}\right)\phi(\mathbf{r}, \omega) = S_0(\mathbf{r}, \omega) & \mathbf{r} \in \Omega \\ \frac{1}{2}R_\phi\phi(\mathbf{r}) - D(\mathbf{r})R_j\hat{n} \cdot \nabla\phi(\mathbf{r}) = 0 & \mathbf{r} \in \bar{\Omega} \setminus \Omega \end{cases} \quad (1)$$

where the diffusion coefficient, $D(\mathbf{r})$ is $\frac{1}{3\mu'_s(\mathbf{r})}$, $\mu_a(\mathbf{r})$ is the absorption coefficient and $\mu'_s(\mathbf{r})$ is the reduced scattering coefficient. The photon fluence is $\phi(\mathbf{r}, \omega)$, a function of position, \mathbf{r} , and modulation frequency, ω . The isotropic source intensity is $S_0(\mathbf{r}, \omega)$ and v is the speed of light in tissue. The dependence of all the parameters on the wavelength, λ , is implicit. The spatial extent of the diffusive region is Ω . The Fresnel boundary reflection coefficients for the photon density and current are R_ϕ and R_j , respectively, and \hat{n} is the direction normal to the boundary.

In a typical DOT experiment, tissue is illuminated at M source positions, and, for each source, measurements are made at N detector positions. This procedure is repeated for a number of wavelengths. It is also possible to make use of a number of modulation frequencies, but we will assume that all experiments use a modulation frequency of 70 MHz. We further make the assumption that the absorption at each point in space is a linear combination of C chromophores whose spectra have been experimentally determined (Li *et al* 2005):

$$\mu_a(\mathbf{r}, \lambda) = \sum_{j=1}^C \epsilon_j(\lambda)c_j(\mathbf{r}) \quad (2)$$

where $\epsilon_j(\lambda)$ is the extinction coefficient of chromophore j for light of wavelength λ and $c_j(\mathbf{r})$ is the chromophore concentration. Specifically, we assume that the absorption at each wavelength is due to the following four chromophores: oxy- and deoxy-haemoglobin, H_2O and lipids. Motivated by Mie scattering theory, the reduced scattering coefficient is modelled as follows:

$$\mu'_s(\mathbf{r}, \lambda) = a(\mathbf{r})\lambda^{-b(\mathbf{r})} \quad (3)$$

where $a(\mathbf{r})$ represents the scattering amplitude and $b(\mathbf{r})$ is related to average particle size.

In our simulations, the following wavelengths of light have been used: 685, 750, 808, 830, 906 and 980 nm. The first four wavelengths are intended to quantify concentrations of HbR and HbO, and the last two are situated near peaks in the spectra for H_2O and lipids, and are intended to discern their concentrations. The discretized absorptions and concentrations have the following relationship:

$$\boldsymbol{\mu}_a = \mathbf{E}\mathbf{c} \quad (4)$$

where $\boldsymbol{\mu}_a$ and \mathbf{c} are stacked vectors of absorption and concentration values for each voxel: $\boldsymbol{\mu}_a = [\boldsymbol{\mu}_a(685) \ \boldsymbol{\mu}_a(750) \ \boldsymbol{\mu}_a(808) \ \boldsymbol{\mu}_a(830) \ \boldsymbol{\mu}_a(906) \ \boldsymbol{\mu}_a(980)]^T$, $\mathbf{c} = [\mathbf{c}_{\text{HbO}} \ \mathbf{c}_{\text{HbR}} \ \mathbf{c}_{\text{H}_2\text{O}} \ \mathbf{c}_L]^T$. The matrix \mathbf{E} transforms an image in terms of chromophore concentrations into an image in terms of absorptions at each wavelength.

2.2. Inverse problem solution

As the inverse problem for DOT is generally severely underdetermined and ill-posed, some form of regularization, or prior information, is necessary in order to stabilize the inversion procedure. Thus, the inverse problem is often reduced to minimization of the following functional (Arridge 1999, Milstein *et al* 2002, Li *et al* 2003):

$$L = \arg \min_{\mathbf{c}, \mathbf{a}} \|\mathbf{y} - \mathbf{h}(\mathbf{c}, \mathbf{a})\|_{\Sigma_n}^2 + r(\|\mathbf{c}\|_2^2 + \|\mathbf{a}\|_2^2) \quad (5)$$

where Σ_n is the measurement noise covariance matrix (assumed to be diagonal in this paper), $\mathbf{y} = [\mathbf{y}^{685} \ \mathbf{y}^{750} \ \mathbf{y}^{808} \ \mathbf{y}^{830} \ \mathbf{y}^{906} \ \mathbf{y}^{980}]^T$, and $\mathbf{h} = [\mathbf{h}^{685} \ \mathbf{h}^{750} \ \mathbf{h}^{808} \ \mathbf{h}^{830} \ \mathbf{h}^{906} \ \mathbf{h}^{980}]^T$, \mathbf{h}^λ being the hypothesized measurements at wavelength λ , which can be generated by means of the finite-difference method. The regularization parameter is denoted by r . We note that, in this work, we are making use of an identity regularization functional.

The adjoint method (Arridge 1995, 1999) can be used to compute, with fairly low computational cost, a first-order linear perturbational approximation for the change in the model solution, \mathbf{h} , with respect to changes in each voxel's chromophore concentrations and scattering amplitude relative to a known background:

$$\mathbf{h} \approx \mathbf{h}_0(\mathbf{c}_0, \mathbf{a}_0) + \mathbf{J}(\mathbf{c}_0, \mathbf{a}_0)\mathbf{p} \quad (6)$$

where

$$\mathbf{J} = [\text{block diag}(\mathbf{J}_{\mu_a}^{685} \ \mathbf{J}_{\mu_a}^{750} \ \mathbf{J}_{\mu_a}^{808} \ \mathbf{J}_{\mu_a}^{830} \ \mathbf{J}_{\mu_a}^{906} \ \mathbf{J}_{\mu_a}^{980})\mathbf{E} \ | \ \mathbf{J}_a] \quad (7)$$

and

$$\mathbf{p} = \begin{bmatrix} \mathbf{c}_p \\ \mathbf{a}_p \end{bmatrix} \quad (8)$$

with $\mathbf{c}_p = [\mathbf{c}_{p,\text{HbO}} \ \mathbf{c}_{p,\text{HbR}} \ \mathbf{c}_{p,\text{H}_2\text{O}} \ \mathbf{c}_{p,L}]$ being the vector of concentration perturbations, and \mathbf{a}_p being the vector of perturbations in scattering amplitude. The vector $\mathbf{h}_0(\mathbf{c}_0, \mathbf{a}_0)$ is the *incident field*, which is dependent on an assumed background chromophore distribution \mathbf{c}_0 and the background scattering amplitude distribution \mathbf{a}_0 . The Jacobian matrix at wavelength λ with respect to absorption perturbations is $\mathbf{J}_{\mu_a}^\lambda(\mathbf{c}_0, \mathbf{a}_0)$, where the dependence on \mathbf{c}_0 and \mathbf{a}_0 will be assumed to be implicit, and the Jacobian matrix with respect to scattering amplitude is

$$\mathbf{J}_a = [(\mathbf{J}_a^{685})^T \ (\mathbf{J}_a^{750})^T \ (\mathbf{J}_a^{808})^T \ (\mathbf{J}_a^{830})^T \ (\mathbf{J}_a^{906})^T \ (\mathbf{J}_a^{980})^T]^T. \quad (9)$$

In order to simplify the exposition of our simulations, we assume in our simulations that all tissues have the same scattering amplitude. The least-squares solution to the linear inverse problem is then

$$\hat{\mathbf{p}} = \mathbf{J}^T \Sigma_n^{-\frac{1}{2}} (\Sigma_n^{-\frac{1}{2}} \mathbf{J} \mathbf{J}^T \Sigma_n^{-\frac{1}{2}} + r\mathbf{I})^{-1} \Sigma_n^{-\frac{1}{2}} (\mathbf{y} - \mathbf{h}_0(\mathbf{c}_0, \mathbf{a}_0)). \quad (10)$$

In equation (10), we note that an error in our knowledge of the background concentrations can cause a systematic reconstruction error, by introducing an error in our estimate of the incident field, $\mathbf{h}_0(\mathbf{c}_0, \mathbf{a}_0)$, and by causing inaccuracies in the Jacobian, \mathbf{J} . It is these systematic errors that are analysed in the simulation results that follow.

The direct spectral reconstruction approach to the inverse problem does impose an additional computational cost compared to more traditional processing schemes. Rather than

solving for the absorption and scattering perturbations at each wavelength independently, we solve for all chromophore concentrations simultaneously, making use of measurements at all wavelengths. In the case of an identity regularization function and our particular configuration of sources and detectors, the solution is still computationally feasible with non-specialized computing hardware, and, if a more complex regularization function is employed, iterative methods, such as the LSQR algorithm (Paige and Saunders 1982), may be employed.

In this paper, we make use of the generalized cross-validation (Golub *et al* 1979, Golub and von Matt 1997) approach to regularization parameter selection, which is based on the Bayesian criterion of minimizing expected reconstruction error.

2.3. Estimation of background chromophore concentrations

Before the linear inverse problem can be solved, we must first estimate the *incident field*, $\mathbf{h}_0(\mathbf{c}_0, \mathbf{a}_0)$, in equation (10). Here, we assume that the breast is piecewise constant, with distinct glandular and adipose regions, and we must estimate the chromophore concentrations in these regions. We solve this optimization problem using the Gauss–Newton method, solving for the parameter vector $\mathbf{g} = [\mathbf{g}_{1,\text{HbR}} \ \mathbf{g}_{1,\text{HbO}} \ \mathbf{g}_{1,\text{H}_2\text{O}} \ \mathbf{g}_{1,L} \ \mathbf{g}_{1,a} \ \mathbf{g}_{2,\text{HbR}} \ \mathbf{g}_{2,\text{HbO}} \ \mathbf{g}_{2,\text{H}_2\text{O}} \ \mathbf{g}_{2,L} \ \mathbf{g}_{2,a}]^T$, where the adipose and glandular regions are labelled region 1 and region 2, respectively.

Now, we define the following indicator vectors, assuming a lexicographic ordering of the breast voxels in the forward problem:

$$(\mathbf{i}_j)_k = \begin{cases} 1, & \text{voxel } k \text{ is in region } j \\ 0, & \text{voxel } k \text{ is not in region } j. \end{cases} \quad (11)$$

The Jacobian with respect to changes in background chromophore concentrations is computed as follows:

$$\mathbf{J}_g = \mathbf{J} \begin{bmatrix} \mathbf{i}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{i}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{i}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i}_2 \end{bmatrix} \quad (12)$$

where \mathbf{J} is the Jacobian with respect to voxelwise changes in concentration and scattering amplitude.

Estimation of the parameter vector, \mathbf{g} , then proceeds by means of the Gauss–Newton algorithm, with a cubic line search.

2.4. Reconstruction error analysis

In this section, we analyse the image reconstruction error in the case of Gaussian noise, decomposing it into deterministic and stochastic components. Firstly, we assume that $\mathbf{y} = \mathbf{y}_m + \mathbf{n}$, where \mathbf{y}_m is known and \mathbf{n} is a zero-mean vector of Gaussian noise, with covariance matrix Σ_n . For the sake of clarity, we assume in the analysis that follows that \mathbf{h}_0 , \mathbf{n} , \mathbf{J} and \mathbf{y} have been pre-multiplied by $\Sigma_n^{1/2}$.

The expected mean-squared reconstruction error for a given regularization parameter can be written as follows:

$$\mathbf{C} = E \|\mathbf{M}(\mathbf{c}_0 + \mathbf{p}) - \mathbf{M}(\hat{\mathbf{c}}_0 + \mathbf{J}^T(\mathbf{J}\mathbf{J}^T + r\mathbf{I})^{-1}(\mathbf{y}_m + \mathbf{n} - \mathbf{h}_0))\|_2^2 \quad (13)$$

where \mathbf{c}_0 is the true background, \mathbf{p} is the true perturbation, $\hat{\mathbf{c}}_0$ is the estimated background and \mathbf{M} is a selection matrix, focusing on a particular region of interest.

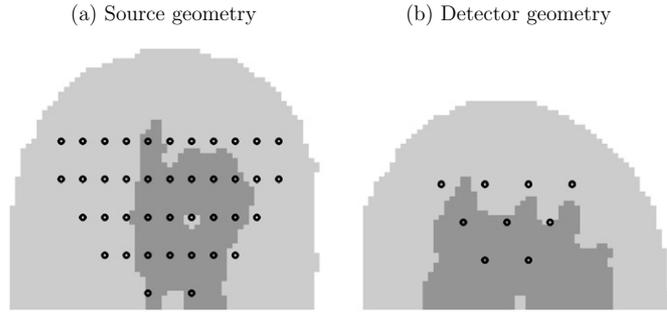


Figure 3. Source and detector configurations used in the simulations. (a) Source configuration and (b) detector configuration.

Assuming that the noise is statistically uncorrelated with \mathbf{c}_0 , \mathbf{p} , $\hat{\mathbf{c}}_0$, \mathbf{y}_m and \mathbf{y}_0 , we can show that

$$\mathbf{C} = \mathbf{C}_{\text{bias}} + \mathbf{C}_{\text{variance}} \quad (14)$$

where

$$\mathbf{C}_{\text{bias}} = \|\mathbf{M}(\mathbf{c}_0 + \mathbf{p}) - \mathbf{M}(\hat{\mathbf{c}}_0 + \mathbf{J}^T (\mathbf{J}\mathbf{J}^T + r\mathbf{I})^{-1} (\mathbf{y}_m - \mathbf{h}_0))\|_2^2 \quad (15)$$

$$\mathbf{C}_{\text{variance}} = \text{tr}((\mathbf{J}\mathbf{J}^T + r\mathbf{I})^{-1} \mathbf{J}\mathbf{M}^T \mathbf{M}\mathbf{J}^T (\mathbf{J}\mathbf{J}^T + r\mathbf{I})^{-1}). \quad (16)$$

Thus, mean-squared error is comprised of two components: \mathbf{C}_{bias} , the *bias*, represents the deterministic component of the error and $\mathbf{C}_{\text{variance}}$, the *variance*, represents the stochastic component of the error, which originates from the amplification of the measurement noise. In the results that follow, we will be primarily interested in the bias that results from making incorrect assumptions about the breast background structure.

3. Results and discussion

An MRI image of a healthy breast, shown in figure 1, was manually segmented into air, glandular and adipose regions and interpolated onto a 2.5 mm uniform grid as shown in figure 2(a). The simulation results presented here were generated by a finite-difference forward model with 2 mm uniform grid spacing. A tumour of diameter 2 cm was simulated, centred at $(x = 13 \text{ cm}, y = 10 \text{ cm}, z = 2.5 \text{ cm})$. The breast thickness is 7.5 cm, and we made use of simulated 70 MHz frequency-domain measurements, using 40 sources and 9 detectors, with a geometry as shown in figure 3. We assumed the following chromophore concentrations for the adipose and glandular regions, respectively, 20 μM HbO, 5 μM HbR, 30% H₂O, 40% lipid and 22 μM HbO, 5.5 μM HbR, 40% H₂O, 10% lipid. The true chromophore distributions of breast tissue are yet to be determined experimentally, but we have chosen values to be consistent with published results (Shah *et al* 2004). The following chromophore concentrations were simulated for the tumour, which has a greater blood volume and a lower oxygenation than the surrounding tissue (Fishkin *et al* 1997): 40 μM HbO, 15 μM HbR, 20% H₂O, 10% lipid. The simulated tumour has a 50% perturbation in H₂O content, and has the same lipid concentration as the surrounding glandular tissue. Amplitude-dependent Gaussian noise with a variance of 1% was added to the measurements, and an identically distributed Gaussian phase uncertainty with a variance of 1° was assumed (Zhang *et al* 2001).

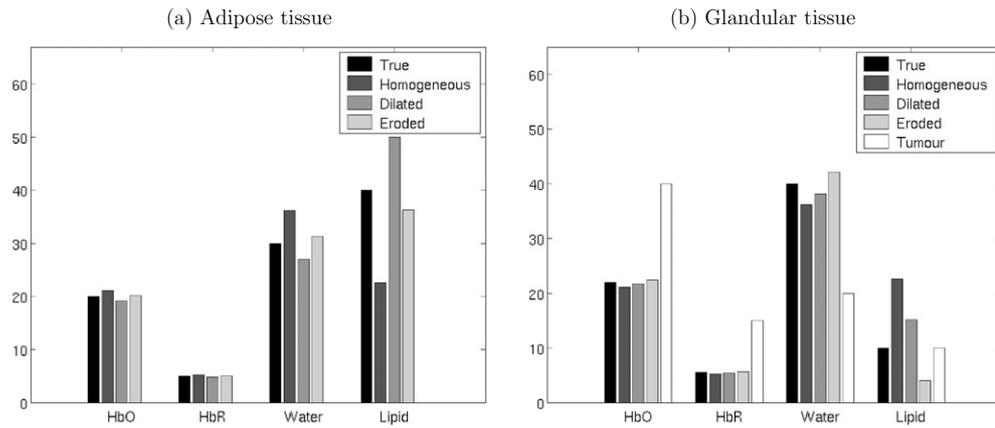


Figure 4. Estimated background chromophore concentrations. (a) Adipose tissue and (b) glandular tissue, with the true chromophore concentrations of the tumour shown for the purpose of comparison.

Table 1. Estimates of background chromophore concentrations.

	HbO (μM)	HbR (μM)	Water (%)	Lipid (%)
Adipose region				
True concentrations	20	5	30	40
Homogeneous assumed	21.09	5.29	36.21	22.63
Dilated glandular assumed	19.18	4.81	27.07	49.96
Eroded glandular assumed	20.20	5.05	31.22	36.30
Glandular region				
	HbO (μM)	HbR (μM)	Water (%)	Lipid (%)
True concentrations	22	5.5	40	10
Homogeneous assumed	21.09	5.29	36.21	22.63
Dilated glandular assumed	21.72	5.43	38.14	15.23
Eroded glandular assumed	22.47	5.64	42.13	4.02

We estimated background chromophore concentrations and reconstructed perturbations for four cases: true knowledge of the breast background structure, an assumed homogeneous background structure, a dilated glandular region and an eroded glandular region. In the latter two cases, we increased and decreased the volume of the glandular part of the breast using erosion and dilation operators, as shown in figures 2(b) and (c). The true volume of the glandular tissue is 355 cm^3 and that of the adipose tissue is 817 cm^3 . We dilated the glandular tissue to a volume of 544 cm^3 and eroded it to a volume of 239 cm^3 .

The results of the background concentration estimation for the four cases are shown in table 1. These results are graphically illustrated in figure 4. We note the fairly large variation in the estimates, particularly in the concentrations of H_2O and Lipids, for which there is a greater difference between the two regions. In figure 4(b), the true tumour chromophore concentrations are also shown for the purpose of comparison.

In each of the above cases, we computed the Jacobian using these estimated background properties and solved the linear inverse problem to reconstruct an image of the perturbation, making use of equation (10). The regularization parameter was chosen using the GCV criterion. The true solution, for the $z = 2.5 \text{ cm}$ slice is shown in figure 5, and the

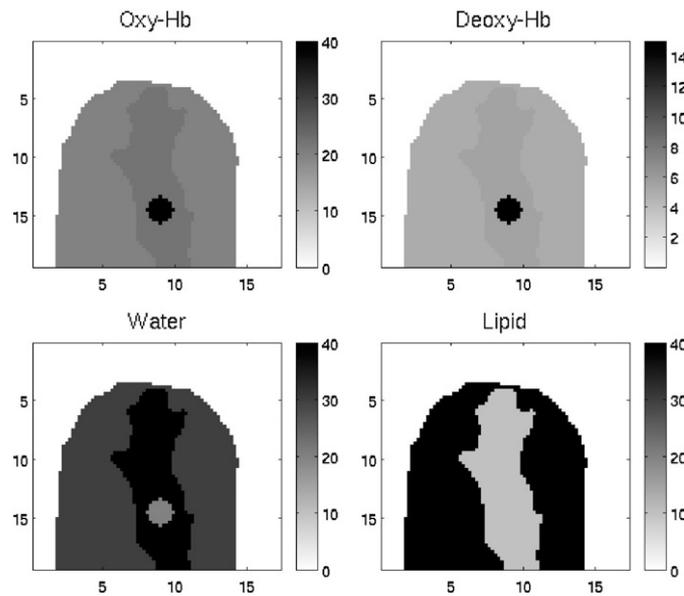


Figure 5. True solution for absorption concentrations.

reconstructions, for the same slice, can be seen in figure 6. We note that an excellent spectroscopic reconstruction of a tumour in an inhomogeneous background can be obtained, provided that the background structure is known with a high degree of accuracy, as shown in figure 6(a). On the other hand, the reconstruction assuming a homogeneous background is plagued by large image artefacts, as shown in figure 6(b). In comparing the dilated and eroded glandular cases, as shown in figures 6(b) and (c), with the homogeneous case, we note that even imperfect prior information is useful in reducing image artefacts. In addition, the reconstructions for HbR and HbO in figures (c) and (d) are superior to the reconstruction for H₂O concentration, which is to be expected given the larger difference between the adipose and glandular tissue in their H₂O content. We believe that the superiority of reconstruction in the case of the assumed eroded glandular structure as compared to the case of the assumed dilated structure is not a general result, but rather is an effect introduced by particular cases studied in our simulations.

In order to further quantify the effect of imperfect background information on spectroscopic image reconstructions, we have computed the image reconstruction bias for each of the four cases, using equation (15), for perturbations in individual chromophores only. This bias is essentially computed by generating a linear reconstruction for each value of the regularization parameter, without any noise added to the measurements (zero-mean random noise manifests itself only in the variance). By using a perturbation in a single chromophore only, we can directly examine the reduction in bias as a function of the regularization parameter and also examine cross-talk, which is the leakage of a perturbation in one chromophore into the reconstruction for another chromophore.

Figure 7 shows the per-voxel reconstruction bias for a 10 μ M perturbation in HbO only as a function of the regularization parameter. The square root of the mean bias for voxels within and outside the tumour is shown using solid and dashed lines, respectively. Each quadrant shows the bias for all four chromophore reconstructions, for a given background structure assumption. In the case of perfect background structural information, as shown in figure 7(a),

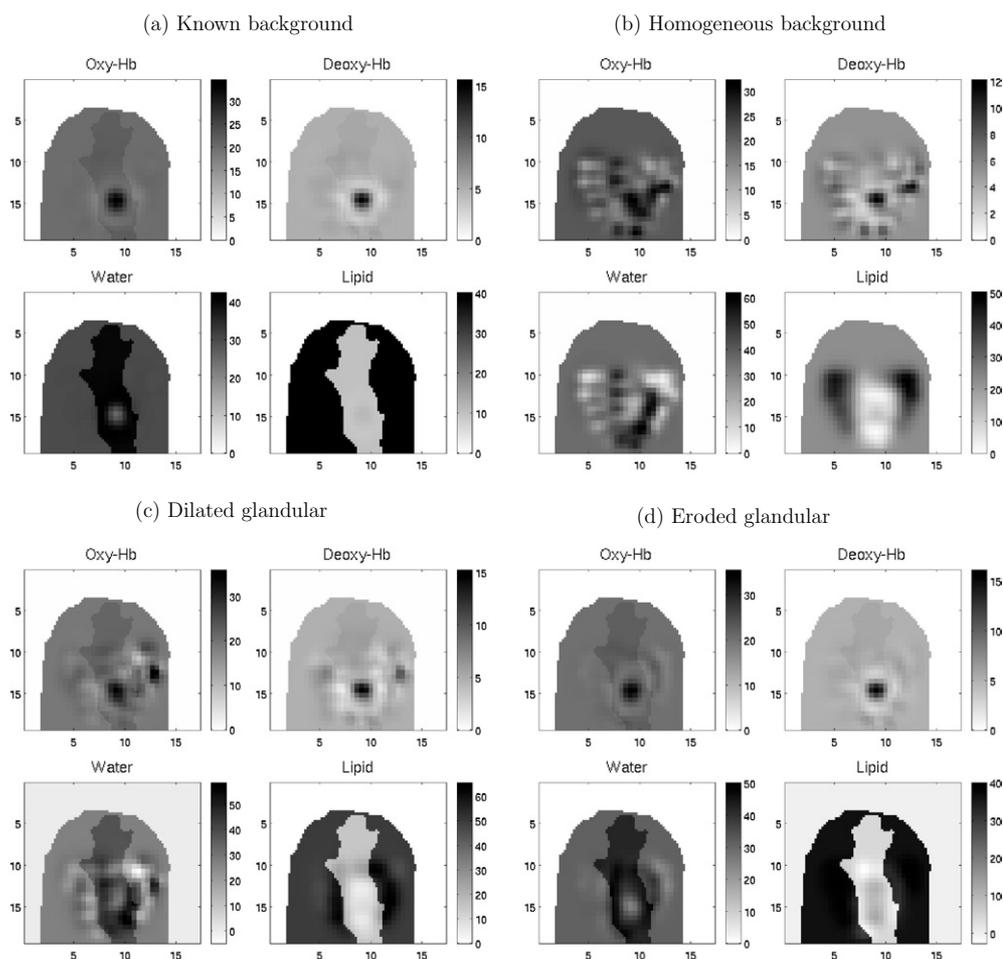


Figure 6. Linear reconstructions of chromophore concentrations given a number of assumptions about the breast background. (a) Reconstruction assuming accurate knowledge of background structure, (b) reconstruction assuming a homogeneous background, (c) reconstruction assuming a dilated glandular region and (d) reconstruction assuming an eroded glandular region.

for the HbO reconstruction and voxels within the tumour we see the following behaviour: for a large value of the regularization parameter, the bias is, as expected, $10 \mu\text{M}$, meaning that no perturbation is reconstructed where one is expected. The bias decreases to a minimum value as we reduce the level of regularization. After this point, discretization error in the iterative, numerical solution of the forward model tends to be amplified, resulting in image artefacts. Strictly speaking, this numerical quantization noise is deterministic, but different forward model implementations will produce quantitatively very different quantization noise realizations. Thus, our bias calculations for small values of the regularization parameter are particular to our finite-difference forward model implementation and grid resolution, although these results are likely to be qualitatively accurate for any numerically computed forward model. An additional reason for the bias tending to increase exponentially, even in the case of perfect background structural information, as we decrease the regularization parameter beyond a certain point, is that the actual measurements used in the inversion make use of a fully nonlinear model, while our inversion procedure assumes that a first-order linearization

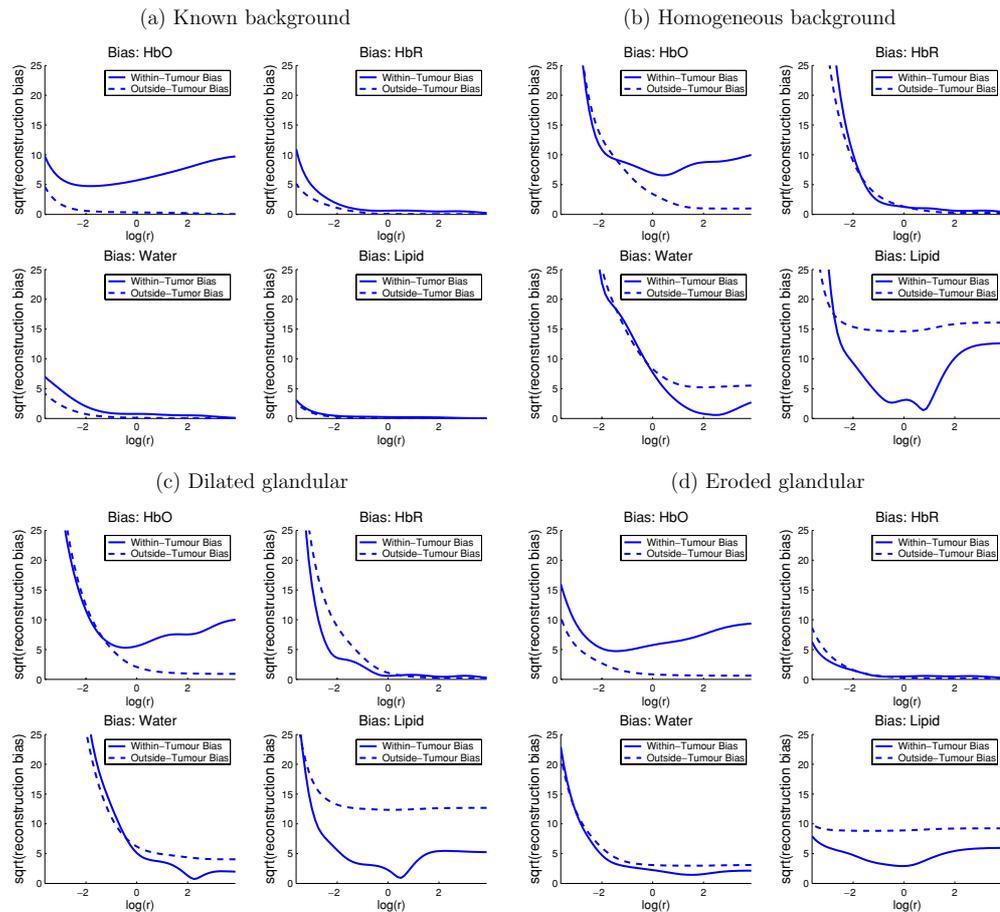


Figure 7. Reconstruction bias for an HbO-only perturbation as a function of the regularization parameter. (a) Reconstructions bias when the background is known exactly, (b) bias when a homogeneous background is assumed, (c) bias assuming a dilated glandular region and (d) bias assuming an eroded glandular region. Note that the bias is given in the units that are natural for each chromophore (i.e. μM for HbO and HbR and percentage for H₂O and lipids). The horizontal axis shows the logarithm, base 10, of the regularization parameter.

is accurate. Thus, this aspect of the model mismatch is amplified. The bias in the HbO reconstruction does not reach a minimum of zero because the regularization used here is not sufficient to overcome the inherent blurring in the forward model. Other regularization schemes may allow for a further reduction in bias.

For the HbO reconstruction, the value of the regularization parameter at which we see a minimum in the bias varies approximately two orders of magnitude, depending on the background structure assumption that is being made. Clearly, the lowest achievable bias is attained when the background structure is known perfectly, as shown in quadrant (a). We also see in this quadrant that even with perfect information, at the point that minimizes bias in HbO there is still a small degree of spectroscopic cross-talk between the HbO reconstruction and the reconstructions for the other three chromophores. In quadrant (b), where we assume a homogeneous background, the minimum achievable bias is somewhat higher than in the case of perfect background information, as expected, and, at the value of the regularization

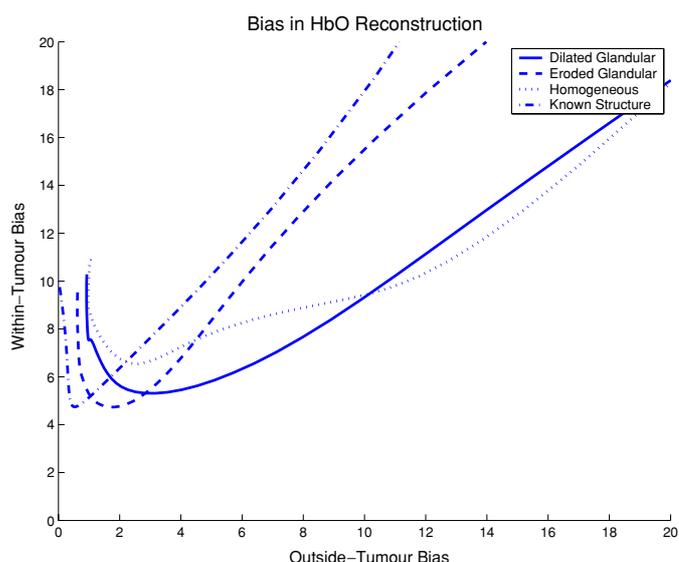


Figure 8. Within-tumour bias versus outside-tumour bias, for a perturbation in HbO only, for each of the background structure assumptions, over a wide range of regularization parameters.

parameter where the bias is minimized in HbO, we are beginning to see considerable bias in the H₂O reconstruction. In figures 7(c) and (d), in which we assume dilated and eroded glandular structures, respectively, it is somewhat surprising that the minimal achievable bias in HbO is not much different than in the case of perfect information, but, at this minimal point, we are beginning to see considerable bias in the reconstructions for the other three chromophores. In all four cases, the GCV algorithm returned a choice of regularization parameter between 0.1 and 1.0.

The square root of the mean bias for voxels outside the tumour is shown using dashed lines, essentially quantifying the degree to which image artefacts are reconstructed by our inversion approach. Given our noise model and background assumptions, the H₂O reconstruction shows the greatest sensitivity to background modelling errors, as the value of the regularization parameter at which the image artefacts tend to increase exponentially is consistently higher than for the remaining chromophores. It is also interesting to observe that the value of the regularization parameter at which the bias seems to increase exponentially differs for the bias within the tumour as compared to the bias outside the tumour. This indicates that when there is background modelling error, we have a trade-off between tumour reconstruction accuracy and background artefact suppression. We note the differences in regularization parameters at which point the artefacts tend to increase exponentially for the four background structure assumptions, which clearly show the greater presence of image artefacts as our knowledge of the background medium worsens. Figures 7(b)–(d) show that a large bias in the estimate of background lipid concentration can be mitigated, to a certain extent, by the image reconstruction. It may be that more sophisticated approaches to the inverse problem solution can reduce the image artefact bias somewhat. For example, penalizing the image gradient rather than the image norm may tend to produce solutions which are more concentrated in a single region.

In order to analyse the minimal bias that is possible, both within and outside the tumour, we plot the reconstruction bias within the tumour versus the bias for voxels outside the tumour in figure 8, over a wide range of regularization parameters. The optimal inversion scheme

would intersect the origin, minimizing both the bias within the tumour and the image artefacts. For the reasons discussed above, even in the case of perfect background structural knowledge, we cannot achieve this goal, but our analysis shows the extent to which the best achievable reconstruction result varies as our knowledge of the background structure worsens. We note that assuming dilated or eroded background structures increases the within-tumour bias only modestly, given an optimal means for choosing the regularization parameter, but these assumptions increase the bias for voxels outside the tumour more substantially. As expected, the worst achievable result is attained when we assume a homogeneous background structure, in which case the minimum achievable bias within the tumour is significantly higher than for the other three cases, and the level of image artefacts is somewhat worse as well.

We have repeated the analysis above for perturbations in HbR, H₂O and lipids, obtaining results that are qualitatively very similar to the case of an HbO perturbation. We have also computed the variance of the image reconstruction for the four cases of spatial prior information mentioned above, as we vary the regularization parameter. As expected from equation (16), in which the background structure estimate does not appear, the variance does not noticeably depend on the prior information used in the reconstruction.

4. Conclusion and future work

We have described a two-step algorithm for combining a high-resolution segmentation of the breast into distinct tissue types with a spectroscopic DOT reconstruction. In our approach, the background chromophore concentrations in the glandular and adipose regions are estimated using nonlinear Gauss–Newton iterations and the perturbation is estimated with a linearized spectroscopic reconstruction. If the segmentation is known with precision, we are able to both localize and characterize the tumour with very high accuracy.

Given that the true structure of the breast, as evidenced by MRI scans, is likely to be heterogeneous, we have explored the importance of this heterogeneity in linearized reconstructions. Specifically, we have reconstructed a perturbation when the breast background is assumed to be heterogeneous, though highly structured. We have examined the case where the breast is assumed to be homogeneous and cases in which the assumed glandular tissue volume is ‘dilated’ or ‘eroded’ with respect to the true structure. In each of these cases, we have first estimated the background chromophore concentrations and then computed a perturbation assuming this estimate of the background optical properties. Our results clearly show that even incorrect knowledge of the background can be useful in localizing anomalies, but that the mismatch between our assumptions and reality can introduce severe bias in image reconstructions.

It is interesting to note that incorrect spatial prior structural information seems to have an asymmetric effect on the image reconstruction. This is evidenced by the fact that the minimum achievable bias within the simulated tumour increases more slowly than the minimum achievable bias outside the tumour (i.e. the level of image artefacts) as the quality of our prior structural information worsens. It may be possible to use this tendency in image reconstruction algorithms, utilizing approaches specifically designed to focus the solution of the inverse problem within a region of limited spatial support. Further research will also examine whether the prior information about the background structure can be applied in a probabilistic manner.

The analysis that we have conducted here gives us a framework for understanding the effect of our assumptions on reconstruction bias. As we improve our algorithms, the statistical approach shown here can quantify the reconstruction accuracy gains that are achievable, potentially justifying the additional computational cost. It is quite clear that, if there is

significant optical heterogeneity within the breast, some sort of structural background prior information may be necessary if we are able to obtain quantitatively accurate images using linearized methods. Future work will address whether this situation is improved if we make use of fully nonlinear optimization methods, and will also address the question of regularization parameter selection in the case of model mismatch. We also plan to examine how the results reported here generalize to the case where not only the constituent components of the breast are heterogeneous, but its scattering coefficient is spatially variant as well.

Acknowledgments

GB, ELM and DHB acknowledge the support of CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers programme of the National Science Foundation (award number EEC-9986821). DAB, TC, QZ and AL were supported by the National Institutes of Health.

References

- Aronson R 1995 Boundary conditions for diffusion of light *J. Opt. Soc. Am. A* **12** 2532–9
- Arridge S R 1995 Photon-measurement density functions: Part 1. Analytical forms *Appl. Opt.* **34** 7395–409
- Arridge S R 1999 Optical tomography in medical imaging *Inverse Problems* **15** R41–93
- Banks E *et al* 2004 Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the million women study: cohort study *Br. Med. J.* **329** 477–503
- Boas D A 1996 Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications *PhD Thesis* University of Pennsylvania, Philadelphia, PA
- Boas D A, Brooks D H, Miller E L, Dimarzio C A, Kilmer M, Gaudette R J and Zhang Q 2001a Imaging the body with diffuse optical tomography *IEEE Signal Process. Mag.* **18** 57–75
- Boas D A, Gaudette T J and Arridge S R 2001b Simultaneous imaging and optode calibration with diffuse optical tomography *Opt. Express* **8** 263–70
- Brooksby B A, Dehghani H, Pogue B W and Paulsen K D 2003 Near-infrared (NIR) tomography breast image reconstruction with *a priori* structural information from MRI: algorithm development for reconstructing heterogeneities *IEEE J. Sel. Top. Quantum Electron.* **9** 199–209
- Cerussi A E, Berger A J, Bevilacqua F, Shah N, Jakubowski D, Butler J, Holcombe R F and Tromberg B J 2001 Sources of absorption and scattering contrast for near-infrared optical mammography *Acad. Radiol.* **8** 211–8
- Chance B, Anday E, Nioka S, Zhou S, Hong L, Worden K, Li C, Ovetsky M T, Pidikiti D and Thomas R 1998 A novel method for fast imaging of brain function, non-invasively, with light *Opt. Express* **2** 411–23
- Cubbedu R, D'Andrea C, Pifferi A, Taroni P, Torricelli A and Valentini G 2000 Effects of the menstrual cycle on the red and near-infrared optical properties of the human breast *Photochem. Photobiol.* **72** 383–91
- Cubbedu R, Pifferi A, Taroni P, Torricelli A and Valentini G 1999 Noninvasive absorption and scattering spectroscopy of bulk diffusive media: an application to the optical characterization of human breast *Appl. Phys. Lett.* **74** 874–6
- Dehghani H, Pogue B W, Poplack S P and Paulsen K D 2003 Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results *Appl. Opt.* **42** 135–45
- Durduran T, Choe R, Culver J P, Zubkov L, Holboke M J, Giammarco J and Yodh A G 2002 Bulk optical properties of healthy female breast tissue *Phys. Med. Biol.* **47** 2847–61
- Elmore J G, Barton M B, Moceri V M, Polk S, Arena P I and Fletcher S W 1998 Ten-year risk of false positive screening mammograms and clinical breast examinations *N. Engl. J. Med.* **338** 1089–96
- Fishkin J B, Coquoz O, Anderson E R, Brenner M and Tromberg B J 1997 Frequency-domain photon migration measurements of normal and malignant tissue optical properties in a human subject *Appl. Opt.* **36** 10–20
- Fletcher S W and Elmore J G 2003 Mammographic screening for breast cancer *N. Engl. J. Med.* **348** 1672–80
- Franceschini M A, Moesta K T, Fantini S, Gaida G, Gratton E, Jess H, Mantulin W W, Seeber M, Schlag P M and Kaschke M 1997 Frequency-domain techniques enhance optical mammography: initial clinical results *Proc. Natl Acad. Sci. USA* **94** 6468–73
- Franceschini M A, Toronov V, Filiaci M, Gratton E and Fantini S 2000 On-line optical imaging of the human brain with 160-ms temporal resolution *Opt. Express* **6** 49–57
- Gaudette R, Brooks D, Dimarzio C and Boas D 1999 Multi-wavelength diffuse optical tomography: experimental results *Proc. Conf. on Lasers and Electro-Optics* pp 253–4

- Golub G H, Heath M H and Wahba G 1979 Generalized cross-validation as a method for choosing a good ridge parameter *Technometrics* **21** 215–22
- Golub G H and von Matt U 1997 Generalized cross-validation for large-scale problems *J. Comput. Graph. Stat.* **6** 1–34
- Grosenick D, Wabnitz H, Rinneberg H, Moesta K T and Schlag P M 1999 Development of a time-resolved optical mammograph and first *in-vivo* applications *Appl. Opt.* **38** 2927–43
- Guyen M, Yazici B, Intes X and Chance B 2005 Diffuse optical tomography with *a priori* anatomical information *Phys. Med. Biol.* **50** 2837–58
- Haskell R C, Svaasand L O, Tsay T T, Feng T, McAdams M S and Tromberg B J 1994 Boundary conditions for the diffusion equation in radiative transfer *J. Opt. Soc. Am. A* **11** 2727–41
- Hebden J C, Gibson A, Yusof R, Everdell N, Hillman E M C, Delpy D T, Arridge S R, Austin T and Meek J H 2002 Three-dimensional optical tomography of the premature infant brain *Phys. Med. Biol.* **47** 4155–66
- Hero A O, Piramuthu R, Fessler J A and Titus S R 1999 Minimax emission computed tomography using high-resolution anatomical information and B-spline models *IEEE Trans. Inf. Theory* **45** 920–38
- Hillman E M C 2002 Experimental and theoretical investigation of near-infrared tomographic imaging methods and clinical investigations *PhD Thesis* University of London, London, UK
- Huang M, Xie T, Chen N G and Zhu Q 2003 Simultaneous reconstruction of absorption and scattering maps with ultrasound localization: feasibility study using transmission geometry *Appl. Opt.* **42** 4102–14
- Intes X, Maloux C, Guven M, Yazici B and Chance B 2004 Diffuse optical tomography with physiological and spatial *a priori* constraints *Phys. Med. Biol.* **49** N155–63
- Jiang H B, Paulsen K D, Osterberg U L and Patterson M S 1997 Frequency-domain optical image reconstruction in turbid media: an experimental study of single-target detectability *Appl. Opt.* **36** 52–63
- Jiang H, Paulsen K D, Osterberg U L, Pogue B W and Patterson M S 1996 Optical image reconstruction using frequency-domain data: simulations and experiments *J. Opt. Soc. Am. A* **13** 253–66
- Li A, Boverman G, Zhang Y, Brooks D, Miller E L, Kilmer M E, Zhang Q, Hillman E M C and Boas D A 2005 Optimal linear inverse solution given multiple priors in diffuse optical tomography *Appl. Opt.* **44** 1948–56
- Li A *et al* 2003 Tomographical optical breast imaging guided by three-dimensional mammography *Appl. Opt.* **42** 5181–90
- Li A, Zhang Q, Culver J P, Miller E L and Boas D A 2004 Reconstructing chromosphere concentration images directly by continuous-wave diffuse optical tomography *Opt. Lett.* **29** 256–8
- McBride T O, Pogue B W, Gerety E D, Poplack S R, Osterberg U L and Paulsen K D 1999 Spectroscopic diffuse optical tomography for the quantitative assessment of hemoglobin concentration and oxygen saturation in breast tissue *Appl. Opt.* **38** 5480–90
- Milstein A B, Oh S, Reynolds J S, Webb K J and Bouman C A 2002 Three-dimensional Bayesian optical diffusion tomography with experimental data *Opt. Lett.* **27** 95–7
- Oh S, Milstein A B, Millane R P, Bouman C A and Webb K J 2002 Source–detector calibration in three-dimensional Bayesian optical diffusion tomography *J. Opt. Soc. Am. A* **19** 1983–93
- Paige C C and Saunders M A 1982 LSQR: an algorithm for sparse linear equations and sparse least squares *ACM Trans. Math. Softw.* **8** 43–71
- Pogue B W and Paulsen K D 1998 High-resolution near-infrared tomographic imaging simulations of the rat cranium by the use of *a priori* magnetic resonance structural information *Opt. Lett.* **23** 1716–8
- Pogue B W, Poplack S P, McBride T O, Wells W A, Osterman K S, Osterberg U L and Paulsen K D 2001 Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast *Radiology* **218** 261–6
- Shah N, Cerussi A E, Eker C, Espinoza J, Butler J, Fishkin J, Hornung R and Tromberg B 2001 Noninvasive functional optical spectroscopy of human breast tissue *Proc. Natl Acad. Sci. USA* **98** 4420–5
- Shah N, Cerussi A E, Jakubowski D, Hsiang D, Butler J and Tromberg B J 2004 Spatial variations in optical and physiological properties of healthy breast tissue *J. Biomed. Opt.* **9** 534–40
- Srinivasan S, Pogue B W, Jiang S, Dehghani H, Kogel C, Soho S, Gibson J J, Tosteson T D, Poplack S P and Paulsen K D 2003 Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured *in vivo* by near-infrared breast tomography *Proc. Natl Acad. Sci.* **100** 12349–54
- Vernieri F, Rosato N, Pauri F, Tibuzzi F, Passarelli F and Rossini P M 1999 Near infrared spectroscopy and transcranial Doppler in monohemispheric stroke *Eur. Neurol.* **41** 159–62
- Zhang Q, Brukilacchio T J, Gaudett T, Wang L, Li A and Boas D A 2001 Experimental comparison of using continuous-wave and frequency-domain diffuse optical imaging systems to detect heterogeneities *Optical Tomography and Spectroscopy of Tissue IV (SPIE, Bellingham, WA)* ed B Chance, R R Alfano, B J Tromberg, M Tamura and E M Sevick-Muraca, pp 219–38