3D Shape-Based Imaging for Diffuse Optical Tomography

Misha E. Kilmer

Tufts University Dept. of Mathematics

Medford, MA 02155

Eric L. Miller

Northeastern University ECE Dept.

Boston, MA 02115

Alethea Barbaro

Tufts University

Medford, MA 02155

David Boas

Athinoula A. Martinos Center for Biomedical Imaging

Massachusetts General Hospital

Harvard Medical School

Charlestown, MA

Abstract

In this work, we present a shape-based approach to 3D image reconstruction from diffuse optical data. Our approach differs from others in the literature in that we jointly reconstruct object and background characterization and localization simultaneously, rather than sequentially process for optical properties and postprocess for edges. The key to the efficiency and robustness of our algorithm is in the model we propose for the optical properties of the background and anomaly: we use a low-order parameterization of the background and another for the interior of the anomaly and we use an ellipsoid to describe the boundary of the anomaly. This model has the effect of regularizing the inversion problem and provides a natural means of including physical properties determined via another imaging modality for additional stability. A Gauss-Newton-type algorithm with linesearch is implemented to solve the underlying non-linear least squares problem and thereby determine the coefficients of the parameterizations and the descriptors of the ellipsoid. Numerical results show the effectiveness of this method.

1. Introduction

In diffuse optical imaging, modulated, near-infrared light is transmitted into the body from an array of detectors placed on the surface of the region to be imaged[3]. The diffuse optical systems then measure the photon fluence that results from the interaction (scattering and absorption) of photons by that region of the body. The goal is to use the diffuse optical data to reconstruct an image of the space varying optical absorption and reduced scattering coefficients in the region of interest. These physical parameters are directly related to the hemodynamic state of the tissue. Because the hemodynamics are directly impacted by e.g. the presence of a tumor in the case of breast imaging[1] or activity in the cortex for functional brain mapping[2], diffuse optical tomography (DOT) offers the hope of providing significant insight in a non-invasive manner into these and related problems.

From an information extraction perspective, the DOT problem amounts to the determination of the subsurface structure of the tissue given a limited quantity of fluence data obtained at the air-tissue interface. Reconstructing a 3D image of the absorption and scattering coefficients is an example of a discrete ill-posed problem, meaning that the quality and accuracy of the reconstruction are especially sensitive to noise in the data and other unmodeled physical effects. Generally, some form of regularization must be used to desensitize the problem to noise. To further complicate matters, the problem is nonlinear, and for our application, breast tissue imaging, the discrete problem is highly under-determined.

The ultimate goal of breast imaging is obviously the localization and characterization of tumorous regions in the tissue. In the case of DOT imaging, the existence of a tumor is signaled more by a change in the optical absorption coefficient than the reduced scattering coefficient. Hence in this paper, we concern ourselves only with the recovery of this parameter

although our methods can easily be generalized to problems wherein scattering perturbations are also of interest. The task of isolating these localized perturbations is significantly complicated by the fact that the "nominal" breast is far from a homogeneous, known background in which the perturbations are embedded. The classical approach to overcoming this difficulty is to voxelate the region of interest and treat the values of the absorption coefficient in each voxel as an unknown. Stable recovery of the thousand or even millions of voxel values is facilitated though the use of Tikhonov-type regularization scheme. Tumor localization is accomplished by post-processing the resulting reconstruction to segment background from anomalous regions.

The nonlinearity of the problem necessitates that a forward solver be incorporated directly into the inversion method. Hence, such an imaging approach can quickly become intractable when the number of voxels becomes large. Further, such methods require the selection of a regularization parameter to lessen the effects of the noise, and choosing the right parameter is far from simple. Some methods proposed in the literature for 3D DOT imaging are not applicable in our case because they assume data can be taken across the entire surface of the object to be imaged (see, for example[4]), whereas in breast tissue imaging, the data can only be collected over a limited surface area.

Newer techniques, presented only in a two-dimensional scenario, treat the absorption coefficient as if it were a piecewise constant function. Under such an assumption, two classes of methods have been proposed. Some assume the absorption values in the background and anomalies are known, and seek only to find a low dimensional descriptor of the absorption perturbation boundaries[5]. Other proposed schemes first generate a pixel-based reconstruction of the 2D region of interest. The resulting image is post-processed to obtain an initial

indication as to the number of anomalies, their boundaries and contrasts. Iterative schemes are employed to improve the estimates of the boundaries and their absorption values[6]. In addition to the fact that these methods have only been examined for two dimensional geometries, the robustness of the piecewise constant (PWC) assumption to data consistent with a spatially inhomogeneous background and tumor has not been examined for 3D problems involving limited data.

In this paper, we present an approach to the tumor localization and characterization problem which is designed directly for full three dimensional diagnostic scenarios and which is sufficiently flexible to allow us to begin the quantitative exploration of the effects of unknown inhomogeneities specifically in the non-tumor regions of the breast. Our work here was motivated by our past experience with "anomaly characterization" problems for two dimensional inverse scattering type of problems. The method we presented in[10] for the 2D case is closely related to the method in[6]. In[10], we used B-spline basis functions to define a low-order parameterization of the boundary of the anomaly; therefore, only a few unknowns had to be recovered to determine the boundary. At the same time, we explicitly modeled the variations in the background as a weighted linear combination of a small (< 5) collection of space-varying basis functions. The problem became one of jointly estimating the control-points specifying the boundary and the few expansion coefficients describing the background and anomaly, respectively.

For three dimensional problems, the use of B splines to model surfaces is not at all natural. While one could resort to other, less parametric modeling approaches (e.g. using level sets as in[7]), again three dimensional implementations are quite involved. Thus, in this paper, we consider the estimation of anomalies which have an ellipsoidal shape. Ellipsoids are

5

defined by the (x, y, z) location of their center, the lengths of their three axes, and a set of three rotation angles. Thus estimating a "best fit" ellipsoidal anomaly requires only that we determine these nine geometric parameters as well as a small (less than 10) number of parameters which model the spatial variations of the absorption coefficient over the support of the ellipsoid and over the background; a far less under-determined task than full voxel estimation or even level-set methods.

In practice, tumors are clearly not ellipsoidal; however use of a parametric model as we are advocating is nonetheless advantageous for a number of reasons. As we demonstrate in § 4, the estimation of an ellipsoid can provide important information concerning the location, size, orientation, and contrast of even a non-ellipsoidal object. Such information in and of itself will have medical benefit. Moreover, this level of localization can be used to focus the effort of other, less parametric methods (level sets, constraint imaging methods, etc.) thereby improving their performance and lowering their computational cost. This is a task we reserve for future effort. Finally, and most importantly, by limiting our attention here to ellipsoidal objects, we are more easily able to examine the far more challenging and relevant problem of anomaly localization in the presence of unknown, volumetric, "background" perturbations. If one cannot recover simple shapes under such a scenario, it is hardly worth the effort to look at more elaborate shape-based inversion schemes.

The results in § 4 do in fact demonstrate the utility of this parametric inversion approach. We show through a large number of numerical examples that it is possible to accurately recover the size, location, and orientation of an ellipsoid-shaped object when the model we employ to describe the background (i.e.non-anomalous) variations in the absorption coefficient is not able to reproduce the true spatial variations. Moreover, even when the true object

6

is not itself an ellipsoid, the structure we recover using our inversion approach does provide strong localization of the true anomaly thereby indicating that our shape-based method has the potential for use in and of itself as well as a "focus of attention" preprocessor for other finer grain imaging algorithms. Thus the numerical experiments in this paper demonstrate that this shape-based inversion scheme is robust to the types of modeling errors which will be encountered when processing real sensor data and justify both the further development of this technique and other related geometries inversion schemes.

This paper is organized as follows. In Section 2, we provide the mathematical formulation of the problem. The description of the model and the correspondingly reformulated non-linear inversion problem is provided in Section 3. Extensive numerical results are presented in Section 4 while conclusions and future work are detailed in Section 5.

2. Mathematical Description and Background

As indicated in § 1, the work in this paper is aimed at absorption imaging for breast analysis. Within this context, we are particularly concerned with DOT sensing systems which can be used as an adjunct to traditional mammography. For this class of problems the breast is constrained by a pair of compression plates. Hence, under the assumption that the optodes are located far from the edges of the compressed breast, it is appropriate to employ a slab-geometry to describe the sensing system.

To model these sensing systems, it is well known that the Boltzman transport equation is the best mathematical descriptor of the underlying physical process relating optical properties to photon fluence/flux[8]. However, a more tractable mathematical model, known to be accurate and also widely used, in a medium such as breast tissue, where scattering dominates absorption, is the diffusion equation[8]. In terms of diffusion modeling, one always

7

has the choice of computing using the partial differential equation (PDE) or an equivalent integral equation (IE) formulation. Generally, the sparse matrix structure arising from the discretization of the PDE is preferable to the requirement of handling large dense matrices which is encountered when using an IE approach. As we discuss in a forthcoming paper, however, for the slab geometry of interest here, the IE matrices possess significant structure which can be exploited to partially erase these differences.

To derive this model we let $r = (x, y, z)$ denote an arbitrary point in space. For a source s located at position r_s we have

$$\phi_{tot,s}(r) = \phi_{inc,s}(r) + \int_\Omega G(r, \tilde{r}) \phi_{tot,s}(\tilde{r}) \gamma(\tilde{r}) d\tilde{r}. \tag{1}$$

Here, the incident field at source s, $\phi_{inc,s}$, and G are known. The volume Ω is the region of interest. The G is an appropriately scaled green's function for a slab geometry based on nominal, constant values of absorption and scattering. A method of images approach is used to construct $G(r, \tilde{r})$[9]. The unknown quantities are $\phi_{tot,s}$ the fluence (or flux, depending on location) due to the input at source s, and g, the function representing the perturbation of the absorption about the aforementioned nominal value.

All quantities depend on the intensity modulation frequency of light: that is, there is one such equation for each modulation frequency. For ease of notation, and because our results are presented for data obtained when the light was not intensity modulated, we have suppressed the dependence on frequency.

Discretizing (1), at *any* position r_k, the total fluence/flux due to source s is given by

$$\phi_{tot,s}(r_k) \approx \phi_{inc,s}(r_k) + \Delta \sum_{r_m \in \Omega} G(r_k, r_m)\phi_{tot}(r_m)\gamma(r_m), \tag{2}$$

where Δ is a constant dependent on the grid spacing.

Let n denote the number of voxels and m denote the number of detectors. Let the subvectors of length n, $\phi_{tot,s}^{(1)}$, $\phi_{inc,s}^{(1)}$ have entries $\phi_{tot,s}(r_k)$, $\phi_{inc,s}(r_k)$ where $r_k \in \Omega$. Then by moving the summation to the other side of (2), the equation can be written in matrix vector form by as

$$(I_n - G^{(1)}\text{DIAG}(\gamma))\phi_{tot,s}^{(1)} = \phi_{inc,s}^{(1)}, \tag{3}$$

where I_n denotes the $n \times n$ identity matrix, $\text{DIAG}(v)$ is a diagonal matrix with v on the main diagonal, and $G^{(1)}$ has entries $\Delta G(r_k, r_m)$, where $r_k, r_m \in \Omega$.

Likewise, if we let $\phi_{tot,s}^{(2)}$, $\phi_{inc,s}^{(2)}$ have entries $\phi_{tot,s}(r_k)$, $\phi_{inc,s}(r_k)$ with r_k a detector location, the corresponding matrix-vector equation can be written, according to (2), as

$$\phi_{tot,s}^{(2)} = \phi_{inc,s}^{(2)} + G^{(2)}\text{DIAG}(\gamma)\phi_{tot,s}^{(1)}. \tag{4}$$

Here, $G^{(2)}$ has entries $\Delta G(r_k, r_m)$ where r_k is a detector location but $r_m \in \Omega$. If there are m detectors, the $\phi_{tot,s}^{(2)}$, $\phi_{inc,s}^{(2)}$ are m-length vectors and $G^{(2)}$ is $m \times n$.

The total field due to source s that is available to us for imaging is only that which is collected over the receivers, namely $\phi_{tot,s}^{(2)}$. Since $\phi_{inc,s}^{(2)}$ is known, the available data according

to (4) is

$$
\begin{aligned}
y_s &= \phi_{tot,s}^{(2)} - \phi_{inc,s}^{(2)} \\
&= G^{(2)} \mathrm{DIAG}(\gamma) \phi_{tot,s}^{(1)} \\
&= G^{(2)} \mathrm{DIAG}(\phi_{tot,s}^{(1)}) \gamma \\
&\equiv h_s(\gamma).
\end{aligned}
\tag{5}
$$

But since $\phi_{tot,s}^{(1)}$ is a function of the vector γ through (3) there is a nonlinear relation between the data y_s and the unknown, γ.

If we stack the subvectors $y_s, h_s(\gamma)$, we obtain the model $y = h(\gamma)$. Thus, the problem we should be solving given data y is the nonlinear least squares problem

$$
\min_{\gamma} \| W (y - h(\gamma)) \|_2^2,
\tag{6}
$$

where the weighting matrix W is often taken as the inverse square root of the covariance matrix of any noise corrupting the data. There are several difficulties with the solution to this nonlinear least square problem. First, the solution is sensitive to small changes in y such as noise, even if there are as many equations as unknowns γ_i. Worse, the problem is usually underdetermined (there are n unknowns γ_i and only mk components in the data vector y, where k is the number of sources,) so there are generally many solutions.

Most traditional methods of estimating the desired solution to this underdetermined and ill-posed problem take a Tikhonov approach. That is, the solution is approximated as the

solution to the regularized problem

$$\min_{g} \|W(y - h(\gamma))\|_2^2 + \lambda\Theta(\gamma)$$

where λ is a positive regularization parameter that balances the fidelity to the data with damping the noise effects and $\Theta(\gamma)$ is a regularization operator such as $\|\gamma\|_2^2$. The differences among DOT imaging methods can be summarized as follows:

- Differences in forward model h: Often, h is taken to be linear because the solution of linear equations is considerably faster. However, nonlinear models, based either on the partial differential equation or the integral equation formulation, are more accurate.

- Choice of Θ: A typical choice is $\|L\gamma\|_p^p$ where L is a matrix and $1 < p \le 2$. The regularization term can have the effect of making the solution unique. On the other hand, the model we describe for g does some of the stabilization for us – if the slab is not too thick, no additional Θ term is needed, which eliminates the additional annoying difficulty of choosing a regularization parameter.

- Definition of γ: Most methods consider γ as a vector of unknown values, and therefore the space over which the problem is minimized is huge (R^n). With our formulation of γ, however, we greatly reduce the dimension of the search space.

3. The Model for Perturbation of Optical Absorption

Our modeling approach is founded on the underlying assumption that there is a single localized anomaly (although if we knew we were searching for more than one, we could readily adapt our model and algorithm to find it) somewhere in a tissue medium which itself

is volumetrically heterogeneous with respect to the assumed "constant" background used to construct the Green's function for our model. Now, if we knew *a priori* the contrast of the anomaly for which we were searching and any volumetric perturbations not associated with the anomaly, the problem would become one of boundary localization. Likewise, if we knew the boundary, we could optimize for the contrasts. In this paper, we assume that neither is known and develop a procedure for jointly estimating low order models for the anomaly shape and contrast as well as more "global," non-anomalous volumetric perturbations required to capture the inherent heterogeneity of nominal, non-tumorous breast tissue.

First, we consider how to model absorption inhomogeneities in the background as well as in the anomaly. Assume for a moment that we do know the boundary of the anomaly. Then we can define $S(r)$ as the characteristic function

$$S(r) = \begin{cases} 1 & r \in \text{anomaly} \\ 0 & \text{otherwise} \end{cases}.$$

Next, we assume that the value of the perturbation at a point in the background tissue can be expressed as the linear combination of a small number, say N_b, of known basis functions. We assume the same about the value inside the anomaly, although the number, N_a of basis functions and the type may be different. This is the approach that we took in[10], except we only considered a two dimensional version of the problem. We now define the value of the perturbation function at the point r according to

$$\gamma(r) = S(r) \underbrace{B_a(r)}_{1 \times N_a} \overbrace{\alpha}^{N_a \times 1} + [1 - S(r)] \underbrace{B_b(r)}_{1 \times N_b} \overbrace{\beta}^{N_b \times 1}.$$

Notice that $B_a(r)$ (likewise $B_b(r)$) is a row vector whose entries are the values of the anomaly (background) basis functions at r, and the components in α (β) are the expansion coefficients. We note that formally, we should define γ as the mollified version of the following equation so that it will be differentiable in the desired parameters (see Sections 3 B,3 C), but at our current level of discretization, the algorithm acts as though we are in fact dealing with the differentiable, mollified version.

Upon discretization, the equation becomes

$$\gamma = SB_1\alpha + (I - S)B_2\beta \tag{7}$$

where S now represents a diagonal $n \times n$ matrix with 1's and 0's on the diagonal, I denotes the identity matrix of the same size and

$$B_1 \in R^{n \times N_a}, \quad B_2 \in R^{n \times N_b}, \quad \alpha \in R^{N_a}, \quad \beta \in R^{N_b}.$$

In other words, B_1 is the discretization of $B_a(r)$ over all voxels r_i in the region of interest while B_2 is the discretization of $B_b(r)$.

Thus, if S is known, there are a total of $N_a + N_b$ unknowns that we need in order to define γ. In the remainder, we discuss how to describe the boundary of the anomaly (and therefore define S) in terms of a small number of unknowns.

A. *Describing the Boundary*

In[10], the DOT problem posed was two-dimensional. In 2D, the problem of parameterizing a boundary was handled by a B-spline basis. Specifically, the boundary was a linear combination of a small number of B-spline basis functions, and the so-called control points ((x, y)

coordinate pairs) were the only unknowns. Generalizing this idea to 3D is not straight-forward so we look for a better 3D alternative. We assume that our anomalies resemble either spheres or ellipsoids, which is not unreasonable in the context of breast imaging given the limited spatial resolution possible from DOT data. The benefit of using such shapes as spheres and ellipsoids is that they can be completely described in terms of only a few unknowns.

We begin by characterizing an ellipsoid, and consider the sphere as a special case. To completely characterize an ellipsoid, we need only the lengths of the semi-axes, a centroid location, and a set of orthogonal basis vectors describing the orientation of the axes. In mathematical notation, we need

- A length-3 vector, c, denoting the centroid of the ellipsoid

- A diagonal matrix, $D = \text{DIAG}(d)$, whose (absolute) diagonal entries d_i are the lengths of the semi-axes.

- An orthonormal matrix U whose three columns are the coordinates of the semi-axes. Alternately, we can store three angles from which U can be defined. We call this three-length vector θ.

Thus, a point $r = (x, y, z)^T$ is determined to be on or in the ellipsoid if

$$\|D^{-1}U^T(r - c)\|_2^2 \leq 1. \tag{8}$$

The matrix U is defined from the vector of angles θ by

$$U = \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) & 0 \\ -\sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{bmatrix} \begin{bmatrix} \cos(\theta_3) & \sin(\theta_3) & 0 \\ -\sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9}$$

We note that there are multiple ways to define a rotation matrix from three angles[11]. The one we use here is referred to as the 'x-convention'. We also tried the 'pitch-yaw-roll convention' and found the representation of U makes little difference as far as the performance of the algorithm.

Therefore, given the characterization for an ellipsoid as noted above, one can use (8) to determine if the center of a voxel lies on or in the ellipsoid; if so, the corresponding diagonal entry in S in (7) is assigned a 1, otherwise, it is set to zero.

The discussion simplifies if we decide to use spheres. In this case, D will be a multiple of the identity. So rather than keeping track of a three-vector d, we keep track of a single number (the radius). That is, d will be a scalar. Since for spheres U remains the identity (since rotations need not be considered), we do not need to keep track of angles, either.

B.   Reformulated problem

Thus, S is defined in terms of a small number of unknowns that characterize the ellipsoid. Given the preceding discussion and our model for γ, (6) simplifies, in the case of ellipsoids, to

$$\min_{\alpha,\beta,c,d,\theta} \|W(y - h(\alpha, \beta, c, d, \theta))\|_2^2. \tag{10}$$

In the special case of spheres, the optimization problem further simplifies to

$$\min_{\alpha,\beta,c,d} \|W(y - h(\alpha, \beta, c, d))\|_2^2, \tag{11}$$

where d is scalar. Notice that we have reduced the search space considerably: instead of optimizing over R^n, we are optimizing over the $N_a + N_b$ entries in α and β and the 9 entries in the vectors c, d, θ (4 entries in the spherical case).

Now h is a nonlinear function of the parameters because h is a nonlinear function of γ (5), and γ depends on α, β explicitly and d, c, θ implicitly through S (7). Hence (10), (likewise (11)) is a nonlinear least squares problem in our parameters. (See the remark in previous section regarding the differentiability of γ in d, c, θ.)

C. *Computational Issues*

According to (5) and (3), in order to compute data generated by a given absorption perturbation γ requires the solution of a large linear system (3) at each source. However, for each source, the matrix involved is the same. Further, the matrix is highly structured, making it possible to develop and apply a preconditioned block iterative method to solve over the sources simultaneously: this is discussed further below.

We propose to solve the minimization problems (10), (11) using a Gauss-Newton type of approach. For simplicity, we discuss only the ellipsoidal case here and note that the simplification in the case of spheres is straightforward.

Define the function $\epsilon(\alpha, \beta, c, d, \theta) = W(h(\alpha, \beta, c, d, \theta) - y)$. Then an equivalent formulation

of the problem is (dropping the dependence on the unknowns for ease of notation)

$$\min_{\alpha,\beta,c,d,\theta} \frac{1}{2}\epsilon^T \epsilon.$$

The Gauss-Newton algorithm is an iterative method for determining the solution to nonlinear least squares problems. That is, given an initial guess, say v_0, of the solution vector v (v is given by $v = [\alpha^T, \beta^T, c^T, d^T, \theta^T]^T$), the algorithm produces a sequence of iterates v_k that converge (under certain assumptions[12]) to v. Each search direction, s, computed by the Gauss-Newton approach to this problem is given by the solution to the normal equations

$$(J^T J)s = -J^T \epsilon, \tag{12}$$

where J denotes the Jacobian of ϵ with respect to the unknowns, evaluated at the current estimates, and ϵ is the current value of the residual. The updated guess for the unknowns is given by $v_{k+1} = v_k + s$. We describe our algorithm as "Gauss-Newton-type" rather than as Gauss-Newton since some of the entries in our Jacobian are computed through the use of finite difference approximations and the columns of the Jacobian are only approximately solved for via an iterative method.

Let p denote one of the parameters that describe γ. From (5) we have:

$$\begin{aligned}
\frac{\partial \epsilon_s}{\partial p} &= W\frac{\partial h_s}{\partial p} \\
&= WG^{(2)}(\frac{\partial}{\partial p}\text{DIAG}(\phi_{tot,s}^{(1)})\gamma) \\
&= WG^{(2)}(\text{DIAG}(\gamma)\frac{\partial \phi_{tot,s}^{(1)}}{\partial p} + \text{DIAG}(\phi_{tot,s}^{(1)})\frac{\partial \gamma}{\partial p})
\end{aligned} \tag{13}$$

But from (3) it is easy to show

$$(I_n - G^{(1)}\text{DIAG}(\gamma))\frac{\partial \phi_{tot,s}^{(1)}}{\partial p} = G^{(1)}\text{DIAG}(\phi_{tot,s}^{(1)})\frac{\partial \gamma}{\partial p} \tag{14}$$

It follows that to compute the necessary derivatives we must proceed as follows:

1. Given current γ, compute $\phi_{tot,s}^{(1)}$ from (3).

2. For each parameter p do:

   (a) Compute (or estimate) $\frac{\partial \gamma}{\partial p}$.

   (b) Solve the matrix-vector equation (14).

   (c) Compute the products in (13) to find $\frac{\partial \epsilon_s}{\partial p}$.

Now note that Step 2a involves the same matrix, regardless of the source (and regardless of the parameter, for that matter). Therefore, we can solve for them "simultaneously" if we solve the block equation

$$AX = B, \qquad X, B \text{ are } n \times k,$$

where $A = (I_n - G^{(2)}\text{DIAG}(\gamma))$, X is an $n \times k$ matrix obtained by placing the vectors $\frac{\partial \phi_{tot,s}^{(1)}}{\partial p}$ side-by-side for each of k sources, and B has columns $G^{(1)}\text{DIAG}(\phi_{tot,s}^{(1)})$, one column for each source s. Note that we can likewise compute $\phi_{tot,s}^{(1)}$ for all sources s in Step 1 in a similar manner, and that the matrices in Steps 1, 2a are the same. The solutions are generated with an iterative method. The iterative method is an approximate solver. Iterations are terminated when the relative residual norm $\|AX - B\|/\|B\|$ is less than some user defined tolerance, and in this sense, the derivative evaluations are approximate for all the parameters. We use a preconditioned block MINRES[13] algorithm for each block system solve. (Since

18

MINRES can only be used for a symmetric system, the equation is first "symmetrized" by multiplying through by the appropriate diagonal matrix. The preconditioner is also a diagonal matrix.) As the computation of columns of the Jacobian is independent, the columns can be computed in parallel if the algorithm is to be run on a parallel computer.

Thus, our approximation to J is as follows. The first N_a columns of J in row i correspond to derivatives with respect to α, the next N_b to derivatives with respect to β, and the remaining 9 columns (or 4 in the case of spheres) correspond to estimates of derivatives in c, d, θ, respectively. Note that no finite difference approximation in Step 2a is needed to compute the derivatives with respect to α and β.

We found that even when the Jacobian matrix was reasonably well conditioned, the full Gauss-Newton step (12) was not appropriate for every iteration. Therefore, we employed a linesearch[12,14] to scale the length of the step. In other words, we updated v according to $v_{k+1} = v_k + \delta_k s$, where δ_k was a positive scalar selected at step k by the linesearch subroutine. The linesearch routine requires the evaluation of the function being minimized (namely $\frac{1}{2}\epsilon^T \epsilon$) and its gradient. Thus, each step of the line search is not cheap, requiring the same number of block linear system solves as are needed to form the Jacobian. However, just as the columns of the Jacobian can be evaluated in parallel, so can the gradient computations. The Gauss-Newton algorithm when implemented with a linesearch is often called "damped" Gauss-Newton[12].

In order to generate our first search direction from (12) we must generate a Jacobian and residual ϵ. This means we must provide a starting guess to the algorithm; that is, we must begin with an initial ellipsoid (sphere) and initial values for α and β. Once the initial ellipsoid is chosen, some care must be taken in choosing the starting guesses for α, β. These values

should not be completely arbitrary because they are related to the initial ellipsoid/sphere. In our approach, we fixed the parameters defining the starting ellipsoid/sphere and used Gauss-Newton with no linesearch (i.e. we are minimizing (10) only with respect to α, β) to determine the initial values for α and β. We continued taking Gauss-Newton steps until relative mean-square error between successive approximations was .01; usually, this required only two or three steps.

4. Numerical Results

All results were computed in Matlab in IEEE floating point double precision arithmetic and were built as an extension of the Photon Migration Imaging (PMI) Toolbox[15]. As discussed in § 2, a slab geometry is considered as a model for current breast imaging scenarios. There were 16 sources on the top of the slab and 16 detectors on the top and 16 on the bottom. Data were collected at DC only (that is, zero frequency intensity modulation was used), making the length of the data vector y equal to 512. The slab was 6 cm thick with the sources and detectors spanning 6 x 6 cm on each surface. The region of interest in which the optical image was reconstructed spanned 6 x 6 x 4 cm centered between the two planes of the slab. The size of each voxel was either 2 x 2 x 2 mm or 2.5 x 2.5 x 2.6 mm as described below. The values of the background optical properties that were used to construct the Green's function matrix were an absorption coefficient (μ_a value) of 0.05 cm^{-1} and reduced scattering coefficient (μ'_s) of 10cm^{-1} .

We used centered differences to compute approximate first derivatives with respect to the descriptors of S. The behavior of our algorithm was somewhat dependent on the spacing we used to define these centered differences: clearly, the spacing cannot be too small, since then a change in S is not observed on the grid we are using and the Jacobian will always be singular.

Of course the spacing defining the differences cannot be too large, or the approximation to the derivatives becomes poor. In our experiments, we used a spacing of 1mm (1.25mm for the coarser grid) for derivatives with respect to change in centroid and change in axis lengths, and $\pi/4$ spacing for the angular derivatives. These values are one half of the grid spacing and were determined to work well by trial and error.

Both shot and electronic noise was added to the data. The fractional standard deviation in the total signal ranged from $10^{-4}\%$ to 200% with a mean standard deviation of 5%. The inverse of the square root of the diagonal noise covariance matrix, W in (10), was then computed and used to whiten the data.

In the first set of experiments, we report results in detail for a total of seven configurations of objects and background structures using a 2 x 2 x 2mm grid. The first six examples are concerned with illustrating the effects of mismatch in our knowledge of the background variations in our ability to recover a spherical inclusion (examples 1–3) and an ellipsoidal anomaly (examples 4–6). In each case, $N_a = 1$ (i.e. the value inside the anomaly was constant) and therefore α is a scalar[1]. Because the primary purpose of these examples is to illustrate and analyze the effects of background model mismatch, the true objects we seek to characterize are in fact spherical in shape for the first set of examples and ellipsoidal for the second. To model a "lumpy" background, in these first six examples B_b is constructed using sinusoidal functions of space. The specific functions used in these examples were *not* chosen particularly to highlight the utility of our approach but rather represent results which are illustrative of the general performance of this method.

---

[1] Given the relatively small size of the anomaly and given the resolution being used, according to the results in[16], we do not expect taking N_a bigger than one to be helpful.

In experiments (examples 7-8), we report results on the performance of the algorithm on a 2.5 x 2.5 x 2.6 mm grid when a) the "lumpy" background is specified by supplying the discrete B_2 directly, rather than discretizing a continuous function $B_b(r)$, and b) the anomaly is neither a sphere nor an ellipsoid.

In all examples, α was a single coefficient with value .15. Due to our model, specifying this coefficient is the same as specifying the perturbation value inside the anomaly as .15cm^{-1}. The values of the expansion coefficients β_i, on the other hand, do not correspond directly to a background perturbation value (unless $B_b(r) = 1$) and hence are unitless. Note that all images are images of the absorption perturbation γ or estimates thereof, and not absolute absorption.

Example Number	B_b used to generate data	B_b used for inversion	Figures
1	1	1	1
2A	$\sin(3x) + 1,$ $\cos(8y)\sin(2y) + 1,$ $\sin(5z) + 1$	1	2 and 3
2B	$\sin(3x) + 1,$ $\cos(8y)\sin(2y) + 1,$ $\sin(5z) + 1$	$\sin(3x) + 1,$ $\cos(8y)\sin(2y) + 1,$ $\sin(5z) + 1$	2 and 4
3A	$\sin(3x) + 1,$ $\cos(8y)\sin(2y) + 1,$ $\sin(5z) + 1$	$\sin(\omega x) + 1,$ $\sin(\xi y) + 1,$ $\sin(\kappa z) + 1$	2
3B	$\sin(3x) + 1,$ $\cos(8y)\sin(2y) + 1,$ $\sin(5z) + 1$	$1,$ $\cos(x) + \cos(y),$ $\cos(z)$	2 and 5

Table 1. Setup for spherical inversion examples

A. Inverting for Spherical Anomalies

The primary purpose of this set of examples is to explore the effects of unknown and perhaps mismatched background heterogeneities on our ability to localize and characterize spheri-

Example Number		Radius cm	Contrast cm^{-1}	Location (cm, cm, cm)	Final voxel error
1	Truth	0.80	0.15	(-0.60, 1.00, 3.40)	11
	Estimate	0.76	0.17	(-0.60, 0.99, 3.34)	
2A	Truth	0.80	0.15	(-0.60, 1.00, 3.40)	98
	Estimate	0.86	0.10	(-0.46, 1.00, 3.46)	
2B	Truth	0.80	0.15	(-0.60, 1.00, 3.40)	12
	Estimate	0.86	0.17	(-0.60, 1.00, 3.39)	
3A	Truth	0.80	0.15	(-0.60, 1.00, 3.40)	N/A
	Estimate	N/A	N/A	N/A	
3B	Truth	0.80	0.15	(-0.60, 1.00, 3.40)	295
	Estimate	0.83	0.08	(-0.39, 1.07, 3.35)	

Table 2. Spherical reconstruction results

cal anomalies. By "mismatched" we mean that the functions in B_b used to generate the background absorption structure, are different from those used during the inversion process. The experimental conditions are summarized in Table 1 where we indicate the background basis functions used for generating the data, those used in the inversion procedure, and the relevant figures in the text in which the performance is illustrated.

Summary results for our method are provided in Table 2. Here we indicate the accuracy to which we can estimate the size, location, and contrast of a spherical anomaly located at $(x_0, y_0, z_0) = (-0.60, 1.00, 3.40)$ cm with a radius of 0.80 cm and a contrast of 0.15 cm^{-1}. The difference between the number of voxels (out of about 20,000) in the estimated sphere and the true sphere is also provided. Each of the three examples is discussed in greater detail below.

1. Example 1: Sphere, PWC matched model

In this example $N_b = 1$ for the true image (i.e. the background is assumed constant) and B_2 is the vector of all ones. Therefore, the true image is piece-wise constant (PWC). In

the reconstruction, the N_b and B_2 used are the same as for the true image, meaning there is no "model mismatch" between the true solution and the computed solution (i.e. the reconstruction is also PWC).

We assumed that no good starting guess for the location of the sphere was known. Therefore, we took the starting guess as the largest sphere that fit inside the region of interest – a sphere of radius 2cm centered at the origin. Thus, there was a 3914 voxel difference between the starting guess and the true anomaly.

The values computed by our algorithm are summarized in the first two lines of Table 2. We see that when the model is well matched, we can recover the spherical structure almost perfectly. The significance of the small, 11 voxel, difference between the real sphere and the estimated one is illustrated in Figure 1 which compares 2D cross-sectional sparsity plots of the entries in S for the exact solution with that of the reconstruction. That is, a colored mark indicates a 1 in the corresponding position in S, and no color indicates a 0. Clearly, our algorithm has done an excellent job of both characterizing and localizing the anomaly. We stress that this performance is fairly consistent over different noise realizations, and can improve – in some experiments, there was only a 5 voxel difference between the true and computed anomalies.

2. *Example 2A: Sphere, model mismatch*

Here, we have taken $N_b = 3$. The function $B_b(r)$, where r denotes an (x, y, z) triple, is defined as

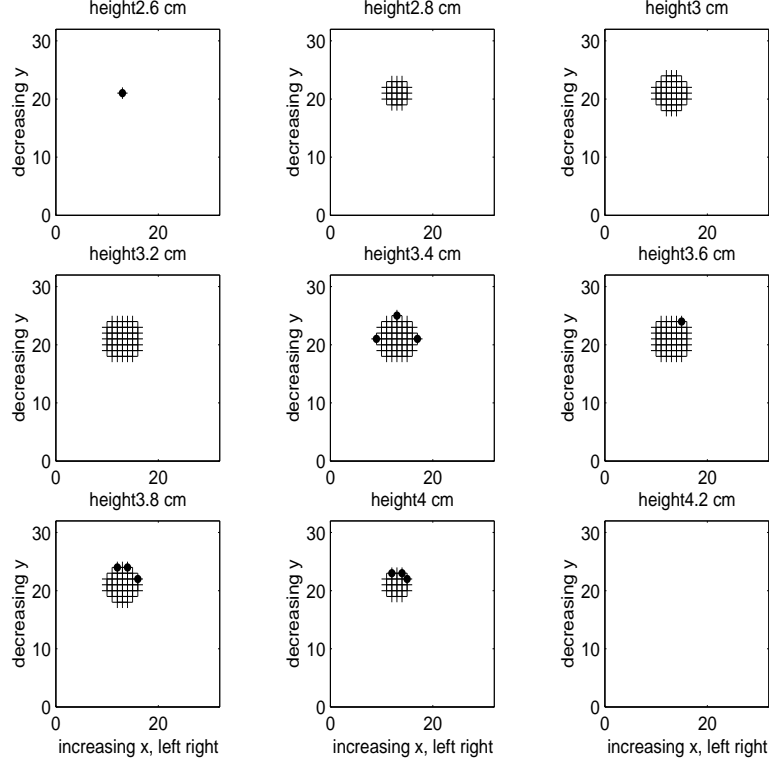$$B_b(r) = [\sin(3x) + 1, \cos(8y)\sin(2y) + 1, \sin(5z) + 1],$$

Fig. 1. Sparsity plots, moving down from the top of the region of interest (where height=1cm), depict the location of true and reconstructed anomalies for Example 1. The plus marks indicate the location of the true anomaly. The black dots indicate voxels missed by the reconstructed anomaly as the reconstructed anomaly was slightly too small. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

and the corresponding matrix B_2 is determined by evaluating B_b at all points r_i inside the region of interest. The true values of the coefficients β_i are 2.0e-3, 2.0e-3, and 1.00e-3. The anomaly itself was the same as in Example 1. We note that the background perturbation was found to have a maximum of 9.97e-3, a minimum of 1.74e-5, and an average value of 4.96e-3.

As discussed previously, one goal of this work was to establish the effect of lumps in the background tissue on reconstruction. Hence here we reconstruct using $N_b = 1$ and take B_2 in the reconstruction is a single vector of all ones (corresponding to $B_b(r) = 1$). The starting guess for the sphere is the same as in Example 1. As we see from Table 2, in spite of the
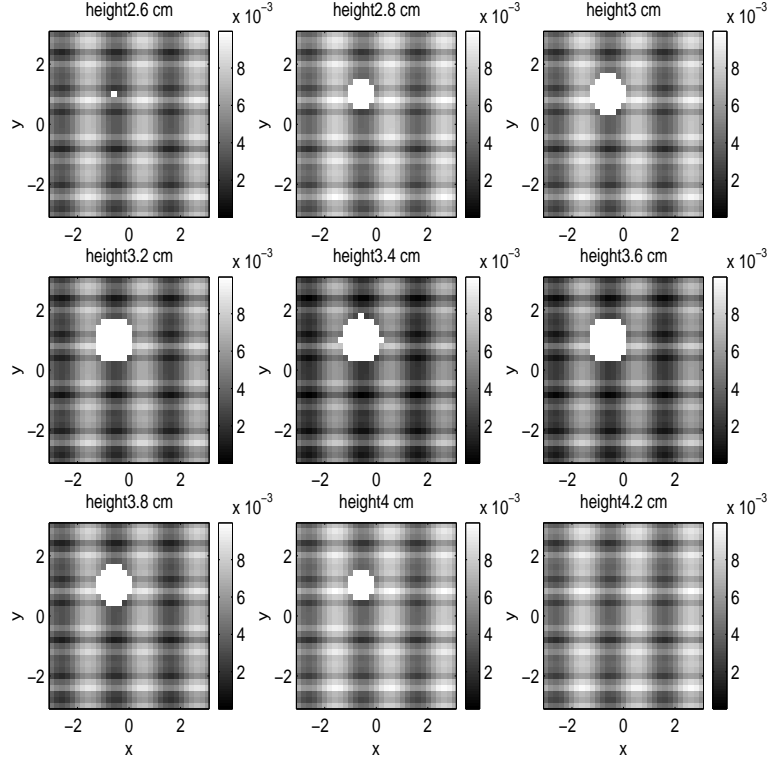
Fig. 2. True image, slices moving down from surface of the region of interest, Examples 2A and 2B. Value inside the anomaly is .15 cm^{-1}: colormap is truncated to show the background variation, so the anomaly appears as a bright white spot.

mismatch in the model, there is only a small degradation in performance. The 98 voxel mismatch which can be attributed to the fact that the radius is slightly overestimated and the x, z centers are slightly off target, as can be seen in the spy plots in Figure 3. Finally, we observe that the value of β in the reconstruction was 4.7e-3 which is very close to the *average* value of the true background perturbation, while the relative error between the true and computed α was 28 percent.

3. Example 2B: Sphere, matched model

We expected an improvement if in fact the same background bases were used in the reconstruction as was used to generate the true image in the preceding example. So here we repeated example 2A, except the $B_b(r)$ used in the reconstruction was the same as that for
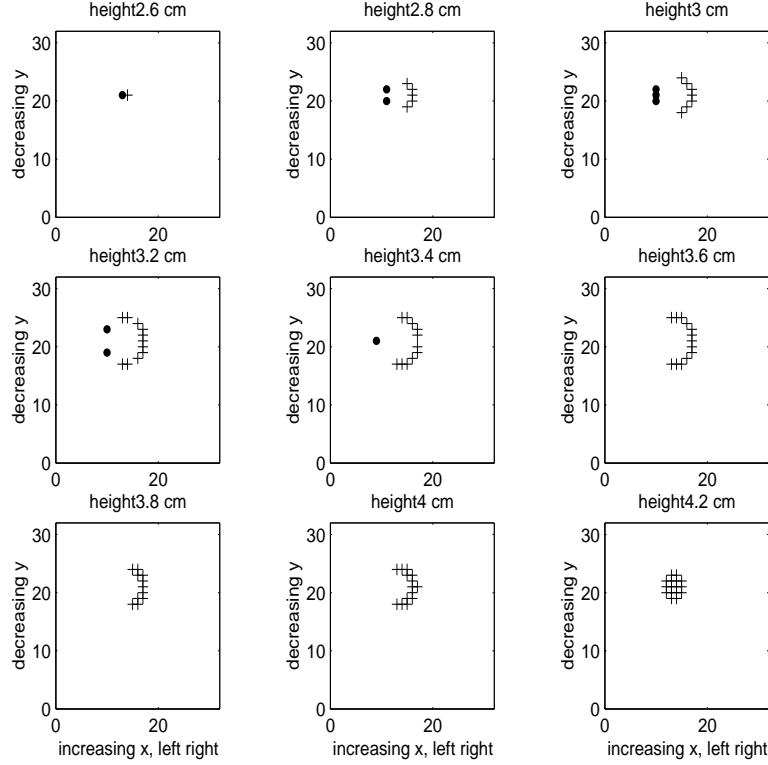
Fig. 3. Sparsity plots depict the reconstruction error in shape for Example 2A. The black plus marks indicate points that were inside the reconstructed anomaly that were not inside the true anomaly. The black dots indicate points that were inside the true anomaly that were not inside the reconstruction. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

the true figure. The estimated parameter values shown in Table 2 lead to a final estimate of S having only a 12 voxel mismatch. Thus, by correctly modeling the background perturbations, we can obtain a significant improvement in localization and in the estimated value of α. Figure 4 shows slices in depth of the reconstruction.

4. *Example 3: Sphere, multiple model mismatches*

In this test we have taken $N_b = 3$ both to generate the true image and to generate the reconstruction. However, $B_b(r)$ for the true image is the same as those in the previous two
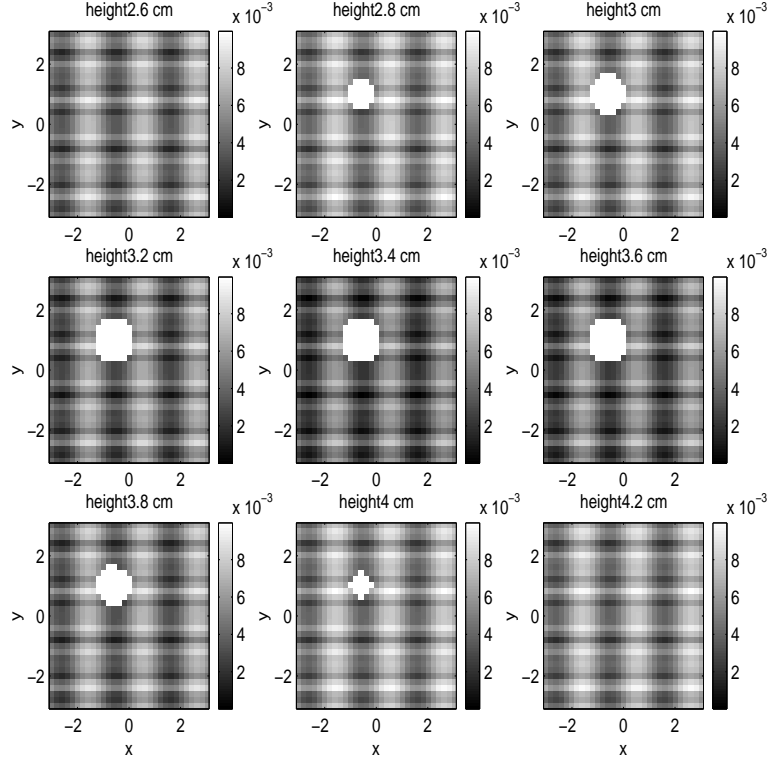
Fig. 4. Reconstruction with matched background basis functions in Example 2B, slices are shown moving away from the top surface. Reconstructed α value is 0.168. The colormap truncation is the same as was used to display the true image, so the reconstructed anomaly appears as a bright white spot.

examples while $B_b(r)$ for the reconstruction is incorrectly assumed to be

$$B_b(r) = [\sin(\omega x) + 1, \sin(\zeta y) + 1, \sin(\kappa z) + 1]$$

and we experimented with different values of the parameters ω, η, and κ. For small values of the parameters (e.g. ≤ 4), the reconstructions were worthless (no better than or worse than the starting guess). This is understandable since no matter what the expansion coefficients, the high frequency bumps in the true background cannot be captured with this basis.

Next we tried $B_b(r) = [1, \cos(x) + \cos(y), \cos(z)]$ in the reconstruction. Again, high frequency bumps cannot be captured in this basis, but since we have allowed for constant

variation, we expect the algorithm to behave similar to 2A in that the background perturbation should converge to roughly the average value of the true background. This is in fact the phenomenon we observed. The reconstructed values for the anomaly are provided in Table 2. The background expansion coefficients were estimated to be $\hat{\beta}$ =[4.84e-4,-3.90e-6,-4.87e-4]. Observe that the maximum, minimum and mean values for the reconstructed background were 5.34e-3, 4.57e-3, 5.04e-3. Hence, the reconstructed background was almost constant again at about the average value of the background perturbation. Sparsity plots in Figure 5 comparing the true sphere with the reconstructed sphere put the 115 voxel mismatch in perspective. Comparing with example 2A, we are lead to believe that a piecewise constant background assumption can be superior to using incorrect basis functions if nothing is known about the background perturbation a priori; this deserves further study.

Example Number	$B_b(r)$ used to generate data	$B_b(r)$ used for inversion	Figures
4	1	1	6
5A	$\sin(7x) + 1$, $\sin(4y) + 1$, $\sin(3z) + 1$	1	7 and 8
5B	$\sin(7x) + 1$, $\sin(4y) + 1$, $\sin(3z) + 1$	$\sin(7x) + 1$, $\sin(4y) + 1$, $\sin(3z) + 1$	7 and 9
6	$\sin(7x) + 1$, $\sin(4y) + 1$, $\sin(3z) + 1$	$\sin(7x) + 1$, $\sin(8y) + 1$, $\sin(10z) + 1$	

Table 3. Setup for ellipsoid experiments

B. *Inverting for Ellipsoidal Anomalies*

In this subsection, we examine the background model mismatch issue for the more complex problem of estimating the structure of an ellipsoidal absorption anomaly. The background
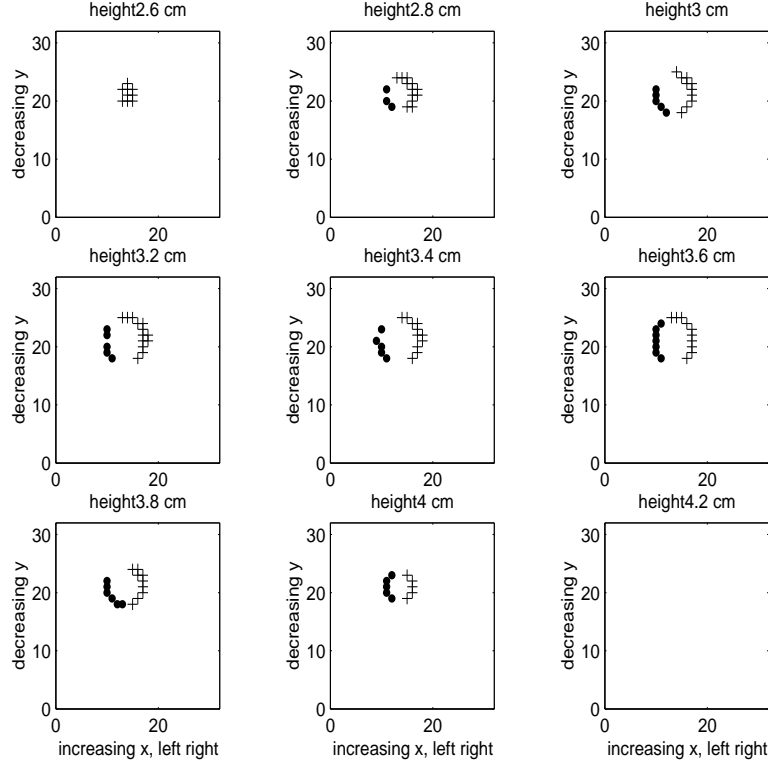
29

Fig. 5. Sparsity plots depict the difference between the location of true and reconstructed anomalies for Example 3. The black dots indicate which voxels are in the true anomaly that are missed by the reconstruction while the plus marks indicate voxels in the reconstruction that are not in the anomaly. Clearly the reconstruction has the anomaly shifted slightly to the right in x. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Example Number		Axis lengths (cm, cm, cm)	Location (cm, cm, cm)	Angles (rad, rad, rad)	Contrast cm^{-1}	Final voxel error
4	Truth	(1.10, 0.50, 0.80)	(0.70, -0.90, 2.40)	(0.79, 0.79, 0.00)	0.15	26
	Estimate	(0.83, 1.11, 0.54)	(0.69, -0.90, 2.42)	(-8.64, 0.81, 1.58)	0.12	
5A	Truth	(1.10, 0.50, 0.80)	(0.70, -0.90, 2.40)	(0.79, 0.79, 0.00)	0.15	3450
	Estimate	(1.99, 2.21, 1.60)	(0.68, -0.79, 3.02)	(-15.4, 0.44, 2.01)	0.014	
5B	Truth	(1.10, 0.50, 0.80)	(0.70, -0.90, 2.40)	(0.79, 0.79, 0.00)	0.15	21
	Estimate	(0.82, 1.10, 0.53)	(0.69, -0.91, 2.41)	(0.53, -1.49, 0.78)	0.13	
6	Truth	(1.10, 0.50, 0.80)	(0.70, -0.90, 2.40)	(0.79, 0.79, 0.00)	0.15	701
	Estimate	(1.75, 1.54, 0.66)	(0.77, -0.78, 2.47)	(-12.0, -0.21, 0.35)	0.036	

Table 4. Ellipsoid results. See comments at the beginning of Section 4 B for further information on how to judge results in this table.

structures used for the experiments are summarized in Table 3 while the inversion results are provided in Table 4.

In Table 4, note that the periodic nature of entries in the rotation matrix U in (9) implies that different angles can affect essentially the same rotation (i.e. the θ_i are not unique), so we report them in Table 4 only for completeness. Further, note that different orderings of axis lengths combined with certain rotation matrices can generate the same ellipsoid, so it is the values of the axis lengths, and not their ordering in the triple, that are important. Therefore, to better gage the quality of the ellipsoidal reconstruction, one should look at the voxel mismatch numbers.

1. *Example 4: Ellipsoid, PWC matched model*

In this example $N_b = 1$ for the true image (i.e. the background is assumed constant) and B_2 is the vector of all ones. The true value for β was taken as 5e-3. In the reconstruction, N_b and B_2 are the same, meaning there is no model mismatch between the true solution and the computed solution.

The values obtained by the reconstruction are shown in the first row of Table 4 and the estimate of β was 5.00e-3. There was a difference of 26 voxels between the true and reconstructed ellipsoid. The sparsity plots, which visually depict the S matrices in each case, are shown in Figure 6. From the plot it is obvious that the axis rotation is accurately captured.

2. *Example 5A: Ellipsoid, model mismatch 1*

Now we specify a lumpy background for the true image using $N_b = 3$ and $B_b(r) = [\sin(\omega x) + 1, \sin(\zeta y) + 1, \sin(\kappa z) + 1]^T$, with $\omega = 7; \zeta = 4; \kappa = 3$. This gives a max background perturbation value of $1.20e - 2$, a minimum perturbation value of $3.07e - 5$ and an average
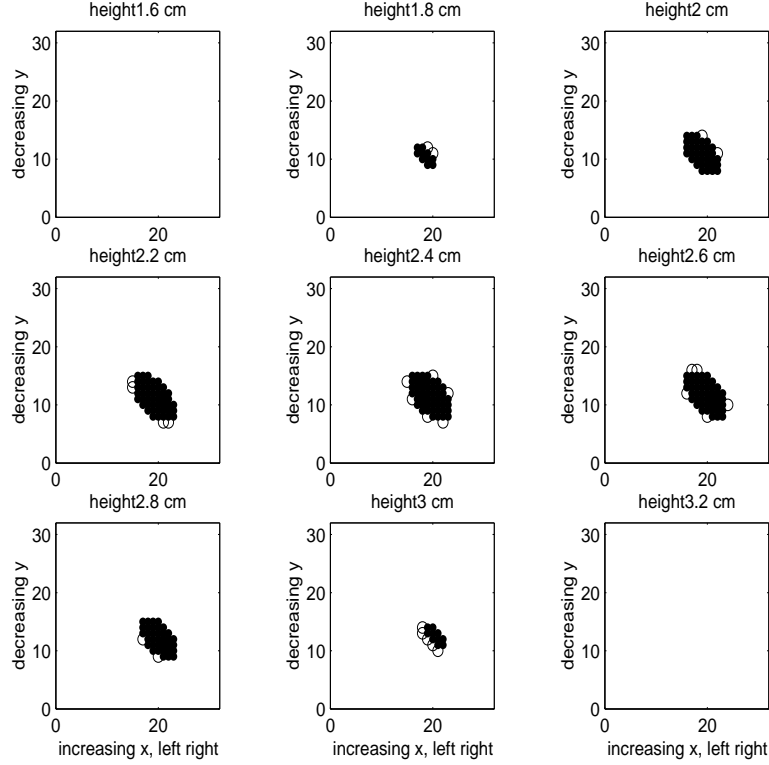
Fig. 6. Sparsity plots depict the location of true and reconstructed anomalies for Example 4. The black dots indicate the true anomaly, the centers of the black circles show areas of the reconstructed anomaly that lie outside the true anomaly (in other words, the reconstruction is represented the dots plus the circles). The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

of $5.98e - 3$. Cross-sectional slices in depth for the true image can be seen in Figure 7.

The reconstructed values for the ellipsoid parameters are provided in Table 4. The estimate of β was; $4.95e - 3$. As in the sphere example, we observe that the background value that is calculated by the algorithm is close to the mean value of the background perturbation and the x coordinate for the center is reasonable but otherwise, the estimate is fairly far off. The algorithm has trouble locating the true center initially (namely the z-coordinate), due to the nature of the background perturbation. Once a bad estimate of the center was obtained by the algorithm, it never seemed to recover and got stuck in a local minimum very early on. This type of phenomenon is related to the location and spread of the "peaks" in the
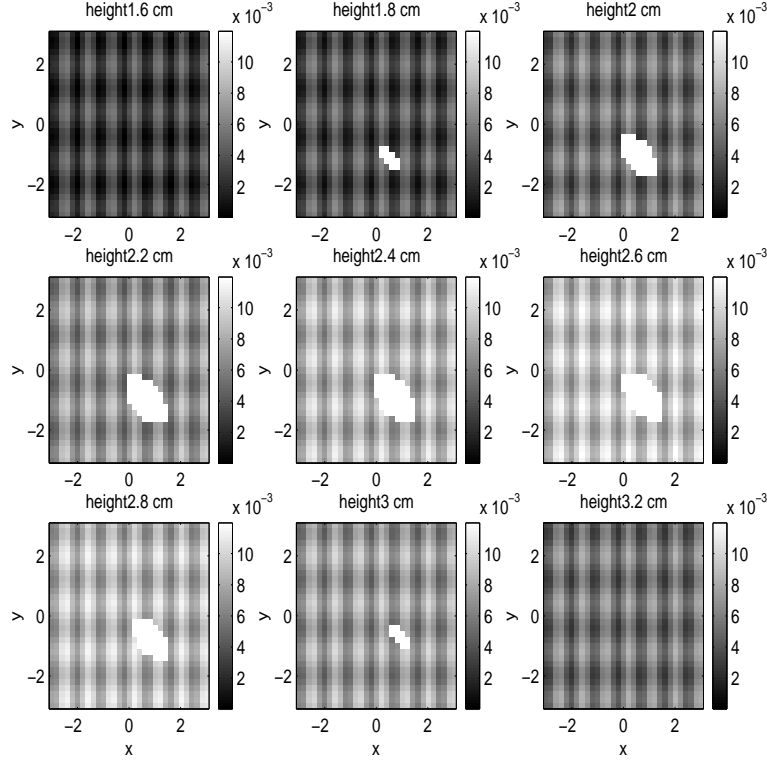
Fig. 7. True image used in Examples 5A, 5B. True α value is .15. The colormap is truncated to show background detail, so the anomaly appears in bright white.

background perturbation (see Section 4 E and Example 6 below) – when the peaks are wide and of high enough value, they obscure the true size and location of the anomaly. There was a 3,450 voxel mismatch (compared to a 8,817 voxel mismatch for the starting guess); see Figure 8.

3. Example 5B: Ellipsoid, matched model

Finally, we use the true image in Example 5A, but we reconstruct with the correct basis. The reconstruction is vastly different than in 5A and fairly close to perfect. Comparing Figures 9 and 7, shows that despite the large difference in the estimates of the rotation angles, when used to construct U, these values do in fact lead to the correct axis orientation.
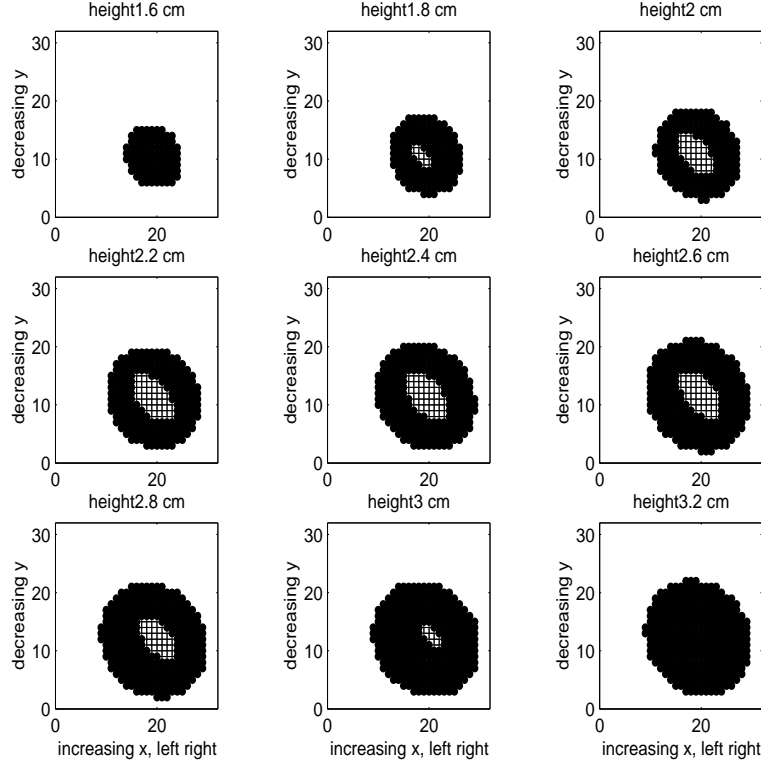
Fig. 8. Cross-sectional display of locations of reconstructed ellipsoid (dots and plus marks) vs. true ellipsoid (plus marks) for Example 5A. The oversized reconstruction is the result of not appropriately accounting for a lumpy background.

4. Example 6: Ellipsoid, model mismatch 2

We performed almost the same experiment as in 5A except that in B_b, the values of ζ and κ were changed to 8 and 10, respectively. This gave a max background perturbation value of 1.20e-2, a minimum of 4.39e-5 and an average of 5.85e-3. The reconstructed α value was only 3.57e-2 – not quite as small as estimated in Example 5A, but still about 1/4 the magnitude it should be. The value for β, 5.18e-3 is close to the average background value again. The centers estimate was somewhat off although still better than Example 5A. The lengths of the axes were too large at $d_1 = 1.75; d_2 = 1.54; d_3 = .66$. There is a 701 voxel mismatch versus to 3,450 in Example 5A. Comparing the results in 5A with these results, we conclude that if an incorrect basis is used in the reconstruction, the texture of the true background
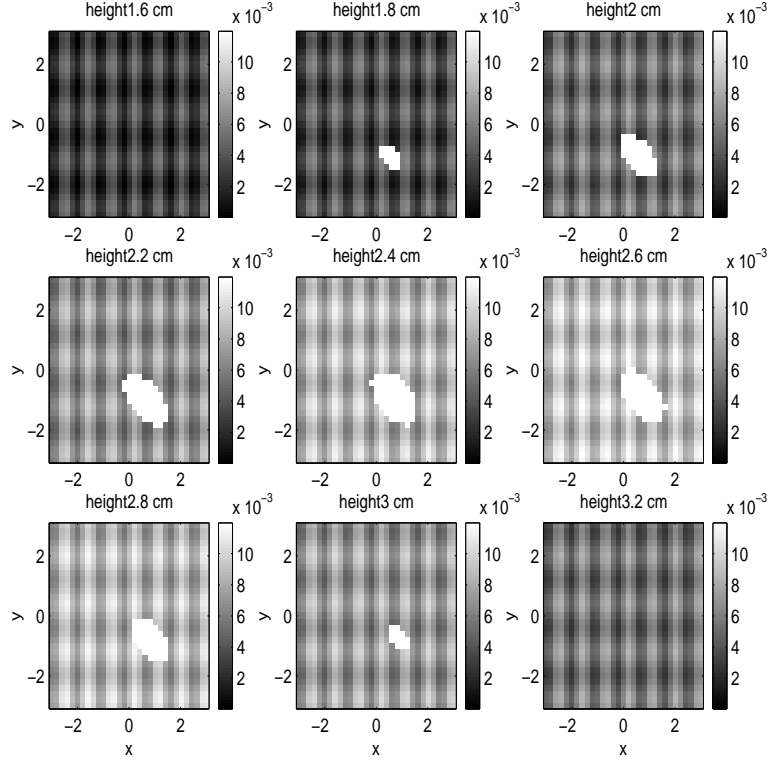
34

Fig. 9. Reconstructed perturbation for example 5B. The reconstructed value of α is 0.13. The colormap truncation is the same as that used in Figure 7.

perturbation can have a significant impact on reconstruction quality.

C.   Example 7: Sphere in Lumpy Background

In these experiments, we explore further the effect of a lumpy background on the reconstruction process, but where the lumpy background does not correspond to a functional expression $B_b(r)$. Instead, we deal only with the discrete formulation of the problem; that is, we specify the background through a single digital image. We take B_2 to have a single column, but to get the values in that column, we apply a 3D highpass filter to a 3D white noise array and unstack the array to obtain a vector. There are two parameters that allow the background to change: one parameter , s_c sets the maximum value of the background, and the other, w, is a negative number in (-15,0) that controls the width of the lumps by controlling the filtering

effect of the highpass filter. For example, when $w \approx -1$, the lumps are small whereas when w is increasingly large and negative, the lumps become smaller (compare Figs 10,11, for example).

The true value of α was 0.15 and the true value of β was 1. The anomaly was a sphere of radius 1 and center located at (0.5,1.1,3.7). The idea was to reconstruct with a sphere and a PWC model for the absorption perturbation in the background and anomaly. In Table 5 we can see that both the maximum value of the background perturbation as well as the width of the lumps does influence the reconstruction. Further experiments, not reported here for the sake of brevity, using increasingly large negative w parameters also support the trend observed in the table. In short, as the lumps widen, the maximum value has to become smaller in order to be able to get a reasonable reconstruction. Figure 10 and 11 show cross-sections of the true solution when the background was created with $s_c = .005, w = -1$ and $s_c = .005, w = -2$, respectively. (Note that the colormap has been truncated so that the background is better seen – the value of the bright white anomaly is actually 0.15 cm^{-1}.)

The last row in Table 5 and the Figure 12 show that it is still possible to get good localization and characterization of the anomaly if the correct background basis vector is used, even when the PWC reconstruction fails. Hence, future work should consider how to estimate the background basis functions (see also further remark in Section 5).

D. *Example 8: Non-ellipsoidal Structure*

With this example we begin to explore the effect of applying our ellipsoidal model and algorithm when the true anomaly is neither a sphere nor an ellipsoid. We used a PWC model both to generate the true image and in the reconstruction (i.e. $B_b(r) = 1$ for both the true absorption perturbation and the reconstructed absorption perturbation.
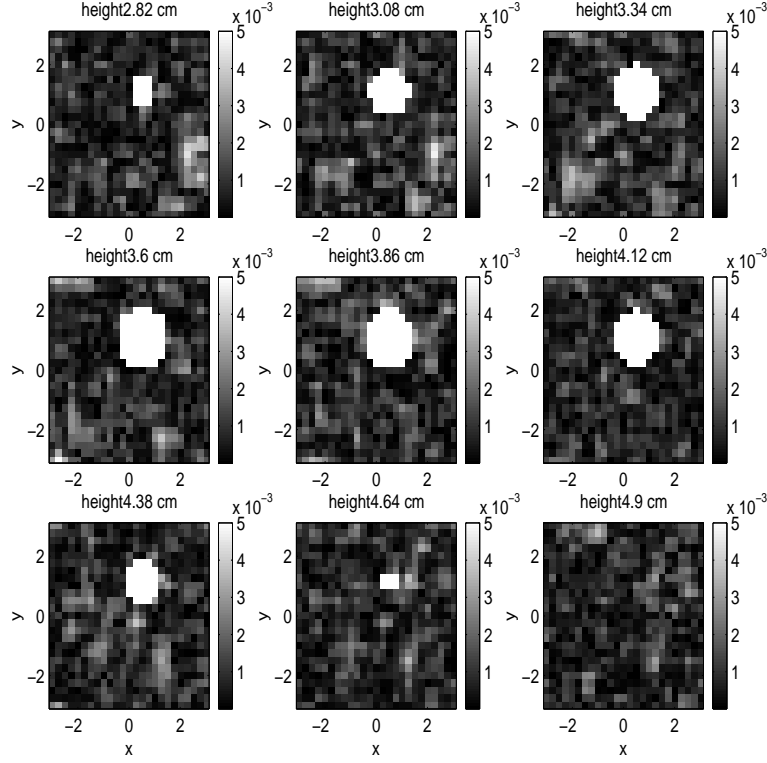
Fig. 10. Cross-sectional slices moving down from the surface for the true solution in Section 4 C with $s_c = .005, w = -1$. True value of α is 0.15: colormap is truncated to show background detail.

For the first experiment, the true anomaly was the concave, non-ellipsoidal shape shown in black in the image in Figure 13, and the perturbation value inside the anomaly was 0.15cm^{-1} while outside it was 0.005. In the reconstructions, we assume S is defined by an ellipsoid. The ellipsoidal shape that is reconstructed by our algorithm is shown in red in Figure 13. There is a 98 voxel mismatch. The reconstructed perturbation value is correct outside the anomaly (0.005cm^{-1}). Inside the anomaly the perturbation value is low at 0.08cm^{-1}, (this represents a 44 percent error) but it is clearly much larger than the background perturbation. The overall localization is very good.

For the second experiment, the shape was the submarine-shaped object in black shown in Figure 14, and the perturbation value inside the anomaly was again 0.15cm^{-1}. As in the first
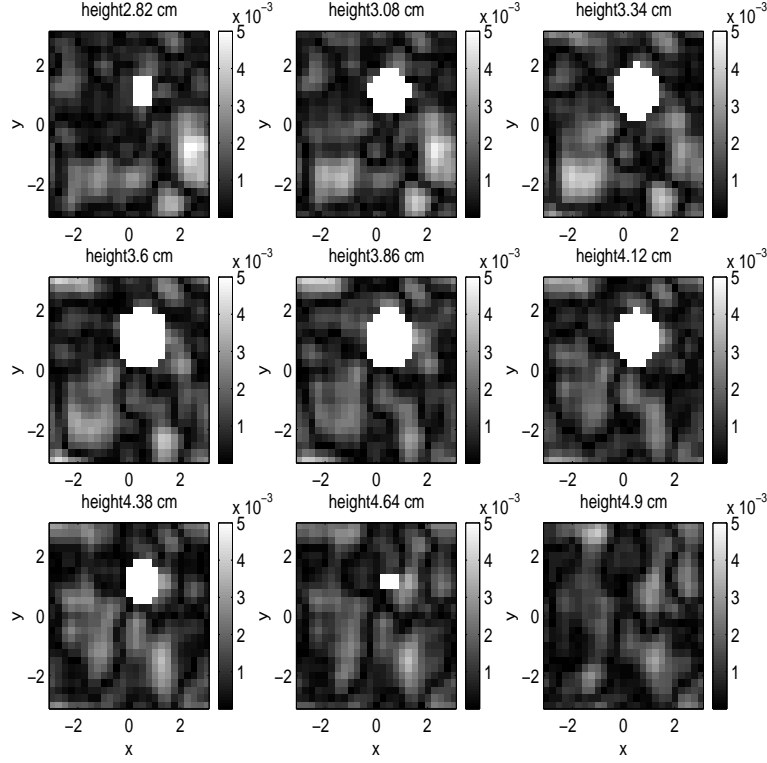
Fig. 11. Cross-sectional slices moving down from the surface for the true solution in Section 4 C for $s_c = .005$, $w = -2$.

		α	B_b generated	β	center	Radius	Voxel mismatch
truth		.15	lumps	1	(.5,1.1,3.7)	1	
bkg values							
s_c	w						
.01	-1	.008	PWC	.002	(-1.25,1.70,4.01)	2.64	2892
.005	-1	.06	PWC	9.8×10^{-4}	(.48,1.12,3.59)	1.15	133
.005	-2	.003	PWC	9.6×10^{-4}	(.28,1.62,4.7)	3.05	3255
.001	-2	.09	PWC	2.02×10^{-4}	(.50,1.08,3.61)	1.08	61
.005	-2	.16	lumps	1.00	(.50,1.10,3.69)	.99	7

Table 5. Results when reconstructing with PWC basis for various lumpy backgrounds. The last row gives values when reconstructing with the correct basis. When interpreting the last column, keep in mind that the voxel mismatch between the starting guess and true solution is 1856.

experiment, in the reconstruction phase we use an ellipsoid to define S. The reconstruction

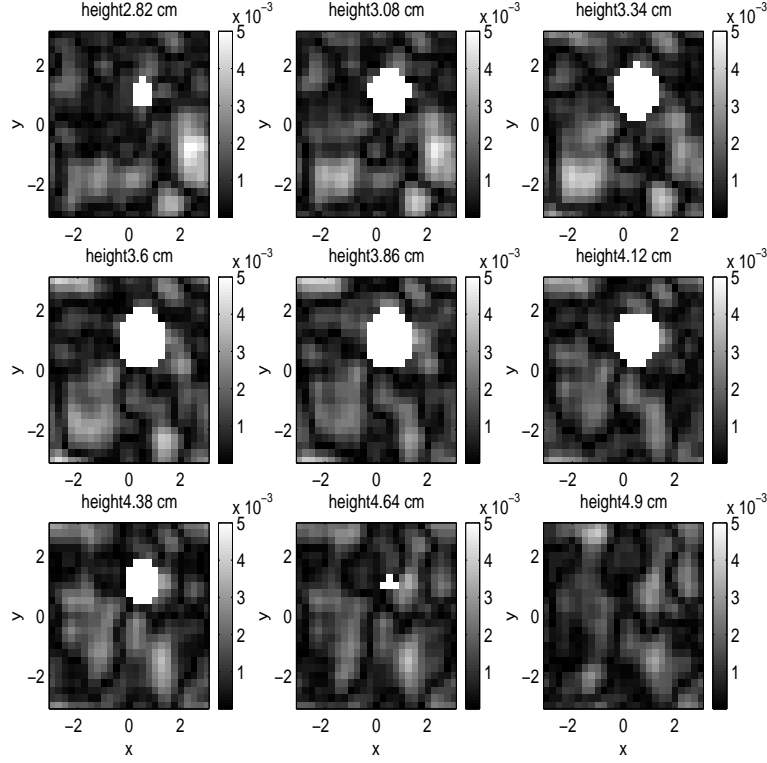is shoen in red in Figure 14. There is a 94 voxel mismatch between the true shape and the

Fig. 12. Cross-sectional slices moving down from the surface for the reconstructed solution in Section 4 C for $s_c = .005, w = -2$ using the correct B_2 in the reconstruction. Reconstructed α value is .16 and the colormap truncation is the same as was used in Figure 11.

reconstruction. The reconstructed perturbation value inside the anomaly is still somewhat low at 0.10cm^{-1} (this represents a 33 percent error), but the reconstructed value in the background is correct (.005) and the localization is very good.

Both results suggest that the ellipsoidal algorithm performs almost as well as can be expected given the model mismatch and that it could be used as a starting point for a related imaging scheme with more flexible parametrically described shapes.

E.   *Summary of Experimental Results*

Although not detailed explicitly here in the interest of space, a number of other experiments were performed. In general, we noticed the following types of behavior, given that such an arbitrary starting guess was used in all cases:
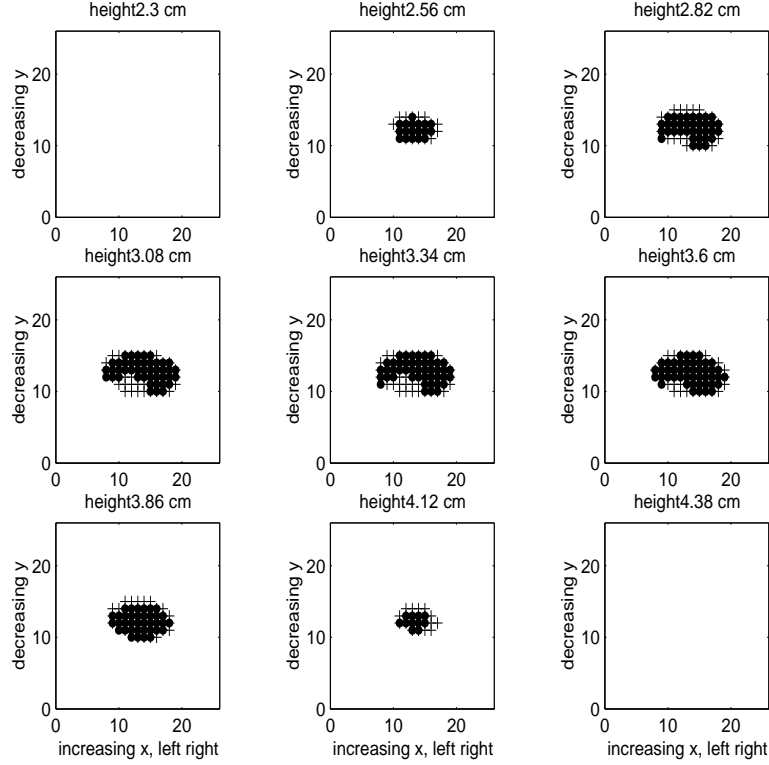
Fig. 13. Example 8A: Cross-sectional plots moving down from the surface comparing the shape/location of the true object with the reconstructed shape. The true anomaly is marked with black dots, the reconstruction is overlaid using plus marks . The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

1. If the anomaly was centered far from the "origin" ([0,0,3]), or if the Thickness of the medium was much larger than 6cm, the algorithm did not perform well: it could often estimate the centroid, but was not effective at shrinking the axis lengths. This is to be expected because the problem effectively becomes more severely underdetermined in that instance and the coverage from the sensors in these regions of space are sparser. A possible help for this would be to use additional data from different modulation frequencies, which we did not consider in the present work. Another possibility is to use an additional regularization term that encourages shrinkage relative to the estimated value inside the anomaly, similar to what we did in[17], but then one needs to
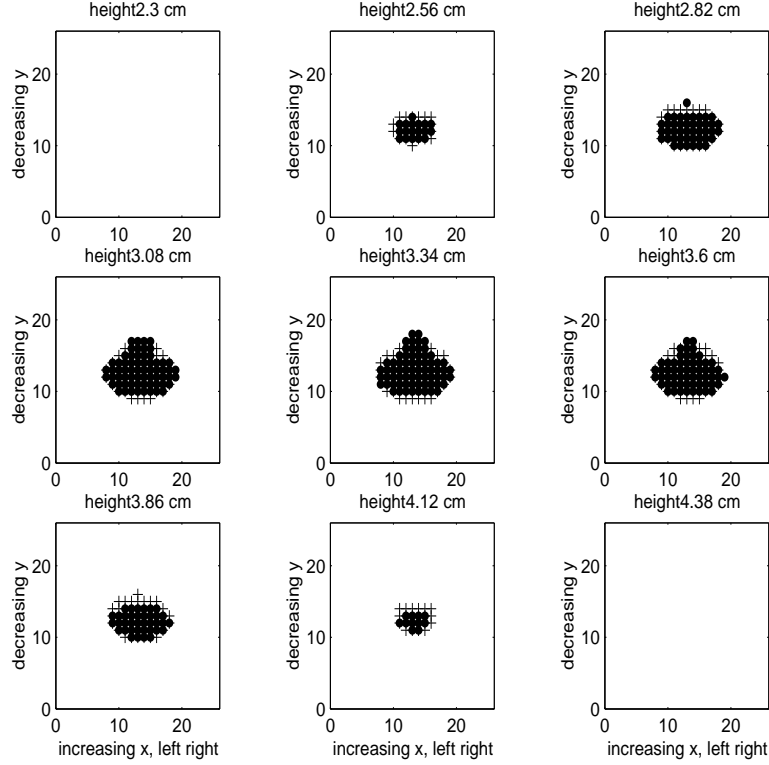
40

Fig. 14. Example 8B: Cross-sectional plots moving down from the surface comparing the shape/location of the true object with the reconstructed shape. The true anomaly is indicated with black dots, the reconstruction is overlaid using plus marks. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

worry about the choice of regularization parameter. Finally, as money/physics allow,

improvements will be gained by increasing the number of sources/detectors at the cost

of increased computation time.

2. If the background perturbation was non-constant but a PWC assumption was used

in the reconstruction, the types of lumps in the background perturbation affected the

reconstruction. More specifically, the wider the peaks of the lumps, the smaller the

peak values needed to be to obtain an accurate reconstruction.

3. Further work needs to be done to determine how much inaccuracy can be tolerated in

the solutions returned by the iterative solver.

4. Spheres can be used as a first step to approximate ellipsoids. When the spherical algorithm is run, it does a relatively good job of identifying the correct center of the ellipsoid, but the resulting sphere tends to have a radius which is smaller than the longest semiaxis. Then if that sphere is used as a starting guess to the ellipsoidal algorithm, the algorithm can get stuck in a local minimum, so care must be taken in using the spherical result to initialize the ellipsoidal algorithm. One possible remedy is to initialize the ellipsoidal algorithm with a sphere or ellipsoid having a larger-than-estimated radius but using the computed center.

5. The results show that localization had to be very good to get a reasonable estimate of the value inside the anomaly.

6. Useful information on localization could be obtained even if the anomaly was not spherical/ellipsoidal.

7. Finally, we note that we expect that as we refine the grid the approximations to the derivatives will correspondingly be improved, and therefore we expect the performance of our algorithm to improve.

5. Conclusions and Future Work

We have described an algorithm, based on a low-order parametric model, for jointly localizing and characterizing the absorption perturbation in breast tissue. Experimental results show the robustness of the algorithm under certain restrictions that may hold true in practice.

The background basis vectors used in our experiments were discretizations of continuous functions and were selected to show the strengths and limitations of our approach – they

42

were not necessarily related to physical properties of breast tissue. However, a key benefit of our approach when viewed in the discrete setting is that other choices of realistic basis vectors are possible. For example, suppose that another imaging modality were able to localize some inhomogeneities in the tissue (blood vessels or fat, for example). A segmented "image" localizing that type of inhomogeneity can serve as a column in the B_2 matrix. This is one area of current interest to us.

Future work includes the extension of our method to more than one anomaly. We claim that if an initial estimate from some other imaging scheme suggests the presence of say k anomalies, it is straightforward to extend our model and algorithm to find all k such anomalies. The question remains whether or not it is possible to find k anomalies without such initial information.

Our approach can be extended to jointly solve for perturbation in both absorption and scattering coefficients by specifying a similar model for the scattering coefficient using possibly different basis functions but the same S. More work is required to discern the value of adding data at other modulation frequencies as well as to identify the limits of resolution of our algorithm as the number of sources and/or detectors are increased. Finally, the application of this method to real sensor data is an area of some interest to us in the near future.

References

[1] V. Ntziachristos, B. Chance, "Probing physiology and molecular function using optical imaging: applications to breast cancer," *Breast Cancer Research*, Vol. 3, 2001.

[2] A. Villringer, B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends in Neuroscience*, Vol. 20, 1997.

[3] D. Boas, D. Brooks, E. L. Miller, C. DiMarzio, M. E. Kilmer, R. Gaudette, Q. Zhang, "Imaging the Body with Diffuse Optical Tomography," *Signal Processing Magazine*, Vol. 18, No. 6, 2001, pp. 57-75.

[4] M. Braunstein and R. Levine, "Three-dimensional tomographic reconstruction of an absorptive perturbation with diffuse photon density waves," J. Optical Soc. Am. A, Vol. 17, No. 1, 2000, pp. 11-20.

[5] M. Schweiger and S. R. Arridge, "Optical tomographic reconstruction in a complex head model using a priori region boundary information," *Phys. Med. Biol.*, Vol. 44, 1999, pp. 2703-2721.

[6] V. Kolehmainen, M. Vauhkonen, J. Kaipio, S. R. Arridge, "Recovery of piecewise constant coefficients in optical diffusion tomography," *Optics Express*, Vol. 7, No. 13, 2000, pp. 468-480.

[7] O. Dorn, E. L. Miller, C. Rappaport, "A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets," *Inverse Problems*, Vol. 16, No. 5, Oct. 2000, pp. 1119-1156.

[8] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems*, 15, R41-R93, 1999.

[9] T. J. Farrell, M. S. Patterson, B. Wilson, "A diffusion theory model of spatially resolved, steady state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo," *Medical Physics*, Vol. 18, 1992.

[10] M. E. Kilmer, E. L. Miller, D. Boas, D. Brooks, "A Shape-Based Reconstruction Technique for DPDW Data," *Optics Express*, Vol. 7, No. 13, 2000, pp. 481-491.

[11] S. Hassani, *Foundations of Mathematical Physics* (Allyn and Bacon, 1991).

[12] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization ad Nonlinear Equations* (Prentice Hall , New Jersey, 1983).

[13] D. P. O'Leary, "The block Conjugate Gradient algorithm and related methods," Vol. 29, 1980, pp. 293-322.

[14] D. P. O'Leary, Matlab translation of MINPACK subroutine cvsrch, http://www.cs.umd.edu/ oleary/m607/cvsrch.m

[15] D. Boas, D. Brooks, R. Gaudette, T. Gaudette, E. Miller, Q. Zhang, *Photon Migration Imaging (PMI) Toolbox*, freely available at http://www. nmr. mgh. harvard.edu/DOT

[16] E. L. Miller, M. E. Kilmer, and C. Rappaport, "A New Shape-Based Method for Object Localization and Characterization from Scattered Field Data," *IEEE Tran. Geosci. Remote Sensing,* Vol. 38, No. 4, 1682-1696.

[17] M. E. Kilmer, E. L. Miller, D. Boas, D. Brooks, C. DiMarzio, R. Gaudette, "Direct object localization and characterization from diffuse photon density wave data," in proceedings *Photonics West: Optical Tomography and Spectroscopy of Tissue III*, San Jose CA, 1999.

List of Figure Captions

Figure 1: Sparsity plots, moving down from the top of the region of interest (where height=1cm), depict the location of true and reconstructed anomalies for Example 1. The plus marks indicate the location of the true anomaly. The black dots indicate voxels missed by the reconstructed anomaly as the reconstructed anomaly was slightly too small. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Figure 2: True image, slices moving down from the surface of the region of interest, Examples 2A and 2B. Value inside the anomaly is .15 cm^{-1}: colormap is truncated to show the background variation, so the anomaly appears as a bright white spot.

Figure 3: Sparsity plots depict reconstruction error in shape for Example 2A. The black plus marks indicate points that werer inside the reconstructed anomaly that were not inside the true anomaly. The black dots indicate points that were inside the true anomaly that were not inside the reconstruction. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Figure 4: Reconstruction with matched background basis functions in Example 2B, slices are shown moving away from the top surface. Reconstructed α value is .168. The colormap truncation is the same as for the true image, so the reconstructed anomaly appears as a bright white spot.

Figure 5: Sparsity plots depict the difference between the location of true and reconstructed anomalies for Example 3. The black dots indicate which voxels are in the true anomaly that are missed by the reconstruction while the plus marks indicate voxels in the reconstruction that are not in the anomaly. Clearly the reconstruction has the anomaly shifted slightly to the right in x. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Figure 6: Sparsity plots depict the location of true and reconstructed anomalies for Example 4. The black dots indicate the true anomaly, the centers of the black circles show areas of the reconstructed anomaly that lie outside the true anomaly (in other words, the reconstruction

is represented by the dots plus the circles). The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Figure 7: True image used in Examples 5A, 5B. True α value is .15. Colormap is scaled to show background detail.

Figure 8: Cross-sectional display of locations of reconstructed ellipsoid (dots and plus marks) vs. true ellipsoid (plus marks) for Example 5A. The oversized reconstruction is the result of not appropriately accounting for a lumpy background.

Figure 9: Reconstructed perturbation for example 5B. The reconstructed value of α is 0.13. The colormap truncation is the same as that used in Figure 7.

Figure 10: Cross-sectional slices moving down from the surface for the true solution in Section 4 C with $s_c = .005, w = -1$. True value of α is 0.15: colormap is truncated to show background detail.

Figure 11: Cross-sectional slices moving down from the surface for the true solution in Section 4 C for $s_c = .005, w = -2$.

Figure 12: Cross-sectional slices moving down from the surface for the reconstructed solution in Section 4 C for $s_c = .005, w = -2$. Reconstructed α value is .16 and the colormap truncation is the same as was used in Figure 11.

Figure 13: Example 8A: Cross-sectional plots moving down from the surface comparing the shape/location of the true object with the reconstructed shape. The true anomaly is marked with black dots, the reconstruction is overlaid using plus marks. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.

Figure 14: Example 8B: Cross-sectional plots moving down from the surface comparing the shape/location of the true object with the reconstructed shape. The true anomaly is marked with black dots, the reconstruction is overlaid using plus marks. The horizontal axis is increasing x from left to right but the tick marks give matrix column index. The vertical axis is decreasing y from top to bottom and the tick marks indicate matrix row index.