Three-Dimensional Diffuse Optical Tomography with a priori Anatomical Information

Murat Guven[a] Birsen Yazici[a] Xavier Intes[b,c] Britton Chance[b]

[a]Electrical and Computer Engineering Department, Drexel University, USA
[b]Department of Biophysics and Biochemistry, University of Pennsylvania, USA
[c]Department of Physics and Astronomy, University of Pennsylvania, USA

ABSTRACT

Diffuse Optical Tomography (DOT) image reconstruction is a challenging 3D problem with a relatively large number of unknowns. DOT poses a typical ill-posed problem usually plagued by under-determination, which complicates the inverse problem. Conventional image reconstruction algorithms can not provide high spatial resolution and may become computationally expensive and unreliable especially in the presence of noise.

In this work, we extend our previous formulation for the 3D inverse DOT problem, where we focus to improve the spatial resolution and quantitative accuracy of 3D DOT images by using anatomical a priori information, which is specific to the medium of interest. Maximum A Posteriori (MAP) estimate of the image is formed based on the formulation of the image's probability density function, which is extracted from the available a priori anatomical information. An "alternating minimization" algorithm, which sequentially updates the unknown parameters, is used to solve the resulting optimization problem. Proposed method is evaluated in a $3D$ simulation experiment. Results demonstrate that the proposed method leads to significantly improved spatial resolution, quantitative accuracy and faster convergence than standard and regularized least squares solutions even in the presence of noise. As a result, the approach demonstrated in this paper both addresses the ill-posedness and balances the computation complexity vs. image quality trade-off in the 3D DOT inverse problem.

Keywords: Three-dimension, Diffuse Optical Tomography, inverse problem, a priori anatomical information, spatial resolution

1. INTRODUCTION

In this work we formulate the inverse problem within a Bayesian framework with a spatially varying a priori probability density function extracted from a priori anatomical information. Proposed approach leads to a clinically applicable image reconstruction algorithm that can be easily incorporated to a combined imaging system, such as the novel MRI-NIR concurrent imaging system developed at the University of Pennsylvania Optical Imaging and Spectroscopy Laboratory.[26]

1.1. Related Work

Incorporation of structural information as the a priori information in inverse medical imaging problems has been suggested in Refs. 18-19. In Ref. 13, MR images were used to generate a finite element mesh to reconstruct a simulated rat cranium, where the available information, about the structure and optical properties, is used for the initial guess in the inversion algorithm similar to the approach in Ref. 32. In Ref. 33, the optical properties of certain regions are constrained to the available values, based on the a priori structural information. In Ref. 14, the structural information is used to reconstruct images at a low resolution level in order to obtain a good initial guess for the high-resolution solution of the same problem.

Use of MR scans has been employed for optical breast imaging[15,17] where the "reference medium" is obtained from accurate optical properties of the tissue, with the anatomy derived from MR images. In Ref. 29, based on MR prior information, it has been suggested to unite pixels that belong to the same tissue

Photon Migration and Diffuse-Light Imaging, David A. Boas, Editor,
Proceedings of SPIE-OSA Biomedical Optics, SPIE Vol. 5138 (2003)

type into one non-uniform pixel to reduce the number of unknowns. While this approach is computationally viable, it is unable to resolve local variations within the same tissue type. A Bayesian formulation of the DOT inverse problem has been suggested in Refs. 8-12 to incorporate a priori information. Introducing penalty functions[11] and uniform[2–5] or spatially varying regularization[12] terms in the inverse problem formulation is another way of incorporating a priori information. However, in these studies, rather than specific information about the unknown image, generic probability density functions or regularization terms have been used. Therefore, the improvements have been relatively limited.

1.2. Proposed Method

We have formulated the $3D$ inverse DOT problem within a Bayesian framework based on a linear forward model. The available high resolution anatomical image is segmented into sub-images that represent major tissue types. Prior probability density function of the image is formulated such that each sub-image is assigned a mean value that need not be equal to its actual optical value; and a "confidence level" is defined in the form of an image variance formulation to allow local variations within sub-images. As a consequence, the overall formulation of the prior information becomes spatially varying, which is specific to the image of interest. Maximum A Posteriori (MAP) estimate of the image is formed based on the formulation of the image's probability density function. An "alternating minimization" algorithm, which sequentially updates the unknown parameters, is used to solve the resulting optimization problem. The proposed method is tested extensively using a $3D$ numerical simulation. Our results show that high resolution anatomical image obtained from a secondary modality can significantly improve the spatial resolution, the quantitative accuracy and the speed of DOT image reconstructions, even in the presence of noise.

Even though the presented results are based on the linearization of the forward model around a homogeneous background; the Bayesian formulation of the inverse problem based on a priori anatomical information and the alternating minimization scheme depicted in this paper allow inclusion of an iterative update scheme for the forward problem solution as the image reconstruction problem is resolved.

1.3. Organization of the Paper

The paper is organized as follows: Section 2.1 defines the forward model and section 2.2 provides the background on the formulation of the inverse problem. Section 2.3 and 2.4 formulate the data likelihood function and prior information, respectively. Section 2.5 formulates the inverse problem within a Bayesian framework. Section 2.6 describes the "alternating minimization" algorithm. Simulation experiment results are demonstrated in section 3. Conclusion is presented in section 4.

2. PROBLEM FORMULATION

2.1. Forward Model

In order to characterize a forward model, we need to model the photon propagation in the scattering medium. In this work, we focus on the reconstruction of absorption coefficients, hence we keep scattering coefficient and consequently the diffusion coefficient constant throughout our calculations. As a result, the use of the following diffusion equation given in frequency domain is sufficient to define the forward model:

$$-\frac{i\omega}{c}\phi(r) - D \bigtriangledown^2 \phi(r) + \mu_a(r)\phi(r) = A\delta(r_s) \tag{1}$$

where $\phi(r)$ represents the spatially varying optical field due to the point source $A\delta(r_s)$ located at $r = r_s$. ω denotes the frequency, c is the speed of light and $i = \sqrt{-1}$. D is the spatially invariant diffusion coefficient and $\mu_a(r)$ stands for the spatially varying absorption coefficient.

Phase $\Phi(r)$ of the optical field $\phi(r)$ can be expressed in terms of $\mu_a(r)$ using the perturbation theory,[20–22] by expressing the spatially varying absorption coefficient as $\mu_a(r) = \mu_{a0} + \delta\mu_a(r)$, where μ_{a0} is the spatially invariant background absorption coefficient and $\delta\mu_a(r)$ is the spatially varying component.

Discretization of the 3D domain into N uniform voxels yield the following relationship[16]:

$$y = W \times x \tag{2}$$

where we have applied Rytov approach[20] using a first order approximation. y denotes the measurement vector holding the perturbative Rytov phase $\Phi_{sc}(r, r_s)$ for each source-detector pair, W represents the linearized forward model (weight matrix) which relates the differential absorption coefficient distribution $x = \begin{bmatrix} \delta\mu_a(r_1) & \delta\mu_a(r_2) & \cdots & \delta\mu_a(r_N) \end{bmatrix}^T$ to the measurement vector y. Note that, linearization of the forward model around a homogeneous background is suitable for a medium, such as human breast, with absorption coefficient values less than 0.3.[1]

2.2. Inverse Problem

The inverse problem in imaging, namely the restoration and reconstruction of images can be defined as the general problem of estimating a 3-D field $x(r)$ from some form of indirect observations related to this field. The problem can be stated as follows:

$$y = T\{x\} + \zeta \tag{3}$$

where y represents the measurements, x denotes the spatially varying 3-D field, ζ is the additive noise in the measurement system and T is an operator that relates the unknown field x to the available measurements y.

According to our forward model, T is replaced by the weight matrix W and x denotes the unknown spatial distribution of the differential absorption coefficient $\delta\mu_a$.

The least squares (LS) solution can be formulated for this inverse problem as:

$$\hat{x}_{LS} = arg\min_x \|y - Wx\|_2^2 \tag{4}$$

where $\|.\|_2$ denotes the L_2-norm. Due to the ill-posed and/or underdetermined nature of the DOT problem, the LS solution is typically not robust. When the available measurement data is noisy, the noise components will be highly amplified in the LS solution, which is not a desired outcome. This difficulty can be resolved by incorporating a priori information of the image, which leads to a reduced mean-square error at the expense of a bias introduced to the solution.[30] The basic idea is to constrain the solution such that the amplified noise effects are avoided.

The Bayesian approach provides a natural framework to incorporate prior information. Maximum a posteriori (MAP) estimate of the image is given by:

$$\hat{x}_{MAP} = arg\max_x \{log\ p(x|y)\} \tag{5}$$

where $p(x|y)$ is the conditional probability density function of x given the measurements y. Equivalently,

$$\hat{x}_{MAP} = arg\max_x \{log\ p(y|x) + log\ p(x)\} \tag{6}$$

where $p(y|x)$ is the data likelihood function and $p(x)$ is the probability density function of the corresponding image. In the following two subsections, we shall discuss modelling the data likelihood, and prior information which is designed to reflect the structural and statistical properties of the image.

2.3. Formulation of the Data Likelihood Function

The measurement vector y is formed as

$$y = \begin{bmatrix} y_{11}^{f_1} & y_{12}^{f_1} & \cdots & y_{1D}^{f_1} & y_{21}^{f_1} & \cdots & y_{SD}^{f_1} & y_{11}^{f_2} & \cdots & y_{SD}^{f_F} \end{bmatrix}^T \tag{7}$$

where S is the number of sources, D is the number of detectors, and F is the number of frequencies associated with each source. The total number of measurements is then equal to $P = S \times D \times F$. Since we are interested in only the real part of the measurements, $y_{ij}^{f_k}$ are all real.

Photon detection can be modelled using shot noise statistics, which originates from Poisson statistics.[6] With a sufficiently large number of detected photons, the Poisson statistics can be approximated by a Gaussian distribution, uncorrelated between measurements and with a variance proportional to the measurement values. Consequently each measurement $y_{ij}^{f_k}$ is assumed to be an independent real Gaussian random variable and the data likelihood is given by

$$p(y|x) = \frac{1}{K|\Lambda_\zeta(\lambda)|^{1/2}} exp(-\frac{1}{2}\|y - Wx\|^2_{\Lambda_\zeta^{-1}(\lambda)}) \tag{8}$$

where $\Lambda_\zeta(\lambda)$ is the covariance matrix of size $P \times P$, K is the normalization constant and $\|z\|^2_\Lambda = z^T \Lambda z$. Under the assumption of statistical independence, $\Lambda_\zeta(\lambda)$ becomes a diagonal matrix:

$$\Lambda_\zeta(\lambda) = \lambda\Lambda_y = \begin{bmatrix} \lambda\sigma^2_{\zeta_1} & 0 & 0 & \cdots & 0 \\ 0 & \lambda\sigma^2_{\zeta_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \lambda\sigma^2_{\zeta_P} \end{bmatrix} \tag{9}$$

where $\sigma^2_{\zeta_p}$ is equal to the absolute value of the corresponding measurement $y_{ij}^{f_k}$, hence $diag(\Lambda_y) = |y|$ and the unknown term λ is related to the noise variance.

2.4. Formulation of the Prior Information

The available a priori anatomical information offers the opportunity to formulate a specific prior probability density function for the unknown image. Prior information comprising the smoothness and the optical boundaries provides additional advantages as compared to generic assumptions which can not take particular structural properties of the image of interest into account. In the case of DOT, such anatomical information can be obtained from concurrently or sequentially acquired Magnetic Resonance (MR) or X-ray images.

We assume that the optical and anatomical edges are the same. Available edge information is utilized to decompose the unknown image x into M homogeneous sub-images. Volume elements of each sub-image are assumed to possess the same second-order statistical properties. Relying on the anatomical prior, each sub-image x_i is assigned a single mean value, μ_i which is approximately known in prior, and a single variance σ_i^2, which is an unknown parameter. Under these assumptions, the probability density function of the i^{th} sub-image becomes

$$p(x_i|\sigma_i) = \frac{1}{(2\pi\sigma_i^2)^{N_i/2}} exp(-\frac{1}{2\sigma_i^2}\|x_i - \mu_i\|^2), \quad i = 1, 2, ..., M \tag{10}$$

where M is the number of sub-regions and N_i is the number of voxels in the i^{th} sub-image. As a result the probability density function of the whole image is modelled as follows:

$$p(x|\sigma) = \frac{1}{(2\pi)^{N/2}|\Lambda_x(\sigma)|^{1/2}} exp(-\frac{1}{2}\|x - \mu\|^2_{\Lambda_x^{-1}(\sigma)}) \tag{11}$$

where σ is the vector holding the variances that belong to each sub-image

$$\sigma = \begin{bmatrix} \underbrace{\sigma_1^2 \cdots \sigma_1^2}_{N_1} & \underbrace{\sigma_2^2 \cdots \sigma_2^2}_{N_2} & \cdots & \underbrace{\sigma_M^2 \cdots \sigma_M^2}_{N_M} \end{bmatrix}^T \tag{12}$$

and $\Lambda_x(\sigma)$ is the covariance matrix of the image x. Further assumption of the uncorrelated sub-images result in an image whose pixels are all uncorrelated. Therefore the covariance matrix, $\Lambda_x(\sigma)$ becomes a diagonal

matrix

$$\Lambda_x(\sigma) = \begin{bmatrix} \sigma_1^2 I_{N_1 \times N_1} & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 I_{N_2 \times N_2} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_M^2 I_{N_M \times N_M} \end{bmatrix} \tag{13}$$

where $I_{N_i \times N_i}$ is an $N_i \times N_i$ identity matrix and μ is the vector holding the mean values assigned to each one of the pixels

$$\mu = \begin{bmatrix} \underbrace{\mu_1 \cdots \mu_1}_{N_1} & \underbrace{\mu_2 \cdots \mu_2}_{N_2} & \cdots & \underbrace{\mu_M \cdots \mu_M}_{N_M} \end{bmatrix}^T \tag{14}$$

The function of the covariance matrix is two-fold. As for the more obvious one, it incorporates the prior information together with mean value vector and serves together with the noise variance as the regularization parameter. The covariance matrix also defines lower and upper bounds for the deviation from the mean values that are assigned to the image. In other words, it determines the "confidence level" for the assigned mean values: lower the variance lower is the deviation that is expected from the mean.

In order to incorporate the confidence level into the statistical reconstruction procedure, we formulate the probability density function of the sub-image variances. We assume that each of these parameters is also Gaussian distributed with a positive mean and a positive variance that ensures a positive value for the parameter. This assumption in fact gives the power to control the sub-image variance by constraining it to realistic values. The probability density function of the sub-image variance becomes

$$p(\sigma_i) = \frac{1}{(2\pi\gamma_i^2)^{N_i/2}} exp(-\frac{1}{2\gamma_i^2}\|\sigma_i - \mu_{\sigma_i}\|^2), \quad i = 1, 2, ..., M \tag{15}$$

where γ_i is the variance and μ_{σ_i} is the mean value of σ_i, which are all known prior. Assuming statistically independent variances joint distribution $p(\sigma) = p(\sigma_1, \sigma_2, \cdots, \sigma_N)$ can be written as

$$p(\sigma) = \frac{1}{(2\pi)^{N/2}|\Lambda_\sigma(\gamma)|^{1/2}} exp(-\frac{1}{2}\|\sigma - \mu_\sigma\|^2_{\Lambda_\sigma^{-1}(\gamma)}) \tag{16}$$

where $\Lambda_\sigma(\gamma) = \Lambda_x(\sigma)|_{\sigma_i=\gamma_i}$ and $\mu_\sigma = \mu|_{\mu_i=\mu_{\sigma_i}}$ where $i = 1, 2, ..., M$.

2.5. Bayesian Formulation of the Inverse Problem

After forming the data likelihood and image probability density functions, the inverse problem becomes the reconstruction of the image x together with estimation of the unknown parameter λ and the unknown vector of parameters σ. Assuming that all of the parameters are on the same level,[27] the joint MAP estimate of these parameters can be stated as

$$(\hat{x}, \hat{\lambda}, \hat{\sigma})_{MAP} = arg \max_{x,\lambda,\sigma} p(x, \lambda, \sigma|y) \tag{17}$$

where

$$p(x, \lambda, \sigma|y) = p(y|x, \lambda, \sigma)p(x, \lambda, \sigma) \tag{18}$$

Since the image variance σ does not affect the measurement distribution, we can write equivalently

$$p(x, \lambda, \sigma|y) = p(y|x, \lambda)p(x, \lambda, \sigma) \tag{19}$$

The first term in Eq. (19) is the data likelihood function which is derived in Eq. (8). The second term in Eq. (19) can be expressed as

$$p(x, \lambda, \sigma) = p(x|\sigma, \lambda)p(\lambda, \sigma) \tag{20}$$

Assuming that the noise variance related parameter λ has no influence on the image distribution, the first term in the above equation becomes the distribution given by Eq. (11). Expanding the second term $p(\lambda, \sigma)$ we get

$$p(\lambda, \sigma) = p(\sigma|\lambda)p(\lambda) \tag{21}$$

If we assume that the image variance and the noise variance are statistically independent, which is a reasonable assumption, we can rewrite the above relationship as

$$p(\lambda, \sigma) = p(\sigma)p(\lambda) \tag{22}$$

Consequently Eq.(20) can be rewritten as

$$p(x, \lambda, \sigma) = p(x|\sigma)p(\sigma)p(\lambda) \tag{23}$$

As a result, $p(x, \lambda, \sigma|y)$ can be expressed as

$$p(x, \lambda, \sigma|y) = p(y|x, \lambda)p(\lambda)p(x|\sigma)p(\sigma) \tag{24}$$

The joint MAP estimates $\hat{x}, \hat{\lambda}, \hat{\sigma}$ are then equal to

$$(\hat{x}, \hat{\lambda}, \hat{\sigma})_{MAP} = arg \max_{x,\lambda,\sigma} p(y|x, \lambda)p(\lambda)p(x|\sigma)p(\sigma) \tag{25}$$

Assuming a uniform probability density function for the noise variance parameter λ and using the corresponding probability density functions, the joint MAP estimation problem becomes

$$(\hat{x}, \hat{\lambda}, \hat{\sigma})_{MAP} = arg \min_{x} \min_{\lambda} \min_{\sigma} \{log(|\Lambda_\zeta(\lambda)|) + \|y - Wx\|^2_{\Lambda_\zeta^{-1}(\lambda)} + log(|\Lambda_x(\sigma)|) + \|x - \mu\|^2_{\Lambda_x^{-1}(\sigma)} + \|\sigma - \mu_\sigma\|^2_{\Lambda_\sigma^{-1}(\gamma)}\} \tag{26}$$

In order to make some observations, let's assume that the estimate $\hat{\lambda}$ of the parameter λ and the estimate $\hat{\sigma}$ of the vector σ are known, then the MAP estimate $\hat{x}_{MAP}$ becomes

$$\hat{x}_{MAP} = arg \min_{x} \{\|y - Wx\|^2_{\Lambda_\zeta^{-1}(\hat{\lambda})} + \|x - \mu\|^2_{\Lambda_x^{-1}(\hat{\sigma})}\} \tag{27}$$

Taking the derivative of the above expression and setting it equal to zero yields an equation which is linear in x. Solution of this equation is the image estimate $\hat{x}$.

$$[W^T \Lambda_\zeta^{-1} W + \Lambda_x^{-1}]\hat{x}_{MAP} = W^T \Lambda_\zeta^{-1} y + \Lambda_x^{-1} \mu \tag{28}$$

Note that, in regularization framework, equation (28) can be interpreted as a zero-order Tikhonov regularization.[23] In the limiting case, as the image variance σ goes to 0, the confidence to the mean values μ improves and the terms Λ_x^{-1} and $\Lambda_x^{-1}\mu$ become dominant in equation (28). Consequently, the estimate $\hat{x}_{MAP}$ is constrained more towards the assigned mean values, which results in separation of the sub-images with respect to their mean values with edges whose sharpness is controlled by the variance parameter σ.

2.6. Inversion using Alternating Minimization

Let $\Psi(x, \lambda, \sigma)$ denote the objective function to be minimized as derived in equation (26), with the unknown parameters λ, σ, and the image x.

$$\Psi(x, \lambda, \sigma) = \{log(|\Lambda_\zeta(\lambda)|) + \|y - Wx\|^2_{\Lambda_\zeta^{-1}(\lambda)} + log(|\Lambda_x(\sigma)|) + \|x - \mu\|^2_{\Lambda_x^{-1}(\sigma)} + \|\sigma - \mu_\sigma\|^2_{\Lambda_\sigma^{-1}(\gamma)}\} \tag{29}$$

An alternating minimization algorithm,[28] which successively updates the unknown parameters while a minimum of the objective function is sought, is employed to compute the image estimate $\hat{x}$. At each iteration of the optimization, the objective function is minimized sequentially with respect to the unknown parameters

λ and σ and a new update for the image x is computed using the updated values of these parameters. More specifically at each iteration, the unknown parameters are computed as

$$\begin{aligned} \hat{\lambda} &= arg\min_{\lambda} \Psi(\hat{x}, \lambda, \hat{\sigma}) \\ \hat{\sigma} &= arg\min_{\sigma} \Psi(\hat{x}, \sigma, \hat{\lambda}) \end{aligned} \tag{30}$$

and the image estimate is updated with one iteration of Conjugate Gradient (CG) algorithm (Table 1)

$$\hat{x} \leftarrow CG\_update\{\Psi(x, \hat{\lambda}, \hat{\sigma})\} \tag{31}$$

where the α parameter in the CG algorithm is computed by the exact line search and Polak-Ribière-Polyak method[24, 25] is used to calculate β parameter. Under certain conditions,[29] the alternating minimization converges monotonically to a global minimum.

Table 1. Conjugate Gradient Algorithm

Initialize the image estimate to zero: $\hat{x}^{(0)} = \vec{0}$
Assign the mean value vector μ its entries corresponding to each individual region x_i
Define a confidence level on the mean and compute the corresponding μ_{σ_i}
Initialize the image variance σ; $\sigma > 0$
Initialize the parameter λ: $\hat{\lambda}^{(0)} = 1$
Find gradient vector: $\vec{g}^{(0)} = -\Omega(\hat{x}^{(0)}) = -\frac{\partial \Psi(x,\lambda,\sigma)}{\partial x}|_{x=\hat{x}^{(0)}}$
Set search direction: $\vec{d}^{(0)} = \vec{g}^{(0)}$
Define termination criterion ε
Set iteration counter: $n = 0$
repeat
Find $\alpha^{(n)}$ such that: $\alpha^{(n)} = arg\min_{\alpha>0} \Psi(x^{(n)} + \alpha d^{(n)}, \hat{\lambda}, \hat{\sigma})$
$x^{(n+1)} = x^{(n)} + \alpha^{(n)}\vec{d}^{(n)}$
Update λ using equation (32)
Update σ using equation (39)
$\vec{g}^{(n+1)} = -\Omega(\hat{x}^{(n+1)})$
$\beta^{(n+1)} = max(\frac{\vec{g}^{(n+1)T}(\vec{g}^{(n+1)} - \vec{g}^{(n)})}{\vec{g}^{(n)T}\vec{g}^{(n)}}, 0)$
$\vec{d}^{(n+1)} = \vec{g}^{(n+1)} + \beta^{(n+1)}\vec{d}^{(n)}$
$n = n + 1$
until $\|\alpha^{(n)}\Omega(\hat{x}^{(n)})\| < \varepsilon$

The minimization of the cost function with respect to the parameter λ given the updated estimates $\hat{x}$ and $\hat{\sigma}$ results

$$\hat{\lambda} = \frac{\|y - W\hat{x}\|^2_{\Lambda_y^{-1}}}{P} \tag{32}$$

In order to find an estimate for the vector σ, we make use of the probability density function formulation for each individual sub-image and rewrite the objective function with σ_i dependent terms:

$$\Psi_{\sigma_i}(\sigma_i) = \{log(\sigma_i^{N_i}) + \frac{1}{2\sigma_i^2}\|x_i - \mu_i\|^2 + \frac{1}{2\gamma_i^2}\|\sigma_i - \mu_{\sigma_i}\|^2\} \tag{33}$$

After taking the derivative of the above expression, the estimate for the sub-image variance σ_i^2 satisfies the following equation:

$$N_i\sigma_i^2 - \|x_i - \mu_i\|^2 + \frac{N_i}{\gamma_i^2}\sigma_i^3(\sigma_i - \mu_{\sigma_i}) = 0 \tag{34}$$

This is a fourth order equation in σ_i, which is rather difficult to solve. In order to simplify the solution, we make use of the following approximation:

$$\mu_{\sigma_i}\sigma_i^3 \approx \mu_{\sigma_i}^2 \tag{35}$$

and equation (34) becomes a quadratic equation of the parameter σ_i^2

$$\frac{N_i}{\gamma_i^2}\sigma_i^4 + N_i\sigma_i^2 - (\|x_i - \mu_i\|^2 + N_i\frac{\mu_{\sigma_i}^2}{\gamma_i^2}) = 0 \tag{36}$$

The discriminant of this equation can be evaluated as

$$\triangle = N_i^2 + 4\frac{N_i}{\gamma_i^2}(\|x_i - \mu_i\|^2 + N_i\frac{\mu_{\sigma_i}^2}{\gamma_i^2}) > 0 \tag{37}$$

from which the positive value for the variance σ_i^2 can be calculated as

$$\sigma_i^2 = \frac{-N_i + \sqrt{\triangle}}{2N_i/\gamma_i^2} > 0 \tag{38}$$

Note that in the limiting case, for $\gamma_i^2 \to \infty$, the sub-image variance estimate σ_i^2 converges to

$$\lim_{\gamma_i^2\to\infty} \hat{\sigma}_i^2 = \lim_{\gamma_i^2\to\infty} \frac{-N_i + \sqrt{\triangle}}{2N_i/\gamma_i^2} = \frac{\|\hat{x}_i - \mu_i\|^2}{N_i}, \quad i = 1, 2, ..., M \tag{39}$$

which is the ML estimate for the variance where one can assume a uniform probability density function instead of the Gaussian.[16]

3. SIMULATION EXPERIMENT

We have performed a $3D$ simulation experiment to demonstrate the performance of the proposed method. We simulated optical measurements using a finite difference code for predetermined geometries, whose cross-sections are shown in figure 1..

We evaluate the performance of image reconstructions in terms of convergence rate and error norm given as

$$\epsilon(n) = \Sigma_{i=1}^{M}\|\hat{x}_i^{(n)} - x_i^*\|^2 \tag{40}$$

where x_i^* is the actual $\delta\mu_a$ value in sub-image x_i, and $\hat{x}_i^{(n)}$ is the estimated image in sub-image x_i after n^{th} iteration.

3.1. Simulation Experiment

The diffusion equation is solved using finite-difference method for the 3D medium whose cross-sectional views are shown in Fig.1. The background absorption coefficient μ_{a0} is equal to 0.05 cm^{-1} and the diffusion constant $D \cong 1/3\mu_s' = 1/30$ cm. The cylinder inclusion has an absorption coefficient $\mu_a = 0.15$ cm^{-1} that corresponds to a $\delta\mu_a$ value of 0.1 cm^{-1}. The heterogeneity shown on the left has an absorption coefficient of 0.22 cm^{-1} leading to a differential absorption coefficient of 0.17 cm^{-1}. The scattering coefficient hence the diffusion coefficient of both heterogeneities are equal to those of the background. The medium is discretized into $5 \times 30 \times 40$ pixels along z, y and x directions each with the dimension $0.60 \times 0.20 \times 0.20$ cm^3. 35 sources that operate at $\omega = 0$ frequency are distributed evenly along xz-plane and 40 detectors are located on the opposite side of the medium (figure 2), as a result the number of total measurements is equal to $35 \times 40 = 1400$. The image is divided into three sub-images for the formulation of prior information, hence $\mu = [\mu_1 \cdots \mu_1\ \mu_2\ \cdots \mu_2\ \mu_3 \cdots \mu_3]^T$, where μ_1 is the mean value of the background image, μ_2 is the mean value of the heterogeneity on the left, and μ_3 is the mean value of the heterogeneity on the right, as shown in figure 1. Two main cases are examined in the reconstruction procedure and LS, and regularized LS solutions are also presented for comparison.

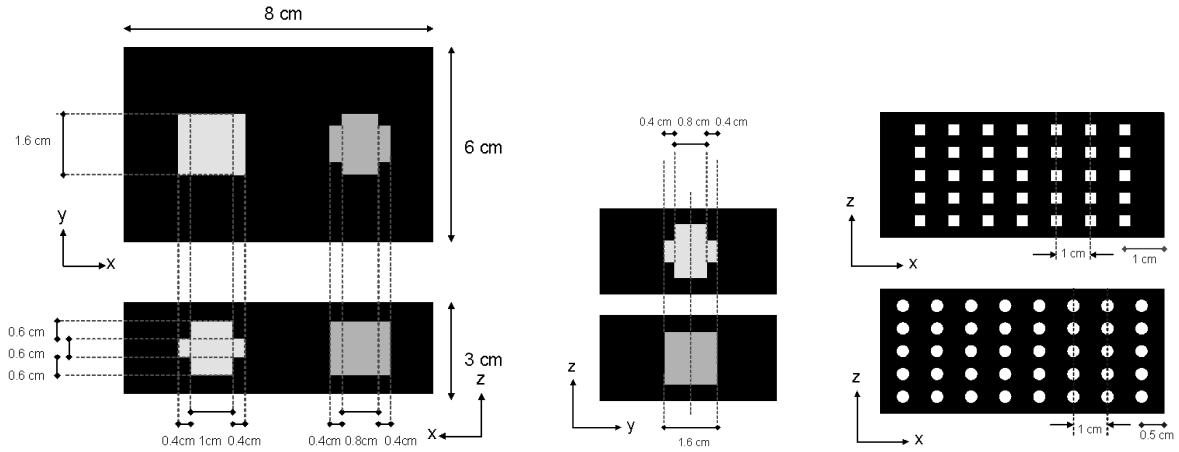


Figure 1. The cross-sectional views of the 3D medium used in the simulation experiment. The cross-section along xy plane is taken at $z = 1.5$ cm; the cross-section along xz plane is drawn for $y = 3$ cm and the cross-sections along zy plane correspond to center of the objects along x-axis. The source (squares) and detector (circles) positions, at $y = 0$ and $y = 6$ cm respectively are shown on the right.

3.1.1. Case 1

First set of experiments are done in the absence of noise. The $3D$ reconstruction results in figure 3 are shown in $2D$ images that correspond to top 3 slices of the $3D$ medium. Each reconstruction has been done for different values of the mean value vector μ and with a confidence level of 30%, which implies that the mean value of the standard deviation μ_{σ_i} in each sub-region x_i is equal to 30% of the assigned mean value. Results show that the proposed algorithm is not sensitive to the mean values assigned in the inverse problem formulation. While the error norm shows that the performance is enhanced as the mean values are closer to the exact values (Figure 5(a)), all of the reconstructed images present high accuracy with preserved edges.

The proposed method achieves much better results when compared to LS and regularized LS solution, even in the absence of noise. Figure 3 shows the top 3 slices of the reconstructed $3D$ images. Regularized LS solution provides a smoother image estimate than the LS solution however their convergence rates follow a similar track (Figure 5(b)).

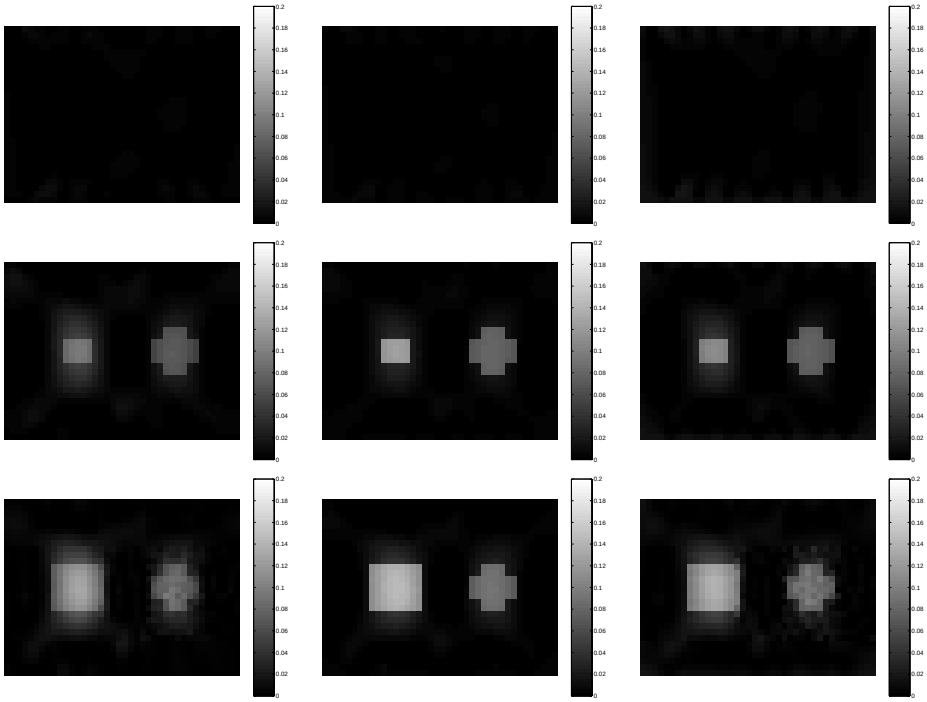
3.1.2. Case 2

For this case, we add zero-mean Gaussian noise to our measurement vector y, with a Signal-to-Noise Ratio (SNR) value of 10 dB where we have defined SNR as

$$SNR = 10 \log_{10}\left(\frac{Mean\ value\ of\ the\ measurements}{Standard\ deviation\ of\ the\ additive\ noise}\right) \tag{41}$$

The LS solution, after some iterations, diverges and can not present reliable results (figure 4). Although the regularized LS solution, with a carefully selected regularization parameter, provides a robust solution, the image is poor in spatial resolution and the solution is not accurate quantitatively (figure 4(b)).

Application of the proposed algorithm with the formulation of the prior information based on exact mean values and 30% confidence level results in superior results with a faster convergence rate compared to regularized LS solution. The achieved performance is slightly worse than the ideal case where there is no noise (figure 6(c)).



(a) (b) (c)

Figure 2. Top 3 slices are shown. All reconstructions are done in the absence of noise and top three slices of the reconstructed 3D medium in the absence of noise are shown (a) The reconstruction with prior information formulation with $\mu_1 = 0.01$, $\mu_2 = 0.08$, $\mu_3 = 0.15$, (b) The reconstruction with prior information formulation with mean values equal to exact $\delta\mu_a$ values (c) The reconstruction with prior information formulation where $\mu_1 = 0.03$, $\mu_2 = 0.12$, $\mu_3 = 0.20$.

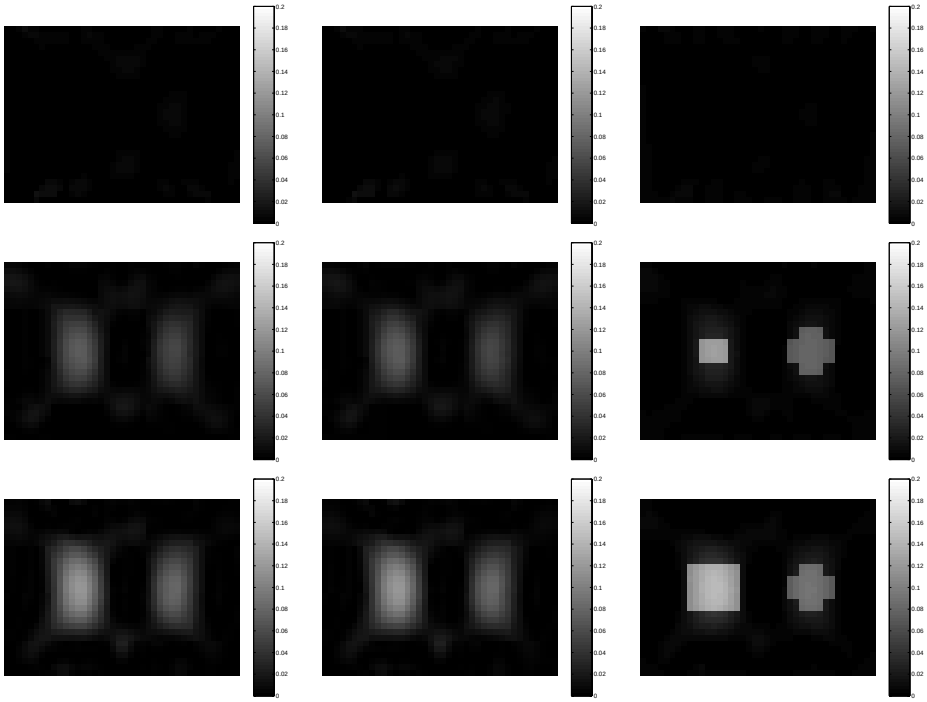
4. CONCLUSION

Previously we had examined the potential utility of incorporating anatomical prior information within a Bayesian framework and demonstrated our method with simulation and phantom experiments for 2D case. In this work, we have extended the previous formulation and applied it for the 3D case. Results show that incorporating specific structural information improves the DOT image reconstruction in spatial resolution, quantitative accuracy and speed of convergence. The high performance achieved especially in the presence of noise is another outcome of the proposed algorithm. Even though the presented results are based on the linearization of the forward model around a homogeneous background, the Bayesian framework formulation and the alternating minimization scheme depicted in this paper allow extension of the inverse problem formulation to include an iterative update scheme for the forward problem solution.

The results obtained in this work are encouraging for our future studies on combined NIR-MRI breast imaging, where the anatomical information will be extracted from the concurrent MR images.

ACKNOWLEDGMENTS

The authors are grateful to M. Holboke for developing the finite difference code. X.I. and B.C. acknowledge partial support from NIH CA 87046.

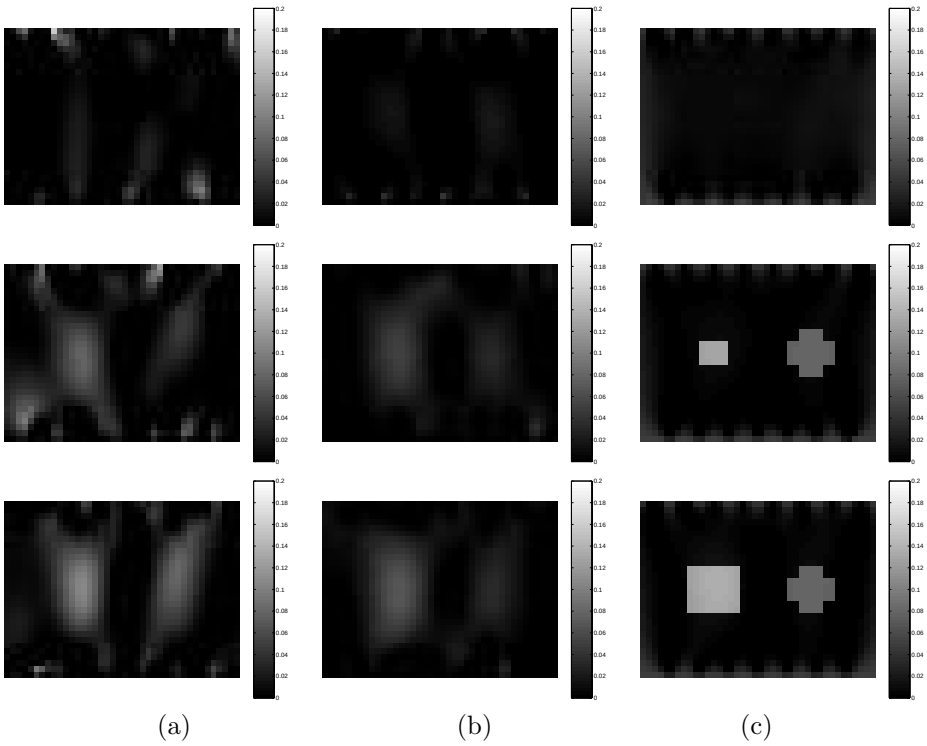


(a) (b) (c)

Figure 3. All reconstructions are done in the absence of noise and the top three slices of the reconstructed 3D medium are shown (a) LS solution (b) regularized LS solution (c) Proposed method with mean values identical to exact $\delta\mu_a$ values.

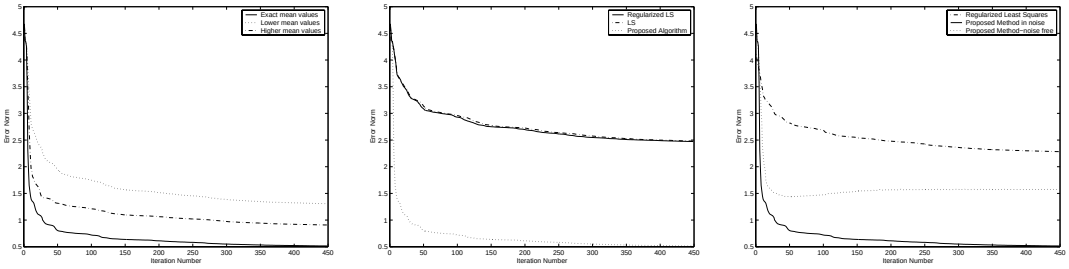
REFERENCES

1. D. A. Boas,"A fundamental limitation of linearized algorithms for diffuse optical tomography," *Optics Express*, Vol.**1**, No. 13, pp.404-413, 1997.
2. K. D. Paulsen and H. Jiang, "Enhanced frequency domain optical image reconstruction in tissues through total variation minimization," *Appl. Opt.* **35**, pp. 3447-3458, 1996.
3. S. R. Arridge, "Forward and inverse problems in time resolved infrared imaging," *Medical Optical Tomography*, G. Muller Ed., Vol. IS11, pp. 53-64, SPIE Optical Engineering, Bellingham, WA, 1993.
4. H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency- domain data: simulation and experiment," *J. Optical Society America A*, vol. **13**, no. 2, pp. 253-266, 1996.
5. Y. Yao, Y.Wang, Y. Pei, W. Zhu, and R. L. Barbour, "Frequency domain optical imaging of absorption and scattering distributions by a Born iterative method," *J. Optical Society America A*, vol. **14**, no. 1, pp. 325-342, 1997.
6. J. C. Ye, C. A. Bouman, K. J. Webb, and R. P. Millane, "Nonlinear Multigrid Algorithms for Bayesian Optical Diffusion Tomography," *IEEE Trans. on Image Processing*, pp. 909-922, vol. **10**, no. 6, 2001.
7. A. B. Milstein, S. Oh, J. S. Reynolds, K. J. Webb, and C. A. Bouman, "Three-dimensional Bayesian optical diffusion tomography with experimental data," *Optics Letters*, Vol.**27**, pp. 95-97, No. 2, 2002.
8. M.J. Eppstein, D.J. Hawrysz, A. Godavarty, and E.M. Sevick-Muraca, "Three dimensional near infrared fluorescence tomography with Bayesian methodologies for image reconstruction from sparse and noisy data sets," Proc. Nat. Acad. Sci., **99**, 9619-9624, 2002.



(a) (b) (c)

Figure 4. (a) LS reconstruction in the presence of noise (b)Regularized LS reconstruction (c) reconstruction with exact mean values and 30% of confidence level



(a) (b) (c)

Figure 5. (a) Error norms for the solution of the proposed algorithm where different values are assigned as the mean values (b) Error norms for solutions of LS, regularized LS and the proposed algorithm (c) Error norm for the solution of the proposed algorithm in the presence of noise; compared to solution given by regularized LS and the proposed algorithm for the ideal case: exact mean values and noise-free measurements

9. M. Guven, B. Yazici, X. Intes, B. Chance, and Y. Zheng, "Recursive Least Squares Algorithm for Optical Diffusion Tomography," *Proceedings of the IEEE 28th Annual Northeast*, pp. 273-274, 2002.
10. S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector Calibration in Three-dimensional Bayesian Optical Diffusion Tomography", *J. Opt. Soc. Am. A*, Vol. **19**, No. 10, pp. 1983-1993, 2002.
11. A. H. Hielscher, S. Bartel,"Use of penalty terms in gradient-based iterative reconstruction schemes for optical tomography," *Journal of Biomedical Optics*, Vol. **6**, pp.183-192, No. 2, 2001.
12. B. W. Pogue, T. O. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, "Spatially Varying

Regularization Improves Diffuse Optical Tomography," *Applied Optics*, Vol. **38**, No. 13, pp.2950-2961, 1999.
13. B. W. Pogue and K. D. Paulsen, "High-resolution near-infrared tomographic imaging simulations of the rat cranium by use of a priori magnetic resonance imaging structural information," *Optics Letters*, Vol. **23**, No. 21, pp. 1716-1718, 1998.
14. M. Schweiger and S. R. Arridge, "Optical tomographic reconstruction in a complex head model using a priori region boundary information," *Phys. Med. Biol.* **44**, pp. 2703-2721, 1999.
15. J. Chang, H. L. Graber, P. C. Koo, R. Aronson, S. L. Barbour, and R. L. Barbour, "Optical imaging of anatomical maps derived from magnetic resonance imaging using time-independent optical sources," *IEEE Trans. Med. Imaging* **16** pp.68-77, 1997.
16. M. Guven, B. Yazici, X. Intes, B. Chance, "Diffuse Optical Tomography with a priori Anatomical Information,"*SPIE*, 2003.
17. R. L. Barbour, H. L. Graber, J. Chang, S-L S. Barbour, P. C. Koo and R. Aronson, "MRI-Guided optical tomography: Prospects and computation for a new imaging method," *IEEE Computational Science and Engineering*, pp.63-76, Winter 1995.
18. C-T Chen, X. Ouyang, W. H. Wong, X. Hu, V. E. Johnson, C. Ordonez, and C. E. Metz, "Sensor fusion in image reconstruction," *IEEE Trans. Nuclear Science* **38** pp.687-692, 1991.
19. G. Gindi, M. Lee, A. Rangarajan, and I. G. Zubal, "Bayesian reconstruction of functional images using anatomical information as priors," *IEEE Trans. Med. Imaging* **12** pp.670-680, 1993.
20. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography," *Opt. Lett.* **20**, pp.426-428, 1995.
21. S. R. Arridge, "Photon Measurement Density Functions. Part 1: Analytical Forms," *Appl. Opt.* **34**, pp. 7395-7409, 1995.
22. A. C. Kak, and M. Slaney,"*Principles of Computerized Tomographic Imaging*," pp. 208-218, IEEE, New York, 1988.
23. A. Tikhonov and V. Arsenin, "*Solutions of Ill-Posed Problems,"* Winston and Sons, New York, 1977.
24. E. Polak, and G. Ribière, "Note sur la convergence de directions conjugèes", *Rev. Francaise Informat Recherche Opertionelle*, 3e annèe, 16 pp. 35-43, 1969.
25. B. T. Polyak, "The Conjugate Gradient Method in Extremem Problems," *USSR Comp. Math. Math. Phys.*, **9**, pp. 94-112, 1969.
26. V. Ntziachristos, A. G. Yodh, M. D. Schnall, and B. Chance, "MRI-guided Diffuse Optical Spectroscopy of malignant and benign breast legions," *Neoplasia*, Volume **4**, Issue 4, pp. 347-354, 2002.
27. A. Mohammad-Djafari, "Joint estimation of parameters in a Bayesian approach of solving inverse problems," in *Proc. IEEE Int. Conf. Image Processinf*, Vol. **2**, Lausanne, Switzerland, Sept. 16-19, 1996, pp.473-476.
28. I. Csiszar and G. Tusnady, "Information geometry and alternating minimization procedures," in *Statistics and Decisions*, Supplement Issue 1:205-237, 1984.
29. S. M. Haas, "The expectation-maximization and alternating minimization algorithms," September 11, 2002.
30. N. P. Galatsanos, A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," in *IEEE Transactions on Image Processing*, Vol. **1**, No. 3, July 1992.
31. H. Xu, H. Dehghani, B. W. Pogue, K. D. Paulsen, J. D. Dunn, "Hybrid MR/Near Infrared Imaging of the Muribe Brain: Optimization of optical fiber arrangement and use of a priori knowledge," Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on , 2002 Page(s): 74 -77.
32. H. Dehghani, B. W. Pogue, K. D. Paulsen,"Development of hybrid NIR/MRI imaging system algorithm: Use of a-priori information for tumor detection in the female breast," Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on , 2002 Page(s): 657 -660.