Normalized–constraint method for minimizing inter–parameter cross–talk in reconstructed images of spatially heterogeneous scattering and absorption coefficients

Yaling Pei¹, Harry L. Graber², and Randall L. Barbour²,
¹NIRx Medical Technologies Corp., 15 Cherry Lane, Glen Head, NY 11545
²SUNY Downstate Medical Center
450 Clarkson Ave., Brooklyn, NY 11203

ABSTRACT

In this report, we present a method to reduce the cross–talk problem in optical tomography. The method described is an extension of a previously reported perturbation formulation related to relative detector values, and employs a weight matrix scaling technique together with a constrained CGD method for imaging reconstruction. Results from numerical and experimental studies using DC measurement data demonstrate that the approach can effectively isolate absorption and scattering heterogeneities, even for complex combinations of perturbations in optical properties. The derived method is remarkably stable to errors originating from an insufficiently accurate estimate of properties of the reference medium.

Keywords: Light propagation in tissue, Optical tomography, Imaging reconstruction techniques, Regularization techniques.

1. INTRODUCTION

The ability to accurately define the absorption and scattering properties of tissue could significantly add to the diagnostic sensitivity and specificity of optical measurements. In the case of imaging studies, a common problem is the issue of parameter cross–talk. This refers to instances where, for example, localized variations in absorption also appear as localized variations in scatter. This issue can arise from fundamental reasons related to the intrinsic information content of a data set, as well as from reasons related the numerical methods used to compute parameter maps. In the case of DC measurement data, Arridge and Lionheart¹ and Hebden et al.² have claimed to have rigorously proven that there is an underlying non–uniqueness in the inverse problem, and have strongly emphasized that total intensity alone is insufficient to distinguish effects of absorption from scatter.

In this report we describe a new algorithm (i.e., the normalized–constraint algorithm) for DC imaging, and use it to demonstrate the ability to separate the effects of absorption from scatter under a wide range of conditions, as evaluated by numerical simulations and experimental laboratory studies. These results clearly demonstrate that, contrary to previous assertions,¹,² DC imaging methods are fully capable of characterizing spatial variations in the absorption and scattering properties of highly scattering media.

2. METHODS

The normalized–constraint algorithm for minimizing coefficient cross–talk employs a three–step process. First, we work with relative detector readings instead of absolute values. We do this in recognition of the practical limits imposed on obtaining reliable absolute measurement data from arbitrary structures such as tissue. We also do this in recognition that perturbation methods, in general, are sensitive to yield grossly incorrect solutions should the selected reference medium differ sufficiently from the actual target background properties. As described in recent reports,³–⁶ we have shown that selection of insufficiently accurate reference media can severely alter the information content of the data vector. Once corrupted, the recovery of images relatively free of artifact can be very difficult or impossible, even with use of full Newton updates. The second step scales the weight matrix by normalizing the column vectors to their respective mean values. This makes the weight matrix more uniform and better conditioned, and serves to suppress numerical errors and accelerate convergence.⁷,⁸ The third step is to impose constraints on iterative solutions of absorption and diffusion coefficients within the CGD method used for solving the resulting system equation. Ordinarily, this option is not applicable to the general case in which the direction of the perturbation is unknown. As is described below, we have found that by adopting a two–step process wherein solutions are obtained for both signs of the constraints and then summed, satisfactory solutions can be obtained. The details of this methodology are described subsequently.
2.1. Forward model

Light propagation in a scattering medium was modeled as a diffusion process. For a domain \( \Lambda \) having a boundary \( \partial \Lambda \), this is represented by the expression

\[
\nabla \left[ D(\mathbf{r}) \nabla I(\mathbf{r}) \right] - \mu_a(\mathbf{r}) I(\mathbf{r}) = -\delta (\mathbf{r} - \mathbf{r}_s), \quad \mathbf{r} \in \Lambda,
\]

where \( I(\mathbf{r}) \) is the photon intensity at position \( \mathbf{r} \), \( \mathbf{r}_s \) is the position of a DC point source, and \( D(\mathbf{r}) \) and \( \mu_a(\mathbf{r}) \) are the position-dependent diffusion and absorption coefficients, respectively. Here we define the diffusion coefficient as

\[
D(\mathbf{r}) = \frac{1}{3 \left[ \mu_a(\mathbf{r}) + \mu_s'(\mathbf{r}) \right]},
\]

where \( \mu_s'(\mathbf{r}) \) is the reduced scattering coefficient.

2.2. The inverse formulation

The optical inverse formulation was based on the normalized difference method \(^{3-6}\) and has the following form:

\[
W_r^{(\mu_a)} \cdot \delta \mu_a + W_r^{(D)} \cdot \delta D = \delta I_r,
\]

where \( \delta \mu_a \) and \( \delta D \) are the vectors of cross-sectional differences between the optical properties (absorption and diffusion coefficients, respectively) of a target (measured) and a “reference” medium (computed or measured) used to generate the initial guess; \( W_r^{(\mu_a)} \) and \( W_r^{(D)} \) are the weight matrices describing the influence that each element has on the surface detectors associated with the absorption and diffusion coefficients of a selected reference medium, respectively; and \( \delta I_r \) represents a normalized difference of detector readings between two sets of data, which is defined by the equation

\[
(\delta I_i)_i = \frac{(I_i - I_0)_i}{(I_0)_i}, \quad i = 1, 2, \ldots, M.
\]

Here, \( I_i \) is the set of computed detector readings corresponding to the selected reference medium, \( I_i \) and \( I_0 \) represent two sets of measured data (e.g., background vs. target, or time-averaged mean vs. a specific time point, etc.) and \( M \) is the number of source-detector pairs in each set of measurements.

2.3. Weight matrix scaling

We have previously described a scaling method that serves to improve the conditioning of the weight matrix. \(^7\) Here we extend this method to simultaneously recover absorption and diffusion coefficients. The effect of scaling the weight matrix is to make it more uniform, which can often improve conditioning of the matrix. A variety of scaling approaches could be adopted, but we have chosen one that scales each column of \( W_r^{(\mu_a)} \) and \( W_r^{(D)} \) to the average value of the column vector. The resulting new weight matrices have the form:

\[
\tilde{W}_r^{(\mu_a)} = W_r^{(\mu_a)} \cdot R_r^{(\mu_a)},
\]

where \( k = \mu_a \) or \( D \), and \( R_r^{(k)} \) is the normalizing matrix with the following entries:

\[
\left( R_r^{(k)} \right)_{ij} = \begin{cases} 
1 & j = i, \\
\frac{1}{M} \sum_{m=1}^{M} (W_r^{(k)})_{mj} & i, j = 1, 2, \ldots, N, \\
0 & j \neq i,
\end{cases}
\]

in which \( N \) is the number of elements used in discretizing the domain \( \Lambda \). The resulting system equation is

\[
\tilde{W}_r^{(\mu_a)} \cdot \delta \tilde{\mu}_a + \tilde{W}_r^{(D)} \cdot \delta \tilde{D} = \delta \tilde{I}_r,
\]

where \( \delta \tilde{\mu}_a = \left[ R_r^{(\mu_a)} \right]^{-1} \cdot \delta \mu_a \) and \( \delta \tilde{D} = \left[ R_r^{(D)} \right]^{-1} \cdot \delta D \).
2.4. Solutions of constrained CGD method

For most measurement geometries, since the weight matrix $\hat{W}$ is not symmetric positive definite, usually a least-squares solution of the system of linear equations shown in Eq. (6) is calculated by minimizing the mean-squared error $E$, which is represented as

$$E = \frac{1}{2} (\hat{W} \cdot \delta \bar{x} - \delta \bar{x}) \cdot (\hat{W} \cdot \delta \bar{x} - \delta \bar{x})$$

where $\delta \bar{x} = \left[ \delta \bar{p}_a \quad \delta \bar{D} \right]$, $\hat{W} = \left[ \hat{W}^{(\mu_a)} \quad \hat{W}^{(D)} \right]$, $\bar{A} = \hat{W}^T \cdot \hat{W}$, and $\bar{b} = \hat{W}^T \cdot \delta \bar{1}$. Formally, such a solution can be obtained by setting the derivative of $E$ to 0, i.e.,

$$\frac{\partial E}{\partial (\delta \bar{x})} = \bar{A} \cdot (\delta \bar{x} - \bar{b}) = 0.$$ (8)

When the CGD method is applied, instead of explicitly solving Eq. (8), we have the following iterative formulation for computing $\delta \bar{x}$:

$$\delta \bar{x}^{(n)} = \delta \bar{x}^{(n-1)} - \alpha^{(n)} \bar{d}^{(n)}.$$ (9)

The iterative procedure can be described as follows:

- Based on an initial estimate of $\delta \bar{x}^{(0)}$, compute $\bar{g}^{(0)} = \bar{A} \cdot \delta \bar{x}^{(0)} - \bar{b}$; $\beta^{(1)} = 0$; $\bar{d}^{(1)} = -\bar{g}^{(0)}$;
- For the $n^{th}$ iteration, compute

$$\alpha^{(n)} = \frac{\left\| \bar{g}^{(n-1)} \right\|^2}{\left\| \hat{W} \cdot \bar{d}^{(n)} \right\|^2},$$

$$\bar{d}^{(n)} = -\bar{g}^{(n-1)} + \beta^{(n)} \bar{d}^{(n-1)},$$

$$\beta^{(n)} = \frac{\left\| \bar{g}^{(n-1)} \right\|^2}{\left\| \bar{g}^{(n-2)} \right\|^2},$$

$$\bar{g}^{(n-1)} = \bar{A} \cdot \delta \bar{x}^{(n-1)} - \bar{b} = \bar{g}^{(n-2)} - \alpha^{(n-1)} \bar{A} \cdot \bar{d}^{(n-1)}.$$ (10)

In our study, positive or negative constraints are separately imposed on the reconstruction results, $\delta \bar{p}_a^{(n)}$ and $\delta \bar{D}^{(n)}$, after each iteration. In some instances we assume prior knowledge of the direction of the perturbation and apply the appropriate constraint. In the more general case, we recognize that more than one approach could be adopted to implement solution constraints. For instance, constraints could be applied in either a homogeneous or heterogeneous manner. In the case of a homogeneous constraint, either a positivity or a negativity constraint could be imposed on both coefficients. For the heterogeneous case, a positivity constraint could be imposed on one coefficient and a negativity constraint on the other. While we have not yet full explored the complete parameter space involving imposed constraints and the direction of the perturbations, it has been our experience thus far that the heterogeneous case provides the best solution. In addition, if one assumes no prior knowledge, then arbitrarily imposing a constraint could result in a grossly incorrect solution. We avoid this problem by adopting a two-step process, first imposing one set of constraints (positivity for one coefficient and negativity for the other), and second reversing the constraint directions, and summing the two solutions.

The CGD method iteratively updates the reconstruction on the basis of the previous reconstruction result $\delta \bar{x}^{(n-1)}$ and the entire set of preceding gradient vectors $\left[ \bar{g}^{(0)}, \bar{g}^{(1)}, \cdots, \bar{g}^{(n-1)} \right]$ and conjugate gradient vectors $\left[ \bar{d}^{(0)}, \bar{d}^{(1)}, \cdots, \bar{d}^{(n)} \right]$. Imposition of constraints on $\delta \bar{x}^{(n)}$ may lead to miscalculation of the gradient and conjugate gradient vectors. It also results in the loss of the
A-orthogonality properties, so that the reconstruction may not converge after a large number of iterations, and may even diverge in some cases. The criterion we use to detect divergence is that the ratio of two consecutive mean-squared errors, 

$$r = \frac{E_n}{E_{n-1}} = \frac{E(\delta \hat{x}^{(n)})}{E(\delta \hat{x}^{(n-1)})},$$

where $E(\delta \hat{x}^{(n)})$ was defined in Eq. (7), is greater than 1. Obviously, if the mean-squared error increases during any one iteration, it will continue to increase in all subsequent ones. Once divergence is detected, we reset the conjugate-gradient vector by taking $d^{(n)} = 0$ and restart the CGD reconstruction using $\delta \hat{x}^{(n-1)}$ as the initial estimate of $\delta \hat{x}$. By doing this, the constrained CGD reduces, in the worst-case limit, to a standard gradient descent algorithm.

After the intermediate solutions $\delta \mu_a$ and $\delta \hat{D}$ are obtained, they are rescaled to get the final results using the formula:

$$\delta \mu_a = R^{(a)} \cdot \delta \mu_a;$$

$$\delta \hat{D} = R^{(D)} \cdot \delta \hat{D}. \quad (11)$$

For all reconstruction results shown, solutions were limited to a first-order computation involving 1000 iterations. The finite element grid used comprised 1296 elements.

### 3. SIMULATION RESULTS

Figure 1 shows a sketch of the target medium considered for the simulation studies. The optical properties of the background medium are $\mu_a = 0.04$ cm$^{-1}$ and $\mu_s' = 10$ cm$^{-1}$. Included are up to four objects having the dimensions and locations shown. Each tomographic simulation considered 6 sources and 18 detectors, providing 108 source-detector pairs. Computed were the detector responses for the target medium and those corresponding to a selected reference medium having the same external geometry, size, and source-detector configuration. Simulated data from these calculations were subsequently analyzed using the methods described above.

Figure 2 shows the original and reconstructed profiles for the target medium considered. Columns (a) and (c) are the original profiles for $\delta \mu_a$ and $\delta \hat{D}$, respectively; columns (b) and (d) are the corresponding reconstructed profiles. In the first row are two objects: left, only $\mu_a$ is increased; right, only $D$ is increased. In the second row, we considered four objects for which only one coefficient was perturbed in any one object (left and right, $\mu_a$ was increased; top and bottom, $D$ was decreased). In the rows 3–7 we considered the more general case in which perturbations in both $\mu_a$ and $D$ can occur simultaneously for any object. In row three, $\mu_a$ and $D$ for both objects were decreased and increased, respectively by an amount indicated by the gray scale. In row four, three objects are present. The absorption coefficient is increased in the left and top inclusions, while the diffusion coefficient is decreased in the right and top inclusions. The geometric arrangements of the objects in rows five and six are the same as in row three and row four, respectively. These cases differ from the former ones in that $\mu_a$ was increased in row five (rather than decreased, as in row three) and $D$ was increased in row six (rather than decreased, as in row four). It is worth noting that the latter two cases are similar to the more difficult examples explored by Arridge and Lionheart, in which the influence of an increase or decrease in absorption can be offset by a decrease or increase in scattering (i.e., increase or decrease in $D$). These are conditions that produced evidence for solution nonuniqueness using intensity-only data. In row seven, $\mu_a$ and $D$ are decreased and increased respectively for the left-hand object, while the opposite trend occurs in the right-hand object.

Inspection of the reconstructed profiles shows that in each case we can effectively isolate perturbations in the absorption and diffusion coefficients, whether or not they are co-located. In the first six cases, this was accomplished by assuming prior knowledge of the direction of the perturbations and applying the appropriate constraints as described in Methods. In practice, this could correspond to situations where the influence of a particular manipulation of tissue (e.g., inflation of a pressure cuff)
would impose an expected response (e.g., venous congestion and hence an increase in hemoglobin absorption). The more general case, shown in row seven, however, is the one in which the perturbation in either coefficient could be positive or negative. To capture this information we applied a two-step process. First we imposed a positivity constraint on one coefficient and a negativity constraint on the other. Second, we reversed the direction of these constraints and sum the two solutions together. The resulting images clearly show that we can achieve parameter isolation with a high degree of fidelity and spatial accuracy. In fact, common throughout is the finding that the spatial localization of the reconstructions is in most cases excellent. This is even more notable when it is considered that the results are limited to a first–order solution.

4. EXPERIMENTAL RESULTS

To validate the normalized–constraint method, we performed a series of laboratory studies on phantom vessels containing one or more objects whose optical properties differ from a homogeneous background primarily in either its scattering or scattering and absorption coefficients. In addition to acting as optical perturbations, we also introduced dynamic behavior by moving the objects in a circular motion as indicated, while all the time acquiring fast tomographic data sets using the DC imager described in an accompanying report.9 Sketches of the target media explored are shown in Figure 3. In case 1, we introduced three white plastic rods (6 mm dia.) composed of white Delrin® into a cylindrical vessel having a diameter of 7.6 cm and filled with 1% (v/v) Intralipid®. It has been reported that the optical properties of white Delrin are $\mu_a = 0.02 \text{ cm}^{-1}$, $\mu'_s = 12 \text{ cm}^{-1}$,10 while those of 1% Intralipid are $\mu_a = 0.02 \text{ cm}^{-1}$, $\mu'_s = 10 \text{ cm}^{-1}$.11 Thus, it is only the scattering coefficient of the rods that is increased (decreased $D$) relative to the background. While we did not independently verify these values, we did observe that the light intensity was increased on the same side of the source, and decreased on the side opposite the source in the presence of the rods, a finding consistent with the reported coefficient values. In case two, a similar experiment was performed, except that a black glossy metal rod of comparable diameter was substituted for one of the white plastic rods. We treated this substitution as introducing a perturbation in both absorption and scatter. In cases three and four we performed similar studies, and increased the complexity of the medium by introducing a 3–mm–thick clear plastic layer, serving as an optical void. Theses cases were intended to determine whether our methods could correctly isolate optical perturbations (and dynamic behavior) under conditions wherein the assumptions underlying the diffusion equation are strongly violated. As a further confirmation, the above–described experiments were repeated in numerous live demonstrations using our DC imaging system being exhibited at this conference. Data acquisition involved 16 equally spaced source positions with 16 co–located detectors (256 source–detector pairs). The illumination wavelength was 785 nm, and full tomographic scans were acquired at 3 Hz for approximately 25 seconds (75 scans).
Figure 4 shows a typical reconstructed profile for absorption and diffusion coefficients, for case 1. Because of format limitations, movies of the reconstructed image time series cannot be presented in this report, and instead we show representative image frames. Inspection of the other image frames reveals that all are similar to those shown here. In every case, almost complete separation of absorption from scatter was obtained. Careful inspection the absorption map reveals that artifact is restricted to the region near the boundary, and the interior is almost completely featureless. The reconstructed diffusion map shows, with high contrast, the correct locations of the inclusions. A quantitative comparison reveals that the amplitude of computed absorption perturbation is quite small in absolute terms and also small compared to the magnitude of the computed perturbation in $D$. That is, there is minimal cross-talk. Another feature we have observed particularly in solutions obtained using Eq. (2) is that whereas the recovered $\delta \mu_a$ and $\delta D$ values scale with the selected reference medium, their ratio does not. Thus we find that the relative magnitude of cross-talk between the coefficients maps is essentially independent of the chosen properties of the reference medium.

Figure 5 shows that similar results are obtained for case 2. Again, because of format limitations, movies of the image maps are not available, and only four of the 75 computed image frames are shown. Results obtained in this case are qualitatively similar to these reported for row 4 in Figure 2. The presence of the plastic rods is revealed only in the reconstructed diffusion
maps, whereas images of the glossy metal rod appear in both reconstructed maps. Note, as with case 1, nearly complete isolation of absorption from diffusion coefficient is obtained for the included plastic rods with only a minimum distortion of object shape and location.

Data in Figures 6 and 7 show results obtained in the presence of a circular optical void for selected image frames. This geometry was considered as a crude representation of the layered structure that occurs in the head. Here, the void is meant to represent the clear fluid space occupied by cerebrospinal fluid in the subarachnoid space. We have considered this case because it has been reported by Dehghani et al.\textsuperscript{12} that reconstruction methods based on the diffusion approximation perform poorly under conditions similar to those examined here. In the case of a single inclusion (white plastic rod, Figure 6), we observe that the occurrence of an intervening optical void does not prevent new complete separation of the coefficient profiles. Data in Figure 7 shows that qualitatively similar results are obtained using three plastic rods in terms of minimizing

*Figure 5: The top and bottom rows are the reconstructed profiles of absorption and diffusion coefficients, respectively, from representative image frames for case 2. Arrows identify spatial coincidence of the reconstructed absorption and diffusion profiles for the glossy metal rod.*

*Figure 6: The top and bottom rows are the reconstructed profiles of absorption and diffusion coefficients, respectively, from representative image frames for case 3.*

*Figure 7: The top and bottom rows are the reconstructed profiles of absorption and diffusion coefficients, respectively, from representative image frames for case 4.*
cross-talk, though increased artifact levels are present. Note that the rods in this case are more centrally located than in case 1, and the positions of the diffusion-coefficient minima in the images correspond closely to the actual spacings of the rods. These findings show that, using the methods described here, we do not encounter the difficulties that Dehghani et al. have reported using diffusion-based imaging codes.

5. DISCUSSION AND CONCLUSIONS

A working hypothesis of our group is that measures of dynamic variations in the optical properties of tissue can significantly extend the utility of optical imaging studies for diagnostic and monitoring purposes. In a series of recent reports, we have shown by numerical simulation and experiment the capacity to identify and accurately characterize complex dynamic states arising from linear and nonlinear dynamic phenomena. Key to these studies has been the use of new a formulation to the inverse problem that relies on the analysis of normalized detector data. As described in Methods, we evaluate the quantity \( \hat{d}_i = [(I - I_0)/I_0]_i \) (i = 1,2, ...,M). Critical to this formulation is appreciation that normalized quantities can be obtained with much greater reliability than can absolute measures. Thus, for a typical case we consider the situation where our data vectors are expressed as perturbations about a temporal mean value. Significantly, we have shown that using this approach we can accurately define dynamic behavior without the need for instrument calibration. We have also observed that this scheme is mainly insensitive to image distortion caused by selection of an insufficiently accurate reference medium. In these previous studies, images were reconstructed without use of the matrix scaling method. Our experience with these methods has been, similar to that of other investigators, that the recovered images are subject to parameter cross-talk.

Matrix scaling methods are known to improve the conditioning of a matrix, thereby leading to more stable solutions. In the current report we have combined this capability with the use of range constraints in an effort to improve separation of the coefficient values. Overall, we have found that there are three steps essential to producing stable solutions with minimal cross-talk. These are (1) use of normalized data, (2) use of the matrix scaling method, and (3) application of appropriate constraints on solutions within the CGD method. In the third step, different approaches can be applied. Most successful are those that employ \textit{a priori} knowledge of the direction of the perturbation. We recognize, however, that in the more general case the directions of the perturbations are unknown, and can involve both coefficients at any location. In these cases we apply a two-step process. First, we impose a positivity constraint on one coefficient and a negativity constraint on the other. Second, we reverse the direction of these constraints and sum the two solutions together. The solutions obtained are frequently qualitatively accurate (i.e., the correct direction of the perturbation is identified, as is target location, and parameter cross-talk is minimized). These findings are contrary to the assertion that total intensity data alone (i.e., DC methods) are insufficient to distinguish effects of absorption and scatter.

A restriction we have imposed in this study is to limit inverse computations to first-order solutions. We have imposed this mainly in recognition of the practical limits that inversion of large number of data sets (e.g., an image time series) would impose should iterative recursive updates be attempted. It is our experience, however, that especially as it relates to defining and characterizing the dynamics of the optical properties of a target, we are nevertheless capable of achieving remarkably accurate estimates of these using the described methods, even when they are restricted to computing first-order solutions.

It is worth noting that results presented here represent only a small portion of a more comprehensive investigation of numerous other test cases, including situations involving complex spatiotemporal coincident variations in absorption and scattering coefficients. In all, we have found that the described procedure produces results that are stable to inaccurate estimates of the reference medium and can effectively isolate variations in the absorption and scattering (diffusion) coefficients, in least in the case of simply structured target media.

6. ACKNOWLEDGMENT

This research was support in part by NIH grant no. R01–CA66184.

7. REFERENCES