Optical tomographic image reconstruction with quasi-Newton methods

Alexander D. Klose and Andreas H. Hielscher
[1]*Depts. of Biomedical Engineering and Radiology, Columbia University*
1010 CEPSR Bldg., MC8904, 530 West 120th Street, New York, NY 10027
Phone: (212) 854-5080 Fax: (212) 854-8725 e-mail: ahh2004@columbia.edu

Abstract: Most of the currently existing image reconstruction algorithms for optical tomography (OT) can be formulated as an optimization problem. To find a minimum of an appropriately defined objective function, researchers in OT mostly rely on conjugate-gradient (CG) methods. In this work we have tested the performance of quasi-Newton methods, which prove to be superior to CG methods, both in terms of conversion time and image quality.

OCIS codes: (170.0170) Medical optics and biotechnology, (170.3010) Image reconstruction techniques

1. Introduction

Most of the currently available reconstruction algorithms for diffuse optical tomography (DOT) use some form of a model-based iterative image reconstruction (MOBIIR) scheme.[1,2,3,4,5,6,7] In these schemes the image reconstruction problem can formulated as a numerical optimization problem, which consists of three major parts. First, a forward model for light transport is used to predict the measured data, assuming a certain initial distribution of optical properties. Second, a scalar objective function, Φ, is evaluated to obtain a measure of difference between the predicted and measured data. In a third step the initial guess is updated in a way that reduces the difference between predicted and measured data as defined by the objective function. These steps are repeated until the objective function is minimized.

Besides the type of forward model being used, currently available algorithms for optical tomography mainly differ in what type of updating schemes are employed. In general the updating procedure can be written as [8,9]

$$\propto_{k+1} = \propto_k + \alpha_k p_k \,, \qquad (1)$$

where $\propto_k$ is a vector containing a set of optical properties from which the new set $\propto_{\square+1}$ is obtained. The vector p_k is a search direction in the N-dimensional space, given a problem with N unknowns. The parameter $\alpha_{\square}$ is the step length in the direction p_k. Reconstruction algorithm for OT, mainly differ in the way p_k and $\alpha_{\square}$ are determined. In the most general from the search direction is given by

$$p_k = A_k \nabla \Phi_k + \beta_k p_{k-1} \qquad (2)$$

For example, using the steepest gradient descent method, $\beta_k = 0$ and A_k equals the identity matrix I. Most commonly applied in OT are conjugate gradient (CG) techniques in which A_k= I and $\beta_k \neq 0$ is a scalar that ensures that p_k and p_{k-1} are *conjugate*. In this work we explore the performance of so-called quasi-Newton methods, in which $\beta_k = 0$ and A_k is chosen as an approximation of the inverse second derivative of the objective function $\nabla^2\Phi = H$ which is also commonly referred to as the Hessian matrix. In general, Quasi Newton methods are often found to be more reliable and faster converging and than CG methods but have not yet been applied to optimization problems encountered in optical tomography.

2. Numerical Background

Quasi Newton methods are derived from Newton's method, which is often used for finding zeros of a nonlinear function. When we consider the minimization of an objective function $\Phi(\propto)$ the tusk becomes to find a zero of the first derivative, $\partial\Phi/\partial\propto = \nabla\Phi = \Phi'$, of the objective function with respect to the optical parameters. Given an estimate $\propto_k$ of the solution, the nonlinear function $\Phi'(\propto_{\square})$ can be approximated by the linear function $q(\propto_{\square})$ which consisting of the first two terms of the Taylor series expansion of $\Phi'(\propto)$ at the point $\propto_k$.:

$$\Phi'(\propto_k + p) \propto q(\propto_k + p) = \Phi'(\propto_k) + \Phi''(\propto_k)p \qquad (3)$$

For $q(\propto_k + p) = 0$ we can solve the Newton equation for p:

$$\Phi''(\propto_k)p = -\Phi'(\propto_k)\,. \qquad (4)$$

The Hessian matrix Φ" consists of second derivatives of the objective function. The parameter p is the search direction to be used in Eq. 1. After determining p from Eq. 4, an updated set of optical properties $\propto_{k+1}$ can be found with Eq. 1 (in general α_k =1 when Newton method is used). Newton's method is rarely used in its "classical" form for nonlinear programming problems, because it is often difficult to obtain the Hessian matrix Φ" for a given problem. Even if Φ" can be calculated, the storage of this matrix requires large amounts of memory. To overcome these problems so-called quasi-Newton methods (QN) are employed.

In the QN approach the Hessian matrix Φ" is replaced by a matrix B_k that approximate Φ" and can be obtain at a lower computational cost. Employing the secant condition the quasi-Newton approximation can be defined as $B_{k+1} s_k = y_k$, with the vectors $s_k = \propto_{k+1} - \propto_k$ and $y_k = \Phi'(x_{k+1}) - \Phi'(x_k)$. Performing a line search according to Eq. 1 the vector s_k is $s_k = \alpha_k p_k$. A new update of the matrix B_{k+1}, having the matrix B_k and the vectors s_k and y_k from the former iteration, is for example given by the Broyden-Fletcher-Goldfarb-Shanno formula (BFGS):

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \qquad (5)$$

Given an approximation to the inverse $A_k = B_k^{-1}$ it is even easier to compute the search direction according to Eq. 2 without having to solve a system of linear equations $B_k(x)p = -\Phi'(\propto_k)$ at all. An iterative way to calculate A is for example given by [7,8]

$$A_{k+1} = A_k + \frac{(y_k - A_k s_k) y_k^T}{y_k^T s_k} - \frac{(y_k - A_k s_k)^T s_k}{(y_k^T s_k)^2} y_k y_k^T \qquad (6)$$

This BFGS formula together with Eq. 1 was implemented for this study. Limted Memory BFGS (lm-BFGS) methods further simply this equation to reduce the memory requirements even more. In this case A_k in Eq. 7 is replaced with the identity matrix.

3. Results

We compared the performance of quasi-Newton methods based on the BFGS formula (Eq. 6) and the lm-BFGS formula (Eq. 6 A_k = I) with to the already widely used conjugate-gradient method. An example of a scattering medium is shown in Fig. 1a. The cross-sectional image to be reconstructed consisted of three objects with scattering coefficients of $\propto_{\square}$ = 2.9 cm^{-1}, $\propto_{\square}$ = 8.7 cm^{-1} and $\propto_{\square}$ = 11.6 cm^{-1}, which were embedded in a 3 cm x 3 cm background medium with $\propto_{\square}$ = 5.8 cm^{-1}. The absorption coefficient $\propto_a$ = 0.35 cm^{-1} did not vary within the isotropically scattering medium (g=0). The medium was surrounded by 8 equally spaced sources and 116 equally spaced detectors. 2 sources and 29 detectors were placed on each side of the medium. Detector readings on the same side as the source, were not used for the reconstruction. The measured data at the detector positions was generated by using the correct spatial distribution of optical properties (Fig. 1). The forward calculations were performed on a 61 x 61 grid with 16 ordinates by using the finite-difference discrete-ordinates method based on the equation of radiative transfer. [10] We used partly reflective boundary conditions based on Fresnel's law with a homogeneous refractive index n=1.54 for the scattering medium. The objective function was defined as the χ^2 error norm, which was iteratively minimized by the optimization scheme using the CG,[11] limited-memory BFGS and full BFGS method. The reconstruction was started with a homogeneous medium $\propto_a$ = .035 cm^{-1} and $\propto_s$ = 5.8 cm^{-1} as an initial guess and the iterative update process was terminated when $|(\Phi_{k+1} - \Phi_k)/\Phi_k| < 10^{-3}$. The results are displayed in Figs. 1b and 2.

Fig. 1b shows the valued of the objective function as a function of number of total number of gradient an d forward calculation. The lm-BFGS method reached the stopping criterion after 92 basic operations, which consisted of 54 forward and 38 gradient calculations (54+38). The BFGS method converged after 172 basic operations (98+74). The CG method took the longest time to reach the stopping criterion, requiring 395 basic operations (363+32). The absolute number of forward runs and gradient calculations differed, depending on the method used. Usually the CG method made more forward runs and fewer gradient calculations than the QN methods. The final value of the objective function was different for all three methods. The BFGS method had the smallest value with $\log_{10}$ = -5.32, whereas the CG method had the largest value with $\log_{10}$ = -4.48. The corresponding reconstructed images are shown in Fig. 2.

4. Summary

The image reconstruction problem in optical tomography can be considered as a nonlinear optimization problem. The goal is to find the minimum of an objective function that evaluates the difference between measured and predicted data. The majority of currently existing algorithm for OT use only first derivative information of the objective function with respect to the optical properties to find the minimum. In this work we presented a quasi Newton technique that employs approximations to the Hessian matrix, which contains second derivative information. Our results show that using the quasi-Newton method considerably reduces convergence time and improves image quality. The disadvantage of the quasi-Newton method is a larger memory requirement as compared to conjugated gradient techniques.

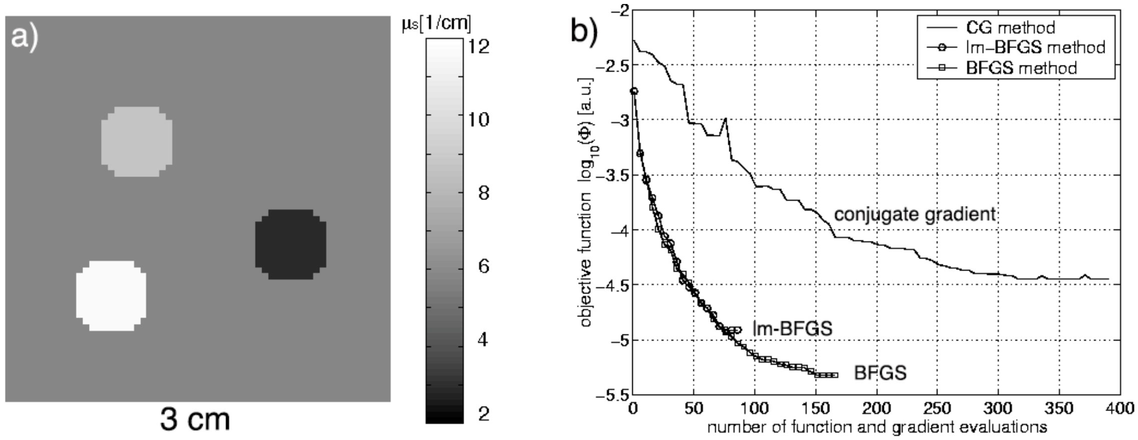

Fig. 1 (a-) Original cross-section of the scattering coefficient μ_s. (b) Value of the objective function (Eq. 7) as function of total number of forward and gradient calculations.

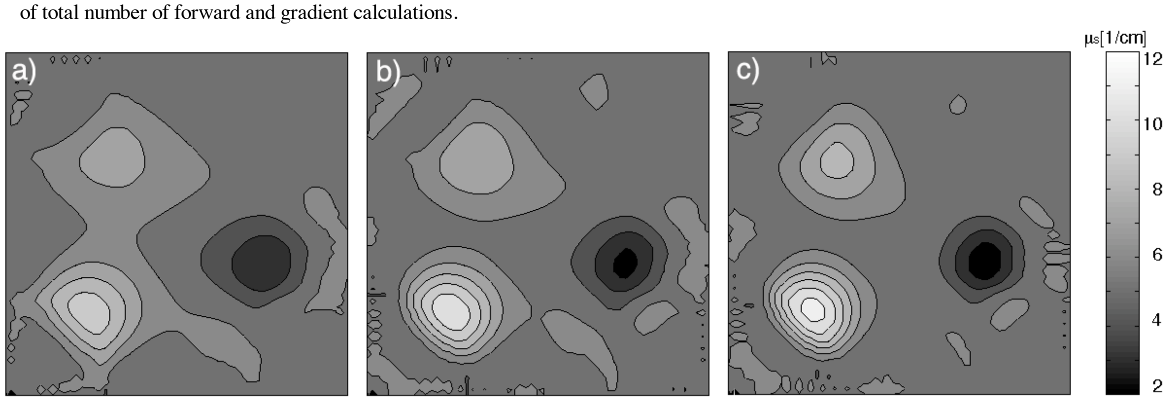

Fig. 2. (a) Reconstructed image using the CG-method. (b) Reconstructed image using the BFGS-method. (c) Reconstructed image using the BFGS-method.

5. Refrences

[1] S.R. Arridge, "Optical tomography in medical imaging," Inverse Problems 15, pp. R41-R93 (1999).

[2] A.H. Hielscher, A.D. Klose, K.M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography," *IEEE Transactions on Medical Imaging* **18**, 262-271 (1999).

[3] R. Roy and E.M. Sevick-Muraca,"Truncated Newton's optimization scheme for absorption and fluorescence optical tomography: Part I Theory and formulation, *Optics Express* **4** (10), 353-371 (1999).

[4] K.D. Paulsen and H. Jiang, "Spatially varying optical property reconstruction using a finite element diffusion equation approximation," *Med. Phys.* , vol. 22, pp. 691-701, 1995.

[5] H. B. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency domain data: Simulations and experiments," *Journal of the Optical Society of America A* , vol. 13, pp. 253-266, 1996.

[6] R.L. Barbour, H.L. Graber, J.W. Chang, S.L. S. Barbour, P.C. Koo, R. Aronson, "MRI-guided optical tomography: Prospects and computation for a new imaging method," *IEEE Computational Science & Engineering* , vol. 2, pp. 63-77, 1995.

[7] A.D. Klose and A.H. Hielscher, "Iterative reconstruction scheme for optical tomography based on the equation of radiative transfer," *Medical Physics* **26** (8), 1698-1707 (1999).

[8] J. Nocedal, S.J. Wright, "Numerical Optimization," Springer, New York, 1999.

[9] Stephen G. Nash, *Linear and Nonlinear Programming*, McGraw-Hill.

[10] A.D. Klose, U. Netz, J. Beuthan, A.H. Hielscher, "Optical tomography using the time-independent equation of radiative transfer. Part 1: Forward model," *Journal of Quantitative Spectroscopy and Radiative Transfer*, in press.

[11] A.D. Klose, A.H. Hielscher, "Optical tomography using the time-independent equation of radiative transfer. Part 2: Inverse model," *Journal of Quantitative Spectroscopy and Radiative Transfer*, in press.

Acknowledgment: This work was supported in part by The Whitaker Foundation (grant # 98-0244) and the National Institute of Arthritis and Musculoskeletal and Skin Diseases, a part of the National Institute of Health (R01 AR46255-02).