Regularized progressive expansion algorithm for recovery of scattering media from time-resolved data

Jenghwa Chang

Department of Pathology, State University of New York Health Science Center, Brooklyn, New York 11203

Wenwu Zhu and Yao Wang

Department of Electrical Engineering, Polytechnic University, Brooklyn, New York 11201

Harry L. Graber

Department of Physiology and Biophysics, State University of New York Health Science Center, Brooklyn, New York 11203

Randall L. Barbour

Department of Pathology and Department of Physiology and Biophysics, State University of New York Health Science Center, Brooklyn, New York 11203

Received April 3, 1996; revised manuscript received July 2, 1996; accepted July 26, 1996

Reconstructions of the absorption cross sections of dense scattering media from time-resolved data are presented. A progressive expansion (PE) algorithm, similar to a layer-stripping approach, is developed to circumvent the underdeterminedness of the inverse problem. An overlapping scheme, which uses detector readings from several consecutive time intervals, is introduced to reduce the propagation of reconstruction errors that occur at shallower depths. To reduce the sensitivity of the PE algorithm to noise, a regularized progressive expansion (RPE) algorithm is proposed, which incorporates regularization techniques into the PE algorithm. The PE and the RPE algorithms are applied to the problem of image reconstruction from time-resolved data. The test media were isotropically scattering slabs containing one or two compact absorbers at different depths below the surface. The data were corrupted by additive white Gaussian noise with various strengths. The reconstruction results show that the PE and the RPE algorithms, when they are combined by proper overlapping, can effectively overcome the underdeterminedness of the inverse problem. The RPE algorithm yields reconstructions that are more accurate and more stable under the same noise level. © 1997 Optical Society of America. [S0740-3232(97)01101-0]

Key words: image reconstruction, layer stripping, regularization, optical tomography, conjugate gradient descent.

1. INTRODUCTION

Great interest has recently developed in recovering anatomical or functional images from tomographic measurements at near-infrared frequencies because the measurement techniques are safe (nonionizing), potentially inexpensive, and portable. Photons of near-infrared light are intensely scattered by biological tissues. Several measurement schemes, involving continuous-wave, time-resolved (TR), or time-harmonic sources,[1–3] have been proposed as potentially suitable for interrogation of highly scattering media. In this paper we consider the recovery of absorption cross sections of scattering media that have optical properties similar to those of tissue from TR near-infrared optical measurements.

For imaging methods such as x-ray computed tomography, in which the path of the detected signal is a straight line, the inverse problem can be accurately formulated as a system of linear equations of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{y}$ is the measured data, $\mathbf{A}$ is an imaging operator, and $\mathbf{x}$ is the unknown.[4] In the case of strong scattering the imaging problem becomes difficult because the photons propagate through the tissue in a highly diffused manner and because the relation between the measured signal and the properties of the media is nonlinear. There is generally no direct method for solving the inverse problem. One way to attack this difficulty is to use the perturbation approach,[5–9] which simplifies the inverse problem to a system of linear equations. The derived image operator, or the weight matrix, is generally underdetermined and ill conditioned. The underdeterminedness results from the fact that the number of detector readings M is less than the number of unknowns N. Even if $M \geqslant N$, the weight matrix may well be rank deficient, which also leads to an underdetermined system. The cause of the ill conditioning is that the weight matrix contains many columns that are nearly proportional. Small variations in detector readings can result in large deviations in reconstruction results.

0740-3232/97/010306-07$10.00

To overcome the underdeterminedness problem Chang *et al.*[10] developed a layer-stripping algorithm that progressively evaluates the data in a way that decreases the depth component of the weight gradients. This algorithm uses the detector readings in different portions of the temporal profile separately and in a progressive manner. In each time interval it considers only the region that may contribute to the detector readings and that is not yet fixed on the basis of the preceding stages of the reconstruction. The contribution to the detector reading in this time interval from the volume elements (voxels) solved previously is first subtracted. The new unknowns are then solved by means of a perturbation equation. For further improvement of the stability of the progressive expansion (PE) algorithm, we have now incorporated Tikhonov–Miller regularization [11,12] into the solution of the subsystem in each time interval. This regularization method can deal with a large-scale system because it is based on an optimization technique that can be solved iteratively. In contrast, regularization based on singular-value decomposition (SVD), in which a small constant is added to all the singular values,[13] is not suitable for large-scale problems because of the rate at which the computation time that it requires grows with increasing problem size[14] (see Section 5 below).

The arrangement of this paper is as follows. In Section 2 we describe the mathematical derivation of the perturbation model and of the PE and the regularized progressive expansion (RPE) algorithms. The experiments are described in Section 3, followed by the reconstruction results in Section 4. Finally, we discuss and conclude this study in Section 5.

2. PROGRESSIVE EXPANSION AND REGULARIZED PROGRESSIVE EXPANSION ALGORITHMS

A. Perturbation Equation

In the past few years iterative perturbation approaches have been developed for solving the inverse problems associated with propagation of light from continuous-wave,[5,6] time-harmonic,[6] and TR[10,15] sources through tissue. These approaches require the solution of a linear perturbation equation at each iteration:

$$\mathbf{W}\Delta\mathbf{x} = \Delta\mathbf{I}, \tag{1}$$

where $\Delta\mathbf{x}$ is a vector of absorption cross-section differences between a reference and a test medium, $\Delta\mathbf{I}$ is a vector of changes in detector readings between the two media, and $\mathbf{W}$ is a weight matrix describing the influence of each voxel on the detector readings, which are essentially the first partial derivatives of the detector readings with respect to the absorption coefficients in the reference medium.

B. Progressive Expansion Algorithm

The PE algorithm evaluates increasing depths within the medium by successively considering signals entering the detector at increasing times following an incident pulse. Instead of solving Eq. (1) directly, the algorithm solves a subsystem represented by

$$\mathbf{W}_l\Delta\mathbf{x}_l = \Delta\mathbf{I}_l \tag{2}$$

at each time interval l. Here $\Delta\mathbf{x}_l$ is a subset of $\Delta\mathbf{x}$ consisting of those elements that correspond to voxels that can contribute signals in this time interval but not earlier, $\Delta\mathbf{I}_l$ is the vector of residual detector readings after the contributions from voxels whose absorption cross sections were fixed during earlier stages of the reconstruction have been subtracted, and $\mathbf{W}_l$ consists of the weights for only the elements of $\Delta\mathbf{x}_l$. As long as the early-arriving signals can be reliably detected, the subsystem for each time interval will be determined or overdetermined. In this case ${\mathbf{W}_l}^T\mathbf{W}_l$ is invertible, and there exists a unique least-squares (LS) solution for each subsystem,

$$\Delta\mathbf{x}_l = ({\mathbf{W}_l}^T\mathbf{W}_l)^{-1}{\mathbf{W}_l}^T\Delta\mathbf{I}_l, \tag{3}$$

which minimizes the error

$$E(\Delta\mathbf{x}_l) = \|\mathbf{W}_l\Delta\mathbf{x}_l - \Delta\mathbf{I}_l\|^2. \tag{4}$$

More specifically, for a time interval t_l the algorithm is as follows:

Step 1. Find all the detectors that may receive signals during t_l. We accomplish this by examining the weights of all the voxels for each detector. Only the detectors that have at least one nonzero weight associated with them are considered. Here the number of detectors found is denoted m_l. The set of detector indices is denoted D_l.

Step 2. Find all the nonfixed voxels that may contribute signals to any of the detectors determined in step 1. We accomplish this by including all the voxels that have nonzero weights for at least one of the selected detectors. Here the number of voxels found is denoted n_l; the set of voxel indices is denoted V_l.

Step 3. From the detector readings and voxels found in steps 1 and 2, form the following set of linear equations:

$$\sum_{i\in V_l} W_{ijk}\Delta x_i = \Delta I_{ijk} - \sum_{i\in V_1,\ldots,V_{l-1}} W_{ijk}\Delta x_i,$$

$$j,k \in D_l \quad \text{or} \quad \mathbf{W}_l\Delta\mathbf{x}_l = \Delta\mathbf{I}_l, \tag{5}$$

where $\Delta\mathbf{I}_l$ consists of the detector readings minus the contribution from the voxels that were fixed in previous reconstruction stages, $\Delta\mathbf{x}_l$ consists of the absorption cross-section perturbations in the voxels to be solved, and $\mathbf{W}_l$ is an $m_l \times n_l$ matrix composed of the weights of voxel i for detectors jk in time interval l, $i \in V_l$, $jk \in D_l$.

Step 4. Find the LS solution, as given by Eq. (3), for the n_l voxels in V_l. In our study the conjugate gradient descent (CGD) method [16,17] is used to minimize the error given by Eq. (4). The principal reason that this algorithm was selected is its relatively fast convergence.

Step 5. Apply a positivity constraint to the reconstructed value. We accomplish this by setting $\Delta x_i = 0$ if $\Delta x_i < 0$. This is appropriate for the test case that we considered for this paper, wherein every element of $\Delta\mathbf{x}$ was known *a priori* to be nonnegative. In general, a less restrictive range constraint would be employed instead.

Step 6. Fix Δx_i for all $i \in V_l$. Go to the next time interval and repeat steps 1–5.

The PE algorithm is highly effective in circumventing the underdetermined nature of the inverse problem. However, because of the ill conditioning of $\mathbf{W}_l$, the LS solution in each time interval is sensitive to noise. The errors in shallower regions that are solved by use of data from the early portion of the TR detector readings can also propagate into deeper regions. This error propagation effect can become especially severe when the algorithm probes deeply beneath the surface. To overcome this problem an overlapping scheme has been developed. In this method a voxel's absorption cross-section perturbation is not fixed the first time that it is solved for. Rather, it is considered in several subsequent time intervals. In each time interval, new voxels that first contribute signal in this time interval as well as a certain number of old voxels (those that were also considered in the previous time interval) are considered. After the perturbation equation that includes all these voxels is solved, a partial set of the old voxels is fixed, and the next time interval is considered. However, overlapping will increase the number of unknowns to be solved in each time interval, and consequently the computation time will increase. Too much overlapping can even make the problem underdetermined in some early time intervals, which will make the reconstructions computed in the subsequent intervals less reliable. Therefore we need to select the overlap interval properly to reach a good compromise. In our experiments we chose an overlap of three time intervals. For convenience we omit the subscript l in the following discussion.

C. Regularized Progressive Expansion Algorithm

Regularization is a well-established technique for dealing with instability in inverse problems and can convert an ill-posed problem into a well-posed problem by incorporating *a priori* knowledge about the image to be recovered. With the Tikhonov–Miller regularization approach the idea is to choose an approximate solution from a set of admissible solutions according to some defined criterion. The class of feasible solutions is defined as the set $S_{\Delta x/\Delta I}(\Delta\mathbf{x}) = \{\Delta\mathbf{x}:\|\mathbf{W}\Delta\mathbf{x} - \Delta\mathbf{I}\|^2 \leqslant \epsilon^2\}$. The bound ϵ^2 depends on the noise level of the observed data. Tikhonov defined the regularized solution to be the one that minimizes a stable functional $\|\mathbf{C}\Delta\mathbf{x}\|^2$ subject to the constraint that $\|\mathbf{W}\Delta\mathbf{x} - \Delta\mathbf{I}\|^2 = \epsilon^2$. Here, $\mathbf{C}$ is a regularization operator and can be selected according to *a priori* knowledge. If the solution is known to be bounded but fluctuating, we can take $\mathbf{C}$ to be an identity matrix. If the solution is known to be continuous, then we can take the first-order differential operator as $\mathbf{C}$. If we know that the solution is smooth, then $\mathbf{C}$ can be chosen to be a second-order differential operator. Then, by use of the Lagrange multipliers method, the problem is to minimize

$$E(\Delta\mathbf{x}) = \|\mathbf{W}\Delta\mathbf{x} - \Delta\mathbf{I}\|^2 + \lambda\|\mathbf{C}\Delta\mathbf{x}\|^2, \tag{6}$$

where the regularization parameter λ can be found from the previously described constraint. The regularized solution is given by

$$\Delta\mathbf{x} = (\mathbf{W}^T\mathbf{W} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\mathbf{W}^T\Delta\mathbf{I}. \tag{7}$$

Miller took a set-theoretic approach and constrained the solution on both $S_{\Delta x/\Delta I}$ and the set $S_{\Delta x}(\Delta\mathbf{x}) = \{\Delta\mathbf{x}:\|\mathbf{C}\Delta\mathbf{x}\|^2 \leqslant E^2\}$, where E is a constant. The two constraints can be combined into a quadratic formula. The solution is identical to Tikhonov regularization with $\lambda = \epsilon^2/E^2$, the so-called Miller criterion. Numerically, we can see that the solution of Eq. (6) is more stable than that of Eq. (3) because $\mathbf{W}^T\mathbf{W} + \lambda\mathbf{C}^T\mathbf{C}$ is better conditioned than $\mathbf{W}^T\mathbf{W}$.

Inasmuch as the dimension of $\mathbf{W}$ can be extremely large, we do not perform matrix inversion directly to obtain the solutions in Eqs. (3) and (6). Rather, we minimize the functional in Eq. (6), using the CGD method. The nth iteration of the CGD algorithm can be described by

$$\Delta\mathbf{x}^{(n+1)} = \Delta\mathbf{x}^{(n)} + \alpha^{(n)}\mathbf{d}^{(n)},$$

where

$$\mathbf{d}^{(n)} = \mathbf{g}^{(n)} + (\|\mathbf{g}^{(n)}\|^2/\|\mathbf{g}^{(n-1)}\|^2)\mathbf{d}^{(n-1)},$$

$$\alpha^{(n)} = \langle\mathbf{d}^{(n)},\mathbf{g}^{(n)}\rangle/(\|\mathbf{W}\mathbf{d}^{(n)}\|^2 + \lambda\|\mathbf{C}\mathbf{d}^{(n)}\|^2),$$

$$\mathbf{g}^{(n)} = \mathbf{W}^T(\mathbf{W}\Delta\mathbf{x}^{(n)} - \Delta\mathbf{I}) + \lambda\mathbf{C}^T\mathbf{C}\Delta\mathbf{x}^{(n)},$$

and $\langle\mathbf{u},\mathbf{v}\rangle$ denotes the inner product of vectors $\mathbf{u}$ and $\mathbf{v}$. In our current implementation we choose the regularization operator $\mathbf{C}$ to be an identity matrix to make use of the bounded nature of the solution. We further assume that the upper bounds on $\|\mathbf{W}\Delta\mathbf{x}^{(n)} - \Delta\mathbf{I}\|^2$ and $\|\mathbf{C}\Delta\mathbf{x}\|^2$ can be estimated from measurement data or from prior knowledge, and we determine the regularization parameter λ by the Miller criterion.

3. METHODS

The source configuration and detector distribution used in our tests are shown in Fig. 1. The medium without absorption perturbations was used as the reference. We calculated the weights for this medium by using Monte Carlo simulations.[5] Values for $\Delta\mathbf{I}$ were first calculated from Eq. (1). We then added Gaussian noise to evaluate the algorithm's sensitivity to noise. The ratio of the noise's standard deviation to the mean value of the difference in detector readings is used as the measure of the noise level. In the first experiment, two closely juxtaposed absorbers of size 1 cubic mean free path (mfp; 1 mfp is the average distance that a photon propagates between successive interactions with the medium), separated by 1 mfp, were buried between 4 and 5 mfp in a 10-mfp-thick slab.[5] In the second experiment, one single absorber of size 8 cubic mfp was buried between 10 and 12 mfp in a 20-mfp-thick slab. In the third experiment, two absorbers, one directly above the other, were buried at depths of 2–4 and 6–8 mfp in a 20-mfp-thick slab. In the fourth experiment, one absorber of size 8 cubic mfp was buried at a depth of 6–8 mfp in the center, and three contiguous absorbers of the same size were buried at a depth of 10–12 mfp in a 20-mfp-thick slab. In all the experiments 10% noise was added in the detector readings. In each reconstruction the regularization parameter λ was chosen according to the Miller criterion, because we usually know the bound of the reconstructed image and noise variance in the detector readings. The regularization operator $\mathbf{C}$ was chosen to be an identity matrix. An overlapping interval of three mean free times (mft; 1 mft = 1

mfp/c, where c is the speed of light in the medium) is used in both PE and RPE algorithms.

4. RESULTS

The reconstructed results for the targets are illustrated in Figs. 2–5. Figures 2(a), 3(a), and 5(a) are the original images; Figs. 2(b), 3(b), and 5(b) are the images reconstructed by the PE algorithm after 17, 26, and 24 mft, respectively; Figs. 2(c), 3(c), and 5(c) are the images reconstructed with the RPE algorithm at the same times as in Figs. 2(b), 3(b), and 5(b). To explain the stability of RPE and to show how effective it is at suppressing noise, in Fig. 4 we show the reconstructed images at different time windows. Figure 4(a) is the original image, Figs. 4(b) and 4(c) are the images reconstructed by PE at times 22 and 28 mft, respectively; Figs. 4(d) and 4(e) are the images reconstructed by RPE at times 22 and 28 mft, respectively.

In all the figures the left-hand column is the $X-Z$ cross section of the medium and the right-hand column is the $Y-Z$ cross section. For display purposes the reconstructed values are quantized into 10 levels. Further, the images have been scaled individually, so the maximum intensity value in each image is represented by the same darkness. Therefore the same gray level in different figures may represent different absorption levels, especially in the reconstructed images without regularization, in which the maximum values are usually much greater than those in the reconstructed images with regularization. In fact, the maximum value obtained without regularization usually reaches the preset upper bound, greatly exceeding the real value in the test medium.

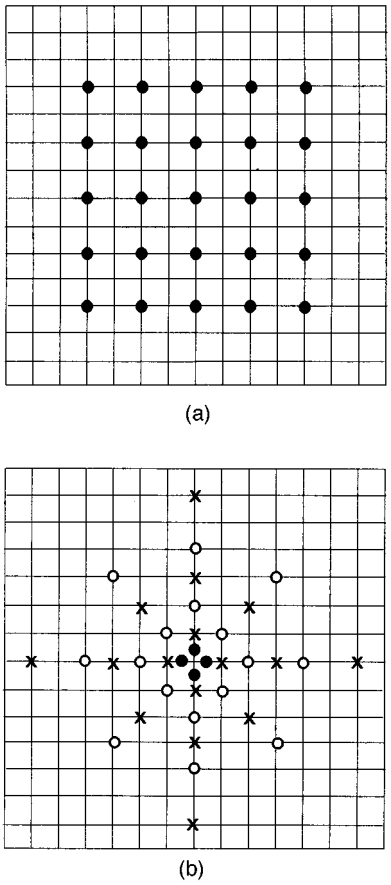

Fig. 1. (a) Source locations for the three-layer simulation; (b) position and orientation of detectors about each source: the source is placed in the center and ○'s and ×'s indicate positions of detectors that are inclined 10° and 80° from the normal, respectively. ●'s, positions where measurements are made in both orientations. The grid size is 3 mfp × 3 mfp.

5. DISCUSSION AND CONCLUSIONS

In this study Tikhonov–Miller regularization has been incorporated into the PE algorithm previously proposed for image reconstruction from TR data. From the results that we obtained we can see that the PE algorithm converged to a solution containing numerical artifacts comparable in magnitude with the target. The RPE algorithm, in contrast, yielded a solution much more closely resembling the original medium. In fact, the PE algorithm tends to diverge after a certain time interval because of error propagation in the presence of noise, as shown in Fig. 4(c), whereas the RPE algorithm can successfully suppress this effect for noise levels up to 10%. The computation time of the RPE algorithm is increased by ~10% over that of the PE algorithm. Thus the regu-

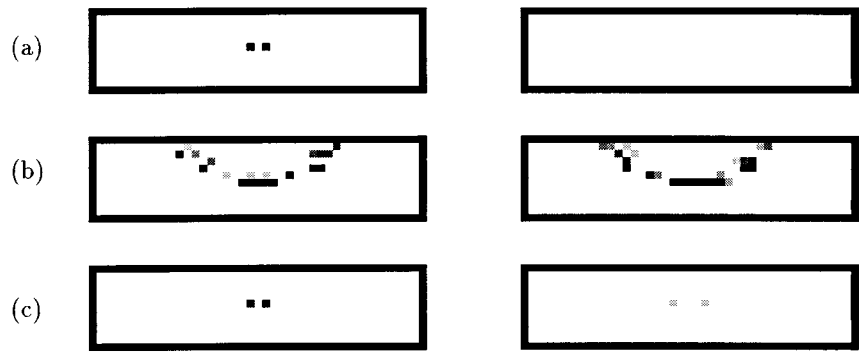

Fig. 2. Reconstruction of a 10-mfp-thick slab medium containing two 1-cubic-mfp absorbers separated by 1 mfp at a depth of 4–5 mfp. Left, $X-Z$ cross sections; right, $Y-Z$ cross sections. (a) Original medium, (b) reconstruction by PE with an overlapping interval of 3 mft, (c) reconstruction by RPE with the same overlapping interval. The noise level was 10%. All the reconstructed images are obtained after 17 time windows of width 1 mft. The maximum-values levels in (a), (b), and (c) correspond to 0.01, 0.1, and 0.0091, respectively.

larization technique is highly effective in stabilizing the PE algorithm and suppressing error propagation.

A critical problem in the use of regularization is the selection of the parameter λ when the upper bounds on $\|\mathbf{W}\Delta\mathbf{x} - \Delta\mathbf{I}\|^2$ or $\|\mathbf{C}\Delta\mathbf{x}\|^2$ or both are unknown. If only one of the bounds is known, a constrained LS approach can be followed.[18] When both bounds are unavailable, the cross-validation method[19] can be used. One disadvantage of these two approaches is that they require an overwhelming amount of computation when the weight matrix is large. When incorporating regularization in our PE algorithm, we would have to repeat this computation for each time window. Recently an algorithm was developed by Kang and Katsaggelos[20] that obtains the regularized LS solution iteratively by a gradient descent method and at each iteration calculates an updated regularization parameter λ based on the previous solution. The algorithm is powerful in that λ on convergence does not depend on the initial estimate of $\Delta\mathbf{x}$. Because λ does not need to be determined in a separate initial step, additional computational overhead is minimal.

Regarding issues of computational overhead, one may ask what advantages an iterative algorithm for computing the LS solution, e.g., the CGD method used in this study, has over the SVD method. When choosing a linear-system solver, one should consider three major factors: (1) the number of operations, (2) storage requirements, and (3) the degree of useful information obtained during intermediate stages of the reconstruction. In the ideal case when round-off error is absent, the CGD method converges in N iterations for a fully determined $N \times N$ linear system. Because there are two vector–

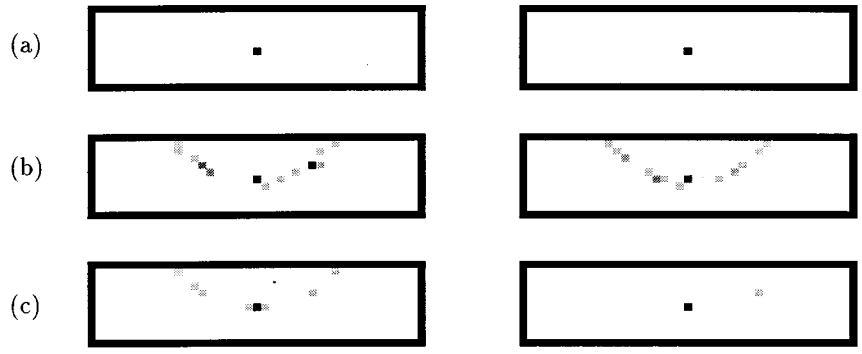

Fig. 3. Reconstruction of a 20-mfp-thick slab medium containing one absorber of size 8 cubic mfp at a depth of 10–12 mfp. Left, X–Z cross sections; right, Y–Z cross sections. (a) Original medium, (b) reconstruction by PE with an overlapping interval of 3 mft, and (c) reconstruction by RPE with the same overlapping interval. The noise level was 10%. All the reconstructed images are obtained after 26 time windows of width 2 mft. The maximum-values levels in (a), (b) and (c) correspond to 0.01, 0.1, and 0.01, respectively.

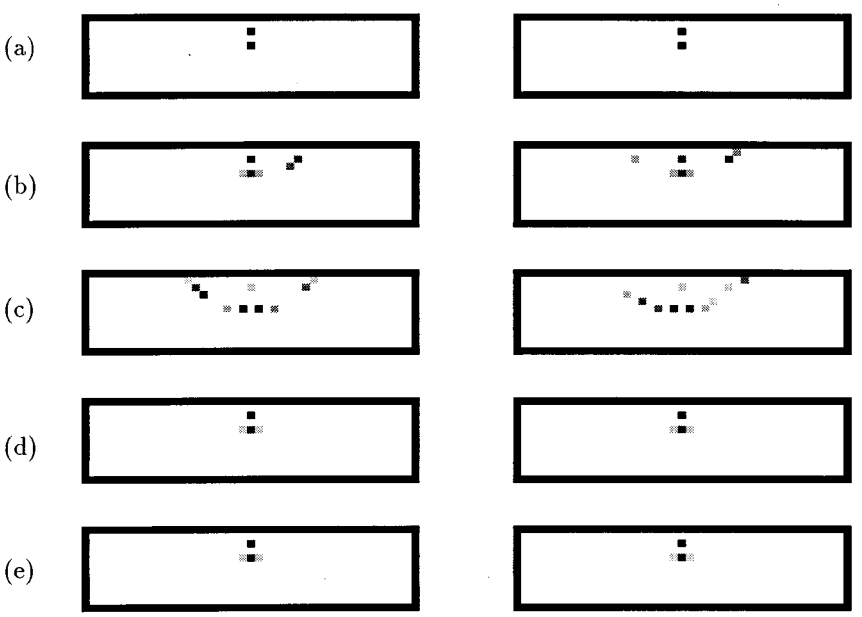

Fig. 4. Reconstruction of a 20-mfp-thick slab medium containing two 8-cubic-mfp absorbers, one above the other, at depths of 2–4 mfp and 6–8 mfp. (a) Original medium; (b), (c) reconstruction by PE with an overlapping interval of 3 mft after 22 and 28 time windows of width 2 mft, respectively; (d), (e) reconstruction by RPE with the same overlapping interval after 22 and 28 time windows of width 2 mft, respectively. The noise level was 10%. The maximum-values levels in (a), (b), (c), (d), and (e) correspond to 0.01, 0.01, 0.1, 0.01, and 0.01, respectively.

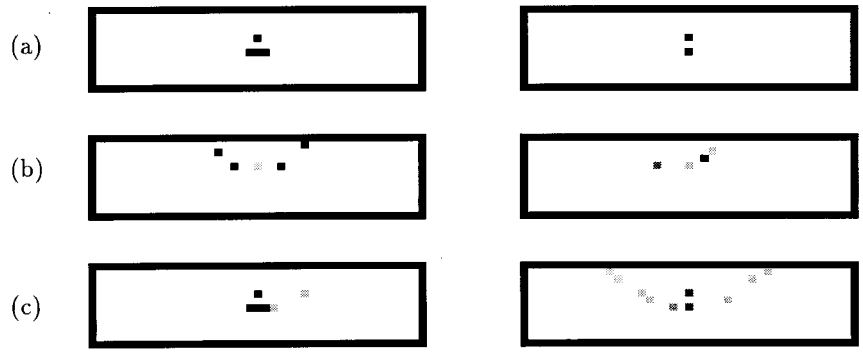

Fig. 5. Reconstruction of a 20-mfp-thick slab medium containing one absorber of size 8 cubic mfp at a depth of 6–8 mfp and three contiguous absorbers of the same size at a depth of 10–12 mfp. (a) Original medium, (b) reconstruction by PE with an overlapping interval of 3 mft, (c) reconstruction by RPE with the same overlapping interval. The noise level was 10%. All the reconstructed images are obtained after 24 time windows of width 2 mft. The maximum-values levels in (a), (b), and (c) correspond to 0.01, 0.1, and 0.01, respectively.

matrix-product operations in each iteration, CGD requires approximately $2N^2 + 5N + 2$ dot-product computations for convergence.[21] For a sparse system the CGD method can be highly efficient, in terms of both computational complexity and storage requirements, if an indexed storage method is used.[22] However, a previous study[6] of the effect of truncating small weight-matrix elements on the reconstructed image showed that at most 40% of the elements can be set to 0. This is not sufficiently sparse to produce substantial savings by use of an indexed storage method for the weight matrix, because these techniques require at least twice as many words of storage as there are nonzero elements.[22] For a dense matrix the CGD method requires $\sim 2N^3$ scalar multiplications, because there are $\sim N^2$ scalar multiplications in each matrix–vector operation. The major storage requirement for the CGD method is just the $N \times N$ weight matrix. In contrast, the major computational effort for the classic SVD method lies in the QR/QL iterations and needs $\sim 6N^3$ scalar multiplications[23] to yield the eigenvalues and the eigenvectors of $\mathbf{W}^T\mathbf{W}$; the major storage requirement is two $N \times N$ matrices needed for UV decomposition. Thus the CGD method saves ~67% in computation time and ~50% in storage space relative to the SVD method.

For imaging an object whose size and optical properties are of practical interest, e.g., a human breast or brain, there typically are ~2,500 unknowns (for a 10.0 cm × 10.0 cm breast with $\mu_s' \approx 1$ mm^{-1} and $\mu_a \approx 0.01$ mm^{-1} and a pixel size of 2.0 mm × 2.0 mm) in a two-dimensional reconstruction. The number of unknowns can easily increase by a factor of 10 or more if a fully three-dimensional reconstruction is attempted. For a full-ranked linear system, i.e., when at least 2500 source–detector pairs are used, the weight matrix would occupy 25 Mbytes of random access memory for single-precision storage, which would exceed the storage capabilities of most PC's or workstations. This requirement would make implementation of the CGD method very difficult, and the SVD method would be even more problematic because an additional 25 Mbytes of random access memory would be needed for UV decomposition. To reduce the computation time and the storage load we could either reduce the number of source–detector pairs or terminate the reconstruction before it has converged and use an intermediate result. The first approach essentially makes the weight matrix underdetermined and more ill conditioned, so regularization methods and other techniques that make use of *a priori* information are required for useful images to be obtained. As for the second approach, the availability of reconstruction results obtained after an intermediate number m (e.g., 10, 100, ...) of iterations is a feature of iterative solvers that does not exist for either Gaussian elimination (or an equivalent method) or SVD. Although the intermediate results are usually blurred, they do provide reasonably accurate information about the locations and the magnitudes of perturbations after a relatively small number of iterations. One more advantage of the early-iteration reconstructed images, as was pointed out in a previous study,[24] is that they are less sensitive than late-iteration images to noise. That is, when noise is added, more iterations may not guarantee better image quality because of the noise effects; the artifacts that result from the noise tend to grow as the number of iterations increases.

In conclusion, this study demonstrates that Tikhonov–Miller regularization greatly improves the quality of images reconstructed from time-resolved data when the PE algorithm is applied. Further studies are planned or are under way to investigate the effect of updating the regularization parameter λ in the RPE as suggested by Kang and Katsaggelos[20] and to compare the results obtained with the zeroth-, first-, and second-order derivatives of the image $\Delta\mathbf{x}$ as the regularization operator.

ACKNOWLEDGMENTS

This research was supported in part by the National Institutes of Health under grant RO1-CA59955, by the U.S. Office of Naval Research under grant N000149510063, and by the New York State Science and Technology Foundation.

Address correspondence to Randall L. Barbour, SUNY Health Science Center at Brooklyn, Department of Pathology, Box 25, 450 Clarkson Avenue, Brooklyn, New York 11203 (tel. 718-270-1284).

REFERENCES

1. B. Chance and R. R. Alfano, eds., "*Optical Tomography: Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation*, Proc. SPIE **2389** (1995).
2. R. R. Alfano, ed., *Advances in Optical Imaging*, Vol. 21 of OSA Proceedings Series (Optical Society of America, Washington, D.C., 1994).
3. G. Müller, B. Chance, R. R. Alfano, S. R. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. R. Masters, S. Svanberg, and P. van der Zee, eds., *Medical Optical Tomography: Functional Imaging and Monitoring*, Vol. IS11 of Institute Series of SPIE Optical Engineering (Society of Photo-Optical Instrumentation Engineers, Bellingham, Wash., 1993).
4. A. V. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (Institute of Electrical and Electronics Engineers, New York, 1988), Chap. 7, pp. 114–134.
5. H. L. Graber, J. Chang, R. Aronson, and R. L. Barbour, "A perturbation model for imaging in dense scattering media: derivation and evaluation of imaging operators," in Ref. 3, pp. 121–143.
6. J. Chang, R. Aronson, H. L. Graber, and R. L. Barbour, "Imaging diffusive media using time-independent and time-harmonic sources: dependence of image quality on imaging algorithms, target volume, weight matrix, and view angles," in Ref. 1, pp. 448–464.
7. S. Feng and F.-A. Zeng, "Perturbation theory of photon migration in the presence of a single defect," in Ref. 2, pp. 217–228.
8. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Simultaneous scattering and absorption images of heterogeneous media using diffusive wave within the Rytov approximation," in Ref. 1, pp. 320–327.
9. S. R. Arridge, "The forward and inverse problems in time resolved infra-red imaging," in Ref. 3, pp. 35–64.
10. J. Chang, Y. Wang, R. Aronson, H. L. Graber, and R. L. Barbour, "A layer-stripping approach for recovery of scattering media from time-resolved data," in *Inverse Problems in Scattering and Imaging*, M. A. Fiddy, ed., Proc. SPIE **1767**, 384–395 (1992).
11. A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems* (Winston, Washington, D.C., 1977), Chap. 2, p. 45.
12. K. Miller, "Least squares method for ill-posed problems with a prescribed bound," SIAM J. Math. Anal. **1**, 52–74 (1970).
13. R. Arridge, P. van der Zee, M. Cope, and D. T. Delpy, "Reconstruction methods for infrared absorption imaging," in *Time-Resolved Spectroscopy and Imaging of Tissues*, B. Chance, ed., Proc. SPIE **1431**, 204–215 (1991).
14. M. T. Silvia and E. C. Tacker, "Regularization of Marchenko's integral equation by total least squares," J. Acoust. Soc. Am. **72**, 1202–1207 (1982).
15. R. L. Barbour, H. L. Graber, Y. Wang, J. Chang, and R. Aronson, "A perturbation approach for optical diffusion tomography using continuous-wave and time-resolved data," in Ref. 3, pp. 87–120.
16. P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, New York, 1981), p. 44.
17. G. Strang, *Introduction to Applied Mathematics* (Wellesley-Cambridge, Wellesley, Mass., 1986), pp. 419–424.
18. B. R. Hunt, "Application of constrained least squares estimation to image restoration by digital computers," IEEE Trans. Comput. **C-22**, 805–812 (1973).
19. G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," Technometrics **21**, 215–223 (1979).
20. M. G. Kang and A. K. Katsaggelos, "Regularized iterative image restoration based on an iterative updated convex smoothing functional," in *Visual Communications and Image Processing '93*, B. G. Haskell and H.-M. Hang, eds., Proc. SPIE **2094**, 1364–1372 (1993).
21. J. L. Volakis, "EM programmer's notebook," IEEE Trans. Antennas Propag. **37**, 94–96 (1995).
22. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 2nd ed. (Cambridge U. Press, Cambridge, 1992), Chap. 2, pp. 63–82.
23. S. V. Huffel and J. Vandewaffle, *The Total Least Squares Problem, Computation Aspects and Analysis* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1991), p. 115.
24. J. Chang, H. L. Graber, and R. L. Barbour, "Dependence of optical diffusion tomography image quality on image operator and noise," in *Conference Record of the 1995 IEEE Nuclear Science Symposium and Medical Imaging Conference* (Institute of Electrical and Electronics Engineers, New York, 1996), pp. 1524–1528.