Published in: *Medical Optical Tomography: Functional Imaging and Monitoring* (SPIE Institutes Vol. IS11, SPIE Optical Engineering Press, 1993)

A Perturbation Approach for Optical Diffusion Tomography Using Continuous–Wave and Time–Resolved Data

Randall L. Barbour† and Harry L. Graber
Departments of Pathology and Biophysics
SUNY Health Science Center, Brooklyn, NY 11203

Yao Wang, Jeng-Hwa Chang, and Raphael Aronson*
Departments of Electrical Engineering and Physics*
Polytechnic University, Brooklyn, NY 11201

†Author to whom inquiries should be addressed.

1 Introduction

In this paper, we discuss approaches our group has developed for the problem of imaging the interior of dense scattering media [1]. While our principal focus is on potential biomedical applications, we believe our methods are sufficiently general to have applications to other imaging problems as well. We begin our consideration of the imaging problem by assuming that the target medium of interest interacts with the penetrating energy source with sufficient strength to cause intense scattering. We further assume that for essentially all practical schemes, only the multiply scattered signal is measurable. One result of multiple scattering is that all the detected photons will have propagated above and below the plane in which the source and detector lie. Thus, it becomes necessary to explicitly consider volume functions whose spatial distribution will depend on the properties and geometry of the medium and on the geometry and type of illumination scheme. Measurement schemes which have been suggested include steady-state [2], ultrafast [3–5], and amplitude modulated [6, 7] sources. Other schemes include holographic methods which have the potential advantages of directly yielding an image without the need for numerical reconstruction [8, 9]. In developing approaches to image reconstruction, our group has emphasized the first two of the four methods [10–16].

For imaging methods such as x-ray CT, 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 $\mathbf{Mx} = \mathbf{y}$; where $\mathbf{y}$ is the measured response, $\mathbf{M}$ is the imaging operator and $\mathbf{x}$ is the unknown [17]. The form of this relationship is equivalent to Beer's law, which follows directly from the radiation transport equation in the limit of negligible scattering [18]. In the case of strong scattering, though, the change in photon intensity with distance becomes a nonlinear function, because of the contribution of previously scattered photons which add to the net flux at a particular point and direction. A consequence of the additional term is that there is no generally applicable direct method for solving the inverse problem. One approach frequently applied to problems of this type is known as perturbation methods. It is assumed that the composition of the unknown medium deviates only by a small amount from a known, reference medium. This approach is appealing, as it has the effect of reducing a highly nonlinear problem to one which is linear with respect to the difference between the target and reference media. The coefficients in the equations, to be referred to as *weights*, are the gradients of the detector readings with respect to the absorption properties of the points inside the medium.

The idea of weights, or importance functions, is commonly used in the control of nuclear reactors [19]. For isotropically scattering media with continuous wave (CW) measurements, the weight of a given point for a chosen source–detector pair is the product of the forward and adjoint fluxes at that point. For time–resolved (TR) measurements, the weight at a point for a given source-detector pair during a chosen time interval is the convolution of the forward and adjoint fluxes during that time at that point. The forward flux is proportional to the collision density at the point for photons injected from the source; the adjoint represents the time-reversed photon path entering the detector, and is proportional to the collision density when photons are launched from the detector. This result is a counterpart of the sensitivity theorem used in electrical impedance tomography (EIT) [20].

In the following, Sec. 2 introduces several notations and describes the perturbation approach. Sec. 3 derives the weights for the CW and TR measurements, respectively. Sec. 4 introduces the general methodology for solving the perturbation equations and describes several iterative methods, including the conjugate gradient descent, POCS, and multi–grid algorithms. Sec. 5 presents the progressive expansion algorithm for reconstruction from TR data. Sec. 6 shows reconstruction results from both CW and TR data for several test media. Finally, Sec. 7 summarizes the main results and compares the methods proposed here with others. Most of the analysis and results given in Secs. 2 to 6 have previously appeared in, respectively, Refs. [10–15].

2 Notations and The Method of Perturbation

Consider an optical imaging system illustrated in Fig. 1. The unknown medium has a volume Ω bounded by a surface $\partial\Omega$. In a CW system, the incident signal is a point source (isotropic or collimated) with constant intensity (a harmonic wave at an NIR frequency) and the measurements are photon fluxes in the steady state. The measurement is accomplished by injecting photons with intensity S_j at J source locations, $j = 1, 2, \ldots, J$, and measuring the photon flux I_{jk} at K_j detectors, $k = 1, 2, \ldots, K_j$, for the j-th incident signal. These detectors may differ in position as well as orientation. In a TR system, the incident light is a short pulse approximating a delta function, and the measurements consist of the photon fluxes I_{jkl} at different time intervals, $t_l, l = 1, 2, \ldots, L_{jk}$ for each source-detector pair (j, k). Let the medium Ω be divided into N small volume elements called voxels, and let $\boldsymbol{x}_i$ represent the optical property to be imaged at voxel i. The goal of the imaging system is to reconstruct the N optical properties $\boldsymbol{x}_i$ from the $M = \sum_{j=1}^{J} K_j$ or $M = \sum_{j=1}^{J} \sum_{k=1}^{K_j} L_{jk}$ measurements I_{jk} or I_{jkl}. The sources and detectors can either be placed around the entire surface of the medium or only portion of it. One special case is when the sources and detectors are placed on one side of the medium and the backscattered field alone is measured. This measuring strategy affords more feasible implementation, but the reconstruction is more difficult.

In general, $\boldsymbol{x}_i$ is a vector consisting of several optical properties to be imaged at voxel i, including the absorption coefficient $\mu_{a,i}$ and the scattering coefficient $\mu_{s,i}$. Because of the close relationship between absorption and oxygenation status in tissues and organs, imaging the absorption property has been the focus of most optical tissue imaging systems. The scattering property in such media is usually

assumed to be isotropic and homogeneous, and is estimated in advance by other means. In this case, μ_{a_i} becomes the only remaining parameter included in x_i.

For an arbitrary object, the relationship between the photon fluxes I_{jk} at the surface and the absorption properties x_i in the interior can be very complicated. A simple approximation to this relation is given by the diffusion equation [18]. However, even with this simplification, closed-form solutions of the forward problem, i.e, determining I_{jk} from the given x_i and S_j, are available only for media having simple and regular geometric structures. The inverse solution, *i.e.*, reconstructing x_i from S_j and I_{jk}, has been considered by many as intractable. In order to overcome difficulties in solving the inverse problem, we have adopted a *perturbation* approach. It assumes that the absorption coefficients of an unknown object, x_i, are very close to those of a reference medium, x_i^r. Suppose the detector readings for the test and reference media are respectively I_{jk} and I_{jk}^r. The perturbation model relates the differences between the absorption properties of the two media ($\Delta x_i = x_i - x_i^r$) with the changes in detector readings ($\Delta I_{jk} = I_{jk}^r - I_{jk}$) by the following first order approximation:

$$\sum_{i=1}^{N} w_{ijk}^{\mathrm{CW}} \Delta x_i = \Delta I_{jk}, \quad k = 1, 2, \ldots, K_j, \; j = 1, 2, \ldots, J. \tag{1}$$

Here, $w_{ijk}^{\mathrm{CW}} = -\frac{\partial I_{jk}^r}{\partial x_i}$ is called the *weight* of voxel i for source–detector pair (j, k). It is proportional to the reduction of the photon flux at detector k due to the increase in absorption at voxel i when photons are injected from source j.

Let $\mathbf{x}^r$, $\mathbf{x}$, $\Delta\mathbf{x}$ represent the vectors composed of, respectively, the elments x_i^r, x_i, and Δx_i for $i = 1, 2, \ldots, N$. Then, $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^r$. Further, let m represent the one dimensional ordering of the two-dimensional source–detector pair indices, and let $\mathbf{w}_m$ be the vector consisting of the weights of the m-th source-detector pair. Eq. (1) can be recast as:

$$\mathbf{w}_m^T \Delta\mathbf{x} = \Delta I_m, \quad m = 1, 2, \ldots, M. \tag{2}$$

Let $\Delta\mathbf{I}$ be the vector consisting of ΔI_m, and $\mathbf{W}$ the matrix containing $\mathbf{w}_m^T$ as row vectors. The above equations become:

$$\mathbf{W}\Delta\mathbf{x} = \Delta\mathbf{I}\,. \tag{3}$$

The vector $\mathbf{w}_m$ will be called the weight vector for the m-th source–detector pair, and the matrix $\mathbf{W}$ the weight matrix. They are functions of the properties of the reference medium, $\mathbf{x}^r$, the geometry of the medium, and the source–detector configurations.

Given $\Delta\mathbf{I}$ and $\mathbf{W}$, $\Delta\mathbf{x}$ can be determined by solving the linear equation (3). Once $\Delta\mathbf{x}$ is found, $\mathbf{x}$ can be readily identified. When the reference medium is very different from the test medium, the approximation in Eqs. (1)–(3) is inaccurate. Hence the above process should be repeated a number of times, with the reference image and the weight matrix being updated after each iteration. The process can be stopped when the difference between two successive estimates becomes smaller than a prescribed threshold. The flow chart of the above perturbation method is shown in Fig. 2.

The above formulations are for a CW measurement system. In a TR system, let I^r_{jkl} represent the detector readings for a reference medium in time interval t_l for the source-detector pair (j,k). Then the perturbation model relates Δx_i to the changes in the detector reading $\Delta I_{jkl} = I^r_{jkl} - I_{jkl}$ by:

$$\sum_{i=1}^{N} w^{\mathrm{TR}}_{ijkl} \Delta x_i = \Delta I_{jkl}, \quad l = 1,2,\ldots,L_{jk},\ k = 1,2,\ldots,K_j,\ j = 1,2,\ldots,J. \tag{4}$$

Here, $w^{\mathrm{TR}}_{ijkl} = -\frac{\partial I^r_{jkl}}{\partial x_i}$ specifies the reduction of the photon flux in time interval t_l at detector k due to the increase in absorption at voxel i when photons are injected from source j. The vector and matrix representations of the above equation are the same as in Eqs. (2) and (3). The index m now represents the one dimensional ordering of the three dimensional indices (j,k,l) and the vector $\mathbf{w}_m$ consists of the weights associated with the m-th source–detector–time triplet.

There are four separate tasks involved in the perturbation approach described above: 1) Determination of a good initial reference medium that is close to the actual medium; 2) Calculation of the detector readings and the weight functions for the reference medium; 3) Solution of the perturbation equation and update of the reference medium; and 4) Update of the detector readings and weights for the new reference medium. Until now, we have concentrated on the investigation of the second and third problems, which are described below.

3 Determination of the Weight Functions

The original idea of using a weight function to describe the "importance" of a voxel to each source/detector pair is due to Aronson [10]. Assume that photons are injected at source j with intensity S_j. Let Ψ_{ij} be the integrated angular intensity per injected photon at voxel i, V_i the volume of voxel i, and Σ_{t_i} the macroscopic total cross–section of voxel i. Then $N_{ij} = S_j \Psi_{ij} V_i \Sigma_{t_i}$ will be the number of collisions occurring in voxel i. If the absorption coefficient at voxel i is μ_{a_i}, then $\mu_{a_i} N_{ij}$ photons will be absorbed at voxel i and the rest will be scattered, or "born." Suppose that the absorption coefficient of this voxel is increased by $\Delta\mu_{a_i}$ and all the other voxels remain the same, then $\Delta N_{ij} = \Delta\mu_{a_i} N_{ij}$ represents the decrease in the number of photons that will be born in voxel i. Let P_{ik} represent the contribution of a single photon born in voxel i to the photon intensity at detector k. Based on the reciprocity property of photon transport, it can be shown that $P_{ik} = D_k \Phi_{ik}$, where Φ_{ik} represents the flux at voxel i due to a photon injected at detector k. The constant D_k depends on the orientation and solid angle of detector k. From the above relations, the reduction of the photon intensity at detector k due to the change in voxel i is given by:

$$\Delta I_{jk}(i) = P_{ik} \Delta N_{ij} = V_i \Sigma_{t_i} S_j D_k \Psi_{ij} \Phi_{ik} \Delta\mu_{a_i}. \tag{5}$$

When the absorption coefficient at every voxel is changed by $\Delta\mu_{a_i}$, each change will cause a reduction in detector readings. Since a strong absorber in one voxel may cast a shadow in the medium, and hence reduce the photon intensities in other voxels, the total reduction is not a simple sum of the reductions caused by each

voxel independently. But when the change in each voxel is small, we can assume the changes are additive, so that:

$$\Delta I_{jk} = \sum_i \Delta I_{jk}(i) = \sum_i w_{ijk}^{\mathrm{CW}} \Delta \mu_{a,} \tag{6}$$

with

$$w_{ijk}^{\mathrm{CW}} = V_i \Sigma_{t,} S_j D_k \Psi_{ij} \Phi_{ik}. \tag{7}$$

Comparing Eq. (6) to Eq. (1), we see that w_{ijk}^{CW} is simply an estimate of the derivative $\frac{\partial I_{jk}}{\partial x_i}$, when $x_i = \mu_{a,i}$. The weight w_{ijk}^{CW} is proportional to the product of the forward and adjoint photon fluxes due to photons launched from source j and from detector k.

The quantities Ψ_{ij} and Φ_{ik} in Eq. (7) can be obtained by Monte Carlo simulations in which photons are launched from source j and successive collision, scattering, or absorption events are simulated according to the properties of the reference medium. Let F_{ij}^S and F_{ik}^D be the average number of collisions occurring at voxel i due to one injected photon from source j and detector k, respectively. Then $\Psi_{ij} = F_{ij}^S / V_i \Sigma_{t,}$, $\Phi_{ik} = F_{ik}^D / V_i \Sigma_{t,}$.

In the above derivation, Ψ_{ij} and Φ_{ik} are steady–state photon intensities and are appropriate only for a CW system. In a TR system where photon flux $I_{jk}(t)$ at different times is evaluated, the reduction $\Delta I_{jk}(i;t)$ at time t due to a change in the absorption coefficient of voxel i is proportional to the percentage of photons launched from source j that travel first to voxel i and then to detector k in a total time of t. Let $\Psi_{ij}(t)$ and $\Phi_{ik}(t)$ be the photon intensities at voxel i at time t due to photons injected from source j and detector k, respectively; then:

$$\Delta I_{jk}(i;t) \;=\; \Delta \mu_{a,i} V_i \Sigma_{t,} S_j D_k \int_{\tau=0}^{t} \Psi_{ij}(\tau) \Phi_{ik}(t-\tau) d\tau.$$

The reduction in a time interval $t_l = t_1$ to t_2 is simply

$$\Delta I_{jkl}(i) = \int_{t_1}^{t_2} \Delta I_{jk}(i;t) dt.$$

Applying the linear approximation as for the CW case, we obtain

$$\Delta I_{jkl} \;=\; \sum_i w_{ijkl}^{TR} \Delta \mu_{a,i}, \tag{8}$$

with

$$w_{ijkl}^{\mathrm{TR}} \;=\; V_i \Sigma_{t,} S_j D_k \int_{t_1}^{t_2} \int_{\tau=0}^{t} \Psi_{ij}(\tau) \Phi_{ik}(t-\tau) d\tau dt. \tag{9}$$

The time–dependent photon intensities $\Psi_{ij}(t)$ and $\Phi_{ik}(t)$ can be determined by Monte Carlo simulations in the same way as for the CW system, except that the collisions should be counted within separate time intervals.

Using Monte Carlo simulations, the weights for several isotropic scattering and homogeneous media have been evaluated. Examples of weights for CW backscatter

measurements are shown in Fig. 3 [12]. These were calculated according to Eq. (7) for a medium having finite thickness (10 mfp), isotropic scattering, $\mu_{a,i} = 0.01$, and detectors located 1 (Panel A) and 5 (Panel B) mfp from the source. For purposes of visual interpretation, planar cross–sections through the weight functions have been displayed as contour plots. Not surprisingly, the voxels lying directly beneath the source and detector have the greatest weights. The weights decreases very rapidly at increasing depths. This suggests that the maximum depth that can be probed by CW signals is quite limited.

Fig. 4 shows weight functions for TR signals calculated according to Eq. (9) [14]. These results are displayed in the same manner as those in Fig. 3. For an early time interval (Panel A: t_l=9–10 mfp), all the detected photons are restricted to a small finite volume. This suggests that use of early time signals alone may resolve the voxels near the source and detector. At later times (Panel B: t_l = 69–70 mfp), this volume has greatly increased, containing voxels at a greater depth than in the case of CW signals (Compare to Fig. 3, Panel B). This suggests that using TR measurement, a greater depth may be probed than with the CW measurement.

In order to evaluate quantitatively the advantages offered by TR over CW measurements, Fig. 5 shows the ratio of the TR weight to the CW weight in each voxel. The ratios for three cases are illustrated, which differ in source-detector distances and the time intervals examined. In all three cases, a triphastic trend is observed. The value of the weight ratio in the vicinity of the source and detector is less than 1.0. As the distance from these points increases, the ratio at first increases, becoming larger than 1.0, and then declines precipitously. This behavior is not unexpected. It is interesting that in early time intervals, for detectors located close to the source (Panel A: r=1 mfp, t_l=19–20 mfp), the position of the region of greatest sensitivity from TR measurements is located outside the region between the source and detector. In this example, the value of the ratio reaches a maximum of approximately 3.0. At later time intervals for the same detector location (Panel B: t_l=69–70 mfp), the maximum value increases to more than 14.0, and occurs at greater depths. Thus, for this particular source-detector pair, a TR measurement affords considerable advantages over a CW measurement. The advantage gained by time-gating in a particular region is, however, highly dependent on the choice of source-detector pair. When a comparison is made for a similar time interval but at a greater source-detector separation (Panel C: r=10 mfp, t_l=65–70 mfp), the maximum value is reduced to only approximately 1.35. In this case, little advantage is gained by employing a TR measurement.

4 Solution of the Perturbation Equation

The difficulties in solving the perturbation equation are three-fold. First, the dimension of the unknown $\mathbf{\Delta x}$, and hence the weight matrix $\mathbf{W}$, can be extremely large. Second, the weights for voxels far below the surface are very small, so that the matrix $\mathbf{W}$ is ill-conditioned. This makes the solution sensitive to measurement noise and numerical errors. The last, and also the most difficult to deal with, is that the problem may be underdetermined, or ill-posed. This is the case when there are fewer detector readings than voxels, *i.e.*, $M < N$, as often will be the case when only the backscattered field is measured. The problem may still be underdetermined even when $M \geq N$, because the voxels far below the surface cannot be

"seen" by any detector. Mathematically, the column vectors in $\mathbf{W}$ corresponding to the weights for deep voxels all are nearly zero and hence are correlated, which makes $\mathbf{W}$ rank-deficient. In these cases, there exists an infinite number of solutions and additional *a priori* information, *e.g.,* of the type described in Refs. [22–24], must be incorporated to yield a correct solution.

Until now, we have concentrated on the application of iterative schemes for solving the perturbation equation. Compared to their counterparts using direct matrix inversion or pseudo-inversion, these permit progressive reconstruction and selective use of detector readings, which is very important for preliminary investigations [25]. Moreover, it is much easier to incorporate *a priori* information in iterative schemes. An effective iterative scheme must meet at least two requirements: 1) it must converge to a correct solution, and 2) the speed of convergence cannot be too slow. In the following, we compare two iterative schemes in these regards: the CGD and POCS methods. We also present a multi–grid reconstruction algorithm which is faster and in most cases can yield better results.

4.1 Projection onto Convex Sets (POCS)

A set is *convex* if for any two elements x_1 and x_2 in the set, the linear combination $\alpha x_1 + (1-\alpha)x_2$ also belongs to the set, where $0 \leq \alpha \leq 1$. Geometrically, convex sets are the sets without holes or concave boundaries. The method of POCS is applicable in our reconstruction problem if each piece of information about the unknown medium, *e.g.,* detector readings or other *a priori* information, confines the unknowns to a convex set, $C_l, l = 1, 2, \ldots, L$, and the desired solution is in the intersection of these convex sets, i.e, $\mathbf{\Delta x} \in \cap_{l=1}^{L} C_l$. The POCS method reaches a point in the intersection by projecting the current estimate of the solution onto each set sequentially and iteratively. Letting $\mathbf{\Delta x}^{(n)}$ represent the estimate at the n-th iteration, each step in POCS can be represented by:

$$\mathbf{\Delta x}^{(n+1)} = \mathcal{P}_L \circ \mathcal{P}_{L-1} \circ \cdots \mathcal{P}_1 \mathbf{\Delta x}^{(n)} \,. \tag{10}$$

Here, $\mathcal{P}_l$ represents the projection operator onto C_l such that $\mathcal{P}_l \mathbf{\Delta x}$ is the element in C_l that is closest to $\mathbf{\Delta x}$. The symbol $\circ$ denotes the sequential concatenation of two projections. The process in Eq. (10) can be relaxed by replacing $\mathcal{P}_l$ with $\mathcal{T}_l = \mathcal{I} + \lambda_l(\mathcal{P}_l - \mathcal{I})$, where λ_l are user-chosen relaxation constants, and $\mathcal{I}$ stands for the identity operation. Youla [26] has proved that, as long as the intersection of the constraint sets is not empty, iterative projections onto these sets using $0 < \lambda_l < 2$ will converge to their intersection. For a more detailed description of POCS and its applications in medical imaging, see Refs. [26, 27].

To solve the perturbation equation, each equation in (2) can be regarded as a convex set constraint described by : $C_m = \{\mathbf{\Delta x} \in \mathcal{R}^N : \mathbf{w}_m^T \mathbf{\Delta x} = \Delta I_m\}$. Here, $\mathcal{R}^N$ denotes the N-dimensional real Euclidean space. The projection onto this set can be accomplished by [27]:

$$\mathcal{P}_m \mathbf{\Delta x} = \mathbf{\Delta x} + \frac{\Delta I_m - \mathbf{w}_m^T \mathbf{\Delta x}}{\|\mathbf{w}_m\|^2} \mathbf{w}_m \,. \tag{11}$$

Note that the update in (11) is equivalent to the backprojection of the error in an individual reading in the algebraic reconstruction technique (ART) [17].

In addition to the measurement constraints, other *a priori* information can also be incorporated, as long as it forms convex sets. One constraint that we have used is the range constraint:

$$C_r = \{\Delta\mathbf{x} \in \mathcal{R}^N : \Delta x_{\min} \leq \Delta x_i \leq \Delta x_{\max}\}.$$

The projection operator is a simple truncation operation [27]. In our simulation, we have assumed that the real medium always has stronger absorption than the reference medium such that $\Delta x_i \geq 0$ or $\Delta x_{\min} = 0$. An appropriate $\Delta x_{\max}$ is also chosen, based on our a *priori* knowledge of the absorption properties of the test medium. The POCS method using only the measurement and range constraints is equivalent to the additive form of the weighted ART [17]. But POCS is much more general, since almost any *a priori* information that may arise in practice forms a convex set constraint and can be handled by the POCS method.

A problem with POCS is that its convergence rate can be quite slow. This is because each projection makes use of one constraint without considering others. To alleviate this problem, appropriate pre-processing should be conducted to reduce the number of unknowns as much as possible. One approach is to make further use of the positivity of Δx_i. Since ΔI_{jk} or ΔI_{jkl} in Eq. (1) or (4) is a summation of positive numbers, each term in the sum should be close to zero if the sum is very small. This constraint is enforced by setting $\Delta x_i = 0$ if $\Delta I_{jk}/w_{ijk} \leq T$ or $\Delta I_{jkl}/w_{ijkl} \leq T$, where T is a preset threshold. When T is selected properly, this process is very effective in reducing the number of unknowns and speeding up the reconstruction process.

Another drawback of POCS is that it may fall into a "dead loop" if the constraint sets formed by the measurements fail to intersect. This is often the case in practice due to measurement noise. This problem can be circumvented by enlarging each measurement constraint set to $C_m = \{\Delta\mathbf{x} \in \mathcal{R}^N : \mid \mathbf{w}_m^T\Delta\mathbf{x} - \Delta I_m \mid < \epsilon_m\}$. The corresponding projection operator can be accordingly modified from Eq. (11). In addition, if it is known that the reconstructed image is close to the reference image, we can add another constraint set: $C_e = \{\Delta\mathbf{x} : \|\Delta\mathbf{x}\| < \epsilon\}$. The projection operator onto this set can be found in Ref. [27]. This projection will regularize the iterative process so that the solution is never too far from the reference image.

4.2 Least Square Solution Using Gradient Descent Methods

An alternative method for solving the perturbation equation is to minimize the following squared error:

$$E(\Delta\mathbf{x}) = \frac{1}{2}\sum_m(\mathbf{w}_m^T\,\Delta\mathbf{x} - \Delta I_m)^2 = \frac{1}{2}\|\mathbf{W}\Delta\mathbf{x} - \Delta\mathbf{I}\|^2. \tag{12}$$

Since E is a quadratic function of $\Delta\mathbf{x}$, the minimal point can be reached by a gradient descent method. The basic update operation is described by

$$\mathbf{g} = \sum_m(\mathbf{w}_m^T\Delta\mathbf{x}^{(n)} - \Delta I_m)\mathbf{w}_m = \mathbf{W}^T(\mathbf{W}\Delta\mathbf{x}^{(n)} - \Delta\mathbf{I}) \tag{13}$$

$$\Delta\mathbf{x}^{(n+1)} = \Delta\mathbf{x}^{(n)} - \alpha\mathbf{g}\,. \tag{14}$$

Here, $\mathbf{g}$ is the gradient at the n-th iteration, and α is called the *step-size,* which must be chosen appropriately to guarantee convergence. We see that the gradient vector in Eq. (13) is equivalent to the summation of the backprojections from all the readings. Hence, each update described by Eq. (14) is based on the total backprojection, similar to the simultaneous iterative reconstruction technique (SIRT) [17].

There are a variety of gradient descent methods. We have concentrated on the conjugate gradient descent (CGD) algorithm, which is particularly suitable for large-scale least–squares problems [28]. It modifies the current estimate in the conjugate gradient direction, which is a linear combination of the previous gradient directions. Theoretically, the CGD method can reach the optimal solution in no more than N iterations. In practice, many more iterations are required due to rounding errors. It is, nevertheless, much faster than the POCS method. In addition, it is more robust in the presence of noise, since each update is based on the total backprojection. The error in one measurement will not have as significant an effect on the final solution as that with POCS–type sequential algorithms. Budinger *et al.* [29] have found that, of the many iterative techniques for image reconstruction in emission computed tomography, the best results and fastest convergence can be achieved with CGD.

One major problem with CGD is that it is reliable only for overdetermined problems. In an underdetermined setting, it simply reaches a solution that is closest to the initial point. It is very difficult to impose non-linear *a priori* constraints to force the CGD process towards the true solution. One method has been to periodically modify an intermediate solution by projecting it onto convex constraints not included in the error measure [30]. But this will slow down the convergence of the method substantially.

4.3 A Multi-Grid Reconstruction Algorithm

Facing the difficulties with POCS and CGD, we have developed a *multi–grid* progressive reconstruction scheme. It first obtains a coarse–grid image in which each cell corresponds to, say, $3 \times 3 \times 3$ voxels in the original medium, and then further reconstructs the regions of interest (voxels with non-zero values, in our case) on a finer grid. The weight for a coarse cell is approximated by the sum of the weights of all the voxels contained in the cell. The reconstruction at each step can be accomplished by either CGD or POCS. In most cases, the reconstruction problem at intermediate levels can be made overdetermined, and the CGD algorithm is preferable since it converges faster. Only the reconstruction at the last level needs to be accomplished by POCS, if the number of remaining unknowns exceeds the number of available measurements.

Because the number of unknowns to be resolved at each level is much smaller than the total number of voxels, the multi–grid method requires much less computation than CGD or POCS. In addition, such an approach can force the iterative process to proceed in the correct direction if good solutions can be obtained at coarse resolutions. Mathematically, this is a way to prevent the iterative process from being trapped in an undesirable local minimum and enforce a solution that is correct at various scales. The success of such a scheme, however, depends on the

composition of the test medium and the way the coarse grids are formed. Multigrid implementation of several iterative algorithms for image reconstruction from straight–line projections has been studied in [31]. A "local smoothness" property has been described as the necessary condition for the method to converge faster than the direct single grid method. This result is applicable to our problem as well. We have applied this method to CW measurements where the problem is severely underdetermined.

5 A Progressive Expansion Algorithm for TR Data

We have previously recognized that the gradient of weight as a function of depth is highly dependent on the tempo–spatial characteristics of the detector. We have consequently developed a layer-stripping algorithm which progressively evaluates the data in order of decreasing depth component of the weight gradients [12, 14]. When applied to TR data, this approach offers the additional advantage of reducing the number of unknowns and dimension of the weight matrix, thereby improving the determinedness and decreasing computation time. A difficulty with this approach is that reconstruction errors encountered early on may propagate to later times and be amplified. A description of this scheme and an approach for reducing the propagation of errors is described in this section.

5.1 The Progressive Expansion (PE) Algorithm

This algorithm makes use of detector readings in different time intervals separately and in a progressive manner. The process is illustrated in Fig. 6. In each time interval, we only consider the region which may contribute to the detector readings and which is not yet fixed from previous reconstruction stages. The contribution to the detector readings in this time interval from the voxels solved earlier is first subtracted. The new unknowns are then solved for by using the perturbation equation. We have found that, if detector readings are available for every mean free time (mft) [1] after the photons are injected, then the number of voxels having zero weights will be less than the number of detector readings in some early time intervals. This makes the perturbation equation overdetermined at these times, and the CGD method can be used reliably. In the meantime, we still apply the positivity constraint to the reconstruction result after each time interval. Specifically, for a time interval t_l, the algorithm is as follows:

Step 1: Find all detectors which may receive signal during t_l for a given source. This is accomplished by examining the weights of all the voxels for each source-detector pair. Only the pairs that have at least one non-zero weight are considered. The number of detectors found is here denoted as M_l. The index set of source-detector pairs is denoted by $\mathcal{M}_l$.

Step 2: Find all the non-fixed voxels which may contribute signals to any of the detectors determined in Step 1. This is accomplished by including all voxels that have non-zero weights for at least one of the selected detectors. The number of voxels found is here denoted by N_l. The set of voxel indices is denoted by $\mathcal{N}_l$.

[1] One mft is the time it takes a photon to propagate a distance of one mean free path (mfp).

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

$$\sum_{i \in \mathcal{N}_l} w_{ijkl} \Delta x_i = \Delta I_{jkl} - \sum_{i \in \mathcal{N}_1, \ldots, \mathcal{N}_{l-1}} w_{ijkl} \Delta x_i, \quad (j,k) \in \mathcal{M}_l$$

or

$$\mathbf{W}_l \Delta \mathbf{x}_l = \Delta \mathbf{I}_l. \tag{15}$$

Step 4: Find the least squares solution of Eq. (15) for $\Delta x_i, i \in \mathcal{N}_l$, using the CGD method.

Step 5: Apply the positivity constraint to the reconstructed values. This is accomplished by setting $\Delta x_i = 0$ if $\Delta x_i < 0$.

Step 6: Fix $\Delta x_i, i \in \mathcal{N}_l$. Go to the next time interval.

5.2 An Overlapping Scheme

In the algorithm described above, a voxel is solved and fixed when its signal is first detected. Since the weight of this voxel at this initial time is usually small, the reconstruction is not reliable. The error in this voxel will further affect the reconstruction of deeper voxels at later times. In order to overcome this problem, an overlapping scheme has been developed. In this method, a voxel is not fixed the first time it is solved. Rather, it is considered in the next several time intervals as well. In each time interval, new voxels that first contribute signals in this time interval and certain old voxels are considered. After the solution of the perturbation equation involving all these voxels, a partial set of the old voxels is fixed. And the next time interval is considered. This overlapping scheme can greatly reduce the propagation of reconstruction errors in the proposed PE algorithm and provide more reliable reconstructions. However, it also leads to several problems. First, the early detected signals are more ballistic and contain information that is not merged with scattered signals from other voxels too much. Overlapping will lessen the influence of this kind of information in the final result. Second, overlapping will increase the number of unknowns to be solved in each time interval and consequently the computation time. Too much overlapping may even make the problem underdetermined at very early time intervals, which will make the reconstruction in the following intervals less reliable. If we overlap many time intervals at once, the reconstruction problem will become similar to that in the CW case. Therefore, we need to select the overlapping interval properly to arrive at a good compromise.

6 Results

The POCS, conjugate gradient descent, multi-grid, and the PE algorithms have been implemented on a Sparc workstation and a parallel computer testbed, *Kilonode* [32]. With POCS, only the measurement and range constraints are used. With the multi-grid method, two levels of grids are used, the original (in which each voxel is a cube of volume 1 mfp^3) and a coarse representation formed by grouping every $3 \times 3 \times 3$ original voxels. The CGD and POCS methods are respectively used for the coarse

and fine reconstructions. For each test medium, a corresponding homogeneous medium was chosen as the reference, and the weights for the reference medium for the CW and TR cases were calculated, using Monte Carlo methods, according to Eqs. (7) and (9). The intensity data used in the reconstruction were of two types. The first is data from Monte Carlo simulations of reflectance measurements. The second type is calcuated directly from Eq. (3), by multiplying the matrix of Monte Carlo derived weights and the vector of the absorption coefficients chosen for all voxels. Reconstruction from calculated data is intended to reveal the achievable resolution of the proposed methods in the ideal case when the perturbation model is valid.

The source and detector configurations for various experiments are shown in Figs. 7 to 10. The reconstructed images are given in Figs. 12 to 20. In all the figures of the original and reconstructed images, the gray level represents the relative increase in absorption from the reference media. The degree of darkening is in proportion to the relative increase. For display purposes, the images have been scaled to the same minimum and maximum darkness. Hence, the same gray level may represent different absorption levels in separate figures. In the following, we present the results for three types of test media separately.

6.1 Media Containing Point Absorbers

To evaluate the capability of the proposed approaches to resolve absorption distributions at different depths from backscatter measurements alone, we have attempted the reconstruction of one or two closely juxtaposed point absorbers of size 1 mfp^3 buried at various depths in an isotropically scattering half-space. The absorption coefficient of the background medium is 1 % (*i.e.,* the macroscopic absorption cross-section is 1% of the total cross-section) and that of the absorber is 5 %. A volume of size $41 \times 41 \times 10$ mfp^3 is reconstructed, with a total of 16,810 unit-size voxels. Only calculated data have been considered for this type of test media.

Results from CW data

Reconstruction results using the CGD, POCS and multigrid methods for a medium containing a single point absorber at depth 2–3 mfp are presented in Figs. 12(a)–(d). The reconstructed images by both POCS and CGD methods correctly reveal the existence of the absorber. But the CGD algorithm does not yield as sharp an image as the POCS method. In fact, some voxels are given values outside the correct range with the CGD method. The POCS method yields a nearly perfect reconstruction, except that the voxel below the absorber is slightly darkened. With the multi-grid method, a good estimate of the true image is obtained with the coarse grid solution, and the final result is similar to that obtained by the direct POCS method. In terms of computation time, the multi-grid method is the fastest of the three. Specifically, the POCS and CGD methods used 50 and 33 hours of computing time of a SUN Sparc workstation, while the multi-grid method required only 5 hours.

To compare the convergence behavior of the above algorithms, Fig. 11 presents their error reduction curves, where the error is as defined in Eq. (12). It is seen that the CGD algorithm has the fastest convergence rate among these methods. It also has a smaller final error, although the reconstructed image is not as close to the true image as those produced by the POCS and multi-grid methods. This implies

that it converges to a legitimate, but wrong, solution. The convergence rates of the POCS and multi–grid methods are comparable to each other, with the latter being more uniform.

Figs. 12(e)–(j) present the results of the multi–grid method for three other point absorber configurations: a single absorber at depth of 4–5 mfp, and two absorbers separated by 1 mfp at depths 2–3 and 4–5 mfp. We see that the results for media containing absorbers at 4–5 mfp are less accurate, with the darkest regions occurring at depth 3–5 mfp. Many more iterations are required to reach better results. For example, starting with the image in Fig. 12(j), the use of additional 500 CGD iterations produced the image in Fig. 12(l). The reason for the requirement of more iterations is that the weights for deeper voxels are much smaller than those for shallower voxels, so the iterative algorithm does not change the deeper voxels until the shallower voxels have been solved.

Results from Calculated TR Data

Reconstruction results for the medium containing a single point absorber at depth 1–2 and 4–5 mfp, respectively, are presented in Figs. 13 and 14. In order to examine the noise sensitivity of the PE algorithm with and without overlapping, we show results from noiseless as well as noise-added data. The ratio of the noise variance to the mean value of the difference in calculated readings is used as the measure of noise level. It can be seen that overlapping produces a higher quality image in the presence of noise. Further, a comparison of Figs. 13(d) and 14(d) reveals that in deeper layers, the reconstructed image is more sensitive to noise. It should be noted that most of the image artifacts seen at locations away from the central axis of the source would be expected to vanish had additional source positions been considered.

Reconstructed images at different noise levels, for the medium containing a point absorber at 1–2 mfp, are presented in Fig. 15. As expected, the reconstruction quality degrades as the noise level increases, but the result at the 20% noise level is still quite reasonable. Figs. 16 and 17 show the results for two point absorbers separated by 1 mfp at depths 1–2 and 4–5 mfp, respectively. The results are similar to the single absorber case — overlapping significantly outperforms the non-overlapping case and deeper layers are more sensitive to noise.

Overall, when comparing the above results to those from CW data, it is obvious that TR data afford better reconstructions.

6.2 A Medium Containing the "T" Absorber

In this section, we present the results for an isotropically scattering medium containing a "T" shaped black absorber. Fig. 18(a) shows the X-Z (left) and Y-Z (right) cross–sections of the medium. Figs. 18(b)–(d) show the reconstruction results by the CGD and multi–grid methods, using simulated CW data. The multi–grid method correctly identified the existence of a strong absorber, although the absorption level in the "cap" of the "T" absorber was underestimated, and the "stem" is smeared.

For TR measurements, we have attempted reconstructions from both calculated and simulated data. For the simulated data, two different detector configurations are used. Fig. 19(b) shows the result by the PE algorithm based on calculated

data. The reconstructed image has a sharp transition at the "cap" of the "T" but there is evidence of image degradation at greater depths. This is due to numerical errors, as well as to the propagation of reconstruction errors. Figs. 19(c) and 19(d) show images reconstructed from the Monte Carlo simulation data but employing the detector distributions in Figs. 10(a) and 10(b), respectively. The poor result seen in Fig. 19(c) is a consequence of too great a spacing between the detectors (Fig. 10(a)). When the denser packing of detectors (Fig. 10(b)) was used, a much better result was obtained, in which the "cap" of the "T" was reconstructed quite accurately but the "stem" was smeared.

Note that the changes in detector readings in the presence of such a large black absorber will not follow the perturbation model accurately. In general, the rate of increase in ΔI_m is less than linear with increasing Δx_i. Nevertheless, as shown here, the reconstructed images based on this model contain a sharp darkening in the region where the "T" absorber is located. This is a very encouraging result as it suggests that even with a reference medium that is quite different from the real medium, one iteration of the perturbation method may yield useful results.

6.3 A Three-Layer Medium

To evaluate the effect of the accuracy of the reference medium on the reconstruction results, we have attempted to reconstruct a three-layer medium using a homogeneous medium as the reference. The original medium is shown in Fig. 20(a). It is 10 mfp thick, and the three layers are located from 0–3, 3–7, and 7–10 mfp, respectively. The absorption coefficient of the first and third layers is 1%, and that of the middle layer is 5%. For simplicity, we assume it is known that the medium has a stratified structure. With this assumption, a perturbation equation with only 10 unknowns can be set up, each representing the absorption property of a single layer. The weight for each layer is obtained by summing the weights for all the voxels in it. A homogeneous medium with an absorption cross–section of 1% of the total cross–section is used as the reference.

Results from Calculated and Simulated CW Data

To evaluate the effect of the accuracy of the weights on the reconstruction results, the weight functions have been calculated for both the target and reference media. These two sets of weights will be referred to as the *three-layer weights* and *half-space weights*, respectively. Two sets of readings are tested: one is calculated, using the three-layer weights; the other is obtained from Monte-Carlo simulations. In each case, the reconstruction is performed using both the three-layer and half-space weights.

The reconstructed images in various cases are shown in Figs. 20(b)–(e). Fig. 20(b) is obtained from the calculated CW data, using three-layer weights. The transition from the first to the second layer is accurately reconstructed, but the transition from the second to the third is blurred. When the half-space weights are used, an incorrect image resulted (Fig. 20(c)), with a darker and narrower layer appearing before the true second layer. This is reasonable, since the half-space weights used in the planes within and below the second layer are larger than the three–layer weights.

With the simulated CW data, when the three-layer weights are used (Fig. 20(d)), the algorithm reveals correctly, with slight blurring, the three-layer structure. But the absorption level of the second layer is overestimated and its thickness is underestimated. With the half-space weights (Fig. 20(e)), the transition from the first to the second layer is further blurred, and as with the calculated data, the thickness of the second layer is underestimated. This is as expected, since the perturbation model is not valid in the presence of a non-localized variation between the true and reference media. In order to obtain a more accurate reconstruction, the perturbation method should be iterated several times, as described in Sec. 2, with the reference medium and weights being updated after each step.

Results from Simulated TR Data

Fig. 20(f) shows the reconstruction result from simulated TR data using the weights calculated for the half-space reference medium. As expected, this reconstruction result is much more accurate than the reconstruction from CW data. The three–layer structure is correctly revealed, with slight blurring. The transition between the first and second layers was correctly reconstructed and the transition from the second to third was less accurate. As with the CW data, the absorption level of the second layer is overestimated and its thickness underestimated.

7 Discussion

On first consideration, image reconstruction based on the analysis of highly scattered signals appears infeasible. It is clear that the radiation transport equation correctly describes the relation between the internal structure of a medium and the angular intensity at its surface, but no closed–form solutions are available to the general inverse problem. In our attempt to develop practical imaging algorithms, we chose to apply a perturbation approach. This allows the problem of solving an integro-differential equation to be recast as a much simpler one of solving a system of linear equations. Several difficulties, both theoretical and numerical, inevitably arise when this approach is taken. Among the important questions are how should one choose a reference medium, and how should the elements of $\mathbf{I}_0$ and $\mathbf{W}$ be measured or calculated? In addition, the linear assumption used in the perturbation model is accurate only when the reference is similar to the target medium. It thus becomes important to examine the degradation of the reconstruction quality when this assumption is increasingly violated. Other important questions are, how sensitive is the reconstruction algorithm to noise in the measurement of $\mathbf{I}_0$ and $\mathbf{I}$, and how can an accurate, or at least approximately accurate solution be obtained? The studies presented here were directed to partially answer some of these questions.

Theoretical derivations were presented to show how weight functions can be practically calculated for a reference medium with any desired set of properties. As the particular examples shown all dealt with model media in which the correct structure for the reference was known in advance, the problem of finding the best reference without *a priori* knowledge of the medium has not been emphasized here. One approach would be to empirically relate MRI data to absorption and scattering coefficients of tissue, thereby permitting consideration of non–homogeneous reference media.

The decision to use iterative rather than direct algorithms was made because these allowed us to incorporate strategies for reducing numerical errors and noise sensitivity, and for transforming an underdetermined problem into one which is overdetermined. Each of the methods described here has some particular advantage.

- POCS is well-suited for underdetermined problems because of its ability to incorporate additional information and constraints to a solution that lies within the reasonable set.

- The multigrid algorithm can reduce the effect of error in the weight calculation by averaging the weights over a larger volume. It also reduces the number of unknowns in two ways. First, because the coarse grid contains many fewer voxels than the finer grids; second, because coarse-level reconstructions can identify large regions that are not of interest, and whose properties can therefore be fixed in finer-level reconstructions. Either or both of these may make the problem overdetermined, allowing use of the faster CGD algorithm instead of POCS.

- The PE algorithm for TR-data reconstructions also reduces the number of unknowns, by considering only the small set of voxels with non-zero weights in early time windows, and fixing the absorption coefficients in these voxels at later times. This also leads to the separation of the contributions of different voxels to the measured elements of $\mathbf{\Delta I}$. This type of algorithms can greatly increase the depth at which reasonable reconstructions can be obtained.

The reconstructions presented here include examples in which a given target medium was compared to more than one reference medium. The results indicate that the linear perturbation model is robust to variation in $\mathbf{I}_0$ and $\mathbf{W}$, at least for the relatively simple structures considered here. While useful images of complex media would likely require iterative updates of these quantities, results shown indicate that a useful image can be obtained from a single step of the perturbation model.

The perturbation approach, and the algorithms presented here, can be used in a complementary manner with other strategies under consideration. Schlereth *et al.* [32] have formulated image reconstruction as a learning problem in a neural net structure. This method iteratively updates the current image estimate (*i.e.* the node properties) until values calculated for photon fluxes (*i.e.* the output of the net) match those measured (*i.e.* the training vector). This method, taken from suggestions by Singer *et al.* [33] and Grünbaum *et al.* [34], offers great flexibility in modeling the forward problem. When provided with a good initial estimate of the structure of the target medium, it yields accurate and numerically stable results. The method by Singer *et al.* [33] and Grünbaum *et al.* [34] use a general mathematical model, derived from the discretization of the transport equation in both the spatial and angular domains. It is applicable to arbitrarily shaped objects, and permits iterative estimation of both the absorption and scattering coefficients of a target medium. However, the overwhelming amount of computation required has limited them, so far, to evaluation of media having few voxels (*i.e.* ~ 100). Using a method similar to the perturbation model described here is the approach of Arridge *et al.* [21,35], which makes use of the diffusion approximation to the transport equation. Closed-form solutions for the scattered field and its gradient have been derived, and used for reconstruction of cylindrical media. They have also adopted a

finite element calculation for computing solutions to the diffusion equation [36]. An important aspect of this work is the emphasis placed on regularization methods for restricting the solutions to those that are physically reasonable. We anticipate that further algorithmic developments by our group will incorporate similar techniques to complement the use of *a priori* information for restricting the set of allowed solutions.

8 References

[1] B. Chance, "Optical method," *Ann. Rev. Biophys. Chem.*, vol. 20, pp. 1–28, 1991.

[2] K. Yoo, G. Tang, and R. Alfano, "Coherent backscattering of light from biological tissues," *Appl. Opt.*, vol. 29, pp. 3237–3239, 1990.

[3] L. Wang, P. Ho, C. Liu, G. Zhang, and R. R. Alfano, "Ballistic 2-d imaging through scattering walls using an ultrafast optical Kerr gate," *Science*, vol. 253, pp. 769–771, 1991.

[4] G. Hebden and R. A. Kruger, "A time-of-flight breast imaging system: spatial resolution performance," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE–1431, pp. 225–231, Jan., 1991.

[5] B. Chance *et al.*, "Comparison of time-resolved and unresolved measurements of deoxyhemoglobin in brain," *Proc. Nat. Acad. Sci.*, vol. 85, pp. 4971–4975, 1988.

[6] J. Fishkin, E. Gratton, M. vandeVen, and W. Mantulin, "Diffusion of intensity modulated near-infrared light in turbid media," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE–1431, pp. 122–135, 1991.

[7] E. Sevick, B. Chance, J. Leigh, S. Nioka, and M. Maris, "Quantitation of time- and frequency-resolved optical spectra for the determination of tissue oxygenation," *Anal. Biochem.*, vol. 195, pp. 330–351, 1991.

[8] K. G. Spears, T. Serafin, N. H. Abramson, X. Zhu, and H. Bjelkhagen, "Chrono-coherent imaging for medicine," *IEEE Trans. Biomedical Eng.*, vol. 36, pp. 1210–1221, 1989.

[9] A. Rebane and J. Feinberg, "Time-resolved holography," *Nature*, vol. 351, pp. 378–380, 1991.

[10] R. L. Barbour, H. Graber, R. Aronson, and J. Lubowsky, "Model for 3-d optical imaging of tissue," in *Proc. 10th Annual Int. Geoscience and Remote Sensing Symposium (IGARSS)*, vol. II, pp. 1395–1399, 1990.

[11] R. Aronson, R. Barbour, J. Lubowsky, and H. Graber, "Application of transport theory to infrared medical imaging," *Modern Mathematical Models in Transport Theory; Advances and Applications*, vol. 51, pp. 64–75, 1991.

[12] R. Barbour, H. Graber, R. Aronson, and J. Lubowsky, "Imaging of subsurface regions of random media by remote sensing," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE-1431, (Los Angeles), pp. 192–203, Jan. 1991.

[13] Y. Wang, J. Chang, R. Aronson, R. Barbour, H. Graber, and J. Lubowsky, "Imaging scattering media by diffusion tomography: An iterative perturbation approach," in *Proc. Physiological Monitoring and Early Detection Diagnostic Methods*, vol. SPIE-1641, (Los Angeles), pp. 58–71, Jan. 1992.

[14] R. L. Barbour, H. Graber, R. Aronson, and J. Lubowsky, "Imaging of diffusing media by a progressive iterative backprojection method using time-domain data," in *Proc. Physiological Monitoring and Early detection Diagnostic Methods*, vol. SPIE-1641, (Los Angeles), pp. 21–34, Jan. 1992.

[15] J. Chang, Y. Wang, R. Aronson, H. L. Graber, and R. Barbour, "A layer-stripping approach for recovery of scattering media from time resolved data," in *Proc. Inverse Problems in Scattering and Imaging*, vol. SPIE–1767, (San Diego), pp. 384–395, July 1992.

[16] J. Chang, H. L. Graber, Y. Wang, R. Aronson, and R. Barbour, "Time–resolved imaging in dense scattering media," in *Proc. Physiological Imaging, Spectroscopy, and Early Detection Diagnostic Methods*, vol. SPIE-1887, (Los Angeles), Jan. 1993. In Press.

[17] A. V. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging.* IEEE Press, 1988.

[18] A. Ishimaru, *Wave Propagation and Scattering in Random Media.* Academic Press, 1978.

[19] J. R. Lamarsh, *Nuclear Reactor Theory.* Adison–Wesley Publishing Co., Reading, MA, 1966.

[20] J. G. Webster, ed., *Electrical Impedance Tomography.* Bristol, Englands: Adam Hilger, 1990.

[21] S. R. Arridge, P. van der Zee and D. T. Delpy, "Reconstruction methods for infra-red absorption imaging," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE–1431, pp. 204–217, 1991.

[22] R. L. Barbour, H. Graber, R. Aronson, and J. Lubowsky, "Determination of macroscopic optical properties of random media by remote sensing," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE–1431, (Los Angeles, CA), pp. 152–162, Feb. 1991.

[23] R. Aronson, "Exact interface conditions for photon diffusion," in *Proc. Physiological Monitoring and Early detection Diagnostic Methods*, vol. SPIE-1641, (Los Angeles), pp. 72–78, Jan. 1992.

[24] H. Graber *et al.*, "Evaluation of steady state, time-, and frequency-domain data for the problem of optical diffusion tomography," in *Proc. Physiological Monitoring and Early detection Diagnostic Methods*, vol. SPIE-1641, (Los Angeles), pp. 6–20, Jan. 1992.

[25] Y. Censor, "Finite series-expansion reconstruction methods," *Proceedings of IEEE*, vol. 71, pp. 409–419, 1983.

[26] D. C. Youla and H. Webb, "Image restoration by the method of convex projections onto convex sets – Part I," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 81–94, Oct. 1982.

[27] P. Oskoui-Fard and H. Stark, "Tomographic image reconstruction using the theory of convex projections," *IEEE Trans. Med. Imaging*, vol. MI-7, pp. 45–58, Mar. 1988.

[28] P. E. Gill *et al.*, *Practical Optimization.* Academic Press, 1981.

[29] T. F. Budinger *et al.*, "Emission computed tomography," in *Image Reconstruction From Projections – Implementation and Applications* (G. T. Herman, ed.), Spring-Verlag, 1979.

[30] A. K. Katsaggekis and S. N. Efstratiadis, "A class of iterative signal restoration algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, pp. 778–786, May. 1990.

[31] T. S. Pan and A. E. Yagle, "Numerical study of multigrid implementation of some iterative reconstruction algorithms," *IEEE Trans. Medical Imaging*, vol. MI-10, pp. 572–588, Dec. 1991.

[32] F. H. Schlereth, J. M. Fossaceca and R. Barbour, "Imaging of diffusion media with a neural net formulation: A problem in large scale computation," in *Proc. Physiological Monitoring and Early detection Diagnostic Methods*, (Los Angeles), pp. 46–57, Jan. 1992.

[33] J. R. Singer, F. A. Grünbaum, P. Kohn, and J. Zubelli, "Image reconstruction of the interior of bodies that diffuse radiation," *Science*, vol. 248, pp. 990–993, 1990.

[34] F. A. Grünbaum, P. Kohn, G. Latham, J. R. Singer, and J.P.Zubelli, "Diffuse tomography," in *Proc. Time-Resolved Spectroscopy and Imaging of Tissues*, vol. SPIE–1431, pp. 232–238, Jan., 1991.

[35] S. Arridge, M. Cope, and D. Delpy, "The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis," *Phys. Med. Biol.*, vol. 37, pp. 1531–1560, 1992.

[36] S. Arridge, M. Schweiger, and D. Delpy, "Iterative reconstruction of near infrared absorption images," *Proc. Inverse Problems in Scattering and Imaging*, vol. SPIE–1767, (San Diego), pp. 372-383, July 1992.

List of Figures

12. Results for media containing a point absorber from calculated data: (a) original: a single absorber at 2–3 mfp; (b)–(c) reconstructions of (a) by CGD (b) and POCS (c), both after 100 iterations; (k) coarse-grid solution of (a) by multi–grid (using CGD); (d) fine–grid solution of (a) by multi–grid (using POCS, 100 iterations); (e) original: a single absorber at 4–5 mfp; (f) fine–grid solution of (e) by multi–grid (using POCS, 200 iterations); (g) original: two absorbers 1 mfp apart at depth 2-3 mfp; (h) fine–grid solution of (g) by multi–grid (using POCS, 200 iterations); (i) original: two absorbers 1 mfp apart at depth 4–5 mfp; (j) fine–grid solution of (i) by multi–grid (using POCS, 200 iterations); (l) fine–grid solution of (i) by multi–grid with additional 500 CGD iterations starting from (j).

13. Reconstruction of the medium containing a point absorber at depth 1–2 mfp, from calculated TR data. The source configuration is as shown in Fig. 10(a). The detectors are uniformly distributed on the surface, separated by 1 mfp in each direction, within a radius of 10 mfp. The images on the left shows the Y-Z cross–section and those on the right the X-Z cross–section. (a) Original; (b) Reconstruction from noiseless data without overlapping; (b) Reconstruction from noise–added data without overlapping; the noise level is 2%. (d) Reconstruction from noise–added data with an overlapping interval of 3 mfp. All the reconstruction results are obtained by considering data received between 1–12 mfp using a time interval of 1 mfp. "v" indicates location of source. Images shown here were limited to within a hemisphere of 6 mfp originating from the position of the photon entry.

14. Reconstruction of the medium containing a point absorber at depth 4–5 mfp, from calculated TR data. The source–detector configuration is the same as in Fig. 13. (a) Original; (b) Reconstruction from noiseless data, without overlapping, after 20 time windows of width 1 mfp (*i.e.,* 0–1, 1–2, ..., 19–20); (c) Reconstruction from noise–added data without overlapping; The noise level is 10%. (d) Reconstruction from noise–added data with an overlapping interval of 3 mfp. The results in (b), (c) and (d) are limited to within hemispheres of 10, 6, and 5 mfp originating from the point of photon entry respectively.

15. Reconstruction of the medium in Fig. 13(a) at different noise levels from calculated TR data. The noise levels are, respectively (a) 1%, (b) 5%, (c) 10%, and (d) 20%. All the reconstruction results are obtained with an overlapping interval of 3 mfp, after 12 time windows of width 1 mfp (*i.e.,* 0–1, 1–2, ..., 11–12). Volume of medium considered in the reconstructed images is the same as in Fig. 13.

16. Reconstruction of the medium containing two point absorbers separated by 1 mfp at depth 1–2 mfp, from calculated TR data. The source–detector configuration is the same as in Fig. 13. (a) Original; (b) Reconstruction from noiseless data without overlapping; (c) Reconstruction from noise–added data without overlapping; the noise level is 2%. (d) Reconstruction from noise–added data, with an overlapping interval of 3 mfp. All the reconstruction results are obtained after 14 time windows of width 1 mfp. The results in (b), (c) and (d) are limited to the regions within hemispheres of 10, 7, and 6 mfp originating from the point of entry, respectively.

17. Reconstruction of the medium containing two point absorbers separated by 1 mfp at depth 4–5 mfp, from calculated TR data. The source and detector configuration is the same as in Fig. 13. (a) Original; (b) Reconstruction from noiseless data, without overlapping, after 20 time windows of width 1 mfp; (c) Reconstruction from noise–added data, without overlapping; the noise level is 10%. (d) Reconstruction from noise–added data, with an overlapping interval of 3 mfp. The result in (c) and (d) are both obtained after 14 time windows. Volume of medium considered in the reconstructed image is the same as in Fig. 16.

18. Reconstruction of the medium containing the "T" absorber, from simulated CW data. Left column: X-Z cross–section, right column: Y-Z cross–section. Within each column: (a) original medium, (b) reconstruction by CGD, 200 iterations; (c) coarse-grid solution by multi–grid (using CGD); and (d) fine–grid solution by multi–grid (using POCS, 200 iterations).

19. Reconstruction of the medium containing the "T" absorber, from calculated and simulated TR data. Left column: Y-Z cross–section, right column: X-Z cross–section. (a) Original; (b) Reconstruction from calculated data after 16 time windows of width 1 mfp (*i.e.*, 0–1, 1–2, ..., 15–16), without overlapping; the source and detector configuration is the same as in Fig. 13. (c) Reconstruction from simulated data after 9 time windows of width 2 mfp (*i.e.*, 0–2, 2–4, ..., 16–18), with an overlapping interval of 2 mfp; the source and detector configurations are as shown in Figs. 9(b) and 10(a), respectively. (d) The same as in (d), except that the detector configuration is as shown in Fig. 10(b).

20. Results for the three–layer medium, assuming planar structure is known: (a) original medium; (b–c) reconstructions from calculated CW data using the three-layer (b) and half-space (c) weights; (d–e) reconstructions from simulated CW data using three-layer (d) and half-space (e) weights; (f) reconstruction from simulated TR data using half–space weights; the result is obtained after 10 time windows of width 2 mfp (*i.e.*, 0–2, 2–4, ..., 18–20), with an overlapping interval of 2 mfp. For the CW cases, the source and detector configurations are as shown in Figs. 9(a) and 9(b). For the TR case, the sources are located uniformly over a 17×17 mfp^2 region in a grid–like manner separated by 1 mfp in each direction; the detector configuration for each source is shown in Fig. 10(c).

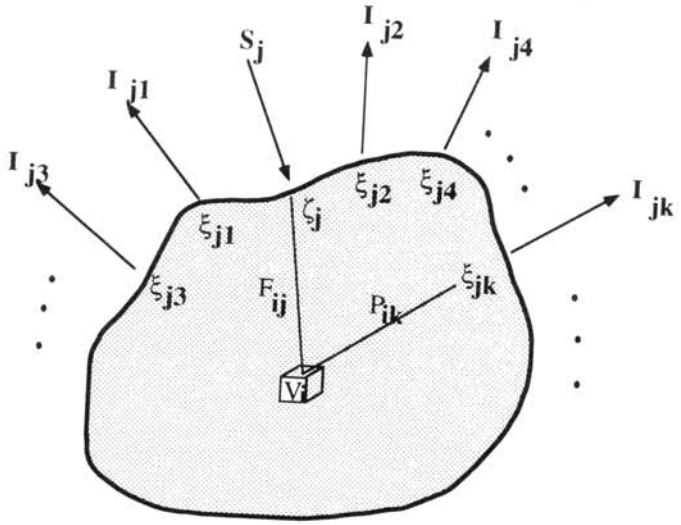

Figure 1.

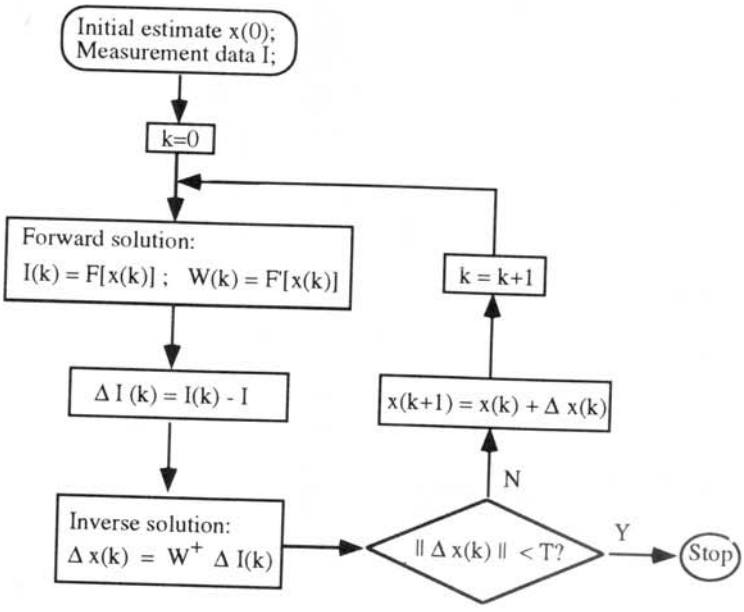

Figure 2.

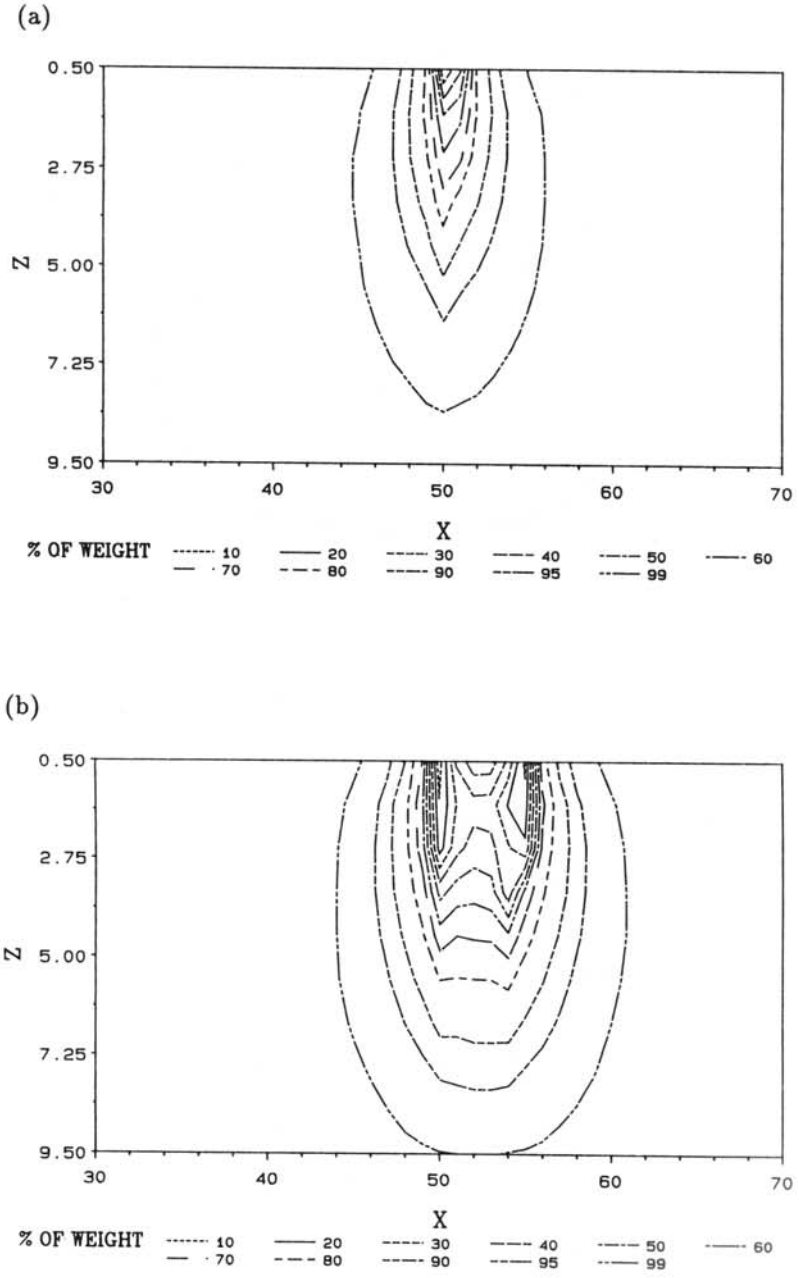
(a)
0.50
2.75
5.00
7.25
9.50
Z
30
40
50
60
70
X
% OF WEIGHT
10
20
30
40
50
60
70
80
90
95
99
(b)
0.50
2.75
5.00
7.25
9.50
Z
30
40
50
60
70
X
% OF WEIGHT
10
20
30
40
50
60
70
80
90
95
99

Figure 3.

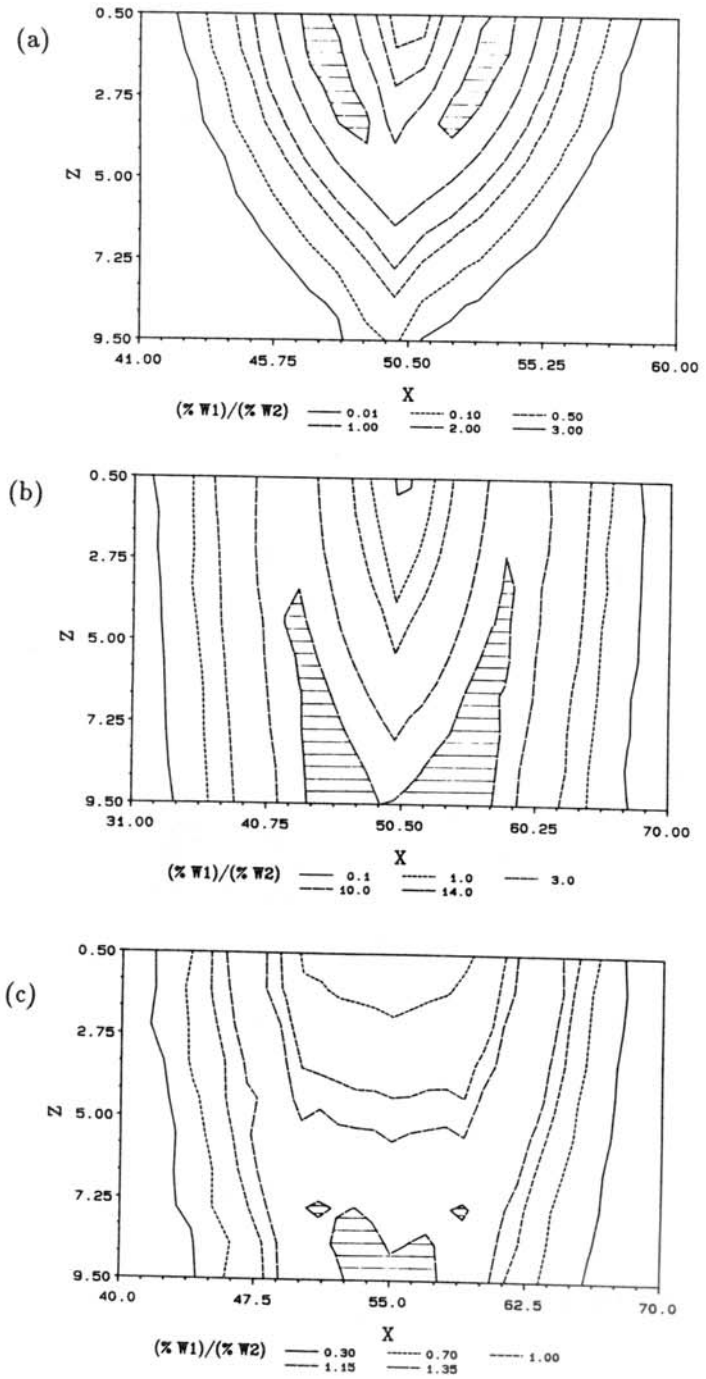
(a)
0.50
2.75
5.00
7.25
9.50
Z
41.00
45.75
50.50
55.25
60.00
X
(% W1)/(% W2)
0.01
0.10
0.50
1.00
2.00
3.00
(b)
0.50
2.75
5.00
7.25
9.50
Z
31.00
40.75
50.50
60.25
70.00
X
(% W1)/(% W2)
0.1
1.0
3.0
10.0
14.0
(c)
0.50
2.75
5.00
7.25
9.50
Z
40.0
47.5
55.0
62.5
70.0
X
(% W1)/(% W2)
0.30
0.70
1.00
1.15
1.35

Figure 5.

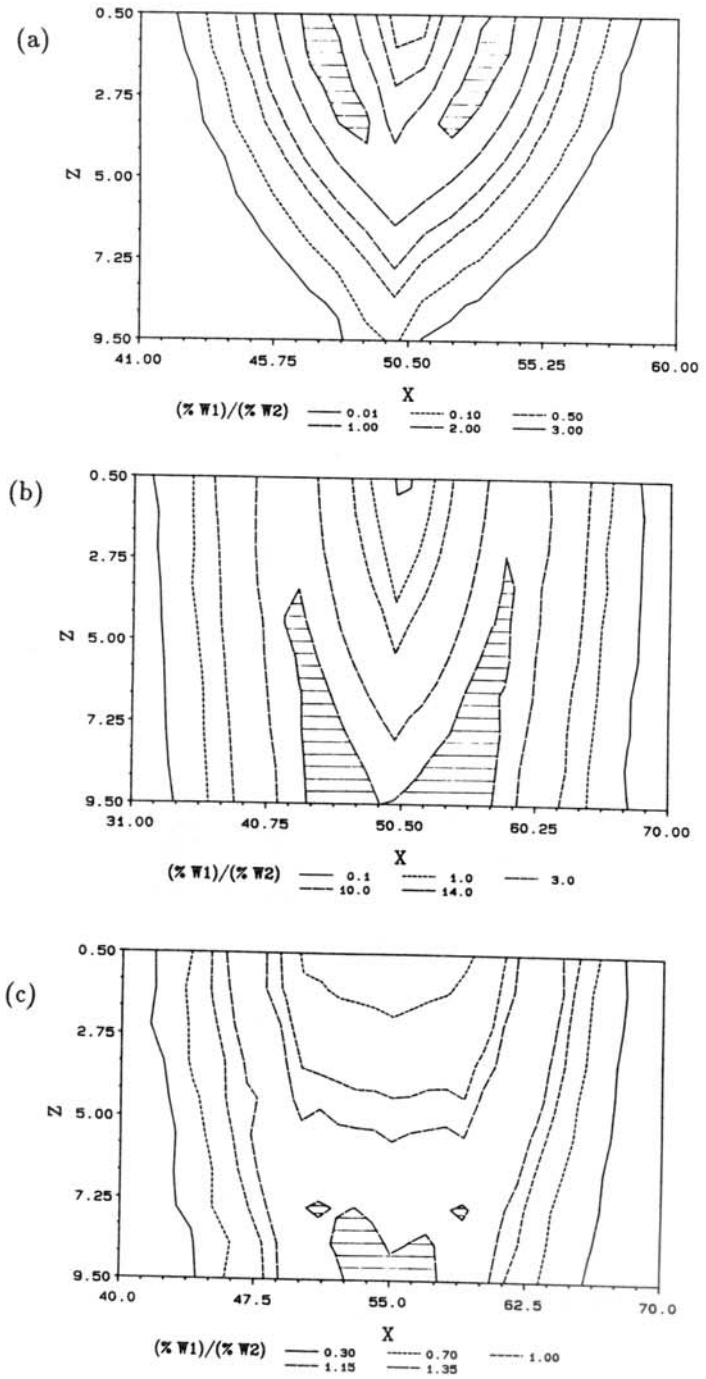
(a)
0.50
2.75
5.00
7.25
9.50
Z
41.00
45.75
50.50
55.25
60.00
X
(% W1)/(% W2)
0.01
0.10
0.50
1.00
2.00
3.00
(b)
0.50
2.75
5.00
7.25
9.50
Z
31.00
40.75
50.50
60.25
70.00
X
(% W1)/(% W2)
0.1
1.0
3.0
10.0
14.0
(c)
0.50
2.75
5.00
7.25
9.50
Z
40.0
47.5
55.0
62.5
70.0
X
(% W1)/(% W2)
0.30
0.70
1.00
1.15
1.35

Figure 5.

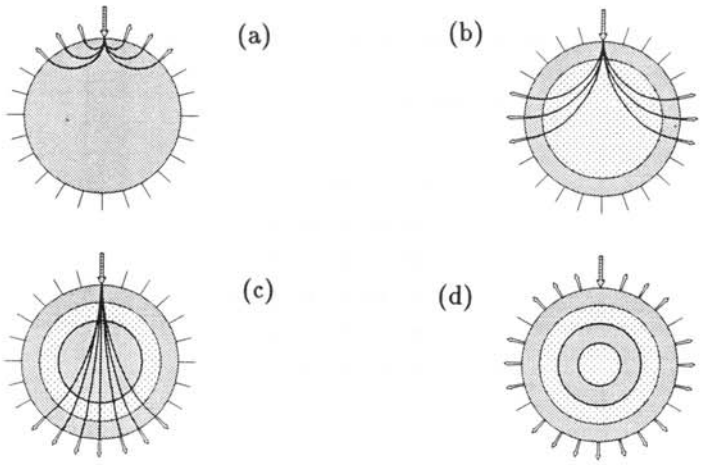

Figure 6.

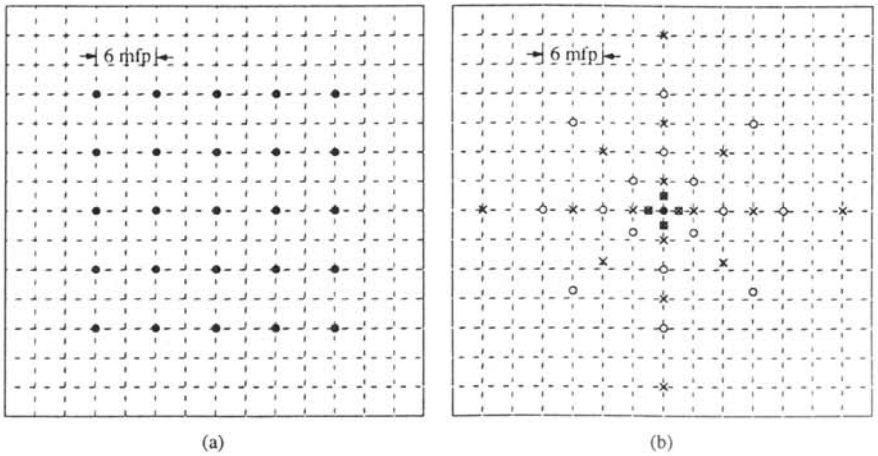

Figure 7.

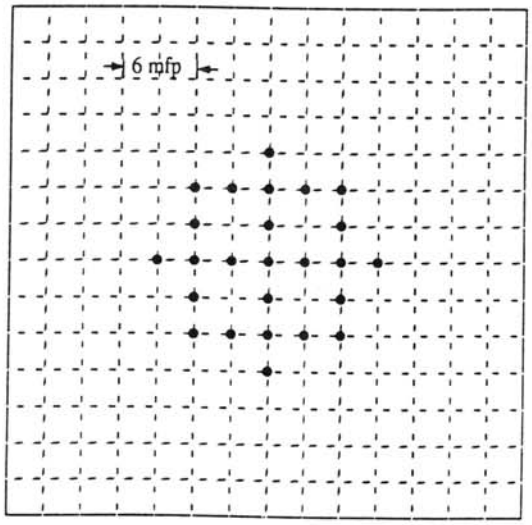

Figure 8

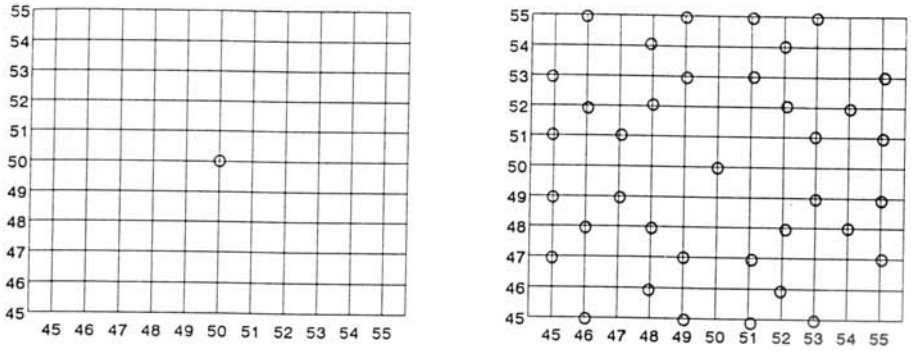

Figure 9

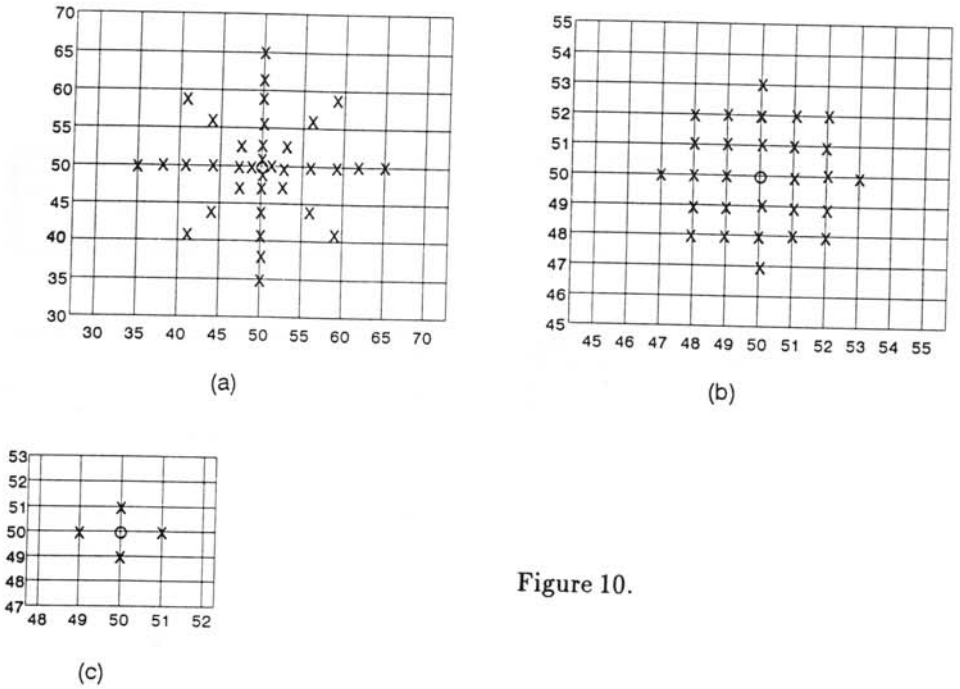

Figure 10.

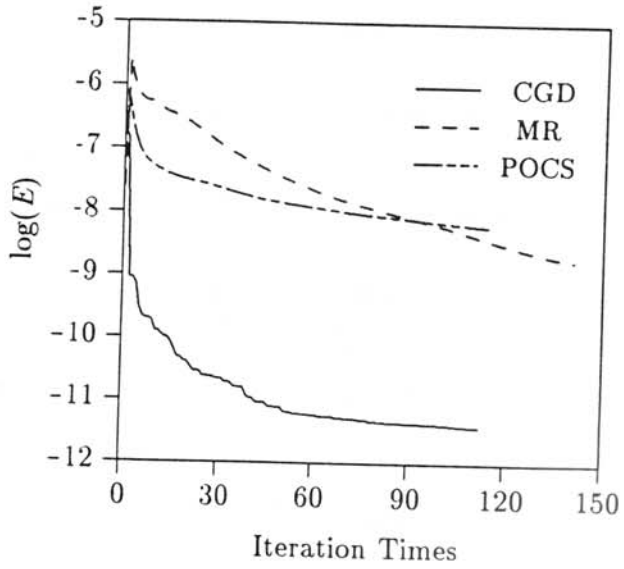

Figure 11.

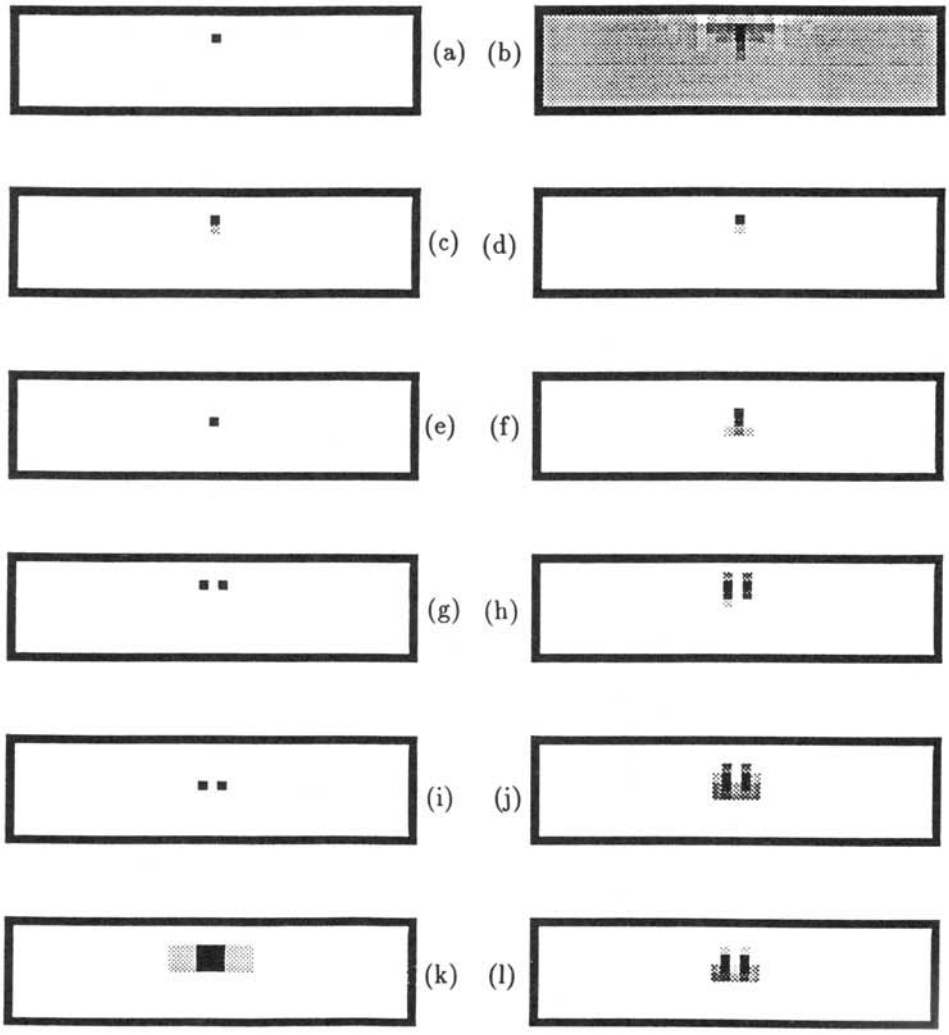

Figure 12.

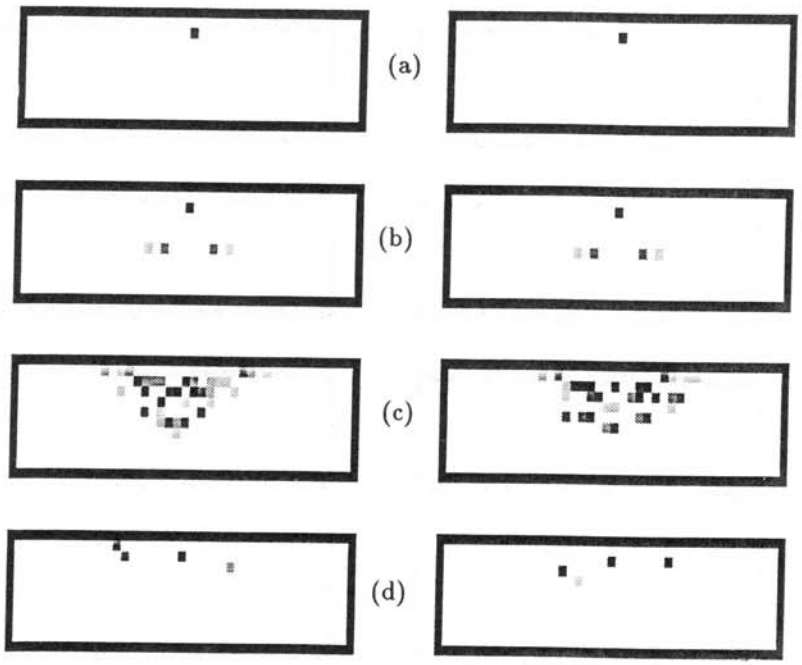

Figure 13.

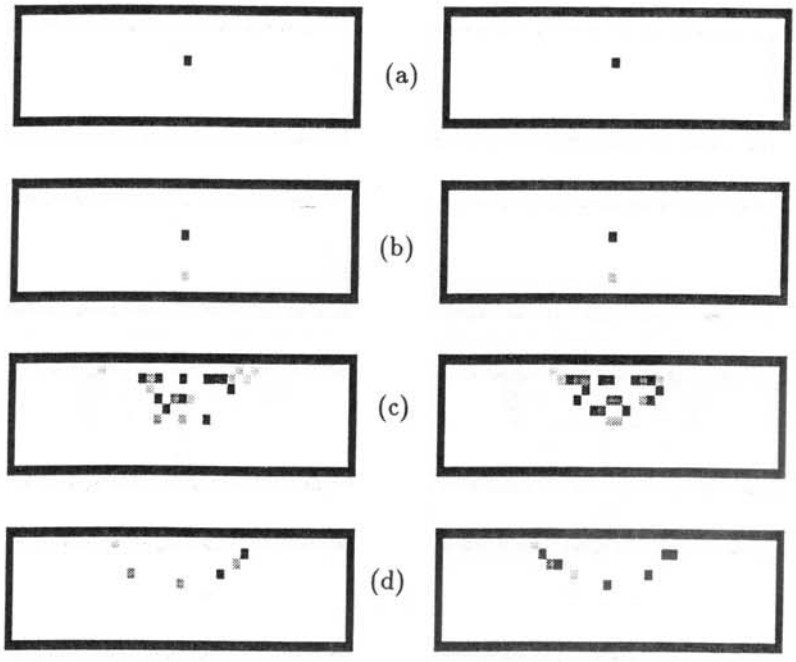

Figure 14.

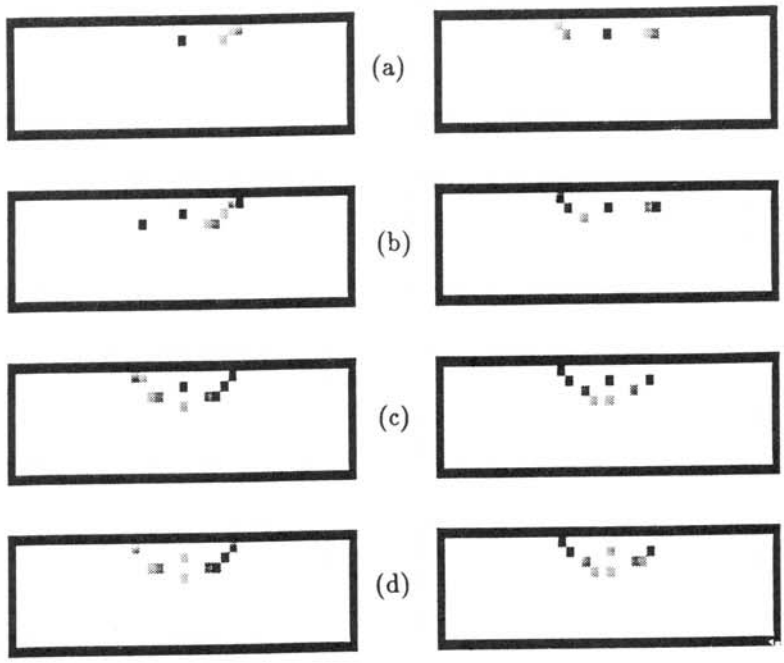

Figure 15.

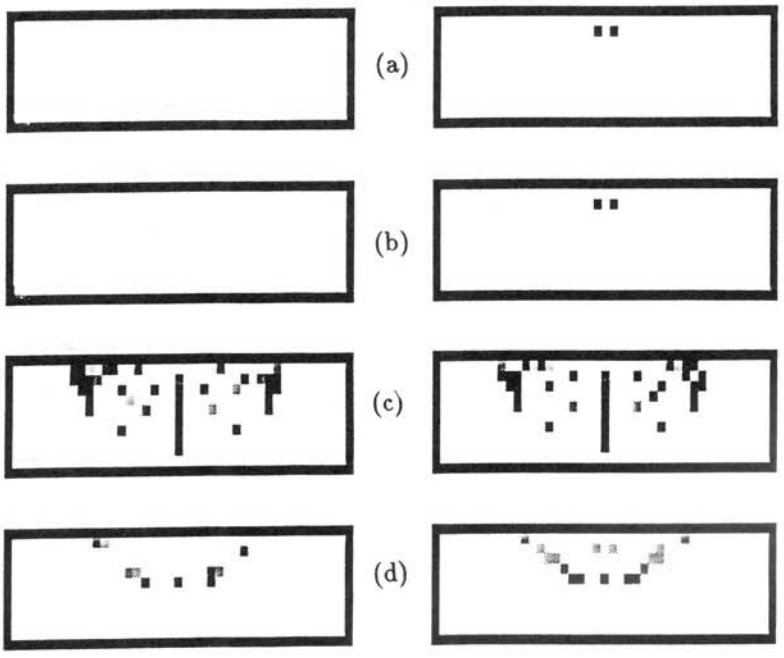

Figure 16.

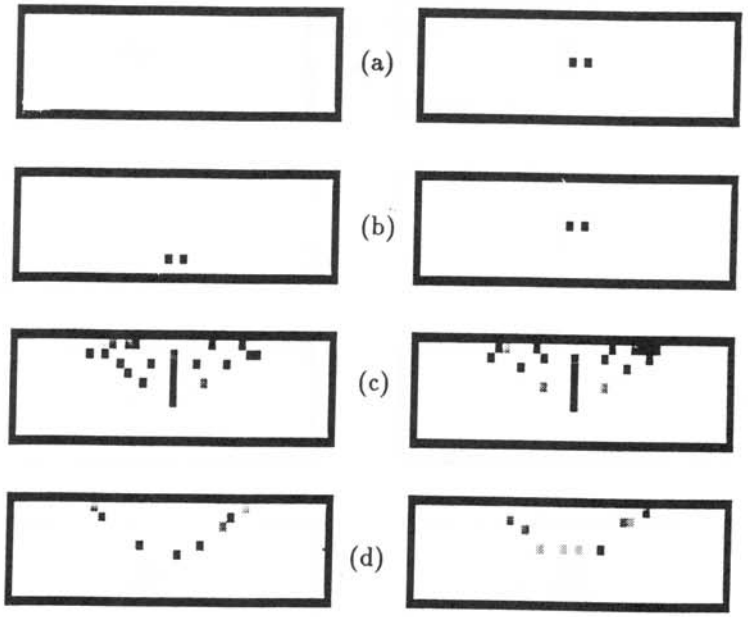

Figure 17.

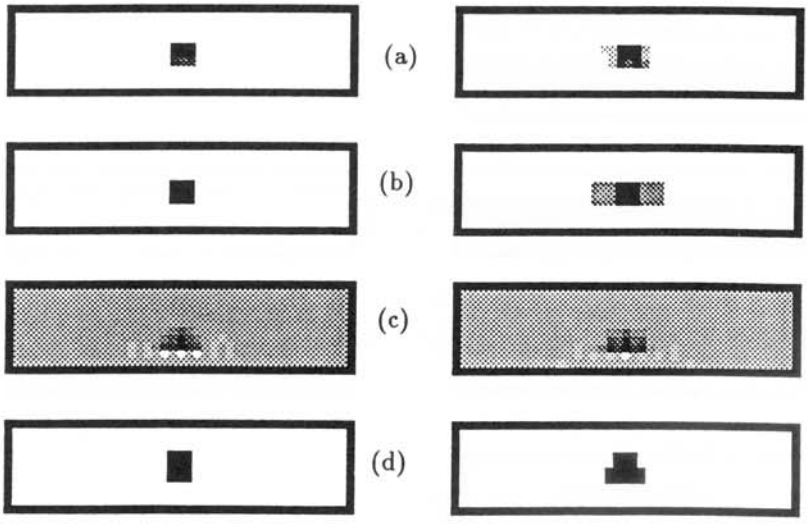

Figure 18.

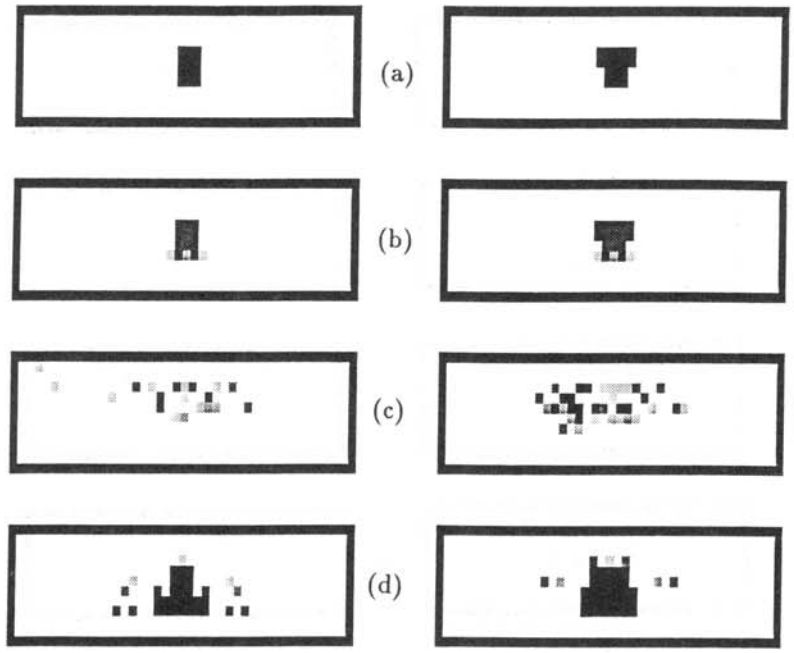

Figure 19.

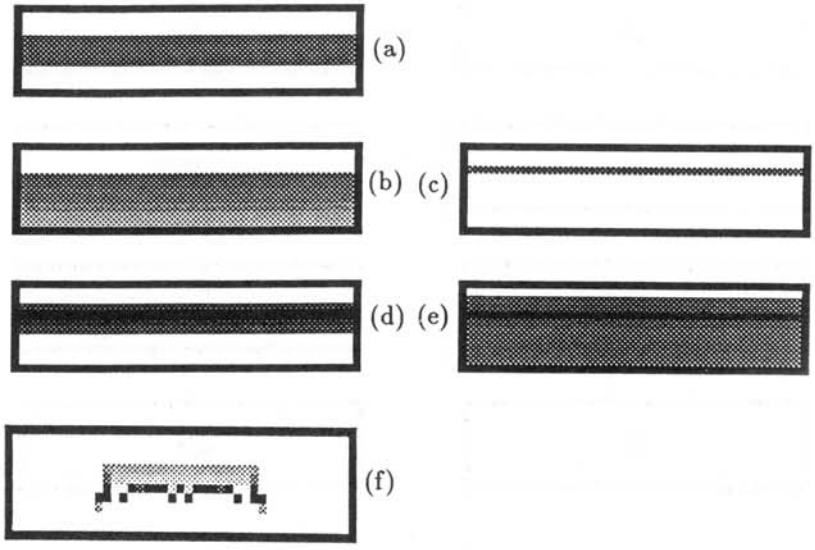

Figure 20.