Imaging in Diffusing Media with a Neural Net Formulation: A Problem in Large Scale Computation

F.H. Schlereth, J.M. Fossaceca, A.D. Keckler,
Department of Electrical and Computer Engineering, 2-189 CST,
Syracuse University, Syracuse, NY 13244

R. L. Barbour, Department of Pathology and Biophysics,
SUNY Health Science Center, Brooklyn, NY 11203

Abstract

A neural network scheme is described, based on a modified backpropagation algorithm for the recovery of images of the interior of objects which diffuse radiation. The method is capable of considering isotropic and any degree of forward directed scatter. The computational requirements of the method are significantly greater than for coherent (i.e. straight-line) based imaging schemes and for this reason we have developed a machine architecture and machine, Kilonode. Results are shown for 2-D simulations of media having simple structure and can readily be extended to 3-D. Training vectors for the neural network are derived from time-independent and time-resolved signals.

1.0 Introduction

Evaluation of the problem of imaging in diffusing media is particularly challenging both in regard to defining a suitable reconstruction algorithm and in terms of identifying an appropriate computing environment for efficient processing. In this paper we have examined both issues and, in particular, describe results of an algebraic technique for imaging the interior of objects which diffuse penetrating radiation using a new multicomputer environment. The problem of imaging in diffusing media differs significantly from other imaging schemes which rely on the detection of coherent signals in that we assume only probabilistic knowledge of the path of the radiation and minimal knowledge of the absorption profile of the medium. Because of this added uncertainty the size of the computation is unavoidably much greater than for coherent-based schemes and is an example of an ultra-large scale computing problem.

Two important issues which arise when evaluating problems of this size are numerical precision and the overall computing efficiency. Both must be considered simultaneously and involve trade-offs. Our approach has been that numerical precision is paramount. In addition we have developed an operating system which evaluates algorithms in a manner which minimize the communication overhead. Because of the efficiency, the system we have developed can allow for the configuration of a large multinode unit. A four-node system is currently available and operates at 320 Mflops. A 100 node system runs at 6400 Mflops, which is an efficiency of 80%. These efficiencies make it feasible and, in fact, desirable to employ numerically stable algorithms which, though computationally more intensive than others, can be evaluated in a time frame which readily permits the development of strategies for the solution of large scale computing problems.

The algorithms we propose are more closely related to algebraic reconstruction algorithms such as ART, SIRT and SART [1] than to algorithms based on the Born and Rytov approximations such as used for tomographic imaging with diffracting sources. In our model we assume that an NIR laser is used to provide the input radiation and suitable detectors are positioned to measure both transmitted and backscatter signals. The present work considers a simple Markov process for the way in which the energy propagates in the medium. It should be noted though, that the reconstruction technique we propose can use any model. Current simulations are in 2-D but are easily extended to 3-D. These studies were motivated by recent reports from our group [2-4] and others, [5-7] which showed promising results for imaging in dense scattering media given only diffusely scattered signals. A preliminary description of the findings has recently been reported. [8]

0-8194-0787-9/92/$4.00

2.0 Modeling of the Forward Problem

In more traditional algebraic methods involving linear tomographic schemes, a matrix, w, is assumed, where $w_{i,j}$ represents the fractional area of the jth image cell intercepted by the ith ray. The equation which is solved is w]*f] = p], where f] represents the absorption of each of the cells and p] the detector readings. It is assumed that [w] is known. Typically the dimension of p] is M and the dimension of f] is N, where M<N in most cases of practical interest. Standard methods are available for the solution of such equations such as least squares, linear programming or the Kaczmarz method[1].

We propose a different model of the physical problem. We assume that the radiation entering the medium travels through the medium according to some well defined probabilistic model which can be simulated using relaxation techniques. An example of a relaxation technique is the solution of Laplace's equation using the standard five point grid template. That computational model is based on a simple discrete approximation to the partial differential equation.

$$4 * f_{i,j}(n+1) = f_{i-1,j}(n) + f_{i,j+1}(n) + f_{i,j-1}(n) + f_{i+1,j}(n).$$

(n) refers to the iteration number.

For our case (2-D) we assume that each cell gives up all of its energy to its eight nearest neighbors at each step of the relaxation iteration, and the portion it loses is modified by an absorption coefficient. For the homogeneous case, the absorption is uniform throughout the medium. The relaxation equation governing this case is as follows:

$$8 * f'_{i,j}(n+1) = f_{i-1,j+1}(n) + f_{i,j+1}(n) + f_{i-1,j}(n) + f_{i+1,j+1}(n) + f_{i+1,j}(n) + f_{i-1,j-1}(n) + f_{i,j-1}(n) + f_{i+1,j-1}(n),$$

and

$$f_{i,j}(n+1) = w_{ij}\, f'_{i,j}(n+1)$$

where w_{ij} is the absorption coefficient for the i,j^{th} cell.

We also consider a directed flow discrete model which accounts for the fact that the energy tends to flow in the same direction in which it is propagating. The extreme case is x-radiation which is assumed to follow a straight line path through the medium. In describing this formulation we will assume that energy is lost to four (n,e,s,w) neighbors. The case for eight neighbors follows easily.

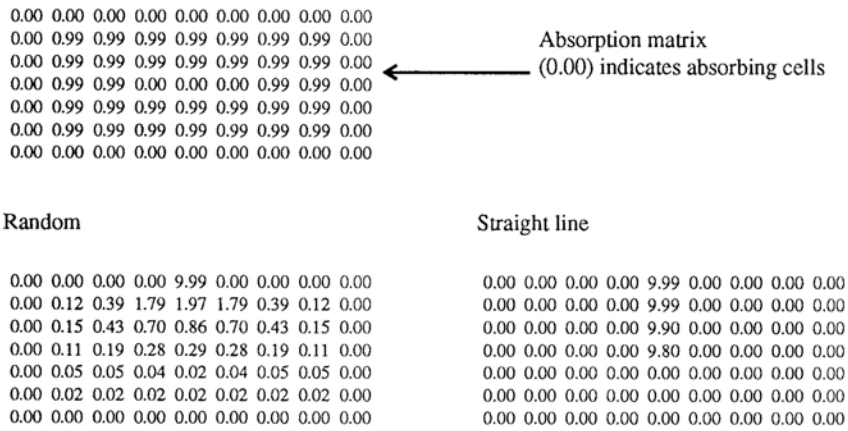
Assume that the flux in each cell is a vector of four components. $F(i,j) = [f^n, f^e, f^s, f^w]^T$. Then the relaxation equations for the directed flow take the following form.

$$4 * F_{i,j}(n+1) = [T] * [[f^n_{i,j-1}(n), f^e_{i+1}(n)_j, f^s_{i,j+1}(n), f^w_{i-1,j}(n)]^T$$

[T] is a transmission matrix defining the path of the radiation through the medium. For example, if [T] is the identity matrix, then the radiation is constrained to flow along the path in which it was initially directed. Using the identity matrix, the above equation expresses the idea that northerly radiation out of (i,j) is due to northerly radiation out of the cell to its south. An example of the use of directed equations is shown in Figure 1. In Figure 1 note the leakage around the absorber in the random case, and the lack of spreading when we use the identity matrix for [T].

Thus we propose a discrete model for the transmission of radiation, not necessarily based on a stochastic differential equation, but rather based directly on a discrete physical model of the medium. The complexity of this model has an effect on computation time, but not on the basic methodology. In generating this model we follow a suggestion in references 5 and 6. Random effects are included by assuming that each of the entries in matrix [T] has a degree of randomness. Mathematical techniques for dealing with stochastic differential equations are discussed by Gard [9]. These relaxation iterations become a basic part of the image estimation method, so it is a requirement that suitable computational facilities be available. In Figure 1 energy is introduced at the top (9.99). The edges are modeled as perfect absorbers, (0.00) because once energy leaves the me-

Figure 1 Flux Transmission - Random and Straight Line for a simple absorber.



dium it is lost. Energy introduced into the medium is by proper selection of the boundary conditions. We assume that energy can be introduced on all six faces of a rectangular body, or a suitable set of faces on an irregular body.

In the following we describe a neural network formulation of our image estimation algorithm, and the computing system we have designed to meet the computational demands. The algorithm we propose is closely related to the system identification problem, which is commonly used in the design of control systems. The reason for using the neural network approach is that it is a technique well suited for estimating absorption in the medium that does not assume a straight line path for the radiation. However as will be seen in the following section, the neural network computations require a means for computing an estimate of the output radiation pattern which is computationally efficient, and which at the same time is a good model of the actual transport phenomena in the medium in question. The relaxation equations shown above meet both requirements. Computations are minimal and the model has enough structure to account for directivity, randomness, reflection and absorption.

Of course we make no claims that these computing models accurately model the physical phenomena in a particular medium. However we do feel that we have a sufficiently rich formulation which could be modified as needed to meet particular transmission phenomena. We hope that others will suggest suitable modifications. The models we employ need not be constrained by computer limitations.

3.0 NEURAL NET FORMULATION

In the following some knowledge of neural networks is assumed. Two references are cited which present an excellent introduction. Wasserman [10] is the simpler, and Pao[11] goes into the mathematical details. We use the LMS algorithm of Widrow[12].

In the case of the physical model we assume that it is possible to introduce radiation from all sides of an object. It seems clear, at this stage of our work, that it will be more difficult to locate an object, such as the one shown in Figure 3, on the basis of backscatter or transmission from a single surface alone. However it is by no means impossible as has been shown by using either time-independent [2-4,13] or time-gated [14] signals. We believe that by locating various sources as shown, we can obtain a better reconstructed image. In addition, the degree to which we assume the radiation is flowing in a directed path will greatly

influence the quality of the reconstruction. Figure 2 shows a block diagram of the neural network we have constructed to perform the image estimation algorithm. Figure 3 shows the physical model. In Figure 2 the inputs x, represent input energy into the medium. The $w_{i,j}$'s represent the absorption of each of the cells. Using these values for absorption we compute the flux, $f'_{i,j}$ in each cell, and in particular the cells at the surface which produce the observables, O_k. These are then fed through the activation function, F, a gain constant η, and the derivative of the activation function F'.

Figure 2 Neural Network Block Diagram

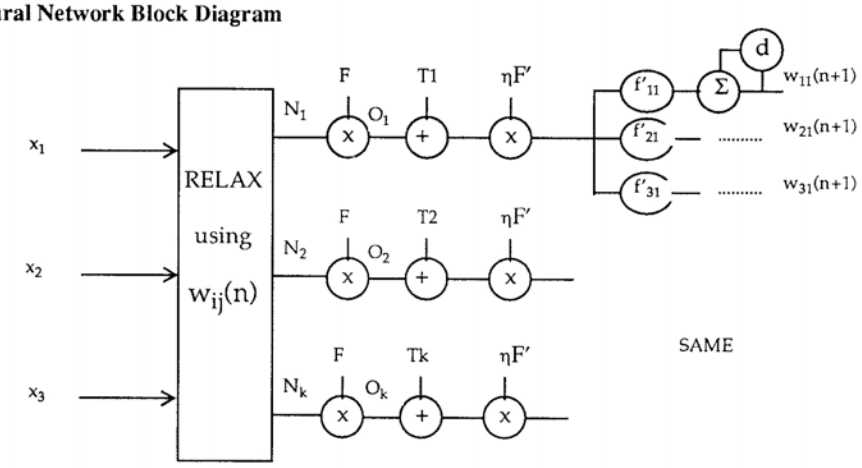
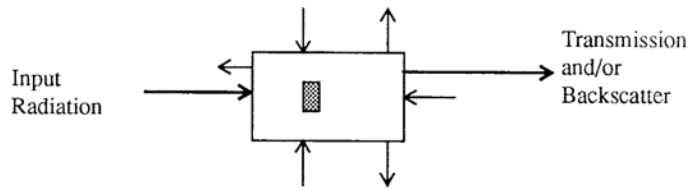


Figure 3 Physical Model



The output of this part of the circuit is fed into a set of loops which update the values for the $w_{i,j}$. The values in the block prior to the summation in the update loop are equal to the flux in the cell, $f'_{i,j}$, prior to multiplication by the absorption coefficient. T_k are the training vectors, which would represent measurements in an actual application, but in our case represent simulated test cases. The challenge is to derive an estimate for the $w_{i,j}$'s, which is, in fact, an estimate of the true image.

In this computational model the relaxation operation is forming an estimate of the radiation pattern at the surfaces of the material under consideration, by calculating the flux at all of the interior points according to the equations in Section 2, using the most recent values for the w_{ij}. This pattern is compared with measured data to form a new estimate for the w_{ij}. The control loops implement a steepest gradient decent algorithm which is described in detail in Pao[11]. The quantity, E, to be minimized is

$$E = \Sigma (T_k - O_k)^2.$$

Incremental changes, Δw_{kj}, are proportional to $-\delta E/\delta w_{kj}$.

It should be noted that there will be many "correct" solutions in the sense that the w_{ij} satisfying a particular set of data are not unique. Therefore there needs to be another set of inputs or a "cost" function to limit the solution to a desired set. There

is begun, the nodes communicate among themselves to maintain synchronism. This style of computing is essential for systems with large numbers of computing nodes. It is the key that makes our system scalable and extensible.

Other components such as hardware, operating system and software are almost a secondary consideration, provided they meet guideline of simplicity and speed. It is, however, important that each of the nodes has a large local memory because this minimizes communication overhead. The Kilonode environment could exist on any number of platforms; Sun, Intel, Ncube, Wavetracer. We have chosen CSPI because it is simple, fast and cost effective.

The performance of Kilonode on a standard benchmark program (1000x1000 matrix inverse) indicates that we achieve an efficiency of almost 50% for a 1000 node machine. Our four node system solves a dense 1000x1000 system of equations in about one minute, using double precision floating point. The relaxation problem is much simpler, requiring much less communication and for this problem Kilonode should have even better performance.

7.0 Example

In our present work we have used a linear activation function in order to gain a better understanding of the stability of the basic neural network control loop. The stability is a function of η. With a suitable choice of η stable operation of the net is assured.

The results of image estimation with the use of backscatter alone indicates that while we are able to discern the presence of an absorber, the degree to which we can localize the absorber is very poor. From this we have concluded that it is necessary to introduce additional information, such as data from other sources, either on the same face or other faces of the object, or from a priori knowledge such as might be obtained from techniques suggested previously by Barbour. [15, 16]

The ability to discern the presence of an absorber based on surface measurements depends on the signal-to-noise ratio and on the sensitivity of the surface data to the position of the absorbing object. At present we have not performed measurements to determine the noise levels, although the experimental setup is in place.

We have however run many 2-D simulations with absorbers at different depths, and results are encouraging. As a typical example, an absorber at a depth of 10 cells, with a length of 5 cells in a medium whose width is 31 cells will produce differences in reading, as compared with the homogeneous case, of about 5%. This, we feel, is very respectable and should produce real data with adequate signal-to-noise ratio.

The introduction of data based on information derived form other sources requires the use of data preprocessing or, in neural network terminology, functional links [10]. For example when we introduce radiation on opposites sides of an object it is natural to use the difference of the backscatter response form each of the faces as primary input to the neural network. In principal the neural network should eventually learn this transformation, but we feel this is an unnecessary computational burden. Similar considerations apply to the use of data introduced into other faces of the object.

Figure 5 shows the results of a 2-D image obtained from backscatter using a single source entering from the top, bottom and each of the sides. In performing the iterations we take advantage of a priori knowledge of the location of the absorber by desensitizing the layers near the surface during the initial learning phases. In this way the network attempts a correction at internal layers first. After this initial phase we allow all of the layers to take part in the learning process.

In Figure 5, the first 1000 iterations were performed with the outer layers desensitized. After that 500 additional iterations were performed allowing all of the cells to adapt. It is important to note that early learning was not "unlearned", although there was some blurring of the edges.

Figures 6 through 11 show a series of 3-D plots which depict the passage of radiation into the medium as a function of the scattering. The "z" direction corresponds to energy magnitude.The amount of scattering is given by the diagram in the left hand corner of each of the figures. The medium contains an absorber in the middle, which is causing the radiation to go to zero. Note that even for the case of uniform scattering, there is still a significant amount of input radiation impinging directly onto the absorber.

Figure 5 Estimated Absorption Matrix for 1000 Iterations

ABSORPTION MATRIX	ESTIMATED MATRIX
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.990.990.990.991.989.991.990.990.990.990.990.990
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.988.988.988.989.991.990.988.988.988.990.990.990
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.988.977.973.969.967.969.973.977.987.990.990.990
.990.990.990.990.990.990.000.000.000.990.990.990.990.990	.990.990.990.987.978.942.934.931.934.942.978.987.990.990.990
.990.990.990.990.990.990.000.000.000.990.990.990.990.990	.989.990.990.987.979.953.831.829.834.953.979.987.990.990.989
.990.990.990.990.990.990.000.000.000.990.990.990.990.990	.991.990.990.987.979.957.854.856.858.958.979.987.990.990.991
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.989.990.990.987.979.953.843.843.847.954.979.987.990.990.989
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.987.978.945.940.938.940.945.977.987.990.990.990
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.988.977.974.969.967.968.972.976.988.990.990.990
.990.990.990.990.990.990.990.990.990.990.990.990.990.990	.990.990.990.988.987.987.988.988.988.988.989.989.990.990.990
	.990.990.990.990.989.989.989.991.991.990.990.990.990.990.990

The numbers in this Table are absorption coefficients.
1.0 is no absorption. 0.0 is full absorption

Figures 12 through 17 shows the results of the neural reconstruction algorithm as a function of the amount of scattering. Note that the results become worse as the amount of scattering increases, which is to be expected. However it is very important to note that the result in Figure 17 is not incorrect as far as the neural network is concerned. In fact there are many solutions to the inverse, given the set of boundary conditions. The solution found depends on the path of the search.

Thus we are led to the idea of directing the search of the neural network by starting the search at a place of high absorption. In a real case this would require additional knowledge which might be gained in any number of ways such as time resolved measurements, or other. If the neural network search is directed in this way, then we obtain the result shown in Figure 18, which is for uniform scattering. This is quite a remarkable result. In Figure 19, we start with the result given in Figure 18, but allow unconstrained search in the neural network. This indicates that the result is stable in the sense that once the network has found a solution, it will not wander off to another solution, once unconstrained operation is invoked. Figures 20 -22 show the case for a more complicated absorption pattern. Our data show similar results as for the one absorber case.

8.0 Conclusions

From the above discussion it should be clear that imaging in diffusing media is bound to be an enormous computing problem, one that will require the resources of the parallel machines which are currently being designed using powerful integrated circuits like the i860 from Intel. It is also clear that use of large parallel machine in itself is not sufficient to meet the requirements. New algorithms have to be devised to take advantage of the hardware, and especially the parallelism. With regard to effective utilization of parallelism, the key requirement is to devise algorithms which have a minimum of communication requirements among processors.

Our approach to this has been to rethink fundamental reconstruction algorithms and concentrate on schemes which will permit the estimation of the medium, with minimal assumptions about the path of the radiation. Neural network schemes offer that possibility. At present we solve the forward problem using relaxation methods, but in the near future we will begin to investigate direct methods using matrix inversion. In either case the computing requirements are large. A typical matrix size for an interesting problem would be tri-diagonal with 100,000 rows and columns. However with the computing machines which are becoming available, this size of problems will be practical for systems in a clinical setting.

New algorithms are based on the idea that communication costs are the dominant factor in large scale computing, and that numerical effects will be critical because of the large numbers of operations involved. It is very easy for limit cycles due to limited word length to dominate the results. The solution to these kinds of problems is two-fold. One is proper choice of algorithm, (Householder's vs. Gaussian elimination, e.g.). The other is the neural network formulation which allows us to intro-

duce other data (e.g., time-gated, frequency domain). We can either employ this to localize the neural network search space and to provide additional information to aid in the preprocessing of data, prior to using the data in the neural network computations. The use of the directed relaxation algorithm allows us to accurately model the actual passage of radiation through the real medium, which greatly decreases the learning time of the network.

Our next step is to compare measured and computed results. The major purpose will be to learn to make effective use of directed radiation in creating images. Using our experimental and computational capability we will be able to experimentally determine the scattering properties of various materials. This knowledge, in conjunction with other measurements such as time resolved methods, will provide us with the clues we need for effective neural network performance.

9.0 References

1. Kak, A.C., Slaney, M., "Principles of Computerized Tomographic Imaging", IEEE Press, 1987.
2. Barbour, R.L., Graber, H., Aronson, R. and Lubowsky, J. Imaging of subsurface regions of random media by remote sensing. SPIE, Vol. 1431, pp. 192-203, 1991.
3. Aronson, R., Barbour, R.L., Lubowsky, J., and Graber, H.L., Application of Transport Theory to NIR Medical Imaging in, Modern Mathematical Models in Transport Theory; Advances and Applications, Vol. 51, p 64-75, Birkhauser Press, 1991
4. Barbour, R.L., Graber, H., Aronson, R., and Lubowsky, J. Model for 3-D Optical Imaging of Tissues. In Proceedings of 10th Annual International Geoscience and Remote Sensing Symposium (IGARSS) Vol. 2, pp. 1395-1399, 1990.
5. Grunbaum, F.A., Kohn, P., Latham, G.A., Singer, R.J., and Zubelli, J.P. Diffuse Tomography. SPIE 1431, Time Resolved Spectroscopy and Imaging of Tissues, pp 232-238, 1991.
6. Grunbaum, F.A. Tomography with Diffusion. In, Inverse Methods in Action, P.C. Sabatier, Ed. Springer Verlag, NY pp 16-21, 1990.
7. Arridge, S.R., Van der Zee, P., Cope, M., and Delpy, D.T. Reconstruction Methods for IR Absorption Imaging. SPIE, 1431, pp 204-215, 1991.
8. Schlereth, F. H., Fossaceca, J.M., Keckler, A.D. and Barbour, R.L., "Multicomputer-based Neural Networks for Imaging in Random Media", IEEE Nuclear Science Symposium, Santa Fe, NM, Nov. 1991. In press.
9. Gard, Thomas C., Introduction to Stochastic Differential Equations", Marcel Dekker, 1988.
10. Wasserman, P.D. Neural Computing, Theory and Practice, Van Nostrand, 1989.
11. Pao, Yoh-Han, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley, 1989.
12. Widrow, B., Stearns, S., Adaptive Signal Processing, Prentice Hall 1985.
13. Wang, Y., Chang, J.H., Aronson, R., Barbour, R.L., Graber, H.L. and Lubowsky, J., "Imaging of Scattering Media by Diffusion Tomography: An Iterative Perturbation Approach", SPIE vol. 1641, accompanying paper in these proceedings, 1992.
14. Barbour, R.L., Graber, H.L., Lubowsky, J., Aronson, R. Das, B.B., Yoo, K.M. and Alfano, R.R. "Imaging of Diffusing Media by a Progressive Iterative Backprojection Method Using Time-Domain Data" SPIE vol. 1641, accompanying paper in these proceedings, 1992.
15. Barbour, R.L., Graber, H.L., Lubowsky, J., "Determination of Macroscopic Optical Properties of Multilayer Random Media by Remote Sensing" SPIE vol. 1431, Time Resolved Spectroscopy and Imaging of Tissues. p52-62, 1991.
16. Graber, H.L.,Barbour, R.L., Lubowsky, J., Aronson, R. Das, B.B., Yoo, K.M. and Alfano, R.R, "Evaluation of Steady-State, Time and Frequency -Domain Data for the Problem of Optical Diffusion Tomography", SPIE vol 1641, accompanying paper in these proceedings, 1992.

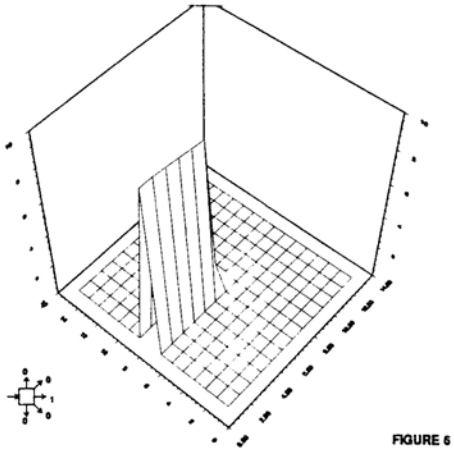


FIGURE 6

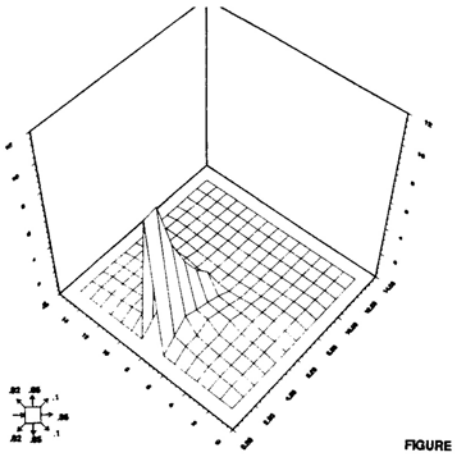


FIGURE 9

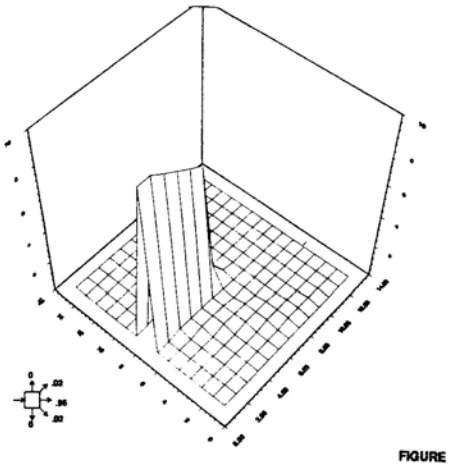


FIGURE 7

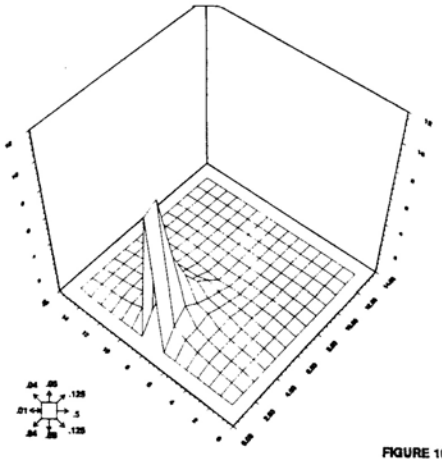


FIGURE 10

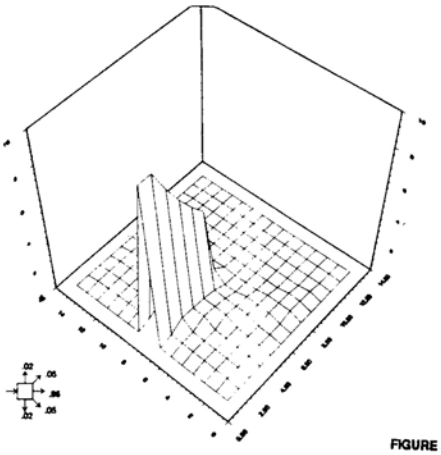
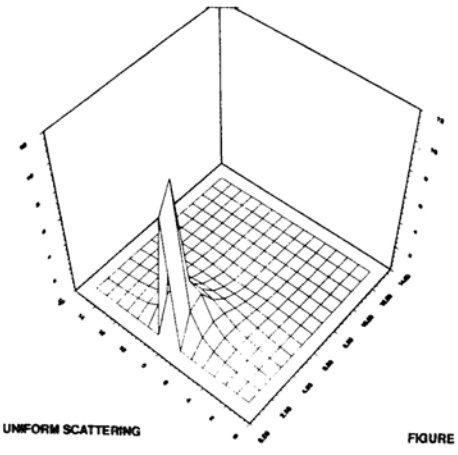
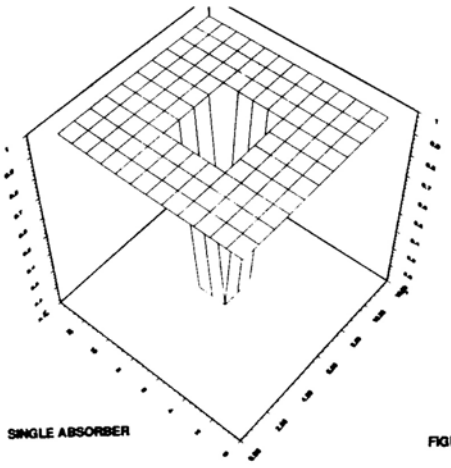


FIGURE 8



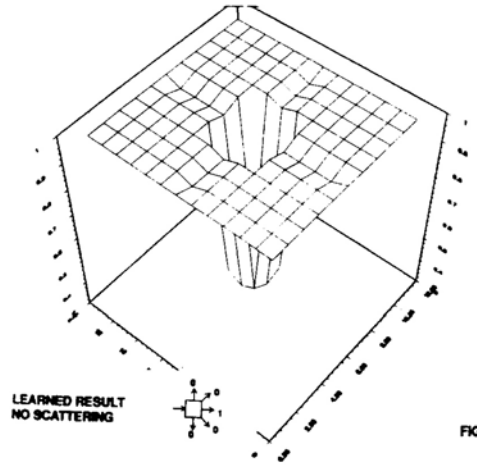
UNIFORM SCATTERING

FIGURE 11



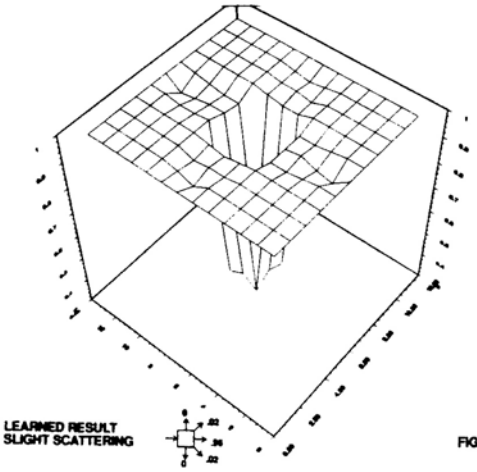
SINGLE ABSORBER

FIGURE 12



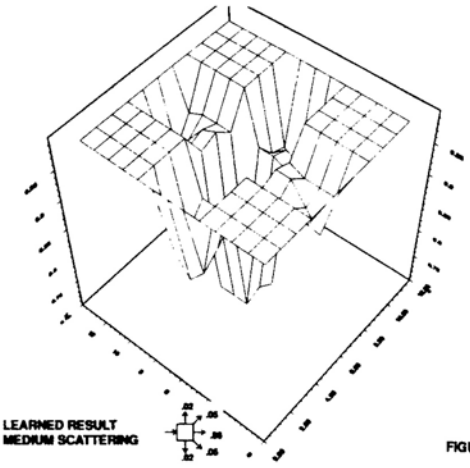
LEARNED RESULT
NO SCATTERING

FIGURE 13



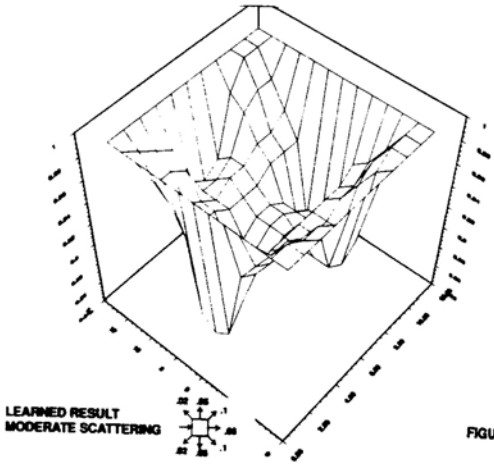
LEARNED RESULT
SLIGHT SCATTERING

FIGURE 14



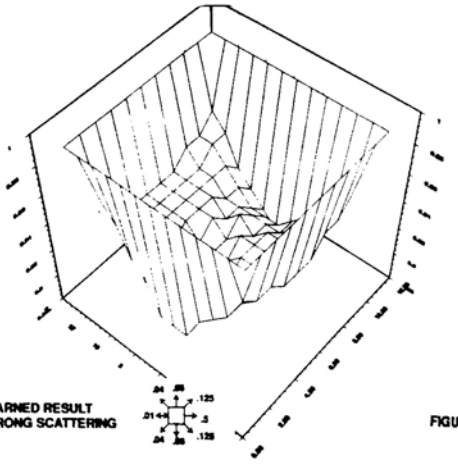
LEARNED RESULT
MEDIUM SCATTERING

FIGURE 15



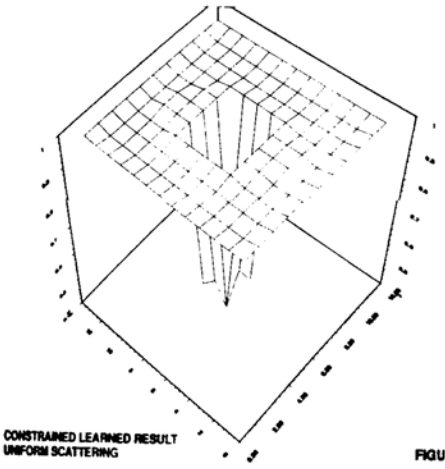
LEARNED RESULT
MODERATE SCATTERING

FIGURE 16



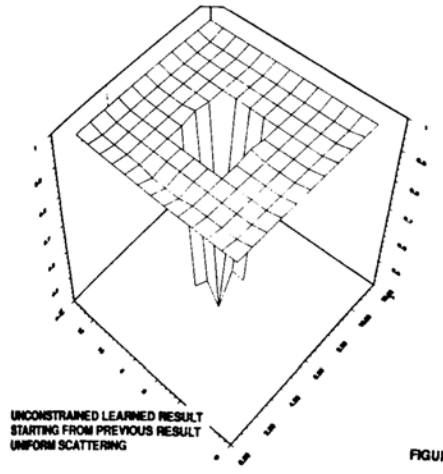
LEARNED RESULT
STRONG SCATTERING

FIGURE 17



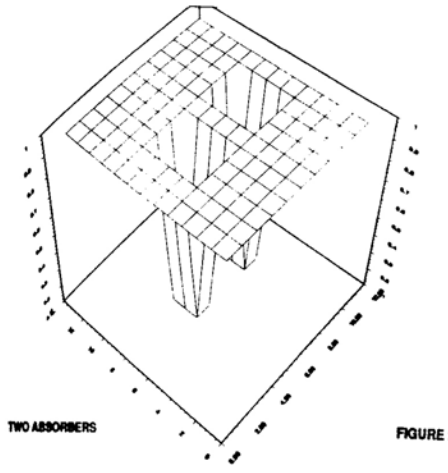
CONSTRAINED LEARNED RESULT
UNIFORM SCATTERING

FIGURE 18



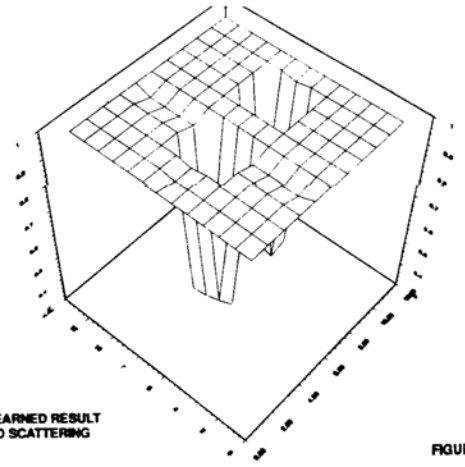
UNCONSTRAINED LEARNED RESULT
STARTING FROM PREVIOUS RESULT
UNIFORM SCATTERING

FIGURE 19



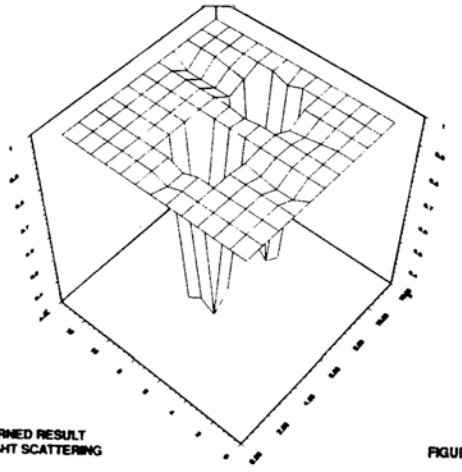
TWO ABSORBERS

FIGURE 20



LEARNED RESULT
NO SCATTERING

FIGURE 21



LEARNED RESULT
SLIGHT SCATTERING

FIGURE 22