An iterative solution to seismic tomography

Kamal Al-Yahya

ABSTRACT

The tomographic problem is posed as an inverse problem. The iterative method, which is practical for geophysical tomography, is investigated and the advantages and limitations of this method are displayed through simulated examples of cross-hole and surface-to-hole geometries. The generalized inverse approach is investigated in another paper (Al-Yahya, this SEP report).

INTRODUCTION

Viewed in a general way, tomography is an inverse problem: given certain observations from outside a medium, we want to determine some properties of that medium. The observations used in tomography are line integrals of some function of the medium. For geophysical tomography, this function can be the attenuation factor, the slowness (the reciprocal of velocity) of sound (Bois et al., 1972; Dines and Lytle, 1979), or electrical properties (Daily et al., 1982). The known solutions to this problem are generally based on ray optics, in which these line integrals represent rays that travel through the medium. It is well known that ray optics is valid only under some restrictions. The main restriction is that the wave length of the travelling energy must be much smaller than any characteristic quantity of length dimension of the medium, like the transmitter-receiver separation and the size of anomalous bodies (Červený et al., 1977).

Furthermore, most of the methods that are based on ray theory assume that the line

integrals (the rays) are straight. This assumption is valid only if the variation in the medium is not large enough to cause appreciable bending of the rays. It is advisable to, whenever possible, use methods that lift this restriction.

Devaney (1982) discusses a wave-equation method of solving the tomographic problem. Because this method is based on the wave equation (actually, an approximation to the wave equation), it is naturally expected to perform better than do the optical methods when the assumptions of ray optics, mentioned above, fail.

The purpose of this study is to look into the details of the problem in order to find the best reconstruction method. In fact, I will discuss techniques that can be classified as *restoration* rather than reconstruction. The difference between reconstruction and restoration is that reconstruction combines the information of the line integrals in order to reconstruct the image, while restoration attempts to compensate for the smearing of those line integrals (Bracewell, 1984).

In all of the following discussion, ray optics is assumed. However, curved, instead of straight, ray paths will be used to lift the restriction of weak velocity contrast. I have not chosen to use wave-equation methods because the current sampling and apertures used in borehole recordings are not suitable for wave-equation methods. This subject will, hopefully, be discussed in a later study. I also have the goal of incorporating anisotropy in the reconstruction algorithm in a later study.

the tomographic problem

For a two-dimensional medium, the tomographic problem can be stated as follows: Determine the function $f(x,y)$ given a set of its projections, or line integrals

$$p(r,\theta) = \int f(x,y)ds$$

for a range of projection angles $\theta = \tan^{-1}(s/r)$, where x, y, r, s, and θ are as shown in Figure 1. The integration is performed from the transmitter to the receiver.

Fourier methods

Fourier methods are based on the projection-slice theorem, originally formulated by Bracewell (1956), which states that the one-dimensional Fourier transform of the projection taken at an angle θ is equal to a slice at the same angle θ of the two-dimensional Fourier transform of the image.

A well known Fourier method is the filtered back-projection method (Deans, 1983).

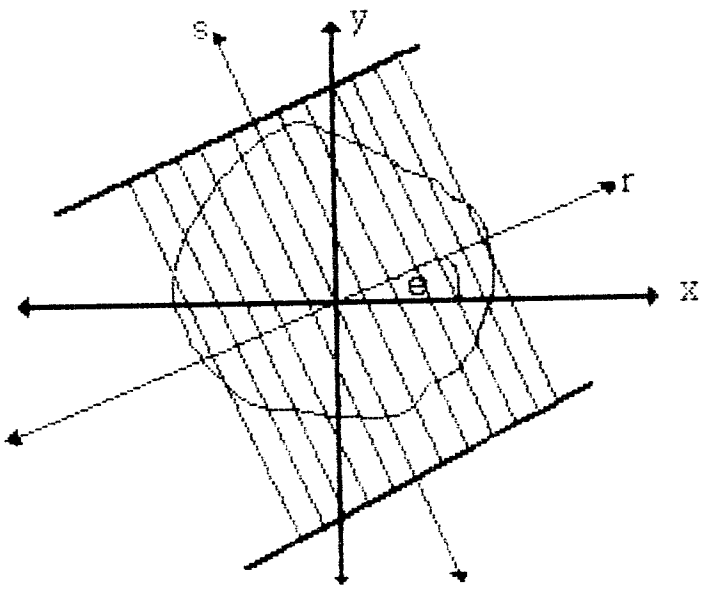

FIG. 1. The geometry of tomography: the r-s axes are rotated around the x-z axes by an angle θ.

When this method is used, the function $f(x,y)$ is reconstructed by

$$f(x,y) = \frac{1}{4\pi^2} \int_0^{\pi} \int_{-\infty}^{\infty} |k| \hat{p}(k,\theta) e^{-ikr} dk d\theta \tag{1}$$

where $|k|$ is a *rho filter* and $\hat{p}(k,\theta)$ is the Fourier transform of the projections $p(r)$ taken at an angle θ. The correspondence between convolution in one domain and multiplication in the transform domain can be used to change the multiplication of the rho filter by $\hat{p}(k,\theta)$ in equation (1) into convolution and the method is thusly known as the convolution method.

It is worth mentioning that the wave-equation methods of tomography, mentioned earlier, use a theorem similar to the projection-slice theorem: the *diffraction*-projection theorem in which semi-circles, instead of slices, of the Fourier transform of the image are assembled (Pan and Kak, 1983).

GEOPHYSICAL TOMOGRAPHY

Various recording geometries are used in geophysical prospecting. Using any of these geometries provides data that can be used in the tomographic inversion. A typical example is cross-hole geometry, in which transmitters are placed in a well and receivers are placed

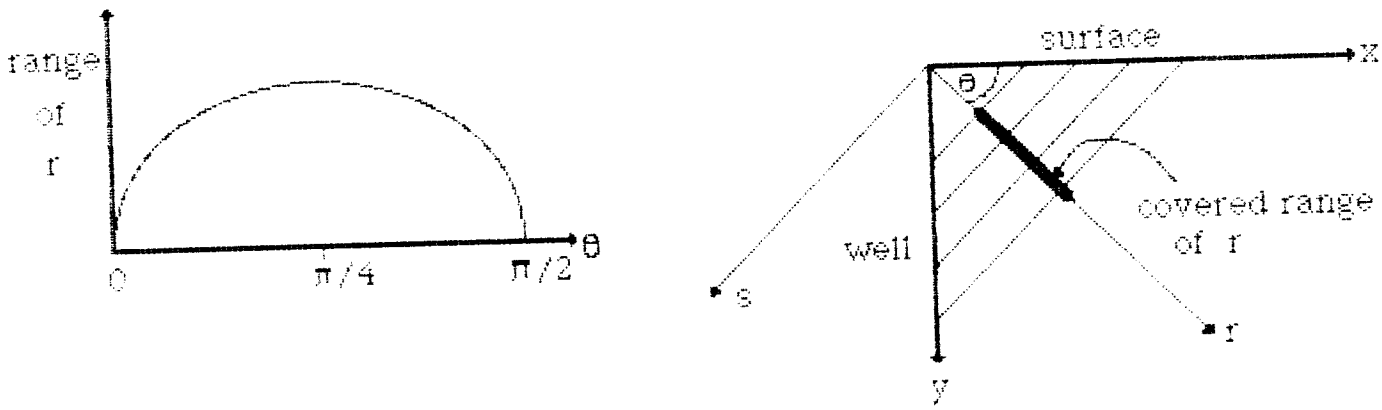

FIG. 2. The limited coverage of the surface-to-hole recording geometry (Figure 3b). Note that the coverage of the r-axis is not uniform.

in another well some distance apart from the first (Figure 3a). Unlike medical tomography, geophysical tomography is often severly limited in coverage, as shown in Figure 2. Furthermore, the integration paths are not straight, so the Fourier method of equation (1) is not applicable. In this paper, the objective is the determination of the velocity of the medium. The projections are therefore the travel times between transmitters and receivers (the first break). If we have n transmitter-receiver pairs, then the travel time for the j-th pair is

$$t_j = \int_{l_j(\mathbf{w})} \mathbf{w} dl_j(\mathbf{w}), \quad j = 1, 2, \ldots, n \tag{2}$$

where $\mathbf{w} = w(x, y)$ is the slowness of the medium. If the medium is subdivided into rectangular cells, the slowness in each cell being constant, then equation (2) can be written as

$$t_j = \sum_i w_i l_i \qquad j = 1, 2, \ldots, n \tag{3a}$$

or in matrix form,

$$\mathbf{t} = \mathbf{L}\mathbf{w} \tag{3b}$$

It is very important to note that equations (2) and (3) are nonlinear because dl_j is strongly dependent on the model $\mathbf{w}$. In our treatment of the tomographic problem, we will need to linearize these equations by replacing $dl_j(\mathbf{w})$ by $dl_j(\mathbf{w}_0)$, where $\mathbf{w}_0$ is an initial guess or an updated smooth version of $\mathbf{w}$.

In equation (3), the summation is along the cells crossed by the ray. To determine these cells we trace the ray through the medium. It should be noted that the ray tracing

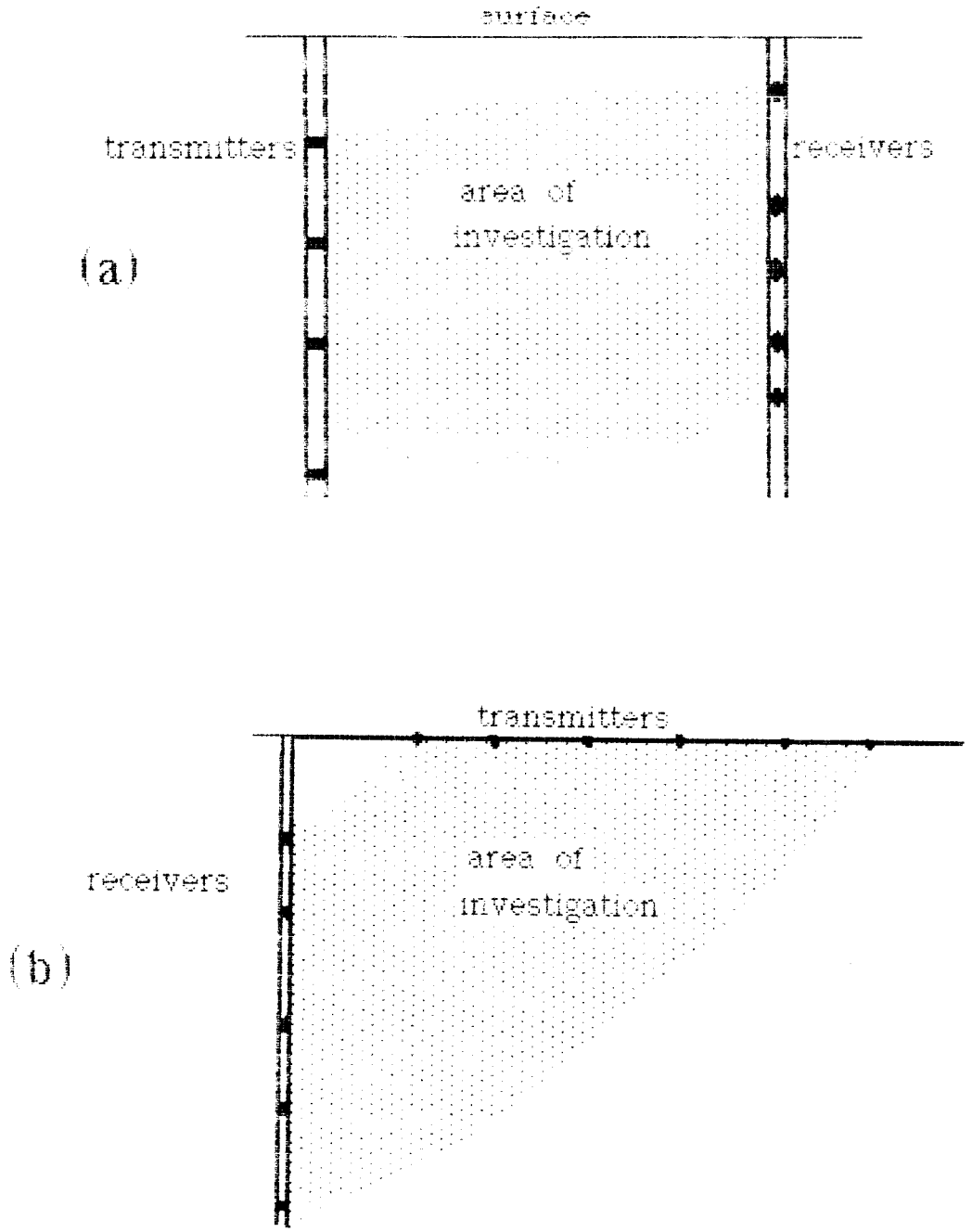

FIG. 3. a) Cross-hole geometry. b) Surface-to-hole geometry.

itself does not have to consider the medium as subdivided into cells. Known ray-tracing schemes in which the ray paths are smooth curves can be used.

There are several theoretical solutions to the linear form of equation (3). The linearity assumption imposes an iterative solution. This paper discusses one iterative solution known as the algebraic reconstruction technique (ART). I also studied the singular value decomposition of the problem. Such decomposition is the basis for the generalized inversed solution. To keep the size of this paper reasonable, the generalized inverse is treated separately in another paper (Al-Yahya, this SEP report).

THE ART METHOD

The algebraic reconstruction technique seems to be the most practical way to solve the problem, because it has the following advantages:

1. It can be used for any recording geometry.
2. It does not require the assumption of straight ray-paths.
3. It does not require large amount of computer memory because only a small number of parameters enter the calculation at any stage.

From equation (3a), the travel time for a given transmitter-receiver pair is

$$\sum_i w_i l_i = t \ .$$

At a given iteration, we assign the cell i a slownesses $\hat{w}_i$. For this particular slowness, we will have a travel time $\hat{t}$, so

$$\sum_i \hat{w}_i l_i = \hat{t} \ .$$

Subtracting the first from the second of the above two equations yields,

$$\sum_i \Delta w_i l_i = \Delta t \tag{4}$$

where $\Delta t = t - \hat{t}$ and $\Delta w_i = w_i - \hat{w}_i$. In order to update our model, we need to determine Δw_i's in equation (4). However, this equation is an under-determined system where we have one equation with many unknowns. To obtain a unique solution for this equation, we regard it as a constraint to a minimization problem in which we minimize some norm. A reasonable norm to use is the Euclidean norm (the L^2 norm),

$$L^2 = \sum_i (\Delta w_i)^2$$

The solution of equation (4) subject to this constraint is easily found by use of the Lagrange multipliers method. Instead of minimizing the norm $\sum_i \Delta w_i^2$, we minimize a combination of the norm and the constraint, that is we minimize

$$K = \sum_i (\Delta w_i^2 - \lambda \Delta w_i l_i) \quad ,$$

where the Lagrange multiplier, λ, is to be determined. When the derivative of K with respect to Δw_i is equated with zero (thus minimizing K), we obtain

$$\Delta w_i = \frac{\lambda}{2} l_i \tag{5}$$

To determine the Lagrange multiplier λ, we substitute for Δw_i in equation (4) to get

$$\lambda = \frac{2\Delta t}{\sum_i l_i^2}$$

and therefore, from equation (5),

$$\Delta w_i = \frac{l_i}{\sum_i l_i^2} \Delta t \tag{6}$$

Dines and Lytle (1979) point out that any *even* norm, L^{2p}, can be used. In this case, equation (6) becomes

$$\Delta w_i = \frac{l_i^{\frac{1}{2p-1}}}{\sum_i l_i^{\frac{2p}{2p-1}}} \Delta t \tag{7a}$$

which is equivalent to equation (12) in their paper (with "typo" corrected!). They also point out that the minimax norm with $p \to \infty$ can be used. In this case equation (7a) becomes

$$\Delta w_i = \frac{\Delta t}{\sum_i l_i} \tag{7b}$$

which is the well known ART algorithm (Gordon 1974).

Equations (7) give the solution of one iteration for one ray; this solution can be used to update those portions of the model passed by the ray. In equation (7a), the error between the actual travel time and the travel time of the guessed function is distributed among the cells so that each cell is given a weight l_i, the length of the ray segment in that cell. In equation (7b), all cells are given the same correction, regardless of the length of the segment. Projecting the error along the ray path is termed *an iteration* in the literature. Dines and Lytle (1979) suggest that the correction be delayed until all rays have been

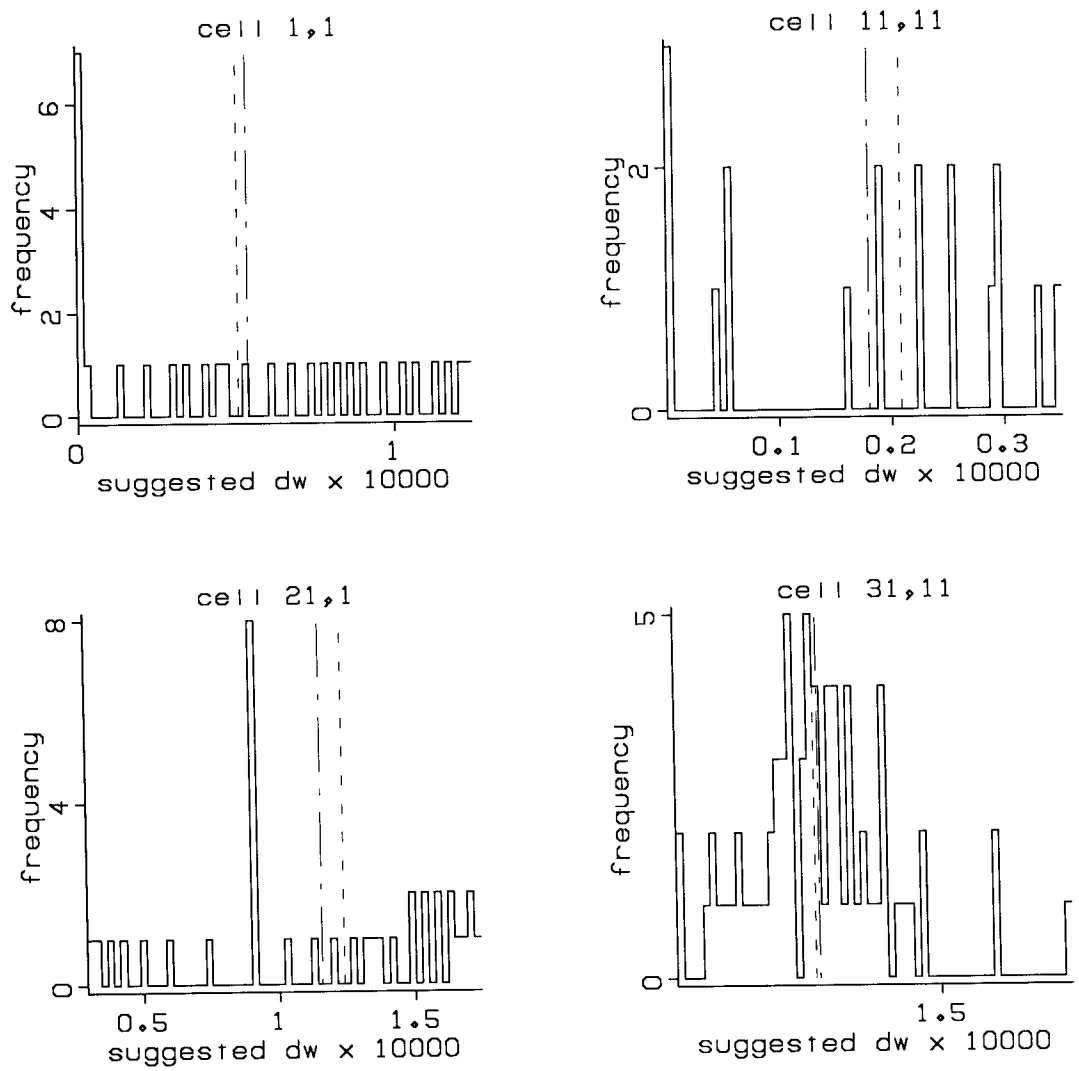

FIG. 4. Histograms for few cells in the model of Figure 6. The row and column number of the cell is shown at the top of each histogram. The dashed line is the median, the broken line is the mean. In the top two cells the mode is better than the mean and the median. In cell 21,1 the mode is the worst and in cell 31,11 all three are close to each other.

processed; the slowness of each cell is then updated by adding to the slowness the mean of the corrections of the passing rays for each cell. This method has the advantage of uncoupling the rays: the error of one ray does not immediately affect other rays.

The iterative method attempts to fit the model to the recorded data, so the misfit between the recorded data and the data corresponding to the reconstructed model should decrease through the iterations. A quantity that reasonably measures this misfit is the

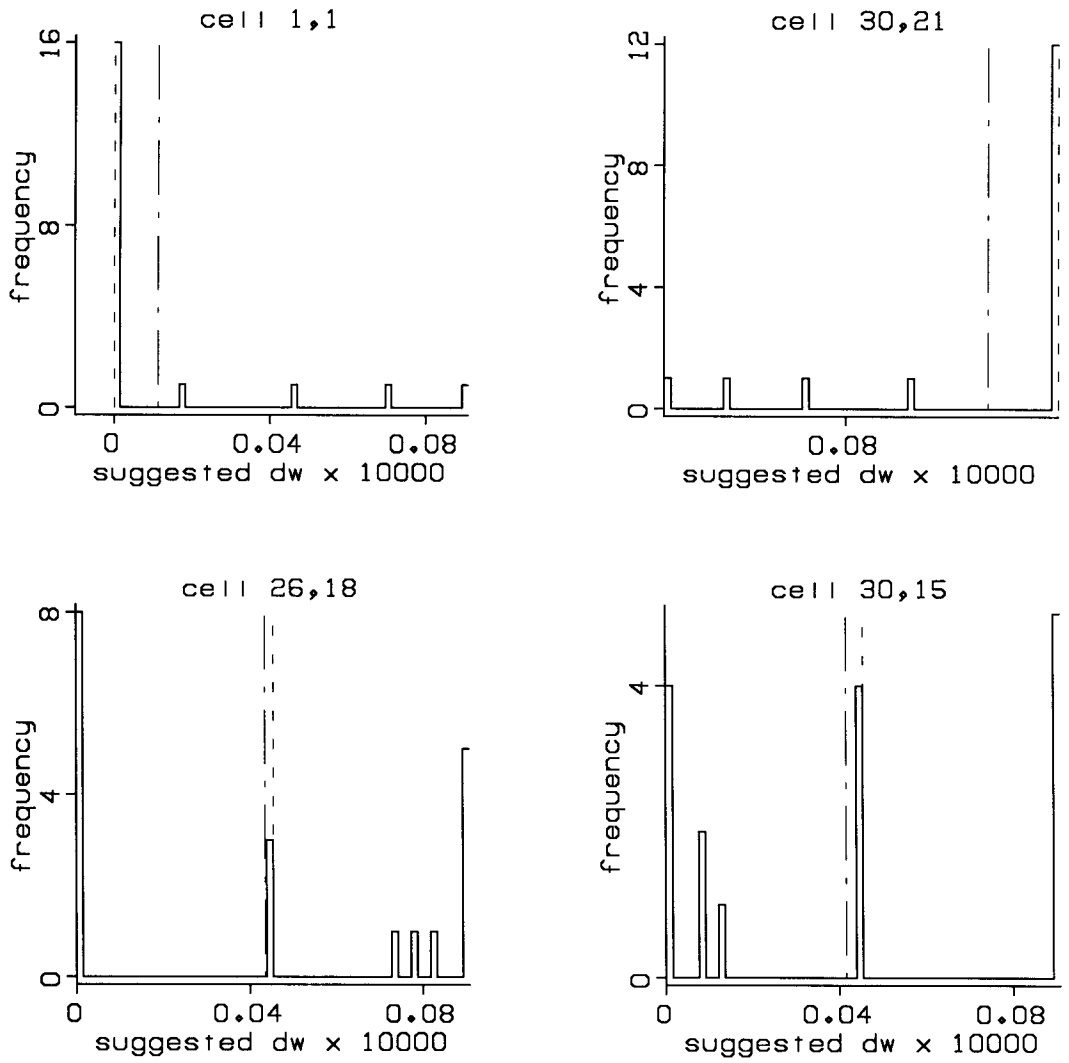

FIG. 5. Histograms for few cells in the model of Figure 16. The row and column number of the cell is shown at the top of each histogram. The dashed line is the median, the broken line is the mean. In the top two cells, the mode and the median are the same. In the bottom two the mode is the best estimate.

square error, E, defined by

$$E = \sum_i (t_i - \hat{t}_i)^2 \tag{8}$$

where t_i is the recorded time and $\hat{t}_i$ is the time corresponding to the reconstructed model for the i-th transmitter-receiver pair. We continue to update our model via equations (7) until the square error, E, becomes very small (compared to a predetermined criterion).

Constraining the solution

Various constraints can be applied to the solution. More weight can be given to rays at a particular angle if the medium has the least variation in that direction. We can also put upper and/or lower limits to the velocity if such limits are known or even fix the velocity at some cells if those velocities have been previously determined. Finally, we can limit the derivatives of the velocity function of the reconstructed image. Of the various constraints I used only the last one which was found to improve the results.

Convergence of the iterative method

It is important to know if the iterative method will converge, and if it does, to what solution. Minimizing the Euclidean norm produces a "unique" answer for an underdetermined system. The iterative method gives a similar minimization: for the *unconstrained* problem the solution is that which minimizes

$$\psi(\mathbf{w}) = \sum_i (w_i - \tilde{w})^2 \ ,$$

where $\tilde{w}$ is the average slowness of the model (Herman et al., 1973). This is a minimum-variance solution: we will obtain the *smoothest* of all possible solutions.

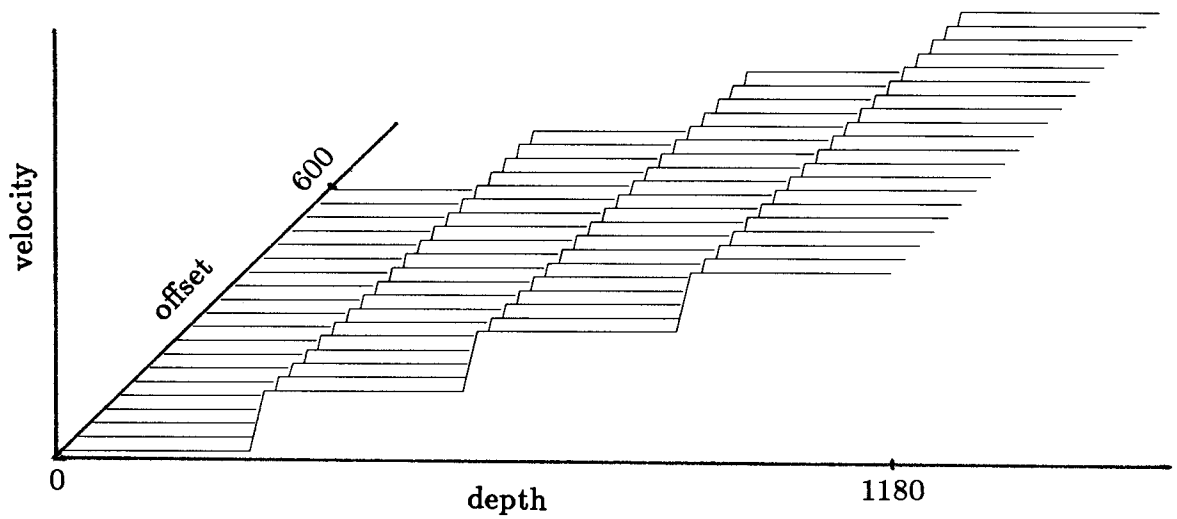

FIG. 6. A layered-earth model. The velocity increase with depth from 1000 m/sec to 1300 m/sec.

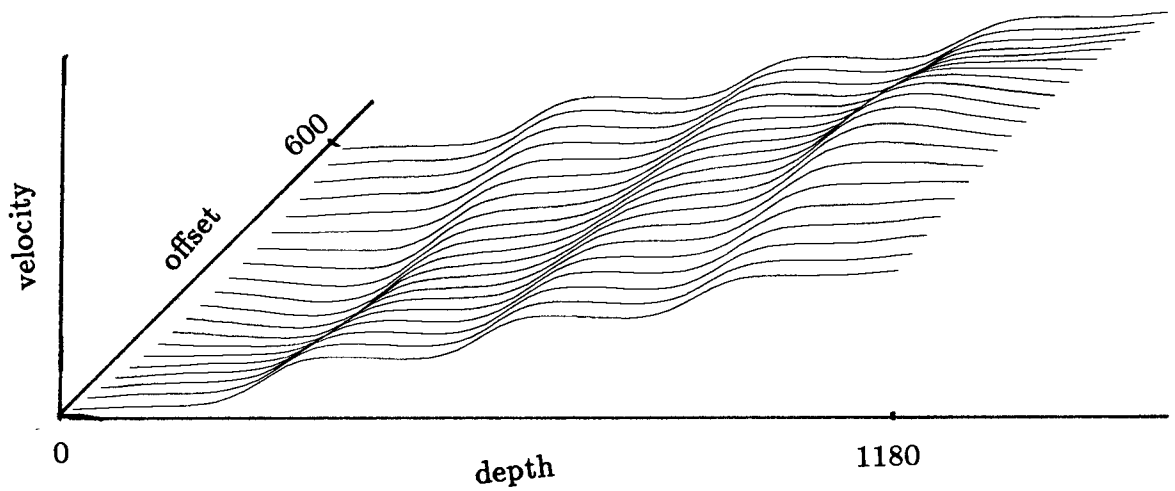

FIG. 7. The reconstructed model of Figure 6 using the mean of suggested values to correct a single cell.

REFINING THE RECONSTRUCTION

Let's closely examine the updating procedure which has two stages. In the first stage, a correction for all cells along the ray path is computed from equation (7a) or (7b). The second stage is done after all rays have been traced. The second stage is a decision-making process in which we use all passing rays to update the slowness of each cell. We can think of each ray as giving a *suggestion* on how to update the slowness of the cell. If one of these rays has passed through a small anomalous body, like that shown in Figure 16 (which will be discussed later), the first stage will suggest that all cells in the path of this ray be updated according to the observed error Δt. An erroneous suggestion can thus be given to those cells. If we use the mean in the second stage, even only one or two erroneous suggestions will affect the updating procedure of the cell. Taking the mean of a number of samples is basically a least-squares operation, which is severely affected by erratic samples. This problem can be solved if the median of the data is taken instead of the mean (Clærbout and Muir 1973).

Let's look at few cells in one iteration and see what suggestions we get for two particular examples. Figure 4 shows histograms of the suggestions made for some cells from the model shown in Figure 6. In this model, the mean and median are close to each other and we expect the difference between the reconstructed images to be small. We notice that each of them is different from the maximum-likelyhood value (the *mode*). Because I created the model and know the desired corrections, I can see that the mode gives a

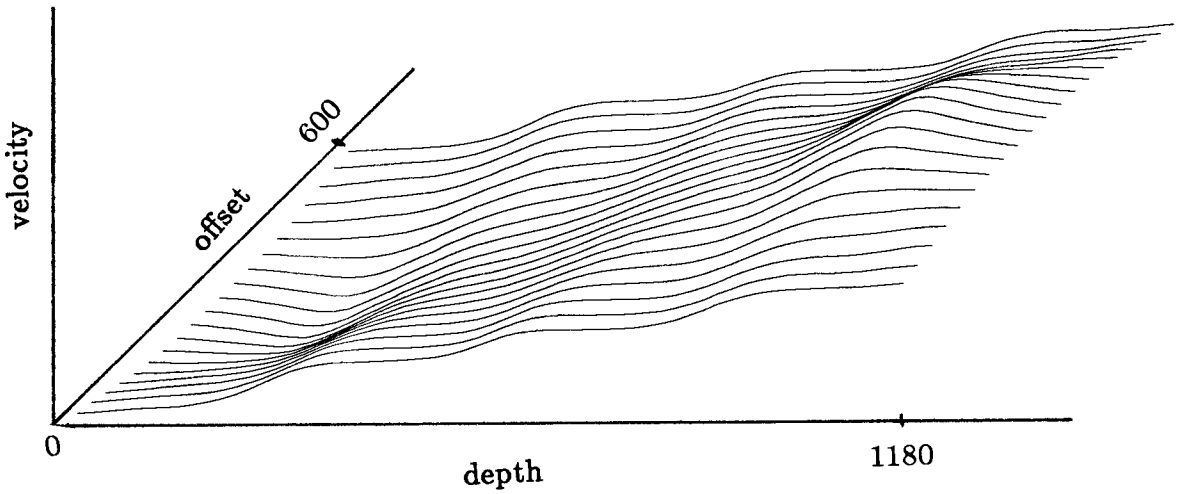

FIG. 8. A reconstructed model of Figure 6. The median of suggested values was used to correct each cell.

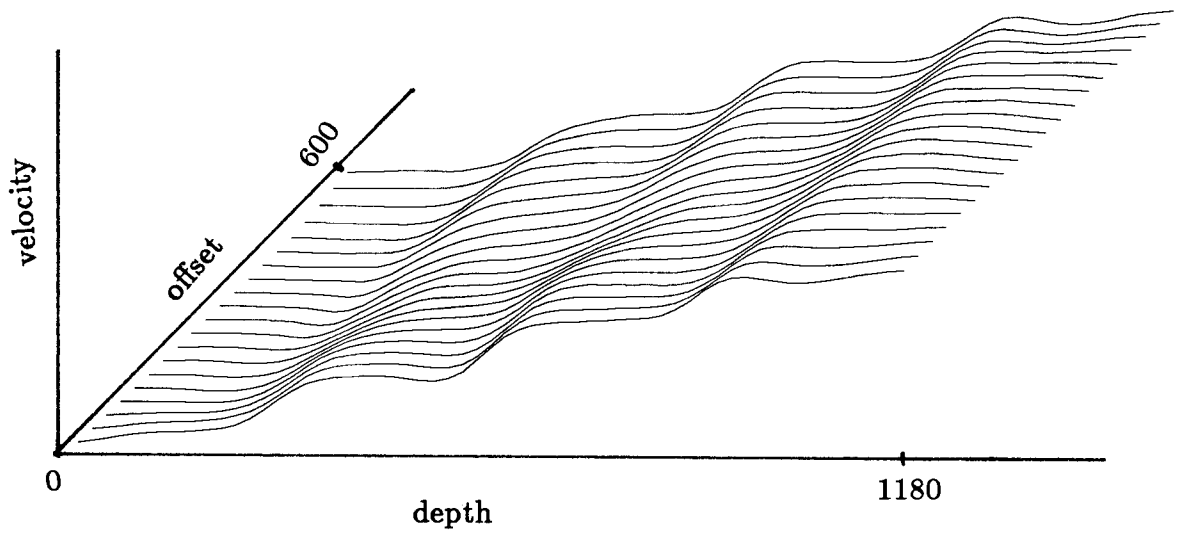

FIG. 9. A reconstructed model of Figure 6. The mode of suggested values was used to correct each cell.

correction better than that produced by either the mean or the median. Figure 5 shows histograms for the suggestions made for some cells from the model shown in Figure 16 (discussed later). Again, knowing the desired corrections, we can conclude from the histograms that the median produces better results than those produced by the mean. It is also clear that the mode in general produces better results than produced by the median or the mean. However, it is not clear that the mode will give the best answer for all models.

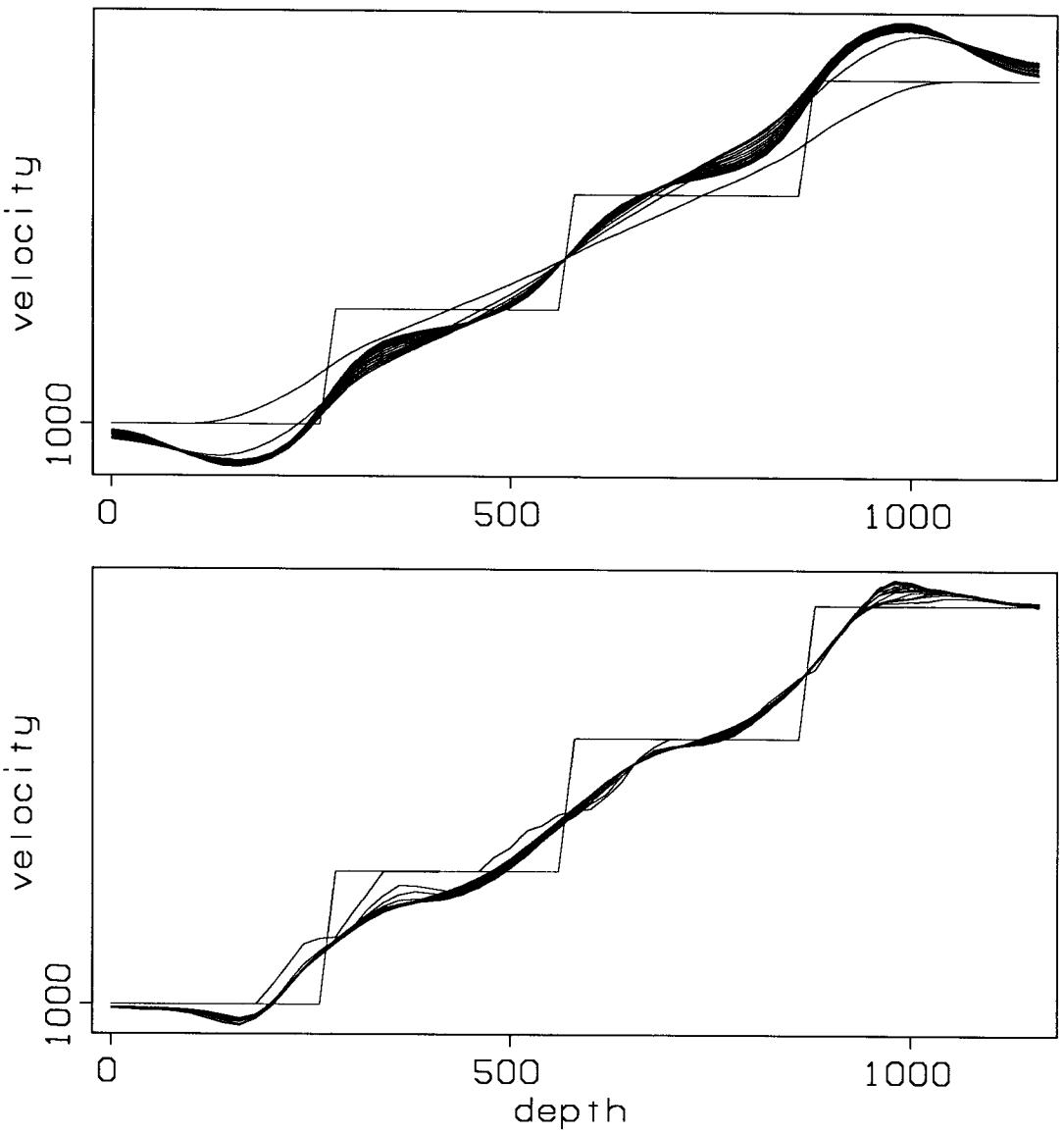

FIG. 10. Profiles taken at x=300 m from Figure 7 (top) and Figure 9 (bottom) for 12 iterations.

Because we have seen the histograms for only one iteration, it is not even clear that, for these two examples, the mode will continue to provide better results than those produced by the mean and the median throughout all iterations.

SOME EXAMPLES

I will now show some results produced with these three estimates: the mean, the median, and the mode using the ART method. In the first example, I attempt to reconstruct the data for the layered earth model of Figure 6. The velocity increases from 1000

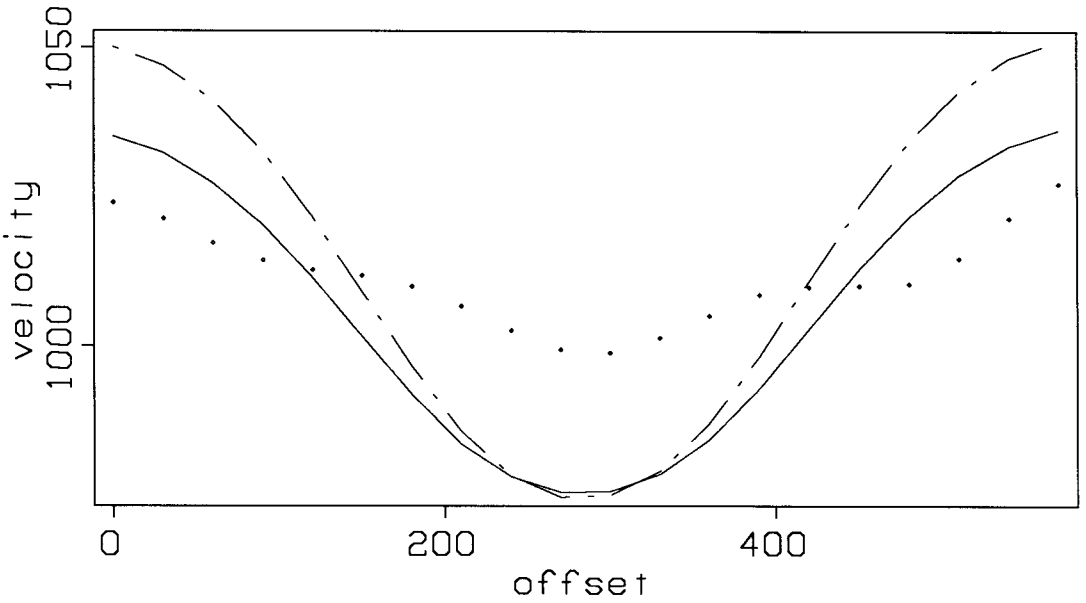

FIG. 11. The reconstructed velocity as a function of offset for the layered-earth model of Figure 6; solid=mean, dashed=median, dotted=mode. The velocity should be a constant function of offset (v=1000 m/sec).

to 1300 m/sec in 100 m/sec steps. The results of reconstruction are shown in Figures 7 to 11 for the cross-hole example. The wells were 1180 m deep and 600 m apart. The region between the two wells was divided into 59 rows, dz=20 m, and 20 columns, dx=30 m. 30 transmitters and 30 receivers were used with both transmitter's and receiver's interval being 40 m. We see that Figures 7 and 8, in which the mean or median was used, show a lateral variation in velocity at the top and bottom parts of the reconstructed image. This "artificial" lateral variation is undesired because it will affect our study of the anisotropy of the medium, a property I would like to include in the tomographic inversion at a later stage. The situation in Figure 9 in which the mode was used is different: the velocity of the reconstructed model has negligible lateral variation. This difference is shown in Figure 11 for the cells at z=200 m. It is clear that, for that depth in this particular model, the result of using the mode is superior to those of the mean and the median.

The square error of equation (8) for the layered-earth model is shown in Figure 12, which shows that the mode gives the fastest decay of the error (However, the error rising slightly at the last few iterations indicates that I could have stopped much earlier).

I now try to reconstruct the image of a surface-to-hole recording geometry for the same model (Figure 3b). The well was 390 m deep in which there were 20 receivers placed at 20 m intervals. On the surface, there were 20 transmitters at 10 m intervals. The cells

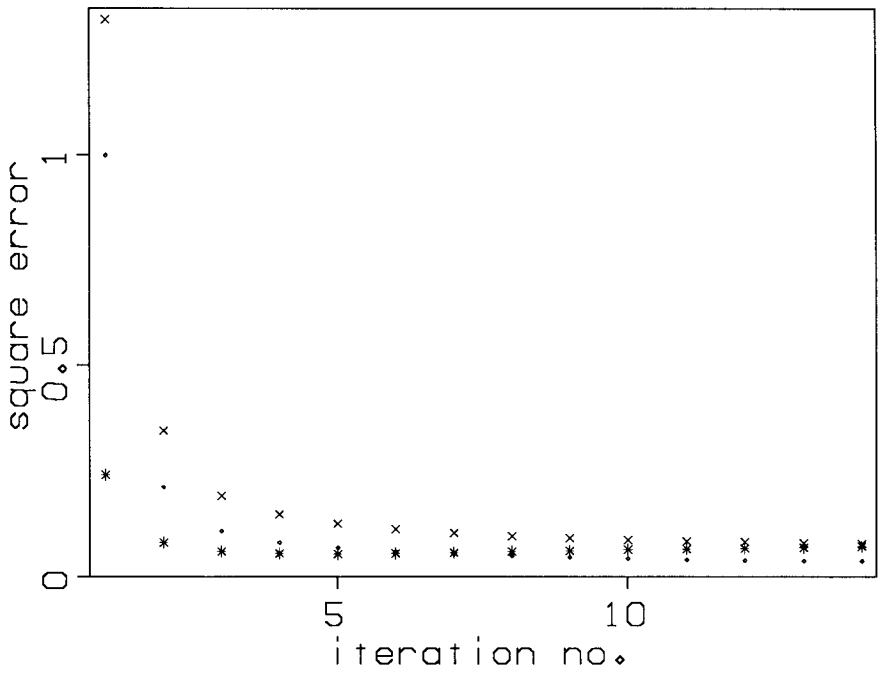

FIG. 12. Square error (arbitrary scale) as a function of iteration number. $\times$ = mean, $\cdot$ = median, $*$ = mode.

were 10 m deep and 10 m wide. The results of the reconstruction are shown in Figures 13 to 15. We see that only the velocity in the area through which the ray passed is corrected; the algorithm is *blind* in the other region. This blindness is directly related to the number of cells in a layer. If there is only *one* cell per layer, then an observation in any offset in a layer will be directly extended to all offsets in that layer.

Comparing the profiles of Figure 15, we see that this geometry requires more iterations for the algorithm to converge than the cross-hole geometry requires. The reason for this difference is that the shallow parts in the surface-to-hole geometry project their low velocity to the deeper parts, which stay under-corrected for the first few iterations. We also see that using the mode gives a slightly better result than that of the mean; the result of the mode also converges faster.

The second example demonstrates the effect of having a compact anomalous body in the middle of the medium. Figure 16 shows a model having a constant velocity of 1000 m/sec everywhere except for a small region in the middle, which has a velocity of 1200 m/sec. The wells were 580 m deep and 600 m apart. The region between the two wells was divided into 58 rows, dz=20 m, and 40 columns, dx=15 m. 30 transmitters and 30 receivers were used with both transmitter and receiver interval being 20 m. The anomalous

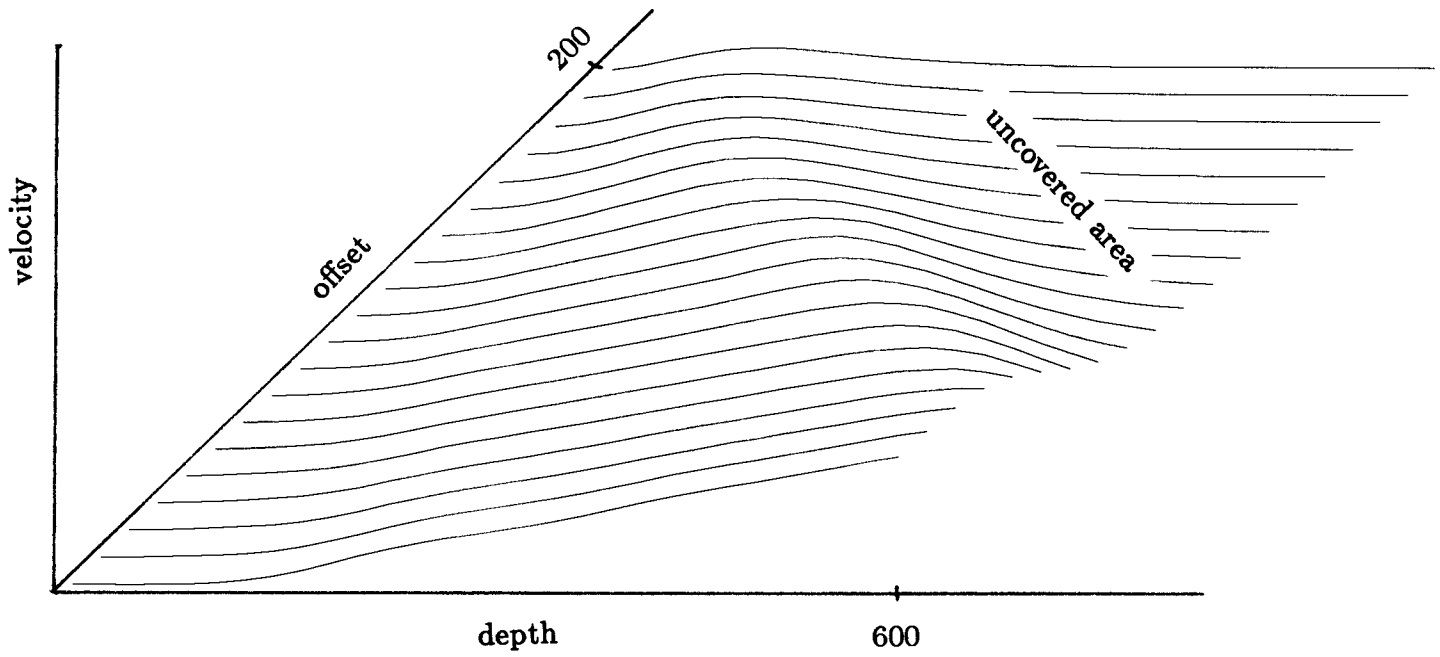

FIG. 13. the reconstructed model of Figure 6 using the mean and surface-to-hole geometry.

body is in the center of the region between the two wells. It is 4 cells wide and 6 cells deep. The reconstructed images made with the three methods discussed above, are shown in Figures 17 to 19. For this example, the median gives the best result; the others give a dispersed anomaly. This dispersion is expected when the mean is used because there is a degree of non-uniqueness, as mentioned above. The median gives a more compact image in this example because those few erratic suggestions are discarded.

It is interesting to note that the mode gives the most compact result for the first iteration, as shown in Figure 20. However, the result changes toward that of the mean as the iterations proceed, while that of the median stays compact. This is probably because the distribution of the suggestions changes toward a gaussian, something expected from the central-limit theorem.

CONCLUSIONS

Because algebraic reconstruction technique methods do not require a large amount of computer memory and can be used for general recording geometry, they are suitable for geophysical tomography.

We saw that looking at the distribution (making histograms) of the suggested correc-

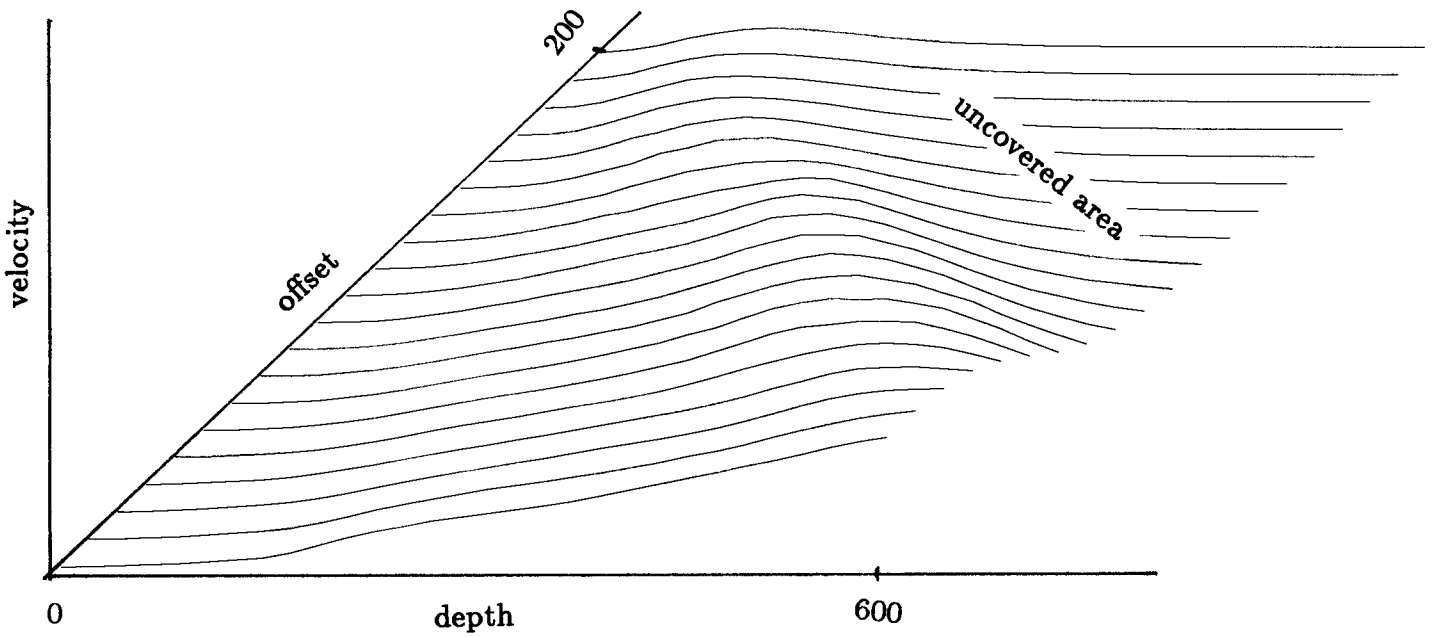

FIG. 14. the reconstructed model of Figure 6 using the mode and surface-to-hole geometry.

tion is useful. For some models, using the highest value in these histograms gives a better result than those produced by other methods. For models having compact anomalous bodies the median proved to be a superior tool.

ACKNOWLEDGMENTS

I benefited from many discussions with and many valuable suggestions from Francis Muir. Several discussions with Joe Dellinger were helpful. I thank John Toldi for many interesting discussions and for his critical review of the paper. Thanks are due to other SEP members who participated in the discussion of the topic.

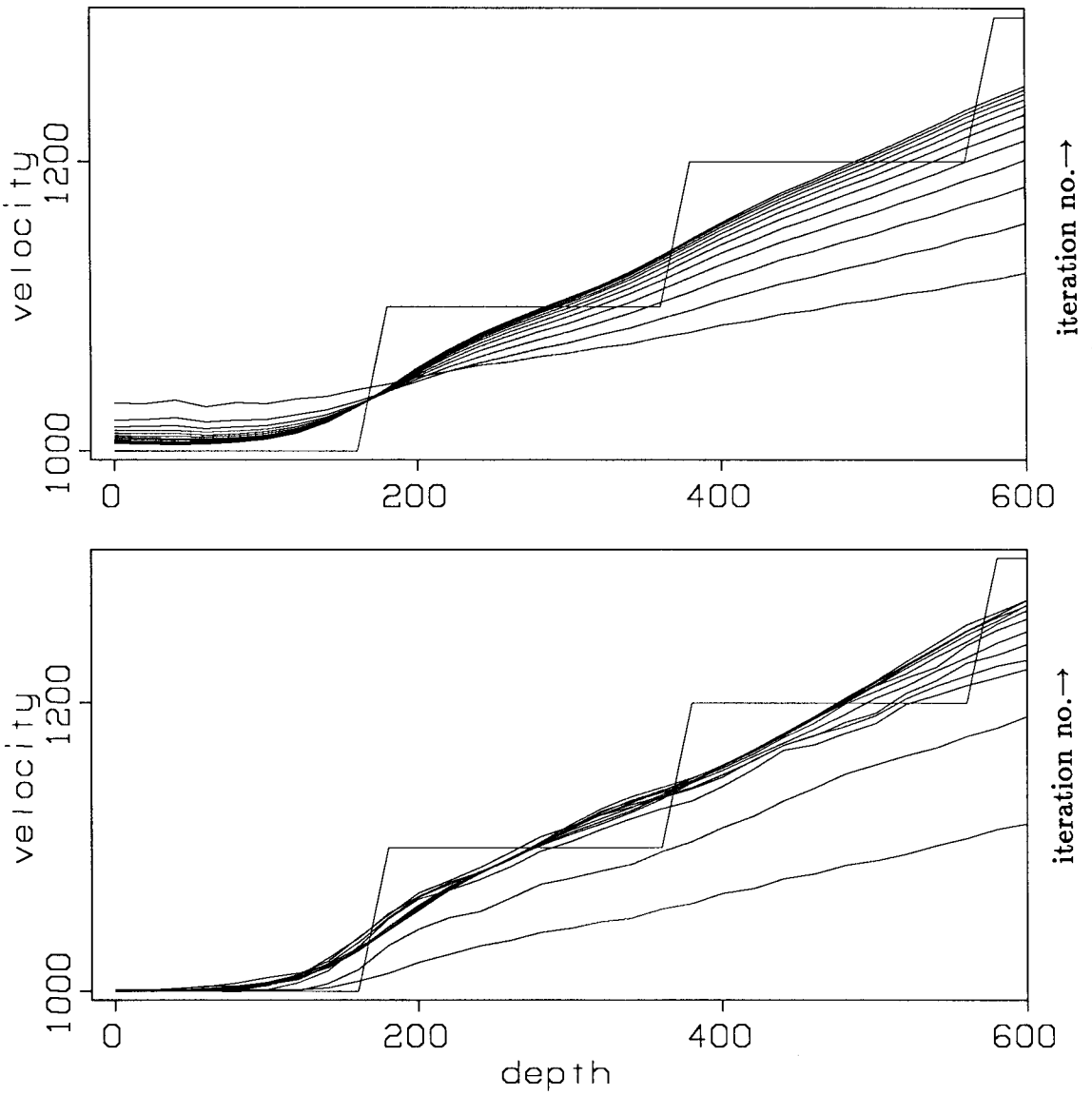

FIG. 15. Profiles taken at x=0 m from Figure 13 (top) and Figure 14 (bottom) for 12 iterations.

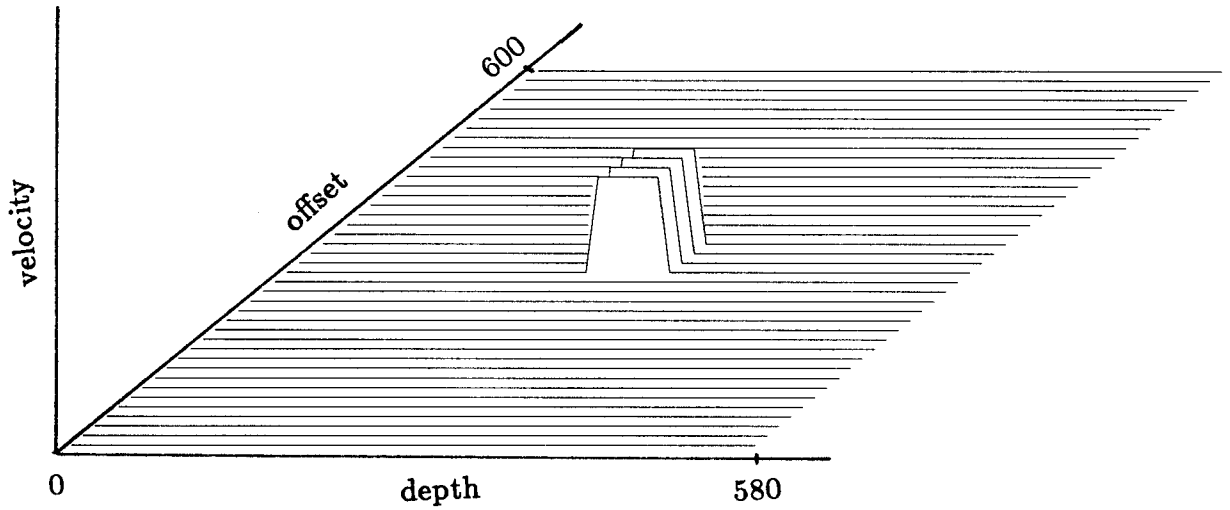

FIG. 16. A model having a high velocity anomaly in the middle of a constant velocity medium.

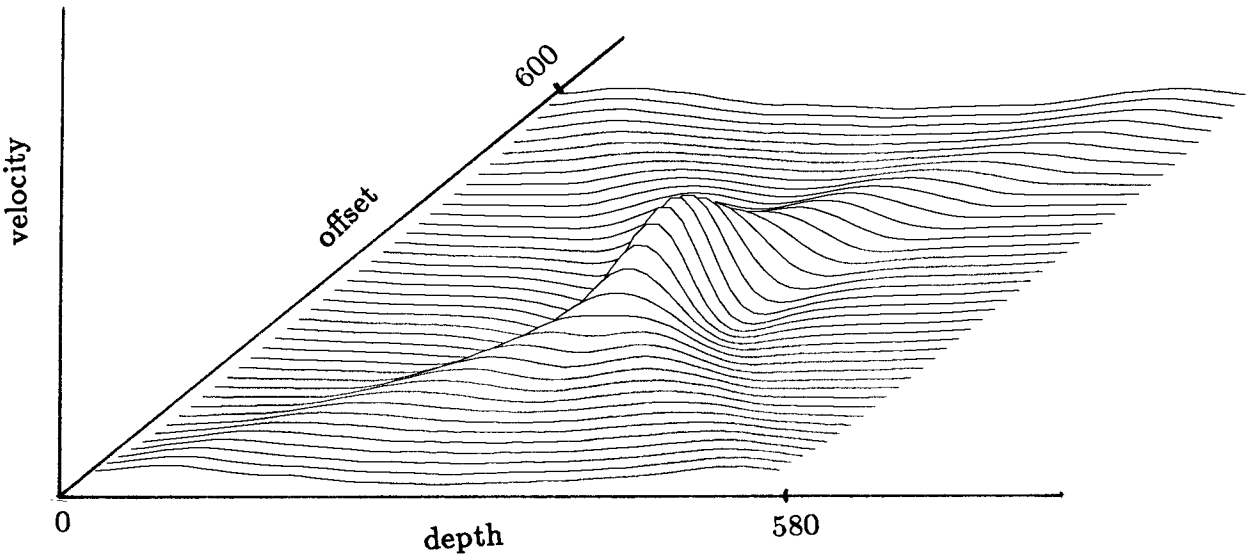

FIG. 17. The model of Figure 16 reconstructed with the mean of suggested values used to correct a single cell.

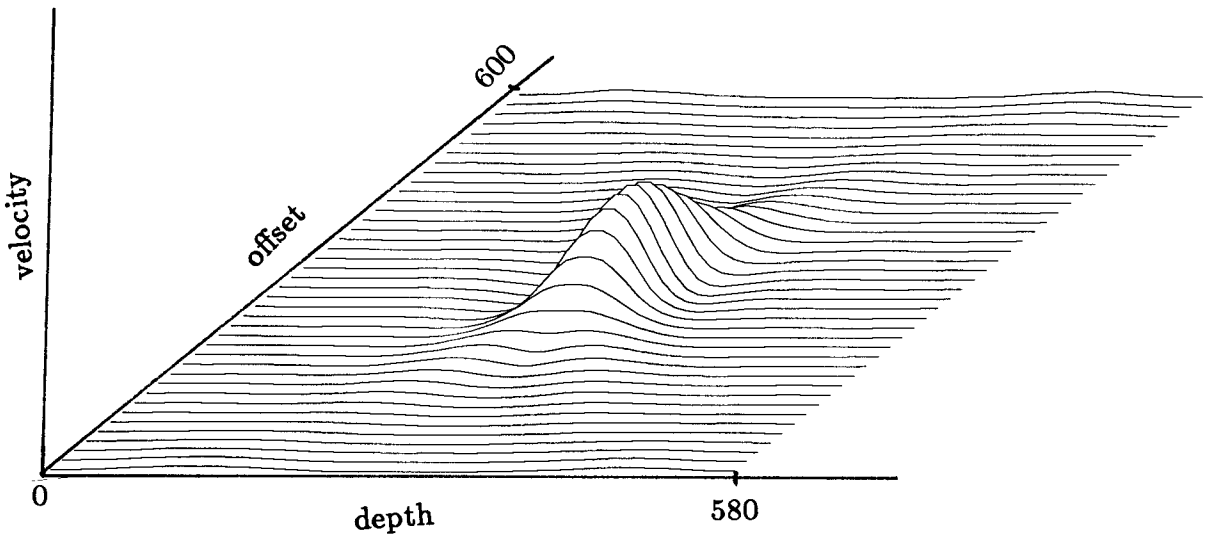

FIG. 18. The model of Figure 16 reconstructed with the median of suggested values used to correct a single cell.

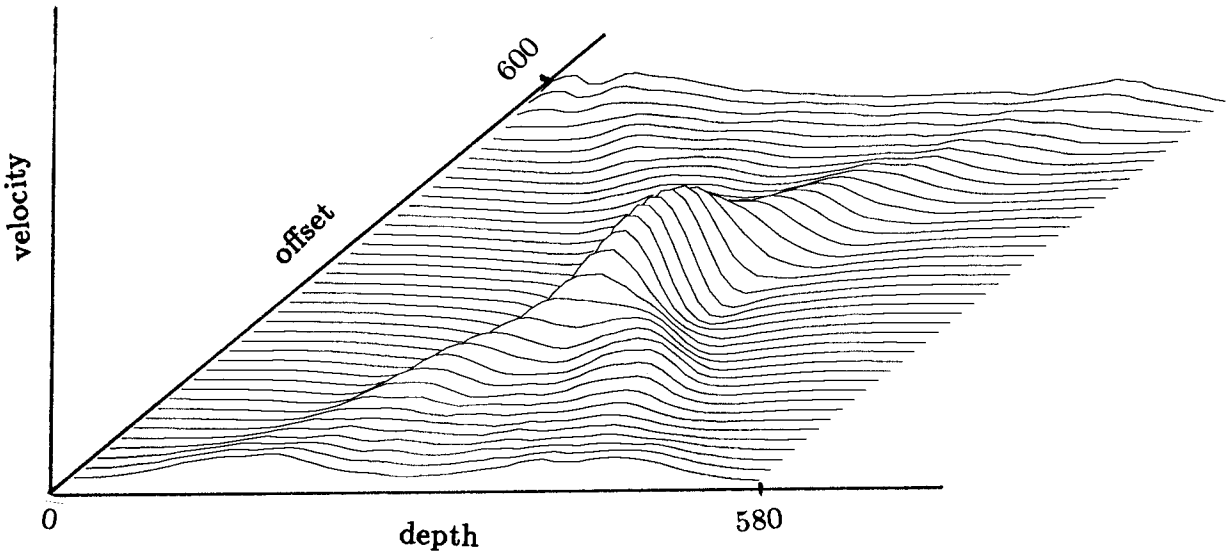

FIG. 19. The model of Figure 16 reconstructed with the mode of suggested values used to correct a single cell.

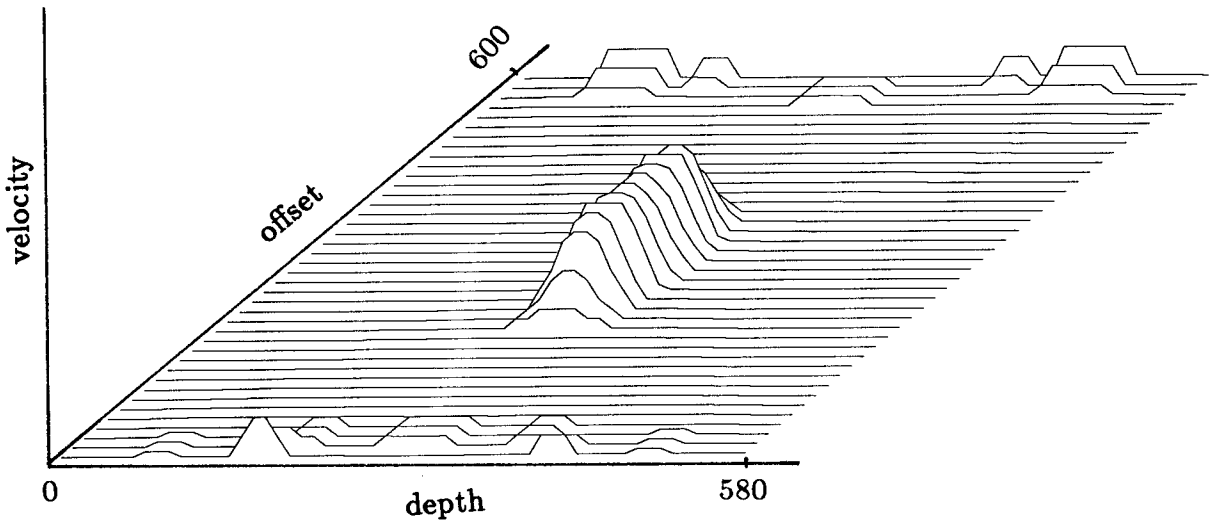

FIG. 20. The result of the first iteration using the mode for the model in Figure 16. Note the compactness of the anomaly compared to Figures 18 and 19.

REFERENCES

Al-Yahya, K., 1985, The generalized inverse approach to seismic tomography, this SEP report.

Bois, P., La Porte, M., Lavergene, M., and Thomas, G., 1972, Well to well measurements, Geophysics, v. 37, p. 471-480.

Bracewell, R. N., 1956, Strip integration in radio astronomy, Aust. J. Phys., v. 9, p. 198-217.

__, 1984, Two-dimensional imaging, class notes.

Červený, V., Moltokov. I. A., and Pšenčík, I., 1977, Ray method in seismology, Univezita Karlova, Prague, Czechoslovakia.

Clærbout , J. F., and Muir, F., 1973, Robust modeling with erratic data, Aust. J. Phys., v. 9, p. 198-217.

Daily, W. D., Lytle, R. J., Laine E. F., Okada, J. T., and Deadrick, F. J., 1982, Geotomography in oil shale, J. Geophysical Res., v. 87, p. 5507-5515.

Deans, S. R., 1983, The Radon transform and some of its applications, Wiley Interscience.

Devaney, A. J., 1982, A Filtered backpropagation algorithm for diffraction tomography, Ultrasonic imaging, v. 4, p. 336-350.

Dines, K. A., and Lytle, R. J., 1979, Computerized geophysical tomography, proceedings of the IEEE, p. 1065-1072.

Gordon, R., 1974, A tutorial on ART, IEEE transection on nuclear science, v. NS-21, p. 78-93.

Herman, G. T., Lent, A., and Rowland, S. W. 1973, ART: Mathematics and applications, J. Theor. Biol., v. 42, p. 1-32.

Pan, S. X., and Kak, A. C., 1983, A computational study of reconstruction algorithm for diffraction tomography, IEEE transection on acoustic, speech and signal processing, ASSP-31, p. 1262-1275.