Seismic inversion using fine grain parallel computers

Peter Mora

ABSTRACT

Seismic recordings depend on the seismic source, the properties of the Earth, the location of the seismic receiver stations, and the physics of seismic wave propagation. It has always been a dream in seismology to predict the Earth properties directly from the seismograms using our knowledge of how seismic waves are affected by the rocks. Thanks to theoretical developments and advances in computer technology, this dream is on the verge of being realized. The Earth properties can be estimated using a least squares conjugate gradient algorithm to solve for the Earth model which predicts seismograms that best match the observed data. A new theory puts the gradient direction required by this algorithm in terms of wave simulations. In the past, wave simulations in realistic Earth models were too CPU intensive for this formulation to be practical but this no longer appears to be the case. A well understood method to do wave simulations in media of arbitrary complexity is by directly solving the discretized wave equation using the method of finite differences. The Earth is parametrized as a grid with each node of the grid associated with the elastic properties governing seismic wave propagation. Finite differences are used at each node to propagate the seismic waves from one instant of time to the next. At any instant of time, the calculations at a given node are independent from the calculations at other nodes. Therefore, the calculations at an instant of time can be done at all nodes simultaneously. Hence, the method is ideally suited to fine grain parallel computer architectures such as that of the "Connection Machine". Results suggest that by using such fine grain parallel computers, realistic sized inverse problems can be solved for the first time!

THE SEISMOLOGISTS' DREAM

Earth images using traditional seismic processing methods

Traditionally, interpretation of reflection seismograms has been based on seismic processing methods. These methods modify the seismograms to obtain a picture of the reflectors in the Earth. It has always been dissatisfying that most of these methods require interpretive steps, approximations and oversimplifications.

The processing methods consist of sequential steps to modify the seismic records to produce an image of the Earth. In oil exploration, the most common processing methods are called velocity analysis, NMO stacking and migration. Normally, each of these steps requires considerable interpretive input from an experienced seismologist. Even then, several attempts at processing may be necessary before the seismologist is satisfied he has obtained a good image of the Earth.

Because of the large quantity of data involved in seismic experiments, the processing steps had to be fast and so the simplest approximations to describe seismic wave propagation were used. For instance, many methods are based on the assumption that seismic waves can be approximated as acoustic waves but the Earth is not a liquid even to first order! Both compressional and shear waves are observed! Even if the waves were acoustic, the most common methods in oil exploration still make simplifying assumptions that restrict the applicability of the different techniques. For example, velocity analysis methods normally assume that there is no refraction of seismic waves in the Earth. This assumption would only be true if the Earth were homogeneous but the Earth has structure and is not homogeneous!

The dream of obtaining Earth properties from seismic observations

Seismologists dream of the day when it will be possible to automatically obtain the Earth's physical properties directly from the recordings of seismic waves with no approximations or oversimplifications. In principle, this can be achieved by inverting the equations of physics describing seismic wave propagation. Instead of computing the data observations (seismograms) from known Earth properties using the wave equation, the Earth properties are computed from a set of seismic observations using an inverse equation. This can be done by finding the Earth model which predicts seismograms that best match the observed seismograms. The measure of match depends on the statistics of the noise in the data and the statistics of the Earth properties.

This inverse problem is not easy to solve because of its immense dimensions. If a 4x4x4 kilometer cube of the Earth is discretized every 10 meters then there are $3\text{x}400^3 \approx 2\text{x}10^9$ parameters required to define an isotropic elastic solid. Even more parameters are required if the Earth is anisotropic in this volume. The size of the data space, the digitized seismic records, for such a volume of the Earth is about 10^{10}. These are the sizes of model and data spaces corresponding to seismic surveys used for oil exploration. Full Earth seismic studies using Earthquake seismograms involve comparable sized spaces.

Even if an Earth model were obtained by solving the seismic inverse problem, its meaning may not be clear because of non-uniqueness.

Head in the cloud dreams

Rather than a single Earth model, a better answer would be the probability of each possible Earth structure (i.e. a range of answers and their meanings). This range of answers and associated meanings could be represented as a multidimensional function giving the probability of every possible distribution of rocks in the Earth. The probability map would tell the seismologists what their answer means! As if the inverse problem to find a single answer was not hard enough, now we need all possible answers to know what the single answer means!

Back to reality

The dream of probability maps of every possible Earth model is of such immense dimensions that it is considered impossible by most humans. To see why, consider how to compute a function giving the probability of each possible Earth model. One way is to generate synthetic seismic data for each realization of our discretized Earth and subsequently measure the probability of each realization. The probability would be measured by comparing the synthetic data with the data observations. When the two data sets look alike, the probability of the corresponding realization of Earth properties is high and when they are dissimilar the probability is low. For seismic inverse problems with 10^9 Earth parameters and say 100 realizations of each parameter then we would require $10^{9^{100}}$ forward modeling runs. Considering wave simulations take from seconds to hours on most computers, it is infeasible to compute probability maps using this brute force approach.

How about trying to solve for the single most likely solution? Is this smaller dream realizable? The answer appears to be yes provided a few assumptions are made. The most crucial assumption is that we can make educated guesses of the Earth properties that are accurate enough that the more generally applicable Monte Carlo and probability map methods are not required. In that case, we can obtain the most probable Earth model by doing only a few seismic wave simulations. Each simulation would determine the probability of the current Earth model contained in the computer. Some other calculations would determine how to change this Earth model to improve the probability.

INVERSION USING SEISMIC WAVE SIMULATIONS

Maximum probability inverse solution

Statistical knowledge is required in order to derive expressions for the most probable Earth properties. I assume Gaussian probability density functions for the Earth parameters and the noise in the seismic data observations. This corresponds to the least squares criterion to measure the fit between the synthetic and observed seismograms. Then the most probable Earth model can be found by iterative least squares which updates an Earth model iteratively until the best fit solution is obtained. Mora (1987a) used the preconditioned conjugate gradient algorithm which updates the Earth model as a linear combination of current and previous model perturbations. The current perturbations are a function of the least squares steepest descent direction which is the set of perturbations that most rapidly decrease the sum of squared difference between observed and synthetic seismograms.

The elastic forward problem

Least squares theory requires forward and adjoint calculations. The forward calculations consist of seismogram synthesis by modeling the propagation of seismic waves and the adjoint calculations that find the "steepest descent" model perturbations. Tarantola (1984) and Mora (1987a) have shown that for elastic waves, the adjoint calculations can be formulated in terms of the forward calculations.

If the Earth is assumed to be perfectly elastic, then the seismic forward problem, that of computing seismic data (seismograms) $\mathbf{d}$ from Earth properties $\mathbf{m}$ denoted $\mathbf{d}(\mathbf{m})$ may be computed by solving the elastic wave equation (Aki and Richards, 1980),

$$\rho \ddot{u}_i - \partial_j c_{ijkl} \partial_l u_k = f_i \ , \tag{1a}$$

$$c_{ijkl} \partial_l u_k n_j = T_i \ , \tag{1b}$$

$$u_i = 0 \ , \quad t < 0 \ , \tag{1c}$$

$$\dot{u}_i = 0 \ , \quad t < 0 \ , \tag{1d}$$

where $u_i = u_i(\mathbf{x}_S, \mathbf{x}, t)$ is the i-th component of displacement resulting from shot S (i.e. body force f_i and/or traction T_i) located at $\mathbf{x}_S$. If the receivers are located at $\mathbf{x}_R$ then digital data recorded every Δt seconds can be represented as:

$$d(S, R, j, i, \mathbf{m}) = u_i(\mathbf{x}_S, \mathbf{x}_R, (j-1)\Delta t, \mathbf{m}) \ , \quad \textit{where } i = 1, 2, 3 \ , \tag{2}$$

so $\mathbf{d}(\mathbf{m})$ is $d(\mathbf{m})$ arranged into a column vector of size $n_S x n_R x n_t x n_i$ where n_i is the number of components recorded by the receivers (e.g. $n_i = 2$ if the ground displacement in both the x and z directions is measured by the receivers).

The elastic inverse problem

From Tarantola (1984) and Mora(1987a) the adjoint calculations can be performed using expressions of form

$$\delta \hat{m}^\gamma(\mathbf{x}) = \sum_S \int dt \left(\Omega^\gamma_{ijk} u_j(\mathbf{x}_S, \mathbf{x}, t) \right) \left(\Omega^\gamma_{ijk} \psi_j(\mathbf{x}_S, \mathbf{x}, t) \right) \ , \tag{3}$$

where Ω^{γ}_{ijk} is an operator that is dependent on the physical property type denoted γ, u_j is the background wavefield computed using Earth model $\mathbf{m}$, and ψ_j, called the "back propagated residual wavefield", will be defined shortly. For an isotropic solid we may choose to specify the Earth properties in terms of the Lame moduli so $\gamma = \lambda$, μ or ρ. In that case we have

$$\Omega^{\rho}_{ijk} = \sqrt{-1}\,\partial_t \ , \ \Omega^{\lambda}_{ijk} = \sqrt{-1}\,\partial_j \ , \ \Omega^{\mu}_{ijk} = \sqrt{-1/2}\,(\delta_{jk}\partial_i + \delta_{ji}\partial_k) \ . \qquad (4)$$

A more experimentally observable and hence preferable choice of model parameters is the compressional and shear wavespeeds and density (see Mora, 1987a for the appropriate formulas). Note that implied sums inside ()'s in the above expression must be performed first.

The wavefield ψ_j is computed by applying the data residuals as a forcing function in the elastic wave equation but backwards in time, i.e.

$$\psi_j(\mathbf{x}_S, \mathbf{x}, t) = \sum_R G_{ij}(\mathbf{x}, -t; \mathbf{x}_R, 0) * \delta u_i(\mathbf{x}_S, \mathbf{x}_R, t) \ . \qquad (5)$$

The inverse calculations

Consider equations (3) though (5) which define the adjoint of the elastic wave equation. These can be calculated by the following steps:

(i) Propagation elastic waves through some Earth model $\mathbf{m}$ using equation (1) to solve for the background wavefield $u_j(\mathbf{x}_S, \mathbf{x}, t)$ and the synthetic seismograms $u_i(\mathbf{x}_S, \mathbf{x}_R, t)$.

(ii) Compute the residual seismograms $\delta u_i(\mathbf{x}_S, \mathbf{x}_R, t) = u_i(\mathbf{x}_S, \mathbf{x}_R, t) - u_{0i}(\mathbf{x}_S, \mathbf{x}_R, t)$. These seismograms measure the difference between the synthetic and observed seismic data.

(iii) Back propagate the residual seismograms by applying $\delta u_i(\mathbf{x}_S, \mathbf{x}_R, t)$ as a forcing function in a time reversed wave equation to evaluate the back propagated residual wavefield $\psi_j(\mathbf{x}_S, \mathbf{x}, t)$. Note that because the elastic wave eauation is time symmetric, equation (1) is also used to do the back propagation.

(iv) Calculate the time integral and shot sum in equation (3). In practice, this step is done during the back propagation to avoid storage of the wavefield ψ_j.

The power of the inverse formulation

Mora (1987c,1987d) pointed out that these inversion formula are like a combination of traditional migration methods of seismic processing that obtain reflector locations and tomographic and velocity analysis methods that obtain wavespeed between the reflectors. All seismic waves would be accounted for using this method and the most likely Earth model obtained. For example, Rayleigh waves which are traditionally treated as noise would be useful in the inversion, especially to resolve the compressional and shear wavespeeds near the Earth's surface. If this worked well, one of the oldest problems of reflection seismology called "the statics problem" would be solved. Namely, to account for traveltime delays of seismic waves caused by near surface wavespeed fluctuations (these delays distort the image of deeper reflectors if they are not taken into account).

To summarize, the real power of the inverse formulation in this paper is that it can theoretically do everything that all the methods of seismic data processing could do combined. It is a unified approach to estimate the Earth properties.

SIMULATING SEISMIC WAVES

In principle, two methods can be used to model seismic waves propagating in a complex inhomogeneous medium like the Earth:

Cellular Automata

Cellular Automata (CA) simulate the physics of wave propagation using a set of rules describing particle interactions on a microscopic level. Therefore, this method can be compared to real physics where molecules interact at a microscopic level giving rise to seismic waves and other macroscopic phenomena. Although research on acoustic wave simulations (e.g. Muir, 1987) based on fluid flow models (see Wolfram, 1986) is progressing rapidly, it has not yet extended to cover elastic waves.

Finite differences

Finite differences simulates the continuous differential equations describing wave motion by using finite difference approximations to the true derivatives. In comparison to the Cellular Automata approach where a microscopic description of molecular interactions is used to statistically simulate the macroscopic phenomena, finite differences directly solves the macroscopic equations. This method is well understood and is routinely applied at the present time. Mora (1986) has derived a fast and accurate isotropic elastic modeling scheme based on Kosloff et al. (1984) but using convolutional operators to perform the spatial derivatives rather than Fourier transforms. The finite difference calculations are done using stress variables and are summarized by:

$$
\begin{array}{l}
\textit{for all time } \{ \\
\quad \begin{array}{lcll}
\sigma(\mathbf{x},t) & \rightarrow & \ddot{u}(\mathbf{x},t) & \textit{from the elastic wave equation} \\
\ddot{\mathbf{u}}(\mathbf{x},t) & \rightarrow & \ddot{e}(\mathbf{x},t) & \textit{from the strain displacement relation} \\
\ddot{e}(\mathbf{x},t) & \rightarrow & \ddot{\sigma}(\mathbf{x},t) & \textit{from Hooke's Law} \\
\ddot{\sigma}(\mathbf{x},t) & \rightarrow & \sigma(\mathbf{x},t+\Delta t) & \textit{step in time using an explicit finite difference scheme}
\end{array} \\
\} \ .
\end{array}
\tag{6}
$$

PARALLELISM IN NATURE AND COMPUTERS

How can finite differences be done fast enough to be useful for inversion? First observe that nature is intrinsically a parallel process (i.e. particles may vibrate simultaneously in different locations of the Earth). Surely, we can build a computer that can simulate wave propagation as fast or faster than they propagate in the Earth! Then, the outlined inversion method using wave propagations would be feasible and an Earth image could be automatically computed in real time! What would be necessary to achieve this kind of speed of calculation? The most obvious answer is a computer that is built to look like the Earth with many particles (processors) that operate simultaneously. The most crucial feature of a fine grain parallel computer is the ability of nodes (processors) to communicate with adjacent nodes. This communication must be done about as rapidly as the processors do a calculation or the parallel computer would be inefficient.

Physical processes such as wave propagations are easily simulated on massively parallel computers. To see this observe that the calculations at all **x** locations at an instant of time in the finite difference algorithm (equation (6)) can be performed simultaneously. The "Connection Machine" of Thinking Machines Corporation (Hillis, 1986) with 64,000 processors has the highest level of parallelism existing today and well suited to solving the seismic inverse problem using equations (1) through (6)). It is interesting that the creation of the "Connection Machine" was motivated by another physical problem, that of simulating the brain. The brain consists of many interconnected neurons and so the creators of the "Connection Machine" put great effort into solving the important processor connectivity problem (hence the computer's name).

Figure 2 shows how some synthetic data was generated by simulating elastic waves propagating through an Earth model using the method of finite differences (equation (6)). The Earth model in this figure represents a typical cross section of a sedimentary basin in an oil producing region. The different frames contain snapshots of waves propagating through the Earth model at an instant of time and the seismograms recorded thus far in the calculations (see Figure 1 for a description of one frame). The finite difference calculations over the entire Earth model are done in parallel so the computer time is proportional to the length of the seismogram time axis rather than the complexity of the Earth model. This is just the way real physics works with time progressing at the same rate whether or not the Earth has a complex structure.

The CPU time for this simulation is about one minute on the CM-2 "Connection Machine" which is only ten times slower than the waves take to propagate through the Earth in reality. Because the simulation is slower than the physical experiment, real time inversions using equations (1) through (6) are not yet possible. However, the simulation is fast enough that the inversion process is feasible using recorded seismic data for the first time!

INITIAL TESTS

Inversion tests done by Mora (1987b) are encouraging and fuel the dream and desire to be parallel. They indicate that very good pictures of the Earth may be obtained by the inversion process. Therefore, the dream of feeding seismograms into computers to obtain the Earth properties may soon become a reality as fine grain parallel architectures become more widespread.

The synthetic seismograms shown in Figure 2 were inverted with a linear with depth initial velocity model. Figures 3 through 5 depict the inversion process and demonstrate its dependence on the wave simulations. Figure 3 shows a forward modeling run using the initial velocity model to generate the background wavefield u. The data residual calculated by subtracting u_0 from u is used as a forcing function in reverse time to calculate the wavefield ψ as shown in Figure 4. As the calculations proceed, the velocity and density perturbations are computed using equations of the form of equation (3). The new Earth model is computed by adding these perturbations to the current model. This summarizes one iteration of the inversion procedure.

In this example, the inversion algorithm converged to a solution Earth model that generated best matching synthetic data after 10 iterations (see Figure 5). The 10 iteration solution shown in Figure 5 looks like the Earth model of Figure 2 verifying that the inversion technique works at least under ideal circumstances.

CONCLUSIONS

Fine grain parallel computers are well suited to simulating physical processes. In seismology, the inverse problem to find the Earth properties using the seismic data observations can be formulated in terms of the physics of wave propagation and is hence suited to parallel computations. Results from an implementation on the "Connection Machine" indicate that realistic sized oil exploration seismic inverse problems can be tackled. This brings the seismologists' dream of feeding seismograms into a computer and waiting for an Earth model to pop out one step closer to becoming a reality.

ACKNOWLEDGEMENTS

Thanks to the sponsors of the Stanford Exploration Project, Jon Claerbout and Thinking Machines Corporation during this research. Thanks also to Albert Tarantola for his inspirational guidance.

REFERENCES

Hillis, W. D., 1986, The connection machine: The MIT Press.

Kosloff D., Reshef M., and Loewenthal D., 1984, Elastic wave calculations by the Fourier method: Bulletin of the Seismological Society of America, **74** , 875-891.

Mora, P., 1986, Elastic finite differences with convolutional operators: Stanford exploration project report **48** , 277.

Mora, P., 1987a, Nonlinear 2D elastic inversion of multi-offset seismic data: Geophysics, **52**.

Mora, P., 1987b, Elastic wavefield inversion of reflection and transmission data: Geophysics, submitted.

Mora, P., 1987c, Elastic wavefield inversion for low and high wavenumbers of the P- and S-wave velocities, a possible solution: in Proceedings for the research workshop on deconvolution and inversion, September 1986, Rome, Italy.

Mora, P., 1987d, Inversion = migration + tomography: Geophysics, submitted.

Muir, F., 1987, Three experimental modeling systems, Stanford exploration report **51**, 119.

Tarantola, A., 1984, The seismic reflection inverse problem, in Inverse problems of acoustic and elastic waves, edited by: F. Santosa, Y.H. Pao, W. Symes, and Ch. Holland, SIAM, Philadelphia.

Wolfram, S., 1986, Theory and applications of cellular automata: World Scientific Press.

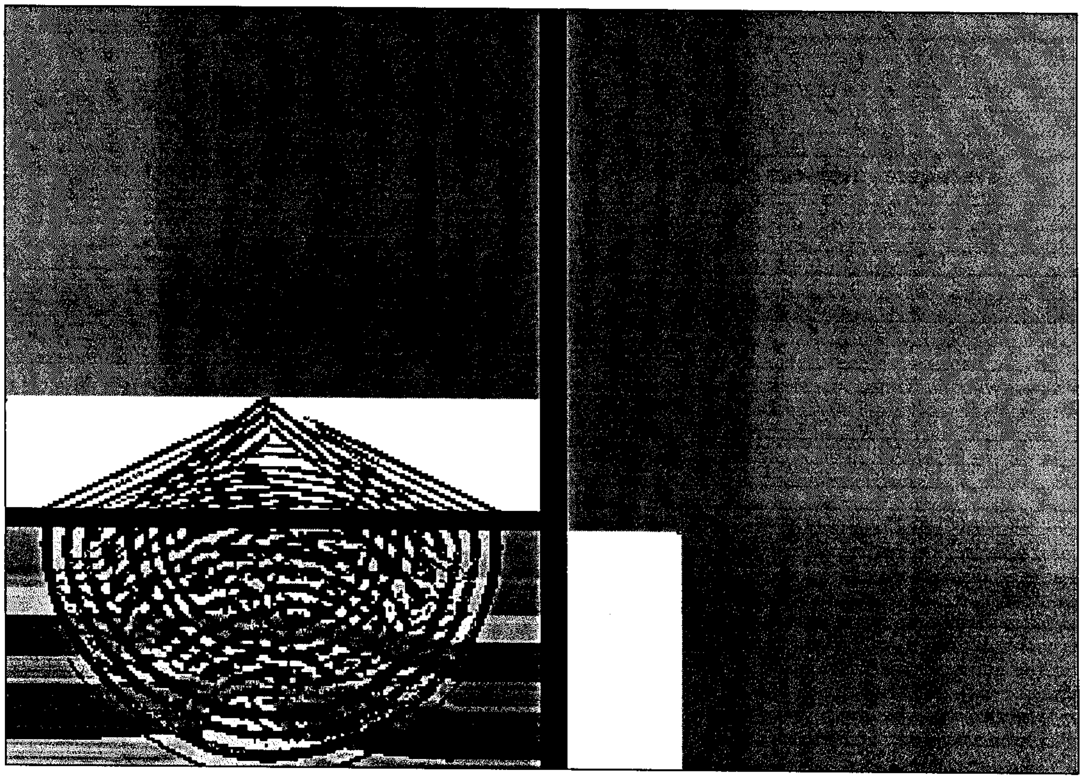

FIG. 1. A frame of an elastic wave simulation .67 seconds after a seismic source was activated. The left half of the plot contains the seismograms recorded at the Earth's surface in the upper portion (the time axis ranges from 0.00 secs at top (gray-white boundary) to 0.67 secs at the bottom) and the velocity model with the seismic waves superposed in the lower portion (the depth axis ranges from 0.0 km at the top to 1.8 km at the bottom). The horizontal axis for the left half of the plot is distance ranging from 0.0 km to 4.0 km. The lower right half of the plot contains VSP seismograms recorded down a well on the right of the model and has a vertical depth axis and horizontal time axis (the time axis ranges from 0.00 secs at right (gray-white boundary) to 0.67 secs at the left).

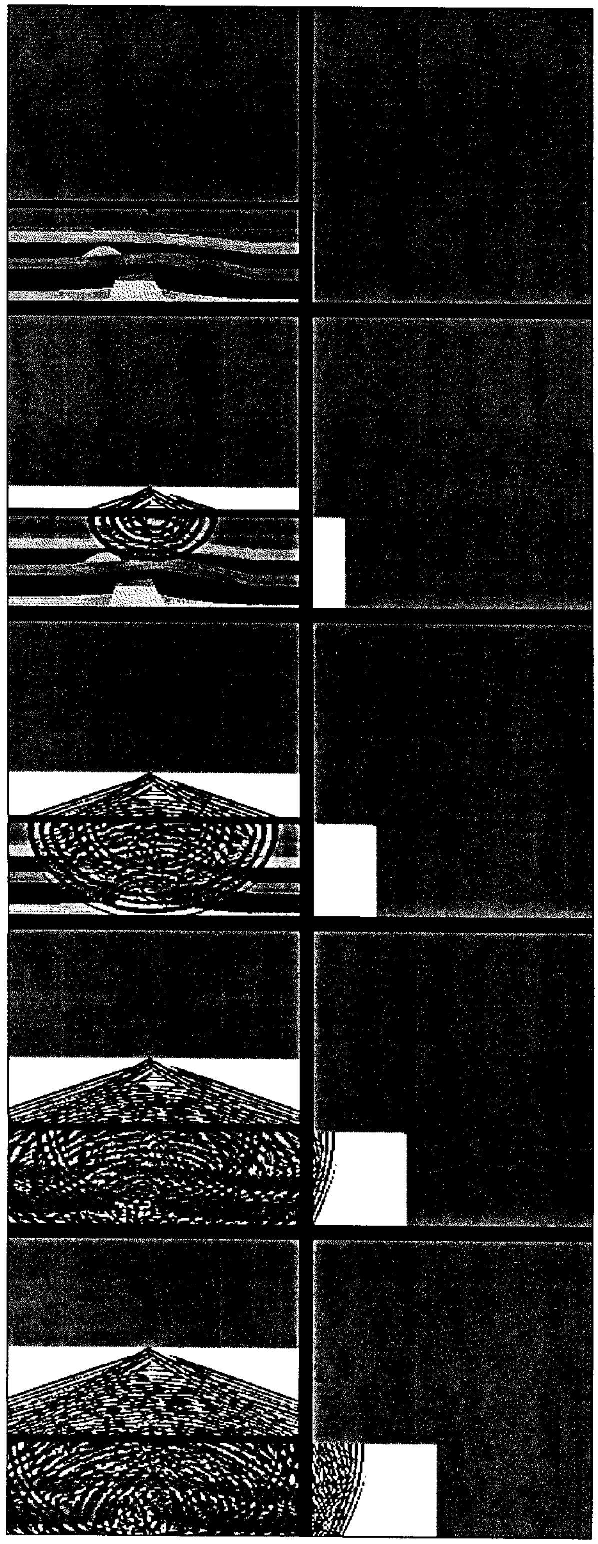

FIG. 2. (a). Snapshot frames every .33 seconds of an elastic wave simulation from t=0.00 secs at the top to t=1.33 secs at the bottom. Refer to Figure 1 for a description of a frame.

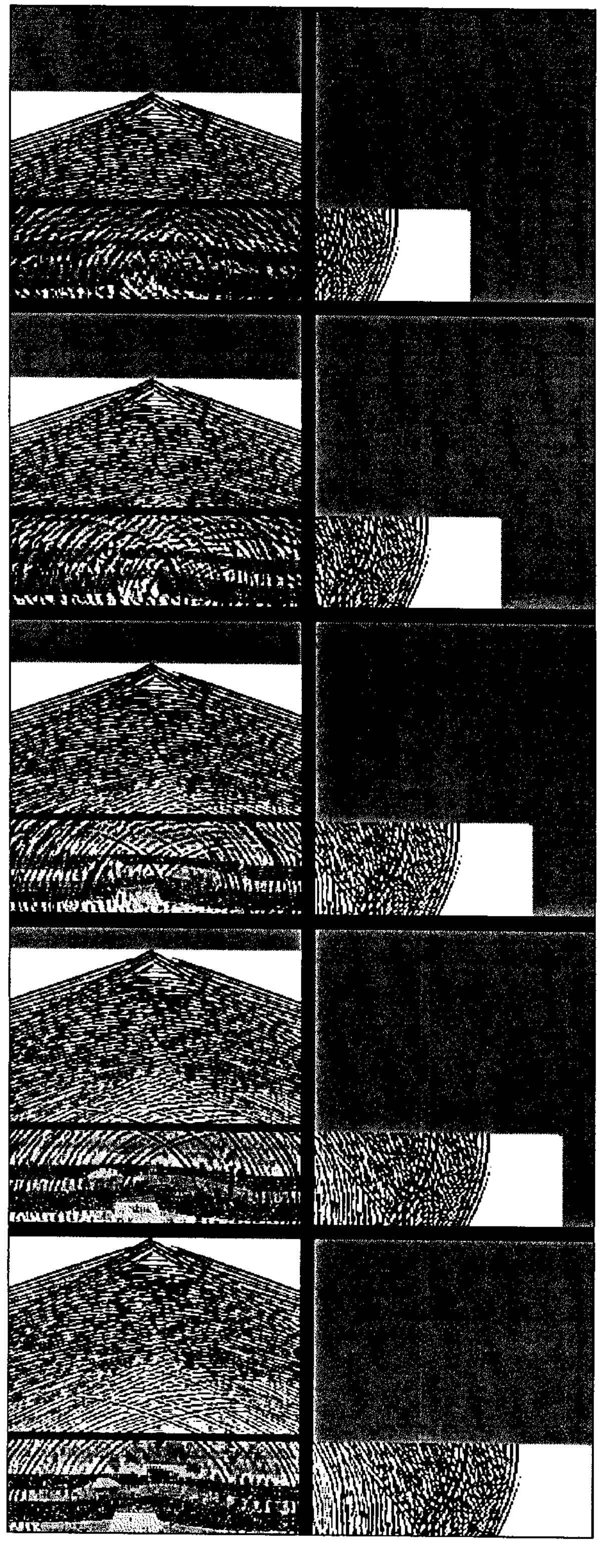

FIG. 2. (b). Snapshot frames every .33 seconds of an elastic wave simulation from t=1.67 secs at the top to t=3.00 secs at the bottom. Refer to Figure 1 for a description of a frame.

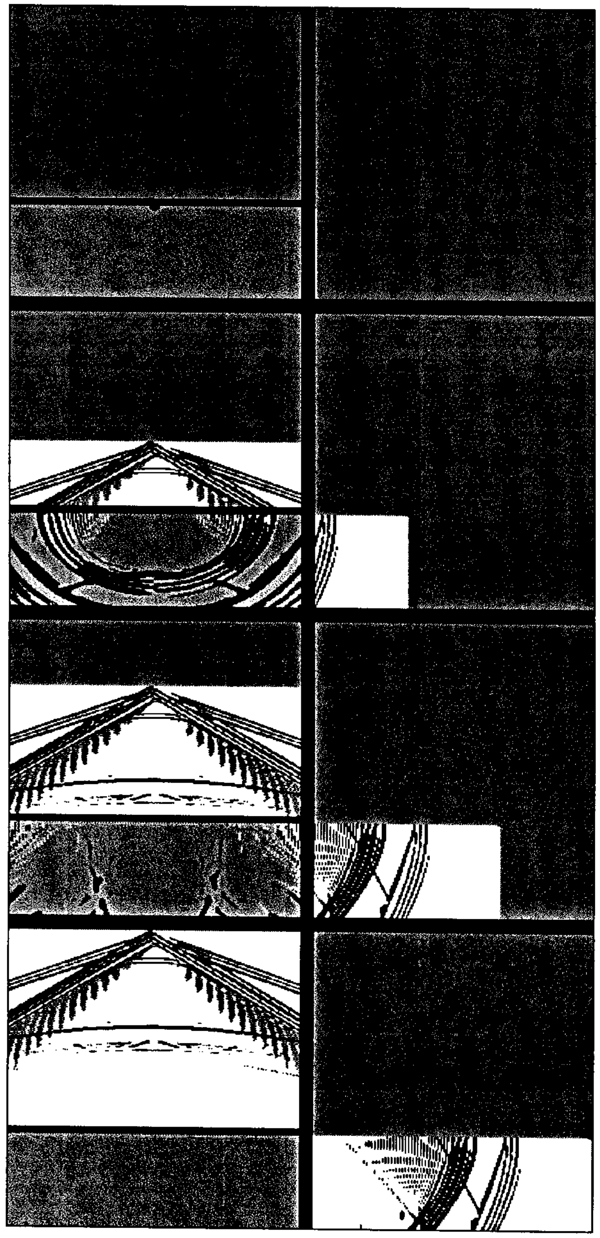

FIG. 3. Snapshots showing waves propagating through the initial velocity model to compute the background wavefield and synthetic data at iteration one. Frames are plotted every 1.0 second from t=0.0 at the top to t=3.0 at the bottom.

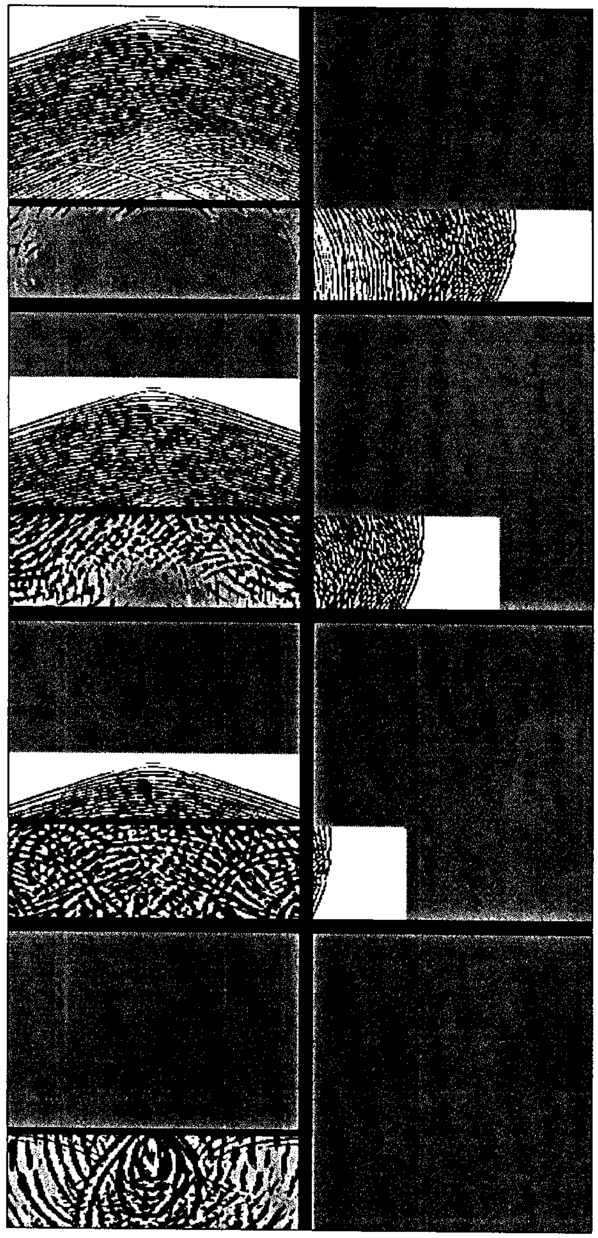

FIG. 4. Snapshots showing the computation of the back propagated residual wavefield at iteration one. Frames are plotted every 1.0 second from t=3.0 at the top to t=0.0 at the bottom. The data residuals are used as a forcing function in reverse time.

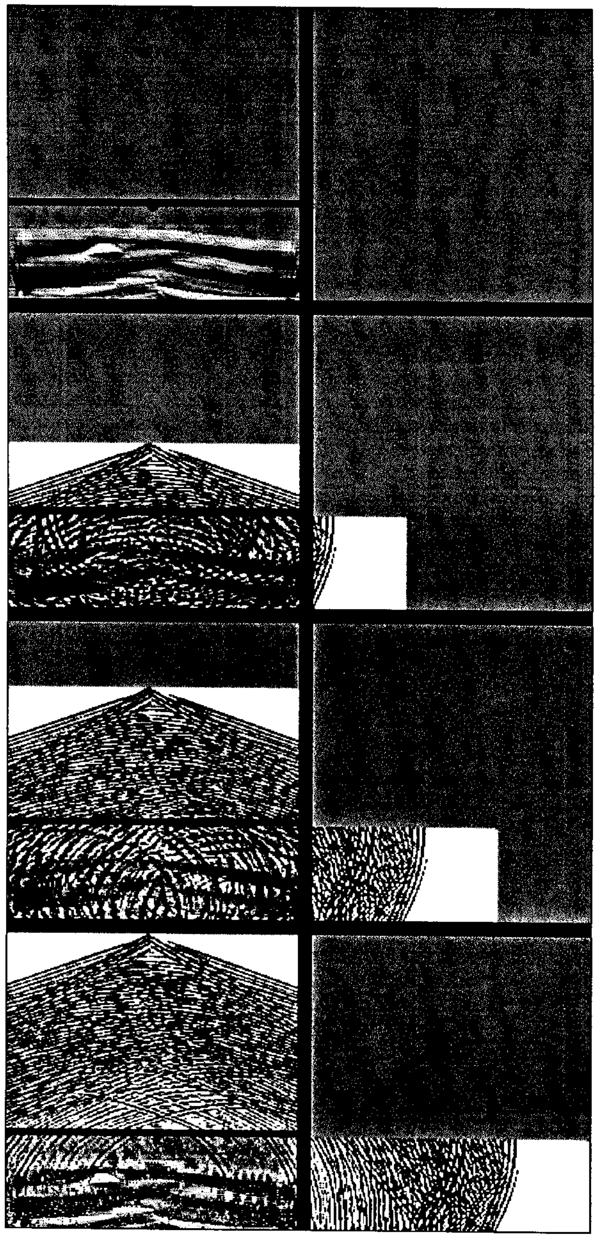

FIG. 5. Snapshots showing waves propagating through the ten iteration velocity model result to compute the synthetic data at iteration ten. Frames are plotted every 1.0 second from t=0.0 at the top to t=3.0 at the bottom. This synthetic data matches well with the data being inverted shown in Figure 2 and the ten iteration velocity model looks like the true model of Figure 2 so the inversion was a success.