Multi-channel inversion
(A draft for thesis chapter 3)

Shuki Ronen

ABSTRACT

Pre stack partial migration and interpolation of missing data can be combined into one process, multi-channel inversion, based on the wave equation and Fourier analysis of aliasing. The process has two functions: (1) Interpolation: finding the model which best fits the aliased data. (2) Offset extrapolation: the model is the ideal zero offset section, the data are collected with finite offset between shot and receiver.

I review the formulation and present results of multi-channel inversion applied to field data, using samples from 2-D data to simulate the cross-line direction in 3-D. A specific design of a 3-D experiment, suitable to multi-channel inversion, is proposed.

The inversion is done with conjugate-gradient, an iterative method which in this case converges in few iterations. The first iteration is equivalent to DMO-stacking with zero data in place of missing data.

INTRODUCTION

Sampling and aliasing

Reflection seismologists use a wavefield they sample on the surface of the earth to image the earth's interior. Their methods require that the wavefield is adequately sampled. Adequate sampling, according to Fourier analysis, is at least two sampling points per the smallest wavelength. If this sampling condition is violated, events may appear as other events (alias): high frequencies alias as low frequencies (Figure 1), steep dips may alias as flat events (Figure 2).

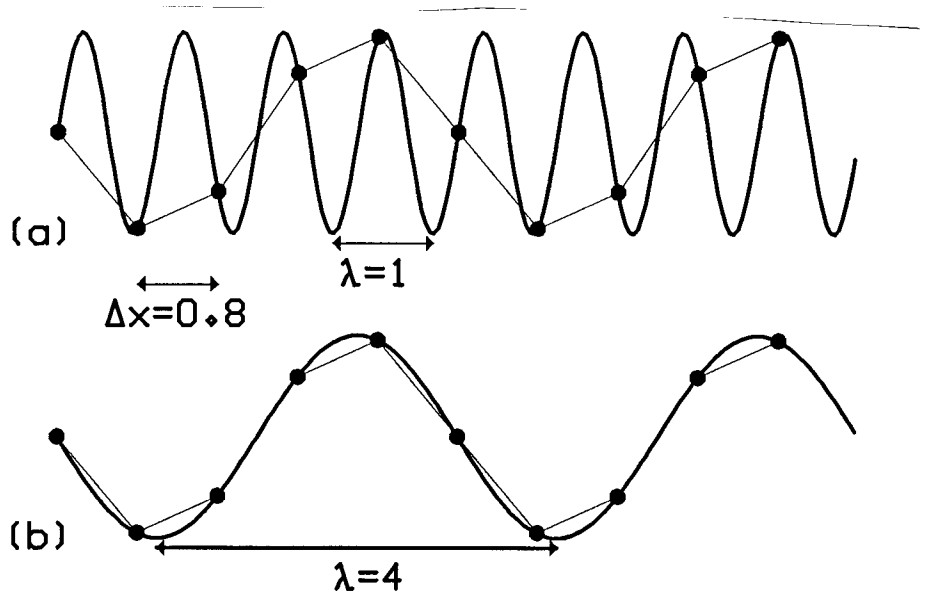


FIG. 1. Aliasing in one dimension. Samples from a high frequency signal (a) cannot be distinguished from samples from a low frequency signal (b). The high frequency (λ_a =1) aliases as low frequency (λ_b =4) because the sampling interval (Δx =0.8) is bigger than half the wavelength (λ_a /2=0.5).

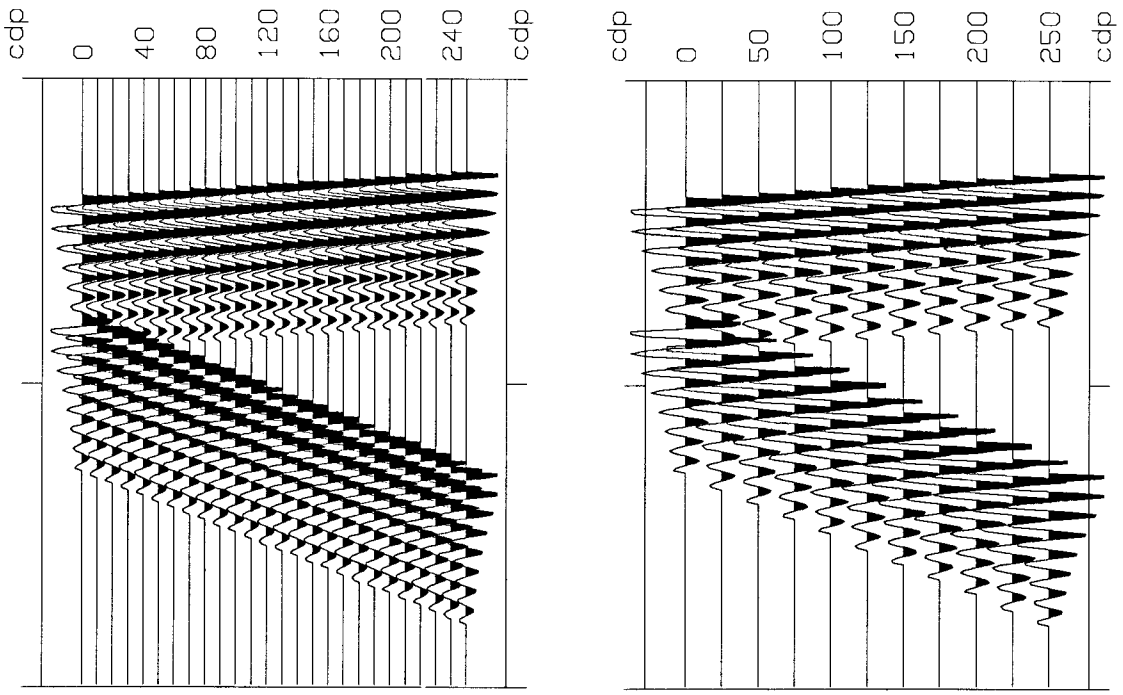


FIG. 2. Spatial aliasing. Left: adequate sampling. Right: spatial aliasing, the dipping events alias as short segments of almost flat events.

The seismic wavefield is sampled in time and space. Sampling in time is never a problem: improving the temporal sampling can be achieved by using faster analog to digital converters which are available and cheap. Not so in space, improving the spatial sampling may be achieved by having more recording and source stations, but this may be an economical problem, especially in 3-D surveys.

Suppose the data have frequencies up to 50 Hz, and half the velocity (half velocity, because it is a reflected wavefield, to be migrated with the exploding reflectors model), is 1000 m per second, so the minimum wavelength is $1000/50 = 20$ m. The adequate sampling interval is 10 m or less. Small sampling intervals like this can be achieved in two dimensional surveys (mostly marine), but for three dimensional surveys it is usually impossible to find a reasonable compromise between recording cost and spatial aliasing; the sampling is not uniform, in general, and midpoint intervals may reach 100 to 200 m in the cross-line direction. Interpolation methods are used to fill in the missing data in between the sampling points.

Missing data or too much data?

On one hand, spatial aliasing is a missing data problem: we need more data to adequately sample the seismic wavefield. On the other hand, it seems that we have too much data: all the offsets in a common midpoint gather, (after NMO and DMO), describe the same reflection point, and they are summed in the stacking process.

In stacking, with or without pre-stack full or partial migration, we loose a good opportunity to overcome spatial aliasing. The pre-stack data show the same earth as seen with different source and receiver locations. Each shot profile or common-offset section may be migrated and provide a partial image of the earth's interior. Stacking is the summation of all those images, but summing is not necessarily the correct way to combine multi-channel information. In a previous report (SEP-42) I described how linearized inversion can be used to overcome spatial aliasing. In this report I show the application of multi-channel inversion to field data.

FORMULATION

In 2-D (review)

The relation between aliased pre-stack data, $\mathbf{d}$, and the ideal zero-offset model, $\mathbf{m}$, (Ronen, 1985), is

$$\mathbf{d}_h(k) = \sum_n \mathbf{D}_h^+(k - n\kappa)\mathbf{m}(k - n\kappa) \ . \tag{1}$$

The nt-long vector, $\mathbf{d}_h(k)$, is a common-offset section at half offset h, in the (k,t) domain, (Fourier transformed over space but not over time). nt is the number of samples in a trace. k is the spatial frequency. The $n\,\omega$-long vector, $\mathbf{m}(k)$, is the ideal zero-offset section, in the (k,ω) domain, (Fourier transformed over space and time). κ is $2\pi/\Delta x$. Δx is the midpoint interval on a common-offset section. The summation over n has one term ($n=0$) if the sampling is adequate, two terms if the sampling interval is bigger than half the wavelength but smaller than the wavelength, and so on: if the sampling interval is 150 m and the minimal wavelength is 25 m, there are 12 terms in the summation. $\mathbf{D}_h^+$ is the matrix,

$$\left[\mathbf{D}_h^+(k)\right]_{\omega,t} = A^{-1}e^{-i\omega At}\ , \tag{2}$$

where,

$$A^2 = 1 + \left(\frac{hk}{\omega t}\right)^2 . \tag{3}$$

Equation (1) is based on Fourier analysis of sampling (via the aliasing summation, $\sum_n$) and on the wave equation (via the operator $\mathbf{D}_h^+$).

If we have five offsets (CDP fold = 5), and the ratio between the sampling interval to half the wavelength (aliasing fold) is 3, the system of equations we need to solve is,

$$\begin{pmatrix} \mathbf{d}_{h_1}(k) \\ \mathbf{d}_{h_2}(k) \\ \mathbf{d}_{h_3}(k) \\ \mathbf{d}_{h_4}(k) \\ \mathbf{d}_{h_5}(k) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{h_1}(k-\kappa) & \mathbf{D}_{h_1}(k) & \mathbf{D}_{h_1}(k+\kappa) \\ \mathbf{D}_{h_2}(k-\kappa) & \mathbf{D}_{h_2}(k) & \mathbf{D}_{h_2}(k+\kappa) \\ \mathbf{D}_{h_3}(k-\kappa) & \mathbf{D}_{h_3}(k) & \mathbf{D}_{h_3}(k+\kappa) \\ \mathbf{D}_{h_4}(k-\kappa) & \mathbf{D}_{h_4}(k) & \mathbf{D}_{h_4}(k+\kappa) \\ \mathbf{D}_{h_5}(k-\kappa) & \mathbf{D}_{h_5}(k) & \mathbf{D}_{h_5}(k+\kappa) \end{pmatrix} \begin{pmatrix} \mathbf{m}(k-\kappa) \\ \mathbf{m}(k) \\ \mathbf{m}(k+\kappa) \end{pmatrix} \tag{4}$$

The matrix has 5 by 3 blocks, every block, $\mathbf{D}_{h_j}$, is a nt by $n\,\omega$ matrix. Each $\mathbf{d}_{h_j}$ is a nt-long vector. Every $\mathbf{m}(k-n\,\kappa)$ is a $n\,\omega$-long vector. There is a system like equation (4) for every spatial frequency $-\kappa/2 < k < \kappa/2$. We can solve equation (4) and find the well sampled zero-offset section $\mathbf{m}(k)$ for $-3\kappa/2 < k < 3\kappa/2$.

In 3-D

In 3-D, there are two spatial axis, x and y. The half offset, $\mathbf{h}$, and the spatial frequency, $\mathbf{k}$, are the vectors,

$$\mathbf{h} = \begin{pmatrix} h_x \\ h_y \end{pmatrix}, \quad \text{and} \quad \mathbf{k} = \begin{pmatrix} k_x \\ k_y \end{pmatrix} .$$

Aliasing may occur in both x and y directions, so the 3-D equivalent of equation (1) is,

$$\mathbf{d_h}(k_x,k_y) = \sum_{n_x}\sum_{n_y}\mathbf{D_h^+}(k_x - n_x\,\kappa_x, k_y - n_y\,\kappa_y)\mathbf{m}(k_x - n_x\,\kappa_x, k_y - n_y\,\kappa_y) \quad . \tag{5}$$

κ_x is $2\pi/\Delta x$. Δx is the sampling interval in x. κ_y is $2\pi/\Delta y$. Δy is the sampling interval in y. $\mathbf{D_h^+}$ is as in equation (2), only now,

$$A^2 = 1 + \left(\frac{\mathbf{h}\cdot\mathbf{k}}{\omega t}\right)^2 . \tag{6}$$

$\mathbf{h}\cdot\mathbf{k}$ replaced the hk in equation (3). The 2-D case is a special case of the 3-D case, where the offset vector is in the x direction (in-line).

Matrix relation, analog to equation (4) can be formulated for the 3-D case. The number of block rows is the CDP fold. The number of block columns is the aliasing fold in x times the aliasing fold in y.

IN AND OUT OF DMO SPACE

A pessimist's view

The cost of matrix inversion is proportional to the cube of the size of the matrix. In SEP-42, I used Gaussian elimination to solve equation (4). It ran overnight for a synthetic example with nt =128 sapling points in time, CDP fold of 5, aliasing fold of 8, covering about half a mile (nx =32 midpoints). Field data size is about 10 times in both time and space. Since the cost is proportional to $nx \times nt^3$, the estimated run time is measured in 10^4 nights (many years) on the SEP's VAX.

An optimist's view

Jon Claerbout suggested I use conjugate gradient. Conjugate gradient is an iterative method. Each iteration involves the application of the forward operator, and the application of its transpose. As I noted in SEP-42, application of the transpose of the matrix of equation (4) is equivalent to DMO-stacking. Application of the forward matrix is transpose DMO-stacking, a' la "What is the transpose?" (Claerbout, 1985). The forward operator is going out of DMO space, its transpose, into DMO space.

Conjugate gradient inversion of an $N \times N$ matrix, converges within N iterations (with infinite machine precision). Often, only few components of the solutions are well determined and only few iterations provide a reasonable solution. For the multi-channel inversion, I found that about four iterations were enough. If each iteration of the conjugate gradient program costs like doing DMO twice, the total cost is like eight DMOs.

Actually, the cost may be less because Fourier transforms are not repeated and operators may be re-used.

Application of the matrix in equation (4) is still a costly operation, proportional to nt^2. Using this matrix is equivalent to using Dave Hale's DMO method (1984). There are other DMO methods whose cost is proportional to $nt \log nt$ or even to nt. The alternative methods involve approximation in one way or another, which may be perfectly acceptable in practice, but may complicate this paper's study of the multi-channel inversion For now, I chose to use matrix multiplication for DMO and transpose-DMO, but in practice more efficient methods can be used. Strong candidates are finite differencing (Bolondi et. al., 1983) for 2-D data, and summation ("Kirchhoff") methods for 3-D data.

In my case, not only the computational cost was in applying DMO and transpose DMO, the programming cost, as well, was limited to those routines, as I used a program written by Paige and Saunders (1982), based on an algorithm by Golub and Kahan, to drive the inversion.

APPLICATION IN 3-D

Cross-line aliasing

Assume that a (marine) 3-D survey is done by performing many 2-D surveys in parallel. The line spacing, Δy, is too big, so there is a spatial aliasing problem in the cross-line direction. The in-line midpoint interval, Δx, is small enough to avoid in-line aliasing: the summation over n_x in equation (5) reduces to the $n_x = 0$ term, and the relation (5) becomes,

$$\mathbf{d_h}(k_x, k_y) = \sum_{n_y} \mathbf{D_h^+}(k_x, k_y - n_y \kappa_y)\mathbf{m}(k_x, k_y - n_y \kappa_y) \quad , \tag{7}$$

with $n_x = 0$ inside A,

$$A^2 = 1 + \left(\frac{h_x k_x + h_y k_y - h_y n_y \kappa_y}{\omega t} \right)^2 . \tag{8}$$

Note that n_y is always multiplied by h_y, the cross-line component of the offset.

Cross-line offset

If each of the 2-D lines in our 3-D survey is a regular 2-D line in which all shots and receivers are on the same line (x axis), then the offset vectors are always in the x direction, $h_y = 0$ in equation (8), and A becomes,

$$A^2 = 1 + \left(\frac{h_x \, k_x}{\omega t} \right)^2 . \tag{9}$$

A , and consequently $\mathbf{D_h^+}$ are independent of n_y ,

$$\mathbf{D_h^+}(k_x, k_y - n_y \, \kappa_y) = \mathbf{D_h^+}(k_x, k_y) \quad , \tag{10}$$

and can be moved out of the summation in equation (7),

$$\mathbf{d_h}(k_x, k_y) = \mathbf{D_h^+}(k_x, k_y) \sum_{n_y} \mathbf{m}(k_x, k_y - n_y \, \kappa_y) \quad . \tag{11}$$

In a matrix form,

$$\begin{pmatrix} \mathbf{d_{h_1}}(k_x, k_y) \\ \cdot \\ \mathbf{d_{h_j}}(k_x, k_y) \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \mathbf{D_{h_1}}(k_x, k_y) & \mathbf{D_{h_1}}(k_x, k_y) & \mathbf{D_{h_1}}(k_x, k_y) \\ \cdot & \cdot & \cdot \\ \mathbf{D_{h_j}}(k_x, k_y) & \mathbf{D_{h_j}}(k_x, k_y) & \mathbf{D_{h_j}}(k_x, k_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \mathbf{m}(k_x, k_y - \kappa_y) \\ \mathbf{m}(k_x, k_y) \\ \mathbf{m}(k_x, k_y + \kappa_y) \end{pmatrix} . \tag{12}$$

The block columns of the matrix are all the same and the system (12) is ill conditioned.

Cross-line offset can be obtained by the recording geometry shown in Figure 3: two boats in parallel, both shoot, alternatively, but only boat 1 records. Boat 1 produces in-line offsets: $h_y = 0$, and h_x varies from the near to the far half offset. Boat 2 produces cross-line offsets: $h_y = 500$ m (if the distance between the boats is 1000 m), and h_x varies along the cable. The cross line offset is odd number times the line spacing, (5 in Figure 3), so the cross line midpoints fall between recording lines.

The matrix relation for this experiment is

$$\begin{pmatrix} \mathbf{d_{h_1}}(\mathbf{k}) \\ \cdot \\ \cdot \\ \mathbf{d_{h_{J/2}}}(\mathbf{k}) \\ \hline \mathbf{d_{h_{J/2+1}}}(\mathbf{k}) \\ \cdot \\ \cdot \\ \mathbf{d_{h_J}}(\mathbf{k}) \end{pmatrix} = \begin{pmatrix} \mathbf{D_{h_1}}(k_x, k_y) & \mathbf{D_{h_1}}(k_x, k_y) & \mathbf{D_{h_1}}(k_x, k_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{D_{h_{J/2}}}(k_x, k_y) & \mathbf{D_{h_{J/2}}}(k_x, k_y) & \mathbf{D_{h_{J/2}}}(k_x, k_y) \\ \hline \mathbf{D_{h_{J/2+1}}}(k_x, k_y - \kappa_y) & \mathbf{D_{h_{J/2+1}}}(k_x, k_y) & \mathbf{D_{h_{J/2+1}}}(k_x, k_y + \kappa_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{D_{h_J}}(k_x, k_y - \kappa_y) & \mathbf{D_{h_J}}(k_x, k_y) & \mathbf{D_{h_J}}(k_x, k_y + \kappa_y) \end{pmatrix} \begin{pmatrix} \mathbf{m}(k_x, k_y - \kappa_y) \\ \mathbf{m}(k_x, k_y) \\ \mathbf{m}(k_x, k_y + \kappa_y) \end{pmatrix} . \tag{13}$$

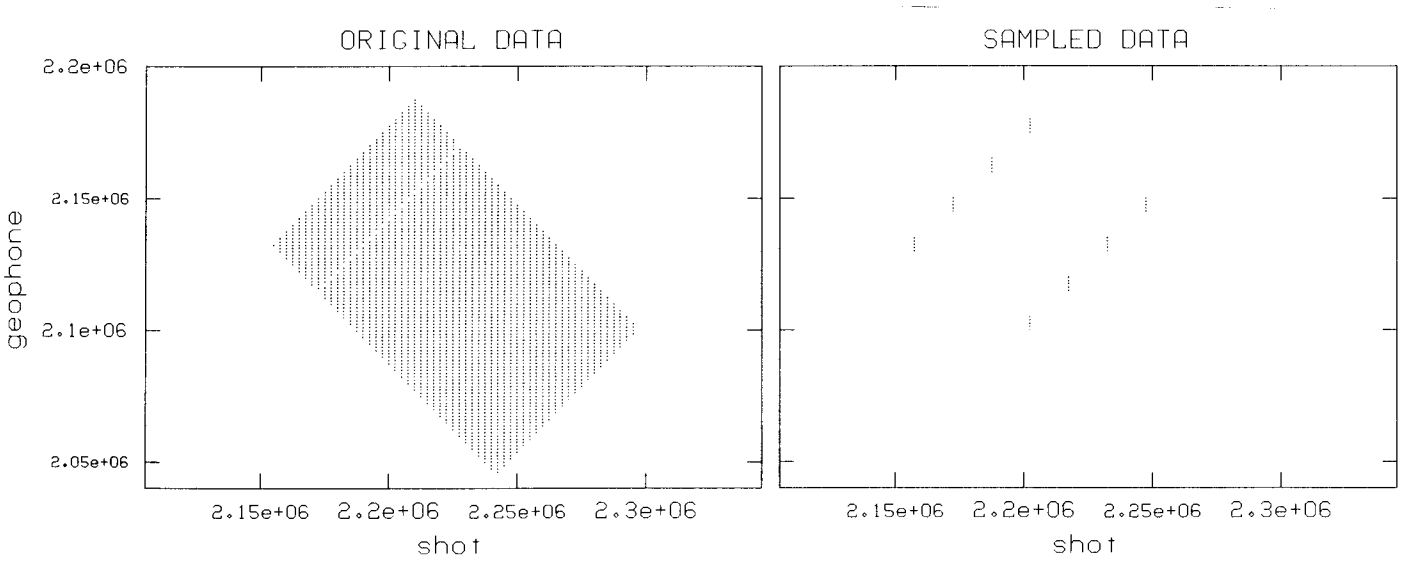
3-D simulated by 2-D

The case, k_x =0, in equation (13) is important for reflectors which do not dip in the in-line direction (they may dip in the cross-line direction). Since h_x is always multiplied by k_x in equation (14), the upper $J/2$ block rows of the matrix of equation (13) are all the same, because they may be different only in their h_x. Similarly, the lower $J/2$ block rows are all the same. The condition of the matrix, for the k_x =0 case is as if we perform a 2-D experiment, using only two offsets: zero and the cross line offset.

The cross line aliasing of the experiment of Figure 3, with five alternating sources per boat, can be simulated by taking a shot interval of 200 m, and ten offsets at 0,±12.5,±25, 975,987.5,1000,1012.5 and 1025 m on each shot profile.

Field data example

A small part of a marine line was the data base. The original data do not have aliasing problem: shot interval of 25 m and 240 channels, 12.5 m apart, provide a 6.25 m midpoint interval on the CDP stack. A near offset section (midpoint interval of 25 m) is shown in Figure 5a.



(a) (b)

FIG. 4. (a) Stacking chart of the original data: 25 m shot interval, 12.5 m channel interval. One offset is missing because of a dead hydrophone-group. (b) Stacking chart of the sampled data that simulates the cross line direction in the experiment of Figure 3. 150 m shot interval, 10 channels per shot in two groups of 5, 750 m apart.

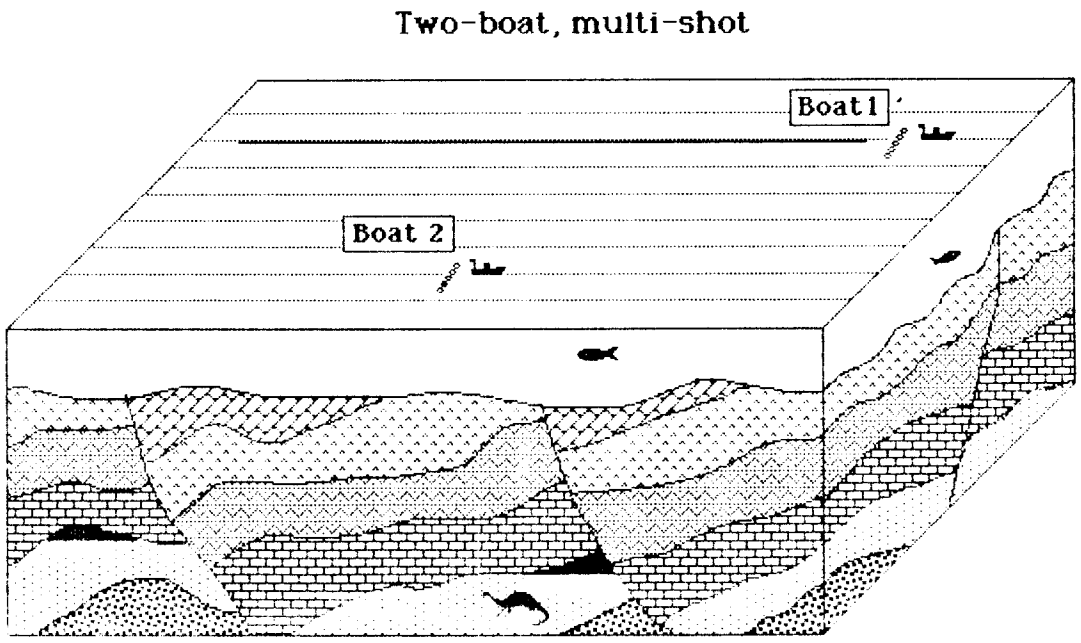


FIG. 3. Two boat recording geometry. One cable, alternating shots. The line spacing in fifth the distance between the boats.

Each block, $\mathbf{D}_{\mathbf{h}_j}(k_x, k_y - n_y \kappa_y)$, depends only on its n_y, where n_y is 1,0, or -1. The dependency is through A,

$$A^2 = 1 + \left(\frac{h_x k_x + h_y k_y - h_y n_y \kappa_y}{\omega t} \right)^2 . \tag{14}$$

The columns are independent because they have different n_y's and $h_y \neq 0$ in some of their elements. The upper $J/2$ block rows in the matrix of equation (13) come from the in-line data, shot and recorded from boat 1, $h_y = 0$ in all of them, but h_x varies from row to row. The lower $J/2$ rows come from the cross-line data, shot from boat 2 and recorded from boat 1, $h_y = 500$ in all of them and h_x varies from row to row. Without boat 2, we would not have the lower $J/2$ equations and the matrix would be ill conditioned because its block columns would all be the same.

Additional independent channels can be obtained if each boat would use alternating multiple sources as shown in Figure 3: each boat tows a linear array of sources, for example, in 5 groups, 12.5 m apart, extending to 25 m on each side of the boat. In every shot point, only one group fires. This generates a range of cross-line offsets. If the distance between boat 2 to the middle of the cable is 1000 m, the following cross-line offsets are recorded: -25,-12.5,0,12.5 and 25 m - from boat 1, 975,987.5,1000,1012.5 and 1025 m - from boat 2.

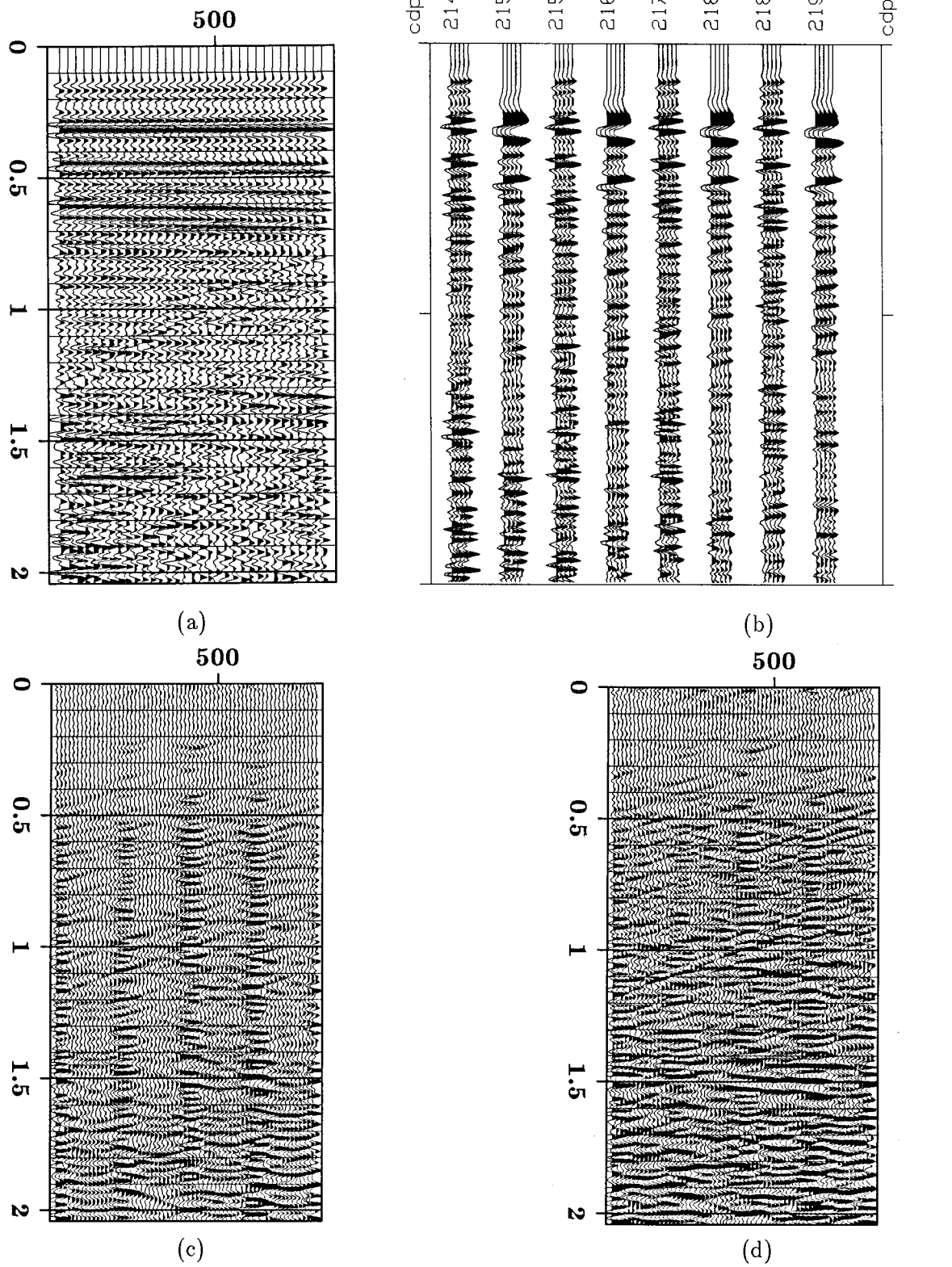


FIG. 5. (a) Approximately the model: near offset section with 25 m sampling interval. (b) The data: the coverage is not uniform, every trace is NMOed and plotted in its midpoint. (c) First iteration: this would be the result of DMO stacking. (d) Fourth iteration: flat and dipping reflectors are well interpolated.

The reason I chose 240 channel data was the 12.5 m group interval that enabled simulation of the alternating multiple shots of Figure 3. I simulated five alternate sources, 12.5 m apart, line spacing of 150 m, distance between boats: 750 m. Stacking chart of the data used is shown in Figure 4.

A familiar problem in 3-D is that the coverage is not uniform, this is also the case in this simulation (Figure 5b). Non uniform coverage is natural in the data space, it is the model, not the data, which should be uniformly and adequately sampled. The model obtained with conjugate-gradient multi-channel inversion is shown in Figure 5d. The multi-channel inversion made a substantial improvement, compared to DMO with zero traces in place of missing data (Figure 5c): the improvement is mainly with respect to the flat reflectors. There is no problem, of-course, in crossing dips.

CONCLUSIONS

Multi-channel inversion with conjugate gradient is applicable to field data. The results show a substantial improvement, from NMO-stacking, through the first iteration (DMO-stacking) to the last iteration (multi-channel inversion).

Cross-line offset in 3-D surveys is useful for the overcoming of spatial aliasing in the cross-line direction. The CDP coverage does not have to be uniform, the multi-channel inversion produces a uniformly sampled model out of non-uniform and aliased data.

REFERENCES

Bolondi G., Loinger E., and Rocca F., 1983, Offset continuation of seismic sections: Geophysical Prospecting, **30**, 813-828.

Claerbout, J.F., 1985, What is the transpose operation? SEP **42**, 113-128.

Hale, D., 1984, Dip move out by Fourier transform: Geophysics, **49**, 741-757.

Paige, C.C., and Saunders, M.A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software **8**, 195-209.

Ronen, J., 1985, Overcoming spatial aliasing in reflection seismology: SEP **42**, 281-329.

DATE: January 14, 1985

TO : Occupants of the Mitchell Building

FROM : Anita Bravman
Building Manager

SUBJECT: University No-Smoking Policy

As most of you probably know, the University has adopted a new no-smoking policy which became effective January 1, 1985. This policy prohibits smoking "in indoor locations where smokers and non-smokers occupy the same area," and will be enforced in the Mitchell Building.

In the Mitchell Building smoking is allowed in private offices (not in offices shared with non-smokers) when the doors are closed "as long as non-smokers are not exposed to second-hand or sidestream smoke--e.g., the ventilation system does not redistribute smoke, or smoke doesn't waft out of open doorways to nearby non-smokers." Smoking is not permitted in the third floor lobby or other common areas.

In an attempt to accommodate smokers in the Mitchell Building who do not have access to private offices, room B-17 has been designated as a "Smoking Permitted" area. This room is located on the basement level near the freight elevator and is used primarily for storage. However, it is the only available space in the building which complies with the guidelines established by the University policy.

Keys to this room may be checked out from the Dean's office. It is sometimes chilly at that end of the building so it would be advisable to have a sweater or light jacket with you if you plan to use the room.

Use of B-17 as a "Smoking Permitted" area will be initially for a trial period of time running from January 15, 1985, through February 28, 1985. If you use the room during that period, please be sure to indicate such use on the sign-in sheet which will be located in the room.

A copy of the policy is attached for those of you who have not yet seen it. Any questions should be directed to me or to Ruth Sloan in the Dean's office.

Anita Bravman

• STANFORD UNIVERSITY • OFFICE MEMORANDUM • STANFORD UNIVERSITY • OFFICE MEMORANDUM • STANFORD UNIVERSITY • OFFICE MEMORANDUM