Stanford Exploration Project, Report 108, April 29, 2001, pages 1–??

Adaptive multiple subtraction with non-stationary helical shaping filters

James Rickett, Antoine Guitton, and Douglas Gratwick[1]

ABSTRACT

We suppress surface-related multiples with a smart adaptive least-squares subtraction scheme in the time-space domain after modeling multiples with a fast but approximate modeling algorithm. The subtraction scheme is based on using a linear solver to estimate a damped non-stationary shaping filter. We improve convergence by preconditioning with a space-domain helical roughening filter.

INTRODUCTION

Both surface-related and internal multiples are a major source of coherent noise in many basins throughout the world. Dragoset and Jeričević (1998) describe a two-step multiple elimination process: prediction based on the Delft autoconvolution model (Verschuur et al., 1992), followed by subtraction with an adaptive noise cancellation (ANC) algorithm. However, practical multiple removal by adaptive noise cancellation is significantly more complex than the one-dimensional theory described in the text books [e.g. Widrow and Stearns (1985)]. In particular the multidimensional nature of prestack seismic data leads to ambiguity in the choice of potential parameters.

This paper describes an alternative to the conventional ANC algorithm also based on the minimum energy criterion. However, our approach leverages helical preconditioning to ensure the non-stationary shaping filters vary smoothly, and do not accidentally remove primary energy. In our description, we place particular emphasis on determining a robust set of parameters for the process.

THEORY

A canonical time series analysis problem (Robinson and Treitel, 1980) is that of shaping filter estimation: given an input time series $\mathbf{b}$, and a desired time series $\mathbf{d}$, we must compute a filter, $\mathbf{f}$, that minimizes the difference between $\mathbf{f}*\mathbf{b}$ and $\mathbf{d}$. Optimal filter theory provides the classical solution to the problem by finding the filter that minimizes the difference in a least-squared

[1]**email:** james@sep.stanford.edu, antoine@sep.stanford.edu, doug@sep.stanford.edu

sense, i.e. minimizing

$$\sum_i \left| \sum_j f_j \; b_{i-j} - d_i \right|^2 . \tag{1}$$

With the notation that **B** is the matrix representing convolution with time series **b**, we can rewrite this desired minimization as a fitting goal [e.g. Claerbout (1998a)],

$$\mathbf{B}\,\mathbf{f} - \mathbf{d} \approx \mathbf{0}, \tag{2}$$

which leads us to following system of normal equations for the optimal shaping filter:

$$\mathbf{B}^T \mathbf{B}\,\mathbf{f} = \mathbf{B}^T \, \mathbf{d}. \tag{3}$$

Equation (3) implies that the optimal shaping filter, **f**, is given by the cross-correlation of **b** with **d**, filtered by the inverse of the auto-correlation of **b**. The auto-correlation matrix, $\mathbf{B}^T\mathbf{B}$, has Toeplitz structure that can be inverted rapidly by Levinson recursion.

For the multiple suppression problem, the vector **d** represents the multiple infested raw data, and the matrix **B** represents convolution with the multiple model. Criterion (1) implies a choice of filter **f** that minimizes the energy in the dataset after multiple removal.

One advantage with working with time-domain filters as opposed to frequency-domain filters is that the theory can be adapted relatively easily to address non-stationarity. Following Claerbout (1998a) and Margrave (1998), we extend the concept of a filter to that of a non-stationary filter-bank, which in principle contains one filter for every point in the input/output space. For a non-stationary filter-bank, **f**, we identify $\mathbf{f}_j$ with the filter corresponding to the j^{th} location in the input/output vector, and the coefficient, $f_{i,j}$, with the i^{th} coefficient of the filter, $\mathbf{f}_j$. The response of non-stationary filtering with **f** to an impulse in the j^{th} location in the input is then $\mathbf{f}_j$.

With a non-stationary convolution filter, **f**, the shaping filter regression normal equations, are massively underdetermined since there is a potentially unique impulse response associated with every point in the dataspace. We need additional constraints to reduce the null space of the problem.

For most problems, we do not want the filter impulse responses to vary arbitrarily, we would rather only consider filters whose impulse response varies smoothly across the output space. This preconception can be expressed mathematically by saying that, simultaneously with expression (1), we would also like to minimize

$$\sum_i \left(\sum_j \left| \sum_k r_k \;\; f_{i,j-k} \right|^2 \right) \tag{4}$$

where the new filter, **r**, acts to roughen filter coefficients along the output axis of **f**.

Combining expressions (1) and (4) with a parameter, ϵ that describes their relative importance, we can write a pair of fitting goals

$$\mathbf{B}\mathbf{f} - \mathbf{d} \approx \mathbf{0}, \quad \text{and} \tag{5}$$
$$\epsilon\, \mathbf{R}\mathbf{f} \approx \mathbf{0}. \tag{6}$$

By making the change of variables, $\mathbf{q} = \mathbf{R}\mathbf{f} = \mathbf{S}^{-1}\mathbf{f}$ (Fomel, 1997), we obtain the following fitting goals

$$\mathbf{B}\mathbf{S}\mathbf{q} - \mathbf{d} \approx \mathbf{0}, \quad \text{and} \tag{7}$$
$$\epsilon\, \mathbf{q} \approx \mathbf{0}\,. \tag{8}$$

which are equivalent to the normal equations

$$\left(\mathbf{S}^T\, \mathbf{B}^T\, \mathbf{B}\mathbf{S} + \epsilon^2 \mathbf{I}\right)\, \mathbf{q} = \mathbf{S}^T\, \mathbf{B}^T\, \mathbf{d}. \tag{9}$$

Equation (9) describes a preconditioned linear system of equations, the solution to which converges rapidly under an iterative conjugate gradient solver.

APPLICATION TO A SYNTHETIC DATASET

We tested the algorithm on the 2-D BP multiple dataset which is based on a sub-salt play in the deep-water Gulf of Mexico.

Although the modeling algorithm is not the focus of this paper, we modeled surface-related multiples with a very fast 1-D algorithm. The multiple model was simply the multi-dimensional autoconvolution of common midpoint (CMP) gathers (Kelamis and Verschuur, 2000). This auto-convolution reduces to a multiplication in the *f-k* domain, and so it can be performed rapidly with multi-dimensional FFT's. More accurate multiple modelling algorithms will better attenuate multiples associated with the dipping salt-flanks. However, in 3-D examples, multiple prediction will always be imperfect, so we were interested in how this algorithm would adapt under less than ideal conditions.

Figures 1 and 2 show common-midpoint and common-offset sections before and after multiple suppression. With an imperfect multiple model, there is always a trade-off between suppressing multiples and preserving primary energy. For these results, we took a conservative approach - although some multiple energy remains in the data, hopefully all the primary energy remains too. In the areas with no salt present (e.g. `cmp_x`>10000 m), the multiples are almost entirely eliminated. However, in areas below the salt, especially where steeply-dipping diffracted multiples are present, some muliple energy remains.

PARAMETER CHOICES

Important practical considerations when solving equation (9) include the choice of ϵ, the number of conjugate-gradient iterations, the choice of roughening filter, and the number and position of adjustable coefficients in filter $\mathbf{f}$.

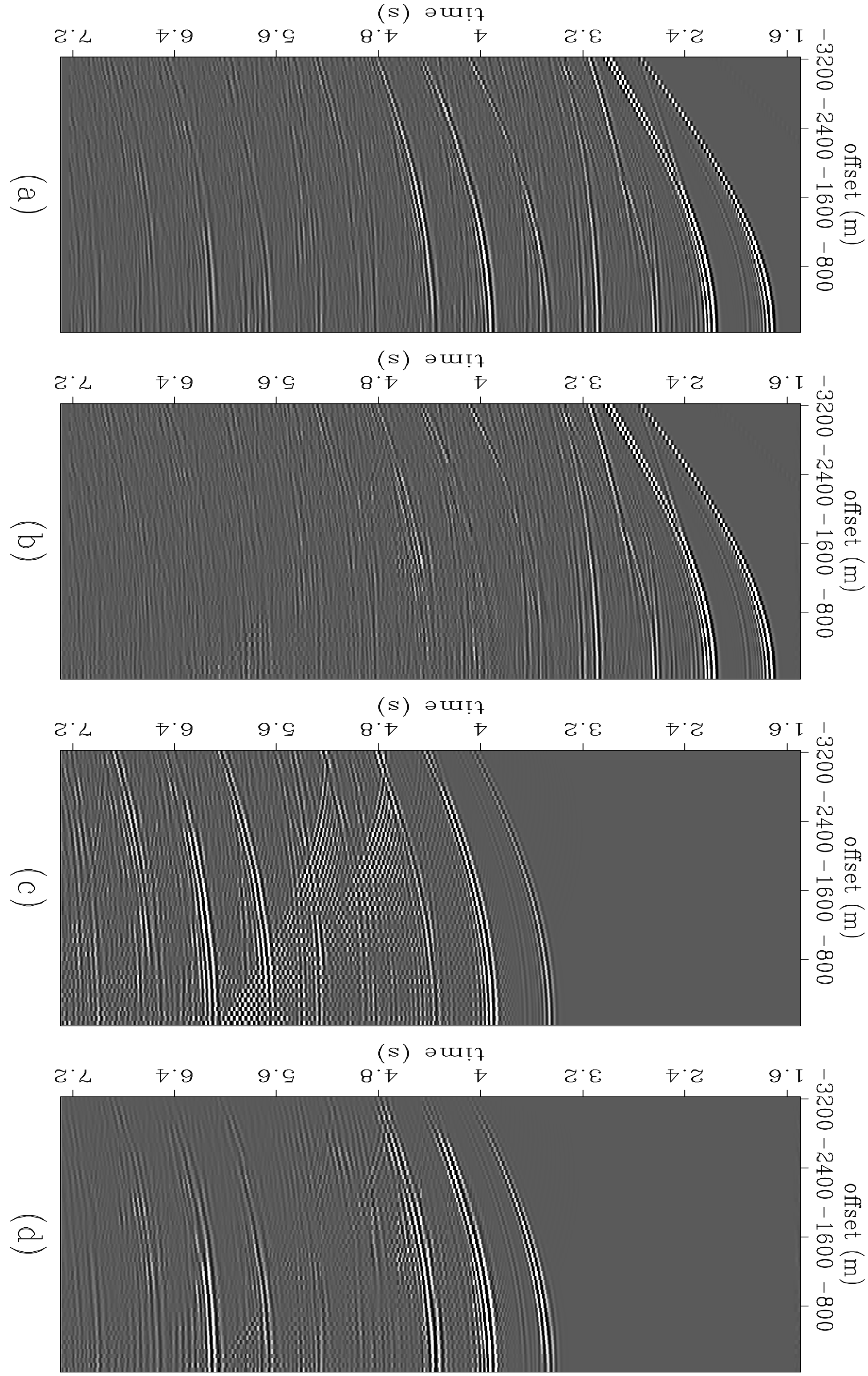

Figure 1: CMP gather with `cmp_x`=7000 m: (a) input gather, (b) gather after surface-multiple attenuation, (c) modeled muliples, and (d) multiples that were removed. james1-cmp7000 [CR,M]

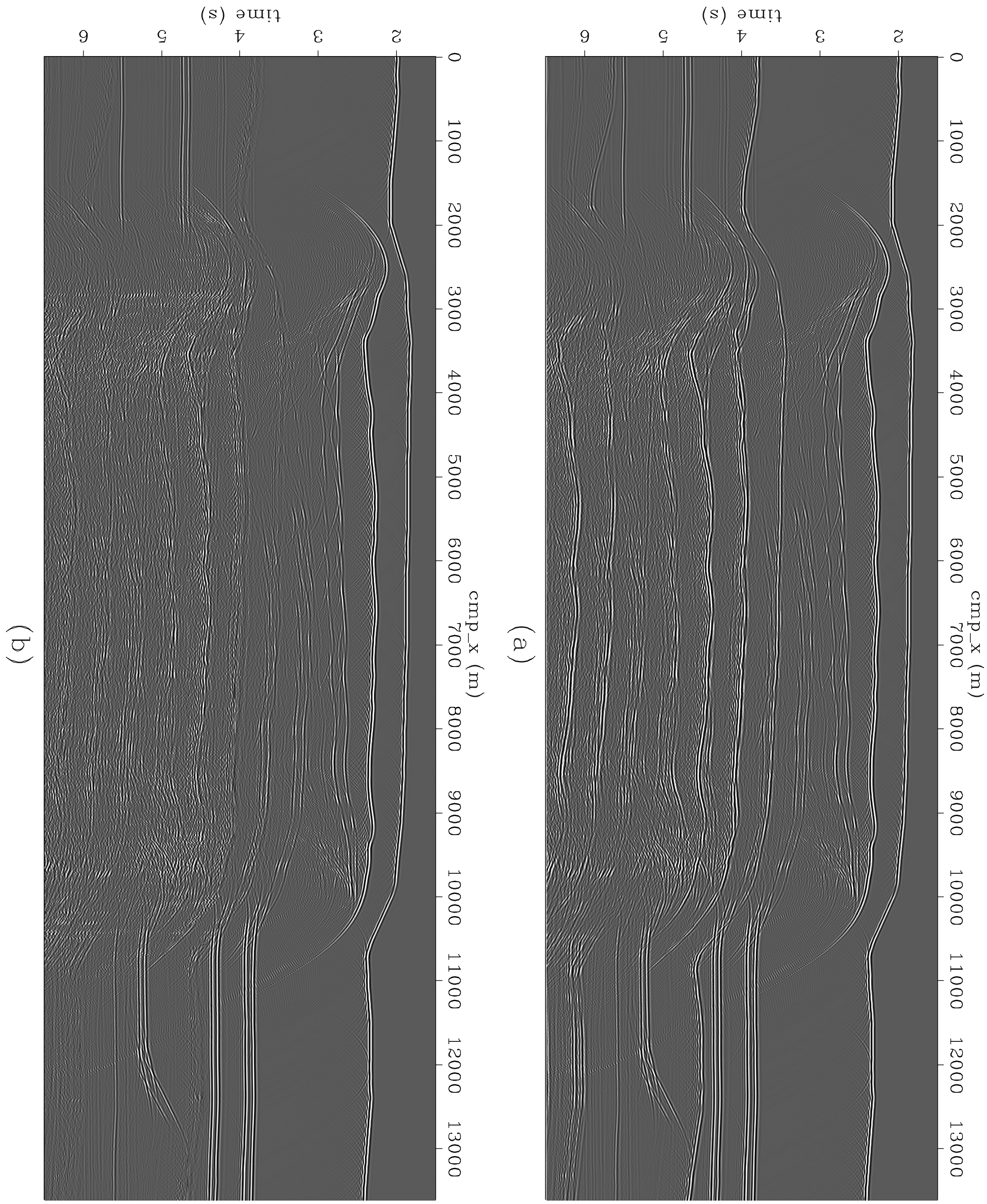

Figure 2: Common offset section with `offset`=1066 m: (a) input section, and (b) after multiple attenuation. james1-coff40 [CR,M]

Choice of ϵ and number of CG iterations

In principle, the number of iterations should not be an important parameter since we should iterate until the solution converges. However, determining the correct value of ϵ is a long-standing difficulty with large exploration-type geophysical inverse problems. Conventional solutions (Menke, 1989) such as picking the knee of misfit vs. model norm curves, or examining the singular-values of the operator matrix are not practical when the model-space is a large multi-dimensional image. If the choice of ϵ is too small, the solution will begin to degrade as the number of iterations increases as poorly resolved eigenvectors leak into the model space. On the other hand, if the choice of ϵ is too large, the solution will converge to a smooth model that does not satisfy our first fitting goal [expression (7)].

Despite these difficulties, with preconditioned problems we often obtain good results after only a few iterations without the solution fully converging, and with little or no dependence on the the choice of ϵ. Well-resolved low-frequency eigenvectors propagate into the solution quickly after only a few iterations.

Therefore, to reduce the dimensionality of the parameter space, we set $\epsilon = 0.$, and keep the filters smooth by restricting the number of conjugate-gradient iterations (Crawley, 1999). After solving the problem only once, we can plot misfit vs. model norm curves for intermediate solutions with varying number of iterations, and choose the best result.

Choice of roughening filter, R

The most important consideration in the choice of roughening filter is that it is easily invertible. A Fourier domain roughener would meet this criterion; however, we apply a time-space operator that is both cheaper, and less prone to Fourier artifacts such as wrap-around and Gibbs' phenomenom. Claerbout (1998b) describes how to construct invertible multi-dimensional time-space operators by applying helical boundary conditions to the problem. Helical operators cost $O(N)$ operations to apply and invert rather than $O(N \log N)$ for an equivalent Fourier operator.

For the results shown in this paper, we choose $\mathbf{R}$ to be the helical derivative operator that roughens isotropically in the midpoint-time plane. A cascade of two one-dimensional derivative filters first along the time axis and then along the midpoint axis also works well. Anisotropic smoothing can be controlled by tweaking the "micropatch" parameters described below.

Choice of non-stationary shaping-filter parameters

As described above, a non-stationary filter can have a different impulse response for each point in the input/output space. For the non-stationary shaping filter estimation procedure, we need to define which coefficients are adjustable for every individual impulse response. For simplicity, we characterize each impulse response with the same set of adjustable coefficients:

for the examples, shown here, they were one-dimensional non-causal symmetric filters about 180 ms long. Tests indicated that if the filters were shorter than the seismic wavelength, the quality of the results decreased. Increasing filter length beyond this length, however, did not alter results significantly, even in cases when the kinematics of the multiples were not accurately predicted.

When implementing the non-stationary filters, it is not strictly necessary to force each point of the input/output space to keep a unique impulse response. Rather, we apply the concept of "micropatches" (Crawley, 2000), in which points within a small neighbourhood share a single impulse response. This cuts computational memory requirements significantly, and provides an alternative method of controling spatial and temporal variabilty of impulse responses.

CONCLUSIONS

We describe a robust methodology for adaptive noise cancellation, based on non-stationary shaping filters and geophysical inverse theory. Helical preconditioning ensures the non-stationary shaping filters vary smoothly, preserving primary energy. Results on synthetic data show that the algorithm successfully attenuates well-predicted multiples, and to a lesser extent poorly-predicted multiples as well.

REFERENCES

Claerbout, J. F., 1998a, Geophysical estimation by example: `http://sepwww.stanford.edu/prof/gee/`.

Claerbout, J. F., 1998b, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 5, 1532–1541.

Crawley, S., 1999, Interpolating missing data: convergence and accuracy, t and f: SEP–**102**, 77–90.

Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. thesis, Stanford University.

Dragoset, W. H., and Jeričević, Ž., 1998, Some remarks on surface multiple attenuation: Geophysics, **63**, no. 02, 772–789.

Fomel, S., 1997, On model-space and data-space regularization: A tutorial: SEP–**94**, 141–164.

Kelamis, P. G., and Verschuur, D. J., 2000, Surface-related multiple elimination on land seismic data—Strategies via case studies: Geophysics, **65**, no. 3, 719–734.

Margrave, G. F., 1998, Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering: Geophysics, **63**, no. 1, 244–259.

Menke, W., 1989, Geophysical data analysis: Discrete inverse theory: Academic Press, San Diego, revised edition edition.

Robinson, E. A., and Treitel, S., 1980, Geophysical signal analysis: Prentice-Hall, Inc.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 9, 1166–1177.

Widrow, B., and Stearns, S. D., 1985, Adaptive signal processing: Prentice-Hall, Inc.