Differential semblance velocity analysis by wave-equation migration

Peng Shen and William W. Symes, Rice University, and Christiaan C. Stolk, École Polytechnique

Summary

Differential semblance measures the deviation from flatness or focus of image gathers. While many other functions have been suggested to measure gather quality, only differential semblance responds smoothly to velocity changes. Therefore gradient descent methods are uniquely attractive for velocity updating by differential semblance optimization. Because of their kinematic fidelity, wave equation (depth extrapolation) migration methods are natural platforms for velocity analysis in complex structure. This paper demonstrates an economical computation of the differential semblance gradient as an addendum to prestack depth extrapolation, and its use in constructing velocity updates by optimization.

Introduction

Flattening or focussing of prestack image gathers underlies many practical velocity analysis methods based on prestack data, from NMO velocity spectra to migrated-domain tomography. This principle fails, however, when strong refraction generates multiple ray paths from sources and/or receivers to scattering points. Multipathing in turn creates *imaging artifacts*, i.e. coherent energy not corresponding to actual reflectors, for many common imaging methods. As a result, the image gathers created by most common migration techniques are not flat even when the velocity is correct (Nolan and Symes, 1996; Nolan and Symes, 1997; Stolk, 2002; Stolk and Symes, 2002).

The present paper is motivated by the recent discovery (Stolk and De Hoop, 2001) that shot-geophone ("wave equation") migration based on the DSR equation (Claerbout, 1985) yields artifact-free image gathers, which are focussed even in the presence of severe multipathing. This fact guarantees that certain functions constructed from image gathers vary smoothly with velocity and attain their global minimum at kinematically correct models (Stolk and Symes, 2003). Minimization of these *differential semblance* ("DS") functions thus provide a systematic method for updating velocity models via gradient descent optimization methods (Symes, 1986; Symes and Carazzone, 1991; Symes, 1998; Chauris and Noble, 2001; Mulder and ten Kroode, 2002; Brandsberg-Dahl et al., 2003).

This paper gives precise definitions for several versions of differential semblance based on DSR migration, and describes an economical method for computing the DS gradient as a extension of the depth extrapolation process at modest additional cost. We have made a preliminary implementation of this computation and applied it to a strongly refracting synthetic example, the upper 0.6 km layer of the Marmousi model. We observe reasonable discrimination of lateral velocity changes and focussing of offset image gathers, as predicted by the theory.

Wave equation common image gathers, focussing, and differential semblance

All examples in this paper are 2D, so we shall adopt 2D notation, though the ideas are equally valid for 3D. Horizontal (or lateral) position will be denoted by x, or when it refers to a source or receiver by s, r respectively, depth by z. The medium velocity is $c = c(x, z)$.

Denote by $u(s, r, t, z)$ the downward continuation of the data $d(s, r, t)$. Provided that all significant energy in the data travels along rays which do not turn, the downward continued field satisfies the DSR wave equation (Claerbout, 1985):

$$\frac{\partial u}{\partial z} = F[c]u, \;\; u|_{z=0} = d \tag{1}$$

where $F[c]$ is (some approximation to) the DSR operator

$$F[c] = -\sqrt{c(s, z)^{-2}\partial_t^2 - \partial_s^2} - \sqrt{c(r, z)^{-2}\partial_t^2 - \partial_r^2}. \tag{2}$$

If the acquisition geometry is incomplete (sparsely sampled source-receiver pairs), as is usually the case, $d(s, r, t)$ is understood to be padded by zero traces to complete the sampling in s, r. Of course then the downward continued field cannot be the *physical* reflected field, as the necessary surface data is lacking: that is, it is not possible to literally "sink the survey". Instead, the DSR migration operator is the *adjoint* of the DSR (prestack) forward modeling operator; for more on this see (Stolk and De Hoop, 2001).

Denote half offset by h: then the prestack image volume I is related to the depth-propagated field u by

$$I(x, z, h) = u(x + h, x - h, 0, z)$$

The corresponding depth image is $I(x, z, 0)$, see (Claerbout, 1985). Other definitions parametrize the image by offset ray parameter p, using either the Radon transform in t, h before restriction to intercept time 0 (de Bruin et al., 1990; Prucha et al., 1999) or the Radon transform in z, h after restriction to $t = 0$ (Sava et al., 2001).

In terms of the image volume $I(x, z, h)$, the characteristic signature of kinematically correct velocity is *focussing* or concentration at $h = 0$, because that is the characteristic feature of the physical reflectivity in prestack DSR modeling, see (Stolk and De Hoop, 2001). If little energy is found away from $h = 0$, then $hI(x, z, h) \simeq 0$. Clearly the image volume $I(x, z, h)$ depends on the data $d(s, r, t)$

and the velocity model $c(x,z)$. Thus the mean square of $hI(x,z,h)$ is a function $J[c,d]$ of velocity and data:

$$J[c,d] = \frac{1}{2}\int \mathrm{d}x \int \mathrm{d}z \int \mathrm{d}h |hI(x,z,h)|^2$$

It can be shown that $J[c,d]$ responds smoothly to (large-scale) changes in velocity c and arbitrary finite energy data perturbations; it shares these characteristic features with other versions of differential semblance (Stolk and Symes, 2003). For data consistent with the velocity c, $J[c,d] \simeq 0$.

Numerical computation of the gradient

Equation (1) is solved numerically by approximating the DSR operator $F[c]$ using one of a wide variety of methods - see for example (Claerbout, 1985; Ristow and Ruhl, 1994; De Hoop et al., 2000). The result is a depth marching scheme of the form

$$u_{n+1} = u_n + \Delta z \Phi_n[c] u_n \tag{3}$$

in which $u_n = u_n(s,r,t) \simeq u(s,r,t,n\Delta z)$ is the downward continued field and $\Phi_n[c]$ is an extrapolation operator which derives from applying eg. a Runge-Kutta formula to a discretization of $F[c]$. [**NB:** it is conventional to carry out depth extrapolation in the frequency domain; we represent all operations in the time domain only for notational convenience.] Note that the depth step may be implicit; equation (3) represents implicit schemes as formally explicit. The depth extrapolation operator $\Phi_n[c]$ depends functionally on $c(x,z)$; the precise dependence depends on the choices of lateral and depth discretization, approximation to the DSR, and depth marching scheme.

After discretization,

$$J[c,d] = \frac{1}{2}\sum_n \sum_{x,h} |Pu_n|^2$$

in which P is the operator: *transform* $(s,r) \mapsto (x,h)$, *restrict to* $t=0$ *and multiply by* h.

The perturbation $\delta J[c,d]$ in $J[c,d]$ resulting from a perturbation δc in c is

$$\delta J[c,d] = \sum_n \sum_{x,h} (Pu_n)(P\delta u_n)$$

$$= \sum_n \sum_{s,r,t} (P^* P u_n)(\delta u_n)$$

in which P^* denotes the *adjoint operator* to P (namely: *multiply by* h *then by* $\delta(t)$, *transform* $(x,h) \mapsto (s,r)$) The perturbation field δu_n solves the linearization of (3):

$$\delta u_{n+1} = \delta u_n + \Delta z \Phi_n[c]\delta u_n + \delta\Phi_n[c] u_n \tag{4}$$

Define the *adjoint field* $w_n(s,r,t)$ as the solution of the *adjoint state equation*

$$w_{n-1} = w_n + \Delta z \Phi[c]^* w_n + P^* P u_n \tag{5}$$

which is to be solved in order of *descending* n (i.e. upward continuation) with initial condition $w_n = 0$ for large n, and the superscript $*$ denotes the *adjoint* or transpose operator. Then

$$\delta J[c,d] = \sum_n \sum_{s,r,t} (w_{n-1} - w_n - \Delta z \Phi[c]^* w_n)\delta u_n$$

$$= \sum_n \sum_{s,r,t} (w_n)(\delta u_{n+1} - \delta u_n - \Delta z \Phi[c]\delta u_n)$$

which is, from (4),

$$= \sum_n \sum_{s,r,t} w_n \delta\Phi_n[c] u_n$$

Now $\delta\Phi[c]u = \partial_c(\Phi[c]u)\delta c$, in which ∂_c denotes partial derivative with respect to c. Denote by $\Psi[c,u]$ the adjoint of the operator $\delta c \mapsto \partial_c(\Phi[c]u)\delta c$. Then

$$\delta J[c,d] = \sum_{x,z} (\sum_n \Psi[c,u_n] w_n)\delta c$$

from which we read off that the gradient of J is

$$\nabla J[c,d] = \sum_n \Psi[c,u_n] w_n$$

This method for computing the gradient of a function depending on the solution of a differential equation goes by the name *adjoint state method*, and originates in control theory (Lions, 1972). It was introduced into the theory of inverse problems by (Chavent and Lemonnier, 1974), and further developed by (Tarantola, 1987) amongst others.

To summarize, to compute the gradient:

1. Downward continue the wavefield (compute $u_n, n = 0,...,N$ by solving (3));
2. Upward continue the adjoint wavefield (compute $w_n, n = N,...0$ by solving (5));
3. Apply the operator $\Psi[c,u_n]$ to w_n, which amounts to a weighted crosscorrelation of u_n and w_n, and sum on n to obtain the gradient.

Note that the velocity should be confined to a space of smooth models, either by representation (basis functions) or by definition of the norm in velocity space. Either option must be taken into account in computing the adjoint operator Ψ. We have used a B-spline basis on a sparse set of nodes, with the Euclidean norm in the coefficients, in the examples below. The projection onto the B-spine basis effectively smooths the gradient, and is necessary for effective velocity updating.

Examples

This section presents a few results of an initial implementation, using Fourier Finite Difference operators (Ristow and Ruhl, 1994) to approximate $F[c]$.

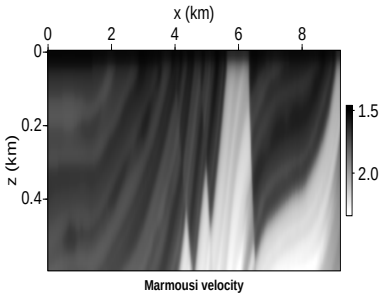

Fig. 1: The Marmousi velocity from surface to 0.6 km in depth.

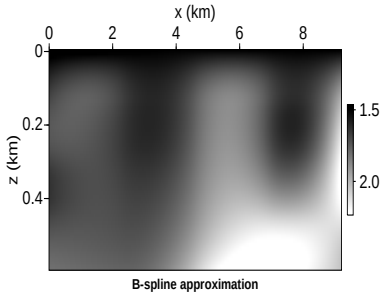

Fig. 2: The best B-spline approximation

The aim of this experiment is recovery of the velocity in the upper 0.6 km of the Marmousi velocity model (Bourgeois et al., 1991), see Figure 1. The data are the original Marmousi traces, with source signature deconvolution and continuation to zero offset applied. We adopted a 6×6 grid of B-spline nodes to represent velocities. The negative gradient at the initial constant (1.8 km/s) velocity model appears in Figure 3 (with the B-spline smoothing implicit in its definition, as explained above). While it lacks the shorter-scale lateral features of the B-spline projection of the velocity (Figure 2), it is certainly a constructive update direction.

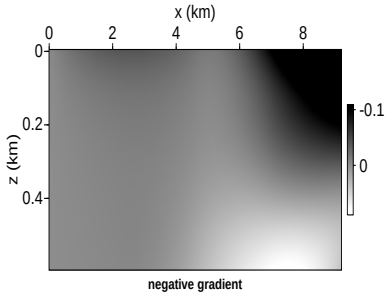

Fig. 3: Gradient at constant velocity of 1.8km/sec.

We used a limited memory BFGS quasi-Newton algorithm (Nocedal and Wright, 1999) to decrease the differential semblance objective $J[c, d]$. This algorithm combines an approximation of Newton's method to determine the search direction, and an efficient backtracking line search to determine the update. Each evaluation of the objective requires remigration. One or sometimes two migrations were required per BFGS iteration. The BFGS method reduced the objective value by roughly a factor of two, the length of the gradient by an order of magnitude in five iterations. The fifth velocity iterate appears as Figure . Most of the shorter-scale lateral features in the velocity have now appeared. Comparison of image gathers at initial vs. iteration 5 velocities shows considerable improvement in focussing (Figures 5, 6).

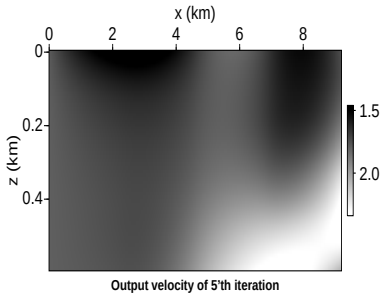

Fig. 4: Velocity estimate recovered after 5 BFGS iterations

Discussion

We have defined a version of differential semblance appropriate to shot-geophone migration by downward wavefield extrapolation ("wave equation migration"). We have also shown how the differential semblance gradient can be computed at modest additional expense. Preliminary results with synthetic data seem promising.

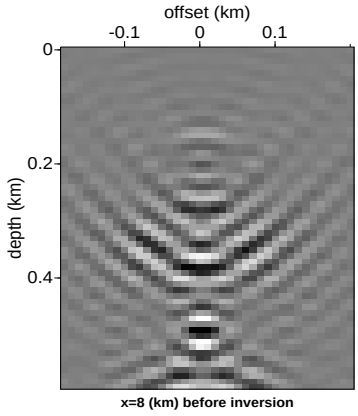

Fig. 5: offset image gather at x=8 km before inversion

Acknowledgement: This work was supported in part by the sponsors of The Rice Inversion Project. The authors are indebted to Scott Morton, Biondo Biondi, Paul Sava, Marteen deHoop and Bee Bednar for insightful conversations.

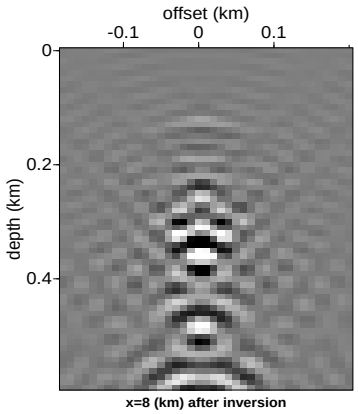

Fig. 6: offset image gather at x=8 km, 5th iteration

References

Bourgeois, A., Lailly, P., and Versteeg, R., 1991, The Marmousi model *in* Versteeg, R., and Grau, G., Eds., The Marmousi Experience:.

Brandsberg-Dahl, S., De Hoop, M., and Ursin, B., 2003, Focusing in dip and AVA compensation on scattering angle/azimuth common image gathers: Geophysics, **68**, 232–254.

Chauris, H., and Noble, M., 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications synthetic and real data sets: Geophys. J. Int., **144**, 14–26.

Chavent, G., and Lemonnier, P., 1974, Identification de la non-linéarité d'une équation parabolique quasi-linéaire: Appl. Math. and Opt., **1**.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications, Oxford.

de Bruin, C. G. M., Wapenaar, C. P. A., and Berkhout, A. J., 1990, Angle-dependent reflectivity by means of prestack migration: Geophysics, **55**, 1223–1234.

De Hoop, M., Le Rousseau, J., and Wu, R.-S., 2000, Generalization of the phase-screen approximation for the scattering of acoustic waves: Wave Motion, **31**, 43–70.

Lions, J., 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag, Berlin.

Mulder, W., and ten Kroode, A., 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191.

Nocedal, J., and Wright, S., 1999, Numerical optimization: Springer Verlag, New York.

Nolan, C. J., and Symes, W. W., 1996, Imaging and coherency in complex structure: 66th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 359–363.

Nolan, C., and Symes, W., 1997, Global solution of a linearized inverse problem for the wave equation: Comm. P. D. E., **22**, 919–952.

Prucha, M., Biondi, B., and Symes, W., 1999, Angle-domain common image gathers by wave-equation migration: 69th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 824–827.

Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration: Geophysics, **59**, 1882–1893.

Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 296–299.

Stolk, C. C., and De Hoop, M. V., December 2001, Seismic inverse scattering in the 'wave-equation' approach, Preprint 2001-047, The Mathematical Sciences Research Institute, http://msri.org/publications/preprints/2001.html.

Stolk, C., and Symes, W., 2002, Artifacts in Kirchhoff common image gathers: Expanded Abstracts, Society of Exploration Geophysicists, 72nd Annual International Meeting, SEG, 1129–1132.

Stolk, C. C., and Symes, W. W., 2003, Smooth objective functionals for seismic velocity inversion: Inverse Problems, **19**, 73–89.

Stolk, C. C., 2002, Microlocal analysis of the scattering angle transform: Comm. P. D. E., **27**, 1879–1900.

Symes, W., and Carazzone, J., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, no. 5, 654–663.

Symes, W., 1986, Stability and instability results for inverse problems in several-dimensional wave propagation *in* Glowinski, R., and Lions, J., Eds., Proc. 7th International Conference on Computing Methods in Applied Science and Engineering:: North-Holland.

Symes, W. W., 1998, High frequency asymptotics, differential semblance, and velocity analysis: 68th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1616–1619.

Tarantola, A., 1987, Inverse problem theory: Elsevier.