

DSR Migration Velocity Analysis by Differential Semblance Optimization

Alexandre Khoury, William Symes (Rice University), Paul Williamson, Peng Shen (Total E&P USA Inc.)

SUMMARY

The relative efficiency of Common Azimuth wave equation depth migration based on the double square root equation makes it an attractive engine for use in automated velocity updating schemes. One such automated velocity analysis procedure is Differential Semblance Optimization, which constructs velocity updates to reduce, and eventually minimize, a quantitative measure of image gather misfocussing. We review the construction of a differential semblance velocity analysis tool based on common azimuth migration, and present some 2D examples. These examples illustrate some of the promise, and some of the potential pitfalls, of this approach to velocity analysis.

MIGRATION IN DSR FORMULATION

Wave equation migration is based on the "survey sinking" idea (Claerbout, 1985). The seismic field $p(x_r, y_r, x_s, y_s, z, t)$ is a function of source and receiver positions (x_s, y_s) and (x_r, y_r) on a horizontal datum at depth z , and time t . It will be convenient to describe the field in terms of midpoint and (half-)offset vectors: $m_x = (x_r + x_s)/2$, $h_x = (x_r - x_s)/2$, similar expressions for y coordinates. The data are modeled as a sampling of this field for a depth near the surface. Assuming the absence of downgoing energy, identical depths for source and receiver, zero phase wavelet, and several other assumptions which must usually be satisfied approximately by preprocessing, the propagator which gives the rate of change in this field with depth is the so-called double square root ("DSR") operator. It is conventional to express this operator by its dispersion relation:

$$k_z = \sqrt{\frac{\omega^2}{v_s^2} - \frac{1}{4} [(k_{m_x} - k_{h_x})^2 + (k_{m_y} - k_{h_y})^2]} + \sqrt{\frac{\omega^2}{v_r^2} - \frac{1}{4} [(k_{m_x} + k_{h_x})^2 + (k_{m_y} + k_{h_y})^2]}, \quad (1)$$

where v_s and v_r are respectively velocity at extrapolated source and receiver positions, and k_h and k_m are offset and midpoint wavenumbers.

Except in the case of constant velocity, the action of this operator must be approximated to achieve reasonable computational cost. A very large literature discusses a variety of approximations. We have followed the so-called Generalized Screen Propagator ("GSP") approach (de Hoop et al., 2000), based on a certain expansion of the DSR operator in a series of products of functions of the space variables and functions of corresponding wavenumbers.

In the 3D case, this approximate depth extrapolation can be applied within the Common Azimuth framework to produce an economical computation of the seismic field (Biondi and Palacharla, 1996). In this report we will present only 2D examples, for which Common Azimuth migration simply amounts to solving the DSR equation or a suitable approximation. Accordingly the field will be written as $p(x_r, x_s, z, t)$ or alternatively in terms of 2D midpoint and offset coordinates $p(m, h, z, t)$.

DIFFERENTIAL SEMBLANCE PRINCIPLE

According to Claerbout (1985), the image is extracted from the seismic field at coincident source and receiver locations and times. Since the depths are already constrained to be the same, and the time variable

t actually represents the lag between source and receiver times, the image or estimated reflectivity of the Earth is given in terms of the notation introduced above by $r(x, z) = p(x, 0, z, 0)$.

Information about velocity is also inherent in the seismic field. The imaging principle introduced by Claerbout (1985) also states that negligible energy should be present in the seismic field at nonzero offset at zero time (or at nonzero time for zero offset (Sava and Fomel, 2005)), provided that the migration velocity is at least kinematically correct. That is, the zero-time offset image volume $r_h(x, z, h) = p(x, h, z, 0)$ should be focused at $h = 0$, and failure to focus indicates need for a velocity correction.

It has become customary to display the velocity diagnostic content of the seismic field using so-called image gathers in the scattering angle or offset ray parameter domains (Prucha et al., 1999; Rickett and Sava, 2002). Focusing in the offset domain is equivalent to flatness of image gathers in the angle domain. In this report we will display exclusively offset image gathers, i.e. $r_h(x, z, h)$ for fixed x .

Several authors have suggested automated methods for velocity inversion (Toldi, 1989; Cao et al., 1990; Biondi and Sava, 2004). Most such suggestions are based on optimization of a quadratic form in the image volume $J(v) = \frac{1}{2} \|P_h r_h\|^2$, which is implicitly a function of the velocity (and the data). The symbol $\|\cdot\|$ denotes the root mean square of the quantity between the upright bars, over all of its dependent variables. The operator P_h is to be chosen so that J takes its extreme value at correct velocity. Since the quadratic form J is nonnegative, velocity estimation is reduced to the optimization problem $J(v_{true}) = \min_v J(v)$.

Use of efficient gradient-based optimization methods related to Newton's method requires that J be smooth. As shown by Stolk and Symes (2003), smoothness of J as function of velocity model and data traces requires that the operator P_h be (pseudo)differential. A natural choice of differential semblance ("DS") operator P_h for wave equation migration is multiplication by h : since correct choice of velocity concentrates energy at $h = 0$, multiplying the offset image volume by h should result in a very small mean square when the velocity is kinematically correct. According to Stolk and Symes (2003), J with this choice of P_h is regular as a function of velocity and data, hence amenable to gradient-based optimization method. Shen et al. (2003) have demonstrated a velocity analysis method using this choice of DS operator in conjunction with shot-record migration.

Gradient Calculation

The gradient of J represents its derivative via the inner (dot, scalar) product $\langle \cdot, \cdot \rangle$:

$$\langle \nabla J(v), \delta v \rangle = \langle r_h, h^2 D_v r_h \delta v \rangle = \langle D_v r_h^* (h^2 r_h), \delta v \rangle, \quad (2)$$

in which $D_v r_h$ denotes the Jacobian or derivative of the offset image volume with respect to velocity, and $D_v r_h^*$ its adjoint or transposed operator. It follows that $\nabla J = D_v r_h^* h^2 r_h$.

The so-called adjoint state algorithm computes $D_v r_h^*$ via an upward marching scheme in depth, from maximal depth until the surface. We write the propagation of the field p from depth z_{i-1} to depth $z_i = z_{i-1} + \Delta z$ as

$$p_i = H(v_{i-1}) p_{i-1}$$

(v_{i-1} being the velocity in this depth layer). By implicit differentiation, the perturbation δp_i at interface i in the seismic field resulting from velocity perturbation δv_{i-1} in layer $i-1$ and field perturbation δp_{i-1} at interface $i-1$ should satisfy

$$\delta p_i = D_v H(v_{i-1}) p_{i-1} \delta v_{i-1} + H(v_{i-1}) \delta p_{i-1}$$

DSR MVA by DSO

In matrix notation,

$$\begin{pmatrix} \delta p_0 \\ \delta p_1 \\ \vdots \\ \delta p_N \end{pmatrix} = \begin{pmatrix} 0 & 0 & & 0 \\ H(v_0) & 0 & & 0 \\ & \ddots & & \\ 0 & & H(v_{N-1}) & 0 \end{pmatrix} \begin{pmatrix} \delta p_0 \\ \delta p_1 \\ \vdots \\ \delta p_N \end{pmatrix} + \begin{pmatrix} 0 & 0 & & 0 \\ D_v H(v_0) p_0 & 0 & & 0 \\ & \ddots & & \\ 0 & & D_v H(v_{N-1}) p_{N-1} & 0 \end{pmatrix} \begin{pmatrix} \delta v_0 \\ \delta v_1 \\ \vdots \\ \delta v_N \end{pmatrix}. \quad (3)$$

Introducing the notations H and $D_v H p$ for the matrices appearing in the preceding equation, and δp and δv for the vectors of field and velocity perturbations, the matrix equation takes the compact form

$$\delta p = H \delta p + D_v H p \delta v$$

Thus

$$\delta p = (I - H)^{-1} D_v H p \delta v$$

Denote by T the operator which samples the field $p(\dots, t)$ at time $t = 0$ (if the depth extrapolation is carried out in the frequency domain, as is commonly done, then T represents summation over frequency). Then $r_h = T p$ whence $D_v r_h \delta v = T \delta p = T(I - H)^{-1} D_v H p \delta v$ whence $\nabla J = D_v r_h^* P_h^2 r_h = D_v H^* ((I - H)^{-1})^* T^* P_h^2 r_h$. The operator T^* is simply the operator which inserts its argument (a spatial field) as the $t = 0$ level of an (otherwise zero) space-time field. Since $I - H$ is lower triangular, $(I - H)^*$ is upper triangular, and we can write by inspection a recursion for the solution w of $(I - H)w = T^* P_h^2 r_h$:

$$w_i - H(v_i)^* w_{i+1} = (T^* P_h^2 r_h)_{z=z_i}, \quad i = 0, \dots, N \quad w_{N+1} = 0 \quad (4)$$

In terms of the *adjoint state field* w , the gradient of J may be expressed as

$$\nabla J_{z=z_i} = (D_v H(v_{i+1}) p_{i+1})^* w_i, \quad i = 0, \dots, N$$

To summarize: the gradient computation consists of

- Downward propagation of the seismic wavefield p until to a maximum depth z_{max} , at each step reading off r_h layer by layer, then
- Upward propagation of the adjoint wavefield w to the surface and evaluation of the gradient per the preceding formula, at each step of the upward propagation.

We emphasize that this computation is *exact*, i.e. produces the actual gradient of the objective J up to a machine precision error, and requires roughly three times the floating point operations of a downward propagation. On the other hand it requires the entire seismic wavefield p to be stored, for all depth levels, as this field is accessed in order inverse to its computation in the adjoint state recursion (4). Griewank (2000) discusses options for avoiding full storage of p , which could be onerous in 3D.

Smoothness and Bound Constraints

The theory of depth migration suggests that migration velocity models should be smooth on the wavelength scale (an often ignored restriction!). Therefore the model must be kept relatively smooth during the optimization. Several schemes for model representation and regularization satisfy this smoothing requirement. We chose to parametrize the model on a regular grid which is coarse on the wavelength scale,

and interpolate the velocity samples to the (fine) regular grid of the wave equation computation by cubic B-spline interpolation: $v = Bm$, where m is the vector of B-spline nodal parameters and B is the interpolation operator. We carry out optimization for these nodal parameters directly using the gradient of J with respect to m :

$$\nabla_m J = B^* \nabla J.$$

Therefore the adjoint B^* of the B-spline interpolation operator is also required.

For a variety of reasons, reasonable behaviour of the optimization requires that all updates of the velocity model satisfy prescribed upper and lower bounds. For example, the GSP remains accurate only if the velocity remains above a specified reference velocity. We used composition of the velocity nodal parameters with a sigmoidal function to assure that the prescribed bounds were maintained at all stages of the iteration, even though our optimization algorithm (limited memory BFGS algorithm ("LBFGS") (Nocedal and Wright, 1999) did not explicitly assure bound constraints.

EXAMPLES

We generated and processed several synthetic tests; we present the results of two such tests. Both are based on the Marmousi model and used the same data geometry as the blind test data set (Versteeg and Grau, 1991). We generated the test data using a time domain finite difference linearized ("Born modeling") code. The synthetic source was in all cases a point dilatation with trapezoidal bandpass filter (4-10-25-35 Hz) zero phase time dependence. As was true of the original Marmousi data, the top surface of the model was free (pressure-release), hence generated a ghost which we ignored, even though the data so acquired violated the theoretically necessary zero-phase constraint.

The first example modeled the perturbation of the data from a linear background velocity (1.5 km/s at surface, 4.5 km/s at 3 km depth) resulting from a short scale velocity perturbation obtained by subtracting from the original model a smoothed version. To accomplish the smoothing, we used the Seismic Unix utility `smooth2` (Stockwell, 2001) with a smoothing length of 20 m. To represent trial velocities, we used a grid of 6 points in x and 5 points in z .

Figure 1 shows migrated images (i.e. $r(x, z) = r_h(x, 0, z)$) at the initial velocity model and after 20 steps of LBFGS. The initial velocity was approximately 30% too high in the near-surface part of the model, resulting in severe mis-focusing and reflectors mapped considerably too deep. The final velocity is sufficiently accurate that the image is almost perfectly focused, and depths are essentially correct.

The second example in contrast combined a more complex background velocity with a simpler reflectivity (velocity perturbation). The model is the Marmousi velocity, smoothed using `smooth2` on a length scale of 200 m. The velocity perturbation consists of fourteen evenly spaced flat horizontal reflectors. Since the model is more laterally heterogeneous, we used a finer grid of ten nodes in x and 15 nodes in z . Figure 2(a) shows the layered initial velocity model, the inverted velocity model after 30 iterations of LBFGS, and for comparison the "true" smoothed Marmousi velocity model used to generate the data. In Figure 2(b) we show the result of the corresponding migrations.

While the velocity estimate (Figure 2(a), middle) shows clear evidence of the fault block structure, which gives this model its kinematic complexity, the migrated image (Figure 2(b), middle) does not reproduce the flat structure of the target reflectivity. Perhaps the lack of data near the edges is responsible for some of the sag at the ends, but the reflector depths in the middle are also not correct.

To assess this result, we created a number of other displays. One of these appears as Figure 2(c), which exhibits an offset-scaled, squared

DSR MVA by DSO

offset image gather from the middle of the model. The left plot shows the image gather from migration with the initial velocity, and displays the lack of focusing one would expect from a kinematically inaccurate velocity. The middle gather is extracted from the migration output at the inverted velocity - the high-offset energy is essentially suppressed. For comparison, we also show on the right the same gather produced by migrating with the “true” velocity, which does not appear more focused than the middle figure. In fact the objective J (the mean of all such gathers) at the inverted velocity is actually slightly *smaller* than the value produced at the “true” velocity!

DISCUSSION

We draw two tentative conclusions from these results. First, the velocity estimation process described in this paper has been implemented successfully: the optimization succeeds in producing focused image gathers, i.e. satisfies the diagnostic test for velocity correctness. Second, velocity estimation involves a considerable degree of nonuniqueness (Stork and Clayton, 1991); therefore the results of an automatic process like the one described here may be steered by noise, imaging artifacts, or data errors to one or another of the many equivalent models, all equally consistent with the data at least in a kinematic sense.

A number of important preprocessing issues strongly affected the success or failure of our tests with DS velocity inversion:

Tapering: The image volume can be severely affected by any sharp cutoff, on any axis. This means that the data must be tapered over several dominant wavelengths in time and all spatial axes. Such tapering can often be ignored in image production, as the artifacts so generated tend to “stack out”, i.e. be small at zero offset. However edge artifacts are coherent noise and can cause large errors in automated velocity estimates.

Filtering and muting: Incoherent noise is generally harmless. However coherent high frequency artifacts can have a very large effect, as can critical reflections and refractions, diving waves, and other often-strong data components not predicted by specular reflection theory.

Proper choice of extrapolation step, reference velocity: The choice of Δz critically affects the quality of the gathers, as does the deviation between reference and migration velocities. Proper selection of Δz depends on the minimum of the migration velocity field; either the velocity models allowed in the search must be constrained to share a common lower bound compatible with a fixed choice of Δz , or the latter must be adjusted as the optimization proceeds. As adaptive gridding enormously complicates the construction of reliable optimization algorithms, we have used a fixed lower bound as a constraint on the velocity search space. Likewise the reference velocity must not stray too far from the velocity update to maintain kinematic accuracy in the GSP.

ACKNOWLEDGMENTS

The authors would like to thank Total E&P USA, Inc for permission to present this work, and Henri Calandra, Scott Morton, Maarten de Hoop and Christiaan Stolk for many useful conversations. The second author (WWS) is grateful to the sponsors of The Rice Inversion Project for their long-term support of related research.

REFERENCES

Biondi, B. and G. Palacharla, 1996, 3-d prestack migration of common-azimuth data: *Geophysics*, **61**, 1533–1543.

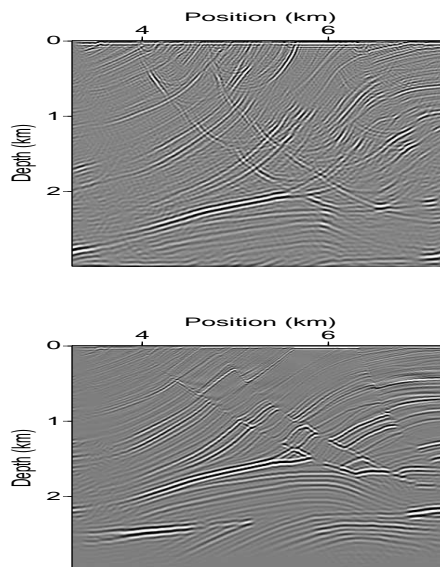
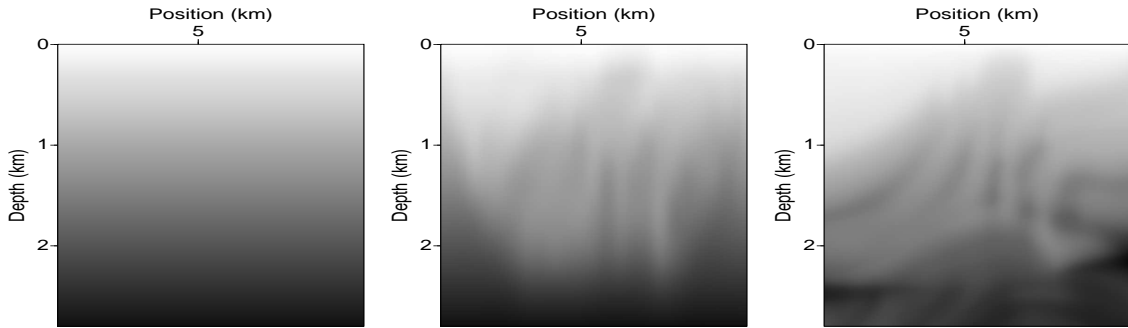


Figure 1: Inversion of a data set generated with the Marmousi reflectivity and a linear background: Top: migration at initial velocity. Bottom: migration at updated velocity.

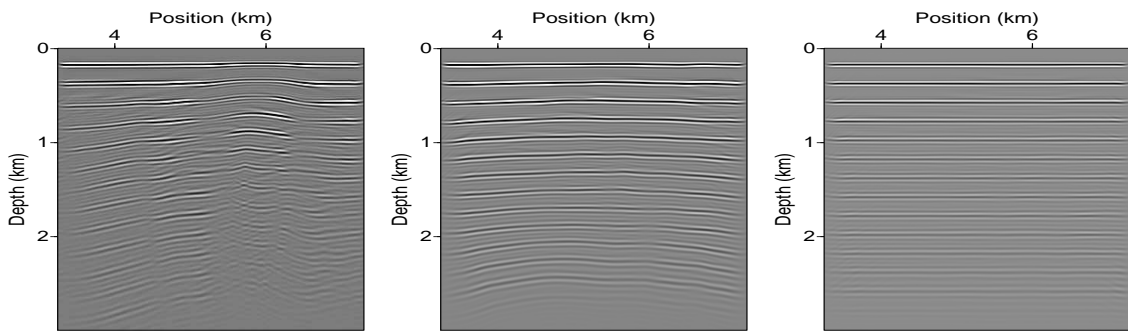
- Biondi, B. and P. Sava, 2004, Wave-equation migration velocity analysis - I: Theory, and II: Subsalt imaging examples: *Geophysics*, **52**, 593–623.
- Cao, D., S. Singh, and A. Tarantola, 1990, Simultaneous inversion for background velocity and impedance maps: *Geophysics*, **55**, 458–469.
- Claerbout, J., 1985, *Imaging the earth’s interior*: Blackwell Scientific Publishers.
- de Hoop, M. V., J. H. Le Rousseau, and R.-S. Wu, 2000, Generalization of the phase-screen approximation for the scattering of acoustic waves: *Wave Motion*, **31**, 43–70.
- Griewank, A., 2000, *Evaluating derivatives: Principles and techniques of automatic differentiation*: SIAM.
- Nocedal, J. and S. Wright, 1999, *Numerical optimization*: Springer Verlag.
- Prucha, M., B. Biondi, and W. Symes, 1999, Angle-domain common image gathers by wave-equation migration: 69th Annual International Meeting, Expanded Abstracts, 824–827.
- Rickett, J. and P. Sava, 2002, Offset and angle-domain common image-point gathers for shot profile migration: *Geophysics*, **67**, 883–889.
- Sava, P. and S. Fomel, 2005, Time-shift imaging condition: 75th Annual International Meeting, Expanded Abstracts, SPMI2–2.
- Shen, P., W. Symes, and C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: Society of Exploration Geophysicists, 73rd Annual International Meeting, Expanded Abstracts, 2135–2139.
- Stockwell, J., 2001, Seismic Unix home page. www.cwp.mines.edu/cwpcodes.
- Stolk, C. C. and W. W. Symes, 2003, Smooth objective functionals for seismic velocity inversion: *Inverse Problems*, **19**, 73–89.
- Stork, C. and R. Clayton, 1991, Linear aspects of tomographic velocity analysis: *Geophysics*, **56**, 483–495.
- Toldi, J., 1989, Velocity analysis without picking: *Geophysics*, **54**, 191–199.
- Versteeg, R. and G. Grau, 1991, Practical aspects of inversion: The Marmousi experience: Proceedings of the EAEG, The Hague.

DSR MVA by DSO

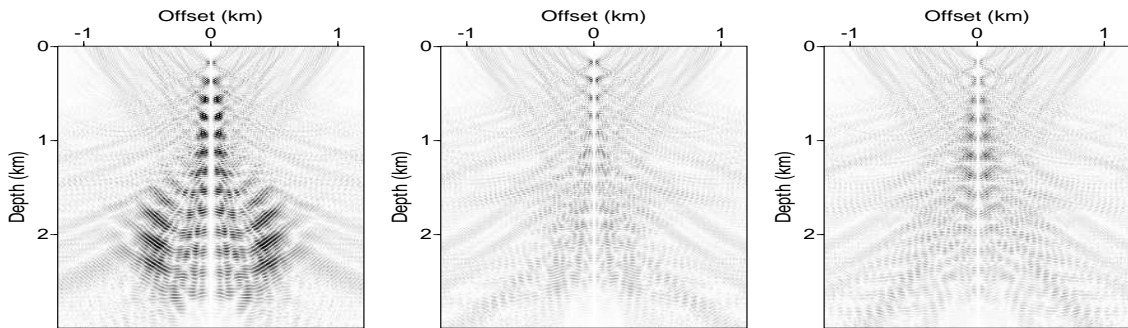
Figure 2:



(a) Velocity model updating: initial velocity (linear) on the left, updated velocity after 30 iterations of LBFGS in the middle, true velocity on the right



(b) Migration: migration at initial velocity on the left, migration at updated velocity in the middle, migration at true velocity on the right



(c) Display of $|h_r h(x, z, h)|^2$ at position $x = 5.250\text{km}$: at initial velocity on the left, updated velocity in the middle, at true velocity on the right