3D CRS analysis: a new data-driven optimization strategy for the simultaneous estimate of the eight stacking parameters

Ernesto Bonomi[1]*, Antonio M. Cristini*[1]*, Daniela Theis*[1] *and Paolo Marchetti*[2*]
[1]*CRS4 Imaging and Numerical Geophysics,* [2]*Eni E&P Division*

SUMMARY

We devised a data-driven strategy for the *simultaneous* estimate of the eight CRS traveltime attributes, solving a global non-linear minimization problem without the need of computing gradients. The essential elements are the following: a *conjugate-direction* method supported by well known convergence properties and an iterative *line-search* implementing the strong Wolfe-Powell rule for the control of the steplength. The resulting algorithm can reach very good solutions in presence of many local minima.

INTRODUCTION

Common-Reflection-Surface (CRS) theory supplies a proper framework to approximate *without* the knowledge of the medium macro-velocity model the kinematics of reflected events moving near a *reference* normal incident ray (Bergler et al., 2002). Leaving the surface at point $\mathbf{x}_0$, the normal incident ray reemerges at the same position with an arrival time t_0 after a specular reflection. The CRS kinematic description of all reflected paraxial rays is fully specified by *eight* geometric *attributes*, each of them estimated at $\mathbf{x}_0$. They are the elevation α_0 and the azimuth β_0 of the reference ray direction $\mathbf{w}$, and the 2x2 curvature matrices, $\mathbf{K}_{NIP}$ and $\mathbf{K}_N$, of two hypothetical waves traveling upward along the reference trajectory. $\mathbf{K}_{NIP}$ characterizes the quadratic approximation of the diffraction wavefront initiated at the normal incidence point. $\mathbf{K}_N$, a matrix related to the curvature of the reflecting surface, characterizes the quadratic approximation of the exploding reflector wavefront around the normal incidence point.

With these eight attributes, CRS analysis provides in the *midpoint-offset* domain centered in $\mathbf{x}_0$ a second-order traveltime *moveout* formula with respect to t_0, valid for each source-receiver paraxial trajectory close enough to the reference ray:

$$
\begin{aligned}
t^2(\mathbf{m},\mathbf{h}) \quad = \quad & \left(t_0 + \tfrac{2}{v_0}\mathbf{w}^T\mathbf{m}\right)^2 \\
& + \tfrac{2}{v_0} t_0 \; \mathbf{m}^T \; \mathbf{H}_{zy} \; \mathbf{K}_N \; \mathbf{H}_{zy}^T \; \mathbf{m} \qquad (1) \\
& + \tfrac{2}{v_0} t_0 \; \mathbf{h}^T \; \mathbf{H}_{zy} \; \mathbf{K}_{NIP} \; \mathbf{H}_{zy}^T \; \mathbf{h} \, .
\end{aligned}
$$

Here v_0 is the near-surface velocity; $\mathbf{H}_{zy}$, a 2x2 matrix depending on α and β, is a rotated projection into the plane perpendicular to the reference normal ray (Höcht, 2001).

For each coordinate $(\mathbf{x}_0, t_0)$ of the zero-offset volume, the moveout formula (1) allows to identify, collect and stack all primary events reflected *within* the aperture in the midpoint-offset domain (Cristini et al., 2002). By repeating this process, we can form a time-domain image based on the specular reflection of normal incident rays and, subsequently, a poststack-migrated section. Since CRS data stacking makes a large use of traces with arbitrary midpoint-offset locations around each position $\mathbf{x}_0$, processing involves much more traces than in the common-midpoint stacking are involved in the processing. This produces an image with increased signal-to-noise ratio and a higher resolution.

The aim of this article is to outline and illustrate a new *data-driven* computational strategy to efficiently and accurately estimate the field of the eight CRS attributes $\{\alpha_0, \beta_0, \mathbf{K}_{NIP}, \mathbf{K}_N\}$. We devised a strategy for the *simultaneous* estimate of the attributes by means of the coherence analysis, solving a global non-linear minimization problem for each coordinate $(\mathbf{x}_0, t_0)$ of the zero-offset volume. The resulting *conjugate-direction* algorithm, based on Powell search method, can reach very good solutions *even* in presence of many local minima, without the need of computing gradients.

The solution of this problem represents a crucial step for the accurate application of CRS theory and the industrial use of CRS analysis tools, designed to assist the interpretation of seismic data.

THE OPTIMIZATION PROBLEM

CRS stacking is a *data-driven* imaging and compression technique that must be implemented as a fully automatic procedure in which the field of stacking attributes $\xi = \{\alpha_0, \beta_0, \mathbf{K}_{NIP}, \mathbf{K}_N\}$ has to be necessarily computed before stacking the data. Assuming that the near surface velocity v_0 is known, the attributes are determined by means of the coherence analysis which essentially consists in maximizing the following *semblance* coefficient:

$$
S(\xi\,|\mathbf{x}_0, t_0) = \frac{1}{M} \; \frac{\displaystyle\sum_{k=-N/2}^{N/2} \left| \sum_{i=1}^{M} a_{i,\;t_i+k} \right|^2}{\displaystyle\sum_{k=-N/2}^{N/2} \; \sum_{i=1}^{M} \left| a_{i,\;t_i+k} \right|^2} \, . \qquad (2)
$$

$a_{i,j}$ is the j-th amplitude sample of the i-th trace belonging to the midpoint-offset aperture centered in $\mathbf{x}_0$. Coefficient (2) is evaluated for a time window of width N centered at $t_i = t(\mathbf{m}_i, \mathbf{h}_i)$, the traveltime prediction (1) measured in sample units. Maximum coherence, $0 \le S(\xi\,|\mathbf{x}_0, t_0) \le 1$, is achieved if there is a *perfect* agreement between the CRS traveltime and the actual arrival time. It is a common practice to solve a totally equivalent problem by minimizing $f(\xi\,|\mathbf{x}_0, t_0) = 1 - S(\xi\,|\mathbf{x}_0, t_0)$, an objective function with the same multimodal landscape.

Since for each new value of α_0, β_0, $\mathbf{K}_{NIP}$ and $\mathbf{K}_N$ calculating the semblance has a high computational cost in terms of floating-point operations and data movement, the search process, if not correctly designed, might be very time consuming.

A common approach (Cristini et al., 2002) consists in splitting the problem in a nested sequence of approximations where for each optimization step the search is limited to a reduced number of attributes on specific data collections. Since for each search either **m** or **h** must be set to zero in (2), the impossibility to gather seismic traces to satisfy these two conditions within the midpoint-offset aperture centered in $\mathbf{x}_0$ gives rise to further inaccuracies.

A global optimization method would certainly overcome all these deficiencies. However, the derivatives of the semblance being not explicitly available, a gradient-based minimization scheme cannot be implemented. An efficient way to keep good convergence properties *without* computing the gradients is to adopt a *conjugate-direction* minimizer, together with a reliable line search algorithm.

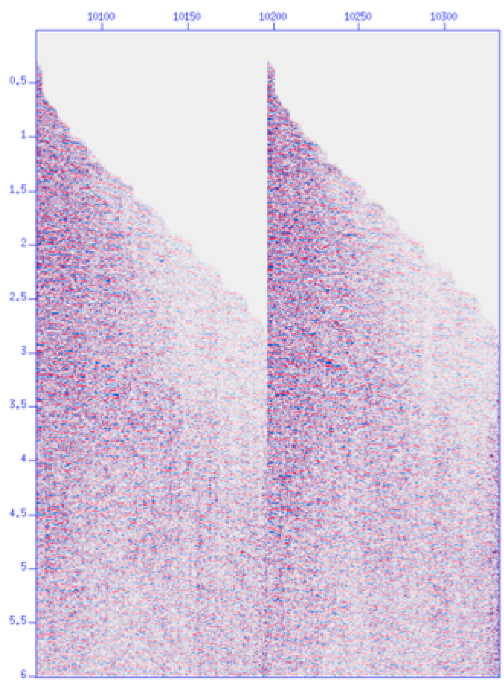

Figure 1: Two input CMP gathers showing a very low S/N ratio

A CONJUGATE-DIRECTION METHOD

Conjugate-direction methods, born to speed up the convergence rate of steepest descent while avoiding using Hessian, are highly efficient iterative search techniques. In case of convex cost functions, the iterations converge quadratically, starting from any real initial guess. Under these conditions the fundamental property of these methods (Pierre, 1986; Powell, 1964) is that the result of the n-th iteration is exactly the *minimizer* of $f(\xi)$ over the set $\mathbf{r}_i$ of conjugate directions in the attribute space. More precisely, if n nontrivial vectors $\mathbf{r}_i$, $i = 0, 1, \ldots, n-1$, are mutually conjugate, the exact minimum of $f(\xi)$ can be obtained by a sequence of n one-dimensional searches: starting at point ξ_0, the final result $\xi_{min} = \xi_n$ is extracted from

$$f(\xi_{i+1}) = \min_{\alpha_i \geq 0} f(\xi_i + \alpha_i \mathbf{r}_i), \qquad i = 0, 1, 2, \ldots, n-1. \qquad (3)$$

However, it can be proved that conjugate-direction methods are *locally* supported by the same good convergence properties even if, as in (2), the cost function $f(\xi)$ is *multimodal*.

Recalling that a conjugate set $\mathbf{r}_i$, $i = 0, 1, \ldots, n-1$ can be generated iteratively starting from a family of n linearly independent vectors $\mathbf{q}_i$, it comes out that the construction of the conjugate directions and the minimizing process can be carried ahead in the same iteration loop. For each coordinate $(\mathbf{x}_0, t_0)$ of the zero-offset volume, the minimization algorithm can be sketched as follows:

Powell Conjugate-direction algorithm

1. Initialize: set ξ_0, $h_0 = f(\xi_0)$ and choose $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{n-1}$ linearly independent
2. FOR $0 \leq i \leq n-1$ DO
   - ***Line Search***: find $\alpha_{LB} \leq \alpha_i \leq \alpha_{UB}$ minimizing $h(\alpha) = f(\xi_i + \alpha\, \mathbf{q}_i)$
   - Define the new point: $\xi_{i+1} = \xi_i + \alpha_i \mathbf{q}_i$
   - Set $h_{i+1} = h(\alpha_i)$
3. Find an integer $0 \leq k \leq n-1$, so that $\Delta = h_k - h_{k+1}$ is maximum
4. Compute: $f_3 = f_s(2\xi_n - \xi_0)$ and define $f_1 = h_0$ and $f_2 = h_n$
5. IF $f_3 < f_1$ AND $(f_1 - 2f_2 + f_3)(f_1 - f_2 - \Delta)^2 < \frac{1}{2}\Delta(f_1 - f_3)^2$ THEN
   - Define the new direction $\mathbf{q}_k = \mathbf{x}_n - \mathbf{x}_0$
   - ***Line Search***: find $\alpha_{LB} \leq \alpha_k \leq \alpha_{UB}$ minimizing $h(\alpha) = f(\xi_n + \alpha\, \mathbf{q}_k)$
   - Define the new point: $\xi_0 = \xi_n + \alpha_k \mathbf{q}_k$

   ELSE
   - Keep all directions $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{n-1}$ for the next iteration and set $\xi_0 = \xi_n$
6. Repeat steps 2 through 5 until convergence is achieved

In agreement with (3), the goal of the line search is to minimize $h(\alpha) = f(\xi + \alpha\mathbf{q})$ along the direction $\mathbf{q}$ varying the steplength $\alpha_{LB} \leq \alpha \leq \alpha_{UB}$. To determine a new point $\mathbf{x} + \alpha\mathbf{q}$, an effective *adaptive* selection rule for the steplength must be carefully implemented. A good choice is the adoption of the strong *Wolfe-Powell* conditions (Minoux, 1983), a robust and reliable two-sided rule coupling a necessary and a sufficient condition. The only pre-condition to be checked is that $h'(0) \leq 0$, so that $\mathbf{q}$ is locally a *descent* direction. Unfortunately, the search of a good solution minimizing a multimodal cost function has to deal with the possibility of a *premature convergence* to a local minimum. Since semblance (2) is characterized by a large number of relative extrema, none optimization method based on a descent algorithm will prevent the *trapping* into local minima without allowing *escape* movements along the *opposite* direction to the descent lines. The introduction of an *uphill* movement along each conjugate line towards regions of lower elevation greatly increases the capability of the line search algorithm to seek for solutions of lower cost. The idea is to steer the search using a modified *Wolfe-Powell* rule to frame admissible solutions also along the counter-descent direction. In particular, to determine whether for a steplength $\alpha < 0$ a new point $\xi + \alpha\mathbf{q}$ is significantly better than the current one ξ, the necessary condition of *Wolfe-Powell* scheme has been properly modified, keeping inalterate the sufficient one.

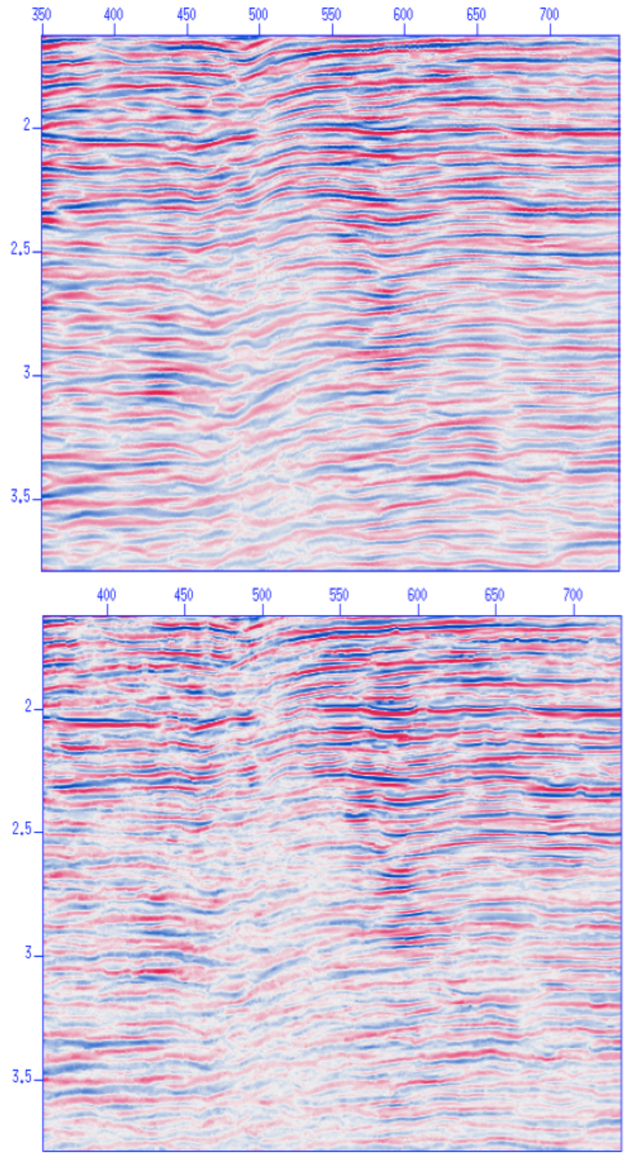

Figure 2: Zero-offset CRS stacking: comparison between the global (top) and multistep (bottom) optimization

3D REAL DATA EXAMPLES

The presented optimization algorithm was tested on several real 3D datasets and today is ready for production purposes. The images presented in this document refer to a dataset characterized by a very low signal-to-noise ratio, high fold (120%) and a relatively low structural complexity. Figure 1 illustrates the input data, a prestack time section strongly conditioned by the features of the acquisition environment (sand dunes). The main problems solved in the preprocessing phase regard the statics and the removal of the ground roll noise. The target zone is hidden by a strong reflector that prevents the energy to propagate in the lower part. The goal of the processing was the identification of the area global structural trend with the definition of two-fault systems.

Figure 2 shows two different CRS stack results. The image on the top illustrates the result of the new global optimization method. The image on the bottom refers to the Eni proprietary application routinely used since 2002 (Cristini et al., 2002). In this code, though significantly improved through the years, the optimization strategy has always been based on a sequence of nested approximations. The difference between the two results is striking and two features can be outlined. In the upper part of the two zero-offset sections, where the reflectors are better defined, the multistep optimization sometimes fails and produces strange cusps that are clearly wrong in a unmigrated time section. These artifacts are completely absent in the top image. The second interesting area is located around 3.4 s where the signal-to-noise ratio is very low. Here the new optimization scheme delivers a much better solution. These improvements in the zero-offset volume will obviously impact on the migrated image. Figure 3 displays the final semblance and demonstrates the accuracy achieved by the new optimizer with respect to the old one.

As a matter of fact, the optimization accuracy is even more evident on the calculated parameters. Just to give an example, Figure 4 displays the maximum component of the diagonalized projected *Fresnel zone* matrix, a derived attribute that directly depends on all eight calculated parameters (Hubral et al., 1993; Cristini et al., 2001): using the global optimizer (top), the improvement in the vertical resolution is impressive.

CONCLUSION

This work demonstrates that the simultaneous estimate of CRS attributes by means of the coherence analysis, *concurrently* solving a global non-linear minimization problem, is not only possible but provides very accurate results. This constitutes an important improvement in the production environment with respect to the previous multistep strategy. Moreover, with the new approach the number of attributes can be increased or reduced without any additional effort, thus opening new prospectives for the geophysical processing based on data-driven analysis. This software is unique and it constitutes a relevant step maintaining the Eni leadership in a very competitive international framework.

ACKNOWLEDGMENTS

The authors would like to thank the management of Eni E&P Division for the permission to present this work.

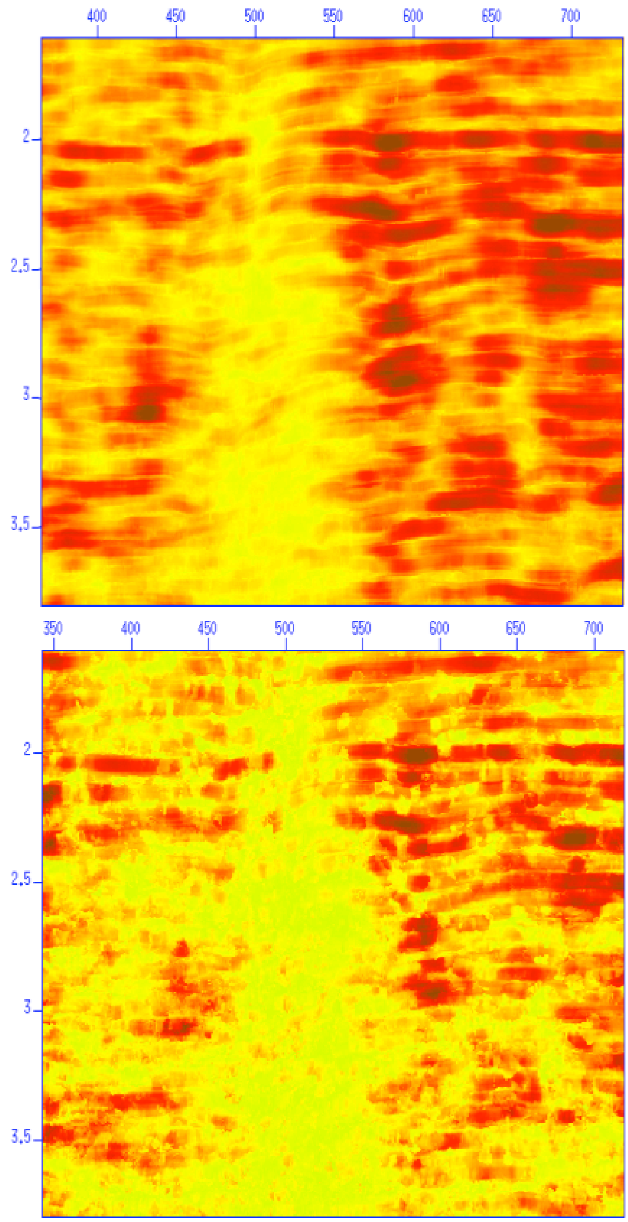

Figure 3: Semblance function: comparison between the global (top) and multistep (bottom) optimization

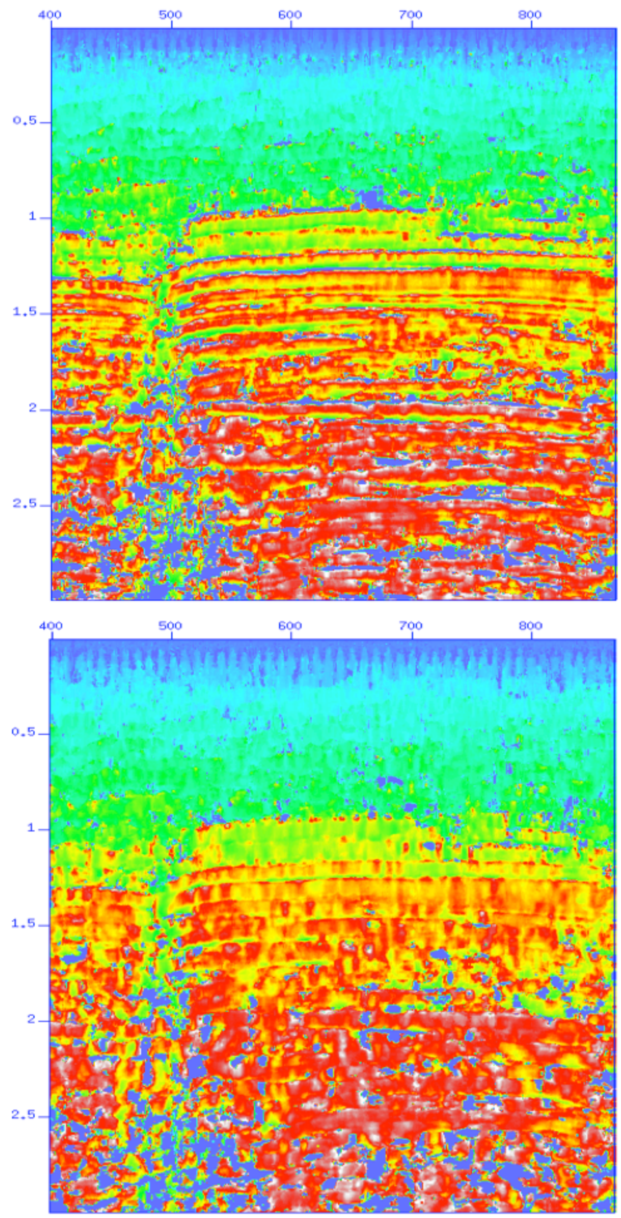

Figure 4: Maximum component of the diagonalized projected Fresnel zone matrix: comparison between the global (top) and multistep (bottom) optimization

APPENDIX A

THE SOURCE OF THE BIBLIOGRAPHY

```
@article{bergler,
  author = {Steffen Bergler and Peter Hubral and Paolo Marchetti and Antonio Cristini and Giovanni Cardone},
  title = {{3D common-reflection-surface stack and kinematic wavefield attributes}},
  year = {2002},
  journal = {The Leading Edge},
  volume = {21},
  number = {10},
  pages = {1010-1015},
  issn = {10},
  keywords = {seismology; data analysis; waveform analysis; classification; acoustic field; acoustic wave reflectio
  url = {http://link.aip.org/link/?LEEDFF/21/1010/1},
  doi = {10.1190/1.1518438},
}
@article{cristini02,
  author = {A. Cristini and G. Cardone and P. Marchetti},
  title = {{3D Zero-Offset Common Reflection Surface Stack for Land Data - Real Data Example}},
  year = {2002},
  journal = {EAGE, Expanded Abstracts},
  pages = {B015},
}
@PHDTHESIS{Hocht02,
  author = {German H\"{o}cht},
  title  = {{Traveltime Approximations for 2D and 3D Media and Kinematic Wavefiled Attributes}},
  school = {Geophysikalisches Institut - Universitat Karlsruhe},
  year   = {2001},
}
@MASTERSTHESIS{Jager99,
  author = {Rainer J\"{a}ger},
  title  = {{The Common-Reflection-Surface Stack Theory and Application}},
  school = {Geophysikalisches Institut - Universitat Karlsruhe},
  year   = {1999},
}
@article{powell,
  author = {M. J. D. Powell},
  title = {{An efficient method for finding the minimum of a function of several variables without calculating deri
  year = {1964},
  journal = {The Computer Journal},
  volume = {7},
  number = {2},
  pages = {155-162},
}
@article{hubral93,
author = {Peter Hubral and J\"{o}rg Schleicher and Martin Tygel and Ch. Hanitzsch},
collaboration = {},
title = {{Determination of Fresnel zones from traveltime measurements}},
publisher = {SEG},
year = {1993},
journal = {Geophysics},
volume = {58},
number = {5},
pages = {703-712},
url = {http://link.aip.org/link/?GPY/58/703/1},
doi = {10.1190/1.1443454}
}
@inproceedings{cristini01,
```

```
  author = {A. Cristini and G. Cardone and P. Chira and P. Hubral and P. Marchetti},
  title = {{3D Zero-Offset Common Reflection Surface Stack for Land Data}},
  year = {2001},
  booktitle = {SEG Workshop "Velocity Model Independent Imaging in Complex Media},
}
@BOOK{pierre,
  author = {Donald A. Pierre},
  publisher={Dover Publications, NY},
  title  = {Optimization theory with applications},
  year   = {1986},
}
@BOOK{minoux,
  author = {Michel Minoux},
  publisher={Dunod, Paris},
  title  = {Programmation math\'{e}matique},
  year   = {1983},
}
```

REFERENCES

Bergler, S., P. Hubral, P. Marchetti, A. Cristini, and G. Cardone, 2002, 3D common-reflection-surface stack and kinematic wavefield attributes: The Leading Edge, **21**, 1010–1015.

Cristini, A., G. Cardone, P. Chira, P. Hubral, and P. Marchetti, 2001, 3D Zero-Offset Common Reflection Surface Stack for Land Data: Presented at the SEG Workshop ”Velocity Model Independent Imaging in Complex Media.

Cristini, A., G. Cardone, and P. Marchetti, 2002, 3D Zero-Offset Common Reflection Surface Stack for Land Data - Real Data Example: EAGE, Expanded Abstracts, B015.

Höcht, G., 2001, Traveltime Approximations for 2D and 3D Media and Kinematic Wavefiled Attributes: PhD thesis, Geophysikalisches Institut - Universitat Karlsruhe.

Hubral, P., J. Schleicher, M. Tygel, and C. Hanitzsch, 1993, Determination of Fresnel zones from traveltime measurements: Geophysics, **58**, 703–712.

Minoux, M., 1983, Programmation mathématique: Dunod, Paris.

Pierre, D. A., 1986, Optimization theory with applications: Dover Publications, NY.

Powell, M. J. D., 1964, An efficient method for finding the minimum of a function of several variables without calculating derivatives: The Computer Journal, **7**, 155–162.