Seismic data processing with curvelets: a multiscale and nonlinear approach

Felix J. Herrmann

joint work with Deli Wang, Gilles Hennenfent and Peyman Moghaddam
Motivation

Exploit *two* aspects of curvelets, namely their
- parsimoniousness
- invariance under certain operators

Formulate
- *non-adaptive* wavefield reconstruction algorithms
- *data-adaptive* matching algorithms

Applications
- *nonlinear* sampling theory for wavefields
- *nonlinear* migration-amplitude recovery
- *nonlinear* primary-multiple separation
Employ parsimoniousness by sparsity promotion.

Exploit behavior of certain operators in phase space
- diagonalization <=> curvelet domain \textit{scaling}
- smoothness <=> \textit{structure} of phase space

Combine \textit{parsimoniousness} with \textit{structure} in phase space
- \textit{diagonal} approximation operators
- \textit{stable} amplitude recovery
- improved \textit{adaptive} separation
Migration-amplitude recovery methods are based on
- diagonal approximation of Pseudo’s
- estimate scaling from a reference vector and demigrated-migrated reference vector
  - Illumination-based normalization (Rickett ’02)
  - Amplitude corrections (Guitton ’04)
  - Amplitude scaling (Symes ’07)

Primary-multiple separation methods are based on
- diagonal approximation in the Fourier domain
- estimate scaling from mismatch pred. multiples & data
  - adaptive subtraction (Verschuur and Berkhout ’97)

We are interested in a formulation that
- estimates the scaling with smoothness control
- prevents overfitting
- allows for conflicting dips
- incorporates curvelet-domain sparsity promotion
The curvelet transform
2-D curvelets

Oscillatory in one direction and smooth in the others!
Obey \textit{parabolic} scaling relation $\text{length} \approx \text{width}^2$
Coefficients Amplitude Decay In Transform Domains

Trace #

Time (s)

Normalized amplitude

Percentage of amplitude-sorted coefficients

Fourier
Wavelets
Curvelets
Partial Reconstruction

Fourier (1% largest coefficients)

SNR = 2.1 dB
Partial Reconstruction

Curvelets (1% largest coefficients)

SNR = 6.0 dB
Non-adaptive curvelet-domain sparsity promotion
Linear quadratic (lsqr):

\[ \tilde{x} = \arg \min_x \|x\|_2 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \epsilon \]

- model Gaussian

Non-linear:

\[ \tilde{x} = \arg \min_x \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \epsilon \]

- model Cauchy (sparse)

Problem:

- data does not contain point scatterers
- not sparse
Our contribution

Model as superposition of little plane waves.

Compound *modeling* operator with curvelet *synthesis*:

\[
K \mapsto KC^T \\
m_0 \mapsto x_0 \\
\tilde{m} = C^T \tilde{x}
\]

Exploit *parsimoniousness* of curvelets on seismic data & images ...
Sparsity-promoting program

Problems boil down to solving for \( \hat{x}_0 \)

\[
y = A \hat{x}_0 + n
\]

with

\[
P_\varepsilon : \begin{cases} 
\hat{x} = \text{arg min}_{\hat{x}} \| \hat{x} \|_1 \quad \text{s.t.} \quad \| A\hat{x} - y \|_2 \leq \varepsilon \\
\tilde{m} = C^T \hat{x}
\end{cases}
\]

- exploit sparsity in the curvelet domain as a prior
- find the sparsest set of curvelet coefficients that match the data, i.e., \( y \approx KC^T \tilde{x} \)
- invert an underdetermined system
Seismic wavefield reconstruction with CRSI
Sparsity-promoting inversion*

Reformulation of the problem

\[ \text{signal} \rightarrow y = \mathbf{RC}^H x_0 + n \rightarrow \text{noise} \]

Curvelet Reconstruction with Sparsity-promoting Inversion (CRSI)

- look for the **sparsest/most compressible**, **physical** solution

\[
P_{\epsilon} : \begin{cases} 
\hat{x} = \arg \min_x \|Wx\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \epsilon \\
\tilde{f} = C^T \hat{x} 
\end{cases}
\]

* inspired by Stable Signal Recovery (SSR) theory by E. Candès, J. Romberg, T. Tao, Compressed sensing by D. Donoho & Fourier Reconstruction with Sparse Inversion (FRSI) by P. Zwartjes
CRSI recovery with 3-D curvelets
Adaptive curvelet-domain matched filtering
Forward model

Linear model for amplitude mismatch:

\[(Bf)(x) = \int_{x \in \mathbb{R}^d} e^{jk \cdot x} b(x, k) \hat{f}(k) dk\]

- \(B\) = Pseudodifferential operator
- \(b(x, k)\) = the symbol
- spatially-varying dip filter
- zero-order Pseudo

After discretization

\[f = Bg\]

- linear operator
- \(f\) and \(g\) known
- matrix \(B\) is full and not known ....
Forward model

Diagonal approximation in the curvelet domain:

\[ f = B g \]

\[ \approx C^T \text{diag}\{w\} C g \]

- curvelet domain scaling
- opens the way to an estimation of \( w \)

Examples:

<table>
<thead>
<tr>
<th></th>
<th>( B )</th>
<th>( f )</th>
<th>( g )</th>
</tr>
</thead>
<tbody>
<tr>
<td>migration</td>
<td>( K^T K )</td>
<td>migrated “image”</td>
<td>“reflectivity”</td>
</tr>
<tr>
<td>multiple removal</td>
<td>obliquity factor</td>
<td>total data</td>
<td>predicted multiples</td>
</tr>
</tbody>
</table>
Key idea

Problems with estimating $w$
- inversion of an underdetermined system
- over fitting
- positivity and reasonable scaling by $w$

Solution:
- use smoothness of the symbol
- formulate nonlinear estimation problem that minimizes

$$J_\gamma(z) = \frac{1}{2} \| d - F_\gamma e^z \|_2^2,$$

with

$$\text{grad} J(z) = \text{diag}\{e^z\} \left[ F^T (F e^z - d) \right]$$

- solve with I-BFGS
Key idea

- East quadrants
- West quadrants
- North quadrants
- South quadrants

16 angles/quad
8 angles/quad

Fine scales
Coarser scales

$D_1$

$D_2$

Fine scales

$D_\theta \theta$

South quadrants

East quadrants

West quadrants

Fine scales

Coarser scales

$D_1 \ x_1$

$D_2 \ x_2$

North quadrants

Fine scales

Coarser scales

$D_\theta \theta$

South quadrants

East quadrants

West quadrants

Fine scales

Coarser scales

$D_1 \ x_1$

$D_2 \ x_2$
Key idea

Impose *smoothness* via following system of equations

\[
f = C^T \text{diag}\{Cg\} w
\]

\[
0 = \gamma L w
\]

with

\[
L = \begin{bmatrix} D_1^T & D_2^T & D_\theta^T \end{bmatrix}^T
\]

first-order differences in *space* and *angle* directions for each *scale*. Equivalent to

\[
\tilde{w} = \operatorname{arg\,min}_w \frac{1}{2} \| b - P[w]\|_2^2 + \gamma^2 \|Lw\|_2^2
\]

with

\[
P = C^T \text{diag}\{Cg\}\]
Smoothness penalty

- reduces overfitting
- scaling is positive and reasonable
Smoothness penalty

\[ \gamma = 0 \]
Smoothness penalty

\[ \gamma = \frac{1}{2} \]
Smoothness penalty

\[ \gamma = 5 \]
Seismic amplitude recovery
Matching procedure

Compute *reference* vector $\leftrightarrow$ defines $\mathbf{g}$
- migrate data
- apply spherical-divergence correction

Create “data” $\leftrightarrow$ defines $\mathbf{f}$
- demigrate
- migrate

Estimate *scaling* by inversion procedure

Define *scaled* curvelet transform

Recover migration amplitudes by *sparsity* promotion.
bandpass-filtered reflectivity

migrated image

reference vector

imaged reference vector
Primary-multiple separation
Matching procedure

Predict multiples \( \iff \) defines \( g \)
- apply conventional Fourier matched filtering

Consider total data as "true" multiples \( \iff \) defines \( f \)
- do not know true multiples
- use total data instead
- minimize energy mismatch

Estimate scaling by an inversion procedure.

Define scaled curvelet-domain threshold.

Separate primaries & multiples by sparsity promotion.
Problem formulation

Signal model for total data

\[ s = s_1 + s_2 \]

Multiple prediction by e.g. SRME may contain amplitude errors, i.e.,

\[ s_2 = B \tilde{s}_2 \]

\[ s_2 \approx C^T \text{diag}\{w\} C \tilde{s}_2 \]

Solve

\[ J_\gamma(z) = \frac{1}{2} \| s - F_\gamma e^z \|_2^2, \]

with \( s \) the total data. Use \( z \) to correct the predicted multiples, i.e.,

\[ \tilde{s}_2 \mapsto C^T \text{diag}\{\tilde{w}\} C \tilde{s}_2 \text{ with } \tilde{w} = e^{\tilde{z}} \]

or correct the thresholding

\[ t = \text{diag}\{\tilde{w}\} |C \tilde{s}_2| \]
Synthetic example

Total data $S$

SRME predicted multiples $\tilde{S}_2$
Synthetic example

SRME predicted primaries

\[ \tilde{s}_1 \]

Curvelet estimated primaries

\[ \tilde{s}_1 = C^T T_t (C_p) \]
\[ t = C \tilde{s}_2 \]
Corrected multiples
\[ \mathbf{s}_2^{\text{corr.}} = \mathbf{C}^T \text{diag}\{\mathbf{w}\} \mathbf{C} \tilde{s}_2 \text{ for } \gamma = 0 \]

Corrected multiples
\[ \mathbf{s}_2^{\text{corr.}} = \mathbf{C}^T \text{diag}\{\mathbf{w}\} \mathbf{C} \tilde{s}_2 \text{ for } \gamma = 0.5 \]
Scaled thresholded primaries

\[ \tilde{s}_1 = C^T T_t (Cp) \]
\[ t = \text{diag}\{w\} |C\tilde{s}_2| \]
Synthetic example

Scaled thresholded primaries
\[ \tilde{s}_1 = C^T T_t(Cp) \]
\[ t = \text{diag}\{w\}|C\tilde{s}_2| \]

Curvelet estimated primaries
\[ \tilde{s}_1 = C^T T_t(Cp) \]
\[ t = C\tilde{s}_2 \]
Real example

SRME predicted multiples
\( \tilde{S}_2 \)

SRME predicted primaries
\( \tilde{S}_1 \)
Real example

Thresholded primaries

\[ \tilde{s}_1 = C^T T_t (Cp) \]
\[ t = C \tilde{s}_2 \]

Scaled thresholded primaries

\[ \tilde{s}_1 = C^T T_t (Cp) \]
\[ t = \text{diag}\{w\} | C \tilde{s}_2 | \]
Conclusions

Combining the parsimonious curvelet transform with phase-space structure allows us to control diagonal estimation \( \iff \) over fitting handle data with conflicting dips stably recover & separate

Application
- improved migration-amplitude recovery
- improved primary-multiple separations

Future
- 3-D
- non-smooth symbols
Acknowledgments

The authors of CurveLab (Demanet, Ying, Candes, Donoho)

Christiaan C. Stolk for his contribution to phase-space smoothness.

The SLIM team Sean Ross Ross, Cody Brown and Henryk Modzeleweski for developing SLIMPy: operator overloading in python

These results were created with Madagascar developed by Sergey Fomel.

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE (334810-05) of F.J.H. This research was carried out as part of the SINBAD project with support, secured through ITF (the Industry Technology Facilitator), from the following organizations: BG Group, BP, Chevron, ExxonMobil and