Seismic imaging and processing with curvelets

Felix J. Herrmann

joint work with Deli Wang
Combinations of **parsimonious** signal representations with nonlinear **sparsity** promoting programs hold the **key** to the next-generation of seismic data processing algorithms ...

Since they

- allow for formulations that are **stable** w.r.t.
  - noise
  - incomplete data
  - moderate phase rotations and amplitude errors

Finding a **sparse** representation for seismic data & images is complicated because of

- wavefronts & reflectors are multiscale & multi-directional
- the presence of caustics, faults and pinchouts
The curvelet transform
Representations for seismic data

<table>
<thead>
<tr>
<th>Transform</th>
<th>Underlying assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>FK</td>
<td>plane waves</td>
</tr>
<tr>
<td>linear/parabolic Radon transform</td>
<td>linear/parabolic events</td>
</tr>
<tr>
<td>wavelet transform</td>
<td>point-like events (1D singularities)</td>
</tr>
<tr>
<td>curvelet transform</td>
<td>curve-like events (2D singularities)</td>
</tr>
</tbody>
</table>

Properties curvelet transform:

- **multiscale**: tiling of the FK domain into dyadic coronae
- **multi-directional**: coronae sub-partitioned into angular wedges, # of angle doubles every other scale
- **anisotropic**: parabolic scaling principle
- Rapid decay space
- Strictly localized in Fourier
- Frame with moderate redundancy (8 X in 2-D and 24 X in 3-D)
2-D curvelets

- Curvelets are of rapid decay in space.
- Curvelets are strictly localized in frequency.
- Oscillatory in one direction and smooth in the others!
- Obey *parabolic* scaling relation: \( \text{length} \approx \text{width}^2 \)
Curvelet tiling & seismic data

# of angles doubles every other scale doubling!
Real data frequency bands example

Data

Data is multiscale!

Decomposition in frequency bands
Single frequency band angular wedges

6th scale image

Data is multidirectional!

Decomposition in angular wedges
Wavefront detection

Significant curvelet coefficient

Curvelet coefficient ~ 0
Extension to 3-D
Cartesian Fourier space

[courtesy Demanet ’05, Ying ’05]

Curvelets live in a wedge in the 3 D Fourier plane...
Curvelets are oscillatory in one direction and smooth in the others.
Coefficients Amplitude Decay
In Transform Domains

Fourier
Wavelets
Curvelets
**Partial Reconstruction**

**Curvelets** (1% largest coefficients)

SNR = 6.0 dB
Curvelet sparsity promotion
Forward model

Linear model for the measurements of a function \( m_0 \):

\[
y = Km_0 + n
\]

with

\[
y = \text{data} \\
K = \text{the modeling matrix} \\
m_0 = \text{the model vector} \\
n = \text{noise}
\]

- inversion of \( K \) either ill-posed or underdetermined.
- seek a prior on \( m \).
Key idea

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

When a traveler reaches a fork in the road, the \(l_1\)-norm tells him to take either one way or the other, but the \(l_2\)-norm instructs him to head off into the bushes.

John F. Claerbout and Francis Muir, 1973

New field “compressive sampling”: D. Donoho, E. Candes et. al., M. Elad etc.

Preceded by others in geophysics: M. Sacchi & T. Ulrych and co-workers etc.
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**:

\[
K \quad \mapsto \quad KC^T
\]

\[
m_0 \quad \mapsto \quad x_0
\]

\[
\tilde{m} = C^T \tilde{x}
\]

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

Problems boils down to solving for $x_0$

$$\text{signal} \rightarrow \begin{bmatrix} y \\ n \end{bmatrix} = \begin{bmatrix} A \\ x_0 \end{bmatrix} + \begin{bmatrix} n \end{bmatrix} \rightarrow \text{noise}$$

with

$$P_\epsilon : \begin{cases} \tilde{x} = \arg \min_{x} \| x \|_1 \quad \text{s.t.} \quad \| Ax - y \|_2 \leq \epsilon \\
\tilde{m} = C^T \tilde{x} \end{cases}$$

- 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
Initialize:

\[ i = 0; \ x^0 = 0; \]

Choose: \( L, \|A^T y\|_\infty > \lambda_1 > \lambda_2 > \cdots \)

while \( \|y - Ax^i\|_2 > \epsilon \) do

for \( l = 1 \) to \( L \) do

\[ x^{i+1} = T^s_{\lambda_i} \left( x^i + A^T (y - Ax^i) \right) \]

end for

\[ i = i + 1; \]

end while

\[ \tilde{f} = C^T x^i. \]
Applications

Problems in seismic processing can be cast in to $P_\epsilon$

- stable under noise
- stable under missing data

Obtain a formulation that

- explicitly exploits compression by curvelets
- is stable w.r.t. noise
- exploits the "invariance" of curvelets under imaging

Applications include

- seismic data regularization
- primary-multiple separation
- seismic amplitude recovery
Seismic data regularization

joint work with Gilles Hennenfent
Motivation
Irregular sub-sampling

Noisy because of irregular sampling ...
Sparsity-promoting inversion*

Reformulation of the problem

\[ \text{signal} \rightarrow y = \mathbf{RC}^H \mathbf{x}_0 + \mathbf{n} \quad \text{noise} \]

Curvelet Reconstruction with Sparsity-promoting Inversion (CRSI)

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

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

* 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
Primary multiple separation

Joint work with Eric Verschuur, Deli Wang, Rayan Saab and Ozgur Yilmaz
Motivation

Primary-multiple separation step is crucial
- moderate prediction errors
- 3-D complexity & noise

Inadequate separation leads to
- remnant multiple energy
- deterioration primary energy

Introduce a transform-based technique
- stable
- insensitive to moderate shifts & phase rotations

Exploit sparsity and parameterization transformed domain
Move-out error

![Graphs showing total data and multiple data](image)
Move-out error

matched filter

Threshold $\lambda = 1.4$
The problem

Sparse signal model:

\[ y = Ax_0 + n, \]

with

\[ A = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \quad \text{and} \quad x_0 = \begin{bmatrix} x_{01} & x_{02} \end{bmatrix}^T \]

- augmented synthesis and sparsity vectors
- index 1 <-> primary
- index 2 <-> multiple
The solution

The weighted norm-one optimization problem:

\[
\begin{align*}
\min_x \|x\|_{w,1} \quad \text{subject to} \quad \|y - Ax\|_2 \leq \varepsilon \\
\end{align*}
\]

\[P_w : \begin{cases} 
\hat{s}_1 = A_1 \hat{x}_1 \quad \text{and} \quad \hat{s}_2 = A_2 \hat{x}_2 \\
given: \ \hat{s}_2 \quad \text{and} \quad w(y, \hat{s}_2) 
\end{cases}\]

with

\[
\begin{align*}
w & := \begin{bmatrix} w_1, w_2 \end{bmatrix}^T \\
A & := \begin{bmatrix} C^T, C^T \end{bmatrix} \\
\hat{s}_2 & := \text{predicted multiples} \\
\hat{s}_1 & := S - \hat{S}_2 
\end{align*}
\]
Solution cont’d

The weights

\[
\begin{align*}
    w_1 &:= \max \left( \sigma \cdot \sqrt{2 \log N}, C_1 |\tilde{u}_1| \right) \\
    w_2 &:= \max \left( \sigma \cdot \sqrt{2 \log N}, C_2 |\tilde{u}_2| \right)
\end{align*}
\]

with

\[
\begin{align*}
    \tilde{u}_1 &\approx C\tilde{s}_1 \\
    \tilde{u}_2 &\approx C\tilde{s}_2
\end{align*}
\]

- during minimization signal components are driven apart
- curvelet compression helps
- separates on the basis of position, scale and direction
Synthetic example

total data

SRME predicted multiples
Synthetic example

SRME predicted primaries

curvelet-thresholded
Synthetic example

SRME predicted primaries

estimated
Real example

total data

SRME predicted multiples
Seismic amplitude recovery

Joint work with Chris Stolk and Peyman Moghaddam
Motivation

Migration generally does not correctly recover the amplitudes.

Least-squares migration is computationally unfeasible.

Amplitude recovery (e.g. AGC) lacks robustness w.r.t. noise.

Existing diagonal amplitude-recovery methods

- do not always correct for the order (1 - 2D) of the Hessian [see Symes ‘07]
- do not invert the scaling robustly

Moreover, these (scaling) methods assume that there

- are no conflicting dips (conormal) in the model
- is infinite aperture
- are infinitely-high frequencies
- etc.
Existing scaling methods

Methods are based on a diagonal approximation of $\Psi$.

- Illumination-based normalization (Rickett ‘02)
- Amplitude preserved migration (Plessix & Mulder ‘04)
- Amplitude corrections (Guitton ‘04)
- Amplitude scaling (Symes ‘07)

We are interested in an ‘Operator and image adaptive’ scaling method which

- estimates the action of $\Psi$ from a reference vector close to the actual image
- assumes a smooth symbol of $\Psi$ in space and angle
- does not require the reflectors to be conormal $\iff$ allows for conflicting dips
- stably inverts the diagonal
Our approach

“Forward” model:

\[ y = K^T K m + \epsilon \]

\[ \approx A x_0 + \epsilon \]

with

- \( y \) = migrated data
- \( A := C^T \Gamma \)
- \( A A^T r \approx K^T K r \)
- \( K \) = the demigration operator
- \( \epsilon \) = migrated noise.

- diagonal approximation of the demigration-migration operator
- costs one demigration-migration to estimate the diagonal weighting
Solution

Solve

\[
P : \begin{cases}
\min_x J(x) \quad \text{subject to} \quad \|y - Ax\|_2 \leq \epsilon \\
\tilde{m} = (A^H)\dagger \tilde{x}
\end{cases}
\]

with

\[
J(x) = \alpha \|x\|_1 + \beta \|\Lambda^{1/2} (A^H)\dagger x\|_p .
\]

sparsity

continuity
Example

SEGAA’ data:
- “broad-band” half-integrated wavelet [5-60 Hz]
- 324 shots, 176 receivers, shot at 48 m
- 5 s of data

Modeling operator
- Reverse-time migration with optimal check pointing (Symes ‘07)
- 8000 time steps
- modeling 64, and migration 294 minutes on 68 CPU’s

Scaling requires 1 extra migration-demigration
Migrated data

Amplitude-corrected & denoised migrated data
Nonlinear data
Conclusions

The combination of the parsimonious curvelet transform with nonlinear sparsity & continuity promoting program allowed us to...

- recover seismic data from large percentages missing traces
- separate primaries & multiples
- recover migration amplitudes

This success is due to the curvelet’s ability to
- detect wavefronts <=> multi-D geometry
- differentiate w.r.t. positions, angle(s) and scale
- diagonalize the demigration-migration operator

Because of their parsimoniousness on seismic data and images, curvelets open new perspectives on seismic processing ...
Acknowledgments

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

William Symes for the reverse-time migration code.

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 Shell.