Velocity Analysis and Waveform Inversion

William W. Symes

The Rice Inversion Project

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

How to access this short course online:

TRIP web site: `www.trip.caam.rice.edu`

go to downloadable materials link, grab (these) slides and reference list from short course on migration velocity analysis and waveform inversion

RICE

Velocity Analysis vs. Waveform Inversion

(Migration) velocity analysis ("MVA"): update velocity parameters to satisfy semblance condition in migrated image volume (usually: flat gathers)

- visual/interactive techniques, from 60's on
- often converted to traveltime inversion (reflection tomography in data, migrated domains) via automated picking $\Rightarrow$ optimization of traveltime misfit
- backproject traveltime residuals

Waveform inversion ("WI"): update model parameters to match predicted to observed data (parameters include p-wave velocity, but maybe much more)

- optimization of waveform misfit directly, without intermediate reduction to traveltime
- backproject waveform residuals

Velocity Analysis vs. Waveform Inversion

Bottom line: today,

- MVA integrated into industrial processing
- WI still academic - intrinsic obstacles

Agenda for this course:

- how to formulate MVA as a waveform-based optimization (update velocity by backpropagating *waveform* residuals)
- MVA = WI based on Born approximation
- proposal: reformulation of WI based on MVA ideas

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

Measuring Semblance

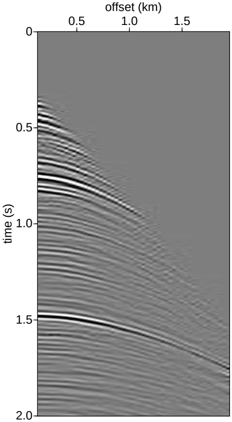
Foundation ("semblance") principle of velocity analysis: adjust velocity to obtain internally consistent image volumes

Example: simplest imaging algorithm = NMO/stack.

Image volume = NMO-corrected CMPs

Internal consistency = flat events

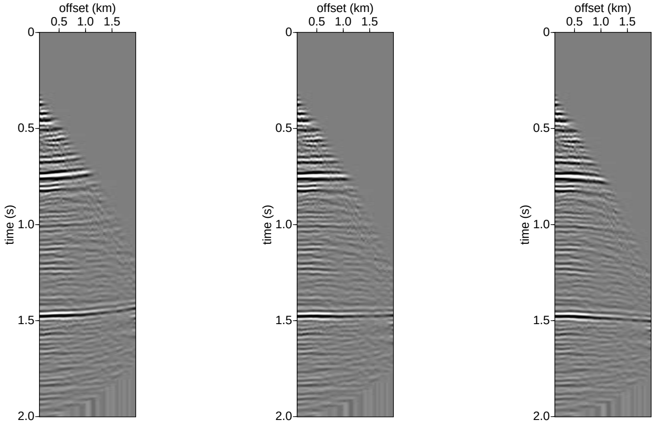
Measuring Semblance



Example:

- CMP from North Sea 2D survey (NAM, courtesy Shell Intl. Research).
- Predictive decon, low pass filter, mute
- Flat lying sediments to 3s+ $\Rightarrow$ convolutional model plausible.

Measuring Semblance



NMO corrected: velocity slightly low, about right, slightly high.

Measuring Semblance

Idea of differential semblance: if velocity is wrong,

- far traces are uncorrelated, but
- correlation of near traces indicates velocity error

Earliest known references: J. Castagna 1986 (inaccessible - cited in Sarkar et al. 2001), S. 1986. Related ideas ("plane wave PEFs") - Claerbout, 80's and 90's (IEI, Fomel 2002).

A Differential Semblance Method

Kinematic quantities:

Interval velocity $v(t_0, x, y)$ - depends on midpoint (x, y) and zero-offset two-way time t_0 (proxy for depth z).

Suppress x, y from notation, for simplicity - consider one midpoint.

Square RMS slowness $u(t_0) = \frac{t_0}{\int_0^{t_0} v^2}$.

Hyperbolic 2-way traveltime approximation $T(t_0, h) = \sqrt{t_0^2 + 4u(t_0)h^2}$. ($h$=half-offset).

Inverse TT function $T_0(t, h)$ - satisfies $T_0(T(t_0, h), h) = t_0$. "t_0 for which 2-way time at offset h is t."

A Differential Semblance Method

Data CMP at midpoint with coords x, y, offset $h = d(t, h)$.

NMO-corrected data

$$I(t_0, h) = d(T(t_0, h), h),$$

Offset divided difference operator

$$D_h I(t_0, h) = \frac{I(t_0, h + \Delta h) - I(t_0, h)}{\Delta h}.$$

A Differential Semblance Method

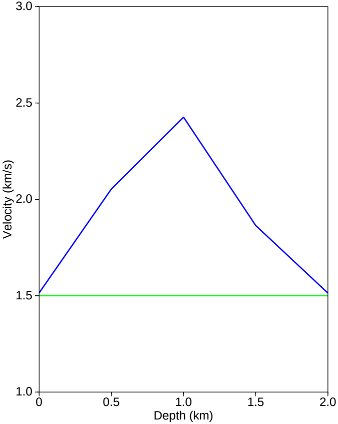
Differential semblance function:

$$J[v,d] = \frac{1}{2}\sum_h \int_0^{t_{\max}} dt |D_h I(T_0(t,h),h)|^2,$$

Algorithm:

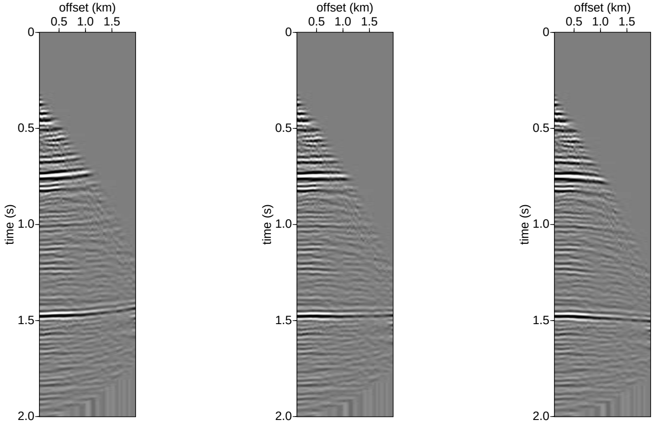
- NMO-correct each midpoint gather: $I(t_0,h) = d(T(t_0,h),h)$;
- form offset derivative (divided difference) $D_h I(t_0,h)$;
- inverse NMO-correct: $I(T_0(t,h),h)$;
- square and sum over t_0 and h (and over midpoints x, y).

A Differential Semblance Method



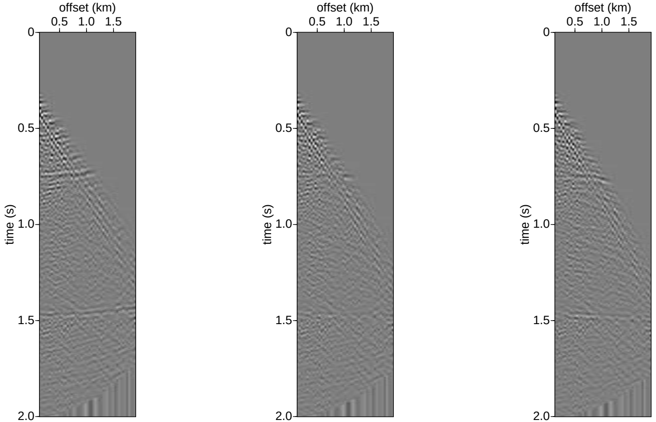
Two interval velocities: reference/initial (green) and estimate (blue).

A Differential Semblance Method



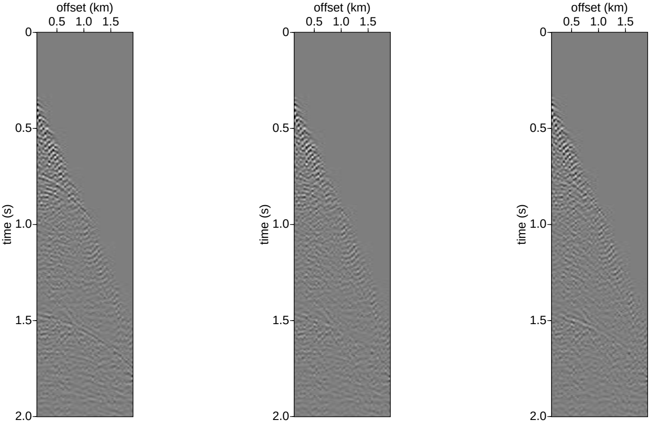
$I(t_0, h)$ = NMO corrected sections. Left = 20% ref, 80% est; Center = 100% est; Right: -20% ref, 120% est.

A Differential Semblance Method



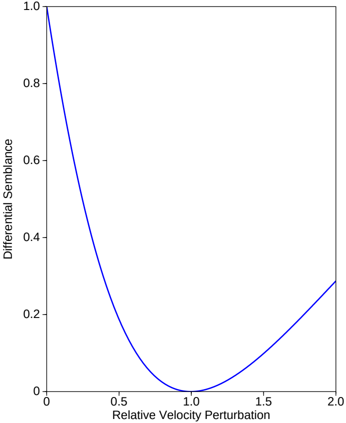
$D_h I(t_0, h) =$ offset divided differences $=$ scaled differences of neighboring traces.

A Differential Semblance Method



$D_h I(T_0(t,h),h)$ = inverse NMO applied to offset divided differences. Mean square = DS objective function.

A Differential Semblance Method



DS = mean square of $D_h I$, as function of relative velocity pert r: velocity $= (1 - r)*\text{ref} + r*\text{est}$.

A Differential Semblance Method

Automated via NMO-based differential semblance: given CMP data d and initial guess v_0, find v to minimize $J[v, d]$ using numerical optimization.

DS is smooth and (apparently) unimodal $\Rightarrow$ gradient-based methods will work, find global min!

Gradient-based optimization algorithm: described in Li & S. 2007. Central issue: computation of gradient $\nabla_v J[v, d]$.

Combine gradient with quasi-Newton optimization algorithm ("LBFGS").

[demos using NAM data - 1, 2, 3]

A Differential Semblance Method

Some questions, within the realm of weak lateral heterogeneity:

- What about other possible objective functions?
- What theoretical support exists?
- What about multiples and other coherent noise?

Total Stack Power

An alternative: *total stack power*

$$J_{\rm TSP}(d,v) = \int_0^{t_{\rm max}} \, dt_0 \, (\sum_h I(t_0,h))^2$$

Reaches *MAX* when v is kinematically correct - then summation of NMO-corrected data along offset axis *interferes constructively*.

See Toldi 1989, Fowler 1986, Shen et al. 2005, Soubaras and Gratacos 2007.

Total Stack Power

Relation with semblance:

Recall *semblance*: simplest (unscaled) definition is

$$S[t_0, v_{\mathrm{RMS}}] = (\sum_h I(t_0, h))^2$$

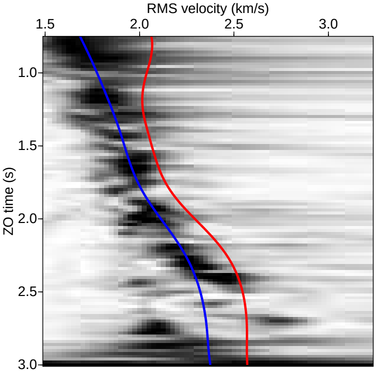
(i.e. a "square of sum" rather than "sum of squares").

NMO correction $d \rightarrow I$ performed with *constant* RMS velocity v_{RMS} - since $I(t_0, h)$ depends only on $v_{\mathrm{RMS}}(t_0)$, can be done economically.

Set $V_{\mathrm{RMS}}[v]$ = RMS velocity from interval velocity v. Then

$$J_{\mathrm{TSP}}[v, d] = \int_0^{t_{\max}} dt_0 \, S[t_0, V_{\mathrm{RMS}}[v](t_0)].$$

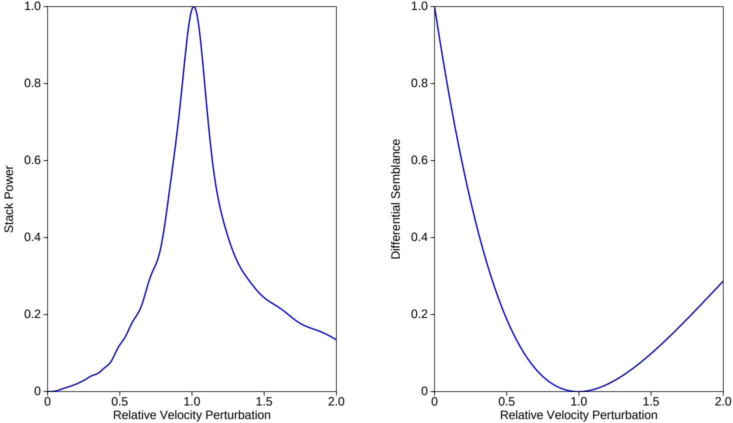
Total Stack Power



$J_{\mathrm{TSP}}[v,d]$ = sum of semblance S along $t_0, V_{\mathrm{RMS}}[v](t_0)$ trajectory.

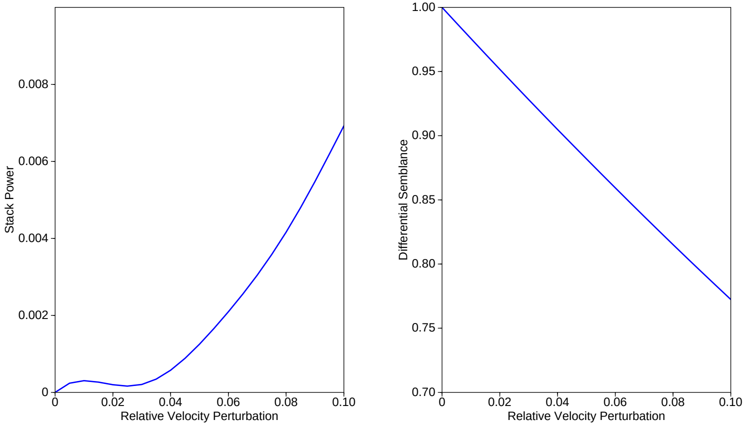
RICE

Total Stack Power



Comparison: Total Stack Power (left), Differential Semblance (right), sampled along line segment in interval velocity space.

Total Stack Power



Detail at left side of previous figures.

Total Stack Power

Summary of accumulated evidence re TSP:

- TSP objective strongly peaked at optimum velocity choice, relatively flat elsewhere.
- TSP prone to local minima at grossly incorrect velocities: see Chauris and Noble 2001 for "proof" that this is intrinsic property of TSP objective.
- Chief consequence: gradient methods to optimize TSP require good initial guess of velocity - and it's hard to say *a priori* exactly how good, or whether any particular velocity is good enough!
- However: peak shape and location insensitive to coherent noise for near-correct velocities, so robust in that sense.

Theory for DS-NMO

Theorem: J is smooth as function of velocity parameters. For data with sufficiently high S/N, all stationary points of J are approximate global minima (S., 1999, 2001).

- Loose translation: no local mins.
- S/N refers to *fidelity to the convolutional model* - best RMS approximation by forward convolutional modeling (relation to least squares data fitting!).
- Stationary points approximate global minima in sense of high-frequency asymptotics - *direct link between bandwidth and velocity resolution.*
- Numerical experience as predicted by Theorem.

Theory for DS-NMO

A step in the proof, of independent interest: assuming sufficient offset sampling,

$$D_h I(t_0, h) = D_h d(T(t_0, h), h)$$

$$\simeq \frac{\partial d}{\partial t}(T(t_0, h), h)\frac{\partial T}{\partial h}(t_0, h) + \frac{\partial d}{\partial h}(T(t_0, h), h)$$

so

$$D_h I(T_0(t, h), h) = \left(P\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h}\right)(t, h),$$

in which $P(t, h) = \frac{\partial T}{\partial h}(T_0(t, h), h)$ is the offset ray parameter.

Theory for DS-NMO

Assume d is convolutional model data for an *target* velocity v^*, for convenience with $\delta(t)$ wavelet. That is,

$$d(t,h) = r(T_0^*(t,h)).$$

Easy calculus exercise:

$$\frac{\partial T_0^*}{\partial h} = -P^*,$$

the offset ray parameter for the target velocity. Thus

$$D_h I(T_0(t,h),h) \simeq \left((P - P^*)\frac{\partial d}{\partial t}\right)(t,h).$$

Theory for DS-NMO

Upshot

$$J[v,d] = \frac{1}{2}\sum_h \int_0^{t_{\max}} dt |D_h I(T_0(t,h),h)|^2,$$

$$\simeq \frac{1}{2}\sum_h \int_0^{t_{\max}} dt \left((P-P^*)\frac{\partial d}{\partial t}(t,h)\right)^2,$$

which leads to a tomographic interpretation of DS: objective is data-weighted error in offset ray parameter.

$\Rightarrow$ connection to stereotomography - Chauris & Noble 2001.

Coherent Noise

DS is *tomographic* $\Leftrightarrow$ detects *moveout* $\Rightarrow$ moveout ambiguity (coherent noise, eg. multiples) must degrade velocity estimate.

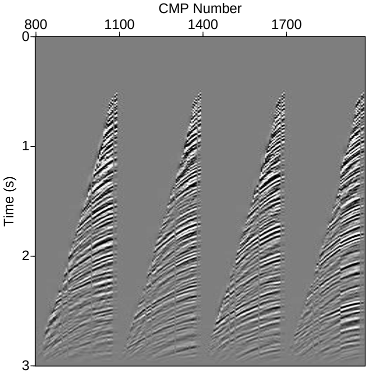
Example: NAM data shown previously consists mostly of multiples - they are nicely flattened, but lead to absurdly slow velocity below 2 s.

Li & S. 2007: Study based on Mobil "AVO" (Viking Graben) data, Keys & Foster 1998.

Upshot:

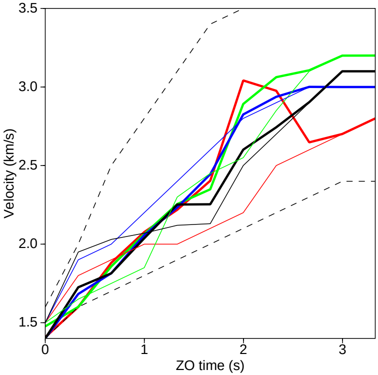
- DS finds *weighted average* of apparent moveout velocities (least squares!);
- if moveout dichotomy exists, can use in conjunction with dip filter to remove noise, enhance velocity estimation.

Coherent Noise



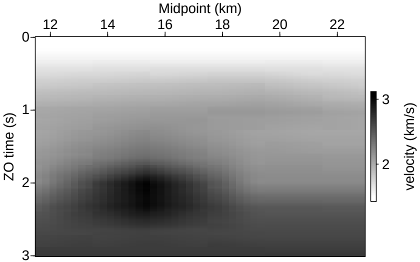
Four CMPs from middle part of Mobil AVO line: hyperbolic Radon demultiple, low-pass filter, mute.

Coherent Noise



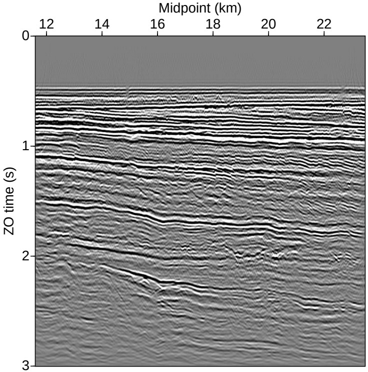
Result of four DS estimations (bold lines) with four initial guesses (thin lines).

Coherent Noise



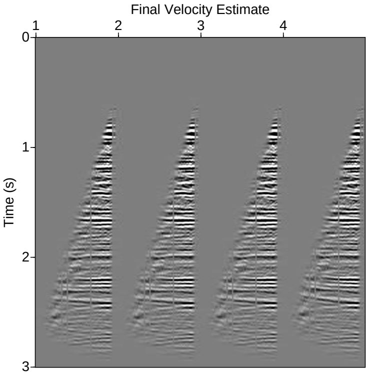
2D velocity model constructed by interpolating estimations from midpoints 800, 1100, 1400, and 1700.

Coherent Noise



Poststack migrated stack of entire line, stacked with 2D velocity model - fails to image clearly graben and fault block below Jurassic-Cretaceous unconform. at about 2-2.4 s.

Coherent Noise



The culprit: DS has *averaged* apparent velocities - deeper slow events undercorrected, fast events overcorrected.

Coherent Noise

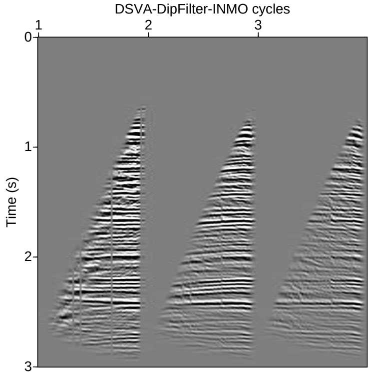
For structure with mild lateral heterogeneity, multiply reflected events are likely *slower* than primary events.

Diagnosis: slow events in Mobil AVO are predominantly residual multiple energy.

Mulder & ten Kroode, 2002: apply dip filter to NMO corrected gathers to remove downward sloping (slow) events. Then remodel data (using convolutional model = inverse NMO), reapply DS. [DS based on Kirchhoff common offset prestack migration]

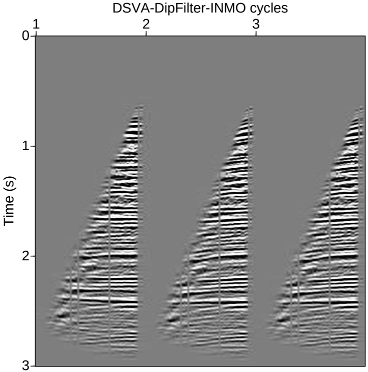
Li & S.: make this into an iteration, apply in NMO setting.

Coherent Noise



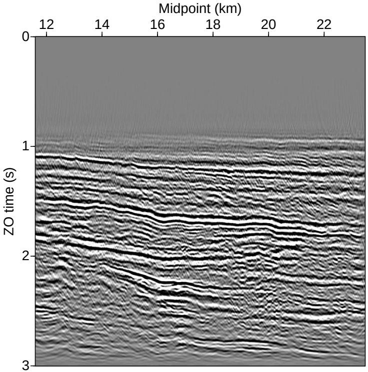
Result of 2 DS-NMO-dip filter-INMO cycles: filtered and NMO corrected. Note that 2.4 s event remains, but most other energy initially flattened by DS has been removed!

Coherent Noise



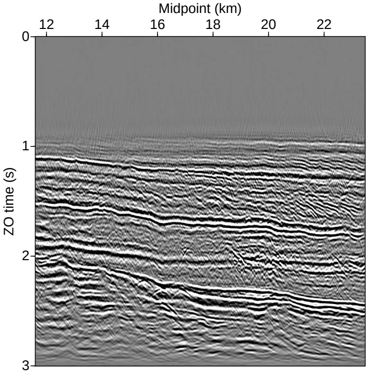
Unfiltered data, NMO-corrected using velocities produced at iterations 1, 2, and 3. Note that most visible energy is now undercorrected!

Coherent Noise



Poststack-migrated far-offset stack, original 2D velocity model from one application of DS.

Coherent Noise



Poststack-migrated far-offset stack, 2D velocity from 3 iterations of DS - dip filter - INMO..

Summary: Differential Semblance VA based on NMO

- VA formulated as optimization problem via DS
- for low-noise data, DSVA converges rapidly to kinematically accurate velocity estimate
- this observation supported by theory: all stationary points are global mins
- result sensitive to coherent noise
- ad hoc coherent noise suppression based on moveout-averaging property of DS successful in some cases

Suggests need for more fundamental and robust approach to coherent noise

RICE

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

RICE

Surface-oriented vs. Depth-oriented Image Volumes

MVA based on prestack depth migration - two major variants. Both produce *image volume* $I(\mathbf{x}, \cdot)$ depending on image point $\mathbf{x}$ and another (redundant) parameter.

(I) Surface oriented - $I_{SO}(\mathbf{x}, \mathbf{h})$: $\mathbf{h} = 0.5$(receiver - source), usually computed by diffraction sum ("Kirchhoff common offset migration"); binwise: offset bin $I(\cdot, \mathbf{h})$ depends only on data traces with offset $\mathbf{h}$.

(II) Depth oriented - $I_{\mathrm{DO}}(\mathbf{x}, \bar{\mathbf{h}})$: $2\bar{\mathbf{h}} =$ difference between subsurface scattering points, $\mathbf{x} =$ their midpoint. Every point in image volume depends on all data traces. Has diffraction sum rep, but usually computed by one-way (shot profile or DSR) or two-way (RTM) wave extrapolation.

Surface-oriented vs. Depth-oriented Image Volumes

Definitions using Green's functions $G(\mathbf{x}, \mathbf{y}, t)$:

$$I_{SO}(\mathbf{x}, \mathbf{h}) = \int d\mathbf{m} \int dt \int d\tau (...)$$

$$G(\mathbf{m}+\mathbf{h}, \mathbf{x}, \tau) G(\mathbf{m}-\mathbf{h}, \mathbf{x}, t-\tau) d(\mathbf{m}, \mathbf{h}, t);$$

$$I_{DO}(\mathbf{x}, \bar{\mathbf{h}}) = \int d\mathbf{m} \int d\mathbf{h} \int dt \int d\tau (...)$$

$$G(\mathbf{m}+\mathbf{h}, \mathbf{x}+\bar{\mathbf{h}}, \tau) G(\mathbf{m}-\mathbf{h}, \mathbf{x}-\bar{\mathbf{h}}, t-\tau) d(\mathbf{m}, \mathbf{h}, t);$$

where $(...)$ = amplitude terms.

Surface-oriented vs. Depth-oriented Image Volumes

Subsitute *asymptotic approximation*

$$G(\mathbf{x},\mathbf{y},t) \simeq A(\mathbf{x},\mathbf{y})\delta(t - T(\mathbf{x},\mathbf{y}))$$

(or sum of such, if multiple arrival times $\mathbf{x} \rightarrow \mathbf{y}$ exist) to get diffraction sum ("Kirchhoff", "Generalized Radon Transform") representations.

True amplitude image volumes for appropriate choice of amplitude terms (Beylkin 1985 etc. etc.)

Surface-oriented vs. Depth-oriented Image Volumes

Mathematical justification: related to *Born scattering* approximation (Cohen & Bleistein 1977, Beylkin 1985, Rakesh 1988,...):

$d \simeq$ *perturbation of acoustic field* caused by *oscillatory* model (velocity, density,...) perturbation about *smooth* background model.

NB: nonsmooth model $\Rightarrow$ much more complicated diffraction and multiple reflection effects, many open theoretical and practical questions.

Kinematic Artifacts

Kinematic artifact = coherent event in image volume not corresponding to physical reflector (i.e. to event in model perturbation).

Nolan & S. 1997: SO image volume for common shot migration typically contains kinematic artifacts when multiple ray paths connect sources, receivers with scatterers (points on reflectors). Used RTM to illustrate $\Rightarrow$ phenomenon has *nothing a priori* to do with Kirchhoff/GRT representation.

Brandsberg-Dahl & de Hoop 2003, Stolk & S. 2004: same phenomenon afflicts Kirchhoff common offset and common scattering angle migration.

Geometry of Reflection

Analysis relies on *geometric optics reflection rule* (Rakesh 1988 - canonical relation of scattering operator):

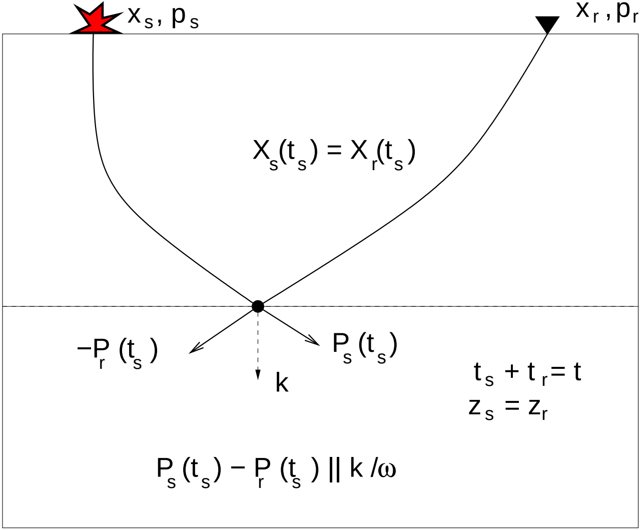
Event at $(\mathbf{m}, \mathbf{h}, t)$ (equiv. to $(\mathbf{x}_s, \mathbf{x}_r, t)$) with (3D) slowness vectors $(\mathbf{p}_m, \mathbf{p}_r)$ (equiv. to $(\mathbf{p}_s, \mathbf{p}_r)$) related to reflector at $\mathbf{x}$ with dip $\mathbf{p} = \mathbf{k}/\omega$ if an only if there exist

- incident ray $\mathbf{X}_s, \mathbf{P}_s$ with $\mathbf{X}_s(0) = \mathbf{x}_s = \mathbf{m} - \mathbf{h}, \mathbf{P}_\mathbf{s}(0) = \mathbf{p}_s = \mathbf{p}_m - \mathbf{p}_h$,
- reflected ray $\mathbf{X}_r, \mathbf{P}_r$ with $\mathbf{X}_r(t) = \mathbf{x}_r = \mathbf{m} + \mathbf{h}, \mathbf{P}_\mathbf{r}(t) = \mathbf{p}_r = \mathbf{p}_m + \mathbf{p}_h$,

so that for some $0 \le t_s \le t$,

- $\mathbf{X}_s(t_s) = \mathbf{x} = \mathbf{X}_r(t_s)$, and
- $\mathbf{P}_s(t_s) - \mathbf{P}_r(t_s) = \mathbf{p}$.

Geometry of Reflection



Kinematic Artifacts in Common Offset Image Volume

Data input for $I_{SO}(\mathbf{x}, \mathbf{h}) =$ offset bin $\{$traces with offset $\mathbf{h}\}$

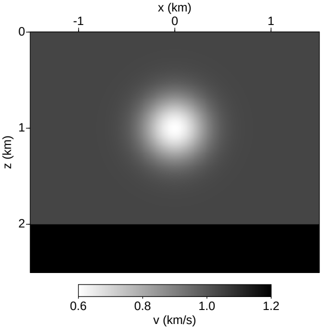
$\Rightarrow$ offset component $\mathbf{p}_h$ of event slowness *not determined by data*

$\Rightarrow$ several ray pairs may satisfy imaging conditions for event in offset bin (same $\mathbf{h}$, different $\mathbf{p}_h$)

$\Rightarrow$ event in offset bin may correspond to *several* scatterers.

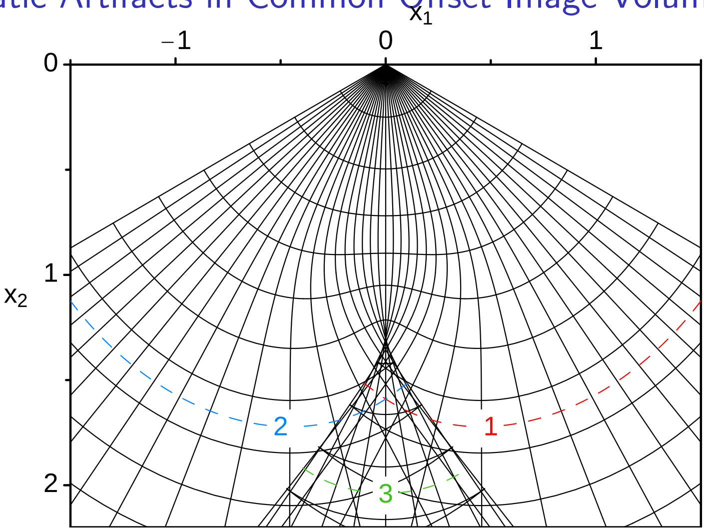
Kinematic Artifacts in Common Offset Image Volume

Example:



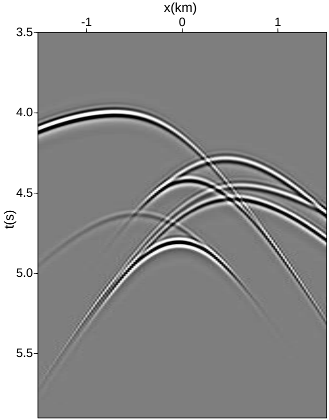
Acoustic lens - Gaussian low-velocity anomaly - over half-space.

Kinematic Artifacts in Common Offset Image Volume



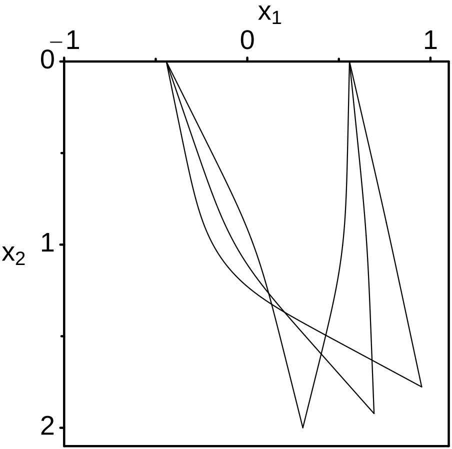
Rays from source at center of spread.

Kinematic Artifacts in Common Offset Image Volume



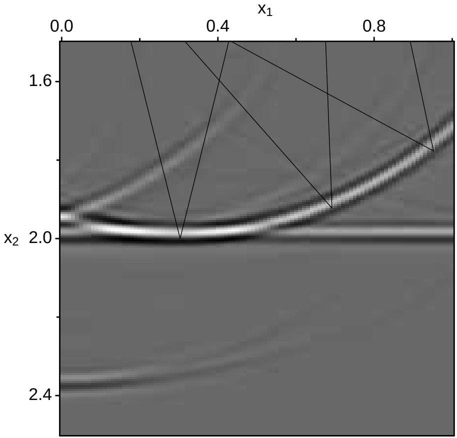
Shot record - source point slightly to right of center. Fixed receiver spread. (2,4) finite difference, zerophase bandpass filter wavelet. Note multiple events at each receiver - ONE reflector!

Kinematic Artifacts in Common Offset Image Volume



Several ray pairs satisfying geometric optics reflection rule for offset = 0.3 km

Kinematic Artifacts in Common Offset Image Volume



Common offset Kirchhoff image for bin at offset = 0.3 km.

Artifact suppression in Depth-Oriented Image Volumes

Stolk & de Hoop 2001 (tech report), 2005, 2006: I_{DO} is free of kinematic artifacts, provided that

- all rays carrying significant energy have monotone depth components (no turning rays, "DSR condition");
- depth offsets $\bar{\mathbf{h}}$ are restricted to horizontal: $\bar{h}_z = 0$;
- data sufficient to determine all components of event slownesses.

"DSR" condition also permits use of *depth extrapolation* algorithms, eg. DSR imaging algorithm.

Generalization beyond DSR imaging, conditions 1, 2: Stolk, de Hoop & S 2005.

RICE

Artifact suppression in Depth-Oriented Image Volumes

3rd condition - relation between acquisition geometry, ray geometry. Satisfied for

- 2D synthetics;
- pure dip shooting;
- true 3D coverage (all azimuths).

Generally I_{DO} derived from narrow azimuth survey will contain artifact energy when subsurface geometry is complex.

Artifact suppression in Depth-Oriented Image Volumes

Ray geometry analysis of I_{DO}:

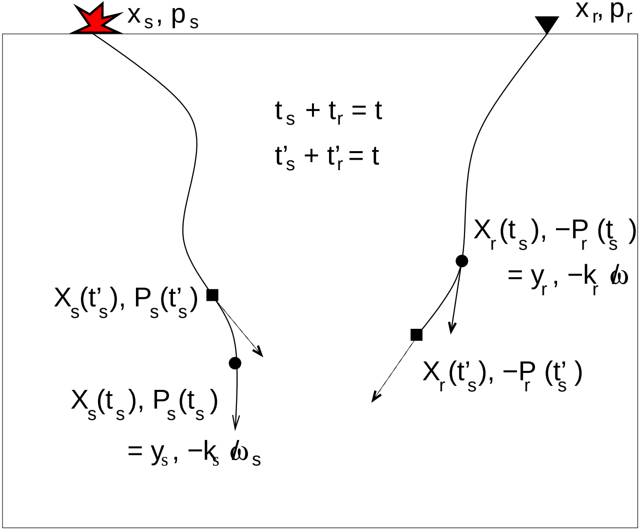
Event in d at $(\mathbf{m}, \mathbf{h}, t)$ (equiv. to $(\mathbf{x}_s, \mathbf{x}_r, t)$) with (3D) slowness vectors $(\mathbf{p}_m, \mathbf{p}_r)$ (equiv. to $(\mathbf{p}_s, \mathbf{p}_r)$) related to event in I_{DO} at $\mathbf{x}, \bar{\mathbf{h}}$ with dip $\mathbf{p} = \mathbf{k}/\omega, \mathbf{p}_{\bar{\mathbf{h}}}$if an only if there exist

- incident ray $\mathbf{X}_s, \mathbf{P}_s$ with $\mathbf{X}_s(0) = \mathbf{x}_s = \mathbf{m} - \mathbf{h}, \mathbf{P}_\mathbf{s}(0) = \mathbf{p}_s = \mathbf{p}_m - \mathbf{p}_h,$
- reflected ray $\mathbf{X}_r, \mathbf{P}_r$ with $\mathbf{X}_r(t) = \mathbf{x}_r = \mathbf{m} + \mathbf{h}, \mathbf{P}_\mathbf{r}(t) = \mathbf{p}_r = \mathbf{p}_m + \mathbf{p}_h,$

so that for some $0 \le t_s \le t$,

- $\mathbf{X}_s(t_s) = \mathbf{x} - \bar{\mathbf{h}}; \mathbf{X}_r(t_s) = \mathbf{x} + \bar{\mathbf{h}}$, and
- $\mathbf{P}_s(t_s) = \mathbf{p} - \mathbf{p}_{\bar{\mathbf{h}}}; -\mathbf{P}_r(t_s) = \mathbf{p} + \mathbf{p}_{\bar{\mathbf{h}}}.$

Artifact suppression in Depth-Oriented Image Volumes



Artifact suppression in Depth-Oriented Image Volumes

No constraint on t_s! Clearly image volume is *too big*: contains *path* of image events for each data event

$$\mathbf{x} = \frac{\mathbf{X}_r(t_s) + \mathbf{X}_\mathbf{s}(t_s)}{2},\ \bar{\mathbf{h}} = \frac{\mathbf{X}_r(t_s) - \mathbf{X}_\mathbf{s}(t_s)}{2}.$$

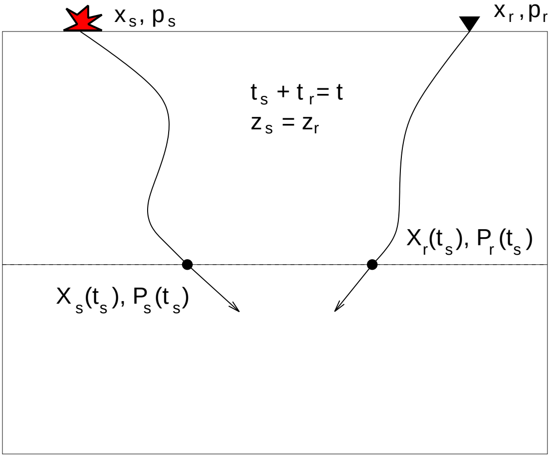
In 3D, image volume is 6D - but data is 5D. Need addtional constraint.

Original idea of Claerbout (1971, 1985): $(\mathbf{x}, \bar{\mathbf{h}})$ represent midpoint, offset of sunken survey.

Natural restriction: offset vector should be horizontal, i.e. $\bar{h}_z = 0$.

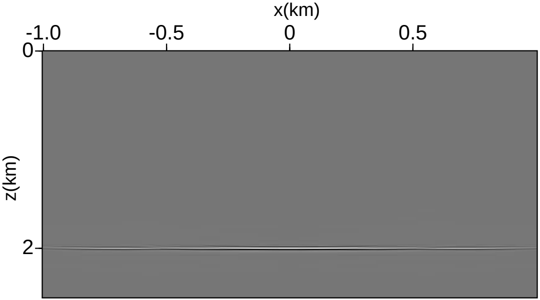
Artifact suppression in Depth-Oriented Image Volumes

"DSR" condition: rays carrying significant energy do not turn - depth is monotone increasing along ray. Then $\bar{h}_z = 0 \Rightarrow$ unique solution of data-image relation.



Artifact suppression in Depth-Oriented Image Volumes

Lens example: Image slice $I_{DO}(\mathbf{x}, \bar{\mathbf{h}})$ at depth offset $\bar{h}_x = 0.3$ km, computed using DSR algorithm (generalized screen propagator) - Stolk, de Hoop and S. 2005.



Artifacts and (stacked) Images

Get image of subsurface from image volume by

- *stacking* surface oriented volume:

$$I(\mathbf{x}) = \int d\mathbf{h}\, I_{SO}(\mathbf{x}, \mathbf{h}),$$

- *extracting zero-offset section* from depth oriented volume:

$$I(\mathbf{x}) = I_{DO}(\mathbf{x}, \mathbf{0}).$$

NB: these are exactly the same!

Artifacts and (stacked) Images

Smit et al. 1998, Nolan & S. 1997: image $I(\mathbf{x})$ contains reflectors only at correct (physical) locations, orientations if velocity satisfies

Traveltime Injectivity Condition ("TIC"): total time uniquely determines intersection point of any (incident, reflected) ray pair.

Clearly implied by "DSR" condition. Not always satisfied: counterexamples involve strong refraction, approximate waveguide geometry.

Essence of proof, for SO volumes: artifacts *stack out*.

Summary: Kirchhoff vs. Wave Equation Image Volumes

- two widely-used methods of image volume formation - differentiated by kinematics, *not* by method of computation! (are there others?)
- surface-oriented imaging (expl: common offset Kirchhoff) tends to produce kinematic artifacts = spurious coherent events, when multiple raypaths carry significant energy
- depth-oriented imaging (expl: DSR prestack migration) avoids kinematic artifacts in many circumstances

RICE

Agenda

Semblance

Semblance condition: expresses consistency between data, velocity model in terms of image volume.

- Surface oriented: $I_{SO}(\mathbf{x}, \mathbf{h})$ independent of $\mathbf{h}$ (at least in terms of phase);
- Depth oriented: $I_{DO}(\mathbf{x}, \bar{\mathbf{h}})$ concentrated (focused) near $\bar{\mathbf{h}} = \mathbf{0}$.

Main principle of waveform MVA: adjust velocity until image volume satisfies semblance condition.

Semblance

Visual assessment of semblance via *image gathers*:

- $I_{SO}(\mathbf{x}, \mathbf{h})$ for fixed $x, y (\mathbf{x} = (x, y, z)) \Rightarrow$ function of $z, \mathbf{h}$ - should be flat, i.e. independent of z (at least in phase);
- $I_{DO}(\mathbf{x}, \bar{\mathbf{h}})$ for fixed $x, y \Rightarrow$ function of $z, \bar{\mathbf{h}}$- should be focused at $\bar{\mathbf{h}} = 0$, so far as bandwidth permits.

Note:

- NMO correction is crude approximation to I_{SO}, and standard semblance criterion (flat NMO-corrected CMPs) is special case;
- I_{DO} can be converted to function of offset ray parameter or (equivalently) angle (Sava & Fomel 2002) - then gathers should be flat, like I_{SO}.

Semblance

Kinematic artifacts violate semblance condition!

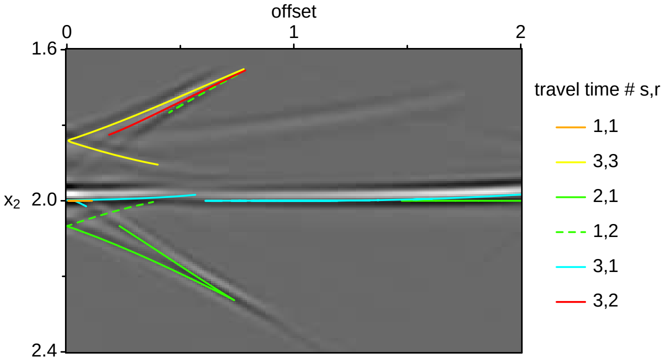
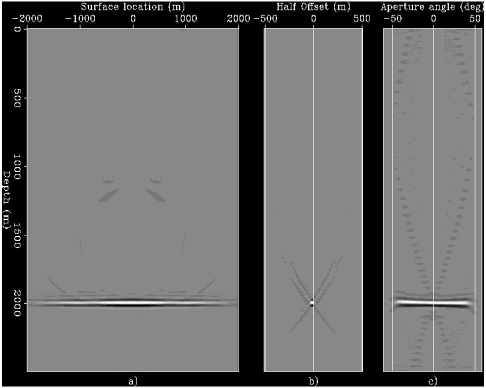


Image gather, Kirchhoff common offset migration of lens data (I_{SO}).

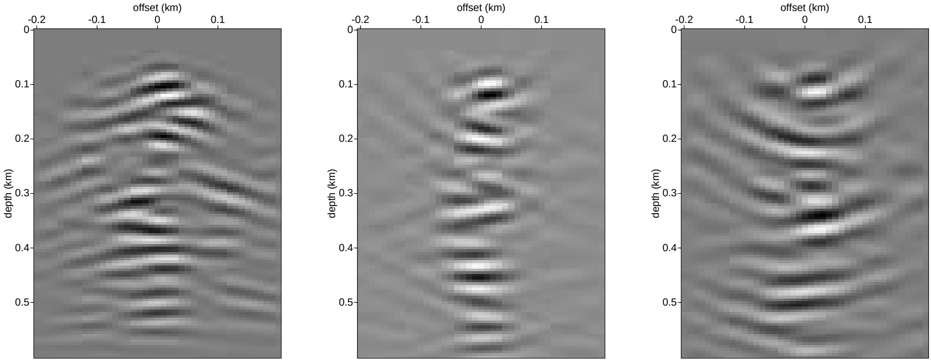
RICE

Semblance



Depth-oriented image volume avoids artifacts: DSR migration of lens data (thanks: B. Biondi). Left: image ($I_{DO}(\mathbf{x}, \mathbf{0})$); Center: image gather - note focus at zero offset; Right: angle-transformed image gather - flat!.

Semblance



OIG, x=1 km: vel 10% low   Offset Image Gather, x=1 km   OIG, x=1 km: vel 10% high

Semblance property has nothing to do with reflector geometry! Depth oriented image gathers ($I_D(\mathbf{x}, \bar{\mathbf{h}})$) via RTM, data synthesized from randomly distributed point diffractors. Left to Right: migration velocity = 90%, 100%, 110% of true velocity.

Implications for MVA

- Kinematic artifacts violate semblance condition;
- Surface oriented image volumes prone to kinematic artifacts in presence of multiple raypaths to scattering points;
- Depth oriented image volumes free of kinematic artifacts, under some circumstances.

Suggests: depth-oriented volume possibly better domain for MVA in complex, refracting subsurface.

RICE

MVA via Optimization

Goal: use all events in data, weighted by strength.

- form objective function of velocity, measuring deviation of image volume from semblance condition - all energy not conforming to semblance condition contributes.
- optimize numerically: gradient = backprojection of semblance-inconsistent energy into velocity update.

Because of problem size, must use gradient-based (Newton-like) method.

Use of gradient $\Rightarrow$ objective function must be *smooth* in velocity - stability of minimizer requires smoothness in data as well.

MVA via Optimization

Form of objective dictated by semblance condition for each type of image volume:

- Surface oriented: minimize over v

$$J[v] = \int dx \int dh \, |\nabla_{\mathbf{h}} I(\mathbf{x}, \mathbf{h})|^2$$

(generalization of DS objective for NMO-based VA);

- Depth-oriented, minimize over v

$$J[v] = \int dx \int dH |\bar{\mathbf{h}} I(\mathbf{x}, \bar{\mathbf{h}})|^2$$

MVA via Optimization

Stolk & S. 2003: Amongst all quadratic forms in image volume, *only* differential semblance is smooth as function of both velocity and data.

[proved for SO, conjectured for DO]

$\Rightarrow$ only DS suitable for large-scale numerical optimization. No other choices possible!

Caveat: this is *high frequency asymptotic* result: other functionals are smooth in velocity for finite frequency data (stack power, least squares). However only DS has *stable shape* as frequency is increased to obtain better resolution.

RICE

DSMVA via Depth Extrapolation

DSR migration:

1. downward continue data by solving

$$\frac{\partial u}{\partial z} = F_2[v]u,\ \ u(x_s, x_r, 0, t) = d(x_s, x_r, t),$$

where $x_s, x_r r$ are horizontal (x, y) coordinates of "sunken source and receivers", and

$$F_2[v] = -\sqrt{v(x_s, z)^{-2}\partial_t^2 - \nabla_{x_s}^2} - \sqrt{v(x_r, z)^{-2}\partial_t^2 - \nabla_{x_r}^2}.$$

2. Extract zero time section: $I(x, z, h) = u(x - h, x + h, z, 0)$.

DSMVA via Depth Extrapolation

Shot profile migration:

1. For each x_s, downward continue source and receiver fields u_s, u_r:

$$\frac{\partial u_s}{\partial z} = F_1[v]u_s,\ \ u_s(x_s, x, 0, t) = \delta(x_s - x)\delta(t);$$

$$\frac{\partial u_r}{\partial z} = F_1[v]u_r,\ \ u_r(x_s, x, 0, t) = d(x_s, x, t);$$

where

$$F_1[v] = -\sqrt{v(x,z)^{-2}\partial_t^2 - \nabla_x^2}.$$

2. Cross-correlate u_s, u_r at zero time lag, sum over sources:

$$I(x,z,h) = \int dx_s \int dt u_s(x_s, x-h, z, t) u_r(x_s, x+h, z, t)$$

DSMVA via Depth Extrapolation

Derivations from Green's function definition: de Hoop and Stolk 2006, my short course on imaging (TRIP web site).

To compute fields, objective, gradient, must discretize. *Huge* literature on approximation of square root operators.

Here: derivation of practical gradient computation (Shen et al. SEG 2003, Khoury & S. SEG 2006) for DSR case.

Shot profile computations similar (Shen et al. SEG 2005).

DSMVA via Depth Extrapolation

Abstractly, $u_n(x_s, x_r, t) \simeq u(x_s, x_r, n\Delta z, t)$ solves

$$u_{n+1} = \Delta z \Phi_n[v] u_n, \;\; n = 0, ..., N_z - 1$$

where Φ represents one of many DSR propagators (PSPI, FFD, GSP,...).

Discrete objective:

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

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

DSMVA via Depth Extrapolation

Perturbation field δu_n solves linearized depth evolution:

$$\delta u_{n+1} = \delta u_n + \Delta z \Phi_n[v]\delta u_n + \delta\Phi_n[v]u_n,\ \ n = 0, ..., N_z - 1$$

Adjoint field $w_n(s, r, t)$ = solution of *adjoint state system*

$$w_{n-1} = w_n + \Delta z \Phi[v]^* w_n + P^* P u_n,\ \ n = N_z, ...1;\ \ w_{N_z} = 0,$$

(*upward continuation!*) - $\Phi[v]^*$ = *adjoint* or transpose of $\Phi[v]$.

DSMVA via Depth Extrapolation

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

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

$$= \sum_n \sum_{x_s,x_r,t} w_n \delta\Phi_n[v] u_n$$

DSMVA via Depth Extrapolation

Note that $\delta\Phi[v]u = \partial_v(\Phi[v]u)\delta v$,

Define $\Psi[v,u] =$ adjoint of $\delta v \mapsto \partial_v(\Phi[v]u)\delta v$. Then

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

whence

$$\nabla_v J[v,d] = \sum_n \Psi[v,u_n]w_n$$

DSMVA via Depth Extrapolation

Summary:

1. downward continue data $\Rightarrow u_n, n = 0, ..., N_z$;
2. upward continue adjoint field $\Rightarrow w_n, n = N_z, ..., 0$;
3. cross-correlate w_n with $\Psi[v, u_n]$, zero lag:

$$\nabla_v J[v, d] = \sum_n \Psi[v, u_n] w_n$$

[version of *adjoint state method* - similar to RTM.]

DSMVA via Depth Extrapolation

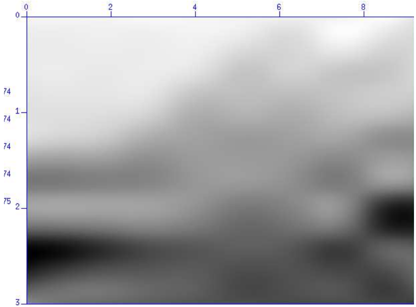
Example (Shen et al. 2005) - uses shot profile migration, computations very similar.

Data synthesis: smoothed Marmousi velocity model v, reflectivity $\delta v =$ difference (Marmousi - smoothed Marmousi), one-way *demigration* by upward continuation (solution of the equation for δu_n above).

Sources, receivers occupy all surface positions (not marine geometry!).

Objective function, gradient computation fed to Limited Memory BFGS algorithm (Nocedal & Wright 1999, available from Netlib).

DSMVA via Depth Extrapolation



Starting velocity model for DS-SP.

DSMVA via Depth Extrapolation

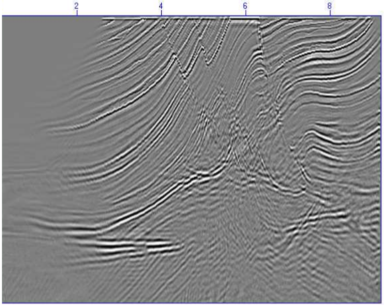
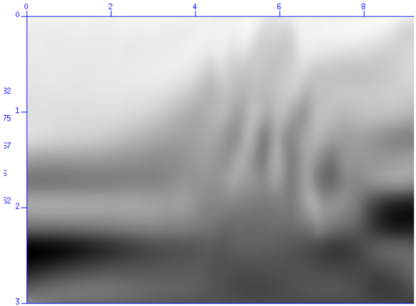


Image ($I_{DO}(\mathbf{x}, \bar{\mathbf{h}} = 0)$) at initial velocity.

DSMVA via Depth Extrapolation



Final velocity (47 iterations of descent method). Note appearance of high velocity fault blocks.

DSMVA via Depth Extrapolation

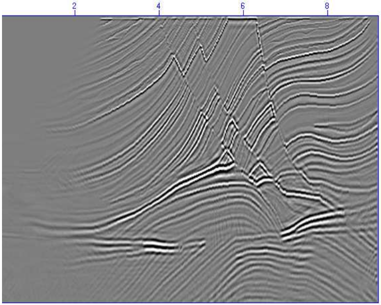


Image ($I_{DO}(\mathbf{x}, \bar{\mathbf{h}} = 0)$) at final velocity.

Sensitivity to Propagator Error

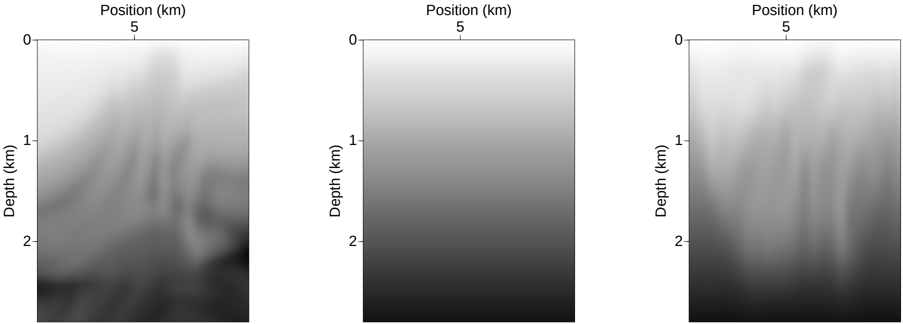
Khoury et al. SEG 2006: DSR-based implementation.

This example: based on Marmousi, with smoothed velocity model and a sequence of flat reflectors

Data generation: time-domain (2,4) FD scheme, bandpass filter wavelet (point isotropic radiator).

Objective function and gradient computation precisely as above. Propagator: GSP. Reference velocity taken to be lower bound for all estimated velocities, velocity bounds implemented via sigmoid representation. LBFGS used to optimize.

Sensitivity to Propagator Error



Velocity: target (Left), initial (Center), and after 30 LBFGS iterations (Right).

Sensitivity to Propagator Error

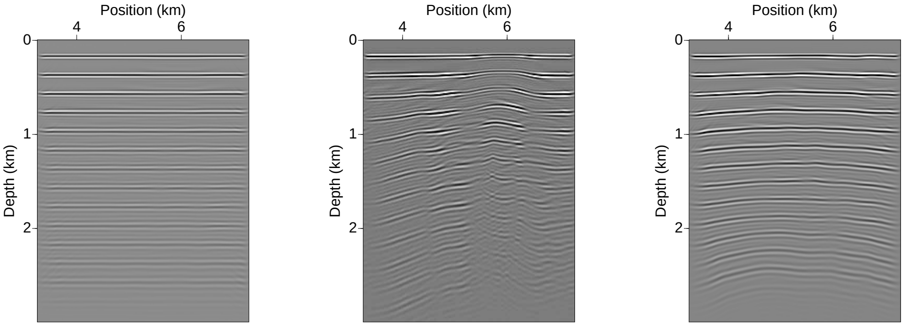


Image ($I_{DO}(x, z, 0)$): target (Left), initial (Center), and after 30 LBFGS iterations (Right).

Sensitivity to Propagator Error

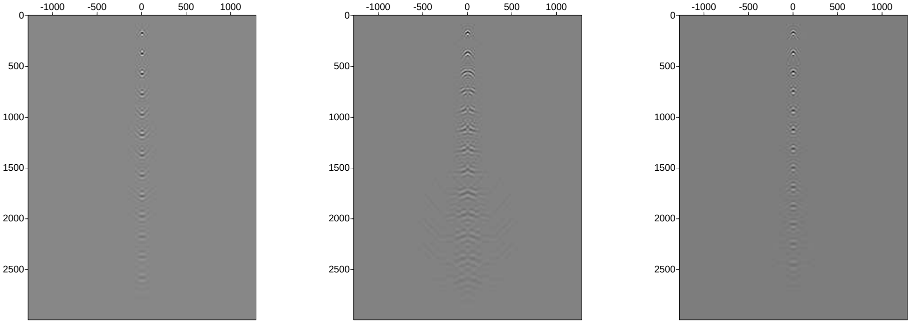
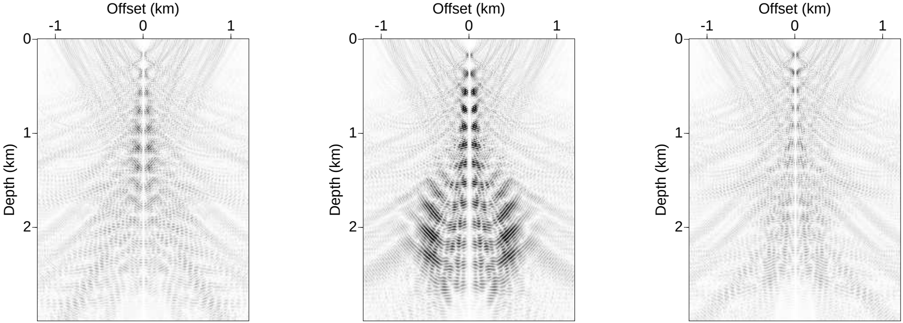


Image gathers (($I(x, z, h)$ for fixed x: with target vel (Left), initial (Center), and after 30 LBFGS iterations (Right).

Sensitivity to Propagator Error



DS gathers ($|HI(x,z,H)|^2$ for fixed x: with target vel (Left), initial (Center), and after 30 LBFGS iterations (Right).

Sensitivity to Propagator Error

Conclusion: DS is sensitive to high-angle propagation error, which acts as coherent noise.

Solutions:

1. produce better image volume (more kinematically consistent) - RTM?

2. modify objective to be less sensitive to this type of noise.

Modified DS

Shen noticed same problem in shot profile case - high-angle errors in FFD propagators led to shifted DS optima. Especially true for nonsmooth background (violates theory!).

His solution: modify DS functional.

Shen's modification - add multiple of image power, robust against imaging noise for near-correct velocity:

$$J_{MDS}[v] = \sum_{x,z,h} |hI(x,z,h)|^2 - \beta^2 \sum_{x,z} |I(x,z,0)|^2$$

See Shen & S. 2008.

Modified DS

Gas chimney example from Shen & S. 2008 (see also Kabir et al. 2007 for similar):

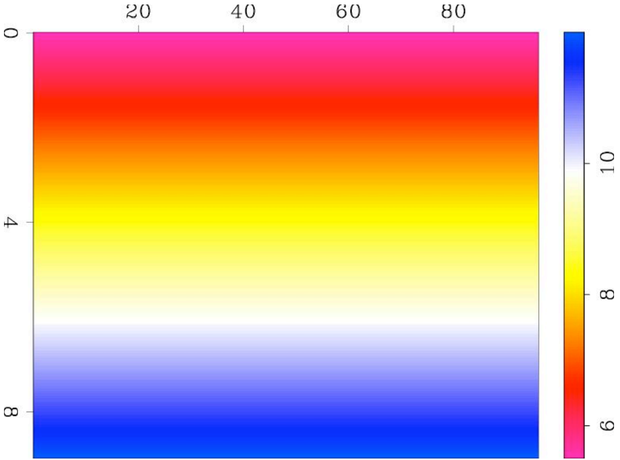
Marine 2D line - preliminary imaging with regional velocity model shows gas-induced sag.

Reflection tomography partially removes sag effect, but interpreters not happy.

MDS to rescue - 20 iterations of Newton-like optimization algorithm produces more interpretable model, image.

[Iterative algorithm follows Shen's PhD thesis - adjoint state method for gradient computation.]

Modified DS



Initial Velocity Model for MDS

Modified DS

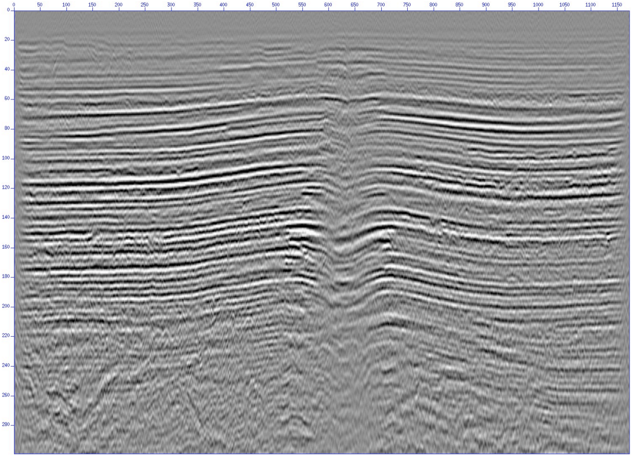
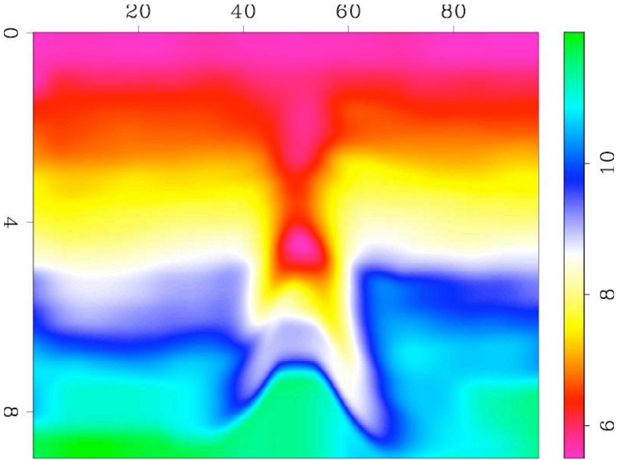


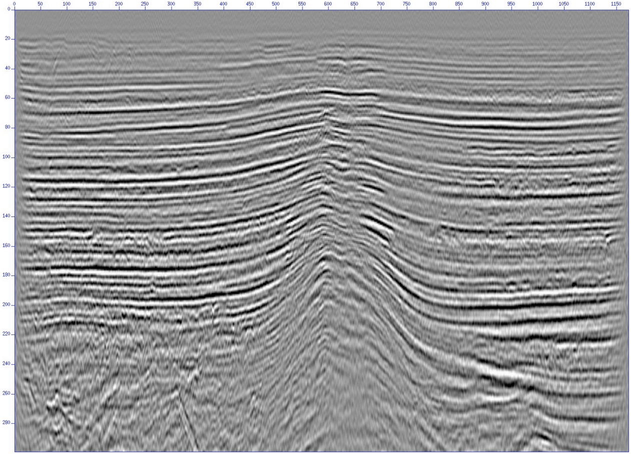
Image at Initial Model

Modified DS



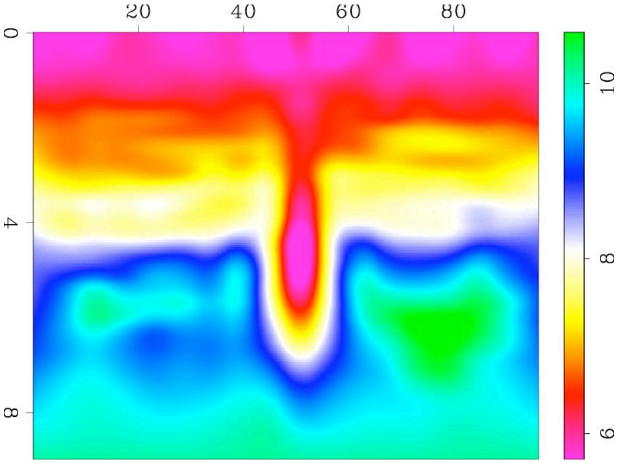
Model produced by Reflection Tomography

Modified DS



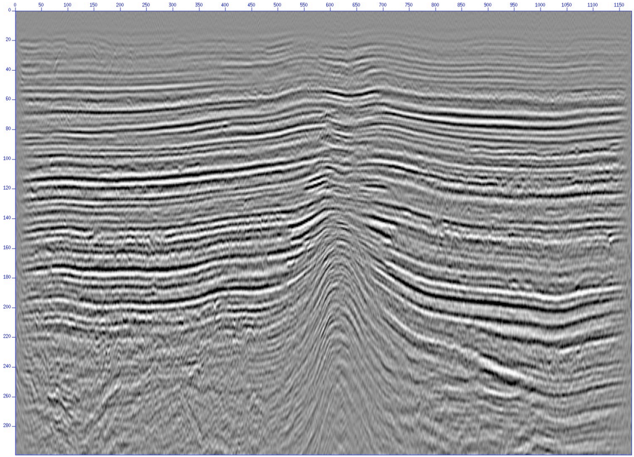
Reflection Tomography Image

Modified DS



Model produced by 20 MDS Iterations

Modified DS



MDS Image

Modified DS

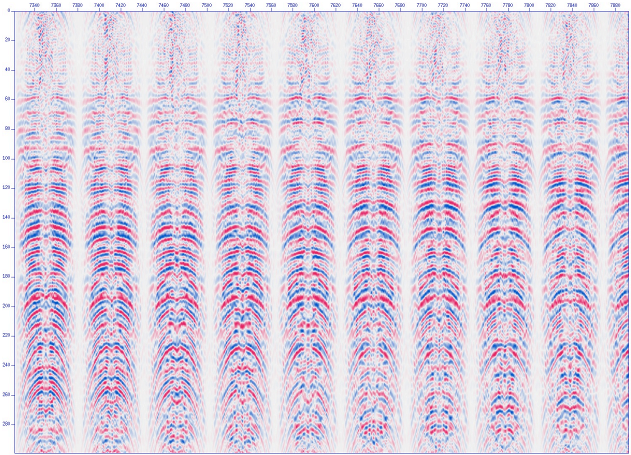
Angle image gathers (Sava & Fomel 03) - Radon transform in depth/offset, should be flat at correct velocity.

Initial velocity - dramatic failure to flatten.

RT velocity - much better, but RMO at larger depths.

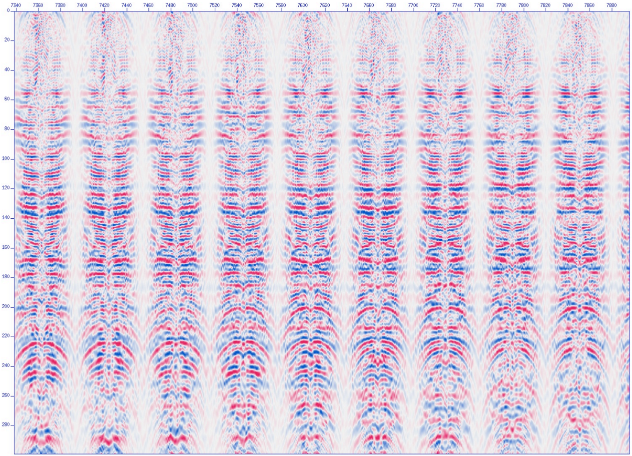
MDS velocity - flat throughout depth range.

Modified DS



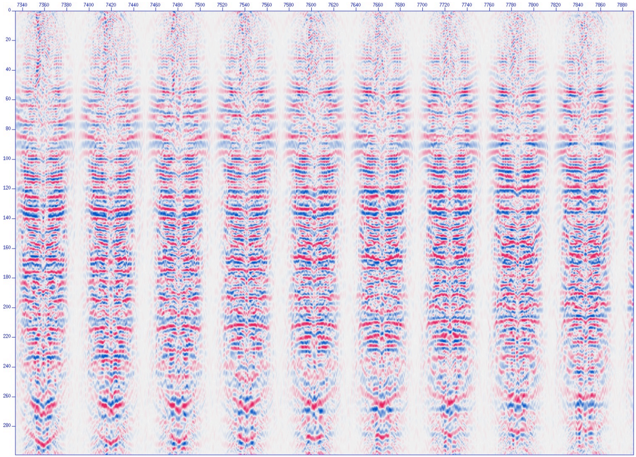
ADCIGs, Initial Model

Modified DS



ADCIGs, RT Model

Modified DS



ADCIGs, MDS Model

Summary: Differential Semblance MVA

- amongst all possible ways to turn MVA into an optimization problem, only DS results in smoothly varying function of velocity, independently of data frequency content
- semblance principle for depth-oriented imaging: *focused* image gathers (in depth-offset - flat in depth-angle)
- depth-oriented imaging preferred over surface oriented imaging for MVA: kinematic artifacts of latter violate semblance principle
- DSMVA based on "multiply by $\bar{\mathbf{h}}$" method of measuring semblance discrepancy: demonstrated for synthetic, field data with significant multipathing
- DSMVA result sensitive to depth extrapolator errors, sensitivity can be reduced by Shen's modification (stack power as penalty term)

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

MVA + WI

Waveform MVA - all very well, but...

- MVA based on *linearized modeling*, generally acoustic, neglects multiple reflections, mode conversions, out-of-plane events (2D), anisotropy, attentuation,...
- DS based on WE migration sensitive to extrapolator error (high angle velocity error & numerical anisotropy - Khoury 2006).
- DS strongly influenced by coherent noise (multiples, mode conversions,...) - Gockenbach & S 1999, Mulder & ten Kroode 2001, Verm & S 2006, Li & S 2007.
- Repeated modeling, migration - s l o w.

A blue-sky approach to overcoming these obstacles: combine DS & *full waveform* 3D modeling, incorporate necessary physics, accelerate convergence.

Waveform Inversion

The usual set-up:

- $\mathcal{M}$ = a set of *models*;
- $\mathcal{D}$ = a Hilbert space of (potential) data;
- $\mathcal{F} : \mathcal{M} \rightarrow \mathcal{D}$: modeling operator or "forward map".

Waveform inversion problem: given $d \in \mathcal{D}$, find $v \in \mathcal{M}$ so that

$$\mathcal{F}[v] \simeq d$$

.

$\mathcal{F}$ can incorporate *any physics* - acoustics, elasticity, anisotropy, attenuation,.... (and v may be more than velocity...).

Waveform Inversion

Least squares formulation: given $d \in \mathcal{D}$, find $v \in \mathcal{M}$ to minimize

$$J_{LS}(v,d) = \frac{1}{2}\|d - \mathcal{F}[v]\|^2$$

$$\equiv \frac{1}{2}(d - \mathcal{F}[v])^T(d - \mathcal{F}[v])$$

Has long and productive history in geophysics (eg. reflection tomography)- but not in reflection waveform inversion.

Problem size $\Rightarrow$ Newton and relatives $\Rightarrow$ find local minima.
BUT....

Waveform Inversion

Albert Tarantola, many others: J_{LS} has lots of useless local minima, for typical length, time, and frequency scales of exploration seismology

$\Rightarrow$ least squares waveform inversion with Newton-like iteration "doesn't work" - can't assure convergence from reasonable initial estimates.

See: Gauthier et al. 1986, Kolb et al. 1986, Santosa & S. 1989, Bunks et al. 1995, Shin et al. 2001, 2006, Chung et al 2007.

RICE

Waveform Inversion

Postmortem on J_{LS}: missing low frequency data is culprit - obstructs estimation of large-scale velocity structure, hence everything else.

Rule of thumb derived from layered Born modeling:

to estimate velocity structure on length scale L, with mean velocity v, data must have significant energy at

$$f_{\min} \simeq \frac{v}{2L}$$

$v \sim 3$ km/s, $L = 3$ km $\Rightarrow$ need good s/n at $0.5 Hz$ - not commonly available.

Waveform Inversion

Caveats: least squares WI computational feasible

- with synthetic data containing very low frequencies ($<< 1$ Hz): Bunks et al. 1995, Shin et al. 2006.
- for basin inversion from earthquake data: target of several major efforts. QuakeShow (Ghattas), SpecFEM3D (Tromp, Komatisch), SPICE (Käser, Dumbser). Typical $L \sim 20$ km, $f_{\min} = 0.1Hz$, $v \sim 4$ km/s - just OK! Will be done, in 3D, in near future.
- for *transmission* waveform inversion (cf Gauthier et al. 1986) with good initial v from traveltime tomography (plus other tweaks) - Pratt 1999, Pratt and Shipp 2004 Brenders and Pratt, 2007.

Waveform Inversion

When you can make it work, does it really work?

That is, can you learn something from WI that you can't learn otherwise?

Affirmative example: Minkoff & S 1997: inverted small part of very clean G-of-M 2D line.

Inverted for all important physics - kinematic accuracy via DSO, layered viscoelastic primaries-only modeling of P-P reflections, source radiation pattern - and 90% fit of data energy in pre-multiple window $\Rightarrow$ identified gas sand (p, but not s, anomaly) - with less complete physics in any way, neither fit to data nor correct gas signature.

Lesson: to gain from inversion, must model all important physics and fit data accurately.

Waveform Inversion

Recap: roughly speaking, with appropriate “fine print”,

- MVA can be successfully cast as an optimization problem - all stationary points are approximate global mins, so can use Newton;
- WI (in usual OLS form) afflicted with spurious local mins - descent methods often fail to fit data, unless you start with a very good initial guess - and it's impossible to know *a priori* how good an initial guess is!

Can WI technology borrow from MVA - conversely, can MVA be recast as an approach to data-fitting?

The crucial WI - MVA link: extended modeling

Extended model $\bar{\mathcal{F}} : \bar{\mathcal{M}} \rightarrow \mathcal{D}$, where $\bar{\mathcal{M}}$ is a *bigger model space*= models depending on $\mathbf{x}$ *and* $\bar{\mathbf{h}}$, i.e. $\bar{v}(\mathbf{x}, \bar{\mathbf{h}})$. [$\bar{\mathbf{h}}$ *may* be offset, or maybe something else (shot coordinates, ray parameter,...) - *redundant degrees of model freedom*.]

Extension map $\bar{\mathcal{E}} : \mathcal{M} \rightarrow \bar{\mathcal{M}}$: identifies physical (normal) model with extended model.

Extension property: $\mathcal{F}[v] = \bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]$.

Annihilator: $\bar{\mathcal{A}} : \bar{\mathcal{M}} \rightarrow ...$ - characterizes physical models: $\bar{\mathcal{A}}[\bar{v}] = 0 \Leftrightarrow \bar{v} = \bar{\mathcal{E}}[v]$, for some $v \in \mathcal{M}$.

The crucial WI - MVA link: extended modeling

Example: surface oriented (common offset,...) extended modeling

$\bar{\mathcal{M}}$ = offset dependent velocities $\bar{v}(\mathbf{x},\mathbf{h})$

Extension op: $\bar{\mathcal{E}} : \mathcal{M} \rightarrow \bar{\mathcal{M}}$ by $\bar{\mathcal{E}}[v](\mathbf{x},\mathbf{h}) = v(\mathbf{x})$ - that is, extended models *don't depend on* **h** (sound familiar?)

Extended modeling op: $\bar{\mathcal{F}}[\bar{v}](\mathbf{h}, t, \mathbf{x}_s) = \bar{p}(\mathbf{x}_s + 2\mathbf{h}, t; \mathbf{x}_s)$, where

$$\frac{1}{\bar{v}^2(\mathbf{x},\mathbf{h})}\frac{\partial^2 \bar{p}}{\partial t^2}(\mathbf{x}, t; \mathbf{x}_s) - \nabla^2 \bar{p}(\mathbf{x}, t; \mathbf{x}_s) = w(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

that is, solve wave equation with (possibly) different velocity for each offset.

Annihilator: $\bar{\mathcal{A}} = \nabla_{\mathbf{h}}$ - *differential semblance*.

The crucial WI - MVA link: extended modeling

The link:

- Waveform inversion: find $v \in \mathcal{M}$ so that $\mathcal{F}[v] \simeq d \Leftrightarrow$

$$\mathrm{find}\, \bar{v} \in \bar{\mathcal{M}}\, \mathrm{so\, that}\, \bar{\mathcal{F}}[\bar{v}] \simeq d\, \mathrm{subject\, to}\, \bar{\mathcal{A}}[\bar{v}] = 0$$

- Nonlinear version of migration velocity analysis:

$$\mathrm{find}\, \bar{v} \in \bar{\mathcal{M}}\, \mathrm{so\, that}\, \bar{\mathcal{A}}[\bar{v}] \simeq 0\, \mathrm{subject\, to}\, \bar{\mathcal{F}}[\bar{v}] = d$$

Same, except that objective and constraint are switched - a form of duality. Continuum of problems in between these (penalty function in Claerbout notation):

$$\mathrm{find}\, \bar{v} \in \bar{\mathcal{M}}\, \mathrm{so\, that}\, \bar{\mathcal{F}}[\bar{v}] \simeq d,\; \epsilon\bar{\mathcal{A}}[\bar{v}] \simeq 0$$

Born extended modeling and MVA

Lailly, Tarantola, Claerbout (80's): migration operator (producing image) is *adjoint* or transpose $D\mathcal{F}[v]^T$. *True amplitude* migration is (pseudo)inverse $D\mathcal{F}[v]^{-1}$.

Same for extended modeling $\bar{\mathcal{F}}[\bar{v}]$:

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^T d(\mathbf{x},\bar{\mathbf{h}}) = I(\mathbf{x},\bar{\mathbf{h}})$$

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^{-1} d(\mathbf{x},\bar{\mathbf{h}}) = \delta\bar{v}(\mathbf{x},\bar{\mathbf{h}}).$$

Born extended modeling and MVA

As for full nonlinear modeling have dual points of view: replace $\bar{v}$ by $\bar{\mathcal{E}}[v] + \delta\bar{v}$, use perturbation theory

- "Partly linearized" waveform inversion: find $v \in \mathcal{M}, \delta\bar{v} \in \bar{\mathcal{M}}$ so that $\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v} \simeq d \Leftrightarrow$

  $$\text{find}\, v \in \mathcal{M},\, \delta\bar{v} \in \bar{\mathcal{M}}\, \text{so that}\, D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v} \simeq d\, \text{subject to}\, \bar{\mathcal{A}}[\delta\bar{v}] = 0$$

- Annihilator form of migration velocity analysis:

  $$\text{find}\, v \in \mathcal{M},\, \bar{v} \in \bar{\mathcal{M}}\, \text{so that}\, \bar{\mathcal{A}}[\delta\bar{v}] \simeq 0\, \text{subject to}\, D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v} = d$$

Continuum of intermediate problems (Gockenbach et al. 1995):

$$\text{find}\, v \in \mathcal{M}, \delta\bar{v} \in \bar{\mathcal{M}}\, \text{so that}\, D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v} \simeq d,\, \epsilon\bar{\mathcal{A}}[\delta\bar{v}] \simeq 0$$

Born extended modeling and MVA

Example, continued: surface-oriented extended Born modeling at physical model v for common offset - $D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\bar{\delta v}(\mathbf{h}, t, \mathbf{x}_s) = \delta\bar{p}(\mathbf{x}_s + 2\mathbf{h}, t; \mathbf{x}_s)$, where

$$\frac{1}{v^2}\frac{\partial^2 \delta\bar{p}}{\partial t^2} - \nabla^2 \delta\bar{p} = \frac{2\delta\bar{v}}{\bar{v}^3}\frac{\partial^2 \bar{p}}{\partial t^2}$$

Express solution via Green's function G:

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\bar{\delta v}(\mathbf{h}, t, \mathbf{x}_s) =$$

$$\int d\mathbf{x} \int d\tau\, G(\mathbf{x}, t-\tau, \mathbf{x}_s + 2\mathbf{h}) G(\mathbf{x}, \tau, \mathbf{x}_s) \frac{2\delta\bar{v}(\mathbf{x}, \mathbf{h})}{v^3(\mathbf{x})}$$

$=$ common offset Born modeling, offset dep. reflectivity $= \frac{2\delta\bar{v}}{v^3}$. [For simplicity, drop convolution with source pulse, i.e. assume $w \simeq \delta$]

Born extended modeling and MVA

Example, continued: read common offset prestack migration from modeling formula:

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^T d(\mathbf{x},\mathbf{h}) = I(\mathbf{x},\mathbf{h}) =$$

$$\int d\mathbf{x}_s dt \int d\tau\, G(\mathbf{x}, t-\tau, \mathbf{x}_s + 2\mathbf{h}) G(\mathbf{x}, \tau, \mathbf{x}_s) d(\mathbf{x}_s + 2\mathbf{h}, t, \mathbf{x}_s)$$

(becomes Kirchhoff CO prestack with geometric optics approx to G). From Beylkin 1985, Rakesh 1988, Bleisten 1987, Nolan & S. 1997, Smit et al. 1998, deHoop & Bleistein 1998: can turn this into inversion ($D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^{-1}$) via amplitude factor under integral sign)

Born extended modeling and MVA

Example, continued: surface-oriented annihilator-based MVA and/or penalty function version:

- Versteeg & S. 1993, Kern & S. 1994: $\mathbf{h} \leftarrow \mathbf{x}_s$ (redundant parameter is shot position), finite difference modeling and reverse time migration
- Chauris & Noble 2001, Mulder & ten Kroode 2002, de Hoop et al. 2003: Kirchhoff common offset or scattering angle migration
- deHoop et al. 2005: VTI P-P and P-S Kirchhoff imaging, angle domain

Only result on *nonlinear MVA = annihilator based WI* so far is theoretical: S. 1991. Computational exploration in thesis project of Dong Sun - stay tuned!

RICE

Summary: combining MVA and WI

- WI admits arbitrary physics, accounts directly for nonlinear wave propagation effects (multiples)
- when WI works, very accurate reconstruction of subsurface models, information not available otherwise (demonstrated at least once using field data!)
- usual OLS formulation of WI susceptible to multiple (spurious) local mins - often does not succeed in fitting data, producing plausible model
- conceptual connection between MVA and WI through *extended model*
- leads to nonlinear MVA = extended WI - familiar as DSMVA in extended Born approximation

RICE

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

Depth-oriented extended modeling

Recall: depth-oriented imaging avoided kinematic artifacts, supports MVA in complex refractive environments

Observe: surface oriented imaging operator $= D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^T$, where $\bar{\mathcal{F}}$ is surface-oriented extended modeling operator.

Question: how to choose extended model, annihilator $(\bar{\mathcal{M}}, \bar{\mathcal{F}}, \bar{\mathcal{E}}, \bar{\mathcal{A}})$ corresponding to depth-oriented imaging?

Depth-oriented extended modeling

Answer: v becomes an *operator*:

$$\bar{v}^{-2}\frac{\partial^2 p}{\partial t^2}(\mathbf{x},t) = \int d\bar{\mathbf{h}}\,\bar{v}^{-2}(\mathbf{x},\bar{\mathbf{h}})\frac{\partial^2 p}{\partial t^2}(\mathbf{x}+2\bar{\mathbf{h}},t)$$

Wave equation still has sensible solutions, defines $\bar{\mathcal{F}}$ as before.

Physical interpretation of operator $\bar{v}^{-2}$: since $v^{-2} = \rho\kappa^{-1}$ and κ is stress per unit strain, $v^{-2}(\mathbf{x},\bar{\mathbf{h}})$ = density-weighted strain at $\mathbf{x}$ due to point stress at $\mathbf{x}+2\bar{\mathbf{h}}$ -represents *action at a distance* [thanks: Scott Morton]

$\bar{\mathcal{M}} = \{$ (depth-)offset dependent velocity $\bar{v}(\mathbf{x},\bar{\mathbf{h}})$ $\}$

$\bar{\mathcal{E}}[v](\mathbf{x},\bar{\mathbf{h}}) = v(\mathbf{x})\delta(\bar{\mathbf{h}}),\ \bar{\mathcal{A}}[\bar{v}] = \bar{\mathbf{h}}\bar{v}$

Depth-oriented extended modeling

Born version of depth-oriented extended modeling: as before, replace $\bar{v}$ by $\bar{\mathcal{E}}[v] + \delta\bar{v}$, use perturbation theory

$$v^{-2}\frac{\partial^2 \delta p}{\partial t^2}(\mathbf{x},t) - \nabla^2 \delta p = -\int d\bar{\mathbf{h}}\delta\bar{v}^{-2}(\mathbf{x},\bar{\mathbf{h}})\frac{\partial^2 p}{\partial t^2}(\mathbf{x}+2\bar{\mathbf{h}},t)$$

Represent using ordinary Green's functions, assume *ramp* source so $\partial^2 p/\partial t^2 = G(\mathbf{x}_s,\cdot,\cdot)$,

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v}(\mathbf{m},\mathbf{h},t) =$$

$$\int d\mathbf{x}\int d\mathbf{h}d\tau\, G(\mathbf{m}+\mathbf{h},t-\tau,\mathbf{x})G(\mathbf{m}-\mathbf{h},\tau,\mathbf{x}-2\bar{\mathbf{h}})\delta\bar{v}^{-2}(\mathbf{x},\bar{\mathbf{h}})$$

Depth-oriented extended modeling

Depth-oriented Born extended modeling/inversion, continued:

Rewrite integral using *midpoint* $\mathbf{x} \leftarrow \mathbf{x} - \bar{\mathbf{h}}$, reflectivity $R(\mathbf{x}, \bar{\mathbf{h}}) = \delta \bar{v}^{-2}(\mathbf{x} + \bar{\mathbf{h}}, \bar{\mathbf{h}})$:

$$\int d\mathbf{x} d\mathbf{h} d\tau \, G(\mathbf{m} + \mathbf{h}, \mathbf{x} + \bar{\mathbf{h}}, t - \tau) G(\mathbf{m} - \mathbf{h}, \mathbf{x} - \bar{\mathbf{h}}, \tau) R(\mathbf{x}, \bar{\mathbf{h}})$$

Adjoint (imaging) operator is exactly the depth oriented imaging op of Claerbout - see p. 45

$$D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]^T d(\mathbf{x}, \bar{\mathbf{h}}) = I_{DO}(\mathbf{x}, \bar{\mathbf{h}})$$

$$= \int d\mathbf{m} \int d\mathbf{h} \int dt \int d\tau \, G(\mathbf{m}+\mathbf{h}, \mathbf{x}+\bar{\mathbf{h}}, t-\tau) G(\mathbf{m}-\mathbf{h}, \mathbf{x}-\bar{\mathbf{h}}, \tau) d(\mathbf{m}, \mathbf{h}, t)$$

Depth-oriented extended modeling

Upshot: annihilator form of depth-oriented MVA (Shen et. al 2003, 2005, Kabir 2007, Shen & S. 2008) derived from depth-oriented extended model

Shen's modified DSMVA (Shen 2005, Shen & S. 2008) essentially same as penalty form:

$$\mathrm{find}\, v \in \mathcal{M}, \delta\bar{v} \in \bar{\mathcal{M}}\, \mathrm{so\, that}\, D\bar{\mathcal{F}}[\bar{\mathcal{E}}[v]]\delta\bar{v} \simeq d,\, \epsilon\bar{A}[\delta\bar{v}] \simeq 0$$

through close relation of semblance maximization and linear least squares minimization.

Conclusion: the Born version of this approach seems to work, so what about...

Depth-oriented extended modeling

Nonlinear MVA, or extended WI

$$\text{find } \bar{v} \in \bar{\mathcal{M}} \text{ so that } \bar{\mathcal{F}}[\bar{v}] \simeq d, \; \epsilon\bar{\mathcal{A}}[\bar{v}] \simeq 0$$

Apparent advantages:

- easy to formulate extended modeling op $\bar{\mathcal{F}}$ for virtually any wave physics - acoustic, elastic, ...
- includes all physical effects, including multiples
- in Born approximation, adjoints and (linear) inverses closely linked because of absence of kinematic artifacts
- MVA extended to elastic modeling in complex structures with multiples, for instance...

Depth-oriented extended modeling

Obvious disadvantages:

- wave equations with operator coefficients -

$$\bar{v}^{-2}\frac{\partial^2 p}{\partial t^2}(\mathbf{x},t) = \int d\bar{\mathbf{h}}\,\bar{v}^{-2}(\mathbf{x},\bar{\mathbf{h}})\frac{\partial^2 p}{\partial t^2}(\mathbf{x}+2\bar{\mathbf{h}},t)$$

  appear to require full matrix multiply at every timestep
- how do you solve the nonlinear extended inverse problem, $\bar{\mathcal{F}}[\bar{v}] \simeq d$, and how do you keep it solved as you reduce the annihilator output $\bar{\mathcal{A}}[\bar{v}]$?
- what is the right way to measure the annihilator output $\bar{\mathcal{A}}[\bar{v}]$?

Source Calibration

Another important issue: source calibration.

Patrick Lailly, Florence Delprat 2003, 2005: nonlinear inversion (any kind!) *demands* good knowledge of source - but for extremely complex media with intense internal multiples, very difficult to invert for source!

Contrast: Minkoff & S 1997, Winslow 1999, Anno et al. 2003: successful linearized inversion for source and reflectivity - moderately complex media

Summary: Challenges

- DSMVA based on depth-oriented Born extended modeling promising for primaries-only data
- Nonlinear MVA = WI based on depth-oriented nonlinear extended modeling formulated, but several obstacles remain to successful deployment
- For any form of WI, source calibration is a first-order issue which must be addressed before any possible successful application in the field.

Agenda

Overview

Differential Semblance VA

Image Volumes: Kirchhoff vs. Wave Equation

Differential Semblance MVA

Effective Waveform Inversion = Nonlinear MVA

Challenges

Conclusions

Conclusions

Takeaway messages of this talk:

- MVA = WI based on *Born extended modeling*
- WI via OLS often fails due to spurious local minima
- Differential semblance MVA - velocity updates via optimization, backprojection of waveform residuals (DS volume), all events constrain velocity updates, much less tendency towards local minima than least squares WI.
- "Kirchhoff" and "Wave Equation" prestack migrations have different *intrinsic* kinematic properties - latter better suited to automated MVA
- Extended modeling: framework for consistent formulation of MVA and WI
- Proposed nonlinear MVA = extended WI:
  - including multiples in MVA
  - removing multiple minima from WI