Exploitation de la cohérence locale des données sismiques pour l'imagerie du sous-sol

Hervé Chauris

Habilitation à Diriger les Recherches
Université Pierre et Marie Curie (Paris VI)
Spécialité : Sciences de la Terre et de l'Univers

Exploitation de la cohérence locale des données sismiques pour l'imagerie du sous-sol

Hervé Chauris
Equipe de Géophysique, Centre de Géosciences, UMR–Sisyphe 7619
Mines ParisTech
35, rue Saint-Honoré, 77300 Fontainebleau, France
e-mail : Herve.Chauris@mines-paristech.fr
http ://www.geophy.mines-paristech.fr

soutenue le 12 Avril 2010, devant le jury composé de

Pr. Michel DIETRICH (Institut Français du Pétrole) ..... rapporteur
Pr. William SYMES (Rice University) ..... rapporteur
Pr. Jean VIRIEUX (Université de Grenoble) ..... rapporteur
Dr. Gilles LAMBARE (CGGVeritas) ..... examinateur
Pr. Raúl MADARIAGA (Ecole Normale Supérieure) ..... examinateur
Dr. Stéphane OPERTO (Université de Nice – Sophia-Antipolis) ..... examinateur
Pr. Alain TABBAGH (Université Pierre et Marie Curie, Paris VI) ..... examinateur

Résumé

Mes travaux de recherche s'inscrivent dans l'imagerie sismique des premiers kilomètres du sous-sol, en particulier dans le domaine pétrolier. Il s'agit de reconstruire les propriétés de la sub-surface, par exemple la vitesse de propagation des ondes de compression, à partir d'enregistrements en surface de déplacements ou de variations de la pression liés au passage d'un train d'ondes acoustiques ou élastiques.

Je me suis attaché, dans ce domaine, à regarder différemment les données sismiques. Les enregistrements sismiques dépendent de la position de la source et d'un ensemble de récepteurs, ainsi que du temps d'enregistrement. Ils sont couramment analysés trace par trace, par exemple pour filtrer les données ou bien pour les migrer, c'est-à-dire retrouver en profondeur les perturbations du milieu, typiquement avec des migrations de type Kirchhoff. Dans le cas de la migration dite "reverse-time migration", les points de tir sont vus comme un ensemble : l'information de tous les récepteurs est rétro-propagée depuis la surface jusqu'en profondeur, également pour retrouver les propriétés du milieu. Je propose de considérer des groupes de traces adjacentes, autour d'une fenêtre en temps, et d'examiner différents algorithmes d'imagerie sous cet aspect. L'aspect cohérence locale est justifié dans le domaine pétrolier par les acquisitions actuelles qui sont de plus en plus denses, et aussi par la notion de zone de Fresnel.

La cohérence latérale des signaux sismiques est traitée sous deux angles. Le premier donne une vision très épurée : les événements localement cohérents sont décrits en 2D par la position de la trace centrale, une position dans la trace, et une pente qui indique la cohérence. Je montre qu'à partir de ce type d'approche, il est possible de retrouver les grandes longueurs d'onde du modèle de vitesse par analyse de vitesse dans le domaine migré. J'ai pu aussi établir le lien entre cette méthode d'analyse de vitesse et la tomographie de pente. Le second angle d'attaque prend en compte la signature des données sismiques, en particulier la bande passante limitée. Les applications sont alors beaucoup plus nombreuses : migration, analyse de vitesse, sensibilité de l'image migrée par rapport au modèle de vitesse, et d'autres tâches liées au pré-traitement comme le débruitage, la prédiction des multiples, ...

Dans chacun des cas, j'analyse les avantages et limitations de l'utilisation de la cohérence locale. Il s'avère que le choix du code de décomposition en événements locaux ("curvelets", ...) est étroitement lié à l'application qui en est faite : si l'objectif est de comprimer les données sismiques, alors les curvelets ne sont pas adaptées. S'il s'agit au contraire de savoir comment l'image migrée dépend du choix du modèle de vitesse pris pour la migration, alors les curvelets ont beaucoup d'avantages. Plus particulièrement, les curvelets offrent la flexibilité de la décomposition des données, très utile pour la suppression des bruits cohérents et pour la résolution de problèmes inverses. De plus, elles diagonalisent presque l'opérateur de démigration/migration qui donne la sensibilité de l'image migrée par rapport au choix du modèle de vitesse. Enfin, elles permettent d'exploiter l'aspect multi-échelle des données, avec des applications sur la suppression de l'aliasing. Je montre aussi qu'il est possible de développer d'autres schémas de décomposition et de reconstruction, comme par exemple les "circlets".

Dans les perspectives de recherche, je tiens à dépasser le cadre asymptotique hautes fréquences dans lequel les curvelets trouvent un cadre naturel. Je propose en particulier une nouvelle formulation du problème d'inversion des formes d'onde (DFWI, Differential Full Wa-

veForm Inversion). Cette méthode se veut générale pour retrouver les propriétés du sous-sol et permettre une optimisation locale, tout en s'affranchissant de la détermination d'un modèle de départ très proche de la solution. Cette idée doit être approfondie. Elle exploite la cohérence locale des données sismiques et suppose que le signal est bien échantillonné en temps et en espace. Les applications sont très nombreuses et touchent le domaine pétrolier, la géotechnique, la sismologie globale et régionale.

Mots-clés : Imagerie sismique, événements localement cohérents, curvelets, circlets, équation des ondes, migration, modélisation, inversion des formes d'onde.

Abstract

Title : Using the local coherency of seismic data for sub-surface imaging

My research activities deal with seismic imaging, mainly in the context of oil and gas exploration. The objective is to derive quantitative images of the sub-surface such as the P-wave velocity model, from seismic measurements recorded at the surface.

In that context, I endeavored to look at seismic data in a different manner. Seismic input gathers are function of the source and receiver positions and the recorded time. Individual traces are usually processed separately, for example in the case of Kirchhoff migration to derive the velocity contrasts at the correct position in the depth domain. For alternative methods, such as reverse-time migration, all traces within a single shot are simultaneously back-propagated from the surface. I propose to select a group of few adjacent traces localized around some time or depth values and to reconsider different imaging processing tasks from this perspective. The local aspect is justified in the context of oil and gas exploration with the increasingly dense acquisition surveys, but also with the notion of Fresnel zone.

The local coherency of seismic data is tackled through two approaches. Firstly, locally coherent events are simply described in 2D by the position of the central trace, a position within the trace and a slope indicating the local coherency. I demonstrate the feasibility to invert for the low frequency components of the velocity models by picking locally coherent events in the depth migrated section. I also show the equivalence between this approach and slope travel time tomography defined in the unmigrated time data. Secondly, I take into account the signature of the data, in particular the band-limited aspect, with the use of curvelets. More applications become then possible : migration, velocity analysis, sensitivity of the migrated section with respect to the velocity model, and other tasks related to the pre-processing steps such as denoising, multiple prediction, ...

In each case, I analyze the advantages and limitations of the notion of locally coherent events. It appears that the choice of the decomposition scheme, e.g. through curvelets, is crucial for subsequent processing. If for example one needs to compress seismic data, then curvelets are not useful. If on the contrary the objective is to derive how a given migrated section depends on the velocity model used for migration, then curvelets are powerful. More precisely, curvelets, with their flexible decomposition schemes, are useful to suppress coherent noise and to solve for inverse problems. They almost diagonalize the demigration/migration operator providing the sensitivity of a migrated section with respect to the velocity model. Finally, curvelets are beneficial to exploit the multi-scale aspect of seismic data, e.g. to remove aliased data. I also show the possibility to develop new decomposition and reconstruction schemes such as "circlets".

As research perspectives, my objective is to go beyond the high frequency approximation for which curvelets are naturally defined. I propose in particular a new formulation for the full waveform inversion, referred as the Differential Full WaveForm Inversion. This method aims at retrieving the sub-surface properties with a local optimization scheme, without the need to derive a precise initial velocity model closed to the solution. More work is needed. This method relies on the local coherency of data and supposes that the signal is well sampled in time and space. Potential applications deal with oil and gas exploration, geotechnics, global and regional seismology.

Keywords : Seismic imaging, locally coherent event, curvelets, circlets, seismic wave equation, migration, modeling, full waveform inversion.

Remerciements

Mes premiers remerciements vont aux membres du jury et tout d'abord au Pr. Alain Tabbagh, président du jury. Dès nos premiers contacts, il m'a toujours chaleureusement encouragé à soutenir l'HDR. Je voudrais le remercier pour sa disponibilité et son soutien, tant sur le plan scientifique que sur le plan administratif. Que le professeur Alain Saliot, président de la commission des HDR de Paris VI (Sciences de la Terre et de l'Univers) soit également remercié pour ses encouragements. Le Pr. Bill Symes était rapporteur de ma thèse de doctorat. Merci, outre ce nouveau rapport, de m'avoir suivi de près ou de loin pendant ces dix dernières années. Que les Pr Jean Virieux et Michel Dietrich soient également remerciés d'avoir volontiers porté une appréciation sur ce mémoire. Merci au professeur Raũl Madariaga qui a su, par des questions, me faire prendre du recul. Merci à Stéphane Operto. Nos chemins se sont croisés à plusieurs reprises et j'aurai du plaisir à le rencontrer de nouveau. Enfin, un remerciement tout spécial pour Gilles Lambaré avec qui, au-delà des contacts scientifiques, j'entretiens des contacts amicaux. Je lui dois beaucoup. Avec Mark Noble, ils m'ont introduit dans le monde de la géophysique, en particulier dans celui de la propagation des ondes et de la résolution de problèmes inverses. Gilles Lambaré a eu la gentillesse de relire une première version du mémoire, et m'a toujours, depuis le début de thèse de doctorat, encouragé et soutenu.

Beaucoup de travaux présentés dans ce mémoire sont récents. Ils n'auraient pas pu voir le jour sans un contexte favorable créé au sein l'équipe de Géophysique du centre de Géosciences de Mines Paristech. Que chacun soit remercié, en particulier Mark Noble cité plus haut, le pilier indispensable de l'équipe. Merci aussi à Alexandrine Gesret et Daniela Donno pour leur enthousiasme permanent. Elles y sont beaucoup pour le nouveau dynamisme de l'équipe. Un merci tout spécial à Véronique Lachasse, toujours très serviable, et qui est le second pilier de cette équipe. Merci à tous ceux qui aujourd'hui ou dans un passé récent ont fait partie de cette équipe : Imen, Sébastien, Jiangbo, Mathieu, Nidhal, Bingbing, Philippe, David, Akram, Jianwei, Mondher, Cédric, Truong, ... C'est à leur contact que j'ai beaucoup appris. Je tiens à souligner que l'équipe peut se développer grâce aux moyens et au support de Jean-Paul Chilès et de Damien Goetz du centre de Géosciences.

De par mon expérience, j'ai eu la chance de cotoyer beaucoup de personnes dans différents endroits. Que celles qui ne sont pas citées ici ne s'offusquent pas ! Merci en particulier à René-Edouard Plessix pour sa rigueur scientifique, à Fons ten Kroode, Frans Kuiper, Boudewijn Salomons, Pierre Garreau, Philippe Craneguy, ... Merci en particulier au Pr. Bernard Beaudoin qui a su développer mon attrait pour le monde de la recherche et de l'enseignement.

Pour conclure, je voudrais simplement dire le plaisir que j'ai eu à écrire ce mémoire. J'espère que le lecteur partagera ce plaisir.

Table des matières

Chapitre 1

Curriculum Vitae

Hervé CHAURIS
Né le 24 avril 1973, à Brest
Marié, 4 enfants
Nationalité Française

Fonction : maître-assistant
Equipe de Géophysique, Centre de Géosciences, Mines ParisTech, UMR–Sisyphe 7619
35, rue Saint-Honoré, 77300 Fontainebleau, France
Tel : +33 1 64 69 49 13, Fax : +33 1 64 69 49 35
e-mail : Herve.Chauris@mines-paristech.fr
http ://www.geophy.mines-paristech.fr, http ://www.geophy.mines-paristech.fr/~herve

1.1 Formation

- Septembre 1997 – Septembre 2000 : Thèse de doctorat, Ecole des Mines de Paris (formation doctorale GRN), mention très honorable, avec félicitations du jury. Directeur de thèse : Mark Noble
- Septembre 1996 – Juillet 1997 : DEA (Master II de recherche) sous la direction de Mark Noble : IFP / Ecole des Mines de Paris / UMPC / IPG / MNHN (Méthodes Quantitatives et Modélisation des Bassins Sédimentaires), mention TB, 1^{er}
- Septembre 1993 – Juillet 1996 : Ingénieur Civil, Ecole des Mines de Paris

1.2 Expérience professionnelle

- Depuis Septembre 2008 (mandat de 2 ans) : responsable de l'équipe de géophysique (centre de Géosciences) de Mines ParisTech
- Depuis Septembre 2005 : maître-assistant à Mines ParisTech
- Octobre 2000 - Août 2005 : Ingénieur de recherche dans le centre de recherche en imagerie sismique de Shell E&P (Rijswijk, Pays-Bas)

1.3 Encadrement d'étudiants

- Encadrement de Daniela Donno (post-doc) entre Mars 2009 et Mars 2011 : "Prédiction des multiples dans le domaine des curvelets" puis "Imagerie sismique des ondes de surface"
- Co-encadrement de Imen Karoui (avec Pierre Garreau, Ifremer, et Philippe Craneguy, Actimar), post-doc, entre Janvier 2009 et Juillet 2010 : "Détection et suivi des tourbillons en océanographie côtière"
- Encadrement de Truong Nguyen (post-doc) entre Juin 2006 et Octobre 2008 : "Imagerie sismique dans le domaine des curvelets : rai+Born"
- Encadrement de Mondher Benjemaa (post-doc) entre Mars 2008 et Août 2009 : "Démigration/migration avec l'équation des ondes"
- Co-encadrement (avec Mark Noble et Caroline Mehl) de Sébastien Penz (thèse de doctorat, Mines ParisTech) depuis Octobre 2009 : "Tomographie électrique"
- Encadrement de Jiangbo Liu (thèse de doctorat, Mines ParisTech) depuis Octobre 2009 : "Inversion de formes d'onde : vers une nouvelle formulation"
- Co-encadrement (avec Jianwei Ma) de Bingbing Sun (thèse de doctorat, Université de Tsinghua, Chine), depuis Octobre 2007 (thèse sur 5 ans) : "Modélisation sismique dans le domaine des curvelets"
- Co-encadrement (avec Mark Noble) de Cédric Taillandier (thèse de doctorat, Mines ParisTech), entre Octobre 2005 et Décembre 2008, sur le sujet "Formulation de la tomographie des temps de première arrivée à partir d'une méthode de gradient : un pas vers une tomographie interactive"

1.4 Enseignements

- Cours de "propagation des ondes sismiques" à l'Ecole du Pétrole et des Moteurs (IFP school), autour de 36 heures par an, niveau Master II
- Cours de "migration et amplitudes préservées" à l'Ecole du Pétrole et des Moteurs (IFP school), autour de 12 heures par an, niveau Master II
- Stage de géophysique de la sub-surface dans la région de Brest (1 semaine par an depuis 2006) pour les élèves de Mines ParisTech (troisième année)
- Stage de géologie de terrain à Digne-les-Bains, Alpes (2 semaines par an depuis 2005) pour les élèves de Mines ParisTech (première année)

1.5 Responsabilités collectives

- Responsable de l'équipe de géophysique depuis Septembre 2008, avec une nouvelle orientation scientifique "imagerie sismique terrestre" et le recrutement de deux nouveaux permanents scientifiques. En Décembre 2009, l'équipe était composée de quatre permanents scientifiques, d'un post-doc, de cinq étudiants en thèse et d'une secrétaire. Plus de détails sont donnés dans le chapitre 2.

- Responsable du projet entre Total et Mines ParisTech (Janvier 2010 – Décembre 2012) sur le sujet : "inversion des formes d'onde : vers une nouvelle formulation"
- Responsable du projet entre Shell et Mines ParisTech (Avril 2008 – Mars 2011) sur le sujet : "migration/démigration par équation d'onde"
- Responsable du projet entre l'Ifremer, Actimar et Mines ParisTech (Janvier 2009 – Juillet 2010) sur le sujet : "détection et suivi de tourbillons en océanographie côtière"
- Responsable du projet entre CGGVeritas et Mines ParisTech (Juin 2008 – Octobre 2008), sur le sujet : "les curvelets et la migration temps, évaluation sur données réelles"
- Responsable du projet entre Shell et Mines ParisTech (Janvier 2006 – Décembre 2007) sur le sujet : "imagerie sismique dans le domaine des curvelets : rai+Born"
- Shell : séminaire interne avec Gérard Herman (tous les 15 jours)

1.6 Divers

- Invitation de Jianwei Ma (professeur assistant à l'université de Tsinghua, Chine), entre Mars 2009 et Septembre 2009
- Invitation par l'EAGE "Distinguished Lecturer Program", 2009, pour donner une série de quatre conférences sur le thème des curvelets
- Organisation avec Huub Douma (Université de Princeton) d'un workshop à Londres lors de la conférence EAGE 2007 (une journée, 15 présentations, plus de 100 participants), sur le thème : "Curvelets, contourlets, seislets, ... in seismic data processing : where are we and where are we going ?"
- Membre de l'EAGE, SEG et du SIAM
- Relecteur pour les revues *Geophysics* et *Geophysical Prospecting*

1.7 Publications

1.7.1 Articles

- Chauris, H., Karoui, I., Garreau, P., Wackernagel, H., Craneguy, P., and Bertino, L. (2010). The circlet transform : a new tool for ocean eddy tracking. *Submitted to Computers and Geosciences*.
- Chauris, H., Donno, D., Nguyen, T., and Ma, J. (2010a). Modeling and imaging seismic data with curvelets : a review. *Submitted to Geophysics*.
- Donno, D., Chauris, H., and Noble, M. (2010). Curvelet-based multiple prediction. *Submitted to Geophysics*.
- Ma, J., Plonka, G., and Chauris, H. (2010). A new sparse representation of seismic data using adaptive easy-path wavelet transform. *Accepted for publication to Geoscience and Remote Sensing Letters*, 2009.
- Nguyen, T., Chauris, H., Liu, Y., and Oraintara, S. (2010). Low bit-rate image coding

based on multiresolution directional filter blanks. *In prep.*

- Nguyen, T. and Chauris, H. (2010). The Uniform Discrete Curvelet Transform. *Submitted to IEEE Transactions on Signal Processing.*
- Chauris, H. and Benjemaa, M. (2010). Seismic wave-equation demigration/migration. *Accepted for publication to Geophysics.*
- Taillandier, C., Noble, M., Chauris, H., and Calandra, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, 74(6) :WCB57–WCB66.
- Mi, T., Ma, J., Chauris, H., and Huizhu, Y. (2009). Multilevel adaptive mesh modeling for wave propagation in layer media. *Accepted for publication to Journal of Seismic Exploration.*
- Sun, B., Ma, J., Chauris, H., and Yang, H. (2009). Solving wave equations in the curvelet domain : a multi-scale and multi-directional approach. *Journal of Seismic Exploration*, 18(4) :385–399.
- Chauris, H. and Nguyen, T. (2008). Seismic demigration/migration in the curvelet domain. *Geophysics*, 73(2) :S35–S46.
- Nguyen, T., Liu, Y., Chauris, H., and Oraintara, S. (2008). Implementation aspects of the contourlet filter blank and application in image coding. *EURASIP Journal on Advances in Signal Processing*, ID 373287 :18 p.
- Le Bégat, S., Chauris, H., Devaux, V., Nguyen, S., and Noble, M. (2004). Comparison of three tomography methods on a 2-D real data set. *Geophysical Prospecting*, 52(5) :427–438.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002a). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I : theoretical aspects. *Geophysics*, 67 :1202–1212.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002b). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II : application on synthetic and real data. *Geophysics*, 67 :1213–1224.
- Xu, S., Chauris, H., Lambaré, G., and Noble, M. (2001). Common angle image gather : a strategy for imaging complex media. *Geophysics*, 66(6) :1877–1894.
- Chauris, H. and Noble, M. (2001). Two-dimensional velocity macro model estimation from seismic reflection data by local Differential Semblance Optimization : applications on synthetic and real data sets. *Geophysical Journal International*, 144 :14–26.
- Chauris, H., Le Rousseau, J., Beaudoin, B., Propson, P., and Montanari, A. (1998). Inoceramid acme/extinction in the gubbio region (Northeastern Apennines of Italy) and relations with environmental changes during the mid-Maastrichtian. *Paleogeography, Paleoclimatology, Paleoocology*, 139 :177–193.

1.7.2 Conférences avec comité de lecture

- Chauris, H., Karoui, I., Wackernagel, H., Garreau, P., Craneguy, P., and Bertino, L. (2009). Ocean eddy tracking with circlets. In *GeoInformatics for Environmental Sur-*

veillance (Milos, Greece). StaGIS.

- Chauris, H. and Ma, J. (2009). Seismic imaging in the curvelet domain : achievements and perspectives. In *Expanded Abstracts, 71st Annual EAGE Meeting, Workshop 1 (Amsterdam)*. European Association of Geoscientists and Engineers.
- Nguyen, T., Chauris, H., and Lambaré, G. (2009). Time demigration/migration using curvelets. In *Expanded Abstracts, 71st Annual EAGE Meeting (Amsterdam)*. European Association of Geoscientists and Engineers.
- Noble, M., Belayouni, N., and Chauris, H. (2009a). First arrival traveltime tomography for complex near-surface velocity strutures. In *Expanded Abstracts, 71st Annual EAGE Meeting, Workshop 10 (Amsterdam)*. European Association of Geoscientists and Engineers.
- Noble, M., Taillandier, C., and Chauris, H. (2009b). First arrival traveltime tomography for complex near-surface velocity structures. In *Expanded Abstracts, 71st Annual EAGE Meeting (Amsterdam)*. European Association of Geoscientists and Engineers.
- Chauris, H. and Nguyen, T. (2008a). Seismic demigration/migration with curvelets. In *SIAM Conference on Imaging Science (San Diego)*. Society for Industrial and Applied Mathematics.
- Chauris, H. and Nguyen, T. (2008b). Seismic velocity estimation in the curvelet domain. In *Expanded Abstracts, 70th Annual EAGE Meeting (Rome)*. European Association of Geoscientists and Engineers.
- Nguyen, T. and Chauris, H. (2008). Uniform discrete curvelet transform for seismic processing. In *Expanded Abstracts, 70th Annual EAGE Meeting (Rome)*. European Association of Geoscientists and Engineers.
- Taillandier, C., Noble, M., Calandra, H., and Chauris, H. (2008). 3-D refraction traveltime tomography algorithm based on adjoint state techniques. In *Expanded Abstracts, 70th Annual EAGE Meeting (Rome)*. European Association of Geoscientists and Engineers.
- Wackernagel, H., Chauris, H., Bertino, L., Garreau, P., and Heyraud, C. (2008). Characterizing discrepancies in SST representation by geostatistical means. In *General Assembly EGU Meeting (Vienna)*. European Geosciences Union.
- Chauris, H. and Nguyen, T. (2007). Towards interactive seismic imaging with curvelets. In *Expanded Abstracts, 69th Annual EAGE Meeting (London), Workshop WO8 "Curvelets, contourlets, seislets, ... in seismic data processing – where are we and where are we going ?"* European Association of Geoscientists and Engineers.
- Nguyen, T. and Chauris, H. (2007). Seismic migration using the complex shiftable pyramid. In *Expanded Abstracts, 69th Annual EAGE Meeting (London), Workshop WO8 "Curvelets, contourlets, seislets, ... in seismic data processing – where are we and where are we going ?"* European Association of Geoscientists and Engineers.
- Taillandier, C., Noble, M., Chauris, H., Calandra, H., and Guilbot, J. (2007a). Refraction travel-time tomography based on adjoint state techniques. In *Expanded Abstracts, 69th Annual EAGE Meeting (London)*. European Association of Geoscientists and Engineers.
- Taillandier, C., Noble, M., Chauris, H., Podvin, P., and Calandra, H. (2007b). Refraction

travel-time tomography based on adjoint state techniques. In *SIAM Conference on Mathematical and Computational Issues in the Geosciences (Santa Fe)*. Society for Industrial and Applied Mathematics.

- Chauris, H. (2006). Seimic imaging in the curvelet domain and its implications for the curvelet design. In *Expanded Abstracts, 76th Annual SEG Meeting*, pages 2406–2410. Society of Exploration Geophysicists.
- Chauris, H. and Lambaré, G. (2005). Seismic velocity analysis : in time or depth domain ? In *IMA annual workshop, Imaging from wave propagation (University of Minnesota)*.
- ten Kroode, A. P. E. and Chauris, H. (2004). On the use of multi-valued migration for sub-surface imaging in complex media. In *Advanced Problems in Mechanics, XXXII International Summer School Conference (St. Petersburg)*. Society of Exploration Geophysicists.
- Bykov, K., Chauris, H., Kashtan, B., and Troyan, E. (2004). Combining refracted and diving waves for seismic anisotropic model building. In *Geomodel'04 (Gelengik, Russia)*.
- Chauris, H. and Salomons, B. (2004). Data-driven Kirchhoff migration or how to take advantage of the slant stack domain. In *Expanded Abstracts, 66th Annual EAGE Meeting (Paris)*. European Association of Geoscientists and Engineers.
- Le Bégat, S., Noble, M., Chauris, H., and Nguyen, S. (2002). 2-D velocity estimation based on locally coherent seismic data. In *Sub-basalt imaging workshop, 9-11 April 2002 (Cambridge)*.
- Chauris, H., Noble, M., Xu, S., and Lambaré, G. (2001). Migration velocity analysis in complex media : application to the Marmousi data set. In *Expanded Abstracts, 63rd Annual EAGE Meeting (Amsterdam)*, pages M–33. European Association of Geoscientists and Engineers.
- Chauris, H. and Noble, M. (2000a). Automatic velocity estimation from locally coherent events picked in the depth migrated domain : application on a 2-D real data set. In *Expanded Abstracts, 69th Annual SEG Meeting (Calgary)*, pages 922–925. Society of Exploration Geophysicists.
- Chauris, H. and Noble, M. (2000b). Tomographic migration velocity analysis : application on a 2-D marine data set. In *Expanded Abstracts, 62nd Annual EAGE Meeting (Glasgow)*. European Association of Geoscientists and Engineers.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (1999). Migration based velocity analysis in 2-D laterally heterogeneous media. In *Expanded Abstracts, 68th Annual SEG Meeting (Houston)*. Society of Exploration Geophysicists.
- Chauris, H. and Noble, M. (1999). DSO for 2-D velocity estimation through local optimization process. In *Expanded Abstracts, SIAM Conference on Mathematical and Computational Issues in the Geosciences (San Antonio)*, pages 102–102. Society for Industrial and Applied Mathematics.
- Chauris, H., Noble, M., and Podvin, P. (1998). Testing the behaviour of Differential Semblance for velocity optimization. In *Expanded Abstracts, 60th Annual EAGE Meeting (Leipzig)*, pages 1–02. European Association of Geoscientists and Engineers.

- Chauris, H. and Noble, M. (1998a). Differential Semblance Optimization for 2-D velocity field estimation. In *EAGE/SEG Summer Research Workshop Depth Imaging of Reservoir Attributes (Boussens, France).*
- Xu, S., Chauris, H., Lambaré, G., and Noble, M. (1998b). Common angle image gather : a strategy for imaging complex media. In *Expanded Abstracts, EAGE/SEG Summer Research Workshop Depth Imaging of Reservoir Attributes (Boussens, France).*
- Chauris, H. and Noble, M. (1998b). Testing the behavior of Differential Semblance for velocity estimation. In *Expanded Abstracts, 68^{th} Annual SEG Meeting (New Orleans)*, pages 1305–1308. Society of Exploration Geophysicists.
- Xu, S., Chauris, H., Lambaré, G., and Noble, M. (1998a). Common angle image gather : a strategy for imaging complex media. In *Expanded Abstracts, 68^{th} Annual SEG Meeting (New Orleans)*, pages 1538–1541. Society of Exploration Geophysicists.
- Le Rousseau, J., Chauris, H., Montanari, A., and Beaudoin, B. (1996). Modifications environnementales à la limite K/T dans le bassin profond des Apennins du NE (Italie). In *Cretaceous-Tertiary boundary : biological and geological aspects.* Congrès de la Société Géologique de France (Paris).

1.7.3 Autres

- Chauris, H.,(2000). *Analyse de vitesse par migration pour l'imagerie des structures complexes en sismique réflexion.* PhD thesis, Ecole des Mines de Paris, 1–225.
- Chauris, H., (1997). Détermination du modèle de vitesse sismique par l'analyse de la focalisation et de ses fluctuations locales. Master thesis (DEA), Ecole des Mines de Paris, 1–48.
- Chauris, H., and Le Rousseau, J., (1996). La disparition des Inocérames au Maastrichtien et les variations environnementales autour de la limite Crétacé-Tertiaire en Italie centrale, Mémoire Ecole des Mines de Paris, 210 p.

Chapitre 2

Introduction : le cadre scientifique

Le contexte scientifique concerne essentiellement le traitement des données sismiques dans le but d'obtenir de meilleures images du sous-sol, c'est-à-dire de mieux caractériser les propriétés de la sub-surface, comme la valeur des contrastes d'impédance ou leur localisation. La plupart des développements et applications concernent de près le domaine pétrolier, de par l'historique du laboratoire de Géophysique de Mines ParisTech, mais aussi en lien avec mon poste d'ingénieur de recherche au centre d'imagerie sismique de Shell E&P à Rijswijk (Pays-Bas) pendant 5 années. Le traitement des données sismiques pétrolières peut être décomposé en deux étapes principales qui suivent la phase d'acquisition.

2.1 Phases de traitement des données sismiques

La phase dite de "pré-traitement" a pour objectif de préparer les données sismiques pour la partie "imagerie" et dépend des traitements appliqués par la suite. Par exemple, la plupart des algorithmes de migration font l'hypothèse que les données correspondent uniquement à des réflexions et des diffractions primaires. Il est donc important dans ce cas d'enlever les réflexions et diffractions multiples et autres bruits cohérents.

La phase d'imagerie au sens large cherche à reconstruire les propriétés du sous-sol à partir des enregistrements sismiques, en général à la surface. Dans l'approche linéarisée [Tarantola, 1984, Beylkin and Burridge, 1990, Jin et al., 1992, Lambaré et al., 1992], le modèle de vitesse recherché est décomposé en deux parties : (1) une partie contenant les grandes longueurs d'onde du modèle de vitesse, qui explique la cinématique de propagation des ondes, et (2) le résidu qui contient les courtes longueurs d'onde du modèle de vitesse et qui correspond à la réflectivité du milieu ou aux contrastes d'impédance (Figure 2.1). L'inversion des formes d'onde, e.g. [Tarantola, 1986, Gauthier et al., 1986, Pica et al., 1990, Pratt et al., 1996], est une méthode beaucoup plus générale pour retrouver les propriétés du sous-sol. Elle ne décompose pas a priori le modèle de vitesse en deux parties. Par contre, son application reste difficile encore aujourd'hui, principalement en raison de la très forte non-linéarité du problème [Symes, 2008]. Dans le chapitre des perspectives, j'indique un certain nombre de thèmes de recherche sur ce sujet.

L'imagerie des données sismiques permet d'obtenir des images qu'il reste à interpréter. Si je n'ai pas d'expérience particulière d'interprétation, il me semble important de mentionner un travail de recherche en sédimentologie (travail de fin d'étude à l'Ecole des Mines de Paris, cycle

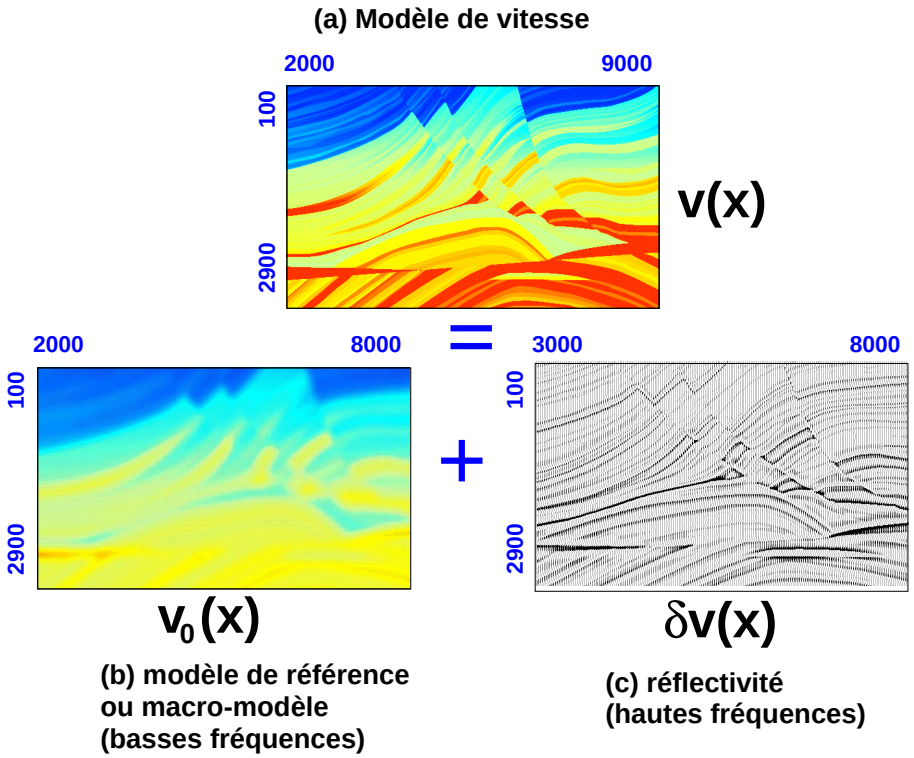


Fig. 2.1 – *Décomposition d'un modèle de vitesse (a) en un modèle lisse (b), typiquement obtenu par tomographie des temps de trajet, et en une partie haute fréquence (c) qui est le résultat de la migration.*

ingénieur) [Chauris et al., 1998, Le Rousseau et al., 1996]. Les conclusions étaient nouvelles à l'époque : il a pu être montré que certaines espèces au Crétacé Supérieur disparaissaient de manière diachrone sur plusieurs millions d'années, probablement en raison de modifications environnementales. Ce travail a remis en cause l'utilisation de l'échelle bio-stratigraphique qui fournit une échelle géologique fondée sur l'apparition et la disparition des espèces animales. Par la suite, ce type de conclusions a été confirmé à des échelles plus restreintes [Desmares et al., 2007].

2.2 Estimation du modèle de vitesse

La détermination du modèle de vitesse basse fréquence pour la migration reste un problème difficile et crucial [Jannane et al., 1989]. Par exemple, le même algorithme d'imagerie appliqué aux mêmes données mais pour deux modèles de vitesse différents donnent deux résultats dont l'un est fortement dégradé (Figure 2.2). Non seulement les profondeurs de focalisation sont différentes, mais surtout la focalisation est affectée. L'origine de ces différences est directement liée à des sommations destructives pour un mauvais modèle de vitesse.

Classiquement, il existe deux méthodes d'estimation de modèle de vitesse : la tomographie des temps de trajet et l'analyse de vitesse par migration. La tomographie des temps de trajet [Bishop et al., 1985] consiste à minimiser l'écart entre (1) les données observées pointées sur des sismogrammes le long d'horizons, et (2) les données calculées par tracé de rais

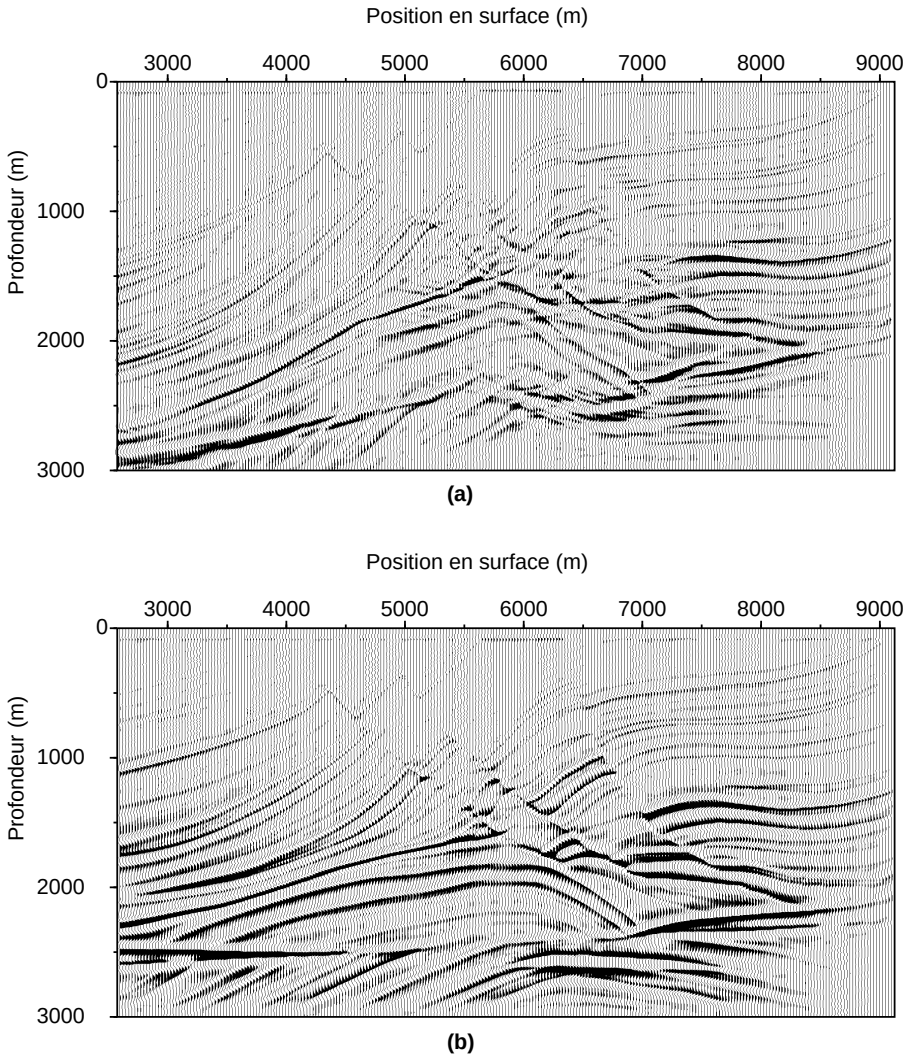


Fig. 2.2 – *Images migrées obtenues avec deux modèles de vitesse différents (migration de type Kirchhoff) : gradient de vitesse constant (a) et milieu exact (b).*

[Farra and Madariaga, 1987]. Les difficultés principales concernent la phase de pointé liée à une phase d'interprétation des données [Lailly and Sinoquet, 1996]. Dans le cadre de mes recherche, j'ai co-encadré la thèse de Cédric Taillandier (e.g. [Taillandier et al., 2009]) sur la tomographie des temps de première arrivée pour la caractérisation de la proche sub-surface. Mon apport principal aura été de promouvoir la technique de l'état adjoint pour la résolution du problème inverse [Plessix, 2006].

Il est également possible de faire l'analyse de vitesse sur des données migrées (Figure 2.3). Les bases de l'analyse MVA ont été données par [Al-Yahya, 1989, Deregowski, 1990]. Le principe est simple : pour des données 2D, fonction du temps d'enregistrement t et de la position des sources et récepteurs s et r, chaque source est migrée indépendamment pour obtenir un cube fonction de la profondeur z, de la position en surface x et de la position de la source. Si le modèle de vitesse utilisé pour la migration explique bien la cinématique de propagation des ondes, alors les panneaux à x donné (CIGs, Common Image Gathers) doivent montrer des

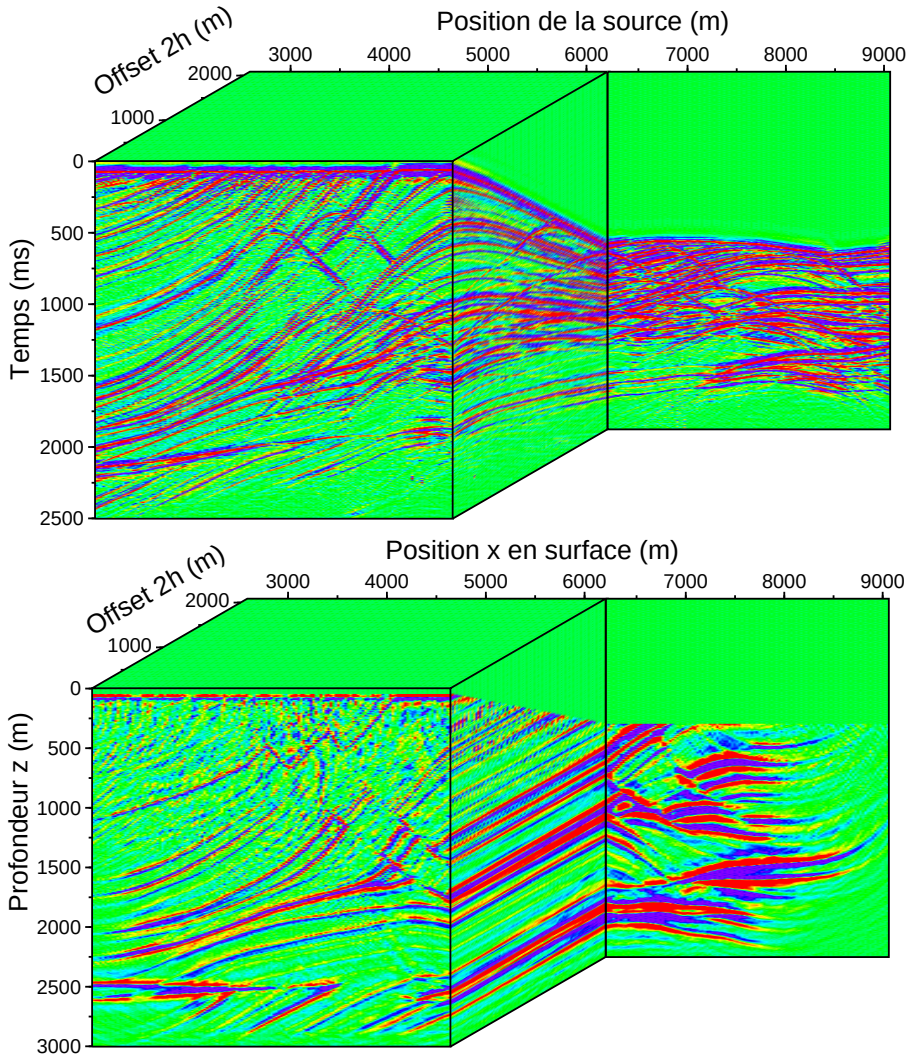


Fig. 2.3 – *Jeu de données 2D, classé en fonction de l'offset (haut), et même jeu de données après migration. Un CIG est défini pour une position x donnée dans le cube migré (bas).*

événements horizontaux, puisque les structures imagées ne dépendent pas de l'acquisition, au moins pour les parties correctement illuminées. Les CIGs sont construits à partir d'un ensemble de sections migrées, soit par point de tir, doit à offset commun, soit pour des mêmes angles d'illumination [Xu et al., 2001], ... Ils sont définis pour une position x donnée à la surface (Figure 2.3). Différents critères existent pour estimer la qualité du modèle de vitesse. Je distingue ici les méthodes selon qu'elles introduisent ou non une phase de pointé.

2.2.1 Méthodes sans pointé

La méthode la plus communément utilisée dans les années 1990–2000 était celle de la semblance [Al-Yahya, 1989, Jin and Madariaga, 1993, Jin and Madariaga, 1994, Docherty et al., 1997, Jin and Beydoun, 2000, Soubaras and Gratacos, 2007]. Il s'agit de maximiser le "stack" des données. Malheureusement, ce type de critère ne permet pas une approche de minimisation locale

par gradient puisque la fonction objective montre des minima secondaires.

A ma connaissance, seules deux approches sont susceptibles de pouvoir utiliser des algorithmes de type gradient pour résoudre le problème d'optimisation. Il s'agit des méthodes MBTT et DSO. Le principe de Migration Based Travel Time [Clément, 1994, Plessix et al., 1995, Plessix, 1996] compare la réflectivité en temps entre les données observées (sismogrammes) et les données calculées *via* une opération de migration/démigration. L'opération de migration et de démigration modifie légèrement les temps de trajet à cause de la sommation sur les offsets lorsque le modèle de vitesse n'est pas parfait. La qualité du modèle de vitesse est estimée essentiellement avec les amplitudes. Depuis, une idée très similaire a été reprise par [van Leeuwen and Mulder, 2007].

La méthode DSO (Differential Semblance Optimization, ou optimisation de la semblance différentielle) a été introduite par [Symes and Carazzone, 1991]. Elle propose de mesurer la qualité d'un modèle de vitesse par une mesure locale. La dérivée horizontale est appliquée aux traces dans les panneaux CIGs. Elle est faible pour des événements horizontaux. En 2000, l'intérêt de cette méthode était démontrée sur un ensemble d'inversions 1D avec quelques tests préliminaires 2D [Symes, 1993, Symes and Kern, 1994, Gockenback and Symes, 1995] et sur le plan théorique [Symes, 1998a] : la DSO est une méthode qui permet des optimisations locales au moins dans les cas 1D. Dans le chapitre 4, j'exposerai ma contribution dans ce domaine ainsi que des travaux plus récents.

2.2.2 Méthodes avec pointés

Beaucoup d'extensions ont été proposées depuis les travaux de [Al-Yahya, 1989], entre autres par [Lafond and Levander, 1993, Eckhardt, 1994, Wang et al., 1995, Liu and Bleistein, 1995, Audebert et al., 1997], et la liste n'est pas exhaustive. Tous font au moins une hypothèse pour remettre à jour le modèle de vitesse, qui peut être : milieu de vitesse latéralement homogène, inversion valide aux faibles offsets, réflecteurs horizontaux... Dans la plupart des cas, le pointé est introduit pour sélectionner un maximum d'amplitude le long de courbes prédéfinies. [Liu, 1997] a proposé une approche valide pour n'importe quel modèle de vitesse 2D. Il a établi la relation entre la perturbation de la profondeur d'un réflecteur et la perturbation d'un paramètre du modèle de vitesse. Des extensions 3D ont même été proposées [Meng et al., 1999b, Meng et al., 1999a].

Il apparaît que quelle que soit l'approche utilisée (avec ou sans pointé, mesure locale ou globale), la question principale est de savoir comment définir la fonction objective à minimiser. La mesure est d'autant plus intéressante qu'elle permet de remettre à jour et inverser efficacement un modèle de vitesse donné. Cette démarche d'analyse des données dépasse la thématique de l'estimation du modèle de vitesse et est au cœur du sujet de ce mémoire.

2.3 Nouvelle perspective : la cohérence locale des données sismiques

Une grande part de mes activités de recherche, depuis la thèse de doctorat, tente de porter un regard différent sur les données sismiques : au lieu de considérer les traces indépendamment les unes des autres, j'ai essayé de voir ce qu'un groupe de plusieurs traces successives pouvait

apporter aux différentes étapes du traitement sismique. La première justification est visuelle : l'œil qui considère une seule trace regarde la forme d'onde. Dans le cas d'un ensemble de traces, c'est d'abord la cohérence des enregistrements qui est perçue (Figure 2.4).

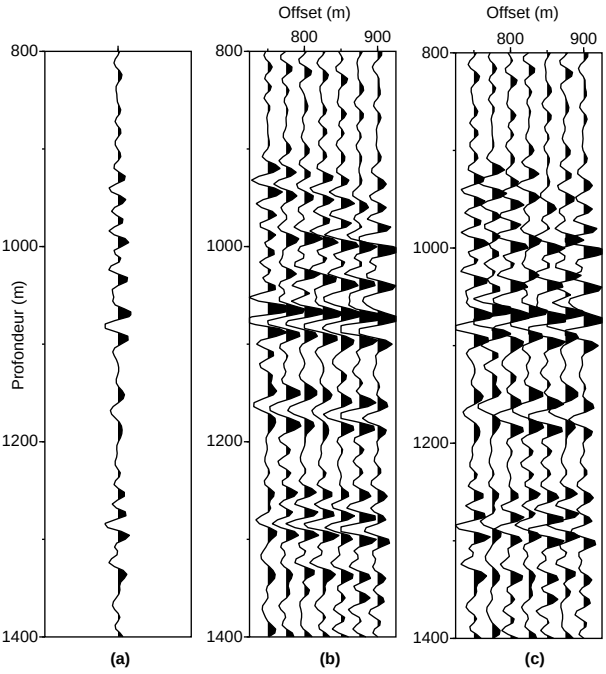


Fig. 2.4 – *Trois images extraites d'un CIG. La trace de gauche (a) correspond à la trace centrale de (b). L'œil s'intéresse à la forme du signal dans le premier cas (a), tandis qu'il est attiré par les zones qui montrent une cohérence latérale sur quelques traces (b). Lorsque les mêmes traces sont classées dans un ordre aléatoire (c), le sismogramme semble moins porteur d'information, car le signal perd sa cohérence latérale.*

Pour un modèle de vitesse donné, il n'est pas toujours facile de suivre des événements continus dans des panneaux CIGs, pourtant à la base de l'analyse de vitesse par migration (Figure 2.5). Il en est en fait de même pour toute section sismique, en particulier sur données réelles (Figure 2.6). Par contre, il est a priori beaucoup plus facile d'identifier un certain nombre d'événements sismiques avec une cohérence latérale sur au moins quelques traces. La définition de cette cohérence et son exploitation fait tout l'objet de ce mémoire.

Pour exploiter la cohérence locale des données sismiques, j'ai développé depuis la période de la thèse de doctorat deux approches :

- La première (chapitre 3) donne une vision minimaliste avec une approche purement cinématique.
- La seconde (chapitre 4) prend en compte la forme d'onde des données, toujours dans le cadre asymptotique hautes fréquences. Je me suis attaché à comprendre en quoi les curvelets [Candès and Donoho, 2004] pouvaient être intéressantes pour le traitement des données sismiques, en particulier autour des problèmes d'imagerie.

Les perspectives (chapitre 5) concernent à la fois des extensions directes des travaux menés, mais aussi de nouvelles pistes, en particulier pour s'affranchir du cadre asymptotique. Avant

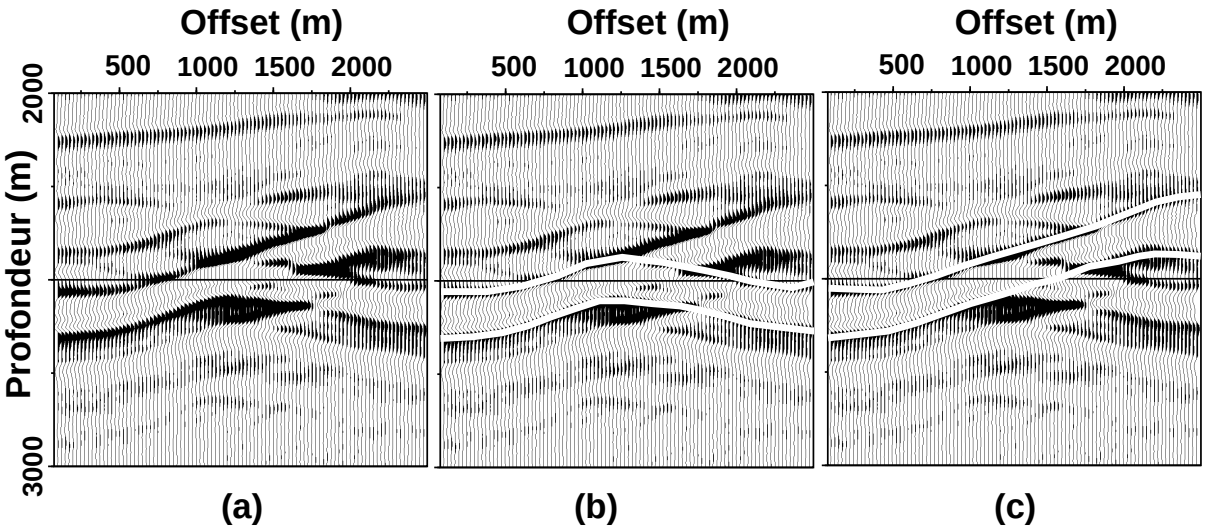


Fig. 2.5 – *Même dans le cas de données synthétiques, il n'est pas toujours facile de suivre continûment un événement le long de tous les offsets dans un CIG (a). Deux interprétations sont proposées, (b) et (c), soulignées par un trait continu blanc. Cet effet apparaît dès que le modèle de vitesse choisi pour la migration n'est pas correct.*

d'exposer les principaux résultats de mes recherches, il me semble important, dans ce chapitre sur le cadre scientifique, de donner quelques mots sur l'évolution récente de l'équipe de Géophysique de Mines ParisTech.

2.4 Direction de l'équipe de Géophysique

Suite au départ de Pascal Podvin pour la direction des études de Mines ParisTech en Septembre 2008, j'ai pris la direction de l'équipe de Géophysique du centre de Géosciences de Mines ParisTech. Il a été convenu que ce serait pour un mandat de deux ans. Mark Noble, présent depuis presque la création du laboratoire, a beaucoup contribué à donner un second souffle à l'équipe. Pour y parvenir, nous avons procédé en deux étapes :

- Définition de la politique scientifique de l'équipe, autour du thème "Imagerie de la proche sub-surface". Quatre thématiques ont été dégagées : (1) Tomographie de la proche subsurface, avec par exemple l'inversion des temps de trajet de première arrivée (thèse de doctorat de Cédric Taillandier [Taillandier et al., 2009]). (2) Sismique passive, avec la micro-sismicité et comme objectif la surveillance des réservoirs, des stockages souterrains de gaz, ..., ou encore l'interférométrie. (3) L'imagerie des structures complexes, avec le développement de nouvelles approches autour de l'inversion des formes d'onde (voir chapitre 5). Et enfin (4), une thématique transverse autour de l'optimisation de le Calcul Haute Performance pour la gestion des grands jeux de données. L'objectif reste les applications sur données réelles à différentes échelles (géotechnique, pétrolière, sismologie).
- Nouveaux recrutements, en particulier de deux jeunes nouveaux permanents scientifiques sur des postes d'assistants de recherche, Alexandrine Gesret en Janvier 2009 et Daniela Donno en Décembre 2009. Fin 2009, l'équipe était donc composée de quatre permanents

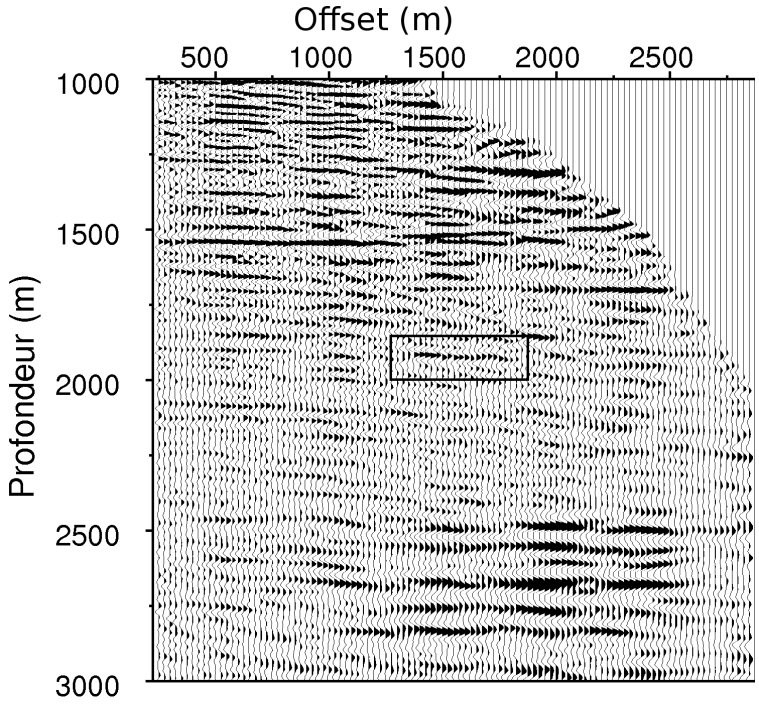


Fig. 2.6 – *CIG obtenu sur des données réelles. Il est plus facile de pointer beaucoup d'événements localement cohérents, par exemple autour de la fenêtre, que de suivre des événements continus depuis les courts jusqu'aux grands offsets.*

scientifiques, un post-doc, cinq étudiants en thèse et une secrétaire.

Les priorités ont d'abord été données aux publications et aux relations avec l'extérieur, tant sur le plan académique que sur le plan industriel. En 2008–2009, l'équipe de Géophysique a eu des projets avec Paris VII, l'IRSN, l'Ifremer et Actimar, l'université de Tsinghua en Chine, CGGVeritas, Total, Shell et Intel.

Au début des chapitres qui suivent, j'indique en quelques lignes mes contributions.

Chapitre 3

Invariants cinématiques

L'objectif est d'exploiter la cohérence locale des données sismiques, avec ici une vision très épurée, puisque, par exemple, dans une section 2D, les événements sont caractérisés par deux positions (temps ou profondeur, et position en surface) et une pente [Billette and Lambaré, 1998]. Par extension, je parle d'*événements localement cohérents*. Dans cette partie, la signature de ces événements n'est pas prise en compte.

Mes contributions autour de l'utilisation de la cohérence locale à partir de pointés concernent :

- Formulation du problème d'estimation du modèle de vitesse dans le domaine migré, avec en particulier un calcul efficace du gradient de la fonction objective [Chauris et al., 2002a] ;
- Application sur des données réelles 2D [Chauris et al., 2002b] ;
- Démonstration de l'équivalence entre l'estimation du modèle de vitesse à partir de pointés sur des images migrées et la tomographie de pente [Chauris et al., 2002a].

3.1 Analyse de vitesse profondeur avec pointés

Le principe de l'analyse de vitesse a été décrit dans la partie 2.2. Les panneaux CIGs doivent présenter des événements horizontaux si la cinématique de propagation des ondes est correctement expliquée. Les données 2D migrées avant sommation (migration pre-stack) dépendent de la profondeur z, de la position en surface x et la position de la source s pour des migrations à source commune ou du demi-offset $h = (s - r)/2$ pour des migrations à offset commun (Figure 2.3). Un événement localement cohérent est caractérisé par trois positions (z, x, s) et deux angles (ξ, φ) définis par

$$\tan\varphi = \frac{\partial z}{\partial s}, \tag{3.1}$$

$$\tan\xi = \frac{\partial z}{\partial x}. \tag{3.2}$$

Les deux pentes $\tan\xi$ et $\tan\varphi$ sont prises le long de la cohérence des données sismiques. Pour le bon modèle de vitesse, l'angle ξ correspond à l'angle de pendage observé dans les données et φ doit être nul. Il est donc naturel de définir comme fonction objective [Chauris et al., 2002a]

$$J[u] = \frac{1}{2} \sum_{\text{pointés}} (w \tan\varphi)^2, \tag{3.3}$$

où u est le modèle de lenteur et w reste une pondération à définir. La plus grande difficulté est de calculer efficacement le gradient de la fonction objective par rapport au modèle de vitesse v ou de lenteur u. Dans des travaux précédents, [Liu, 1997] avait établi le lien entre la profondeur d'un réflecteur et le modèle de vitesse. Ici, c'est la relation entre la pente $\tan\varphi$ dans les CIGs et le modèle de vitesse qui est recherché. Il faut noter que le modèle de vitesse est décrit par un modèle de vitesse a priori lisse sans interface. Dans les données, les événement localement cohérents sont a priori non corrélés.

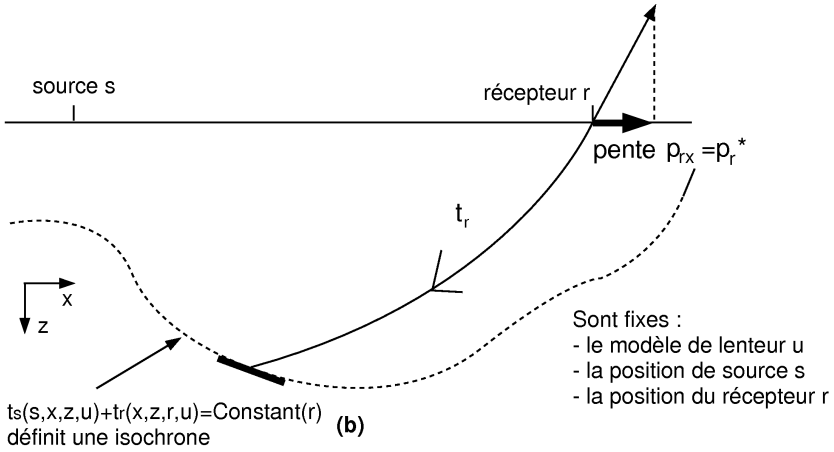


Fig. 3.1 – *Dans le cas d'une migration par point de tir, la position spéculaire, autour de laquelle l'énergie focalise, est donnée par l'intersection de l'isochrone et du rai dont la pente initiale est contrôlée par les données [Chauris et al., 2002a].*

Le temps de trajet calculé par tracé de rais entre la source et le point image (x, z) est noté $\tau(s, x, z)$. Si le temps observé dans les données est appelé $t(s, r)$, alors le principe de migration veut que (Figure 3.1)

$$\tau(s,x,z) + \tau(r,x,z) = t(s,r), \tag{3.4}$$

$$\frac{\partial \tau(s,x,z) + \tau(r,x,z)}{\partial r} = \frac{\partial t(s,r)}{\partial r}. \tag{3.5}$$

La seconde équation s'ecrit aussi

$$p_r = p_r^*, \tag{3.6}$$

où $p_r = \frac{\partial \tau}{\partial r}(r, x, z)$ et $p_r^* = \frac{\partial t}{\partial r}(s, r)$. L'approximation de la phase stationnaire (sommation selon les récepteurs pour des données hautes fréquences) indique que la dérivée des temps par rapport à la position du récepteur, pour une migration à source commune, est un invariant cinématique. La variable p_r correspond à la composante horizontale du rai joignant le point image (x, z) au

récepteur, et p_r^* à la pente pointée dans les données. Pour obtenir une nouvelle expression de $\tan\varphi$, il suffit de différencier l'équation 3.4 par rapport à z, s et r pour x fixé, et utiliser la relation 3.5. Nous obtenons alors

$$2u\cos\theta\cos\xi\cdot\tan\varphi = p_s^* - p_{sx}, \tag{3.7}$$

où θ est le demi-angle d'ouverture entre les rais qui lient le point image au couple source récepteur. La lenteur u est prise au point image. Pour un calcul facilité du gradient de la fonction objective, nous posons $w = 2u\cos\theta\cos\xi$ déterminé dans le modèle courant, et donc

$$\frac{\partial J[u]}{\partial u} = \sum_{\text{pointés}} (p_{sx} - p_s^*)\frac{\partial p_{sx}}{\partial u} = -2\sum_{\text{pointés}} u\cos\theta\cos\xi\tan\varphi\frac{\partial p_{sx}}{\partial u}, \tag{3.8}$$

puisque p_s^* ne dépend pas du modèle u choisi pour la migration. Il s'agit alors de calculer $\frac{\partial p_{sx}}{\partial u}$ pour des conditions particulières où $(s, r, \tau(s,x)+\tau(r,x), x, p_r)$ restent constants. Les détails sont donnés dans [Chauris et al., 2002a]. Le calcul revient à combiner des valeurs issues de tracé de rai paraxial et de dérivées de Fréchet. Depuis, cette méthodologie s'est développée dans l'industrie pétrolière, e.g. [Guillaume et al., 2008].

3.2 Lien avec la tomographie de pentes

Cette analyse permet de faire le lien entre deux méthodes d'analyse de vitesse a priori distinctes puisque formulées dans des domaines différents [Chauris et al., 2002a] : l'analyse de vitesse par migration avec les données pointées et la tomographie de pentes, plus spécifiquement la Stéréotomographie [Billette and Lambaré, 1998].

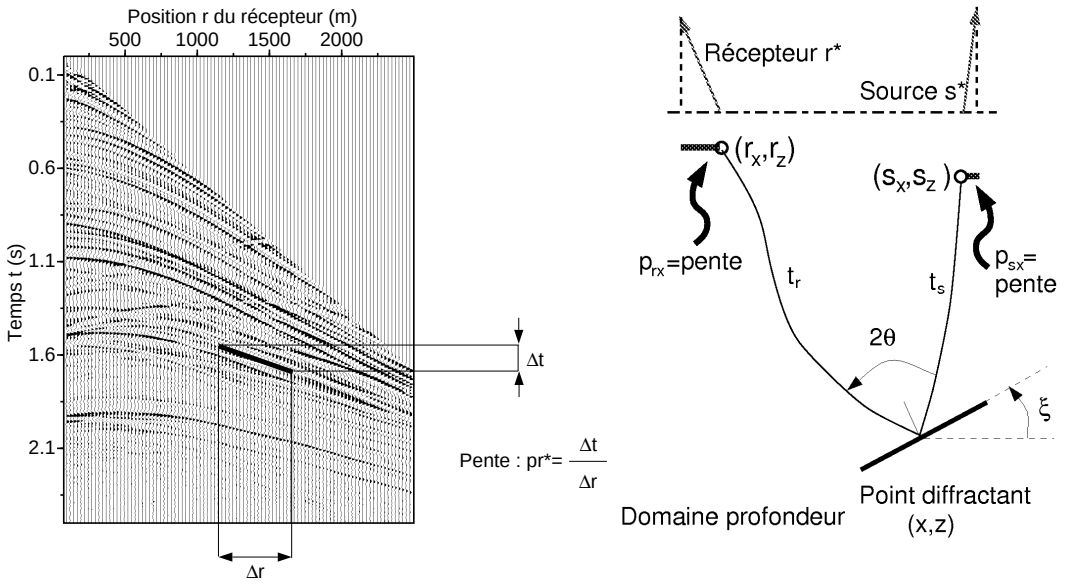


Fig. 3.2 – *Evénement localement cohérent dans le domaine temps (gauche) et représenté avec des rais (droite) [Billette and Lambaré, 1998, Chauris et al., 2002a].*

Dans le cadre de la Stéréotomographie, le modèle à optimiser est décrit par un modèle de vitesse v et la position (x, z, ξ, θ) d'événements localement cohérents, ainsi que deux temps de trajet τ_s et τ_r depuis (x, z) vers la surface. L'angle ξ correspond au pendage et θ au demi angle d'ouverture. Il s'agit de trouver les paramètres optimaux de telle sorte que les rais issus de (x, z) avec les bons angles atteignent les positions $(s^*, r^*, t, p_s^*, p_r^*)$ pointées dans les données observées. La fonction objective est donc

$$J_{\mathrm{St}}[u] = \frac{1}{2}\sum_{\text{pointés}}(s[u]-s^*)^2 + \frac{1}{2}\sum_{\text{pointés}}(r[u]-r^*)^2 + \frac{1}{2}\sum_{\text{pointés}}(\tau_s[u]+\tau_r[u]-t)^2 + \frac{1}{2}\sum_{\text{pointés}}(p_s[u]-p_s^*)^2 + \frac{1}{2}\sum_{\text{pointés}}(p_r[u]-p_r^*)^2. \tag{3.9}$$

La phase de migration dans l'analyse MVA simplifie automatiquement l'expression de la fonction J_{St}. En effet, tous les termes sont nuls, sauf $p_s[u]-p_s^*$ pour une migration à source commune. C'est donc un résultat important qui montre la stricte équivalence de deux méthodes a priori différentes. Il est possible de combiner les deux approches en pratique [Nguyen et al., 2008].

3.3 Les applications

Les applications sont tirées de [Chauris et al., 2002b] où plus de détails peuvent être trouvés. Dans le premier exemple, des données synthétiques ont été générées avec l'approximation de rai+Born, donc sans multiple. Le modèle de référence est une version lisse du modèle de Marmousi [Versteeg and Grau, 1991], tandis que la perturbation préserve la réflectivité originale. A chaque étape d'inversion du modèle de vitesse, en partant d'un modèle de vitesse à gradient constant, les données sont migrées, puis pointées et enfin la fonction objective et son gradient sont évalués. Pour rendre le schéma plus efficace, seules quelques sections à offset constant, tous les 200 m, sont calculées, ainsi que des CIGs tous les 200 m également (Figure 3.3). Il faut noter que puisque l'étape de pointé suit la migration, ce n'est pas exactement la même fonction objective qui est évaluée à chaque itération. Cependant, au fur et à mesure que le modèle de vitesse est amélioré, la focalisation est meilleure, et donc le pointé facilité.

Le résultat de l'inversion produit un modèle de vitesse qui explique la cinématique de propagation des ondes, pour le critère choisi sur les CIGs. Les valeurs de la fonction objective dans les modèles inversés et de référence sont équivalentes (Figure 3.4). Par contre, les profondeurs des réflecteurs varient, car le problème n'est pas bien contraint, principalement en raison de l'acquisition limitée, mais également de la détermination difficile en sismique réflexion des couches superficielles (Figure 3.5). Ce dernier aspect est abordé dans [Taillandier et al., 2009].

Le second exemple choisi est celui d'une application 2D sur données réelles marines dans un contexte tectonique de bassin en extension (Figure 3.6). L'inversion est réalisée en deux étapes [Chauris et al., 2002b] : dans un premier temps, le meilleur modèle 1D est recherché à partir d'un modèle à gradient constant, puis des perturbations 2D. Si ces perturbations restent limitées en valeur absolue, autour de quelques 100 m/s, l'image migrée finale est bien améliorée (Figure 3.7). Il est plus intéressant de faire plusieurs itérations non-linéaires pour éviter de pointer à chaque itération.

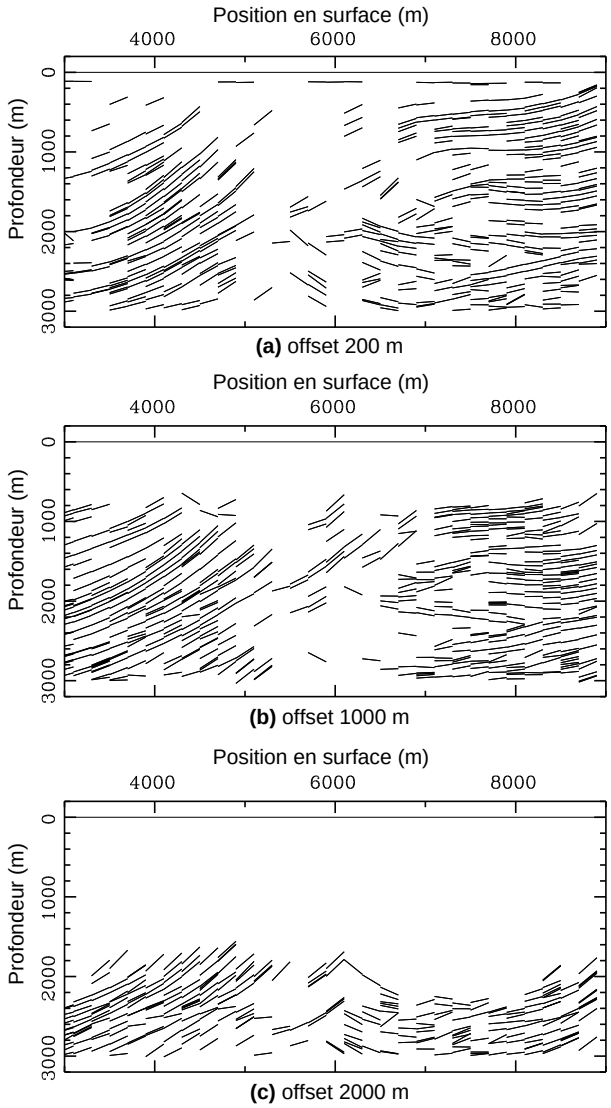


Fig. 3.3 – *Disposition des segments de rais pointés dans trois sections à offset commun (offsets* 200 *m (a),* 1000 *m (b) et* 2000 *m (c)) [Chauris et al., 2002b].*

3.4 Conclusions

J'ai montré qu'il était possible de faire une tomographie de pente dans le domaine migré et qu'elle est équivalente à une tomographie de pente dans le domaine non-migré. L'étape clé est la démigration cinématique dont j'ai indiqué les équations. Dans la pratique, les résultats sont différents [Le Bégat et al., 2004], en raison du pointé qui est appliqué soit sur des données avant migration, soit après migration. La migration, qui peut être vue comme une sommation le long d'isochrone, améliore beaucoup le rapport signal sur bruit. Par ailleurs, les pointés sont faits en pratique autour de traces régulièrement espacées et donc leur distribution dans les domaines temps et profondeur est très différente.

Plus généralement, les événements localement cohérents montrent une vision très épurée des données sismiques. Ils sont cependant utiles pour l'analyse de vitesse (voir par exemple

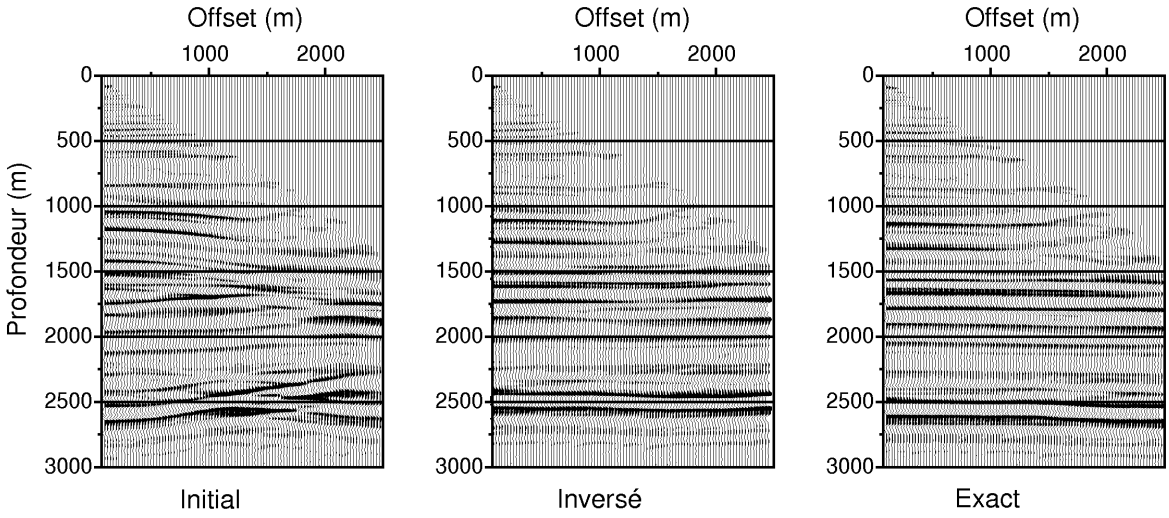


Fig. 3.4 – *CIGs calculés avec les modèles de vitesse initial, inversé et exact (modèle synthétique 2D). Des événements compliqués, comme par exemple à la profondeur* $z = 2500$ *m dans le modèle initial, disparaissent dans les deux autres modèles [Chauris et al., 2002b].*

une synthèse [Lambaré, 2008]), y compris pour des milieux anisotropes [Alerini et al., 2007]. Le chapitre suivant prend en compte la signature des données sismiques.

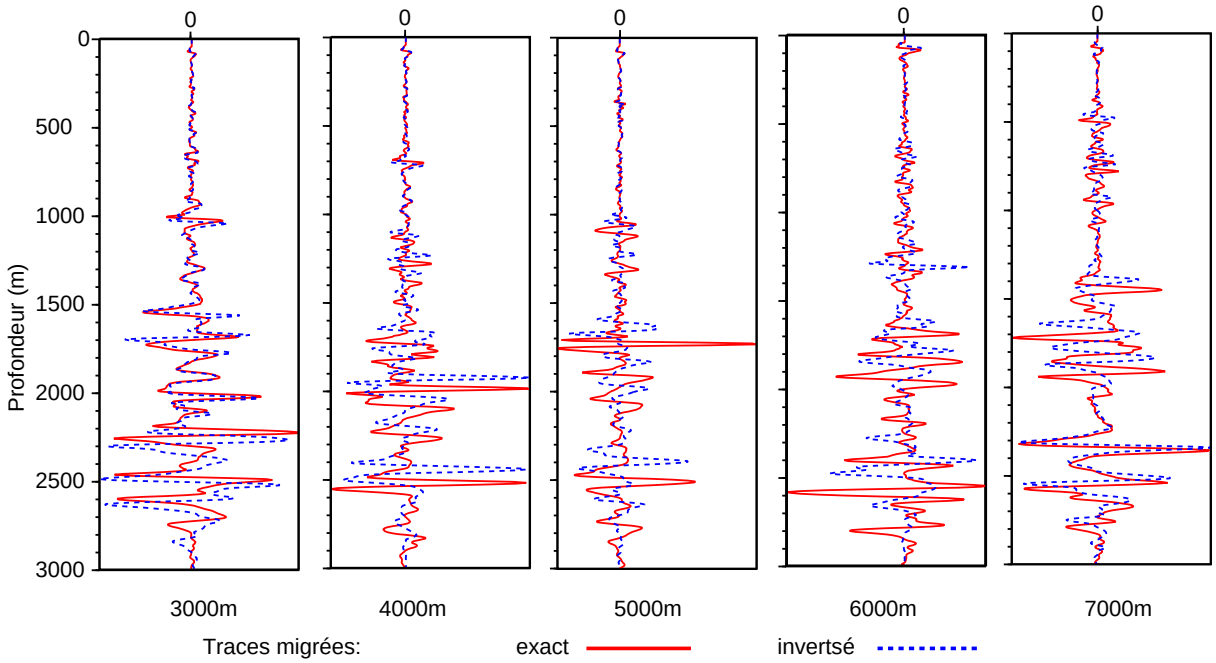


Fig. 3.5 – *Comparaison des profils extraits des images migrées calculées avec le bon modèle de vitesse et avec le modèle de vitesse inversé. Les valeurs de la fonction objective pour les CIGs calculés dans les modèles inversés et exacts sont équivalentes, même si les profondeurs des réflecteurs sont différentes [Chauris et al., 2002b].*

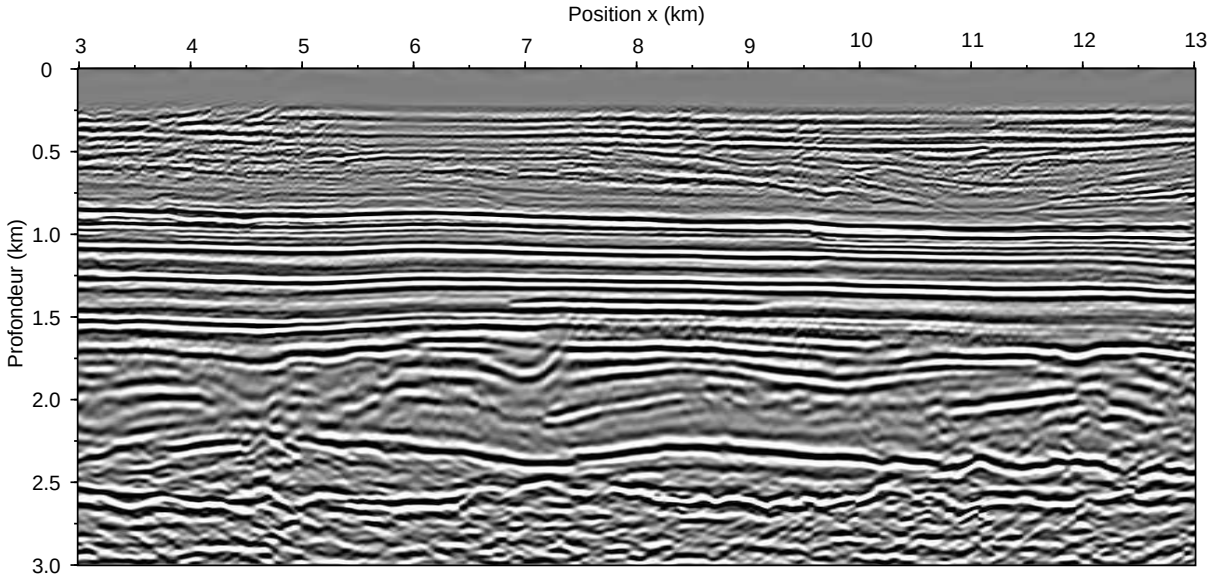


Fig. 3.6 – *Image sur données réelles, migrée avec le modèle de vitesse inversé. Du point de vue géologique, nous pouvons distinguer une couverture sédimentaire stratifiée, une partie avec des blocs basculés délimités par des failles et une zone salifère [Chauris et al., 2002b].*

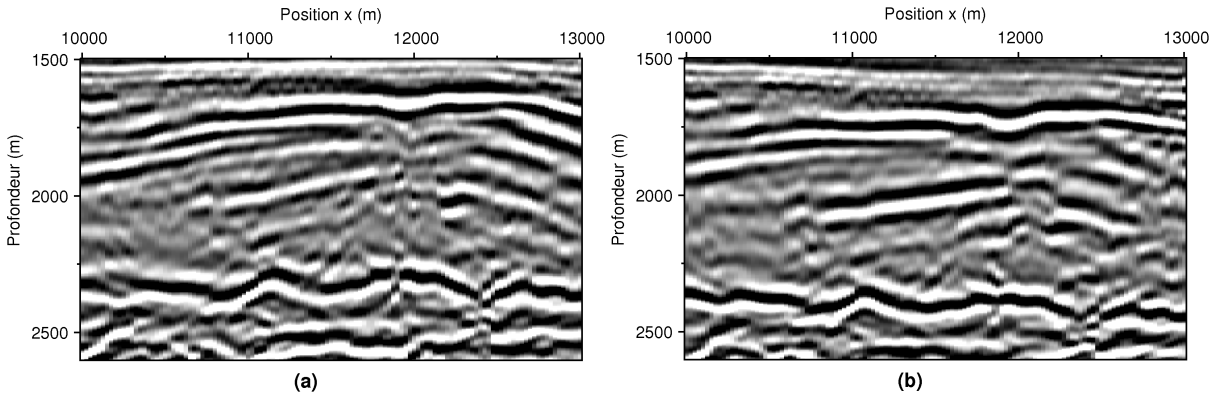


Fig. 3.7 – *Zooms sur les images migrées (Figure 3.6) obtenues (a) avec le modèle inversé 1D, et (b) avec le modèle final 2D. La focalisation de l'image est grandement améliorée, même si le modèle de vitesse ne change pas fondamentalement (autour de quelques 100 m/s) [Chauris et al., 2002b].*

Chapitre 4

Prise en compte de signature des données

La prise en compte de la signature des données sismiques dans la notion d'événements localement cohérents peut se faire de différentes manières. Mes contributions les plus importantes dans ce domaine ont été :

- Migration "slant stack" et introduction de la semblance locale pour l'interprétation des images dans des zones complexes [Chauris and Salomons, 2004]. Typiquement, les algorithmes de Kirchhoff avec arrivées multiples permettent en principe de mieux restituer la position des interfaces en profondeur. En pratique, ils sont aussi source d'artefacts. La semblance locale est un moyen très simple d'épurer les images.
- Application sur données réelles 2D de la semblance différentielle [Chauris and Noble, 2001] (DSO : Differential Semblance Optimization) dont l'idée venait de [Symes and Carazzone, 1991].
- Développement de deux nouveaux codes de décomposition en curvelets [Chauris and Nguyen, 2008a, Nguyen and Chauris, 2010]
- Meilleure compréhension de l'applicabilité des curvelets pour l'imagerie sismique [Chauris and Ma, 20
- Idée de la démigration/migration avec des applications à l'analyse de vitesse [Chauris and Nguyen, 20 Chauris and Nguyen, 2008b] et sur données réelles [Nguyen et al., 2009b].

4.1 Migration par "local slant stack"

De nombreux progrès ont été faits pour améliorer la migration de Kirchhoff, en particulier avec la prise en compte des triplications du champ d'onde et une migration quantitative (voir par exemple [Operto et al., 2000, Albertin et al., 2001, Xu and Lambaré, 2004]). Cependant, les arrivées multiples sont potentiellement source de bruits cohérents additionnels dans les images migrées, dès que par exemple les opérateurs de migration ne sont pas bien apodisés.

Pour y remédier, j'ai proposé de modifier la migration dite "local slant stack migration" [Ottolini and Claerbout, 1984]. Cette migration comporte deux phases : d'abord une transformée $\tau - p$ locale est appliquée aux données enregistrées autour de quelques traces de référence, comme pour la migration par faisceaux Gaussiens [Hill, 2001]. Cette première phase est indépendante du choix du modèle de vitesse utilisé pour la migration. Ensuite, les panneaux $\tau - p$ sont migrés trace par trace en utilisant la pente au récepteur pour une migration par source.

J'ai proposé d'introduire un poids supplémentaire dans la migration qui est la valeur de

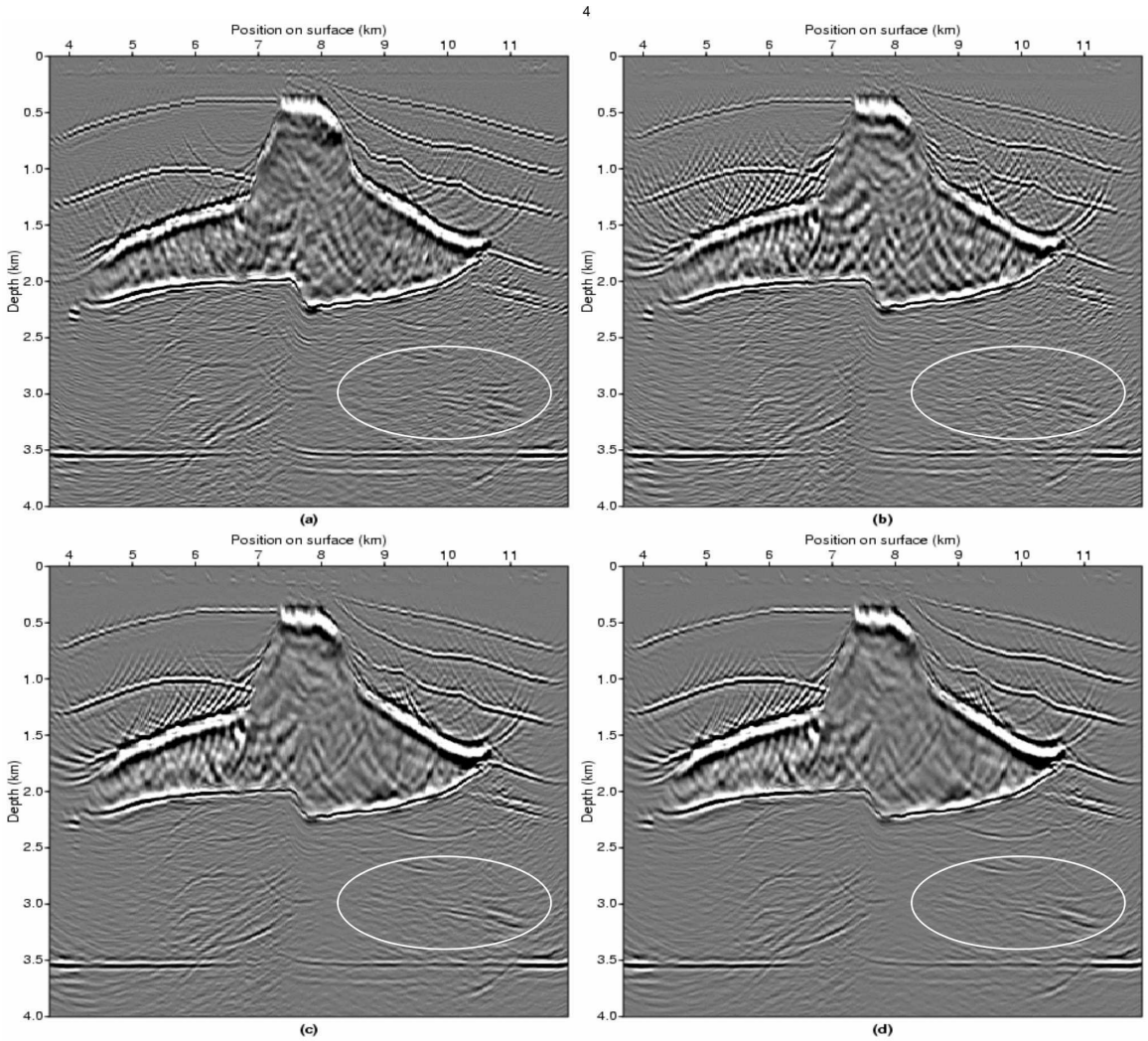


Fig. 4.1 – *Images migrées 3D sur le modèle dôme de sel SEG/EAGE pour y=5.5 km, avec des fonctions de Green multivaluées (a), migration par slant stack (b), migration par slant stack avec l'utilisation de la valeur de la semblance pour clarifier l'image (c et d) [Chauris and Salomons, 2004].*

la semblance, normalisée entre 0 et 1, et obtenue pendant la transformée de Radon locale [Chauris and Salomons, 2004]. Pour les événements spéculaires, la valeur de la semblance est proche de 1, et donc leur image après migration est renforcée. La Figure 4.1 montre des résultats sur les données 3D SEG/EAGE. Sous le dôme de sel, les réflecteurs sont plus visibles. Cependant, dans la partie superficielle, le résultat de migration est moins bon, car le choix des traces de référence pour la transformée $\tau - p$ n'est pas adapté pour cette zone. Dans la partie 4.5, je montrerai une meilleure décomposition des données avec les curvelets [Candès and Donoho, 2004].

La migration "slant stack" locale utilise la cohérence locale des données. Elle a un impact sur la qualité des résultats dans le cas de migration de Kirchhoff avec arrivées multiples, mais aussi pour prendre en compte l'aliasing spatial [Baina et al., 2003, Nguyen et al., 2009a].

4.2 Méthode d'analyse de vitesse automatique (DSO)

La méthode d'analyse de vitesse présentée au chapitre 3 a son équivalent sans pointé, d'abord proposé par [Symes and Carazzone, 1991, Symes, 1993]. Il s'agit de la "Differential Semblance Optimization" (DSO). Cette approche s'applique sur l'ensemble des traces dans le cadre de données réfléchies. Elle exploite la cohérence locale des données : chaque section à offset constant est migrée indépendamment dans un modèle de vitesse choisi. Ensuite, les résultats pour des offsets successifs sont comparés un à un, par le biais de la dérivée horizontale :

$$J_{\mathrm{DSO}} = \frac{1}{2} \iiint \mathrm{d}x \, \mathrm{d}z \, \mathrm{d}h \left| \frac{\partial M}{\partial h} \right|^2 , \tag{4.1}$$

où $M(x, z, h)$ est l'image migrée et h le demi-offset en surface [Chauris and Noble, 2001]. La fonction objective classique dans les années 2000 était la semblance

$$J_{\mathrm{SEM}} = \frac{1}{2} \iint \mathrm{d}z \, \mathrm{d}h \left| \int \mathrm{d}x M \right|^2 . \tag{4.2}$$

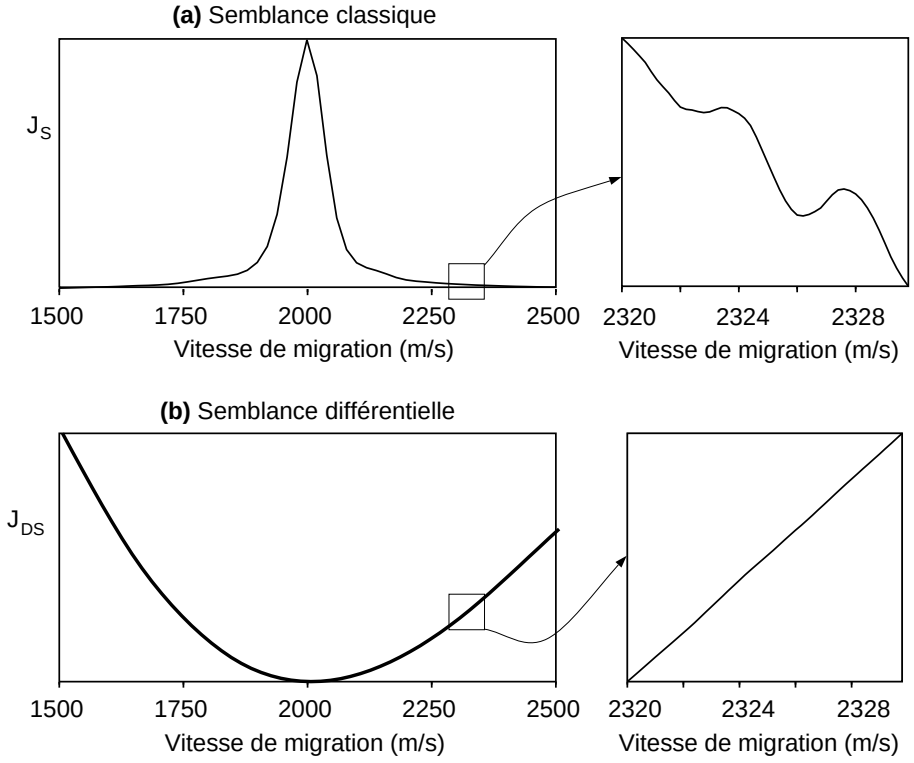


Fig. 4.2 – *Topologie des fonctions objectives semblance classique et semblance différentielle [Chauris and Noble, 2001].*

L'attrait principale de la semblance différentielle est qu'elle permet a priori d'utiliser une approche locale de type gradient pour minimiser la fonctionnelle et obtenir le meilleur modèle de vitesse [Symes, 1998a, Symes, 2008]. En revanche, la semblance classique montre un certain nombre de maxima secondaires (Figure 4.2). La théorie de la DSO a été développée

par [Symes and Kern, 1994, Symes, 1998b, Plessix et al., 2000, Mulder and ten Kroode, 2002, Symes, 2008]. En particulier, ils ont montré comment calculer le gradient de la fonction objective par la méthode de l'état adjoint [Plessix, 2006]. Mon apport principal a été d'appliquer la DSO sur des jeux de données réelles en 2D [Chauris and Noble, 2001].

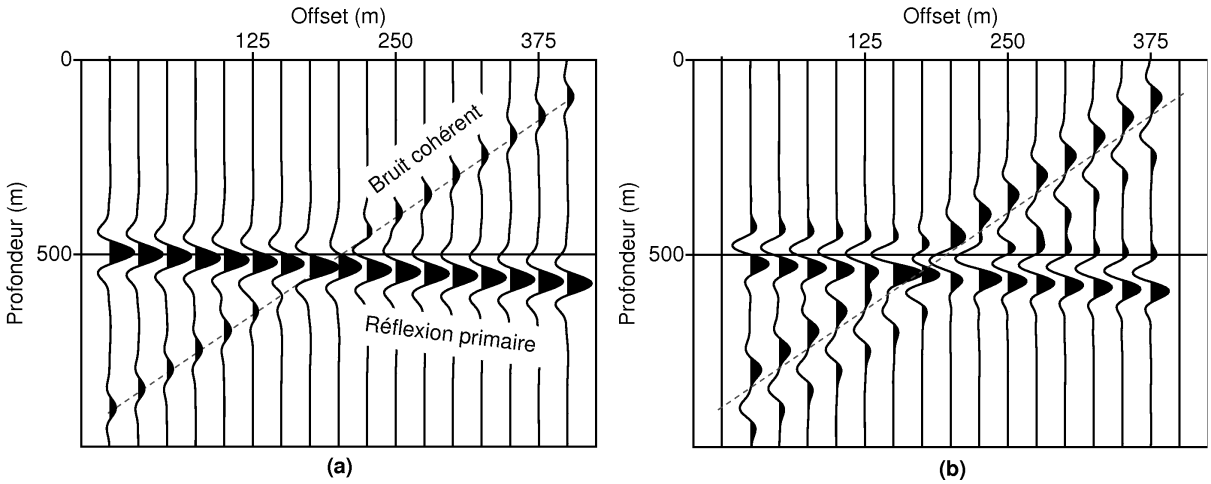


Fig. 4.3 – *Limite principale de la semblance différentielle. Le CIG (a) comprend un événement majeur correspondant à une réflexion primaire, et du bruit cohérent de plus faible amplitude. Le panneau (b) est le résultat du calcul de la dérivée horizontale de (a). La semblance différentielle donne un poids équivalent aux deux événements.*

Il s'avère que l'approche DSO est très sensible aux bruits cohérents (Figure 4.3). Dans la pratique, un filtre doit être appliqué dans les CIGs pour qu'après application de la dérivée horizontale, l'énergie résiduelle donne bien une information sur le modèle de vitesse utilisé pour la migration (Figure 4.4).

Le résultat présenté ici est celui de données marines de la mer du Nord, dans un régime tectonique en extension, autour du champ pétrolier Oseberg [Hagen and Kvalheim, 1992, Johnstad et al., 1995]. Les détails sont indiqués dans [Chauris and Noble, 2001]. Après inversion par une méthode de gradient conjugué, les CIGs montrent des événements horizontaux (Figure 4.5). Une partie de l'image (Figure 4.6) dans la zone centrale montre des réflecteurs plus continus pour l'approche DSO que pour l'analyse classique NMO/DMO ou pour la Stéréotomographie (voir chapitre 3).

En conclusion, la DSO a permis sur des données réelles 2D d'obtenir un modèle de vitesse basse fréquence qui explique la cinématique de propagation des ondes. Cette approche reste sensible aux bruits cohérents, et en pratique un filtrage doit être appliqué aux données. Récemment, une nouvelle définition de l'offset, en profondeur, a été proposée, avec des applications de la DSO sur des données 2D synthétiques et réelles [Shen and Symes, 2008]. L'image migrée est obtenue avec

$$M(x,z,h) = \iiint \mathrm{d}s\,\mathrm{d}r\,\mathrm{d}\omega(-\omega^2)G(s,x-h,z,\omega)D(s,r,\omega)G(r,x+h,z,\omega), \qquad (4.3)$$

où h est l'offset en profondeur, G les fonctions de Green et D les données enregistrées pour lesquelles l'arrivée directe est retirée. Chaque section M correspond à la sommation de toutes

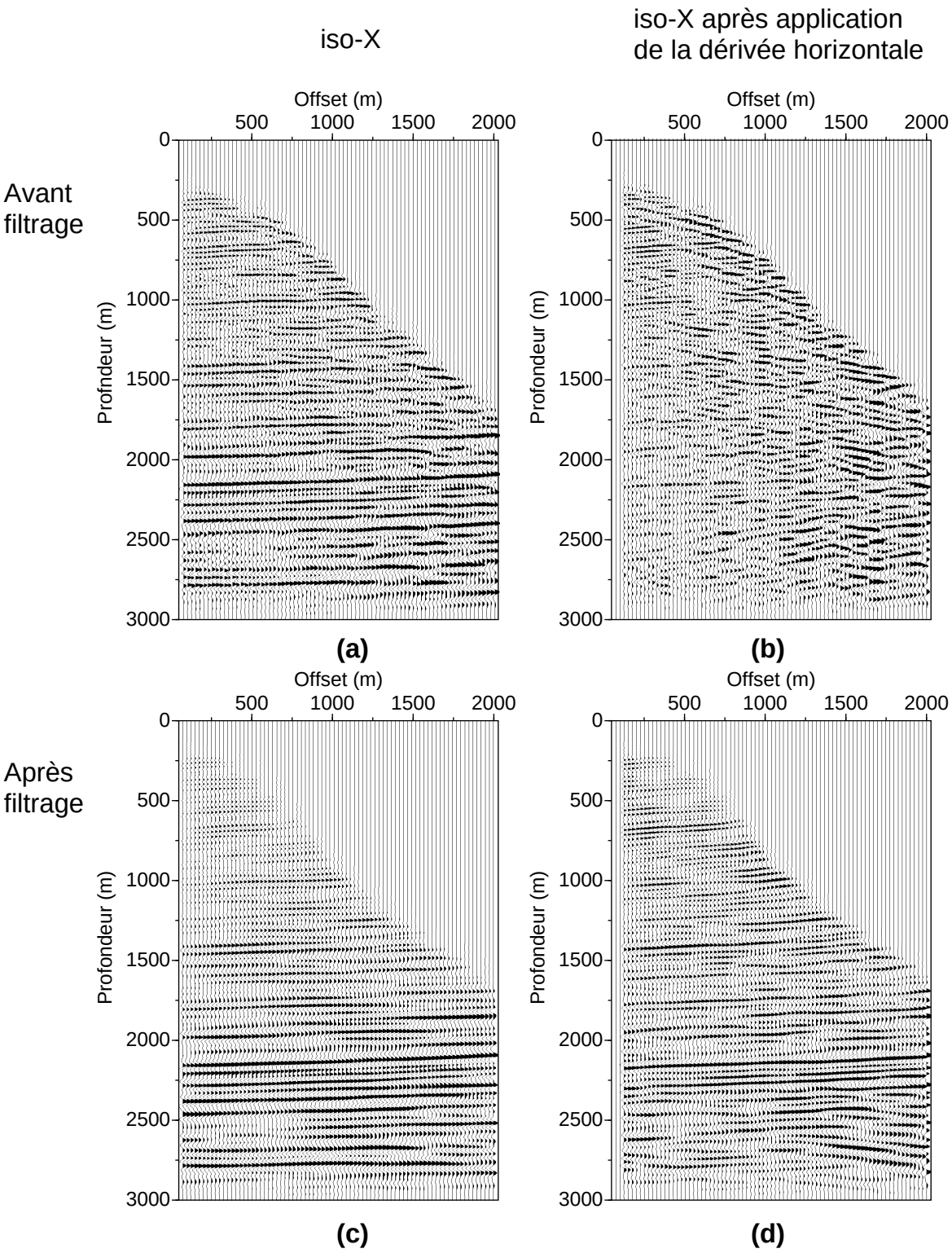


Fig. 4.4 – *Exemples de CIGs avant (gauche) et après (droite) l'application de la dérivée horizontale. Le bruit cohérent est amplifié (b). Après filtrage dans le domaine de Fourier, les CIGs montrent des événements plus continus (c et d) [Chauris and Noble, 2001].*

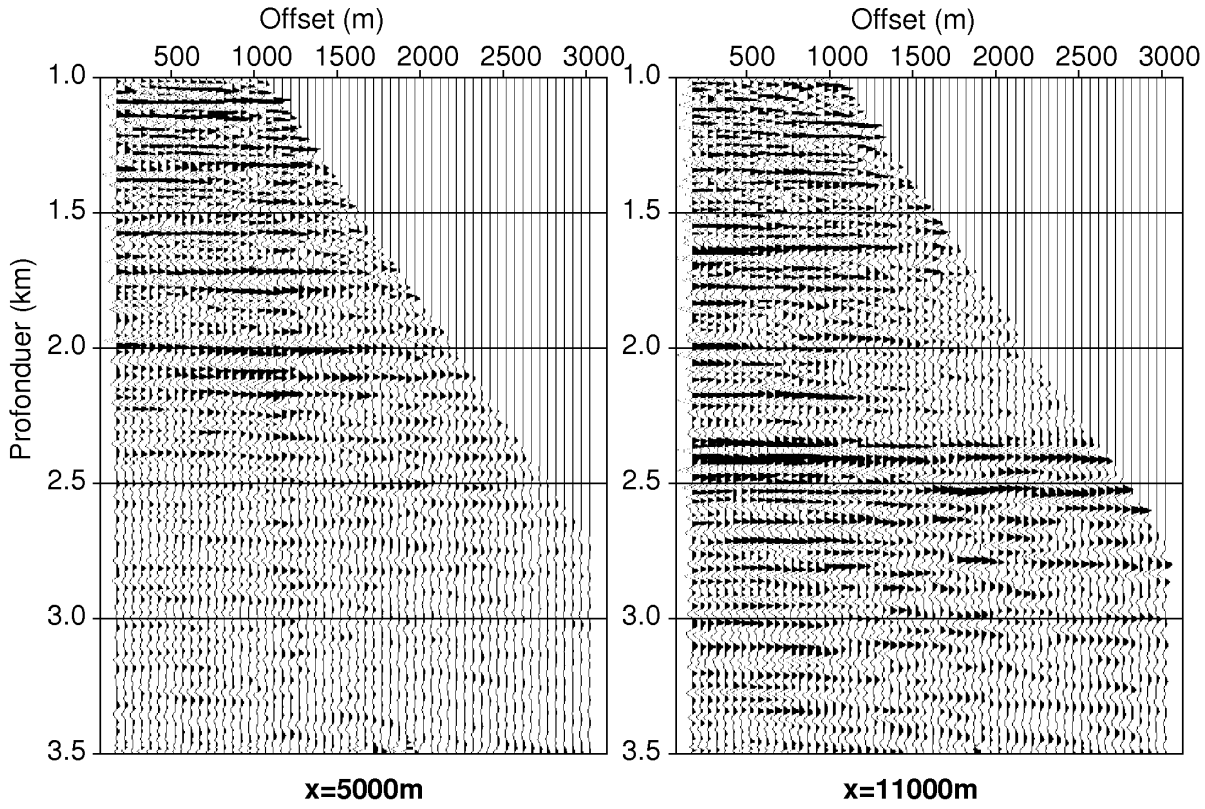


Fig. 4.5 – *CIGs calculés dans le modèle inversé [Chauris and Noble, 2001].*

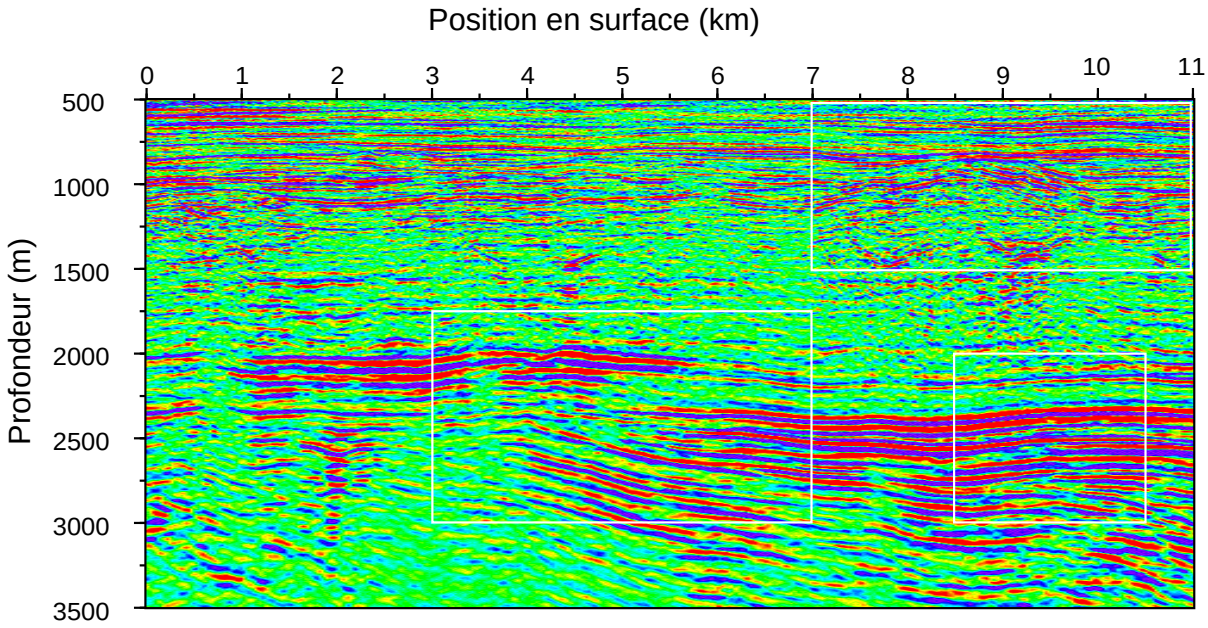


Fig. 4.6 – *Image migrée finale avec le modèle inversé [Chauris and Noble, 2001].*

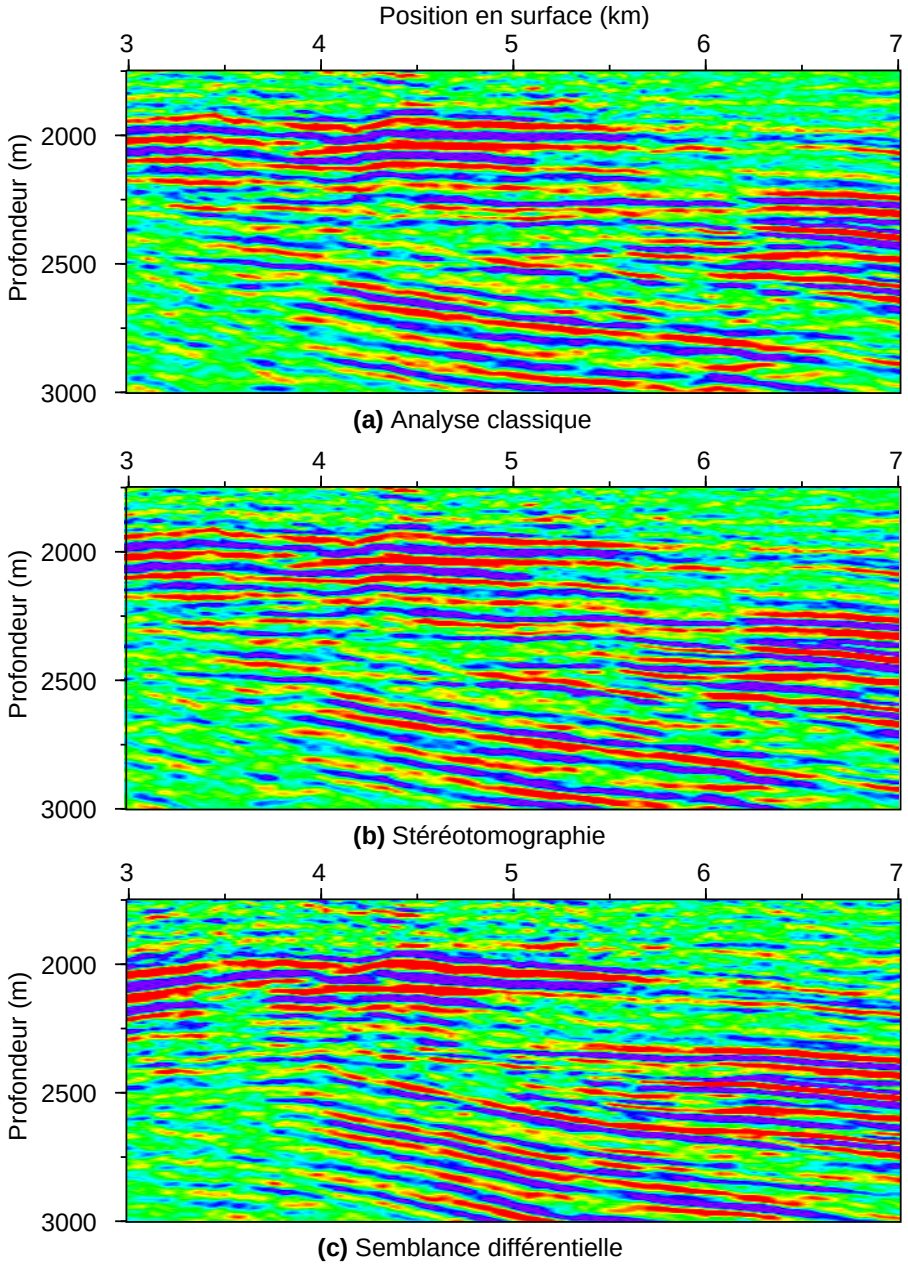


Fig. 4.7 – *Images migrées avec trois modèles de vitesse différents. La continuité des réflecteurs est meilleure pour la semblance différentielle, en particulier dans les blocs basculés [Chauris and Noble, 2001].*

les sources et de toutes les récepteurs. La corrélation des champs est faite avec un décalage en profondeur de valeur h. La nouvelle fonction objective est définie par

$$J_{\mathrm{DSO\_profondeur}} = \frac{1}{2} \iiint \mathrm{d}x\, \mathrm{d}z\, \mathrm{d}h\, |h \cdot M(x,z,h)|^2 . \tag{4.4}$$

Pour le bon modèle de vitesse, l'énergie est concentrée autour de $h = 0$ et donc il s'agit bien de minimiser $J_{\mathrm{DSO\_profondeur}}$ qui a de bonnes propriétés pour l'optimisation. La définition de

l'offset en profondeur est avantageuse dans le cas de modèles complexes avec des triplications [Stolk and Symes, 2004].

4.3 Les curvelets : introduction

Dans le chapitre 3 et au début du chapitre 4, j'ai discuté des avantages et inconvénients de l'approche avec pointés ou sans pointé de type DSO pour l'analyse de vitesse. Il en resort que les curvelets offrent une approche intermédiaire qui est analysée maintenant.

Pour les articles fondateurs, j'en réfère à [Candès and Donoho, 2004] et [Candès et al., 2006]. Les curvelets ont d'abord été construites pour la compression et le débruitage d'images 2D, pas nécessairement sismiques, dès lors que ces images sont lisses avec des discontinuités également lisses [Do, 2001, Do and Vetterli, 2003, Candès and Donoho, 2004]. Une telle image 2D $f(x,z)$ peut être décomposée en une série de curvelets. Pour l'image $f_n(x,z)$ reconstruite à partir des n coefficients de plus forte amplitude, nous avons

$$||f - f_n||^2_{L_2} \sim \frac{\log^3 n}{n^2}. \tag{4.5}$$

Le taux de convergence optimal est de $1/n^2$. Par comparaison, les décompositions dans le domaine de Fourier et des ondelettes [Mallat, 1989, Daubechies, 1992, Meyer, 1993] donnent des représentations moins efficaces, puisque les taux de convergence sont respectivement en $1/\sqrt{n}$ et $1/n$. Les curvelets ont trois propriétés importantes. Premièrement, les curvelets forment une "tight frame", c'est-à-dire que comme pour une base orthonormale, toute fonction $f(x,z)$ peut être décomposée en une série de curvelets $c_\mu(x,z)$ avec la formule de reconstruction

$$f(x,z) = \sum_\mu < f, c_\mu > c_\mu(x,z), \tag{4.6}$$

avec $< f, c_\mu >= \iint \mathrm{d}x\,\mathrm{d}z f(x,z) c_\mu(x,z)$. Deuxièmement, les curvelets sont par essence allongées (Figure 4.8). Dans le domaine spatial, la largeur de la curvelet est proportionnelle au carré de la longueur. Enfin, toute curvelet peut être déduite d'une curvelet de référence, soit par translation, par rotation ou cisaillement, ou par dilatation (Figure 4.9). Le paramètre μ indique donc la position centrale de la curvelet, sa direction et son contenu fréquentiel.

La relation parabolique entre largeur et longueur est expressément choisie pour les curvelets et donne la relation 4.5. Elle peut être perçue comme un compromis : des curvelets très alongées seraient immédiatement déformées lors de la propagation. D'un autre côté, les curvelet trop courtes ne seraient plus vraiment directionnelles et après migration, l'image d'une curvelet serait très différente.

L'objectif du reste de ce chapitre est de montrer en quoi les curvelets peuvent être utiles pour le traitement des données sismiques, mais aussi quelles sont les limitations de leur utilisation. Je donnerai quelques éléments sur deux codes qui ont été développés au sein de l'équipe de Géophysique.

Deux groupes (Mines ParisTech et l'université de British Columbia dirigée par Felix Herrmann), ont beaucoup travaillé sur les curvelets et leurs applications pratiques au traitement sismique, avec deux approches au final très différentes. Le groupe de Felix Herrmann a surtout

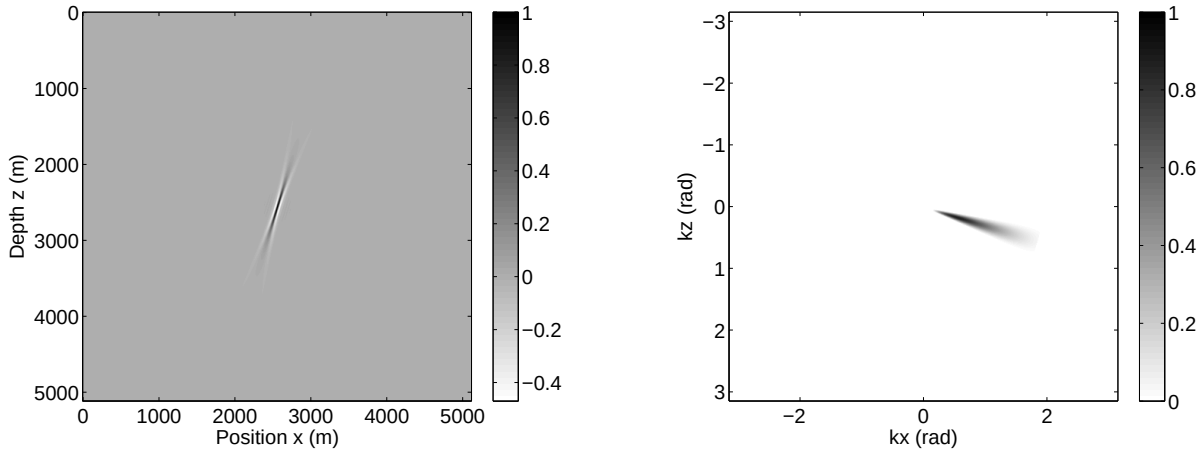


Fig. 4.8 – *Représentation d'une curvelet dans le domaine spatial (a) et dans le domaine des nombres d'onde (b) où la curvelet est localisée [Chauris and Nguyen, 2008a].*

pour objectif d'utiliser les curvelets pour des tâches de compression de données, de débruitage et d'interpolation [Herrmann et al., 2008a]. On peut citer les travaux sur le débuitage [Neelamani et al., 2008, Shan et al., 2009], sur la compression et la régularisation des données [Lin and Herrmann, 2007, Hennenfent and Herrmann, 2008, Tang et al., 2009, Herrmann et al., 2009b], sur la séparation des primaires et des multiples [Herrmann et al., 2008b] et sur l'amplitude préservée en migration [Herrmann et al., 2009a]. Je me suis plus intéressé au lien entre curvelets et propagation des ondes.

4.4 Autres ...lets

Je renvoie au workshop EAGE 2007 à Londres "Curvelets, contourlets, seislets, ... in seismic data processing : where are we and where are we going ?", que j'ai organisé avec Huub Douma de l'Université de Princeton, ainsi qu'au workshop EAGE 2009 à Amsterdam et aux articles de [Ma and Plonka, 2009, Chauris et al., 2010a].

Un certain nombre de transformées a été proposé avec des objectifs différents. Les premières ondelettes très anisotropes sont appelées ridgelets [Candès and Donoho, 1999b]. Elles sont optimales pour représenter des singularités le long de droites, ce qui est rarement le cas dans la nature. La première génération de curvelets est issue de la décomposition en ridgelets d'une série de petites zones dans une image [Candès and Donoho, 1999a]. Plus tard, la seconde génération de curvelets s'est beaucoup simplifiée avec un découpage de l'espace fréquentiel comme expliqué plus bas [Candès and Donoho, 2004]. En dehors de ces transformées, d'autres fonctions ont été proposées. Je les classe en trois catégories.

Les transformées générales

- les "steerable wavelets" [Freeman and Adelson, 1991, Simoncelli et al., 1992] et les ondelettes de Gabor [Lee, 2008] sont les premières ondelettes directionnelles ;
- les shearlets [Guo and Labate, 2007] ont pour opérations élémentaires les translations, les

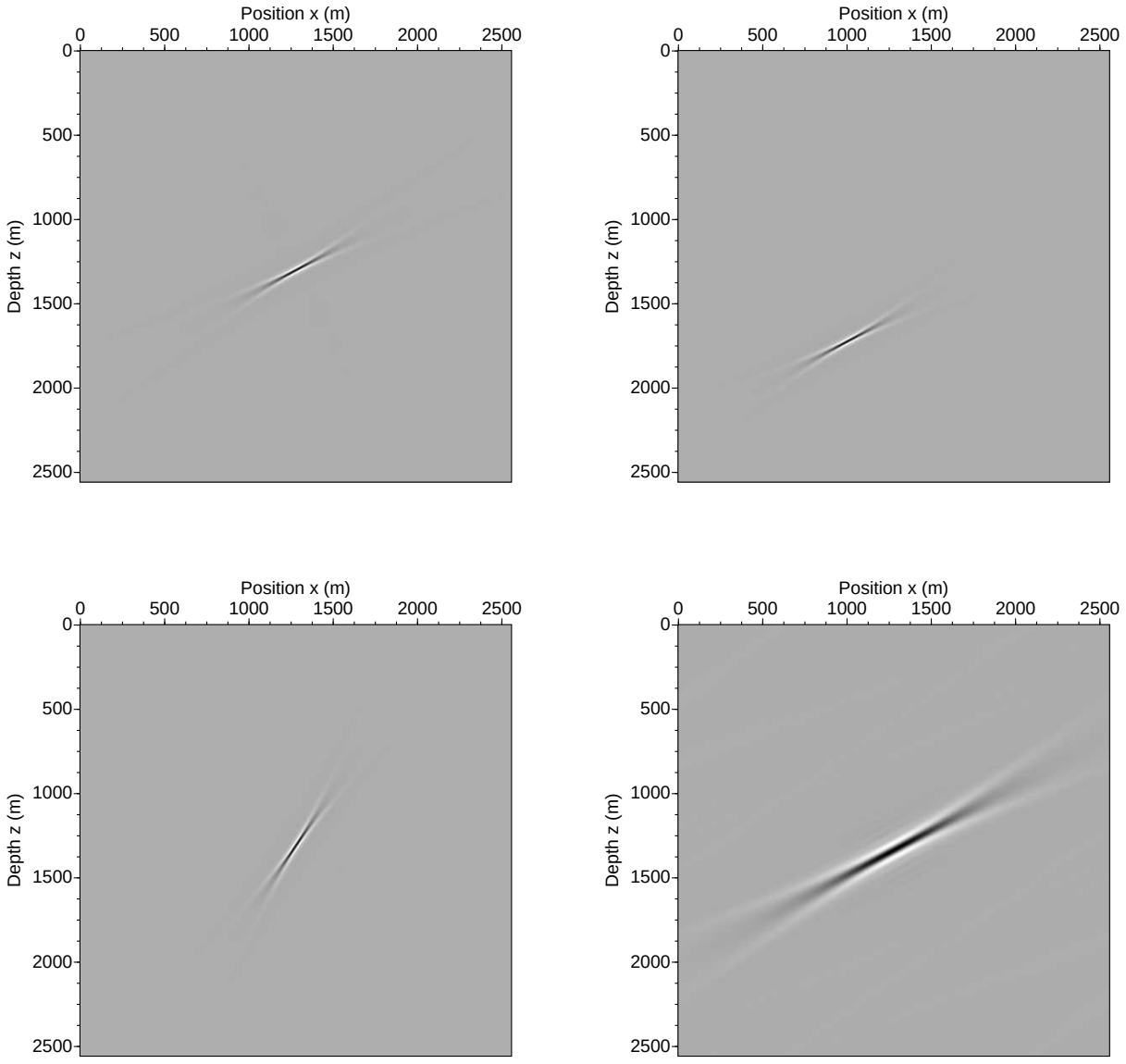


Fig. 4.9 – *Les curvelets dans le domaine spatial peuvent être déduites d'une curvelet de référence (a) par translation (b), par rotation (c) et dilatation (d) [Chauris and Nguyen, 2008a].*

cisaillements et les dilatations, et sont donc très proches des curvelets ;

- les contourlets [Do, 2001, Do and Vetterli, 2003] citées plus bas, utilisent une cascade de filtres ingénieusement agencés avec lesquels les données sont convoluées pour obtenir la transformée ;
- les surfacelets [Lu and Do, 2007] sont les extensions des contourlets en 3D ;
- Les "wavelet atoms" ne sont pas aussi directionnels que les curvelets [Demanet and Ying, 2007]. Ils sont en pratique beaucoup plus oscillants et sont utilisés pour la compression des don-

nées.

Les transformées adaptatives

- les bandlets [Le Pennec and Mallat, 2005, Mallat and Peyré, 2007] utilisent une technique adaptative : la transformée est différente pour chaque image en fonction de la structure de cette image ;
- les seislets [Fomel, 2006] sont dans l'esprit proches des bandlets. Les directions de décomposition sont données comme information a priori ;
- les wedgelets [Donoho, 1999] comprennent deux décompositions : d'abord, des ondelettes classiques divisent l'image en une série de blocs dyadiques, puis ces blocs sont approximés par des "wedgelets", c'est-à-dire des fonctions constantes par morceaux. Les wedgelets ne sont pas très adaptées pour le traitement des données sismiques en raison de leur caractère oscillatoire ;
- les platelets [Willet and Novak, 2003] décomposent les images en zones délimitées par des singularités, tout comme les "wedgelets", et sont a priori robustes vis-à-vis des bruits non Gaussiens.

Les transformées pour l'équation des ondes

- les faisceaux Gaussiens fournissent une solution asymptotique de l'équation des ondes, avec un profil Gaussien perpendiculaire au rai central [Hill, 2001]. L'amplitude décroit autour de ce rai central au cours de la propagation et en fonction du modèle de vitesse ;
- les paquets Gaussiens [Zacek, 2003, Zacek, 2005] sont semblables aux faisceaux Gaussiens, mais localisés en temps ;
- les beamlets [Chen et al., 2006, Wu et al., 2008] sont utilisées en particulier pour des corrections d'illumination et pour des propagations "one-way" [Wu and Chen, 2006, Cao and Wu, 2009].

La liste n'est pas exhaustive. Je discute plus bas d'une autre transformée "Easy Path Wavelet Transform" pour la compression des données sismiques [Ma et al., 2010].

4.5 Les curvelets : deux nouveaux codes

Dans les transformées générales, deux grands types d'implémentation existent : les curvelets [Candès et al., 2006] et les contourlets [Do and Vetterli, 2003]. Dans le premier cas, la décomposition en curvelets est obtenue par filtrage dans le domaine de Fourier. Les données $f(x, z)$ sont d'abord transformées dans le domaine de Fourier $f(k_x, k_z)$. Pour chaque direction et pour chaque dilatation est défini un filtre $F_i(k_x, k_z)$ dans le domaine de Fourier. L'ensemble des coefficients correspondant à ce filtre est obtenu en appliquant la transformée de Fourier inverse au produit $f(k_x, k_z) \cdot F_i(k_x, k_z)$. La transformée inverse en curvelet est très similaire. D'abord les coefficients pour chaque direction et chaque dilatation sont transformés dans le domaine de Fourier. Ensuite, le conjugué du même filtre F_i est appliqué. Finalement, l'image finale est obtenue en sommant toutes les contributions et en prenant la transformée de Fourier inverse. Pour obtenir une image identique, une condition simple existe sur les filtres F_i : pour tout (k_x, k_z),

$$\sum_i |F_i(k_x, k_z)|^2 = 1, \tag{4.7}$$

puisque le filtre est appliqué deux fois [Candès et al., 2006, Chauris and Nguyen, 2008a]. Le choix des filtres particuliers, avec la relation parabolique entre largeur et longueur, definit les curvelets. Si d'autres filtres respectent la relation 4.7, alors la même procédure de décomposition et de reconstruction s'applique. C'est selon ce schéma que je définis une nouvelle transformée (chapitre 5).

Dans le cas des contourlets, l'idée est la même, mais au lieu de multiplier dans le domaine de Fourier par des filtres, il s'agit d'appliquer des convolutions dans l'espace (x, z). Afin d'optimiser le nombre de calculs, les filtres sont remplacés par une cascade de filtres très ingénieusement agencés. Dans la pratique, les contourlets souffrent de problèmes de bord [Nguyen and Oraintara, 2006].

Ma contribution, avec l'aide de Truong Nguyen pendant sa période post-doctorale, aura été de définir deux nouvelles transformées en curvelets. La première est proche des "steerable wavelets" [Simoncelli et al., 1992] dans l'esprit des curvelets [Candès et al., 2006]. Au lieu de travailler sur des coordonnées cartésiennes, nous avons préféré les coordonnées polaires. Ensuite, les filtres ont été choisis pour faciliter par la suite un certain nombre d'opérations élémentaires dans le domaine des curvelets. Par exemple, à une certaine gamme de fréquences, il existe 16 directions différentes. Pour obtenir une rotation d'une curvelet d'un angle multiple de $\pi/16$, il suffit de changer l'indice de la curvelet. Par contre, pour une rotation de $\alpha\pi/16$ avec $\alpha \in]0, 1[$, c'est bien moins évident. Des filtres ont été spécialement conçus pour faciliter la tâche [Chauris, 2006, Chauris and Nguyen, 2008a].

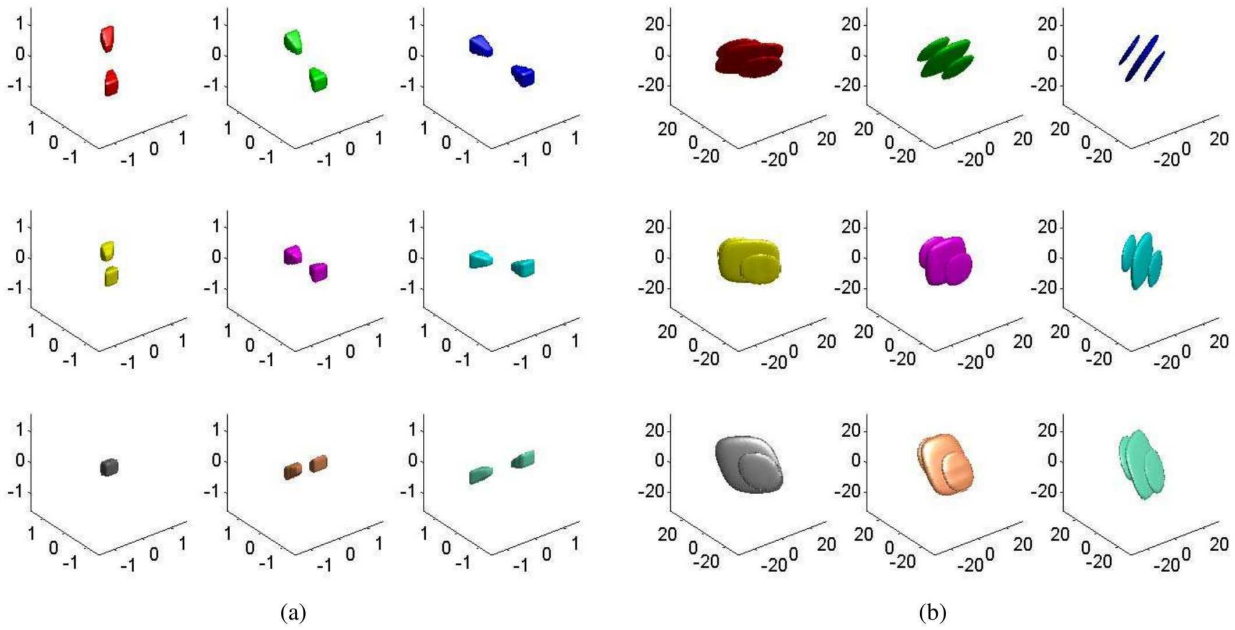


Fig. 4.10 – *Exemples de curvelets 3D dans le domaine spatial (a) et dans le domaine des nombres d'onde (b) [Nguyen and Chauris, 2010].*

Un second code a été développé en parallèle [Nguyen and Chauris, 2010]. L'objectif était de pouvoir appliquer la transformée en curvelet pour des dimensions supérieures à deux. L'idée principale est de combiner l'approche des contourlets pour définir les filtres et limiter la redondance des données, et d'utiliser l'idée des curvelets pour appliquer ces filtres par multiplication dans le domaine de Fourier. Un exemple de ces curvelets est donné sur la Figure 4.10. La redondance obtenue en 3D est de 7.6, tandis que pour les curvelets sensu stricto elle est de 27.0. Pour les surfacelets, elle n'est que de 3.4, mais la transformée souffre d'artefacts sur les bords [Lu and Do, 2007]. Par ailleurs, les tailles des volumes et le nombre de directions peuvent être différents selon x, y et z.

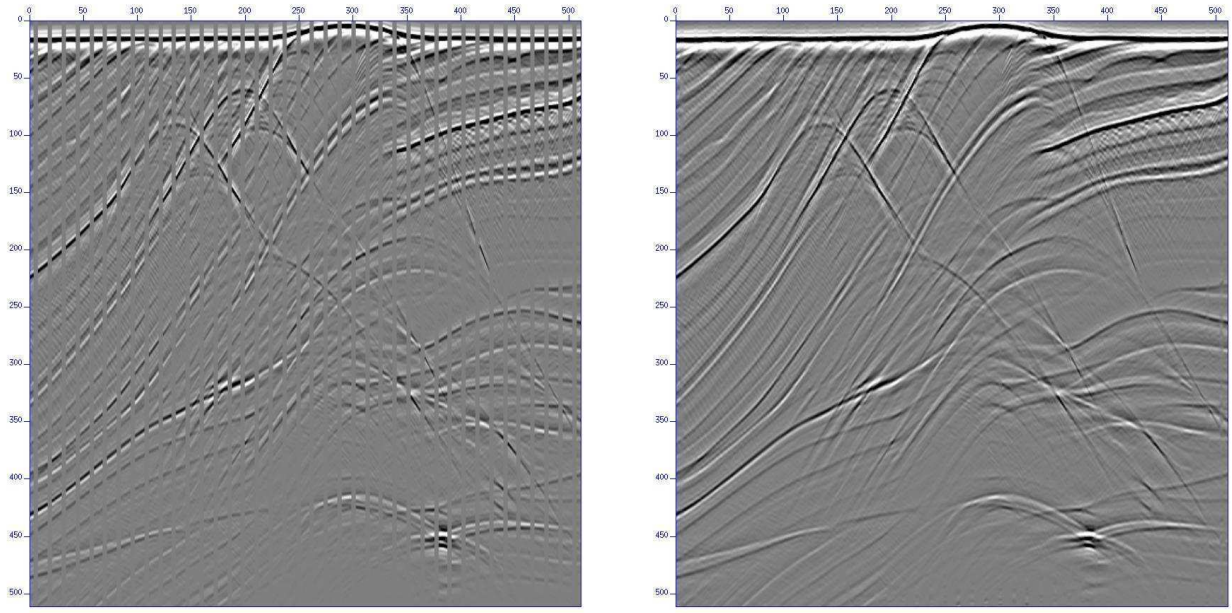


Fig. 4.11 – *Interpolation de traces avec les curvelets (avant (a) et après (b)) [Nguyen and Chauris, 2010].*

Une application très simple de la nouvelle transformée 3D est présentée sur la Figure 4.11. Les curvelets exploitent la continuité latérale pour interpoler les traces manquantes. Je discute maintenant de l'intérêt des curvelets au travers d'une série d'applications.

4.6 Curvelets et multiples : prédiction et élimination

La suppression des multiples est une tâche importante dans la chaîne de traitement sismique. En effet, les multiples peuvent générer après migration des événements perçus comme des réflexions primaires. La méthode SRME (Surface-Related Multiple Elimination) est la méthode la plus fiable pour prédire les multiples à partir des données enregistrées, sans connaissance a priori de la Terre [Verschuur et al., 1992]. Son principe est simple : la prédiction des multiples est réalisée en convoluant les données avec elles-mêmes.

Nous montrons comment cette prédiction peut être réalisée dans le domaine des curvelets [Donno et al., 2010]. Il est possible d'utiliser le principe de Fermat à la surface libre pour ne

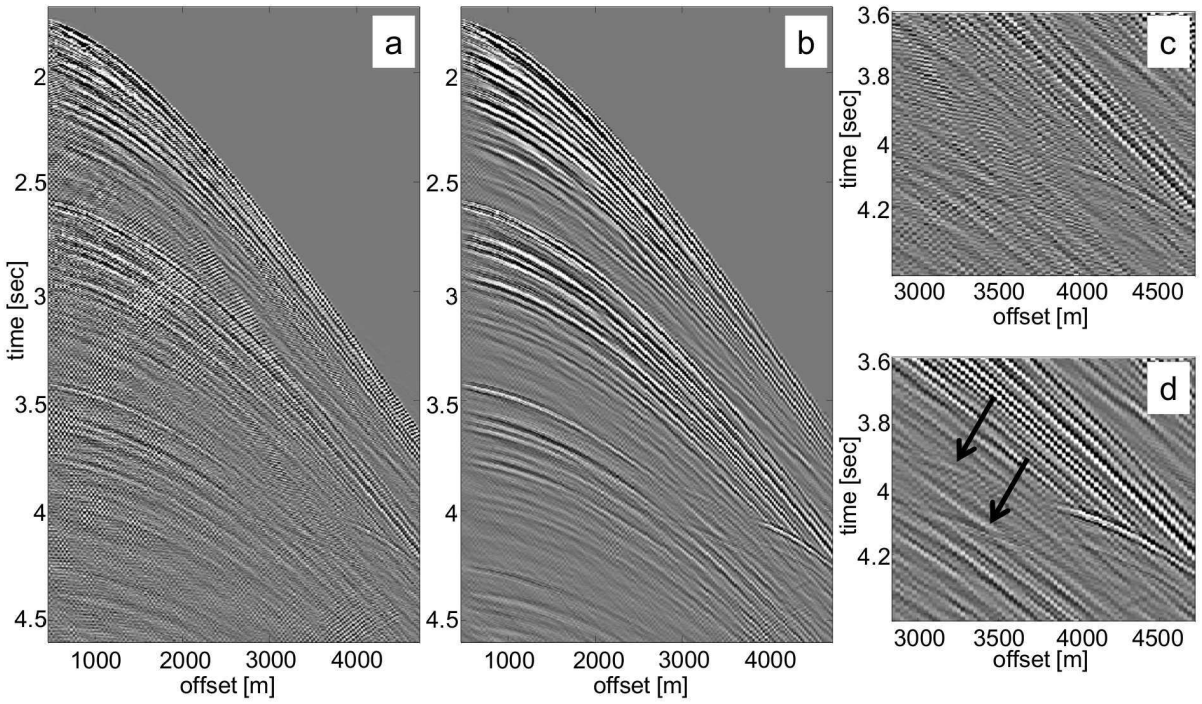


Fig. 4.12 – *Prédictions des multiples pour un jeu de données réelles, (a et c) avec la méthode classique SRME, Surface-Related Multiple Elimination [Verschuur et al., 1992], (b et d) avec la prédiction dans le domaine des curvelets et la prise en compte de l'aliasing [Donno et al., 2010].*

combiner que les curvelets qui ont des directions opposées. Mais après sommation, le résultat est le même pour une bonne illumination. Par contre, les curvelets jouent un rôle dans le cas de données aliasées : la prédiction des multiples est pondérée par la valeur des coefficients de curvelets pour la même position et la même direction, mais pour des fréquences centrales plus faibles pour lesquelles l'aliasing est réduit (Figure 4.12). L'approche multi-échelle contenue dans les curvelets apporte une vraie différence.

La partie élimination des multiples avec les curvelets a été adressée par [Herrmann et al., 2008b].

4.7 Curvelets et migration

Les travaux théoriques de [Candès and Donoho, 2004] montrent que les curvelets sont optimales en tant que vecteurs propres de l'opération de modélisation, au moins pour des milieux lisses. Il est donc naturel de décomposer l'opérateur de modélisation ou de migration et de tester les limites de l'approche. Au même moment, plusieurs articles sont parus [Chauris, 2006, Douma and de Hoop, 2006, Douma and de Hoop, 2007]. Douma a montré que pour un milieu homogène, l'effet de l'opérateur de migration pouvait s'exprimer comme une combinaison de translation, rotation, cisaillement et de dilatation. En parallèle, mon approche a été de réduire l'opérateur de migration à des combinaisons d'opérations élémentaires, c'est-à-dire (selon le code employé) translation, rotation et dilatation. L'élément clé est, pour chaque curvelet considérée, d'exprimer la migration du coefficient dans le domaine de Fourier (k_x, k_z) et de tirer parti

de la localisation de la curvelet dans ce domaine, selon un axe k_x/k_z donné.

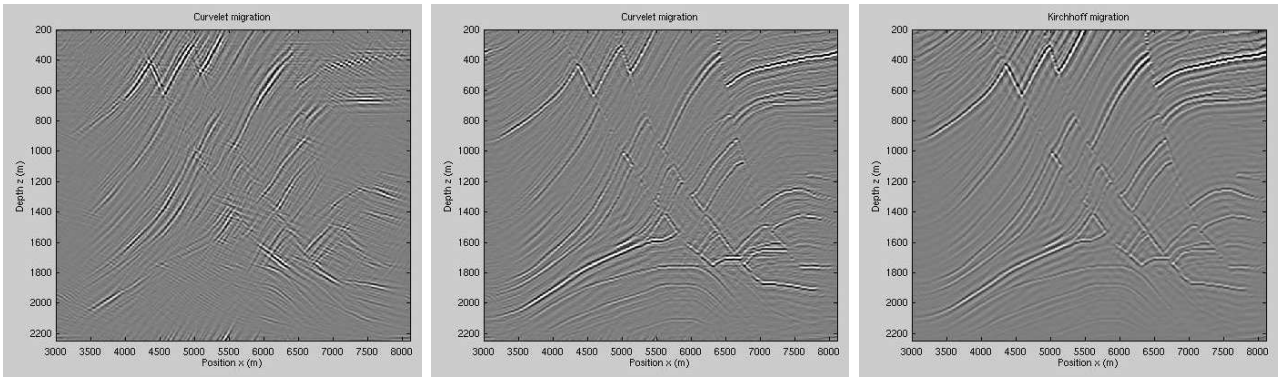


Fig. 4.13 – *Migration des données Marmousi rai+Born pour l'offset zéro et un milieu homogène : migration dans le domaine des curvelets sans interpolation (gauche), avec interpolation (milieu) et migration classique de Kirchhoff [Chauris, 2006].*

Les applications pratiques montrent que la migration dans le domaine de curvelets peut bien être réduite à une combinaison élémentaire d'opérations dans le cas d'un milieu de vitesse homogène (Figure 4.13). Cette combinaison s'exprime dans le domaine des curvelets comme un simple changement d'indices. Cependant, une interpolation est nécessaire car les directions des curvelets sont discrètes. Pour pouvoir le faire de manière efficace, des filtres ont été spécialement définis [Chauris, 2006, Chauris and Nguyen, 2008a].

Par contre, pour la migration dans des milieux lisses et pour des parties profondes, la migration des curvelets restreinte à une combinaison d'opérations élémentaires ne donne pas la qualité du résultat espéré (Figure 4.14). Ce n'est pas en contradiction avec les travaux théoriques, puisque ils sont valables pour des propagations sur un faible nombre de longueurs d'onde. Sur un plan plus quantitatif, les travaux sur les faisceaux Gaussiens montrent que la courbure des événements après migration doit être prise en compte, ce qui n'est pas le cas lorsque la migration est restreinte à des opérations élémentaires [Hill, 2001]. Il est possible de combiner la décomposition des données sismiques en somme de curvelets avec un code de migration par faisceaux Gaussiens (Figure 4.15) [Jilek et al., 2007]. L'avantage de la décomposition en curvelets par rapport à un "slant stack" local est que le nombre de pentes et la largeur de la fenêtre sur laquelle la décomposition est faite dépendent de la bande passante du signal. Pour les basses fréquences, plus de traces sont prises en compte pour moins de pentes. Pour les hautes fréquences, c'est le contraire. Dans une décomposition classique pour les faisceaux Gaussiens, le nombre de traces est fixe et le nombre de pentes est dicté par la fréquence maximale du signal. Ce n'est pas très adapté pour l'imagerie à la fois de zones peu et très profondes puisque la courbure des données dans les sismogrammes est variable. La décomposition en curvelets offre dans ce cadre une décomposition efficace et a priori mieux adaptée.

4.8 Curvelets et démigration/migration, analyse de vitesse

Avec les limites de la migration dans le domaine de curvelets, même dans un cas asymptotique, il est venu l'idée de combiner migration et démigration (modélisation) dans deux modèles

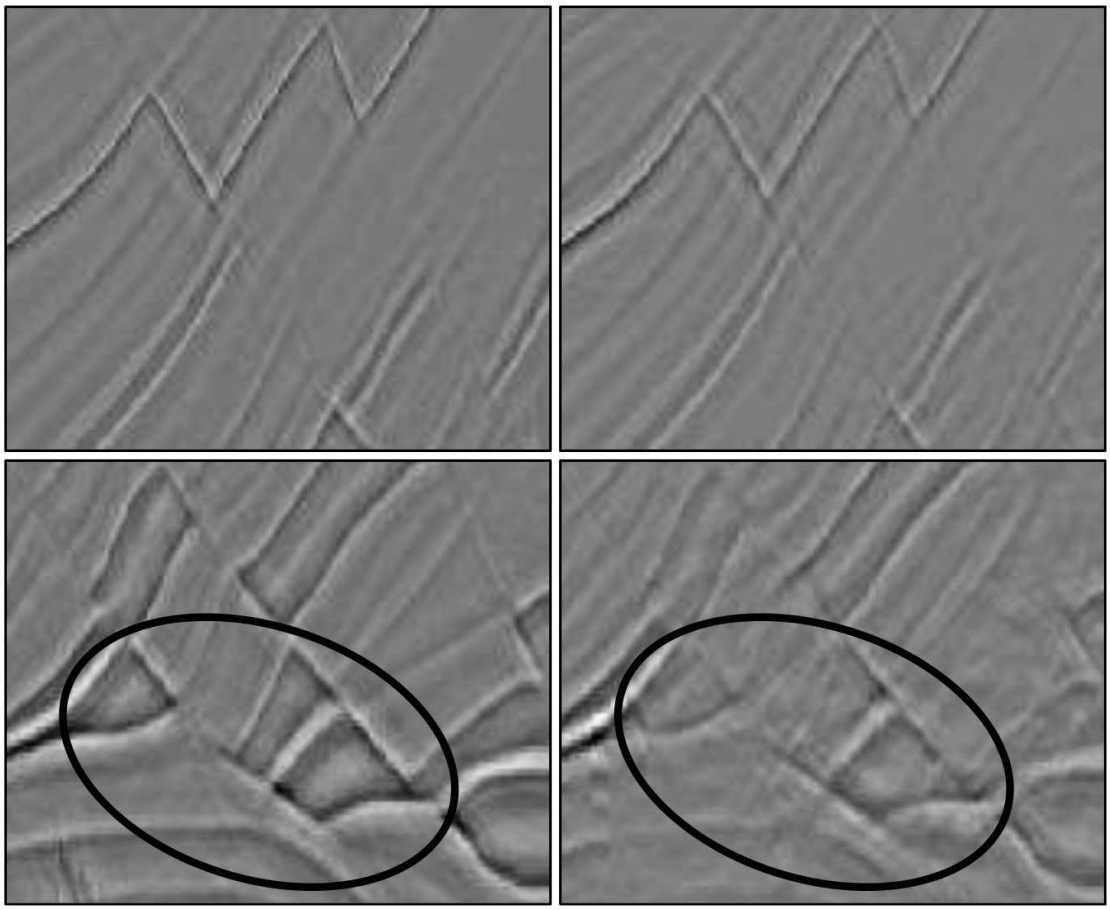


Fig. 4.14 – *Migration de Kirchhoff (gauche) et migration dans le domaine des curvelets (droite) pour les données Marmousi rai+Born, dans un milieu hétérogène lisse. La migration au premier ordre dans le domaine des curvelets n'est pas satisfaisant pour la partie profonde (bas).*

de vitesse localement différents [Chauris and Nguyen, 2008a]. Dans un modèle de vitesse hétérogène donné, une première migration est réalisée, typiquement avec un code de migration de Kirchhoff. L'image obtenue est considérée comme image de référence. Dans [Chauris and Nguyen, 2008a], nous montrons comment les curvelets peuvent avantageusement prédire l'image migrée qui aurait été obtenue dans un modèle de vitesse perturbé. L'image migrée de référence est d'abord décomposée dans le domaine des curvelets. Ensuite, les curvelets sont démigrées dans le modèle de référence puis migrées dans le modèle perturbé. En réalité, c'est l'opérateur combiné qui est appliqué. La démigration/migration est restreinte à une combinaison d'opérations élémentaires dans le domaine des curvelets (translation, rotation et dilatation). Enfin, l'image perturbée est obtenue en appliquant la transformée inverse en curvelets.

Un résultat théorique a pu être montré : dans le cas de démigration/migration et une perturbation locale, la courbure de la curvelet après démigration/migration est contrôlée par la valeur de la perturbation de vitesse. Dans la pratique, des perturbations de 10% de la vitesse peuvent être prises en compte. La courbure induite par la démigration est compensée par la migration qui suit. Il est donc aisé d'obtenir la partie de l'image migrée affectée par une

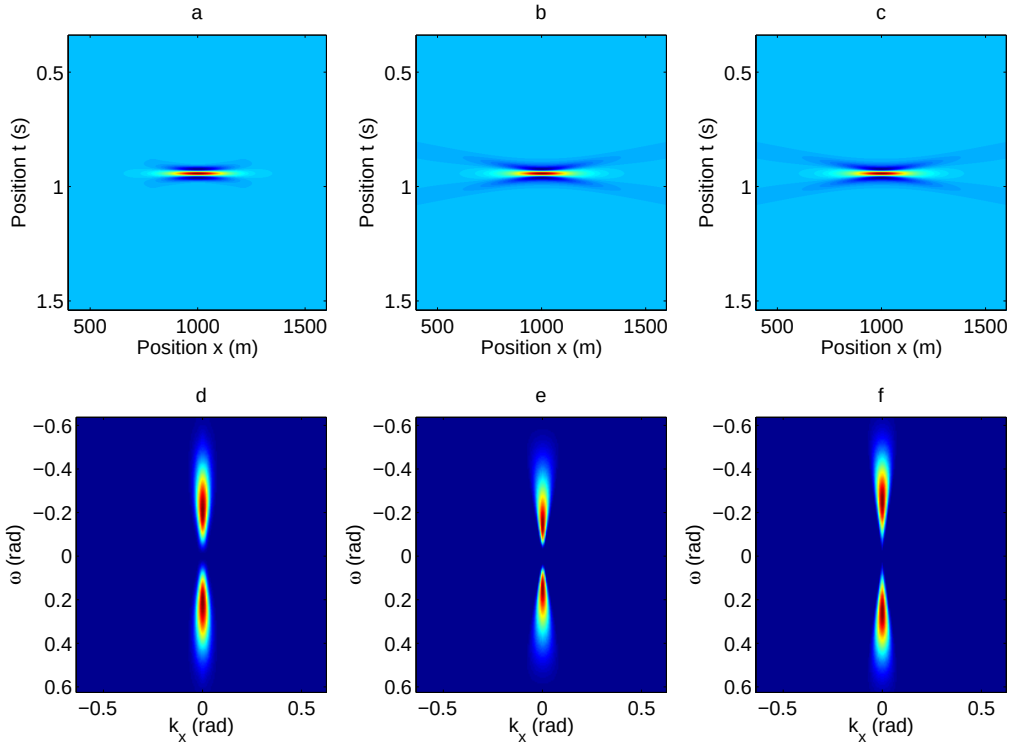


Fig. 4.15 – *Comparaison entre les faisceaux Gaussiens (a et d) et les curvelets (b, c, e et f). En haut : dans le domaine (t, x) et en bas dans le domaine (ω, k_x). Pour (b) et (e), les curvelets remplissent la condition dyadique sur les fréquences. Pour (c) et (f), la source est un Ricker, comme pour (a) et (d). Les faisceaux Gaussiens sont plus localisés en espace. Cependant, le nombre de pentes est dicté par la fréquence maximale du signal, alors que pour les curvelets, le signal est décomposé en plusieurs bandes de fréquences et pour chaque bande, le nombre de pentes change.*

perturbation de vitesse (Figure 4.16). Elle dépend de l'anomalie de vitesse, mais également de la réflectivité du milieu. En répétant la même opération pour un ensemble de sections à offset commun, il est possible de prédire des CIGs (Figure 4.17).

Dans le contexte de la migration en temps, la même stratégie de démigration/migration s'applique. A partir d'une section sismique migrée, les curvelets permettent de prédire efficacement le résultat de la migration correspondant à un autre modèle de vitesse. Des résultats sur données réelles ont été obtenus [Nguyen et al., 2009b] pour des perturbations de vitesse qui affectent l'ensemble du modèle, par exemple de 10% inférieures (Figure 4.18).

L'extension à l'analyse de vitesse est alors faisable [Chauris et al., 2008] avec une approche de type DSO (voir le début du chapitre). L'opérateur de dérivée horizontale selon la dimension offset est appliqué dans le domaine des curvelets en utilisant la relation de Parseval. Il a été montré comment calculer le gradient de la fonction objective. Les premiers résultats sont encourageants et mériteraient d'être étendus (Figure 4.19). L'attrait principal de cette approche est de combiner l'analyse de vitesse avec une décomposition flexible des données pour laquelle

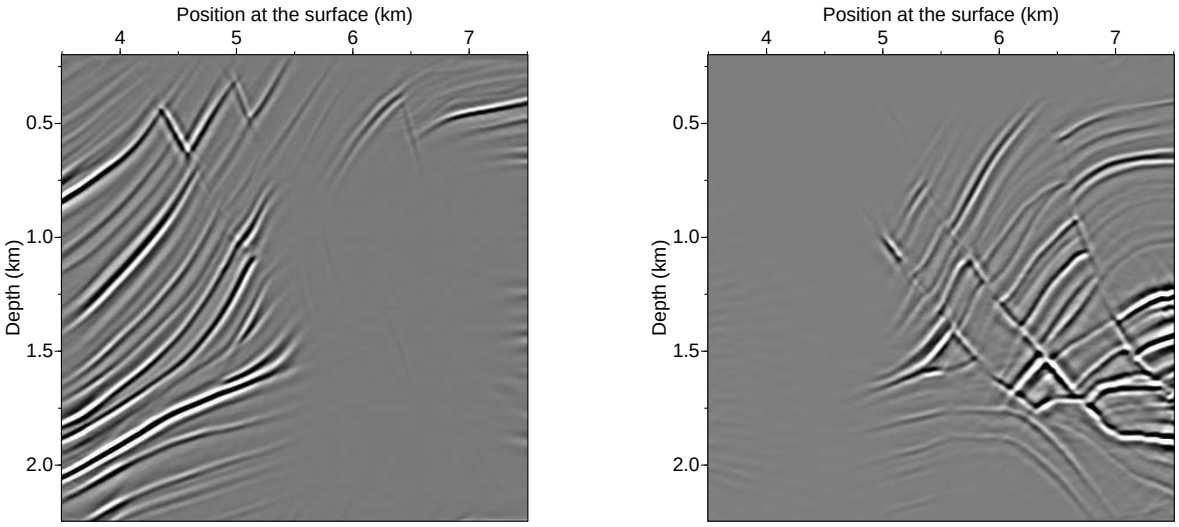


Fig. 4.16 – *Partie migrée non-perturbée (a) et perturbée (b), prédites par démigration/migration dans le domaine des curvelets [Chauris and Nguyen, 2008a].*

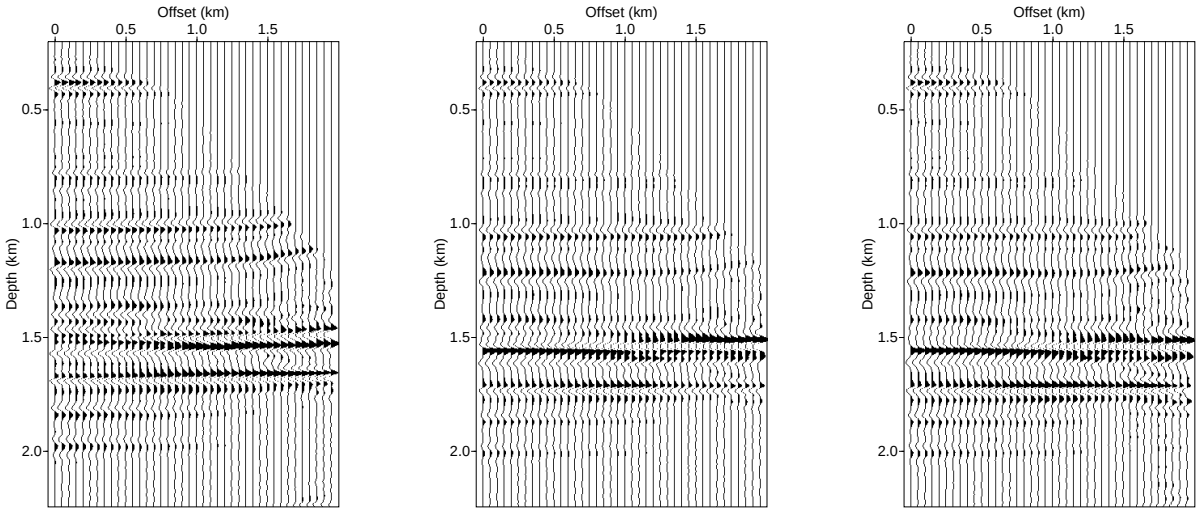


Fig. 4.17 – *CIGs pour la position x=6400 m (Figure 4.16), dans le modèle de vitesse initial (a), dans le modèle de référence (b) et après prédiction dans le domaine des curvelets [Chauris and Nguyen, 2008a].*

il est possible de sélectionner des événements selon des critères à préciser.

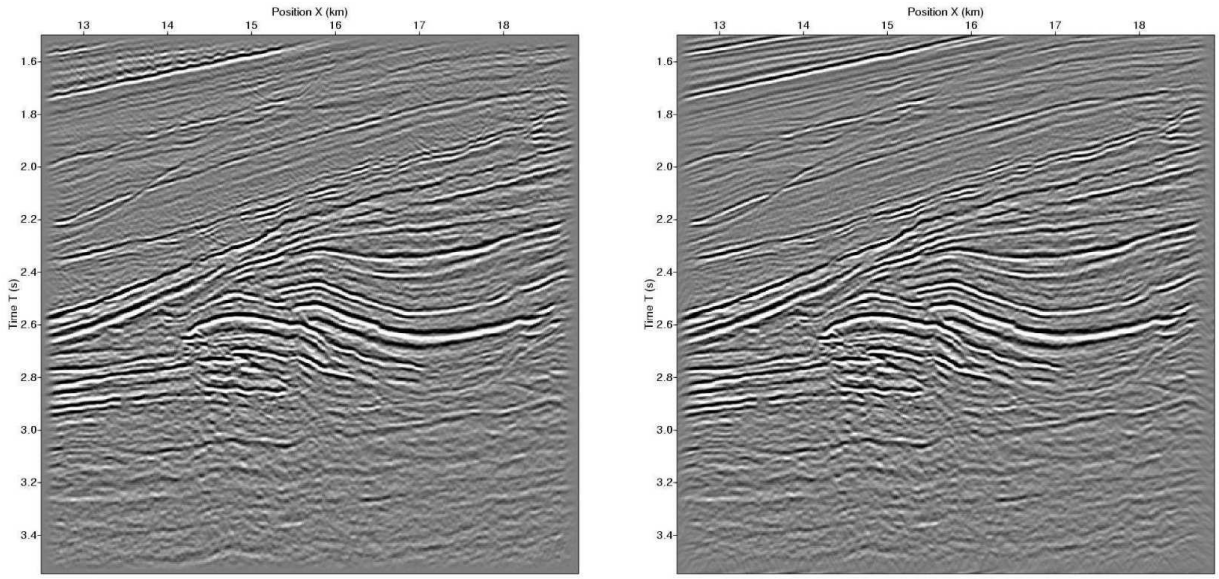


Fig. 4.18 – *Migration temps pour une section à offset commun (1550 m) pour un modèle de vitesse égal à 90% du vrai modèle de vitesse. Gauche : par migration classique de Kirchhoff en temps, droite : par démigration/migration dans le domaine des curvelets, en partant du modèle de référence [Nguyen et al., 2009b].*

4.9 Propagation dans les milieux complexes

Les résultats précédents sur la migration et la modélisation dans le domaine des curvelets ont été réalisés dans le cadre de la théorie des rais. La première possibilité de sortir du cadre asymptotique est celle proposée par [Jilek et al., 2007]. Il s'agit de décomposer les données dans les curvelets, de sélectionner les coefficients "significatifs" en fonction du modèle de vitesse, de l'information a priori, ..., puis de recomposer les données et de les utiliser avec un code de migration par différences finies. Dans le cas de [Jilek et al., 2007], les données sont décomposées en "slant stacks".

Une alternative est de résoudre l'équation des ondes par différences finies dans le domaine des curvelets. Des premiers résultats ont été obtenus par [Sun et al., 2009]. Nous avons montré que l'opérateur de Laplace pouvait s'exprimer simplement dans le domaine des curvelets. Il reste néanmoins à prendre proprement en compte l'effet de l'hétérogénéité du modèle de vitesse.

4.10 Conclusions

Au travers de différentes applications, il s'avère que les curvelets ne sont pas adaptées pour compression des données sismiques en raison de la trop grande redondance de la transformée. Par contre cette redondance est favorable pour le débruitage et pour l'interpolation de traces manquantes.

Pour la compression des données, des transformées alternatives telle que la transformée "Easy Path Wavelet Transform" sont plus adaptées (Figure 4.20) [Ma et al., 2010]. Cette transformée

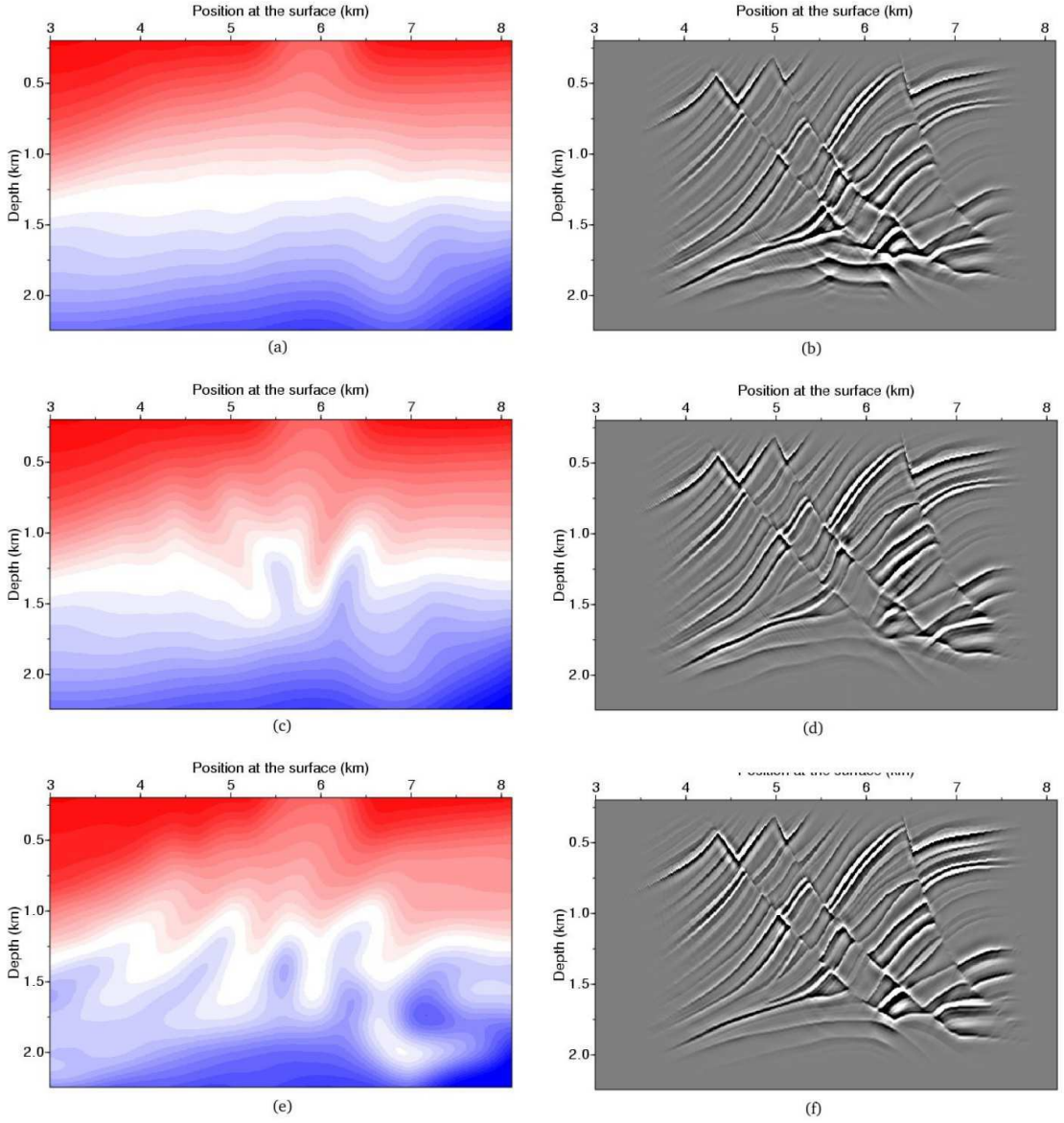


Fig. 4.19 – *Image initiale (haut), image correspondant à la première itération (milieu) et image pour le bon modèle de vitesse (bas) [Chauris et al., 2008].*

recherche d'abord un chemin au travers des données en minimisant l'écart entre les valeurs de deux points image successifs. Puis pour chaque chemin, les données sont réorganisées en fonction de leur amplitude et transformées dans le domaine des ondelettes 1D. La plus grande partie de l'information est contenue dans la partie lisse du signal, et donc très peu de coefficients suffisent pour bien représenter le signal. Bien-sûr, des coefficients supplémentaires sont nécessaires pour mémoriser les chemins.

Les résultats théoriques montrent que les curvelets sont adaptées pour la propagation des ondes dans des milieux lissees, au moins pour des propagations sur quelques longueurs d'onde. Dans la pratique et pour des propagations sur un plus grand nombre de longueurs d'onde, il

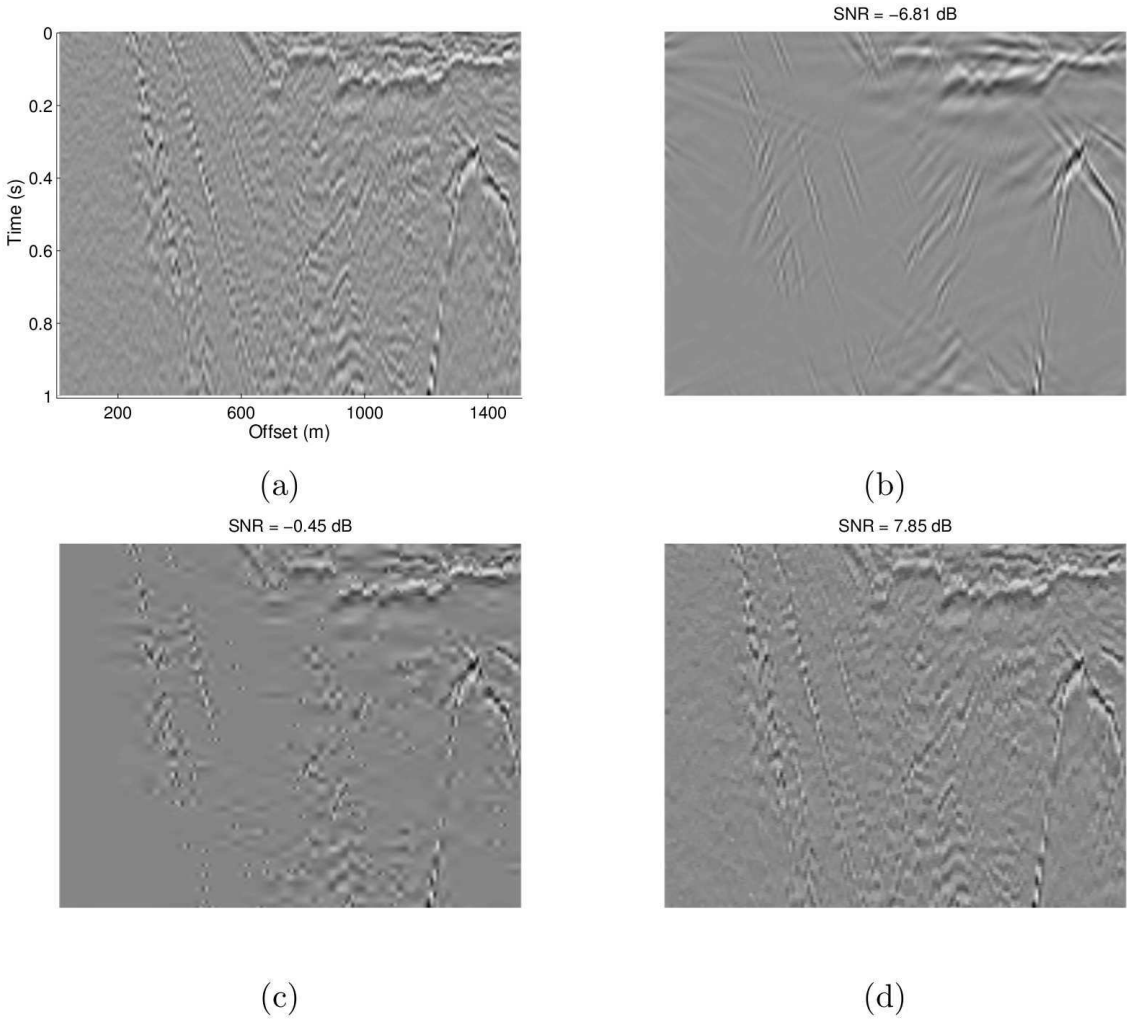


Fig. 4.20 – *Reconstruction des données sismiques réelles (a) avec la sélection des 1024 coefficients les plus forts, pour les curvelets (b), les ondelettes (c) et l'EPWT [Ma et al., 2010]. 1024 coefficients représentent un coefficient sur 160 par rapport à la taille de l'image originale.*

est essentiel de prendre en compte la courbure des fronts. La décomposition des données en curvelets reste cependant intéressante si elle est couplée avec des migrations de type faisceaux Gaussiens. Les curvelets peuvent être directement utilisées pour la démigration/migration dans deux modèles de vitesse localement différents ou pour l'analyse de vitesse. Des travaux sont en cours pour la résolution de l'équation des ondes par différences finies dans le domaine de curvelets.

En conclusion, il n'existe pas de transformée universelle, dans le sens où différents traitements sismiques sont facilités et/ou améliorés par une transformée particulière. D'un point de vue général, les curvelets apportent la flexibilité de la décomposition des données, pour supprimer par exemple les bruits cohérents ou pour ne sélectionner qu'un sous-ensemble des données dans l'espace des phases position/direction/contenu fréquentiel. C'est très important pour la

résolution de problèmes inverses. De plus, elles diagonalisent presque l'opérateur de démigration/migration pour deux modèles de vitesse différents : elles permettent de mieux comprendre les effets de propagation et comment le résultat de la migration dépend du choix du modèle de vitesse. Enfin, elles exploitent l'effet multi-échelle des données, par exemple pour supprimer l'aliasing lors de la prédiction des multiples.

Chapitre 5

Extensions et perspectives

Les extensions des curvelets et de leur utilisation ont abouti au développement d'une transformée en cercle appelée "circlet" [Chauris et al., 2010b] et à de nouvelles applications en océanographie côtière pour la détection et le suivi de tourbillons sur des images satellitaires de température de surface.

Mes travaux récents portent sur la thématique d'inversion des formes d'onde. Dans un premier temps et dans le cadre de l'inversion linéarisée, j'ai repris l'idée de démigration/migration issue des travaux sur les curvelets [Chauris and Nguyen, 2008a] pour la transposer à la migration par différences finies [Chauris and Benjemaa, 2010]. La problématique actuelle la plus importante me semble être celle de la connaissance du modèle de vitesse contenant les grandes longueurs d'onde pour l'imagerie des milieux complexes. J'ai lancé un nouveau projet de thèse pour reformuler le problème d'inversion des formes d'onde, en particulier pour s'affranchir de la détermination d'un modèle initial précis de vitesse. L'idée est de considérer toutes les ondes enregistrées, et pas seulement les ondes réfléchies, avec une approche d'inversion des formes d'onde différentielle qui exploite la cohérence locale des données.

5.1 Les "circlets"

Les curvelets ne sont en rien limitées à l'étude des données sismiques. L'article de [Ma and Plonka, 2009] donne un bon aperçu des différents champs d'application. Par exemple, les curvelets trouvent des applications dans l'étude des turbulences [Ma et al., 2006, Ma and Hussaini, 2007, Bermejo-Moreno and Pu
Dans le contexte de mesures de température de surface (SST) sur des images satellite en océanographie côtière, il s'avère que les prédictions des courants marins issues de modèles ne correspondent pas toujours parfaitement avec les observations. Un moyen d'ajuster les modèles est de s'intéresser aux tourbillons. Les curvelets ne sont pas optimales pour détecter des vortex. En effet, les curvelets peuvent souligner des fronts. Mais pour les structures circulaires, il faut arriver à faire le lien entre plusieurs curvelets qui pointent par exemple vers le même centre.

Pour ce type d'applications, j'ai défini une nouvelle transformée baptisée "circlets" : les éléments de base correspondent à la convolution d'un cercle avec une ondelette [Chauris et al., 2010b, Chauris et al., 2009]. Le principe de la transformée est le même que pour les curvelets. La différence vient de la définition des filtres dans le domaine fréquentiel. Il s'avère que la transformée Fourier 2D d'une structure circulaire reste circulaire. Il est donc possible d'obtenir un pavage

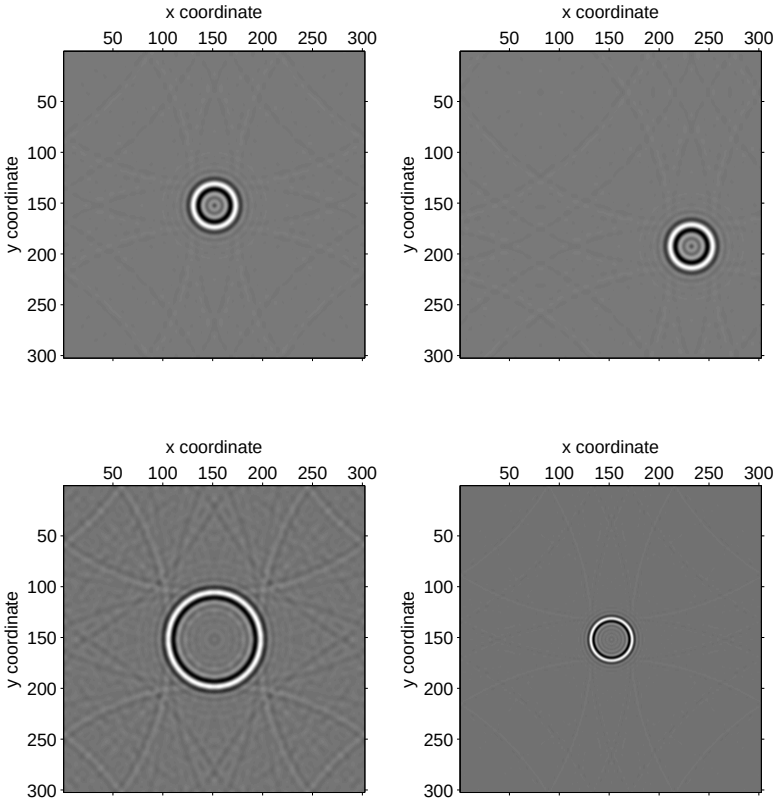


Fig. 5.1 – *Représentation d'une circlet de référence (haut à gauche), après translation (haut à droite), après changement de rayon (bas à gauche) et après changement de contenu fréquentiel (bas à droite) [Chauris et al., 2010b].*

de l'espace de Fourier. Chaque élément dans le domaine spatial a une structure circulaire. Les éléments peuvent se déduire les uns des autres, soit par translation, par changement de contenu fréquentiel ou par changement de rayon (Figure 5.1). La transformée admet une tranformée inverse sans perte d'information et est très efficace (i.e. le coût de quelques transformées de Fourier 2D). Les détails de la transformée sont données dans [Chauris et al., 2010b].

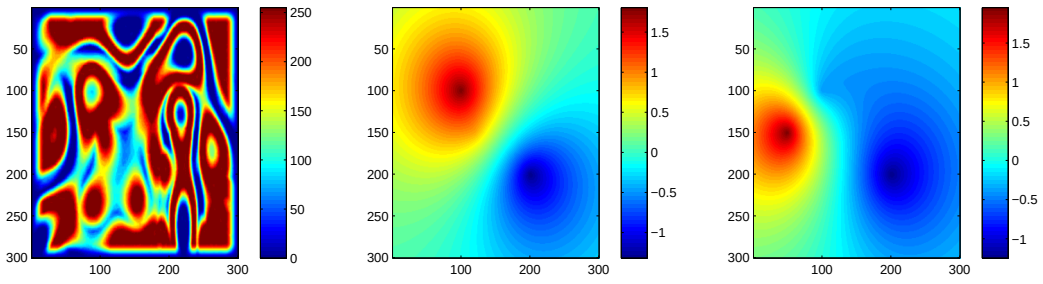


Fig. 5.2 – *Champ de températures de surface (SST, gauche), et gradients horizontaux et verticaux de déplacement des masses d'eau.*

Les premières applications en océanographie côtière donnent des résultats encourageants [Chauris et al., 2010b, Chauris et al., 2009]. Une extension de ce travail est de suivre les tourbillons au cours du temps. Deux images successives sont décomposées dans le domaine des circlets. Les coefficients les plus importants représentent des structures circulaires. Ensuite est déterminé un champ de déplacement dans le domaine des circlets pour expliquer au sens des moindres carrés la déformation des images. Ce problème inverse est résolu itérativement par une méthode de gradient (Figures 5.2 et 5.3). Pour des images qui ne différent par trop, les circlets permettent de reconstituer, en dehors des bords, le champ de déplacement des masses d'eau.

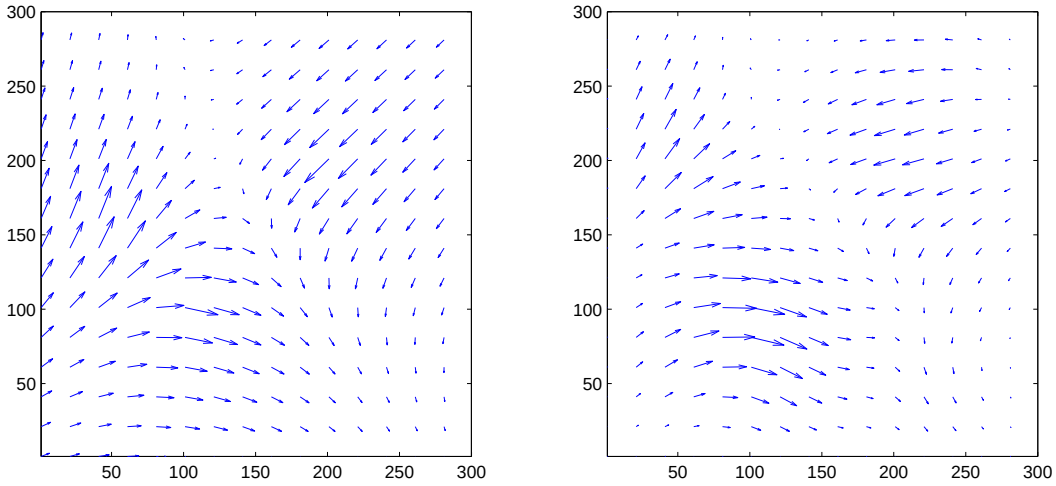


Fig. 5.3 – *Gradient exact (gauche) et après inversion (droite) à partir de la décomposition en circlets et suivi des éléments.*

En conclusion, il est possible d'imaginer de nouvelles transformées pour la résolution de problèmes spécifiques, avec ici une application en océanographie côtière.

5.2 Reverse-time migration et démigration/migration

Mes activités principales de recherche resteront encore pendant quelques années autour de l'imagerie sismique. Plus particulièrement, c'est sur le sujet de l'inversion des formes d'onde que je travaille actuellement. Je voudrais discuter de deux approches nouvelles dans ce domaine.

L'inversion des formes d'onde est une méthode générale pour retrouver les propriétés du milieu à partir d'observations [Tarantola, 1984, Tarantola, 1986, Pratt et al., 1996]. Dans la pratique, le problème est très fortement non linéaire, et il faut partir d'un modèle proche de la solution pour pouvoir converger avec une méthode de gradient [Bunks et al., 1995, Symes, 2008].

La formulation du problème en fréquence permet, avec les basses fréquences, de s'affranchir d'un certain nombre de minima secondaires [Bunks et al., 1995, Mulder and Plessix, 2008]. Il n'en demeure par moins qu'actuellement, les basses fréquences significatives enregistrées ne sont pas encore assez faibles, au moins pour les données réfléchies (Figures 5.4 et 5.5). En partant d'un résultat de tomographie de temps de trajet des premières arrivées, l'algorithme

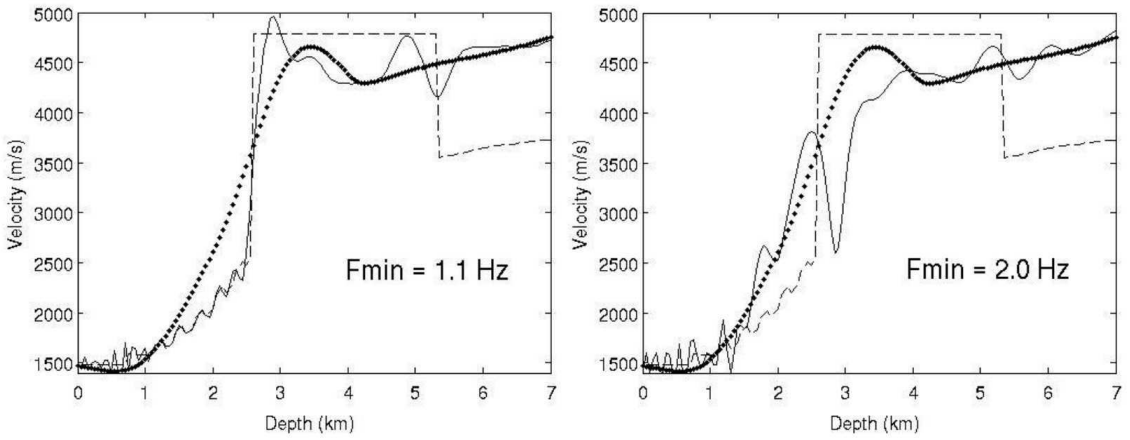


Fig. 5.4 – *Inversion 1D par formes d'onde à partir de résultats de tomographie de temps de trajet des premières arrivées [Taillandier et al., 2009] pour des contenus fréquentiels des données différents. En pointillés : le modèle exact, en pointillés gras : le modèle de départ, et en traits continus : le modèle inversé. [Chauris et al., 2008].*

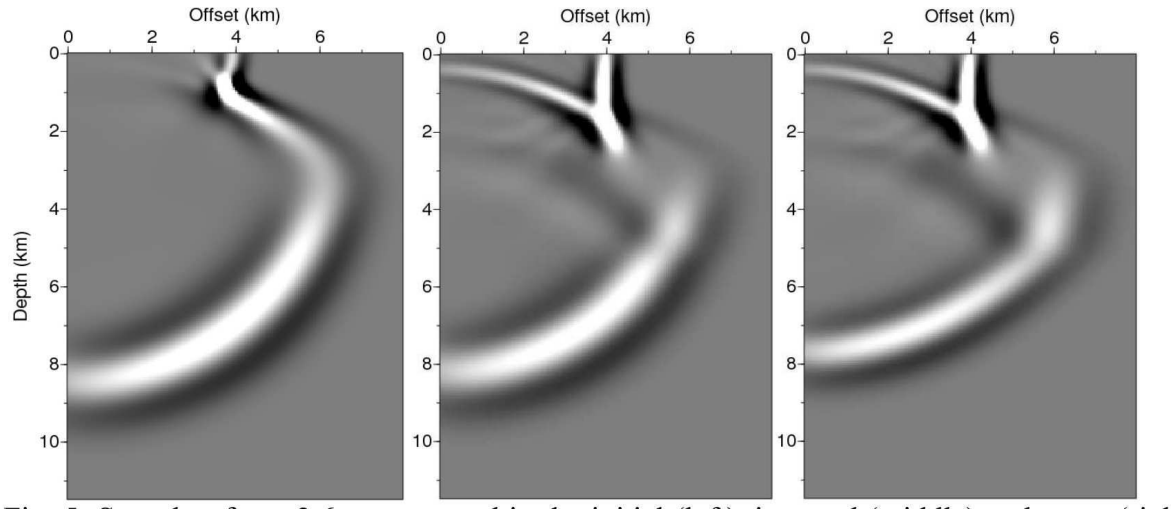


Fig. 5.5 – *Snapshots pour $t = 2.6s$ dans le modèle initial (gauche), inversé (milieu) et exact (droite) [Chauris et al., 2008].*

converge pour des fréquences minimales autour de 1.1 Hz, mais pas pour des fréquences à 2.0 Hz. L'utilisation du grand angle est un moyen d'obtenir de plus basses fréquences et donc de converger dans un certain nombre de cas [Mora, 1988, Pratt et al., 1996, Ben-Hadj-Ali et al., 2008], y compris sur données réelles, e.g. [Hicks and Pratt, 2001, Dessa et al., 2004, Ravaut et al., 2004, Plessix, 2009]. Lorsque le modèle de référence avec les grandes longueurs d'onde est fixé, l'inversion linéarisée correspond à la migration (voir par exemple [Plessix and Mulder, 2004]).

Une nouvelle approche vient directement d'une idée issue des travaux sur les curvelets : la migration ou la propagation des curvelets dans des milieux hétérogènes ne peut se résumer

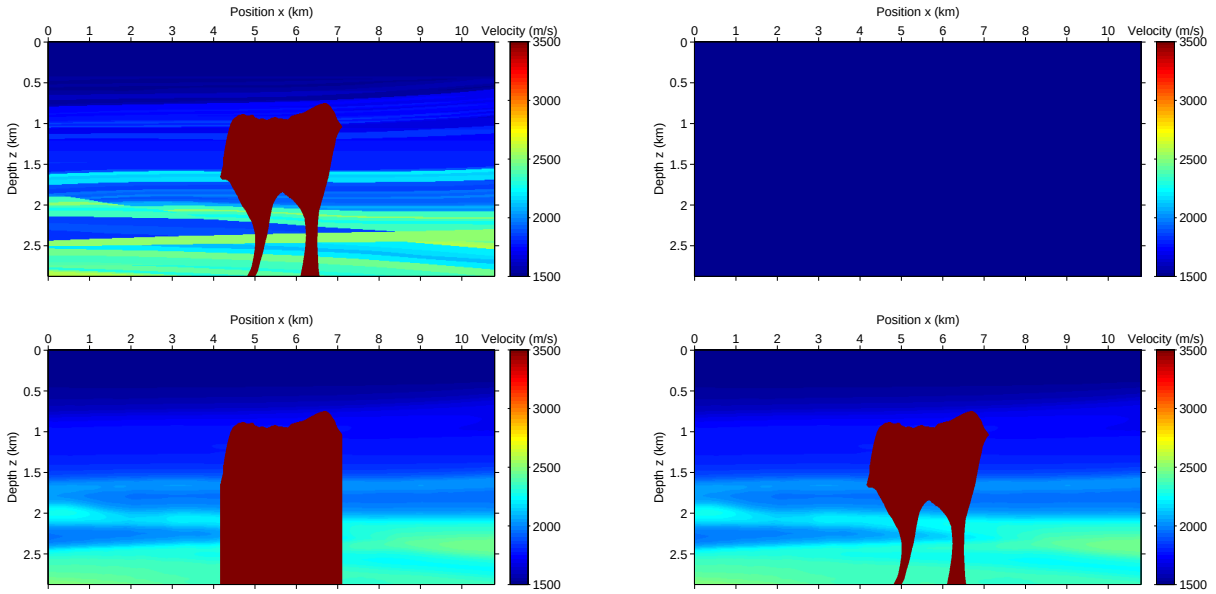


Fig. 5.6 – *Modèles exact, initial, intermédiaire et final (de gauche à droite et de haut en bas) [Chauris and Benjemaa, 2010].*

à une combinaison d'opérations élémentaires (voir chapitre 4). Par contre, la combinaison de modélisation et migration (démigration/migration) dans deux modèles de vitesse localement différents peut se simplifier avec les curvelets, au moins dans le cas de modèles lisses. Sans utiliser les curvelets, il est possible de transposer l'idée de démigration/migration dans le cas où les opérateurs sont calculés par différences finies [Chauris and Benjemaa, 2010]. Le cadre est alors celui de la "reverse-time migration". A partir d'une image de référence obtenue par migration dans un modèle de référence, d'autres images sont directement prédites pour d'autres modèles de vitesse (Figure 5.6, 5.7 et 5.8). L'idée est d'obtenir une série d'images pour différents délais en temps τ [Sava and Fomel, 2006]. Chaque image correspond à la sommation de toutes les sources et de tous les récepteurs.

$$M(x,z,\tau) = \int \mathrm{d}\omega\, e^{i\omega\tau} \int \mathrm{d}s\, G(s,x,z,\omega) \int \mathrm{d}r\, D(s,r,\omega) G(r,x,z,\omega). \tag{5.1}$$

Après simplification, la sommation sur les sources est remplacée par une sommation sur les délais en temps [Chauris and Benjemaa, 2010], ce qui réduit le temps de calcul. En dépit de cette approximation, il est possible de prédire à un coût beaucoup plus faible des images correspondant à de larges perturbations de vitesse (plusieurs kilomètres par seconde). Les réflecteurs sous le sel sont bien retrouvés (Figure 5.8, en bas), alors que dans l'image initiale, les événements étaient très peu visibles (Figure 5.7, en haut). La prédiction de l'image pour un modèle différent utilise l'information correspondant à différents délais en temps.

S'il est possible de prédire des amplitudes correctes, c'est-à-dire que pour le bon modèle de vitesse l'énergie du stack soit maximale, alors il est possible de transposer l'approche de [Shen and Symes, 2008] pour obtenir le modèle de vitesse optimal.

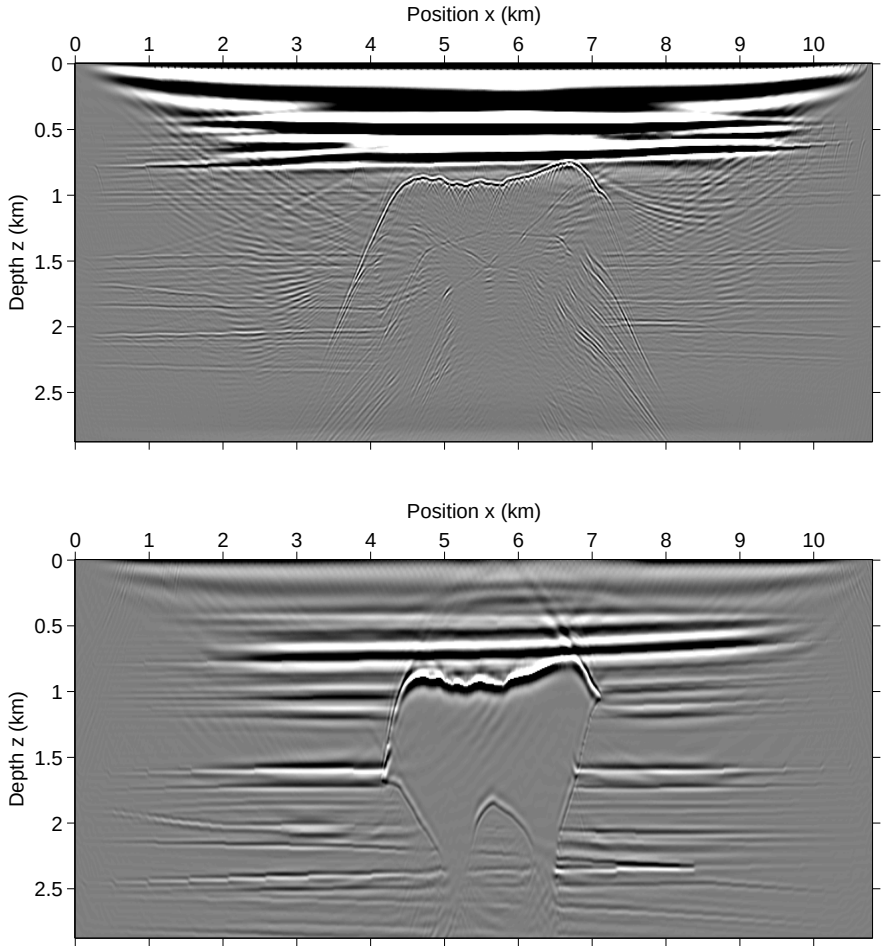


Fig. 5.7 – *Migration par équation d'onde dans le modèle initial (haut) et dans le modèle final pour les données stackées (bas) [Chauris and Benjemaa, 2010].*

$$J_{\text{DSO\_profondeur}} = \frac{1}{2} \iiint \mathrm{d}x\, \mathrm{d}z\, \mathrm{d}h\, |\tau \cdot M(x, z, \tau)|^2 \,. \tag{5.2}$$

C'est une nouvelle formulation pour l'analyse de vitesse pour laquelle la sommation sur les sources est remplacée par une sommation sur les délais en temps.

5.3 Nouvelle formulation de l'inversion des formes d'onde

Pour conclure, je propose également une nouvelle approche de l'inversion des formes d'onde qui utilise la cohérence locale des données. Cette approche se veut générale, dans le sens où les données considérées ne sont pas seulement les réflexions primaires, mais bien toutes les ondes enregistrées par les sismogrammes. La formulation classique des moindres carrés est donné par

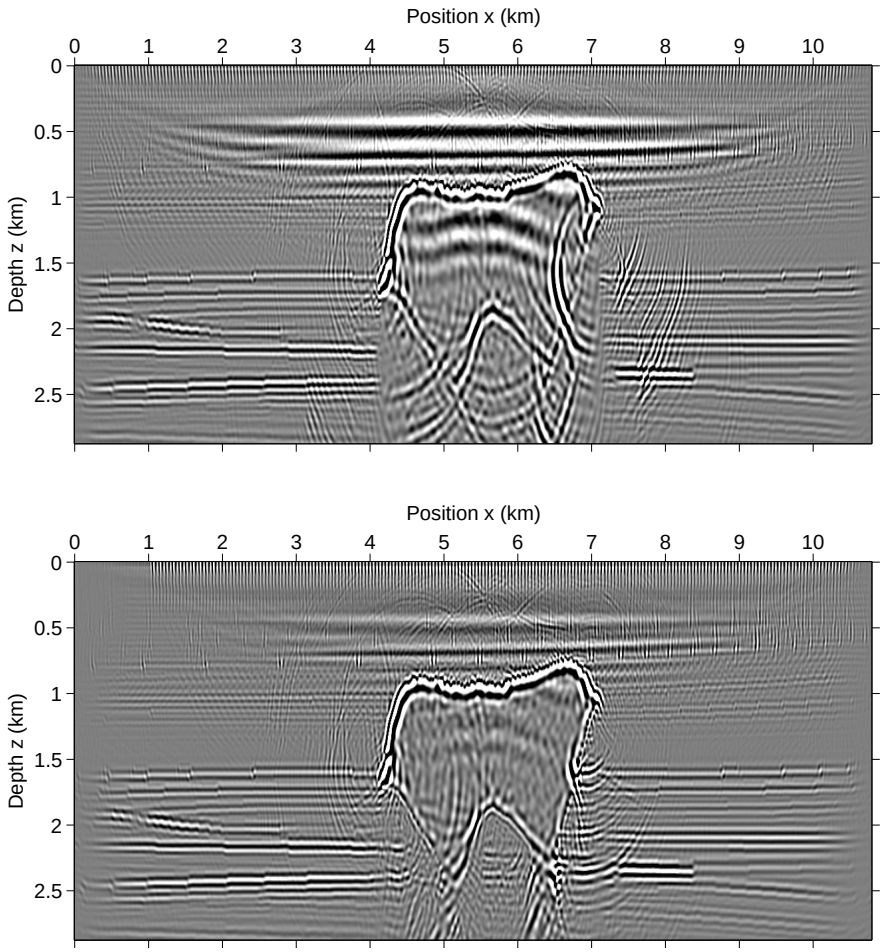


Fig. 5.8 – *Démigration/migration dans le modèle intermédiaire (haut) et dans le modèle final (bas) [Chauris and Benjemaa, 2010].*

$$J_{\mathrm{LS}}[c] = \frac{1}{2} \sum_s \parallel d_s[c] - d_s^{obs} \parallel^2, \tag{5.3}$$

où c est la modèle de Terre recherché. Comme indiqué plus haut, la fonctionnelle exhibe beaucoup de minima secondaires. La nouvelle formulation est définie en deux étapes. Pour une source donnée et un modèle de départ c_0,

$$J_1[\delta c] = \frac{1}{2} \parallel d_s[c_0 + \delta c] - d_s^{obs} \parallel^2, \tag{5.4}$$

$$J_2[c_0] = \frac{1}{2} \parallel d_{s+ds}[c_0 + \delta c(c_0, s)] - d_{s+ds}^{obs} \parallel^2 . \tag{5.5}$$

L'équation 5.4 est la même que l'approche classique, ici pour une seule source. Elle permet de trouver la meilleure perturbation de vitesse δc qui dépend de la source s et du modèle de départ c_0. Pour des données réfléchies, c'est la réflectivité, mais la même démarche prévaut pour tout

type d'ondes. Lorsque J_1 est minimisée, alors d_s^{obs} est correctement prédit dans $c_0 + \delta_c$ pour la source s considérée. La seconde équation 5.5 compare alors les données observées et les données calculées pour le point de tir suivant $s + ds$. Si les données sont correctement échantillonnées, alors, $d_{s+ds}[c_0 + \delta c(c_0, s)]$ et d_{s+ds}^{obs} diffèrent moins d'une demi-longueur d'onde et leur différence contient de l'information sur le modèle de vitesse. Potentiellement, il n'y a pas de distinction à faire entre les ondes réfléchies, réfractées, ... L'opération est répétée pour plusieurs sources.

Cette approche reste ici au stade des hypothèses, et peut être considérée comme une version différentielle de l'inversion des formes d'onde (DFWI : Differential Full WaveForm Inversion). Il faut noter que pour le calcul du gradient de l'équation 5.5 par rapport à c_0, la variable δc_0 dépend aussi de c_0, ce qui complique le calcul. La méthode de l'état adjoint [Plessix, 2006] permet de résoudre le problème (résultat non encore publié). L'approche est différente de la méthode MBTT (voir chapitre 2) pour laquelle la perturbation de vitesse δc est obtenue pour toutes les sources, et non pas par source comme ici [Plessix et al., 1995, van Leeuwen and Mulder, 2007]. C'est une différence importante si le modèle de départ n'est pas correct puisque la moyenne sur les sources de δc engendre des décalages en temps pour la partie modélisation.

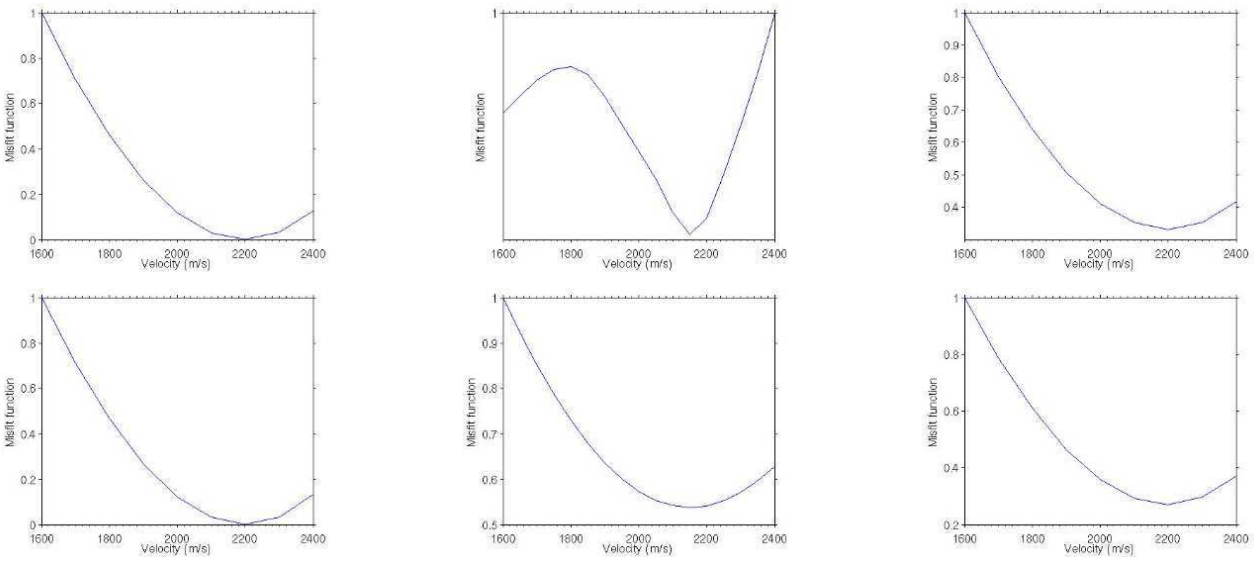


Fig. 5.9 – *Fonctions objectives classique (haut) et différentielle (bas) pour différents types de données : les premières arrivées (gauche), les réfléchies (milieu) et l'ensemble du point de tir (droite). Le modèle c_0 exact a pour valeur* 2200 *m/s (gradient vertical).*

Un premier exemple montre que l'approche est potentiellement intéressante. Dans un milieu synthétique 2D avec une réflectivité dense, les fonctions objectives classique et différentielle (équations 5.3, 5.4 et 5.5) sont calculées pour plusieurs valeurs initiales de c_0 (gradient vertical de vitesse). Trois tests sont réalisés, soit en prenant seulement le signal autour des premières arrivées, soit en ne sélectionnant que les ondes réfléchies, soit en considérant l'ensemble du point de tir (Figure 5.9). Les ondes directes sont plus fortes que les ondes réfléchies et donc les résultats de la troisième colonne sont proches de ceux de la première. Dans les trois cas, la fonction DFWI ne montre qu'un seul minimum, ce qui n'est pas le cas pour la formulation classique.

5.4 Conclusions

La prise en compte de la continuité latérale des données sismiques a permis de reconsidérer la plupart des opérations dans la chaîne du traitement sismique et de mettre en lumière des liens entre par exemple la tomographie de pente et l'analyse de vitesse sur des données migrées en profondeur. Il a été montré que dans le cadre de modèles lisses, les curvelets étaient adaptées pour obtenir une série d'images migrées à partir d'une image de référence. Il s'avère qu'il n'y a pas de transformée générale adaptée pour tous les problèmes. Si l'objectif est de débruiter les données, alors la décomposition optimale n'est pas la même que pour l'analyse de vitesse. Les curvelets au sens large offrent à la fois une flexibilité dans la décomposition des données, mais aussi une nouvelle vision des opérateurs de démigration/migration, une nouvelle formulation pour l'analyse de vitesse, et une approche multi-échelle pour par exemple le traitement des données aliasées. J'ai montré qu'il est possible de définir de nouvelles transformées, comme par exemple la transformée en circlet.

Mes nouvelles activités de recherche veulent dépasser le cadre asymptotique dans lequel les curvelets apparaissent comme naturelles. J'ai proposé deux pistes (démigration/migration avec des opérateurs calculés par différences finies et l'inversion des formes d'onde différentielle) qui, si elles semblent prometteuses, doivent être approfondies. La méthode DFWI exploite elle aussi la continuité latérale des traces sismiques, tout en sortant du cadre de la théorie hautes fréquences. L'élément clé est très certainement la notion de zone de Fresnel et de bande passante limitée du signal [Allen et al., 1995, Favretto-Cristini et al., 2009] qui font que même pour des données très complexes, la méthode DFWI pourrait s'appliquer si le signal est correctement échantillonné à la fois en temps et en espace.

Bibliographie

[Al-Yahya, 1989] Al-Yahya, K. (1989). Velocity analysis by iterative profile migration. *Geophysics*, 54 :718–729.

[Albertin et al., 2001] Albertin, U., Yingst, D., and Jaramillo, H. (2001). Comparing common-offset Maslov, Gaussian beam, and coherent state migrations. *SEG Technical Program Expanded Abstracts*, 20(1) :913–916.

[Alerini et al., 2007] Alerini, M., Lambaré, G., Baina, R., Podvin, P., and Le Bégat, S. (2007). 2D PP / PS- Stereotomography : P- and S-waves velocities estimation from OBC data. *Geophysical Journal International*, 170 :725–736.

[Allen et al., 1995] Allen, J. L., Giardinelli, T., and Reynolds, W. T. (1995). Three Fresnel zone examples. *SEG Technical Program Expanded Abstracts*, 14(1) :120–123.

[Audebert et al., 1997] Audebert, F., Diet, J. P., Guillaume, P., Jones, I., and Zhang, X. (1997). CRP scan : 3-D preSDM velocity analysis via zero offset tomographic inversion. In *Expanded Abstracts*, pages 1805–1808. Soc. Expl. Geophys.

[Baina et al., 2003] Baina, R., Nguyen, S., Noble, M., and Lambaré, G. (2003). Optimal anti-aliasing for ray-based kirchhoff depth migration. *SEG Technical Program Expanded Abstracts*, 22(1) :1130–1133.

[Ben-Hadj-Ali et al., 2008] Ben-Hadj-Ali, H., Operto, S., and Virieux, J. (2008). Velocity model building by 3d frequency-domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73(5) :VE101–VE117.

[Bermejo-Moreno and Pullin, 2008] Bermejo-Moreno, I. and Pullin, D. (2008). On the non-local geometry of turbulence. *J. Fluid Mech.*, 603 :101–135.

[Beylkin and Burridge, 1990] Beylkin, G. and Burridge, R. (1990). Linearized inverse scattering problems in acoustics and elasticity. *Wave motion*, 12 :15–52.

[Billette and Lambaré, 1998] Billette, F. and Lambaré, G. (1998). Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, 135(2) :671–680.

[Bishop et al., 1985] Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., and Spinder, D. A. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50 :903–923.

[Bunks et al., 1995] Bunks, C., Salek, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5) :1457–1473.

[Candès et al., 2006] Candès, E., Demanet, L., Donoho, D., and Ying, L. (2006). Fast discrete curvelet transform. *SIAM Multiscale Modeling and Simulation*, 5 :861–899.

[Candès and Donoho, 1999a] Candès, E. and Donoho, D. (1999a). Curvelets – a surprisingly effective nonadaptive representation for objects with edges. *in Curves and Surface Fitting : Saint-Malo, A. Cohen, C. Rabut, L. Schumaker (Eds), Vanderbilt Univ. Press Nashville*, pages 105–120.

[Candès and Donoho, 1999b] Candès, E. and Donoho, D. (1999b). Ridgelets : a key to higher-dimensional intermittency ? *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357 :2495–2509.

[Candès and Donoho, 2004] Candès, E. and Donoho, D. (2004). New tight frames of curvelets and optimal representations of objects with C^2 singularities. *Comm. on Pure and Appl. Math.*, 57 :219–266.

[Cao and Wu, 2009] Cao, J. and Wu, R.-S. (2009). Fast acquisition aperture correction in prestack depth migration using beamlet decomposition. *Geophysics*, 74(4) :S67–S74.

[Chauris, 2006] Chauris, H. (2006). Seimic imaging in the curvelet domain and its implications for the curvelet design. In *Expanded Abstracts, 76th Annual SEG Meeting*, pages 2406–2410. Society of Exploration Geophysicists.

[Chauris and Benjemaa, 2010] Chauris, H. and Benjemaa, M. (2010). Seismic wave-equation demigration/migration. *Accepted for publication to Geophysics.*

[Chauris et al., 2010a] Chauris, H., Donno, D., Nguyen, T., and Ma, J. (2010a). Modeling and imaging seismic data with curvelets : a review. *Submitted to Geophysics.*

[Chauris et al., 2010b] Chauris, H., Karoui, I., Garreau, P., Wackernagel, H., Craneguy, P., and Bertino, L. (2010b). The circlet transform : a new tool for ocean eddy tracking. *Submitted to Computers and Geosciences.*

[Chauris et al., 2009] Chauris, H., Karoui, I., Wackernagel, H., Garreau, P., Craneguy, P., and Bertino, L. (2009). Ocean eddy tracking with circlets. In *GeoInformatics for Environmental Surveillance (Milos, Greece)*. StaGIS.

[Chauris et al., 1998] Chauris, H., Le Rousseau, J., Beaudoin, B., Propson, P., and Montanari, A. (1998). Inoceramid acme/extinction in the gubbio region (Northeastern Apennines of Italy) and relations with environmental changes during the mid-Maastrichtian. *Paleogeography, Paleoclimatology, Paleoocology*, 139 :177–193.

[Chauris and Ma, 2009] Chauris, H. and Ma, J. (2009). Seismic imaging in the curvelet domain : achievements and perspectives. In *Expanded Abstracts, 71st Annual EAGE Meeting, Workshop 1 (Amsterdam)*. European Association of Geoscientists and Engineers.

[Chauris and Nguyen, 2008a] Chauris, H. and Nguyen, T. (2008a). Seismic demigration/migration in the curvelet domain. *Geophysics*, 73(2) :S35–S46.

[Chauris and Nguyen, 2008b] Chauris, H. and Nguyen, T. (2008b). Seismic velocity estimation in the curvelet domain. In *Expanded Abstracts, 70th Annual EAGE Meeting (Rome).* European Association of Geoscientists and Engineers.

[Chauris and Noble, 2001] Chauris, H. and Noble, M. (2001). Two-dimensional velocity macro model estimation from seismic reflection data by local Differential Semblance Optimization : applications on synthetic and real data sets. *Geophysical Journal International*, 144 :14–26.

[Chauris et al., 2002a] Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002a). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I : theoretical aspects. *Geophysics*, 67 :1202–1212.

[Chauris et al., 2002b] Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002b). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II : application on synthetic and real data. *Geophysics*, 67 :1213–1224.

[Chauris et al., 2008] Chauris, H., Noble, M., and Taillandier, C. (2008). What initial velocity model do we need for full waveform inversion ? In *Expanded Abstracts, 70th Annual EAGE Meeting (Rome), Workshop*. European Association of Geoscientists and Engineers.

[Chauris and Salomons, 2004] Chauris, H. and Salomons, B. (2004). Data-driven Kirchhoff migration or how to take advantage of the slant stack domain. In *Expanded Abstracts, 66th Annual EAGE Meeting (Paris)*. European Association of Geoscientists and Engineers.

[Chen et al., 2006] Chen, L., Wu, R.-S., and Chen, Y. (2006). Target-oriented beamlet migration based on gabor-daubechies frame decomposition. *Geophysics*, 71(2) :S37–S52.

[Clément, 1994] Clément, F. (1994). *Une formulation en temps de parcours par migration pour la détermination des vitesses de propagation acoustique à partir de données sismiques bidimensionnelles*. PhD thesis, Université Paris IX Dauphine.

[Daubechies, 1992] Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.

[Demanet and Ying, 2007] Demanet, L. and Ying, L. (2007). Wave atoms and sparsity of oscillatory patterns. *Appl. Comput. Harmon. Anal.*, 23(3) :368–387.

[Deregowski, 1990] Deregowski, S. M. (1990). Common-offset migrations and velocity analysis. *First Break*, 8 :225–234.

[Desmares et al., 2007] Desmares, D., Grosheny, D., Beaudoin, B., Gardin, S., and Gauthier-Lafaye, F. (2007). High resolution stratigraphic record constrained by volcanic ashes layers at the cenomanian-turonian boundary in the western interior basin, usa. *Cretaceous Research*, 28 :561–582.

[Dessa et al., 2004] Dessa, J.-X., Operto, S., Kodaira, S., Nakanishi, A., Pascal, G., Virieux, J., and Kaneda, Y. (2004). Multiscale seismic imaging of the Eastern Nankai trough by full waveform inversion. *Geophysysical Research Letters*, 31 :L18606.

[Do, 2001] Do, M. N. (2001). *Directional multiresolution image representations*. PhD thesis, Swiss Federal Institute of Technology Lausanne.

[Do and Vetterli, 2003] Do, M. N. and Vetterli, M. (2003). The finite ridgelet transform for image representation. *IEEE Transactions on Image Processing*, 12 :16–28.

[Docherty et al., 1997] Docherty, P., Silva, R., Singh, S., Song, Z., and Wood, M. (1997). Migration velocity analysis using a genetic algorithm. *Geophysical Prospecting*, 45 :865–878.

[Donno et al., 2010] Donno, D., Chauris, H., and Noble, M. (2010). Curvelet-based multiple prediction. *Submitted to Geophysics*.

[Donoho, 1999] Donoho, D. (1999). Wedgelets : nearly minimax estimation of edges. *Ann. Statistics*, 27(3) :859–897.

[Douma and de Hoop, 2006] Douma, H. and de Hoop, M. V. (2006). Leading-order seismic imaging using curvelets. *SEG Technical Program Expanded Abstracts*, 25(1) :2411–2415.

[Douma and de Hoop, 2007] Douma, H. and de Hoop, M. V. (2007). Leading-order seismic imaging using curvelets. *Geophysics*, 72(6) :S231–S248.

[Eckhardt, 1994] Eckhardt, W. (1994). Velocity model updating using image gathers. *Geophysical Prospecting*, 42 :975–986.

[Farra and Madariaga, 1987] Farra, V. and Madariaga, R. (1987). Seismic waveform modeling in heterogeneous media by ray perturbation theory. *J. Geophys. Res.*, 92 :2697–2712.

[Favretto-Cristini et al., 2009] Favretto-Cristini, N., Cristini, P., and de Bazelaire, E. (2009). What is a seismic reflector like ? *Geophysics*, 74(1) :T13–T23.

[Fomel, 2006] Fomel, S. (2006). Towards the seislet transform. *SEG Technical Program Expanded Abstracts*, 25(1) :2847–2851.

[Freeman and Adelson, 1991] Freeman, W. and Adelson, E. (1991). The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9) :891–906.

[Gauthier et al., 1986] Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms : Numerical results. *Geophysics*, 51(7) :1387–1403.

[Gockenback and Symes, 1995] Gockenback, M. C. and Symes, W. W. (1995). Waveform inversion for velocity : Where have the minima gone ? In *Expanded Abstracts*, pages 1095–1098. Soc. Expl. Geophys.

[Guillaume et al., 2008] Guillaume, P., Lambaré, G., Leblanc, O., Mitouard, P., Le Moigne, J., Montel, J.-P., Prescott, T., Siliqi, R., Vidal, N., Zhang, X., and Zimine, S. (2008). Kinematic invariants : an efficient and flexible approach for velocity model building. *SEG Technical Program Expanded Abstracts*, 27(1) :3687–3692.

[Guo and Labate, 2007] Guo, K. and Labate, D. (2007). Optimally sparse multidimensional representation using shearlets. *SIAM J. Math. Anal.*, 39 :298–318.

[Hagen and Kvalheim, 1992] Hagen, J. and Kvalheim, B. (1992). *Giant Oil and Gas fields of the decade 1978-1988*, volume 54, chapter 26, pages 417–428. AAPG memoir.

[Hennenfent and Herrmann, 2008] Hennenfent, G. and Herrmann, F. J. (2008). Simply denoise : Wavefield reconstruction via jittered undersampling. *Geophysics*, 73(3) :V19–V28.

[Herrmann et al., 2009a] Herrmann, F. J., Brown, C. R., Erlangga, Y. A., and Moghaddam, P. P. (2009a). Curvelet-based migration preconditioning and scaling. *Geophysics*, 74(4) :A41–A46.

[Herrmann et al., 2009b] Herrmann, F. J., Erlangga, Y. A., and Lin, T. T. Y. (2009b). Compressive simultaneous full-waveform simulation. *Geophysics*, 74(4) :A35–A40.

[Herrmann et al., 2008a] Herrmann, F. J., Wang, D., Hennenfent, G., and Moghaddam, P. P. (2008a). Curvelet-based seismic data processing : A multiscale and nonlinear approach. *Geophysics*, 73(1) :A1–A5.

[Herrmann et al., 2008b] Herrmann, F. J., Wang, D., and Verschuur, D. J. (2008b). Adaptive curvelet-domain primary-multiple separation. *Geophysics*, 73(3) :A17–A21.

[Hicks and Pratt, 2001] Hicks, G. J. and Pratt, R. G. (2001). Reflection waveform inversion using local descent methods : Estimating attenuation and velocity over a gas-sand deposit. *Geophysics*, 66(2) :598–612.

[Hill, 2001] Hill, R. (2001). Prestack Gaussian-beam depth migration. *Geophysics*, 66(4) :1240–1250.

[Jannane et al., 1989] Jannane, M., Beydoun, W., Crase, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S., Snieder, R., Tarantola, A., and Trezeguet, D. (1989). Wavelengths of Earth structures that can be resolved from seismic reflection data. *Geophysics*, 54(7) :906–910.

[Jilek et al., 2007] Jilek, P., Brandesberg-Dahl, S., Zhou, M., and Albertin, U. (2007). Beam steering in ray-based beam and wave-equation migration. *SEG Technical Program Expanded Abstracts*, 26(1) :2205–2209.

[Jin and Beydoun, 2000] Jin, S. and Beydoun, W. (2000). 2D multiscale non-linear velocity estimation. *Geophysical Prospecting*, 48 :163–180.

[Jin and Madariaga, 1993] Jin, S. and Madariaga, R. (1993). Background velocity inversion with a genetic algorithm. *Geophysical research Letters*, 20(2) :93–96.

[Jin and Madariaga, 1994] Jin, S. and Madariaga, R. (1994). Nonlinear velocity inversion by a two-step Monte Carlo. *Geophysics*, 59(4) :577–590.

[Jin et al., 1992] Jin, S., Madariaga, R., Virieux, J., and Lambaré, G. (1992). Two-dimensional asymptotic iterative elastic inversion. *Geophysical Journal International*, 108 :575–588.

[Johnstad et al., 1995] Johnstad, S. E., Seymour, R. H., and Smith, P. J. (1995). Seismic reservoir monitoring over the Oseberg field during the period 1989-1992. *First Break*, 13(5) :169–183.

[Lafond and Levander, 1993] Lafond, C. F. and Levander, A. R. (1993). Migration moveout analysis and depth focusing. *Geophysics*, 58(1) :582–598.

[Lailly and Sinoquet, 1996] Lailly, P. and Sinoquet, D. (1996). Smooth velocity models in reflection tomography for imaging complex geological structures. *Geophys. J. Int.*, 124 :349–362.

[Lambaré, 2008] Lambaré, G. (2008). Stereotomography. *Geophysics*, 73(5) :VE25–VE34.

[Lambaré et al., 1992] Lambaré, G., Virieux, J., Madariaga, R., and Jin, S. (1992). Iterative asymptotic inversion in the acoustic approximation. *Geophysics*, 57 :1138–1154.

[Le Bégat et al., 2004] Le Bégat, S., Chauris, H., Devaux, V., Nguyen, S., and Noble, M. (2004). Comparison of three tomography methods on a 2-D real data set. *Geophysical Prospecting*, 52(5) :427–438.

[Le Pennec and Mallat, 2005] Le Pennec, E. and Mallat, S. (2005). Sparse geometrical image approximation with bandlets. *IEEE Trans. Image Process.*, 14(4) :423–438.

[Le Rousseau et al., 1996] Le Rousseau, J., Chauris, H., Montanari, A., and Beaudoin, B. (1996). Modifications environnementales Ã la limite K/T dans le bassin profond des Apennins du NE (Italie). In *Cretaceous-Tertiary boundary : biological and geological aspects.* Congrès de la Société Géologique de France (Paris).

[Lee, 2008] Lee, T. (2008). Image representation using 2D Gabor wavelets. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(10) :1–13.

[Lin and Herrmann, 2007] Lin, T. T. Y. and Herrmann, F. J. (2007). Compressed wavefield extrapolation. *Geophysics*, 72(5) :SM77–SM93.

[Liu, 1997] Liu, Z. (1997). An analytical approach to migration velocity analysis. *Geophysics*, 62 :1238–1249.

[Liu and Bleistein, 1995] Liu, Z. and Bleistein, N. (1995). Migration velocity analysis : Theory and an iterative algorithm. *Geophysics*, 60 :142–153.

[Lu and Do, 2007] Lu, Y. and Do, M. N. (2007). Multidimensional directional filter banks and surfacelets. *IEEE Transactions on Image Processing*, 16(4) :918–931.

[Ma et al., 2006] Ma, J., Antoniadis, A., and Le Diment, F.-X. (2006). Curvelet-based snake for multiscale detection and tracking of geophysical fluids. *IEEE transactions on geoscience and remote sensing*, 44(12) :3626–3638.

[Ma and Hussaini, 2007] Ma, J. and Hussaini, M. (2007). Three-dimensional curvelets for coherent vortex analysis of turbulence. *Appl. Phys. Lett.*, 91(18) :184140 :1–3.

[Ma and Plonka, 2009] Ma, J. and Plonka, G. (2009). A review of curvelets and recent applications. *Submitted to IEEE Signal Processing Magazine*.

[Ma et al., 2010] Ma, J., Plonka, G., and Chauris, H. (2010). A new sparse representation of seismic data using adaptive easy-path wavelet transform. *Accepted for publication to Geoscience and Remote Sensing Letters*.

[Mallat, 1989] Mallat, S. (1989). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Pattern Anal. and Machine Intell.*, 11 :674–693.

[Mallat and Peyré, 2007] Mallat, S. and Peyré, G. (2007). A review of bandlet methods for geometrical image representation. *Numerical Algorithms*, 44(3) :205–234.

[Meng et al., 1999a] Meng, Z., Bleistein, N., and Valasek, P. (1999a). 3-D analytical migration velocity analysis II : velocity gradient estimation. In *Expanded Abstracts*, pages 1731–1734. Soc. Expl. Geophys.

[Meng et al., 1999b] Meng, Z., Bleistein, N., and Wyatt, K. (1999b). 3-D analytical migration velocity analysis I : two-step velocity estimation by reflector-normal update. In *Expanded Abstracts*, pages 1727–1730. Soc. Expl. Geophys.

[Meyer, 1993] Meyer, Y. (1993). *Wavelets : algorithms and applications*. SIAM.

[Mora, 1988] Mora, P. (1988). Elastic wave-field inversion of reflection and transmission data. *Geophysics*, 53(6) :750–759.

[Mulder and Plessix, 2008] Mulder, W. A. and Plessix, R.-E. (2008). Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting*, pages 827–841.

[Mulder and ten Kroode, 2002] Mulder, W. A. and ten Kroode, A. P. E. (2002). Automatic velocity analysis by Differential Semblance Optimization. *Geophysics*, 67(4) :1184–1191.

[Neelamani et al., 2008] Neelamani, R., Baumstein, A., Gillard, D., Hadidi, M., and Soroka, W. (2008). Coherent and random noise attenuation using the curvelet transform. *The Leading Edge*, 27(2) :240–248.

[Nguyen et al., 2008] Nguyen, S., Baina, R., Alerini, M., Lambaré, G., Devaux, V., and Noble, M. (2008). Slope tomography assisted by migration of attributes. *Geophysical Prospecting*, 56(5) :613–625.

[Nguyen et al., 2009a] Nguyen, S., Baina, R., Alerini, M., Lambaré, G., Devaux, V., and Noble, M. (2009a). Slope tomography assisted by migration of attributes. *Submitted to Geophysical Prospecting*.

[Nguyen and Chauris, 2010] Nguyen, T. and Chauris, H. (2010). The Uniform Discrete Curvelet Transform. *Submitted to IEEE Transactions on Signal Processing*.

[Nguyen et al., 2009b] Nguyen, T., Chauris, H., and Lambaré, G. (2009b). Time demigration/migration using curvelets. In *Expanded Abstracts, 71st Annual EAGE Meeting (Amsterdam)*. European Association of Geoscientists and Engineers.

[Nguyen and Oraintara, 2006] Nguyen, T. T. and Oraintara, S. (2006). On the aliasing effect of the contourlet filter banks. *European Signal Processing Conference EUSIPCO (Italy)*.

[Operto et al., 2000] Operto, S., Xu, S., and Lambaré, G. (2000). Can we quantitatively image complex structures with rays ? *Geophysics*, 65(4) :1223–1238.

[Ottolini and Claerbout, 1984] Ottolini, R. and Claerbout, J. F. (1984). The migration of common midpoint slant stacks. *Geophysics*, 49(3) :237–249.

[Pica et al., 1990] Pica, A., Diet, J. P., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, 55(3) :284–292.

[Plessix, 1996] Plessix, R.-E. (1996). *Détermination de la vitesse pour l'interprétation de données sismiques très haute résolution à l'échelle géotechnique*. PhD thesis, Université Paris IX Dauphine.

[Plessix, 2006] Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167 :495–503.

[Plessix, 2009] Plessix, R.-E. (2009). 3-D frequency-domain full-waveform inversion with an iterative solver. *Geophysics, in press.*

[Plessix et al., 1995] Plessix, R.-E., Chavent, G., and De Roeck, Y. (1995). Automatic and simultaneous migration velocity analysis and waveform inversion of real data using a MBTT/WBKBJ formulation. In *Expanded Abstracts*, pages 1099–1101. Soc. Expl. Geophys.

[Plessix and Mulder, 2004] Plessix, R.-E. and Mulder, W. A. (2004). Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157 :975–987.

[Plessix et al., 2000] Plessix, R.-E., Mulder, W. A., and ten Kroode, A. P. E. (2000). Automatic cross-well tomography by semblance and differential semblance optimization : theory and gradient computation. *Geophysical Prospecting*, 48 :913–935.

[Pratt et al., 1996] Pratt, R. G., Song, Z. M., and Warner, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124 :323–340.

[Ravaut et al., 2004] Ravaut, C., abd S[] Importa, S. O., Virieux, J., A., A. H., and dell' Aversana, P. (2004). Multi-scale imaging of complex structures from multi-fold wide-aperture seismic data by frequency-domain full-waveform inversions : application to a thrust belt. *Geophysical Journal International*, 159 :1032–1056.

[Sava and Fomel, 2006] Sava, P. and Fomel, S. (2006). Time-shift imaging condition in seismic migration. *Geophysics*, 71(6) :S209–S217.

[Shan et al., 2009] Shan, H., Ma, J., and Yang, H. (2009). Comparisons of wavelets, contourlets and curvelets in seismic denoising. *Journal of Applied Geophysics*, 69(2) :103–115.

[Shen and Symes, 2008] Shen, P. and Symes, W. W. (2008). Automatic velocity analysis via shot profile migration. *Geophysics*, 73(5) :VE49–VE59.

[Simoncelli et al., 1992] Simoncelli, E., Freeman, W., Adelson, E., and Heeger, D. (1992). Shiftable multiscale transforms. *IEEE Trans. Inform. Theory*, 38(2) :587–607.

[Soubaras and Gratacos, 2007] Soubaras, R. and Gratacos, B. (2007). Velocity model building by semblance maximization of modulated-shot gathers. *Geophysics*, 72(5) :U67–U73.

[Stolk and Symes, 2004] Stolk, C. C. and Symes, W. W. (2004). Kinematic artifacts in prestack depth migration. *Geophysics*, 69(2) :562–575.

[Sun et al., 2009] Sun, B., Ma, J., Chauris, H., and Yang, H. (2009). Solving wave equations in the curvelet domain : a multi-scale and multi-directional approach. *Journal of Seismic Exploration*, 18(4) :385–399.

[Symes, 1993] Symes, W. W. (1993). A differential semblance criterion for inversion of multioffset seismic reflection data. *Journal of Geophysical Research*, 98 :2061–2073.

[Symes, 1998a] Symes, W. W. (1998a). All stationary points of differential semblance are asymtotic global minimizers : layered acoustics. Technical report, The Rice Inversion Project.

[Symes, 1998b] Symes, W. W. (1998b). High frequency asymptotics, differential semblance, and velocity estimation. In *Expanded Abstracts*, pages 1616–1619. Society of Exploration Geophysicists.

[Symes, 2008] Symes, W. W. (2008). Migration velocity analysis and waveform inversion. *Geophysicsal Prospecting*, 56 :765–790.

[Symes and Carazzone, 1991] Symes, W. W. and Carazzone, J. J. (1991). Velocity inversion by Differential Semblance Optimization. *Geophysics*, 56 :654–663.

[Symes and Kern, 1994] Symes, W. W. and Kern, M. (1994). Inversion of reflection seismograms by differential semblance analysis : algorithm structure and synthetic examples. *Geophysical Prospecting*, 42 :565–614.

[Taillandier et al., 2009] Taillandier, C., Noble, M., Chauris, H., and Calandra, H. (2009). First-arrival travel time tomography based on the adjoint state method. *Geophysics*, 74(6) :WCB57–WCB66.

[Tang et al., 2009] Tang, G., Shahidi, R., Ma, J., and Herrmann, F. (2009). Design of two-dimensional randomized sampling schemes for curvelet-based sparsity-promoting seismic data recovery. *submitted to Geophysical Prospecting*.

[Tarantola, 1984] Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8) :1259–1266.

[Tarantola, 1986] Tarantola, A. (1986). A strategy for non linear inversion of seismic reflection data. *Geophysics*, 51(10) :1893–1903.

[van Leeuwen and Mulder, 2007] van Leeuwen, T. and Mulder, W. A. (2007). Data-correlation for velocity analysis. In *Expanded Abstracts, 77th Annual SEG Meeting*, pages 1800–1804. Society of Exploration Geophysicists.

[Verschuur et al., 1992] Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A. (1992). Adaptive surface-related multiple elimination. *Geophysics*, 57(9) :1166–1177.

[Versteeg and Grau, 1991] Versteeg, R. J. and Grau, G., editors (1991). *The Marmousi experience*. Proceedings of the 1990 EAEG workshop on Pratical Aspects of Seismic Data Inversion, Eur. Ass. Expl. Geophys.

[Wang et al., 1995] Wang, B., Pann, K., and Meek, R. A. (1995). Macro velocity model estimation through model-based globally-optimized residual-curvature analysis. In *Expanded Abstracts*, pages 1084–1087. Soc. Expl. Geophys.

[Willet and Novak, 2003] Willet, R. and Novak, K. (2003). Platelets : a multiscale approach for recovering edges and surfaces in photon-limited medical imaging. *IEEE Trans. Med. Imaging*, 22(3) :332–350.

[Wu and Chen, 2006] Wu, R.-S. and Chen, L. (2006). Directional illumination analysis using beamlet decomposition and propagation. *Geophysics*, 71(4) :S147–S159.

[Wu et al., 2008] Wu, R.-S., Wang, Y., and Luo, M. (2008). Beamlet migration using local cosine basis. *Geophysics*, 73(5) :S207–S217.

[Xu et al., 2001] Xu, S., Chauris, H., Lambaré, G., and Noble, M. (2001). Common angle image gather : a strategy for imaging complex media. *Geophysics*, 66(6) :1877–1894.

[Xu and Lambaré, 2004] Xu, S. and Lambaré, G. (2004). Fast migration/inversion with multivalued ray fields : Part 1—method, validation test, and application in 2D to Marmousi. *Geophysics*, 69(5) :1311–1319.

[Zacek, 2003] Zacek, K. (2003). Decomposition of the wave field into optimized gaussian packets. *SEG Technical Program Expanded Abstracts*, 22(1) :1869–1872.

[Zacek, 2005] Zacek, K. (2005). Gaussian packet pre-stack depth migration of the marmousi data set. *SEG Technical Program Expanded Abstracts*, 24(1) :1822–1825.

Table des figures

Chapitre 6

Annexes

Une copie des cinq articles suivants a été mise en annexe :

- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002a). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I : theoretical aspects. *Geophysics*, 67 :1202–1212.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002b). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II : application on synthetic and real data. *Geophysics*, 67 :1213–1224.
- Chauris, H. and Nguyen, T. (2008). Seismic demigration/migration in the curvelet domain. *Geophysics*, 73(2) :S35–S46.
- Nguyen, T. and Chauris, H. (2009). The Uniform Discrete Curvelet Transform. *Submitted to IEEE Transactions on Signal Processing*.
- Chauris, H. and Benjemaa, M. (2009). Seismic wave-equation demigration/migration. *Accepted for publication to Geophysics*.

GEOPHYSICS, VOL. 67, NO. 4 (JULY-AUGUST 2002); P. 1202–1212, 12 FIGS.
10.1190/1.1500382

Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects

Hervé Chauris*, Mark S. Noble‡, Gilles Lambaré‡, and Pascal Podvin‡

ABSTRACT

We present a new method based on migration velocity analysis (MVA) to estimate 2-D velocity models from seismic reflection data with no assumption on reflector geometry or the background velocity field. Classical approaches using picking on common image gathers (CIGs) must consider continuous events over the whole panel. This interpretive step may be difficult—particularly for applications on real data sets. We propose to overcome the limiting factor by considering locally coherent events. A locally coherent event can be defined whenever the imaged reflectivity locally shows lateral coherency at some location in the image cube.

In the prestack depth-migrated volume obtained for an a priori velocity model, locally coherent events are picked automatically, without interpretation, and are characterized by their positions and slopes (tangent to the event). Even a single locally coherent event has information on the unknown velocity model, carried by the value of the slope measured in the CIG. The velocity is estimated by minimizing these slopes.

We first introduce the cost function and explain its physical meaning. The theoretical developments lead to two equivalent expressions of the cost function: one formulated in the depth-migrated domain on locally coherent events in CIGs and the other in the time domain. We thus establish direct links between different methods devoted to velocity estimation: migration velocity analysis using locally coherent events and slope tomography.

We finally explain how to compute the gradient of the cost function using paraxial ray tracing to update the velocity model. Our method provides smooth, inverted velocity models consistent with Kirchhoff-type migration schemes and requires neither the introduction of interfaces nor the interpretation of continuous events. As for most automatic velocity analysis methods, careful preprocessing must be applied to remove coherent noise such as multiples.

INTRODUCTION

We address the problem of 2-D estimation of the background or velocity macromodel from seismic reflection data by migration velocity analysis (MVA). This approach of velocity estimation basically uses the flatness of events in common image gathers (CIGs) as a criterion for velocity quality (Al-Yahya, 1989) (Figure 1). Most approaches that do not require picking or any assumption on the velocity field require a global optimization process to minimize or maximize a cost function defined on CIGs (Stoffa, 1989; Jin and Madariaga, 1993, 1994; Jervis et al., 1996; Docherty et al., 1997; Varela et al., 1998; Jin and Beydoun, 2000). No feasible 3-D extensions exist at the moment for data sets of realistic size. Only two methods seem to converge with a local approach: migration-based traveltime inversion (Clément, 1994; Plessix, 1996) and differential semblance optimization (Symes and Carazzone, 1991; Symes, 1993, 1998; Chauris and Noble, 1998, 2001). But even in two dimensions, these approaches remain expensive and are probably not suitable for 3-D applications.

When picking is introduced in the depth-migrated domain, approaches become tractable in three dimensions. Many developments have been proposed. However, the updating formulas to invert the velocity model are generally based on at least one of the three following simplifying assumptions: laterally invariant velocity, small offset, or horizontal reflectors. The earliest MVA methods are based on all these simplifications (Al-Yahya, 1989). Many improvements have been proposed, but all require some simplifying assumptions (Deregowski, 1990; Cox and Wapenaar, 1992; Lee and Zhang, 1992; Lafond and Levander, 1993; Liu and Bleistein, 1995; Wang et al., 1995;

Manuscript received by the Editor June 12, 2000; revised manuscript received July 26, 2001.
*Formerly Ecole des Mines de Paris, Centre de Géophysique 35, rue Saint Honoré, 77305 Fontainebleau Cédex, France; presently Shell International, Volmerlaan 8, P.O. Box 60, 2280 AB Rijswijk, The Netherlands. E-mail: h.chauris@siep.shell.com.
‡Ecole des Mines de Paris, Centre de Géophysique 35, rue Saint Honoré, 77305 Fontainebleau Cédex, France. E-mail: mark.noble@geophy.ensmp.fr; gilles.lambare@geophy.ensmp.fr; pascal.podvin@geophy.ensmp.fr.

Audebert et al., 1997; Woodward et al., 1998). In most cases, picking is introduced to select the maximum of the stack power along predefined (offset, depth) curves. Only Liu's (1997) approach, valid for any 2-D velocity field, relates the perturbations of the reflector depth to the perturbations of the velocity model. In his method, the macromodel is described by velocities and interfaces.

For most approaches proposed in the literature, picking is performed only on continuous events that must be tracked continuously over a large offset range, as illustrated in Figure 2,

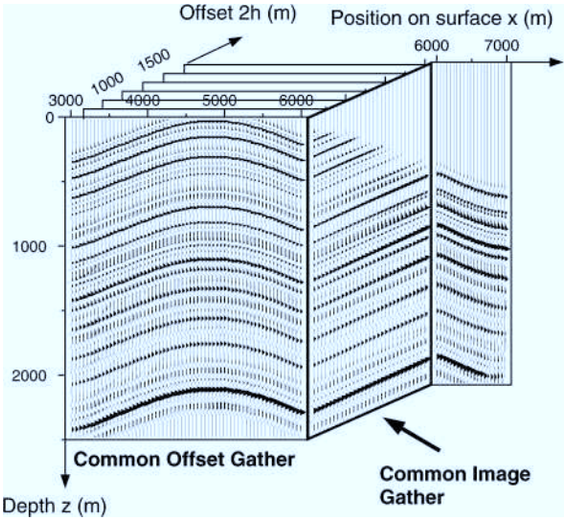


FIG. 1. Definition and use of a CIG. For 2-D data, the migrated volume is obtained by prestack depth migration of each time common-offset section with the same given velocity model. A CIG is a section in the cube, defined by a fixed position on the surface. If the exact velocity model is used for migration, then the CIG should present flat events, independent of the structure of the reflectors in the background. Events in the data are indeed replaced after migration in their exact location in the depth domain, which does not depend on acquisition parameters such as offset.

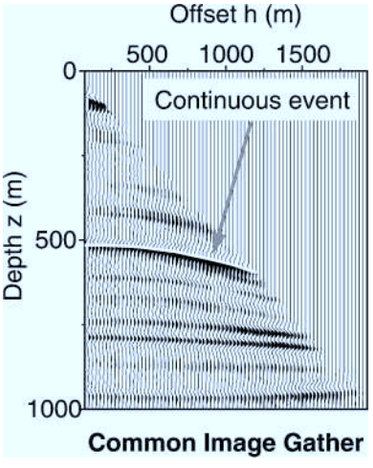


FIG. 2. Continuous event over all offsets, indicated by the white solid line. Such events are selected for inversion in the classical approaches, as in Liu (1997), where depth is picked at some sparse locations along each selected event.

and attached to a particular reflector in the model. This interpretive step is certainly the main limiting factor of present methods. It is impossible to follow continuous events in a CIG in many cases, e.g., when the tested velocity model is far from exact or in the presence of noise [see examples of a synthetic case (Figure 3) and a real case (Figure 4)].

We present a new approach valid for any 2-D velocity field and based on picking locally coherent events (Figure 5) in the depth-migrated domain. As a basic definition, a locally coherent event is defined when the reflectivity displays lateral coherency over a few neighboring traces.

Introducing locally coherent events is motivated by two reasons. From a practical point of view, it is far easier to pick many locally coherent events than to follow continuous events (Figures 3 and 4). From a theoretical point of view, velocity inversion does not require continuous events because a single locally coherent event bears information on the unknown velocity model. The initial idea is that the slope of a single locally coherent event in a CIG provides a constraint on the unknown velocity field. It should be null when the velocity field is correct.

We propose estimating the velocity field by flattening a set of locally coherent events treated as uncorrelated events (i.e., not attached to a particular reflector). Because our approach does not tie picked events to interfaces, the macromodel can be parameterized with smooth, nonblocky basis functions. A blocky parameterization could naturally be considered but is not required.

For our purpose, the main difficulty is to design a way to update the velocity model. We are thus interested in the perturbations of the depth of an event but more generally want to learn how the event is distorted (displacement and rotation) when velocity changes.

We first define a locally coherent event in a depth-migrated domain. We then explain why energy focuses in the depth-migrated domain (i.e., where it comes from) and derive a general link between the time and depth domains. These relations help specify the cost function. We finally show how to update the velocity model from the information contained in these locally coherent events using paraxial ray tracing.

Our only assumption is that the data contain only primary reflections/diffractions. In particular, we assume that careful preprocessing has removed multiples, refracted waves, and other coherent noise. This is the main limitation of the method because it is entirely automatic, including the picking step. But this difficulty is not specific to our method and is encountered by most velocity analysis processes. Apart from this assumption, we want our theory to be valid for any 2-D velocity field and an arbitrary distribution of reflectors in the subsurface, provided it is correctly sampled in the data.

DEFINING A LOCALLY COHERENT EVENT IN THE PRESTACK DEPTH-MIGRATED DOMAIN

Let us define what we call a locally coherent event in the depth-migrated volume. For a 2-D data set, migration is performed in a given velocity model, producing, in the depth domain, CIGs and common-offset gathers (COGs) (Figure 1). Local coherence is sought simultaneously around a given location (x, h) in the two sections (x, z) and (h, z) of the migrated

volume (x, z, h), where h is the half-offset. A locally coherent picked event is thus characterized by five parameters (x, z, ξ, φ, h) (Figure 5), where (x, z) is the event location in depth, ξ is its apparent geological migrated dip (as measured in the COG), and φ is its residual slope (as measured in the CIG). By definition, for each locally coherent event migrated in a given slowness 2-D model u,

$$\tan\varphi = \left.\frac{\partial z}{\partial h}\right|_{x,u}, \tag{1}$$

$$\tan\xi = \left.\frac{\partial z}{\partial x}\right|_{h,u}. \tag{2}$$

For the exact velocity model, the CIGs should present flat events (Al-Yahya, 1989). Mathematically, φ should be zero after convergence so that we use a least-squares misfit function of the form

$$J[u] = \frac{1}{2}\sum_{picks}(w\tan\varphi)^2, \tag{3}$$

where w is a weighting coefficient still to be defined. After convergence, (x, z) represents the actual location of the reflection/diffraction point in depth and ξ is the real geologic dip of the corresponding reflector. Angle φ is basically needed to determine the cost function, whereas ξ is needed to compute its gradient, as explained below.

We want to give evidence that the definition of our cost function has potentially good properties in terms of process optimization. Through the generic cost function [equation (3)], our method can be directly related with differential semblance optimization (DSO) (Symes and Carazzone, 1991; Symes, 1993, 1998; Chauris and Noble, 1998; Chauris and Noble, 2001). The associated differential semblance function J_{DS} also quantifies the flatness of seismic events in CIGs and is based on horizontal derivatives of these panels, i.e.,

$$J_{DS}[u] = \frac{1}{2}\sum_{selected\ x}\iint dz\,dh\left[\frac{\partial R}{\partial h}(x, z, h, u)\right]^2, \tag{4}$$

where $R(x, z, h, u)$ is the value (amplitude) of the prestack depth-migrated image at point (x, z, h) migrated with the slowness model u (Figure 1).

Let us consider a zoom on a migrated CIG (as in Figure 6b) with small dimensions but containing a nonzero reflectivity. The amplitude of the central trace h_0 inside the rectangle may be written as $A\omega[k(z - z_0)]$, where ω is a wavelet displayed around depth z_0, A is the amplitude of the reflectivity, and k is a stretching factor. Because we assume that the box is small enough, the other traces can be deduced from the central trace by translation as follows:

$$R(x, z, h, u) \sim A\omega[k(z - z_0 + p\cdot(h - h_0))], \tag{5}$$

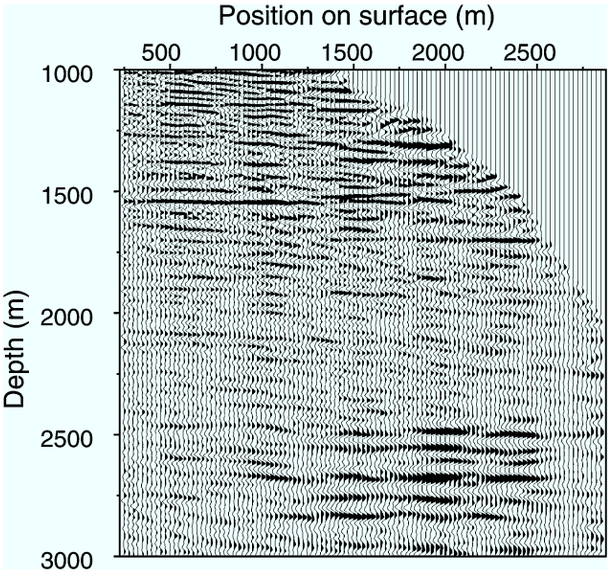


FIG. 4. CIG obtained on a real data set. It is clearly easier to pick many events with local coherency (for example, around depths 2100 or 2500 m) than to follow continuous events over all offsets, as needed in many approaches.

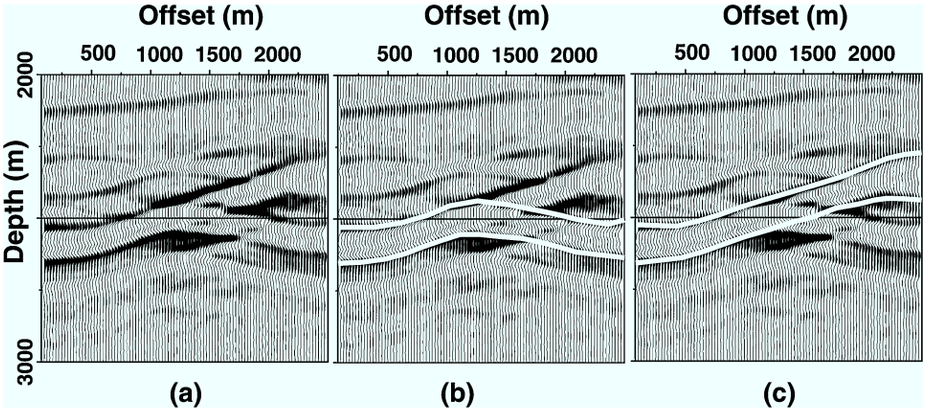


FIG. 3. (a) Even on synthetic data, it is not always easy to follow continuous events over all offsets in a CIG. (b, c) Two interpretations are proposed, indicated by the white solid curve. This effect appears, for example, when the migration velocity is not correct, even when the data only contain primary reflections/diffractions (this case) because they were generated by ray + Born approximation (Lambaré et al., 1996).

where $p = \tan\varphi$ is the slope defining how the depth of the locally coherent event changes with offset. The derivative with respect to the offset gives

$$\left.\frac{\partial R(x,z,h,u)}{\partial h}\right|_{x,z,u} \sim p \cdot kA\omega'[k(z - z_0 + p \cdot (h - h_0))]. \quad (6)$$

From that expression, it becomes clear that minimizing J_{DS} is equivalent to minimizing the residual slope $p = \tan\varphi$ in the CIG as in equation (3).

Symes (1999) demonstrates in the 1-D case that the J_{DS} cost function has a unique minimum. Thus, a gradient-type algorithm may be applied to converge to the (unique) solution. This property has not yet been proven in two (or three) dimensions. However, successful applications of DSO in two dimensions with synthetic and real data sets (Chauris and Noble, 2001) give practical evidence that the J_{DS} cost function remains well behaved.

Because of the formal analogy between J_{DS} and our generic cost function, we consider a reasonable conjecture that our problem can converge with a gradient-type optimization process.

In the following section, we demonstrate that, by establishing a link between locally coherent events in depth and time, we can choose the weighting function w in equation (3) such that the gradient of the cost function can be explicitly computed with the help of paraxial ray tracing done in the tested velocity model, whatever its complexity.

COST FUNCTION

This section is devoted to the appropriate development of the expression of the cost function. We need to understand why energy focuses in the depth-migrated domain. For this purpose, we derive the relationship between the two picked angles in the depth domain and the slopes in the time data (traveltime derivatives in the seismograms). For simplicity and clarity, all calculations are developed in the common-shot domain and are then extended to the common-offset case, which generally is preferred for applications. Because mathematical computations are rather involved, we only mention the crucial points and the final results and try to provide a physical understanding of the different steps. More details are presented in Appendix A. To derive these calculations, we must understand which parameters must be kept fixed and which can be variable.

Focusing energy in the depth-migrated domain (common-shot case)

To better understand the information contained in the residual slope $\tan\varphi$ measured in the CIG (Figure 5), we have to know why energy focuses at a certain position (x, z, h) in the depth-migrated domain. For that purpose, we need to establish the relationship between the time and depth domains in a general 2-D case (Figure 6).

To introduce the formalism we need to derive these relations, we temporarily switch to another domain. The energy focused in the image cube results from focusing the seismic events recorded in the data cube (seismograms) (s, r, t^*), where s and r denote shot and receiver positions and t^* is the recording time.

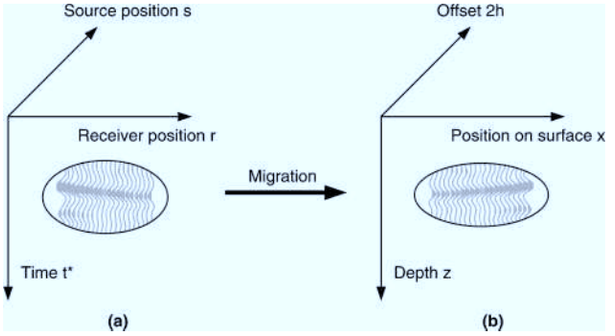


FIG. 6. Definition of (a) time and (b) depth-migrated domains. The aim is (1) to understand why a locally coherent event focuses in the depth domain and then (2) to minimize the residual slope in the CIG.

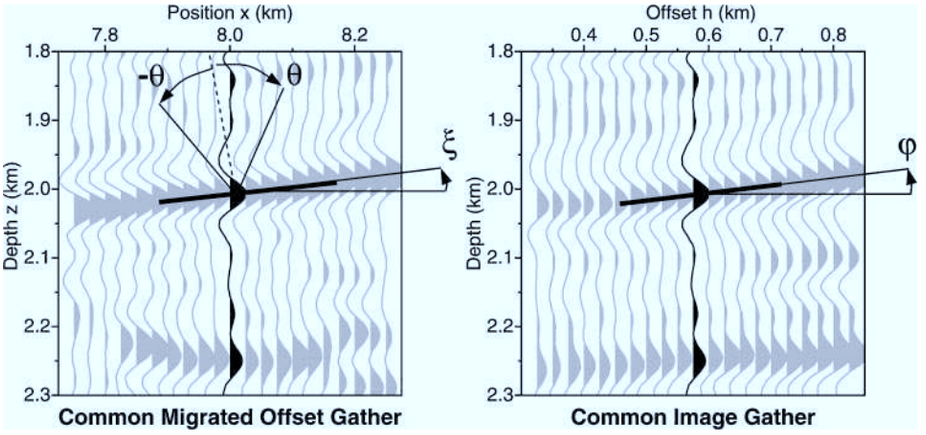


FIG. 5. The five values characterizing a locally coherent event in the 2-D prestack depth-migrated domain (x, z, ξ, φ, h). Migration at common offset is performed to obtain the COG (left) and CIG (right), where picking is performed simultaneously around the same location (x, h). The half-aperture angle θ is defined such that the two rays reach the given offset $2h$ on the surface, starting symmetrically around the normal of the dip.

In this cube, it is possible to define, by analogy, locally coherent events in time. Such an event is a wave packet displaying local coherency near the location (s, r, t^*) in the cube. Once more, defining the local coherency amounts to measuring slopes—for example, the derivatives

$$p_s^* = \left.\frac{\partial t^*}{\partial s}\right|_r, \tag{7}$$

$$p_r^* = \left.\frac{\partial t^*}{\partial r}\right|_s \tag{8}$$

of the event traveltime with respect to shot and receiver locations (Figure 7). This event is described by five parameters $(s, r, t^*, p_s^*, p_r^*)$, analogous to (x, h, z, φ, ξ). Parameters describing the event in time are objective, i.e., experimental. On the contrary, (x, h, z, φ, ξ) describing the counterpart in the image cube depend on the velocity model used for migration. In the following equations, we emphasize objective parameters with a star (e.g., p_s^*).

Image focusing is best understood in the framework of Kirchhoff migration (Schneider, 1978; Bleistein et al., 1987; Audebert et al., 1997). We first cover the case of common-shot migration. Consider a locally coherent event picked in the common-shot gather for source location s. In Figure 7, the event is characterized by the parameters (r, t^*, p_r^*). The parameter p_s^* is not accessible; it can only be measured using several shot gathers simultaneously. In common-shot migration, this event focuses somewhere along the diffraction curve, namely, the isochron (Figure 8a), implicitly defined by

$$t_s(s, x, z, u) + t_r(x, z, r, u) = t^*(s, r), \tag{9}$$

where t_s and t_r are (model-dependent) one-way traveltimes from s and r to the image point (x, z). Energy builds up constructively at one location, provided it also belongs to the envelope of the isochron (i.e., its derivative with respect to receiver position). The two focusing equations for the selected event (r, t^*, p_r^*) thus read

$$t_s(s, x, z, u) + t_r(x, z, r, u) = t^*(s, r), \tag{10}$$

$$\left.\frac{\partial t_r}{\partial r}\right|_{x,z,u} = \left.\frac{\partial t^*}{\partial r}\right|_s = p_r^*. \tag{11}$$

We now must understand what $(\partial t_r/\partial r)|_{x,z,u}$ stands for in terms of modeling. At location (x, z) on the isochron, consider the ray shot toward the receiver in the velocity model used for migration (Figure 8b). This ray reaches the surface at location r and at time t_r by construction. The second focusing equation (11) states that the event in time will focus in depth at location (x, z) such that the ray emerges at the surface with a slowness vector $\mathbf{p}_r = \nabla t$ whose horizontal component p_{rx} exactly matches the slopes p_r^* of the event in the data (Figure 8b), i.e.,

$$p_{rx} = p_r^*. \tag{12}$$

In this construction, the second slope p_s^* of the event in the data cube has not been considered. Conversely, the migrated event location and dip (x, z, ξ) are fixed by the focusing equations, but $\tan\varphi$ in the CIG remains underdetermined (because we only migrated one shot gather).

Since angle φ is measured in a CIG (i.e., at constant x), we must understand how the focusing equations change for

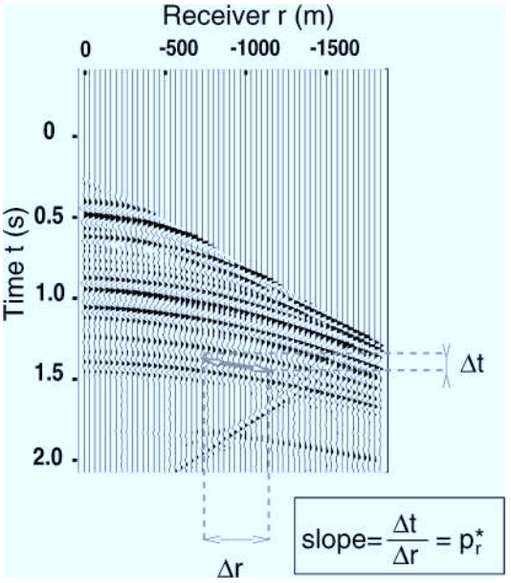


FIG. 7. Locally coherent event in the time domain. In the data, the slope is the tangent of a locally coherent reflector, measured here in a (time) common-shot gather, thus providing a slope at the receiver position.

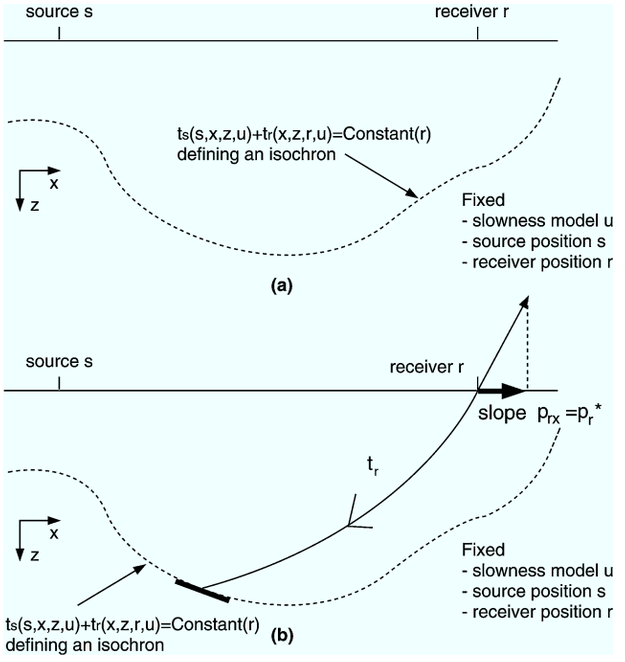


FIG. 8. Focusing a locally coherent event in the depth-migrated domain. (a) The isochron is defined in the depth domain by a constant two-way traveltime for fixed velocity-model, source, and receiver positions. (b) For migration at a constant source position, the ray is entirely defined by the position of the receiver and the slope on the surface. This horizontal component of the slowness vector corresponds to the slope in the time data (see text for explanations) as defined in Figure 7. The locally coherent event in the time domain focuses in the depth domain at the intersection of the isochron and the ray.

different s values at constant x position and constant velocity model u. All other parameters (z, s, r) in the focusing equations must be allowed to vary in the general case. Remember that $p_{rx} = p_{rx}(x, z, r, u)$ and $p_r^* = p_r^*(s, r)$. By differentiating the focusing equations (10) and (12), we obtain the following system:

$$\left(\left.\frac{\partial t_s}{\partial z}\right|_{s,x,u} + \left.\frac{\partial t_r}{\partial z}\right|_{x,r,u}\right)\delta z + \left.\frac{\partial t_s}{\partial s}\right|_{x,z,u}\delta s + \left.\frac{\partial t_r}{\partial r}\right|_{x,z,u}\delta r = p_s^*\delta s + p_r^*\delta r, \tag{13}$$

$$\left.\frac{\partial p_{rx}}{\partial z}\right|_{r,x,u}\delta z + 0\cdot\delta s + \left.\frac{\partial p_{rx}}{\partial r}\right|_{x,z,u}\delta r = \left.\frac{\partial p_r^*}{\partial s}\right|_{r}\delta s + \left.\frac{\partial p_r^*}{\partial r}\right|_{s}\delta r. \tag{14}$$

Using equation (11), the last term on the left side of equation (13) is balanced by the second term on the right side, providing a new system of linear equations:

$$\left(\left.\frac{\partial t_s}{\partial z}\right|_{s,x,u} + \left.\frac{\partial t_r}{\partial z}\right|_{r,x,u}\right)\delta z + p_{sx}\delta s = p_s^*\delta s, \tag{15}$$

$$\left.\frac{\partial p_{rx}}{\partial z}\right|_{r,x,u}\delta z + \left.\frac{\partial p_{rx}}{\partial r}\right|_{x,z,u}\delta r = \left.\frac{\partial p_r^*}{\partial s}\right|_{r}\delta s + \left.\frac{\partial p_r^*}{\partial r}\right|_{s}\delta r, \tag{16}$$

where

$$\left.\frac{\partial p_r^*}{\partial s}\right|_{r} = \left.\frac{\partial^2 t^*}{\partial s^2}\right|_{r} \quad \text{and} \quad \left.\frac{\partial p_r^*}{\partial r}\right|_{s} = \left.\frac{\partial^2 t^*}{\partial r^2}\right|_{s}$$

are directly linked to the curvature of the locally coherent event in the time domain. Fortunately, they do not appear in the expression needed for $z(s)$ that we are interested in. Indeed, the definition from equation (1) leads to

$$\tan\varphi = \left.\frac{\partial z}{\partial s}\right|_{x,u} = \frac{p_s^* - p_{sx}}{\left.\dfrac{\partial t_s}{\partial z}\right|_{s,x,u} + \left.\dfrac{\partial t_r}{\partial z}\right|_{r,x,u}}. \tag{17}$$

The denominator is computed in Appendix A, and we finally obtain

$$\tan\varphi = \frac{p_s^* - p_{sx}}{2u\cos\theta\cos\xi}, \tag{18}$$

where u is the value of the slowness at the scattering point (x, z). We define $\alpha = 2u\cos\theta\cos\xi$. This term equals zero for vertical dips and direct (transmitted) arrivals that are generally not taken into account during migration. It is related to the stretching factor occurring in migration (Tygel et al., 1994).

To summarize, we obtain two equations valid for any 2-D velocity field and any reflector geometry:

$$p_r^* - p_{rx} = 0, \tag{19}$$

$$p_s^* - p_{sx} = \alpha\tan\varphi. \tag{20}$$

Thus, $p_s^* = p_{sx}$ for the exact velocity field (i.e., $\tan\varphi = 0$). Velocity estimation is equivalent to adjusting the horizontal slowness provided by ray tracing to the actual slopes of the seismic event in the (time) data space. This important point is discussed further below. In spite of their very simple form, equations (19) and (20) have been established for the general case. In particular, no hypothesis on the curvature of the event is necessary.

Common-offset case

The approach is very similar to the common-shot case. We define the half-offset $h = (s - r)/2$ and the midpoint $m = (s + r)/2$. For simplicity, we introduce the slope components in the relevant (h, m) domain,

$$p_{hx} = \frac{p_{sx} - p_{rx}}{2}, \tag{21}$$

$$p_{mx} = \frac{p_{sx} + p_{rx}}{2}, \tag{22}$$

and the equivalent equations for the slopes measured in the time data,

$$p_h^* = \frac{p_s^* - p_r^*}{2}, \tag{23}$$

$$p_m^* = \frac{p_s^* + p_r^*}{2}. \tag{24}$$

During the migration and for each offset the summation is performed over all midpoints m. By replacing the roles of s and r by the variables h and m, we obtain equations similar to the common-shot case:

$$t_s(s, x, z, u) + t_r(x, z, r, u) = t^*(h, m), \tag{25}$$

$$\left.\frac{\partial(t_s + t_r)}{\partial m}\right|_{x,z,h,u} = \left.\frac{\partial t^*}{\partial m}\right|_{h}. \tag{26}$$

Equations (25) and (26) indicate where energy focuses in the depth-migrated domain. To obtain the relation equivalent to equation (18), we must differentiate equations (25) and (26) with respect to z, h, and m for constant x and u. Using exactly the same approach as developed for the common-shot case, we finally obtain

$$p_m^* - p_{mx} = 0, \tag{27}$$

$$p_h^* - p_{hx} = \frac{\alpha}{2}\tan\varphi, \tag{28}$$

or, equivalently,

$$(p_s^* - p_{sx}) + (p_r^* - p_{rx}) = 0, \tag{29}$$

$$(p_s^* - p_{sx}) - (p_r^* - p_{rx}) = \alpha\tan\varphi. \tag{30}$$

Focusing relation (29) states that the sum of the two horizontal slowness components (at source and receiver) is constant and independent of the velocity model used for migration. Conversely, the difference between the slopes is directly related to the slope measured in the CIG [equation (30)].

Cost functions

To more easily compute the gradient of the cost function, we introduce $w = \alpha = 2u\cos\theta\cos\xi$ as the weighting term to apply in equation (3). Our cost function thus reads in the common-shot case

$$J[u] = \frac{1}{2}\sum_{picks}[\alpha\tan\varphi]^2 \tag{31}$$

$$= \frac{1}{2}\sum_{picks}(p_{sx} - p_s^*)^2. \tag{32}$$

To calculate the cost function in a given velocity model, let's use the picked data in the migrated domain and relation (31). Equation (32) is only used to compute the gradient of the cost function because p_s^* does not depend on the velocity chosen for migration.

The common-offset case is very similar to shot profile migration. Once again, a convenient cost function can be designed from relation (28):

$$J[u] = \frac{1}{2}\sum_{picks}[\alpha \tan\varphi]^2 \tag{33}$$

$$= 2\sum_{picks}(p_h^* - p_{hx})^2 = 2\sum_{picks}(p_{sx} - p_s^*)^2$$

$$= 2\sum_{picks}(p_{rx} - p_r^*)^2, \tag{34}$$

where the formulations in equation (34) are obtained using equations (29) and (30).

Whatever the migration scheme (common shot or common offset), the criterion compares computed slopes and observed slopes, although no slope is picked in the time data. We now illustrate this point in the following section.

Time versus depth domain for velocity estimation

Different but equivalent expressions have been obtained for the cost function for both the common-shot case with equations (31) and (32) and the common-offset case with equations (33) and (34). Minimizing residual slopes in the CIGs or fitting computed slopes to observed slopes in the time data is equivalent. These relations establish direct links between different methods devoted to velocity estimation, namely, our method, working in the depth-migrated domain on locally coherent events (Stork, 1992; Wang et al., 1995; Woodward et al., 1998), and slope tomography methods, also using locally coherent events but in the time domain (Sword, 1987; Biondi, 1992; Billette and Lambaré, 1998).

With slope tomography methods, velocity is estimated from traveltimes and slopes picked in the time domain on locally coherent events. With stereotomography, a general slope tomography method (Billette and Lambaré, 1998), the model is described by the velocity model itself and scattering points. These points are characterized by their position, dip, and corresponding aperture angles (Figure 9). For each picked event, data (two-way traveltime, positions, and slopes at the source and receiver) are calculated by shooting rays from the scattering point to the surface. The associated cost function is the difference between observed and calculated data. It can be expressed as a weighted summation:

$$J_{stereo}[u] = \frac{1}{2}\sum_{picks}\big[w_{s_x}(s_x - s^*)^2 + w_{s_z}(s_z - 0)^2$$
$$+ w_{r_x}(r_x - r^*)^2 + w_{r_z}(r_z - 0)^2$$
$$+ w_{ps}(p_{sx} - p_s^*)^2 + w_{pr}(p_{rx} - p_r^*)^2$$
$$+ w_t(t_s + t_r - t^*)^2\big]. \tag{35}$$

The inversion consists of simultaneously retrieving the ray segments $(x, z, \theta, \xi, t_s, t_r)$ and the slowness model u. In MVA, the positions are obtained automatically by migration (even for a wrong velocity model). Thus, the total traveltime and the positions of the source and receiver associated with a ray segment match by construction. For common-shot migration, we established in equation (12) that $p_{rx} = p_r^*$. The cost function thus becomes

$$J_{stereo}[u] = \frac{1}{2}\sum_{picks} w_{ps}(p_{sx} - p_s^*)^2, \tag{36}$$

as proposed in equation (32).

These relations let us unify totally different approaches for velocity estimation: MVA in the depth domain and slope tomography in the time domain. However, their implementations are rather different. For example, the coverage of the picked events in the depth domain is more uniform for the MVA method; picking should be easier because it is done in the migrated domain. Since we have defined in detail the cost function that we adopt here and have shown the physical meaning of the residual slope in the CIG, we can now explain how to update the velocity model.

COST FUNCTION GRADIENT

An easy way to obtain the gradient of the cost function could be to compute it by finite differences, requiring at least as many migrations as the number of parameters describing the velocity model. However, this solution is in efficient. We therefore develop an explicit formulation of the expression of the gradient valid for any 2-D velocity fields. Like the cost function, we first present the common-shot case.

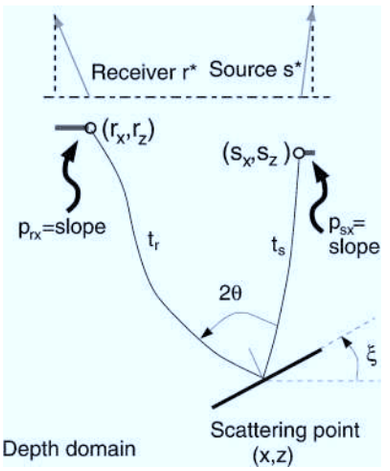


FIG. 9. In stereotomography (Billette and Lambaré, 1998), a general slope tomography method, the observed data consist of $(s^*, r^*, t^*, p_s^*, p_r^*)$, picked on seismograms. The model is given by the velocity model itself and the positions of the scattering points, defined by $(x, z, t_s, t_r, \theta, \xi)$. The cost function compares the characterized values at the extremities of the two ray segments from the scattering points to the data. In our method, the positions of the ray segment are automatically retrieved by migration, i.e., $s_x = s^*$, $s_z = 0$, $r_x = r^*$, $r_z = 0$, $t_s + t_r = t^*$. For migration at a common source, we also have $p_{rx} = p_r^*$. Only the slope at the source position may differ from the observed slope, providing a simpler cost function given by equation (36).

Common-shot case

The computation is performed for a given velocity model, which can be far from the exact velocity field. We have to understand how to relate the perturbations of the positions (x, z) and angles (ξ, θ, φ) to perturbations of the velocity model. Since p_s^* is independent of the tested velocity field, the gradient of the cost function can be formulated as

$$\frac{\partial J[u]}{\partial u} = \sum_{picks}(p_{sx} - p_s^*)\frac{\partial p_{sx}}{\partial u}\bigg|_{s,r,t,p_{rx}} \tag{37}$$

$$= -\sum_{picks}\alpha \tan\varphi \frac{\partial p_{sx}}{\partial u}\bigg|_{s,r,t,p_{rx}}. \tag{38}$$

Equation (38) is obtained using equation (20). In a given velocity model, we have to understand how an elementary part of the image—namely, a locally coherent event in the depth domain—is distorted when a perturbation of the slowness field is added to the model (Figure 10). All variables in the depth domain—position x, depth z, and scattering angles $\theta_s = \xi + \theta$ and $\theta_r = \xi - \theta$—depend on the tested velocity model. The perturbations δx, δz, $\delta\theta_s$, and $\delta\theta_r$ are constrained by two conditions at the surface, which are illustrated in Figure 10:

1) The locally coherent event is the same in the time domain, meaning that the rays shot toward the surface reach exactly the same source and the same receiver with the same double traveltime: $ds = 0$, $dr = 0$, and $dt = d(t_s + t_r) = 0$.
2) The ray toward the receiver reaches the surface with exactly the same horizontal slowness vector: $dp_{rx} = 0$ [focusing condition (12)]. Indeed, p_s^* does not depend on the velocity model chosen for migration and thus p_{rx} is constant.

In a given velocity model, the perturbations of the final conditions ds, dr, dp_{sx}, dp_{rx}, and dt can be expressed using the paraxial ray theory (Farra and Madariaga, 1987) from the perturbations of initial conditions δx, δz, $\delta\theta_s$, and $\delta\theta_r$ and the slowness field δu as explained in Appendix B. Solving this linear system, we finally obtain for the Fréchet derivatives

$$\frac{\partial p_{sx}}{\partial u}\bigg|_{s,r,t,p_{rx}} = \frac{\det\left|\dfrac{\partial(p_{sx}, s, r, t, p_{rx})}{\partial(u, x, z, \theta_s, \theta_r)}\right|}{\det\left|\dfrac{\partial(s, r, t, p_{rx})}{\partial(x, z, \theta_s, \theta_r)}\right|_u\Bigg|}. \tag{39}$$

Conditions where the gradient of the cost function becomes infinite are discussed later.

Common-offset case

Analogous to the development of the cost function, the common-offset case is very similar to the common-shot approach. The gradient of the cost function with respect to velocity can be expressed as

$$\frac{\partial J[u]}{\partial u} = -\frac{1}{2}\sum_{picks}\alpha \tan\varphi \frac{\partial p_{hx}}{\partial u}\bigg|_{h,m,t,p_{mx}}. \tag{40}$$

The gradient is obtained as before with two conditions, giving $\delta\theta_s$, $\delta\theta_r$, δx, and δz (Figure 11):

1) The locally coherent event is the same in the time domain, meaning that the rays toward the surface reach exactly the same offset and the same midpoint with the same double traveltime: $dh = 0$, $dm = 0$, and $dt = 0$ (or, equivalently, $ds = 0$, $dr = 0$, and $dt = 0$).
2) The summation of the horizontal slowness vectors at the surface is constant: $dp_{sx} + dp_{rx} = 2dp_{mx} = 0$ [focusing condition (27)].

The first two conditions are the same as in the common-shot case (same locally coherent event in the time data).

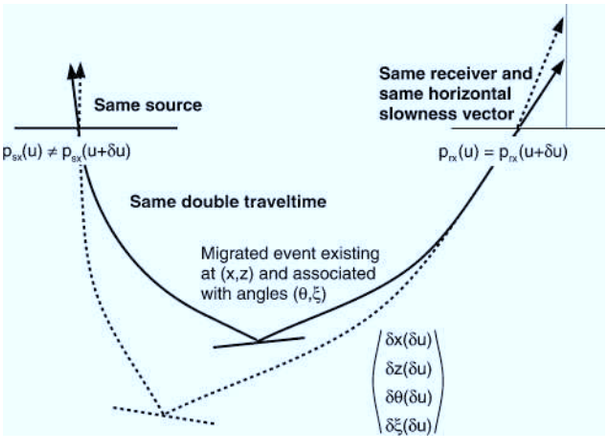


FIG. 10. Position of the migrated locally coherent events in two slightly different velocity models in a common-shot migration. All variables defined in depth (x, z, ξ, and θ) change. Their perturbations δx, δz, $\delta\xi$, and $\delta\theta$ are constrained by (1) constant position of the source and the receiver, constant double traveltime (fixed locally coherent event in the time domain), and (2) constant slope at the receiver position at the surface (common shot migration scheme). See text for more details.

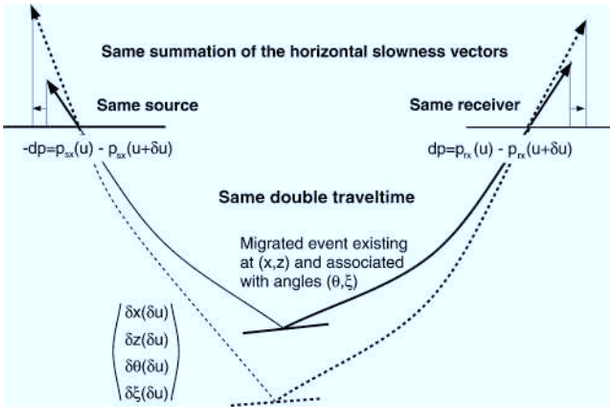


FIG. 11. Position of the migrated locally coherent events in two slightly different velocity models in a common-offset migration. The perturbations δx, δz, $\delta\xi$, and $\delta\theta$ must be evaluated to estimate the gradient of the cost function. They are constrained by (1) fixed position of the source and the receiver, constant double traveltime (same locally coherent event in the time domain), and (2) constant summation of the slopes at the source and receiver positions at the surface (common-offset migration scheme). More details are provided in the text.

We finally obtain, after computations (see Appendix B for details),

$$\left.\frac{\partial p_{hx}}{\partial u}\right|_{h,m,t,p_{mx}} = \frac{\det\left|\dfrac{\partial(p_{hx},h,m,t,p_{mx})}{\partial(u,x,z,\xi,\theta)}\right|}{\det\left|\left.\dfrac{\partial(h,m,t,p_{mx})}{\partial(x,z,\xi,\theta)}\right|_u\right|}. \qquad (41)$$

The expression of the gradient of the cost function is a combination of first-order paraxial quantities, which can be calculated using the classical ray theory (Farra and Madariaga, 1987).

Imaging condition

The final expression of the gradient of the cost function is given by equation (39) for migration at a common source. Here we examine the validity of this expression; in other words, in which cases the denominator is not equal to 0:

$$\det\left|\left.\frac{\partial(s,r,t,p_{rx})}{\partial(x,z,\theta_s,\theta_r)}\right|_u\right| \neq 0. \qquad (42)$$

An equation similar to equation (42) has been interpreted by Xu et al. (1999) as an imaging condition required to compute and use CIGs. Indeed, common-shot migration is not always possible for complex velocity fields with caustics (Nolan and Symes, 1996; ten Kroode and Smit, 1997; Xu et al., 1999). Spurious effects may be observed, such as nonflat events in CIGs when the exact velocity field has been used for migration (Nolan and Symes, 1996; Xu et al., 2001). The approach developed by Xu et al. (1999) may be extended easily to a wrong migration velocity model, leading to the same conclusion: A CIG is artifact free when relation (42) is verified. This property is known as the imaging condition.

In conclusion, we encounter the same theoretical difficulties to obtain an artifact-free CIG or to compute the gradient of the cost function. We may encounter some difficulties in very complicated velocity models with triplicated rayfields. In this case, other migration schemes must be considered, such as common angle migration (Xu et al., 1999).

THE OPTIMIZATION SEQUENCE

Now we move from theoretical aspects to a more practical point of view. The velocity is estimated from the picked events by looping over the following steps.

First, we prestack depth migrate the 2-D data set to compute some selected CIGs and COGs (Figure 1). The velocity model used for migration is the tested velocity field.

Second, we pick locally coherent events [i.e., $\xi(x,z,u)$ and $\varphi(x,z,u)$] in the computed CIGs and COGs. If $\varphi \neq 0$ for all picked events, the velocity model used for migration should be updated. For a given migrated trace, picking is performed by computing local slant stacks (Schultz and Claerbout, 1978). The slant stack, weighted by a Hamming window centered on the trace, measures at every depth the local coherence simultaneously in the two panels: CIG and COG (Figure 5). Picking automatically selects depths and associated angles where the local coherency is maximum. An equivalent tool for the time domain has been developed by Billette et al. (1998) and automated. Working in the depth-migrated domain has the advantage that the S/N ratio is generally higher than in the time data, even after migration with a wrong velocity model (Adler, 1996).

Finally, we ray trace from all picked events up to the surface to compute the gradient of the cost function and to update the velocity field. This step involves searching for two specular rays starting symmetrically with respect to the normal to the dip defined by ξ and reaching the surface with the offset associated to the picked event. As directly adapted, paraxial ray theory (Červený et al., 1977; Farra and Madariaga, 1987) is used to compute the needed expressions—in particular, the Fréchet derivatives [equation (41)].

Refer to Chauris et al. (2002) for applications on 2-D synthetic and real data sets, including more details on practical aspects and implementation.

CONCLUSIONS

Our method is related directly to MVA and is valid for any 2-D velocity fields. We have shown how to use information from locally coherent events picked in the migrated cube. These events are treated independently up front, i.e., not attached to a specific reflector. If extra information such as correlations between these events is introduced, the inversion should be better constrained. Our method provides a smooth velocity model consistent with Kirchhoff-type migration schemes, but the theory could be developed equivalently for blocky models. The method could be called tomographic migration velocity analysis or stereotomography in depth because it links two different velocity estimation methods working in different space, time, and depth domains.

We know that the occurrence of coherent noise such as multiples or refracted waves remains a problem because it could bias the velocity inversion, but this aspect is not specific to our method. The details of the practical aspects are developed in Chauris et al. (2002) that also includes applications to 2-D synthetic and real data sets, showing the efficiency of the method.

ACKNOWLEDGMENTS

This work was partially funded by the European Commission within the JOULE project, 3D-Focus (contract JOF3-CT97-0029).

REFERENCES

Adler, F., 1996, Tomographie de réflexion à partir des images migrées avant addition: Ph.D. thesis, Université de Pau et des Pays de l'Adour.

Aki, K., and Richards, P. G., 1980, Quantitative Seismology: Theory and Methods: W. H. Freeman & Co.

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

Audebert, F., Diet, J. P., Guillaume, P., Jones, I., and Zhang, X., 1997, CRP Scan: 3-D preSDM velocity analysis via zero offset tomographic inversion: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1805–1808.

Billette, F., and Lambaré, G., 1998, Velocity macro-model estimation from seismic reflection data by stereotomography: Geophys. J. Internat., **135**, 671–680.

Billette, F., Podvin, P., and Lambaré, G., 1998, Stereotomography with automatic picking: Application to the Marmousi dataset: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1317–1320.

Biondi, B., 1992, Velocity estimation by beam stack: Geophysics, **57**, 1034–1047.

Bleistein, N., Cohen, J., and Hagin, F., 1987, Two-and-one-half dimensional Born inversion with an arbitrary reference: Geophysics, **52**, 26–36.

Burridge, R., 1976, Some mathematical topics in seismology: Courant Inst. of Math. Sciences.
Červený, V., Molotkov, I. A., and Psencik, I., 1997, Ray theory in seismology: Charles Univ. Press.
Chauris, H., and Noble, M., 1998, Differential semblance optimization for 2D velocity field estimation: Soc. Expl. Geophys./Eur. Assn. Geosci. Eng., Extended Abstracts for Workshop on Depth Imaging of Reservoir Attributes X012.
——— 2001, Two-dimensional velocity model estimation from seismic reflection data by local differential semblance optimization: Application to synthetic and real data sets: Geophys. J. Internat., **144**, 14–26.
Chauris, H., Noble, M. S., Lambaré, G., and Podvin, P., 2002, Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II: Applications on synthetic and real data: Geophysics, **67**, 1213–1224.
Clément, F., 1994, Une formulation en temps de parcours par migration pour la détermination des vitesses de propagation acoustique à partir de données sismiques bidimensionnelles: Ph.D. thesis, Université Paris IX Dauphine.
Cox, H. L. H., and Wapenaar, C. P. A., 1992, Macro model estimation by common offset migration and by shot record migration: J. Seismic Expl., **1**, 29–37.
Deregowski, S. M., 1990, Common-offset migrations and velocity analysis: First Break, **8**, 225–234.
Docherty, P., Silva, R., Singh, S., Song, Z., and Wood, M., 1997, Migration velocity analysis using a genetic algorithm: Geophys. Prosp., **45**, 865–878.
Farra, V., and Madariaga, R., 1987, Seismic waveform modeling in heterogeneous media by ray perturbation theory: J. Geophys. Res., **92**, 2697–2712.
Jervis, M., Sen, M., and Stoffa, P., 1996, Prestack migration velocity estimation using nonlinear methods: Geophysics, **61**, 138–150.
Jin, S., and Beydoun, W., 2000, 2D multiscale non-linear velocity estimation: Geophys. Prosp., **48**, 163–180.
Jin, S., and Madariaga, R., 1993, Background velocity inversion with a genetic algorithm: Geophys. Res. Lett., **20**, No. 2, 93–96.
——— 1994, Nonlinear velocity inversion by a two-step Monte Carlo: Geophysics, **59**, 577–590.
Lafond, C. F., and Levander, A. R., 1993, Migration moveout analysis and depth focusing: Geophysics, **58**, 582–598.
Lambaré, G., Lucio, P. S., and Hanyga, A., 1996, Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field: Geophys. J. Internat., **125**, 584–598.
Lee, W., and Zhang, L., 1992, Residual shot profile migration: Geophysics, **57**, 815–822.
Liu, Z., 1997, An analytical approach to migration velocity analysis: Geophysics, **62**, 1238–1249.
Liu, Z., and Bleistein, N., 1995, Migration velocity analysis: Theory and an iterative algorithm: Geophysics, **60**, 142–153.
Nolan, C., and Symes, W., 1996, Imaging in complex velocities with general acquisition geometry: The Rice Inversion Project Technical Report TR96-02.
Plessix, R. E., 1996, Détermination de la vitesse pour l'interprétation de données sismiques très haute résolution à l'échelle géotechnique: Ph.D. thesis, Université Paris IX Dauphine.
Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76.
Schultz, P. S., and Claerbout, J. F., 1978, Velocity estimation and downward continuation by wavefront synthesis: Geophysics, **43**, 691–714.
Stoffa, P. L., 1989, Tau-p: A plane wave approach to the analysis of seismic data: Kluwer Academic Publishers.
Sword, C. H., 1987, Tomographic determination of interval velocities from reflection seismic data: The method of controlled directional reception: Ph.D. thesis, Stanford University.
Stork, C. H., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, 680–692.
Symes, W. W., 1993, A differential semblance criterion for inversion of multioffset seismic reflection data: J. Geophys. Res., **98**, 2061–2073.
——— 1998, High frequency asymptotics, differential semblance, and velocity estimation: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1616–1619.
——— 1999, All stationary points of differential semblance are asymtotic global minimizers: Layered acoustics: The Rice Inversion Project Technical Report TR99-09.
Symes, W. W., and Carazzone, J., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.
ten Kroode, A. P. E., and Smit, D. J., 1997, A microlocal analysis of a linearized inversion problem: SIAM Inverse Problems in Geophysical Applications, 146–162.
Tygel, M., Schleicher, J., and Hubral, P., 1994, Pulse distortion in depth migration: Geophysics, **59**, 1561–1569.
Varela, C. L., Stoffa, P. L., and Sen, M. K., 1998, Background velocity estimation using non-linear optimization for reflection tomography and migration misfit: Geophys. Prosp., **46**, 51–78.
Virieux, J., 1996, Seismic ray tracing, *in* Boschi, E., Ekstrom, G., and Morelli, A., Eds., Seismic modelling of earth structure: Istituto Nazionale di Geofisica, 223–300.
Wang, B., Pann, K., and Meek, R. A., 1995, Macro velocity model estimation through model-based globally-optimized residual-curvature analysis: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1084–1087.
Woodward, M., Farmer, P., Nichols, D., and Charles, S., 1998, Automated 3D tomographic velocity analysis of residual moveout in prestack migrated common image point gathers: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1218–1221.
Xu, S., Chauris, H., Lambaré, G., and Noble, M., 2001, Common angle migration: A strategy for imaging complex media: Geophysics, **66**, 1877–1894.

APPENDIX A

COST FUNCTION

We explain how to derive equation (18) from equation (17) to express the residual slope with local quantities. As illustrated (Figure A-1), we have

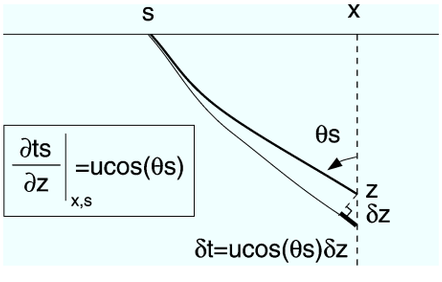


FIG. A-1. For a given source position and a velocity model, the derivative of the traveltime from the source to the scattering point with respect of the final depth depends on the arrival angle and the value of the slowness field.

$$\left.\frac{\partial t_s}{\partial z}\right|_{s,x,u} = u(x,z)\cos\theta_s, \tag{A-1}$$

$$\left.\frac{\partial t_r}{\partial z}\right|_{x,r,u} = u(x,z)\cos\theta_r, \tag{A-2}$$

where $u(x,z)$ is the slowness value at the scattering point. Using

$$\cos\theta_s + \cos\theta_r = 2\cos\theta\cos\xi, \tag{A-3}$$

we obtain

$$\left.\frac{\partial t_s}{\partial z}\right|_{s,x,u} + \left.\frac{\partial t_r}{\partial z}\right|_{r,x,u} = 2u(x,z)\cos\theta\cos\xi = \alpha, \tag{A-4}$$

which is needed for the equivalent equations (18) and (20).

APPENDIX B
GRADIENT COMPUTATION

We now compute the gradient of the cost function using the paraxial ray theory, first for the common-shot case and then for the common-offset case. As expressed in equation (38), the gradient of the cost function in common-shot migration is given by

$$\frac{\partial J[u]}{\partial u} = 2\sum_{picks} u\cos\theta\cos\xi\tan\varphi\left.\frac{\partial p_{sx}}{\partial u}\right|_{s,r,t,p_{rx}}, \quad \text{(B-1)}$$

The perturbations $\delta s, \delta r, \delta p_{sx}$, and δp_{rx} of the final conditions on surface are linked through the paraxial ray theory to the initial perturbations δz, $\delta\theta_s$, and $\delta\theta_r$ at the scattering point and to the slowness field δu. They can be expressed along the rays using the propagator matrix (Aki and Richards, 1980) as follows:

$$\begin{pmatrix}\delta s\\ \delta s_z\\ \delta p_{sx}\\ \delta p_{sz}\end{pmatrix} = P_s(\tau,\tau_0)\begin{pmatrix}\delta x\\ \delta z\\ \delta p_{sx0}\\ \delta p_{sz0}\end{pmatrix} + P_s(\tau,\tau_0) \times \int_{\tau_0}^{\tau} d\tau' P_s^{-1}(\tau',\tau_0)\begin{pmatrix}0\\ \nabla u\delta u\end{pmatrix}, \quad \text{(B-2)}$$

$$\begin{pmatrix}\delta r\\ \delta r_z\\ \delta p_{rx}\\ \delta p_{rz}\end{pmatrix} = P_r(\tau,\tau_0)\begin{pmatrix}\delta x\\ \delta z\\ \delta p_{rx0}\\ \delta p_{rz0}\end{pmatrix} + P_r(\tau,\tau_0) \times \int_{\tau_0}^{\tau} d\tau' P_r^{-1}(\tau',\tau_0)\begin{pmatrix}0\\ \nabla u\delta u\end{pmatrix}, \quad \text{(B-3)}$$

where δs_z and δr_z should equal zero as the ray segments end at the surface. The perturbation $\delta\theta_s$ is linked to the perturbations of the slowness vectors δp_{sx0} and δp_{sz0} at the scattering point by $p_{sx0} = u\sin\theta_s$ and $p_{sz0} = u\cos\theta_s$. In practice, we used a second-order Runge–Kutta method to integrate the rays and the perturbations. The parameter along the ray is τ (Burridge, 1976; Virieux, 1996), given by $d\tau = vd\sigma$ (where σ is the curvilinear abscissa). The propagator is a 2×2 matrix, determined by (Farra and Madariaga, 1987)

$$\frac{\partial P}{\partial\tau} = \begin{pmatrix}0 & I\\ \frac{1}{2}\nabla\nabla u^2 & 0\end{pmatrix}P, \quad \text{(B-4)}$$

$$P(\tau_0,\tau_0) = I, \quad \text{(B-5)}$$

where I is the 2×2 identity matrix. The four focusing conditions state that the source s and receiver r positions, the total traveltime t, and the slope at the receiver p_{rx} are constant. They all depend on x, z, $\delta\theta_s$, $\delta\theta_r$, and u. Remember that the differentiation of a variable can be expressed as follows:

$$d\cdot = \left.\frac{\partial\cdot}{\partial x}\right|_{u,z,\theta_s,\theta_r}\delta x + \left.\frac{\partial\cdot}{\partial z}\right|_{u,x,\theta_s,\theta_r}\delta z + \left.\frac{\partial\cdot}{\partial\theta_s}\right|_{u,x,z,\theta_r}\delta\theta_s + \left.\frac{\partial\cdot}{\partial\theta_r}\right|_{u,x,z,\theta_s}\delta\theta_r + \left.\frac{\partial\cdot}{\partial u}\right|_{x,z,\theta_s,\theta_r}\delta u. \quad \text{(B-6)}$$

The gradient of the cost function needs to differentiate the slope at the source position and finally express it as a function of δu. Conditions $[ds=0; dr=0; dt=0; dp_{rx}=0; dp_{sx}=K\delta u]$ can be expressed as

$$\begin{pmatrix}dp_{sx}\\ ds\\ dr\\ dt\\ dp_{rx}\end{pmatrix} = \begin{pmatrix}dp_{sx}\\ 0\\ 0\\ 0\\ 0\end{pmatrix} = \left(\frac{\partial(p_{sx},s,r,t,p_{rx})}{\partial(u,x,z,\theta_s,\theta_r)}\right)\cdot\begin{pmatrix}\delta u\\ \delta x\\ \delta z\\ \delta\theta_s\\ \delta\theta_r\end{pmatrix}. \quad \text{(B-7)}$$

Solving the linear system given by equation (B-7), we obtain the final formula (39).

The common-offset case is very similar to the previous case; p_{hx} and p_{mx} are defined in equations (21) and (22). The equations $dh=0$, $dm=0$, $dt=0$, and $dp_{mx}=0$ define a system of four linear equations which let us determine $\delta x, \delta z, \delta\xi$, and $\delta\theta$. We thus solve

$$\begin{pmatrix}dp_{hx}\\ dh\\ dm\\ dt\\ dp_{mx}\end{pmatrix} = \begin{pmatrix}dp_{hx}\\ 0\\ 0\\ 0\\ 0\end{pmatrix} = \left(\frac{\partial(p_{hx},h,m,t,p_{mx})}{\partial(u,x,z,\xi,\theta)}\right)\cdot\begin{pmatrix}\delta u\\ \delta x\\ \delta z\\ \delta\xi\\ \delta\theta\end{pmatrix} \quad \text{(B-8)}$$

to obtain the final formula (41).

GEOPHYSICS, VOL. 67, NO. 4 (JULY-AUGUST 2002); P. 1213–1224, 24 FIGS.
10.1190/1.1500383

Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part II: Applications on synthetic and real data

Hervé Chauris*, Mark S. Noble‡, Gilles Lambaré‡, and Pascal Podvin‡

ABSTRACT

We demonstrate a method for estimating 2-D velocity models from synthetic and real seismic reflection data in the framework of migration velocity analysis (MVA). No assumption is required on the reflector geometry or on the unknown background velocity field, provided that the data only contain primary reflections/diffractions. In the prestack depth-migrated volume, locations where the reflectivity exhibits local coherency are automatically picked without interpretation in two panels: common image gathers (CIGs) and common offset gathers (COGs). They are characterized by both their positions and two slopes. The velocity is estimated by minimizing all slopes picked in the CIGs.

We test the applicability of the method on a real data set, showing the possibility of an efficient inversion using (1) the migration of selected CIGs and COGs, (2) automatic picking on prior uncorrelated locally coherent events, (3) efficient computation of the gradient of the cost function via paraxial ray tracing from the picked events to the surface, and (4) a gradient-type optimization algorithm for convergence.

INTRODUCTION

Estimating the velocity model required for prestack depth migration is a critical step in depth imaging. Initial work on migration velocity analysis (MVA) (Yilmaz and Chambers, 1984; Al-Yahya, 1989) demonstrated the potential of common image gathers (CIGs) in the velocity inversion process. As more complicated cases were studied, nonlinear inversions were performed, requiring MVA to iteratively combine velocity estimation and migration. Thus, a velocity model must be consistent with these two dependent steps: (1) MVA based on Kirchhoff migration coupled with (2) a method for providing smooth velocity models. Smoothing inverted blocky models for migration changes the kinematics of wave propagation (Lailly and Sinoquet, 1996; Léger, 1996) and may destroy the benefits of a subtle inversion.

Another direct advantage of using a smooth velocity model is the parameterization of the model, which is far easier than that for blocky models. The model only consists of velocities with no need to introduce interfaces. On the other hand, most MVA methods providing smooth velocity models are highly nonlinear and require a global and costly optimization process (Jin and Madariaga, 1994; Jervis et al., 1996; Varela et al., 1998; Jin and Beydoun, 2000).

We discuss implementing and testing a method introduced in Chauris et al. (2002) (hereafter referred to as Part I). We outline the method, describe what is needed for its implementation, and present test results with synthetic and real data sets. We end with remarks concerning the feasibility of extending this approach to 3-D analysis.

METHODOLOGY

Our method pertains to the MVA philosophy, i.e., it is based on flattening events in CIGs obtained after common-offset prestack depth migration of the data set with some (presumably incorrect) velocity model. In the image cube (x, z, h), where x denotes the location, z the depth and h the half-offset, events considered in our method are only locally coherent. What we call a locally coherent event can be defined at any location (x, z, h) in the data cube where the imaged reflectivity shows good local coherency. Such an event is represented as an elementary planar surface centered at (x, z, h) whose orientation is defined by the two slopes (see Figure 1):

Manuscript received by the Editor June 12, 2000; revised manuscript received July 26, 2001.
*Formerly Ecole des Mines de Paris, Centre de Géophysique 35, rue Saint Honoré, 77305 Fontainebleau Cédex, France; presently Shell International, Volmerlaan 8, P.O. Box 60, 2280 AB Rijswijk, The Netherlands. E-mail: h.chauris@siep.shell.com.
‡Ecole des Mines de Paris, Centre de Géophysique 35, rue Saint Honoré, 77305 Fontainebleau Cédex, France. E-mail: mark.noble@geophy.ensmp.fr; gilles.lambare@geophy.ensmp.fr; pascal.podvin@geophy.ensmp.fr.

$$\tan\xi = \left.\frac{\partial z}{\partial x}\right|_h, \quad (1)$$

$$\tan\varphi = \left.\frac{\partial z}{\partial h}\right|_x, \quad (2)$$

where ξ defines the apparent geologic dip of the event in the migrated image at offset h [common offset gather (COG)] and φ is the residual slope measured in the CIG at location x. This residual angle should be null when the velocity model used for migration is correct.

Our method consists of simultaneously minimizing the residual slopes of a set of locally coherent events picked in the image cube. In Part I, we show that a single locally coherent event in the image cube bears information on the migration velocity model. This is why events can be treated as uncorrelated in the inversion process.

This theory is at the origin of two main practical advantages of our approach. First, locally coherent events can easily be picked automatically in the image cube because the method does not require them to be interpreted in terms of continuous horizons—neither in the images nor along offsets. Second, the velocity model to be estimated can be parameterized with smooth basis functions without introducing interfaces (and estimate their depth together with velocities, as is done, for instance, in reflection tomography).

In the following section, we briefly introduce the key points of our method, focusing on practical considerations. We then test the method against 2-D synthetic and real data. The synthetic data test allows us to discuss how densely events must be picked in the image cube, according to the parameterization of the problem. The real data test enables the study of the sensitivity of the approach to the occurrence of noise in the image cube. These tests are also discussed in terms of practicality, resulting in preliminary remarks about the feasibility of extending the approach to the 3-D case.

Computing the cost function and its gradient

The inverse problem to be solved ideally amounts to annihilating the residual slopes $\tan\varphi$ of a set of locally coherent events picked in the data cube. The generic form of the cost function we wish to minimize is a weighted least-squares sum of the form

$$J[v] = \frac{1}{2}\sum_{picks}(w\tan\varphi)^2. \quad (3)$$

In Part I, we show that the weighting function w can be chosen in such a way that the gradient of the cost function can be computed explicitly with the help of paraxial ray tracing in the tested velocity model.

The weighting term for each pick (x, z, h, φ, ξ) reads $w = 2u\cos\theta\cos\xi$, where u is the value of slowness at point (x, z), $\tan\xi$ is the apparent local geologic dip $(\partial z/\partial x)|_h$, and θ is the half-aperture angle, i.e., half of the angle between the two specular rays shot from point (x, z) according to dip ξ such that the rays reach the surface at locations separated by offset h (Figure 2).

This complicated weighting form has a simple geometric interpretation because it is the stretching factor of the migration operator (Tygel et al., 1994). The choice of w enables an easier computation of the cost function gradient. Indeed, in Part I we demonstrate that $J[v]$ can be rewritten in a completely different way that can be interpreted in relation to the mapping to the image cube of locally coherent events that could have been picked in the time-domain data cube (focusing equations). In this form, the gradient appears as a simple combination of paraxial ray quantities that can be computed easily by ray tracing (Farra and Madariaga, 1987), without making any simplifying assumption on the velocity model or on the reflector geometry.

This second form of $J[v]$ differs according to the migration scheme used. For common-offset migration, we prove that the gradient of the cost function is

$$\nabla J[v] = -\sum_{picks} w\tan\varphi \cdot \frac{\det\left|\dfrac{\partial(p_{hx}, h, m, t, p_{mx})}{\partial(u, x, z, \xi, \theta)}\right|}{\det\left|\dfrac{\partial(h, m, t, p_{mx})}{\partial(x, z, \xi, \theta)}\right|_u}, \quad (4)$$

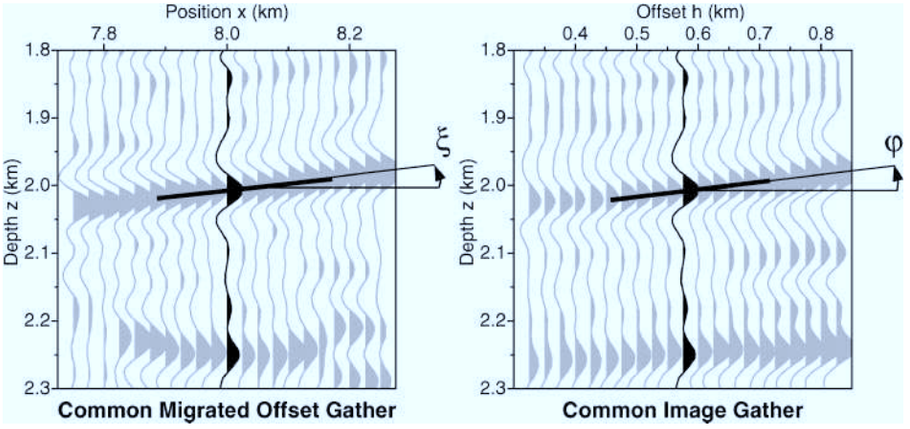


FIG. 1. The five values characterizing a locally coherent event in the prestack depth-migrated domain: (x, z, ξ, φ, h). Migration at common offset h is performed to obtain the two panels, CIG and COG, where picking is simultaneously performed around the same trace (x, h).

where m is the midpoint. The two slopes are defined by

$$p_{hx} = \frac{p_{sx} - p_{rx}}{2}, \quad (5)$$

$$p_{mx} = \frac{p_{sx} + p_{rx}}{2}, \quad (6)$$

where p_{sx} and p_{rx} are the horizontal components of the slowness vector at the extremity of the ray segments on the surface (Figure 2).

The optimization sequence

In practice, we estimate the velocity from the picked events by looping over three steps.

First, we prestack depth migrate the 2-D data set to compute some CIGs and common offset gathers (COGs) at select locations. Velocity model used for migration is the tested velocity field.

Next, we pick locally coherent events [i.e., $\varphi(x, z, v)$ and $\xi(x, z, v)$] in the computed CIGs and COGs. If $\varphi \neq 0$ for all picked events, the velocity model used for migration should be updated.

Finally, we ray trace from all picked events up to the surface to compute the gradient of the cost function and to update the velocity field. In addition to computing the Fréchet derivatives, this step involves searching for two specular rays starting symmetrically with respect to the normal to the dip defined by ξ and reaching the surface with the offset associated to the picked event.

We now present the details and justify the main practical aspects adopted here.

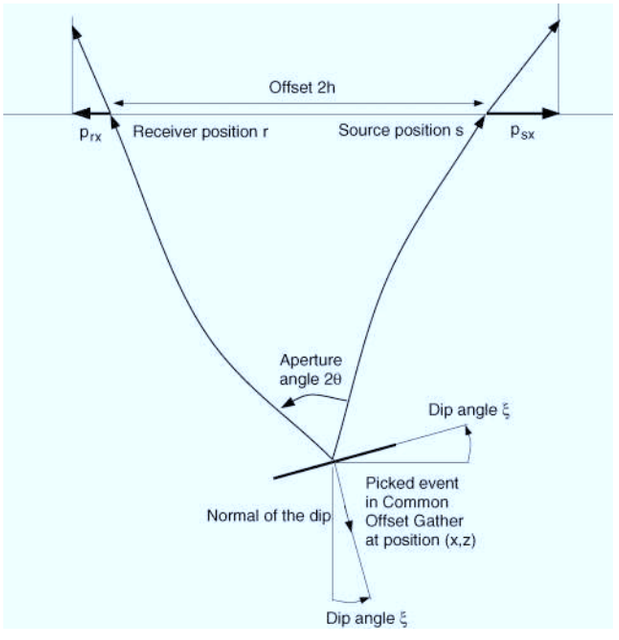


FIG. 2. Aperture angle and other characteristic values for each locally coherent event. The computation of the cost function gradient requires one to find the aperture angle 2θ for each picked event, such that the two ray segments pointing toward the surface reach the corresponding offset $2h$. These two rays start symmetrically around the normal of the dip segment, indicated by angle ξ.

PRACTICAL ASPECTS

Model parameterization

As underlined in the introduction, the model consists of only smooth velocity models without interfaces in spite of the picking step. The method does not require the introduction of blocky models and thus allows a simpler parameterization with a unique class of parameters: velocities.

We used cardinal cubic B-splines (de Boor, 1978) to describe all inverted velocity models. The number of unknowns, defined on regular nodes, is rather limited (around 1000 for the applications). Cubic B-splines are especially adapted for paraxial ray tracing needed to compute the gradient of the cost function (Chapman, 1985; Farra and Madariaga 1987; Lambaré et al., 1992). Such models are also directly useful for migration (Versteeg, 1993; Thierry et al., 1999a,b) because the optimization sequence has been chosen to be consistent with the imaging step.

Migration

For all migrations and inversions, we used a Kirchhoff-type migration scheme. Since the velocity estimation is based on a kinematic criterion, we used for simplicity an eikonal solver (Podvin and Lecomte, 1991) to compute the traveltime maps. The amplitudes were chosen to approximately compensate the geometric spreading with a velocity-independent correction. The choice of a Kirchhoff-type algorithm was motivated by the possibility of easily computing a single CIG. For efficiency, we only calculated some of the CIGs and COGs at predefined locations (Figure 3).

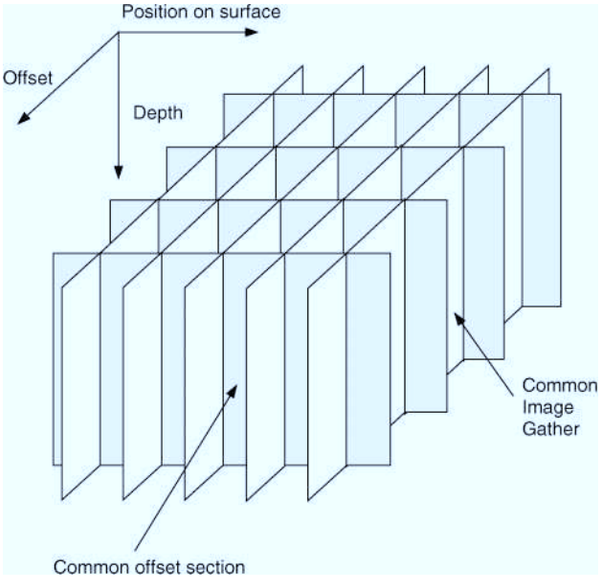


FIG. 3. Sparse migrated volume where the picking is performed. Blank and gray sections are, respectively, CIGs and COGs. Only some selected panels are computed for efficiency. Kirchhoff migration is especially adapted to obtain this type of volume. Picking is performed on traces that belong to a computed CIG and COG. The distance between two adjacent COGs and CIGs depends on the data set itself and the required parameterization of the velocity model.

Picking step

The picking step is the most crucial point. Picking is performed in the depth-migrated domain. The picks for 2-D data depend on depth, position on surface, and offset (Figure 3). From a practical point of view, it is clearly easier to automatically determine many locally coherent events than to follow continuous reflectors in the whole data set. Such evidence has already been pointed out—particularly in tomography. Classical traveltime tomography (Bishop et al., 1985; Chiu and Stewart, 1987; Farra and Madariaga, 1988) requires as input continuous events picked over the entire data set and interpreted as reflectors. This difficult picking stage has been a limitation of the method (Grau and Lailly 1993; Le Stunff and Grenier, 1998). Alternatives have been proposed by mapping the time data into the depth-migrated domain where the picking is performed (Adler, 1996; Apostoiu-Marin and Ehinger, 1997). Even in this case, the same limitations are encountered. Looking over events with local coherency instead of global continuity does not represent the same difficulty.

We pick a depth as well as the associated local slopes. For this purpose, we use local slant stacks as follows (Figure 4): for each selected trace and each depth, the local slant stack applied on the migrated signal measures the local coherency (Schultz and Claerbout, 1978; Milkereit, 1987). The local weighting is ensured by a Hamming window (Kanasewich, 1981). Figure 5 presents two migrated panels—a CIG and a COG—extracted from the migration of a 2-D real data set (equivalence of Figure 4a) and the result of the local stack performed around the same reference trace (equivalence of Figure 4c). To evaluate the slope, we compute the normalized square of the envelope (Taner et al., 1979) of the stacked traces:

$$A(p, z; x_{ref}) = \frac{\left| \sum_x H(x - x_{ref}) \quad \hat{m}(x, z - p \cdot (x - x_{ref})) \right|^2}{\sum_x H^2(x - x_{ref}) \quad |\hat{m}(x, z - p \cdot (x - x_{ref}))|^2}. \quad (7)$$

Here, x denotes the position of the trace in the migrated cube, x_{ref} is the reference trace, p is the direction of the stack, $H(x)$ denotes the Hamming window, and $\hat{m}(x, z) = (1 + i\ Hilb)\ m(x, z)$ is the analytical trace. We consider $m(x, z)$ for a COG and $m(h, z)$ for a CIG. Picked events are then selected as a function of the quality of the local slant stack. The precision required in the picking to obtain information on the unknown velocity inversion is difficult to be known beforehand. This particular aspect is discussed with the application on the 2-D synthetic data set.

The slope-picking tool was automated. Automatic picking is driven by several control parameters:

1) the spacing of reference traces,
2) the spacing in depth of the picks,
3) the width of the Hamming window associated with the local slant stack, and
4) the value of the amplitudes picked simultaneously on the CIG and COG for cross-checking.

Finally, and most importantly, the automatic picking procedure is very fast. Typically it requires about 15 minutes on a standard workstation for a 10-km 2-D line. This time is reduced to a few minutes if the picking is performed on sparse selected sections (Figure 3), as applied here.

Inversion

The inversion step first involves the search for the aperture angle for each picked event. We must ensure that the two rays starting from the locally coherent event reach the surface with

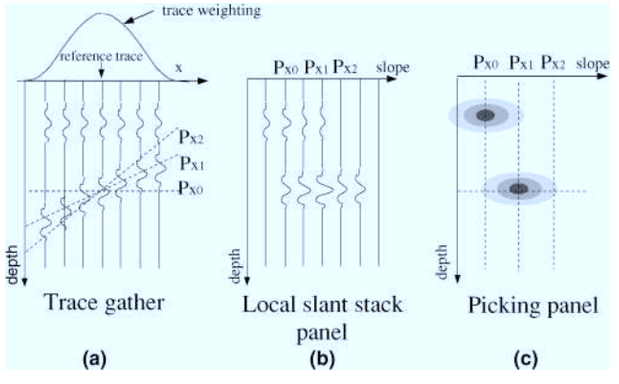


FIG. 4. Computation of the slope panels (right) from a trace gather (left). First, a local slant stack panel (middle) is obtained by slant stacking the trace gather, weighted around the reference trace by an Hamming window. Then, the slope panel is obtained by taking the normalized square envelope of the local slant stack panel.

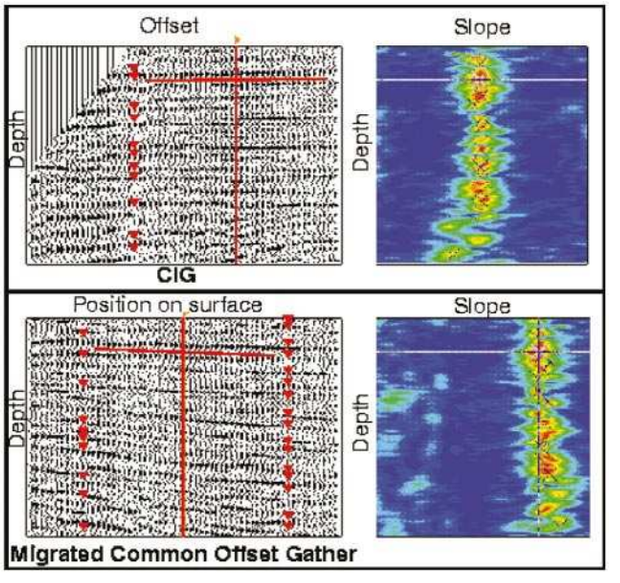


FIG. 5. Picking locally coherent events on a real data set. (left) CIG and migrated COG. (right) Semblance panels for a given trace (red line in CIG and COG) to determine local coherency and thus the position of the events and their two associated slopes. In this case, the slopes are almost horizontal. For a given migrated trace, local slant stacks measure at every depth the local coherence in two panels. This operation is totally automatic and only requires a few minutes for a realistic 2-D data set on a Sun Ultrasparc 10 workstation. No interpretation is introduced because there is no need to identify and follow interfaces. An equivalent tool for the time domain has been developed by Billette et al. (1998).

a consistent offset (Figure 2). A shooting method (Virieux and Farra, 1991) combined with the use of paraxial derivatives (Farra and Madariaga, 1987) efficiently obtains the aperture angle (less than five shooting steps).

The cost function is minimized with a local optimization process. The nonlinearity of the problem is overcome by migrating the data set at each iteration. Theoretical reasons for using a local optimization process are explained in Part I. The gradient of the cost function is given by equation (5) and is computed using paraxial ray tracing (Farra and Madariaga, 1987). For that purpose, we use a second-order Runge–Kutta method (Press et al., 1992) to integrate the dynamic ray equations. The inversion of the Fréchet derivatives is performed by a classical LSQR algorithm (Paige and Saunders, 1982). We also apply a Laplacian regularization term to stabilize the inversion (Lailly and Sinoquet, 1996).

Parameterization details, such as the number the velocity parameters needed for the inversion, the distance between selected CIGs, and the number of picked events required for a stable inversion, are very difficult to estimate ahead of time. They are case dependent and thus are defined based on the data only.

APPLICATION ON SYNTHETIC DATA

We apply the inversion on synthetic data to test the algorithm. In particular, we can compare the results to the best one obtained by using the exact velocity model.

Data modeling

Ray + Born approximation (Lambaré et al., 1992) was used to compute the 2-D synthetic seismograms in the acoustic case. The major advantage of this modeling is that data only contain primary reflections/diffractions and do not contain coherent noise such as multiples, which could bias the inversion process. The Born approximation requires that the modeling define a smooth velocity model—namely, the background velocity model giving the propagating terms—and a high-frequency part corresponding to the reflectivity.

The background velocity model chosen here for the modeling is a smooth version of the original 2-D Marmousi velocity model (Versteeg and Grau, 1991) (Figure 6). This new velocity model is parameterized by 615 cardinal B-spline nodes every 240 m in the horizontal and vertical directions. Smoothing is performed with an additional Gaussian filter to have a caustic-free medium. Imaging can thus be obtained by only using first-arrival traveltimes in the migration algorithm (Thierry et al., 1999a; Operto et al., 2000). Even with the smoothing, the interpolated velocity model has strong lateral velocity variations (Figure 6). The associated reflectivity is the same as in the exact Marmousi data set to create many reflectors with different dips. Figure 7 presents the near-offset section on this synthetic data. The signal source is the second derivative of the Gaussian function.

The geometry of the acquisition is very similar to the exact Marmousi case (Versteeg and Grau, 1991). Here, 261 shot points were generated from position $x = 2600$ m (Figure 6) with 96 receivers per shot, at a constant depth $z = 10$ m. The spacing between shots and receivers is 25 m and the first offset is 100 m. The sampling rate is 4 ms.

Figure 8 presents the final migrated image obtained with the exact velocity model. Our migrated sections were obtained using a kinematic Kirchhoff migration scheme associated with

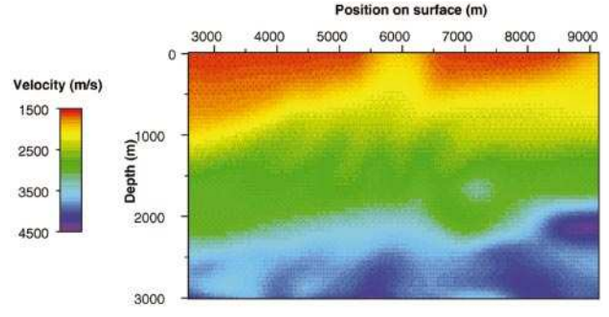


FIG. 6. Exact background velocity model used to generate the data. This model is a smooth version of the original 2-D Marmousi velocity model as described in Versteeg and Grau (1991). High lateral velocity variations exist, in particular up to 1400 m/s at depth $z = 2100$ m.

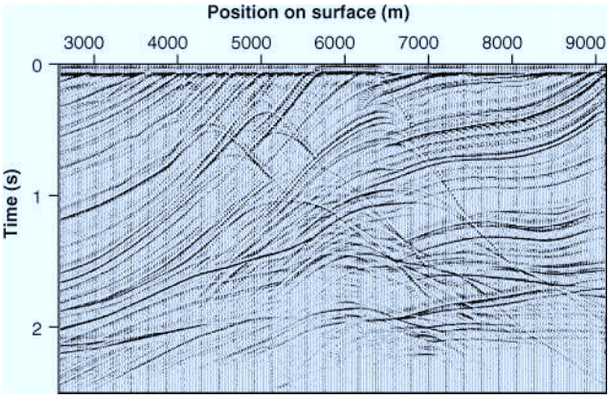


FIG. 7. Data generated by the ray + Born approximation using the smooth velocity model presented in Figure 6 and the reflectivity of the exact Marmousi 2-D data set. Only the first offset corresponding to 100 m is displayed. Because of the choice of model, data only contain primary diffractions/reflections and could be considered perfect data for the inversion process.

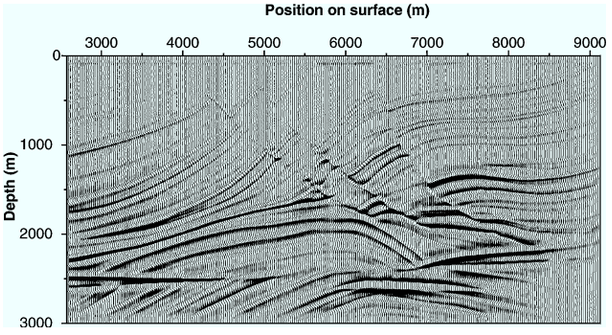


FIG. 8. Final migrated section of the data, using the exact velocity model. Only the first arrival is used for the migration.

an eikonal solver. This image provides the reference image for the inversion.

Inversion

For the velocity estimation, we parameterized the velocity model with 420 B-spline nodes (15 in z and 28 in x), spaced vertically by 240 m and horizontally by 400 m. All velocity nodes were simultaneously inverted, starting from a constant gradient velocity model. The migration of the data in this initial velocity model gives a poor idea of the exact reflectivity, especially in the deeper part around $x = 6000$ m (Figure 9). The bad quality of the image essentially comes from the destructive summation off all common-offset sections as the result of nonflat events in the CIGs.

To constrain the inversion, we computed at each iteration a sparse migrated volume as described in Figure 3. For this particular case, sixteen CIGs every 375 m and seven COGs every 375 m were selected. Automatic picking was performed on these panels. In spite of the distance between two neighboring CIGs, the inversion was stable because of the regularization term added in the cost function. At each iteration, about 3000 locally coherent events were picked, providing seven times more constraints than unknowns.

To illustrate the robustness of the picking step, the dip segments picked in three different COGs are displayed in Figure 10. Each event has a corresponding slope in the CIG. The lateral continuity of picked events is well retrieved, despite the absence of such constraints in our algorithm (the locally coherent events are treated as uncorrelated).

Figure 11 presents the distribution of the angles picked in the CIGs in two migrated cubes obtained with the exact and initial velocity models. These angles should be close to 0 with the exact model (flat events in CIGs). The distribution is not perfectly symmetric because of the use of ray tracing (Lambaré et al., 1992) for modeling and eikonal solver (Podvin and Lecomte, 1991) for migration, creating systematic numeric errors that remain negligible.

Performing the inversion, we converge after ten iterations to a model which shows great similarities with the exact velocity model (Figures 6 and 12). The deeper part of the velocity model is far smoother than for the exact velocity field. The final migrated section (stack of all offsets, Figure 13) shows very good focusing. In particular, faults are clearly defined and a good continuity of the reflectors has been obtained, even if this notion was never taken into account during the inversion (compare with Figure 9). The complex zone around $x = 7000$ m and $z = 2500$ m is also well focused.

The new distribution of the residual angles is displayed in Figure 14. It is symmetric around 0 with a Gaussian shape resulting from a least-squares minimization process. Taking into account this symmetry, the inverted result can be considered as good as the exact velocity model. This conclusion is enhanced by looking at the CIGs computed in three (initial, inverted, and exact) velocity models at two surface positions (Figure 15). The CIGs in the central and right part present the same high quality confirmed by the associated cost function. The main difference is the depth where the reflectors focus.

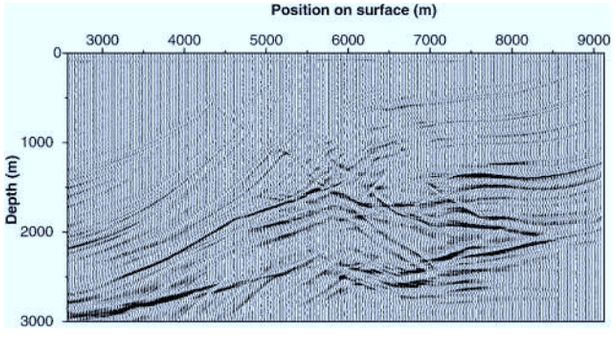


FIG. 9. Migration of the data in the initial constant (gradient velocity) model.

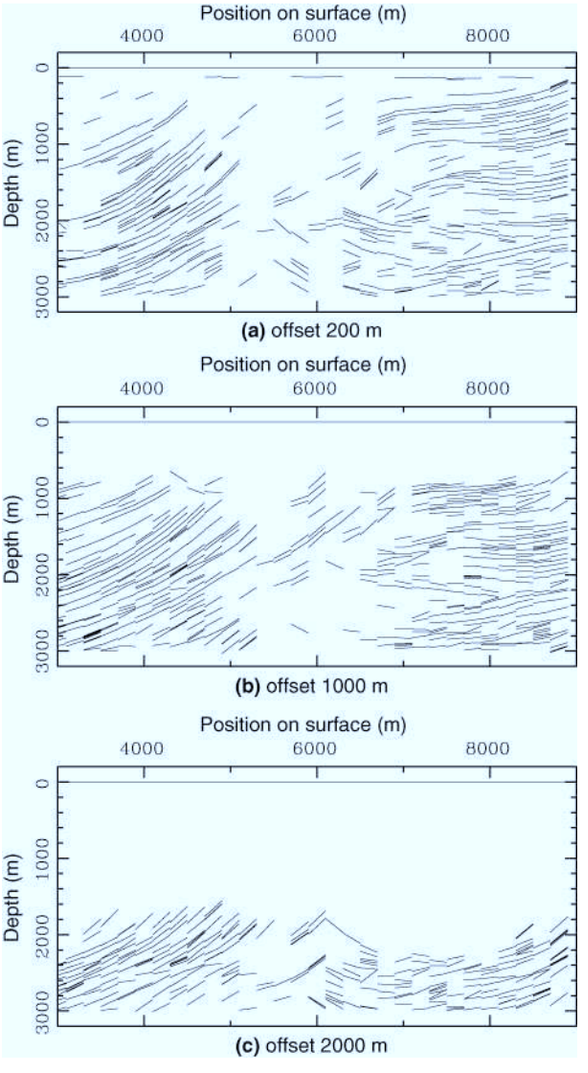


FIG. 10. Position of the picked ray segments in three common-offset sections corresponding to (a) offset = 200 m, (b) offset = 1000 m, and (c) offset = 2000 m. The velocity model used here was obtained at the second iteration in the inversion process. To illustrate the picking step, selected positions on the surface have been chosen every 200 m.

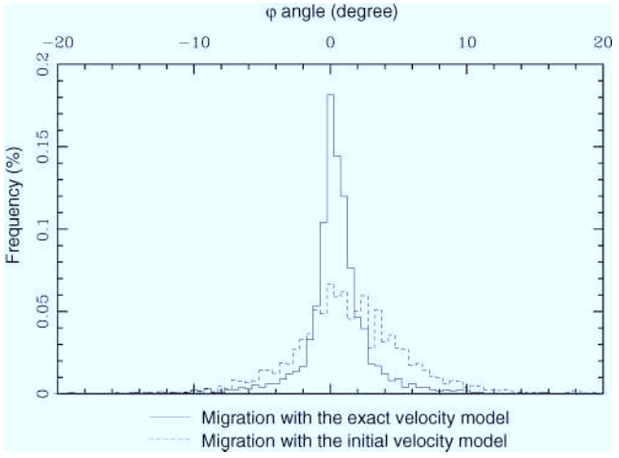


FIG. 11. Distribution of angles picked in the different CIGs. Each class represents 0.5°. The migrations are performed with the exact velocity model (solid line) and with the initial velocity model (dot line). The distribution is not perfectly symmetric because of differences between the modeling code (ray front construction, Lambaré et al., 1996) and the migration code (eikonal solver by finite differences, Podvin and Lecomte, 1991).

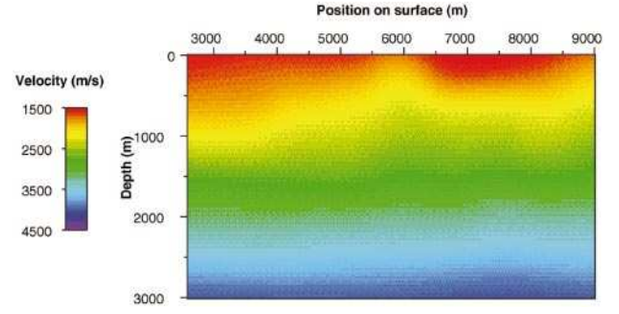


FIG. 12. Final inverted velocity model. The first CIG used for the inversion is at $x = 2600$ m, corresponding to the position of the first source in the data set. The velocity model shows great similarities with the exact velocity field (Figure 6), especially in the first 1.5 km. The deeper part is far smoother.

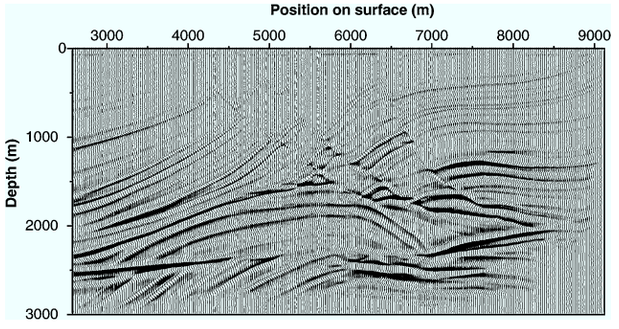


FIG. 13. Migration of the data in the inverted velocity model. The image is focused as well as in Figure 8. The main difference between the two images is the depth where reflectors focus.

Before discussing this aspect in more detail, we want to emphasize again that we were able to obtain a well-focused migrated image with an automatic, local optimization process.

Discussion

A close look at the depths where reflectors focus is provided by directly comparing five logs every 1 km extracted from the migrated images (Figure 16) (Figure 8 for the exact model and Figure 13 for the inverted model). These logs show that the maximum difference between the depths of reflectors is 100 m for an equivalent quality of focusing. The same migration code provides the different logs, allowing a direct comparison of the amplitudes. They are in most cases equal, proving that the events are equivalently flat in the inverted and exact velocity models.

Some differences, especially for $x = 6000$ m around depth $z = 1300$ m, are from lateral deplacements of the inverted image, particularly visible in the faulted regions. Several explanations are possible for such differences. The first reason is the lack of precision in picking locally coherent events (dips and residual slopes). The second reason clearly appears in Figure 17. All of the dips picked in different COGs are displayed on the final migrated section for the inverted velocity model. A lack of picking in the shallower part, especially around $x = 5000$ m is because of the mute applied to the data and the strong dips in the migrated section. It can explain the differences in the depths of reflectors. A solution for this could be to determine the velocity model in the shallower part by an inversion method based on refracted waves (Landa et al., 1995).

We were thus able to use a local optimization process for the inversion. Our aim was to demonstrate that automatically picked locally coherent events let us invert the velocity model and obtain flat events in CIGs. The algorithm has also shown its ability to retrieve lateral velocity variations. In the next section,

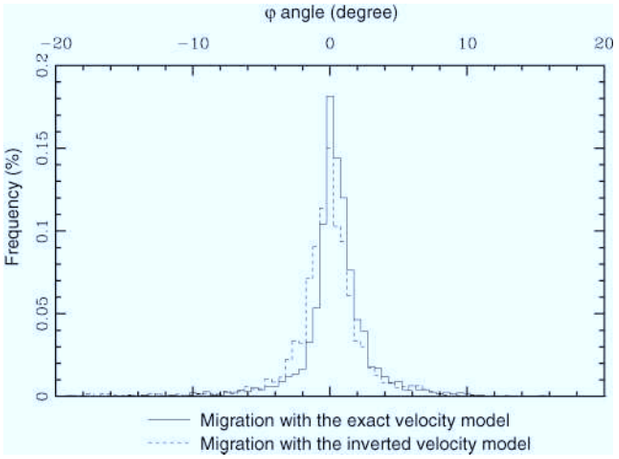


FIG. 14. Distribution of angles picked in the different CIGs. Each class represents 0.5°. The migrations are performed with the exact velocity mode (solid line) and with the inverted velocity model (dotted line). The shape of the distribution of the inverted model is Gaussian because of least-squares minimization and is centered around zero. The associated cost function, a weighted summation of the tangent of all angles, is equivalent for the two distributions.

the robustness of the method is tested with a real marine data set.

APPLICATION ON MARINE DATA

Data

The 2-D marine data set was provided and preprocessed by TotalFina. In particular, the preprocessing includes multiple attenuation. The line contains 505 shot points with an offset range from 112.5 to 2387.5 m and a record length of 4 s. Shot and receiver spacing is approximately 25 m. We applied a simple bandpass filter [5, 20, 45, 65 Hz] to the data before testing the inversion scheme. Figure 18 displays a common-offset time section before migration, showing many diffraction curves—especially around 2 s. This marine data set contains tilted blocks with faults and salt bodies, providing essentially a strong vertical velocity contrast.

Inversion

The inversion process is very similar to the optimization on synthetic data. The inversion was performed in a two-step

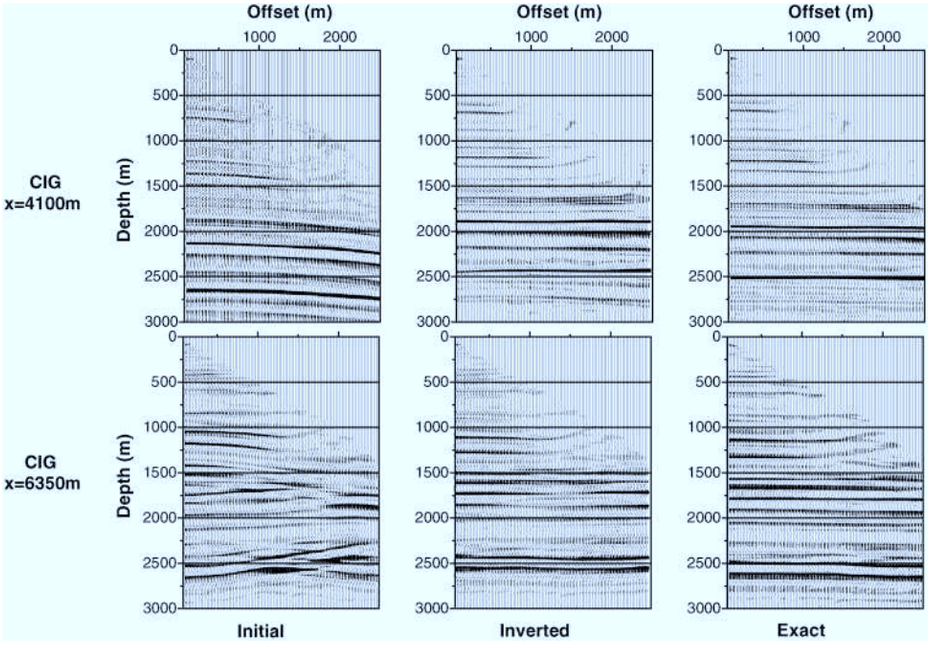


FIG. 15. CIGs computed at two positions on the surface ($x = 4100$ m and $x = 6350$ m) with the initial, inverted, and exact velocity models. The associated cost functions for the CIGs computed with the inverted and exact velocity models are equivalent, even if the depth of the reflectors differs. Complicated events around depth $z = 2500$ m for CIG $x = 6250$ m in the initial velocity model disappear in the two other models.

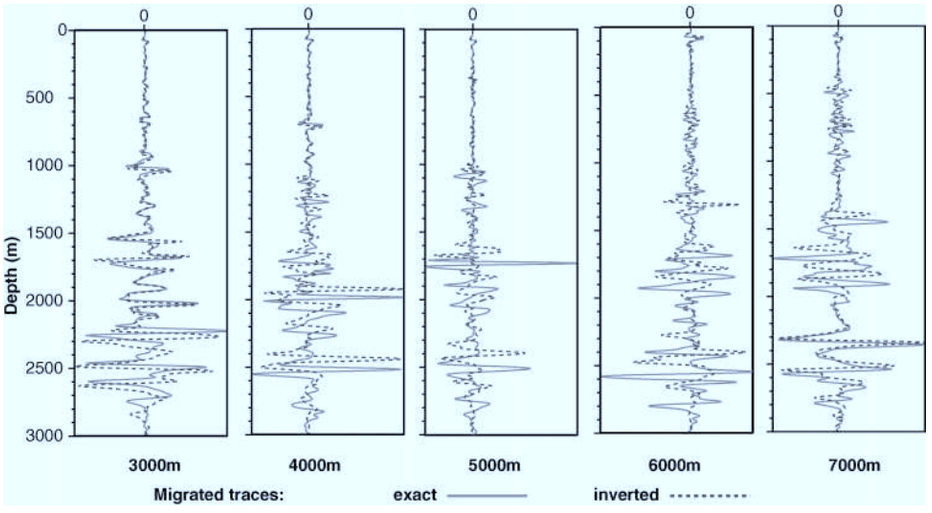


FIG. 16. Comparison of logs extracted from migrated images computed with the exact (Figure 8) and inverted (Figure 13) velocity models. The maximum difference between the depths where reflectors focus is 100 m.

multiscale approach to help convergence. Starting from a constant velocity gradient, we first inverted for a 1-D velocity model described by 29 *B*-spline nodes every 125 m in depth. A single CIG in the central part of the acquisition at position $x = 8000$ m and six COGs every 375 m were computed to perform the first inversion step. Using this result as a new starting model, we then inverted for a new 2-D velocity field with 1015 *B*-spline nodes every 125 m in z and every 500 m in x. To constrain the inversion, we used nineteen CIGs every 500 m and six COGs every 375 m. For the first step (1-D inversion), about 120 picks were used; approximately 1850 were used for the second step (2-D inversion). Their initial distribution is displayed in Figure 19.

Three iterations were needed for the first inversion and ten for the second step. The inverted velocity model is smooth as defined by cubic *B*-splines and presents small 2-D variations around a 1-D profile (Figure 20). The success of the inversion is shown by flat events in CIGs after inversion (Figure 21). The final distribution of the residual slopes (Figure 19) is very similar to the distribution obtained after inversion on the 2-D synthetic data set (Figure 14). It demonstrates the good quality of the result. The final depth-migrated image is built with the same migration code as the inversion process (Figure 22) and shows good continuity for the shallower part over well-defined tilted blocks and faults.

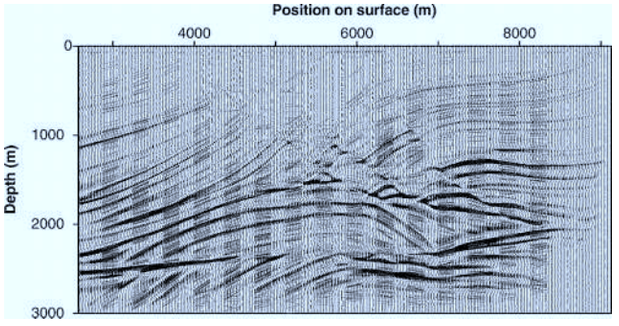


FIG. 17. Locally coherent events picked in different common-offset sections, superimposed on the migrated image computed with the final inverted velocity model.

Discussion

We encountered the same difficulties as when we used synthetic data. The evolution of the shape of the events and the associated picking are displayed in Figure 23. Picked events from different COGs are superimposed on the inverted velocity model (Figure 20). From Figures 20 and 23, it clearly appears that events are not picked in the shallower part (0–500 m). Because of this lack of information, the method does not ensure that the reflectors focus at their exact depth. One way to overcome this difficulty would be to introduce prior information such as wells (Le Stunff and Grenier, 1998) or, as mentioned before, to find the shallower velocity model with an adapted method such as refraction tomography (Landa et al., 1995).

In spite of a strong velocity jump (salt body), a smooth parameterization of the velocity field enabled us to model the kinematics of wave propagation. Migration and velocity estimation cannot be disassociated, and we proved that the velocity model can be consistent with the two steps.

If the 1-D inversion process gave the main low-frequency components of the velocity model, the 2-D inversion step is

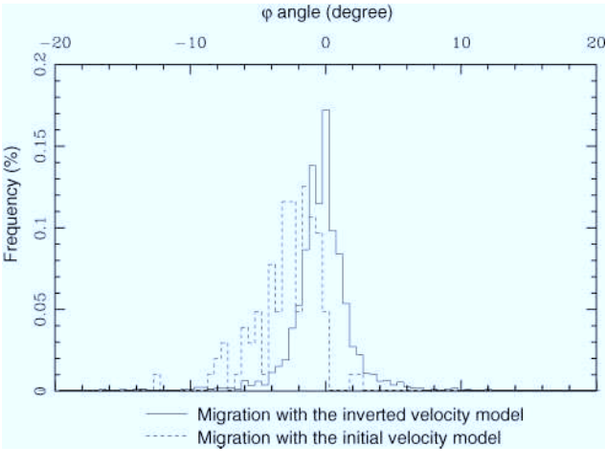


FIG. 19. Distribution of the residual slope before and after inversion. Each class represents 0.5°. The final shape (solid line) is very similar to the shape plotted (dotted line) in Figure 14.

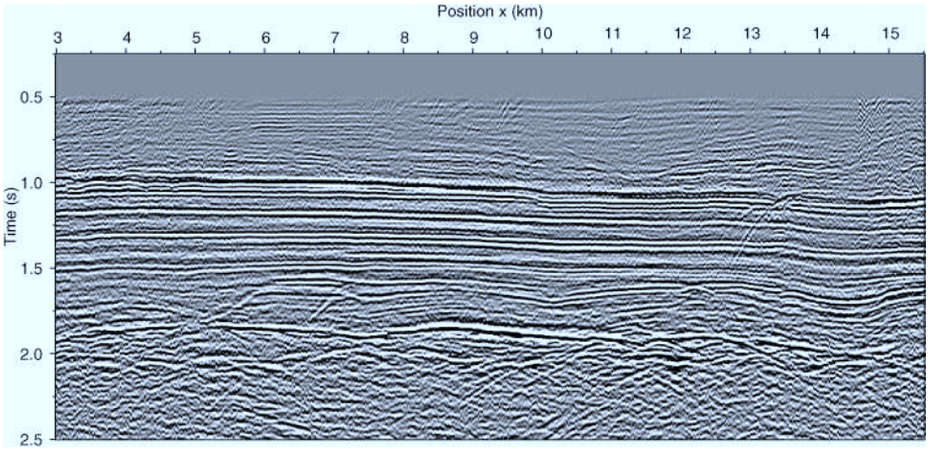


FIG. 18. Common-offset section for offset = 500 m. Only the first 2.5 s are displayed. We can distinguish a stratified overburden, a lower part with faults and tilted blocks over a deeper salt body. Many diffractions occur around the faults.

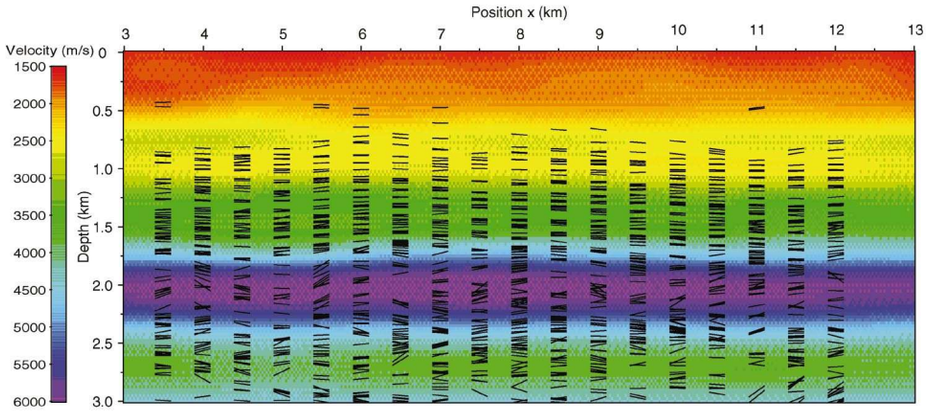


FIG. 20. The locally coherent events (>1800) used at the last iterations, superimposed on the inverted velocity model parameterized by 1015 cardinal B-splines spaced by 500 m in x and 125 m in z. The vertical high-velocity variations around $z = 2$ km are from salt.

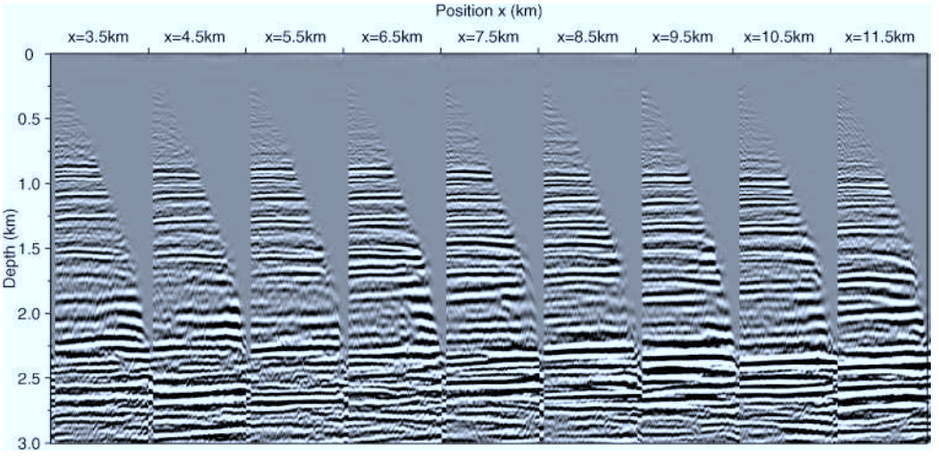


FIG. 21. Nine CIGs every 1 km obtained with the inverted velocity model. The flatness of events shows the good quality of the result.

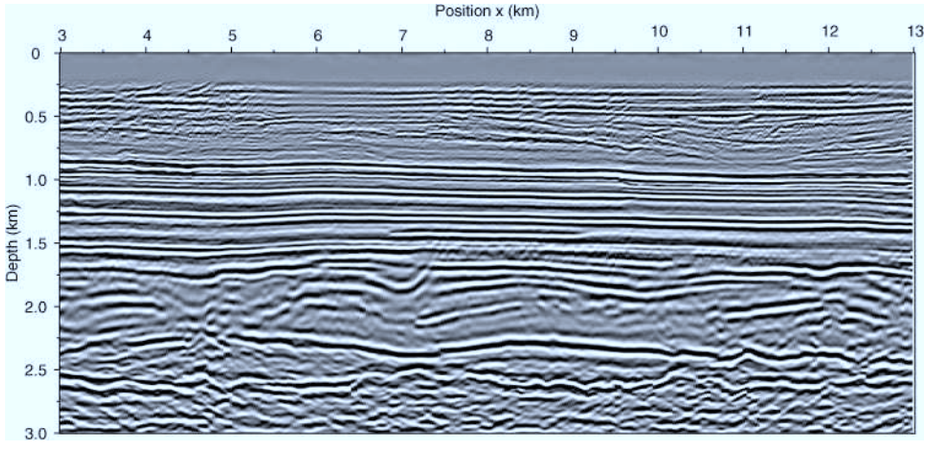


FIG. 22. The final migrated image after velocity inversion. From a geological point of view, we can distinguish a stratified overburden, a lower part with faults, and tilted blocks over a deeper salt body.

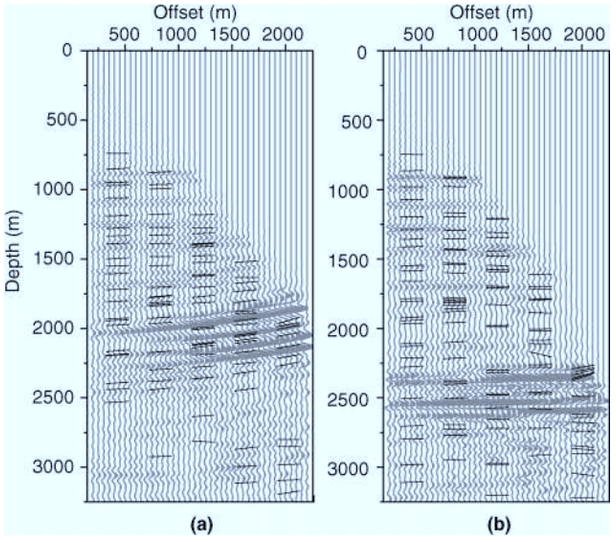


FIG. 23. (a) CIG computed in the initial velocity model with the picked segments for five different offsets. (b) CIG in the final inverted velocity model with the associated picked segments. The two CIGs correspond to the same $x = 8$ km. This CIG was also used for the 1-D inversion.

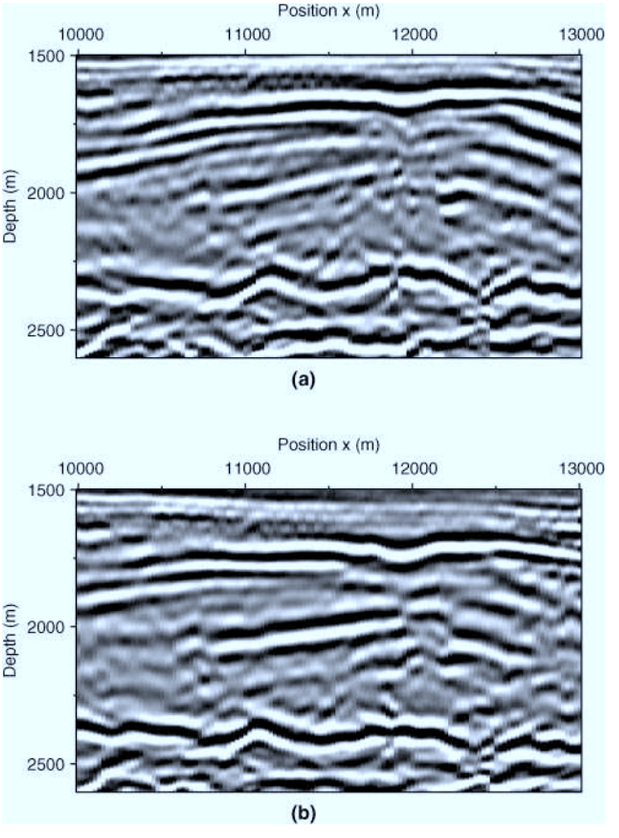


FIG. 24. Zooms on migrated images using (a) the velocity model after 1-D inversion and (b) the velocity model after 2-D inversion. The focusing of the image has been clearly enhanced after 2-D inversion, especially along the faults. The same migration code has been used to obtain these two images. Only the velocity model differs.

rather important, even if the additional velocity updates seem rather weak. As shown in Figure 24, the 2-D inversion indeed improves the quality of the final migrated image, compared with the result obtained after 1-D inversion. The tilted blocks appear more clearly, and some faults become visible. This demonstrates the good sensitivity of CIGs for velocity estimation (if needed).

An automatic inversion succeeded on this real data set. However, other case studies should be performed to confirm the robustness of the method, in particular with respect to noise.

CONCLUSIONS

The success of our method was illustrated on a 2-D synthetic data set and a marine data set, showing in particular its ability to handle real data. There are four key points for its efficiency. First, it allows automatic picking on locally coherent events in the depth-migrated domain on CIGs and COGs without interpretation. Coherent noise such as multiples still remains a difficulty. Second, it uses a gradient-type algorithm during inversion. Third, the method provides efficient computation of the cost function gradient using paraxial ray tracing, without any assumptions on the velocity model or on the reflector geometry. Dynamic ray tracing is simplified because the model does not contain interfaces. Finally, as mentioned above, only selected CIGs and COGs or computed at each iteration step. In the 2-D application on real data, the CPU cost for an iteration—including migration, picking, and velocity updating—corresponds to half the cost of the migration of the whole 2-D data set (here, 25 minutes on a Ultra 10 Sparc Sun station for an iteration).

Extensions to three dimensions will only be possible if such a sparse migration (*via* a Kirchhoff migration scheme) is used for inversion. The ratio between CPU cost for inversion and for migration of the whole data set should be much better. In fact, the more difficult aspect is certainly the possibility to compute CIGs in three dimensions due to the classical acquisition geometry, especially on marine data sets.

ACKNOWLEDGMENTS

This work was partially funded by the European Commission within the JOULE project, 3D-Focus (contract JOF3-CT97-0029). The authors thank TotalFina for providing the real data set and for permission to present the results. They especially thank Vincent Devaux and Yves Le Stunff for fruitful discussions.

REFERENCES

Adler, F., 1996, Tomographie de réflexion à partir des images migrées avant addition: Ph.D. thesis, Université de Pau et des Pays de l'Adour.

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

Apostoiu-Marin, I., and Ehinger, A., 1997, Kinematic interpretation in the prestack depth-migrated domain: Geophysics, **62**, 1226–1237.

Billette, F., Podvin, P., and Lambaré, G., 1998, Stereotomography with automatic picking: Application to the Marmousi dataset: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1317–1320.

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., and Spinder, D. A., 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, **50**, 903–923.

Chapman, C. H., 1985, Ray theory and its extensions: WKBJ and Maslov seismogram: J. Geophys., **58**, 27–43.
Chauris, H., Noble, M., Lambaré, G., and Podvin, P., 2002, Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects: Geophysics, **67**, 1202–1212, this issue.
Chiu, S. K. L., and Stewart, R. R., 1987, Tomographic determination of three-dimensional seismic velocity structure using well-logs vertical seismic profiles and surface seismic data: Geophysics, **52**, 1085–1098.
de Boor, C., 1978, A practical guide to splines: Springer-Verlag, New York, Inc.
Farra, V., and Madariaga, R., 1987, Seismic waveform modeling in heterogeneous media by ray perturbation theory: J. Geophys. Res., **92**, 2697–2712.
———1988, Non-linear reflection tomography: Geophys. J., **95**, 135–147.
Grau, G., and Lailly, P., 1993, Sequential migration-aided reflection tomography: An approach to imaging complex structures: J. Appl. Geophys., **30**, 75–87.
Jervis, M., Sen, M., and Stoffa, P., 1996, Prestack migration velocity estimation using nonlinear methods: Geophysics, **61**, 138–150.
Jin, S., and Beydoun, W., 2000, 2D multiscale non-linear velocity estimation: Geophys. Prosp., **48**, 163–180.
Jin, S., and Madariaga, R., 1994, Nonlinear velocity inversion by a two-step Monte Carlo: Geophysics, **59**, 577–590.
Kanasewich, E., 1981, Time sequence analysis in geophysics, 3rd ed.: Univ. of Alberta Press.
Lailly, P., and Sinoquet, D., 1996, Smooth velocity models in reflection tomography for imaging complex geological structures: Geophys. J. Internat., **124**, 349–362.
Lambaré, G., Lucio, P. S., and Hanyga, A., 1996, Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field: Geophys. J. Internat., **125**, 584–598.
Lambaré, G., Virieux, J., Madariaga, R., and Jin, S., 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.
Landa, E., Keydar, S., and Kravtsov, A., 1995, Determination of a shallow velocity-depth model from seismic refraction data by coherence inversion: Geophys. Prosp., **43**, 177–190.
Léger, M., 1996, Conversion de profils de vitesses: Institut Français du Pétrole Report 42777.
Le Stunff, Y., and Grenier, D., 1998, Taking into account a priori information in 3D tomography: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1875–1878.
Milkereit, B., 1987, Decomposition and inversion of seismic data—An instantaneous slowness approach: Geophys. Prosp., **35**, 875–894.
Operto, S., Xu, S., and Lambaré, G., 2000, Can we quantitatively image complex structures with rays?:, Geophysics, **65**, 1223–1238.
Paige, C., and Saunders, M. A., 1982, LSQR: Sparse linear equations and least squares problems, part I and part II: ACM Trans. Math. Soft., **8**, 43–71.
Podvin, P., and Lecomte, I., 1991, Finite difference computation of traveltimes in very contrasted velocity model: A massively parallel approach and its associated tools: Geophys. J. Internat., **105**, 271–284.
Press, W., Teukolsky, S., Vetterling, W., and Flannery, B., 1992, Numerical recipes in Fortran: The art of scientific computing, 2nd ed.: Cambridge Univ. Press.
Schultz, P. S., and Claerbout, J. F., 1978, Velocity estimation and downward continuation by wavefront synthesis: Geophysics, **43**, 691–714.
Taner, M. T., Koehler, F., and Sheriff, R., 1979, Complex seismic trace analysis: Geophysics, **44**, 1041–1063.
Thierry, P., Lambaré, G., Podvin, P., and Noble, M., 1999a, 3-D preserved amplitude prestack depth migration on a workstation: Geophysics, **64**, 222–229.
———1999b, Fast 2-D ray-Born inversion/migration in complex media: Geophysics, **64**, 162–181.
Tygel, M., Schleicher, J., and Hubral, P., 1994, Pulse distortion in depth migration: Geophysics, **59**, 1561–1569.
Varela, C. L., Stoffa, P. L., and Sen, M. K., 1998, Background velocity estimation using non-linear optimization for reflection tomography and migration misfit: Geophys. Prosp., **46**, 51–78.
Versteeg, R., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, **58**, 873–882.
Versteeg, R. J., and Grau, G., Eds., 1991, The Marmousi experience: 1990 EAEG workshop on Practical Aspects of Seismic Data Inversion, Eur. Assn. Expl. Geophys., Proceedings.
Virieux, J., and Farra, V., 1991, Ray tracing in 3-D complex isotropic media: An analysis of the problem: Geophysics, **16**, 2057–2069.
Yilmaz, O., and Chambers, R., 1984, Migration velocity analysis by wavefield extrapolation: Geophysics, **49**, 1664–1674.

GEOPHYSICS, VOL. 73, NO. 2 (MARCH-APRIL 2008); P. S35–S46, 20 FIGS.
10.1190/1.2831933

Seismic demigration/migration in the curvelet domain

Hervé Chauris[1] and Truong Nguyen[1]

ABSTRACT

Curvelets can represent local plane waves. They efficiently decompose seismic images and possibly imaging operators. We study how curvelets are distorted after demigration followed by migration in a different velocity model. We show that for small local velocity perturbations, the demigration/migration is reduced to a simple morphing of the initial curvelet. The derivation of the expected curvature of the curvelets shows that it is easier to sparsify the demigration/migration operator than the migration operator. An application on a 2D synthetic data set, generated in a smooth heterogeneous velocity model and with a complex reflectivity, demonstrates the usefulness of curvelets to predict what a migrated image would become in a locally different velocity model without the need for remigrating the full input data set. Curvelets are thus well suited to study the sensitivity of a prestack depth-migrated image with respect to the heterogeneous velocity model used for migration.

INTRODUCTION

A constant issue that interpreters must deal with is, given a migrated seismic image obtained in a particular velocity model, how the image distorts if the velocity model is modified locally (Murphy et al., 1993; Gray et al., 2000; Boschetti and Moresi, 2001). Interpreters can, for example, question the position of a reflector or might be uncertain about the uniqueness of the velocity model. In other words, there is a need for quantitative evaluation of image sensitivity with respect to the velocity model used for migration. The obvious solution is to remigrate the full data set in the new velocity model. However, for large data sets and despite growing computer capabilities, this might not be interactive enough. The main related difficulty, especially in the case of complex geology, is to select the appropriate input data that contribute to changes in the seismic image.

An alternative approach based on concepts proposed by Stolt (1996) and Sava (2003) is to distort a given migrated section without remigrating the input data. It combines within a single step both demigration and migration in two different velocity models. This is known as image waves (Hubral et al., 1996; Fomel, 2003) or Kirchhoff image propagation (Adler, 2002). Only Adler (2002) considers a depth-migration scheme in heterogeneous models. However, there exists a need to precompute, during the first migration, additional quantities required to later derive how the seismic image depends on the velocity model.

We revisit this approach for a prestack depth Kirchhoff migration scheme; an initial seismic image is obtained in a given velocity model, but there is no need for additional computed parameters during initial migration. In the demigration/migration process, the initial image is first demigrated and then migrated in a perturbed model. A new way to tackle this problem is to study the demigration/migration process through its effect on curvelets (Do, 2001; Candès and Donoho, 2004). The curvelet transform is a general operation that decomposes any image into elements that can be seen from the geophysical perspective as a representation of local plane waves (Figure 1). With image propagation, the depth gathers migrated in the initial velocity model are decomposed as a weighted sum of curvelets. Each curvelet is then distorted according to the demigration/migration process. The final image is obtained by recombining all modified curvelets. Thus, we need to quantitatively derive how the curvelets are affected by the demigration/migration operation.

Some studies indicate that curvelets are potentially interesting in this context. Theoretical results prove that for any given smooth image containing smooth discontinuities, the curvelet decomposition is almost optimal (Candès and Donoho, 2004) in the sense that only few curvelet coefficients are needed to represent such an image efficiently. This is particularly true for seismic gathers (Herrmann, 2003), where curvelets can be identified in local patches of reflected events (Figure 1a).

Beyond this aspect related to data decomposition, the curvelet functions are close to the eigenfunctions of the migration operator (Smith, 1998; Candès and Donoho, 2004; Candès and Demanet, 2005; Douma and de Hoop, 2005, 2007; Chauris, 2006). Instead of propagating energy in all possible directions as in a classical Kirchhoff migration scheme, the migration of a curvelet is controlled by a single direction, reducing the operation to a map-migration process

Manuscript received by the Editor 10 January 2007; revised manuscript received 29 October 2007; published online 7 February 2008.
[1]Ecole des Mines de Paris, Fontainebleau, France. E-mail: herve.chauris@ensmp.fr; truong.nguyen@ensmp.fr.

(Douma and de Hoop, 2007). Still, in the case of migration only, first applications have been presented on synthetic data for homogeneous velocity models (Douma and de Hoop, 2005, 2006; Chauris, 2006). The demigration operator can be seen as the pseudoinverse of the migration operator, and a priori the demigration/migration process is also well organized in the curvelet domain.

The potential interest of curvelets thus resides in their efficiency to decompose the input seismic gathers and in the possibility to sparsify the migration or demigration operators, at least in homogeneous models. Valuable applications require heterogeneous models that might create triplicated rayfields and large bending of the wavefront during migration. This might limit the usefulness of curvelets. However, combining demigration and migration within a single step has interesting properties; for example, if demigration and migration are performed in the same velocity model, then the curvelet is invariant, at least for a properly illuminated subsurface.

In this paper, we discuss whether the bending or curvature of a curvelet can be limited by the local velocity perturbation in the case of combined demigration/migration. This control is essential to know how curvelets sparsify the demigration/migration operator. We first review curvelet construction. Then we derive the first-order leading term that indicates how the demigration/migration process in two different velocity models affects a single curvelet. We also study the expected curvature of a curvelet after demigration/migration and compare it with the expected curvature after migration only.

The theoretical developments are illustrated by an application on a 2D synthetic data set generated in a smooth heterogeneous velocity model and with a complex reflectivity. A series of common-offset migrated sections is obtained by independently distorting the initial images with the help of curvelets. They finally provide common-image gathers (CIGs) in the perturbed velocity model. The results are compared to the one obtained by migrating the synthetic data in the modified model with a classical Kirchhoff migration scheme.

CURVELETS

We give a brief overview of curvelets, pointing out the main properties, and refer to Candès and Donoho (2004) and Candès et al. (2005) for a formal description and more details on the implementation.

Curvelets initially were designed for (nonseismic) image compression and denoising when data contain geometrical structures (Do, 2001; Do and Vetterli, 2003; Candès and Donoho, 2004). Such a 2D image $f(x,z)$ with smooth discontinuities can be decomposed as a series of curvelets. For the reconstructed image $f_n(x,z)$, keeping the first n most significant coefficients, we have

$$\|f - f_n\|_{L^2}^2 \sim \frac{(\log n)^3}{n^2}. \quad (1)$$

The optimal convergence rate would be $1/n^2$. By comparison, wavelet decomposition (Mallat, 1989; Daubechies, 1992; Meyer, 1993) and Fourier transform provide much less efficient representations, as the convergence rates are, respectively, $1/n$ and $1/\sqrt{n}$.

Beyond the capability of curvelets to efficiently decompose images with geometrical structures, curvelets have three characteristic properties (Candès and Donoho, 2004; Candès et al., 2005). First, the curvelet family forms a tight frame. Like in an orthonormal basis, a function $f(x,z)$ can be decomposed as a series of curvelets $c_\mu(x,z)$ with the reconstruction formula

$$f(x,z) = \sum_\mu \langle f, c_\mu \rangle c_\mu(x,z), \quad (2)$$

where $\langle \cdot,\cdot \rangle$ denotes the scalar product. The weight $\langle f, c_\mu \rangle$ is the amplitude associated with the curvelet c_μ. Second, the curvelets are essentially elongated (Figure 1). In the spatial domain, the width of the curvelet is proportional to the square of the length (parabolic scaling). Third, a curvelet $c_\mu(x,z)$ can be deduced from a reference curvelet $c_0(x,z)$ by combining a translation, a rotation, and a dilation (Figure 2). To satisfy the parabolic scaling, the dilation is expressed as

$$\begin{bmatrix} \sqrt{a} & 0 \\ 0 & a \end{bmatrix}.$$

The parameter μ denotes the 2D central position of the curvelet, its direction, and the scaling parameter a.

The parabolic scaling relationship can be understood as a compromise; very elongated curvelets would immediately be distorted during wave propagation. On the other hand, too short curvelets would not be truly directional and thus would be spread in all directions after migration or demigration.

From a practical point of view, we implement the curvelet transform in the following way (Candès and Demanet, 2005). First, we apply a 2D Fourier transform. Then, for all scales and all directions, we filter by the curvelet functions (Figures 1b and 3). Finally, we apply an inverse Fourier transform. This provides the curvelet coefficients. The inverse curvelet transform consists of applying a forward Fourier transform, filtering once more, adding the contributions for all scales and all directions, and performing an inverse Fourier transform. The curvelet transforms developed in Candès and Demanet (2005) provide efficient schemes of the order $O(n^2 \log n)$ for an input image of size $n \times n$. The number of directions is doubled at every two scales to satisfy the parabolic scaling.

In this paper, we use polar coordinates instead of Cartesian coordinates. This naturally leads us to consider rotations instead of shears as in Candès et al. (2005) and avoids special considerations for handling edge effects at $\pi/4$ in the wavenumber domain. The design of the filters is

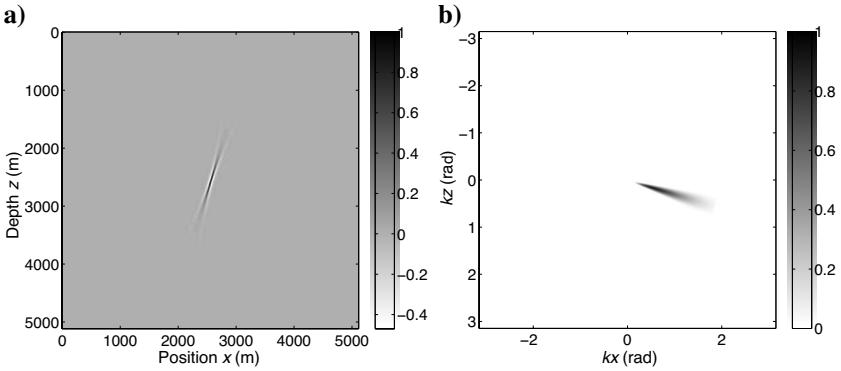


Figure 1. Example of a single curvelet (a) in the spatial domain and (b) in the wavenumber domain where it is perfectly localized.

based on the approach proposed by Simoncelli et al. (1992).

Each filter $F_{k,l}(\theta,r)$ is defined in the polar coordinate system (θ,r) and constructed as the product of two filters, $G_k(\theta)$ and $H_l(r)$ (Figure 3). Suppose we have 2^{N+1} different directions, with $N \geq 1$. The directional part $G_k(\theta)$ is defined for $k \in [1,2^{N+1}]$ as follows:

$$G_k(\theta) = \frac{2}{\sqrt{5}} \cos^3\left[N\theta - \frac{\pi}{4}(k-1)\right], \quad (3)$$

if $\theta \in [-\pi/2N + \pi/4N(k-1), \pi/2N + \pi/4N(k-1)]$ and zero otherwise. The factor $2/\sqrt{5}$ is a normalization factor, and one could check that the sum of all squared filters $\Sigma_k G_k^2(\theta) = 1$. This ensures a perfect reconstruction scheme. Examples of these filters are displayed in Figure 4. The radial filters $H_l(r)$ are defined in a similar way, with a change of variable from r to $\log r$, as described in Simoncelli et al. (1992).

The almost optimal convergence rate is particularly valid for seismic images. As underlined by Herrmann (2003), it is plausible that curvelets can decompose any seismic image efficiently.

THEORY FOR IMAGE DISTORTION IN THE CURVELET DOMAIN

The objective of this section is to derive, with the use of curvelets, how a seismic migrated image is distorted after demigration in the initial 2D smooth velocity model v followed by a migration in a perturbed velocity model $v + \delta v$. We suppose that the amplitude of the velocity perturbation is small compared to the background velocities and only concentrate on the kinematic aspects of the migration and demigration. In essence, the approach is similar to the one conducted by Douma and de Hoop (2005, 2007). We extend the formalism to heterogeneous velocity models by replacing straight rays (Douma and de Hoop, 2006) by curved rays. More importantly, we combine demigration and migration in a single step, whereas Douma and de Hoop (2006) are interested only in migration.

The first step consists of decomposing the initial common-offset migrated section as a weighted sum of curvelets. This is done by applying a curvelet transform onto the image. The depth reflectivity section $R(x,z)$ depends on the x-position and the depth z. We denote by $\mu = (x_i, z_i, \xi_i, f_i)$ the curvelet index corresponding to a central position (x_i, z_i), an angle ξ_i, and a scale or central frequency f_i. The amplitude associated with the curvelet c_μ is given by $A_\mu = \langle R, c_\mu \rangle$, yielding

$$R(x,z) = \sum_\mu A_\mu \cdot c_\mu(x,z). \quad (4)$$

The migration and demigration operators are linear operations, so we can restrict the study to the distortion of a single curvelet c_μ. The derivations are performed for common-offset sections. The half-offset h is not written explicitly in the following equations. The prestack depth demigration formula is given by Bleistein et al. (1987) and Thierry et al. (1999):

$$D_\mu(m,\tau) = \int \mathrm{d}\xi\, c_\mu[x(\tau,m,\xi,v), z(\tau,m,\xi,v)], \quad (5)$$

where m is the midpoint position and where τ is the traveltime between the source $s = m + h$, the subsurface position (x,z), and the receiver $r = m - h$ (Figure 5). We consider here only the kinematic

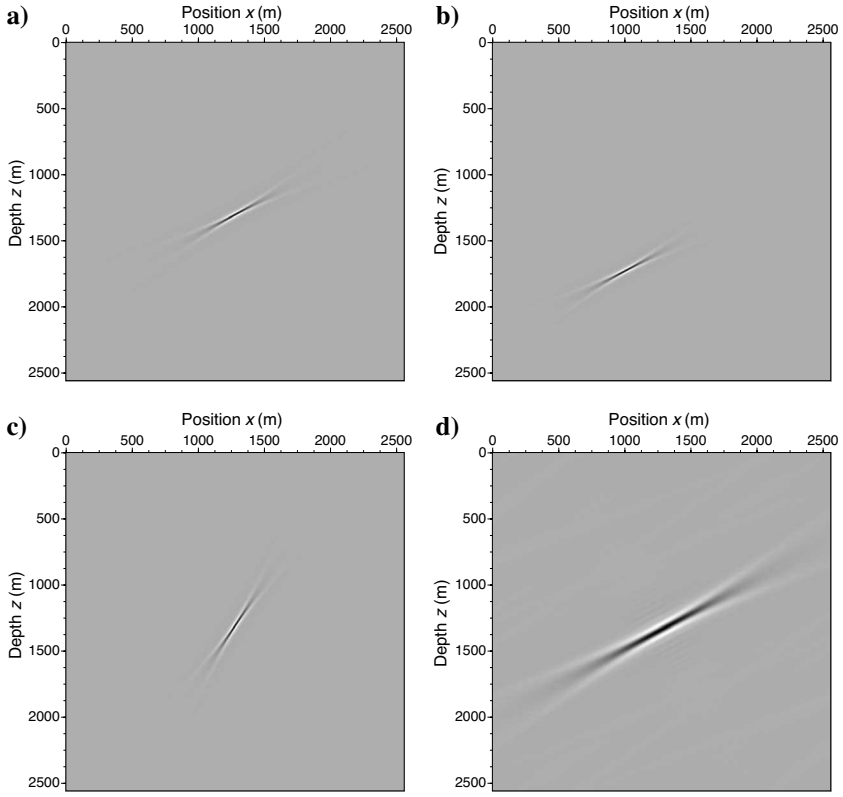


Figure 2. Curvelets in the spatial domain that can be deduced from (a) the reference curvelet by (b)translation, (c) rotation, and (d) dilation.

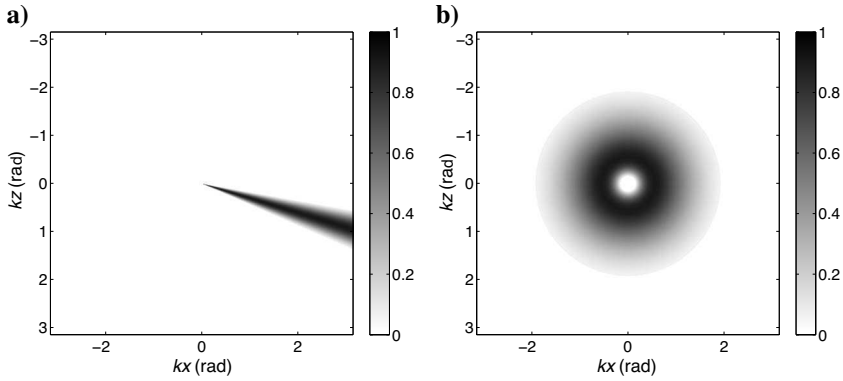


Figure 3. (a) Directional and (b) radial filters. The product of the two is used to construct the curvelet filter displayed in Figure 1.

aspects of demigration and do not take into account any amplitude terms. We suppose a valid imaging condition (Nolan and Symes, 1996; ten Kroode et al., 1998; Stolk and Symes, 2004) so the subsurface coordinates x and z can be expressed as a function of (τ, m, ξ, v).

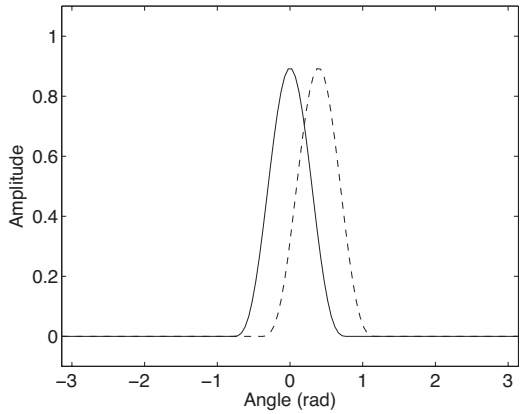


Figure 4. Example of filters $G_1(\theta)$ (solid line) and $G_2(\theta)$ (dashed line) used to design the curvelets.

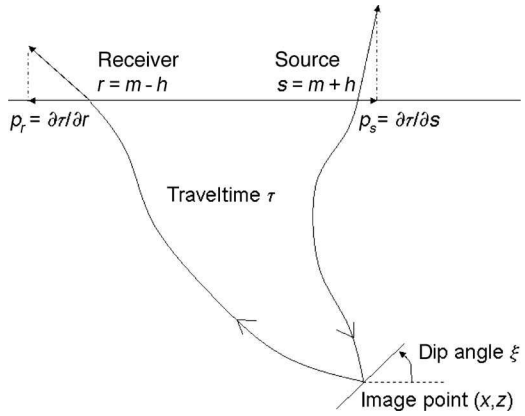


Figure 5. Schematic representation of specular rays for the Kirchhoff migration.

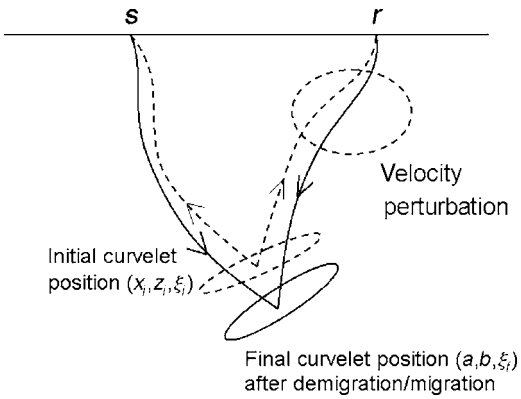


Figure 6. Illustration of image distortion. The migrated image is decomposed in the curvelet domain. Each curvelet is associated with a position (x_i, z_i) and a dip tan ξ_i. It is first demigrated using the initial velocity model and remigrated in the perturbed velocity model.

The prestack depth migration in the perturbed velocity $v + \delta v$ (Bleistein et al., 1987; Thierry et al., 1999), still neglecting the amplitude terms, is expressed by

$$R_\mu(a,b) = \int \mathrm{d}m D_\mu[m, \tau(a,b,m,v + \delta v)], \qquad (6)$$

where τ is the traveltime between the source, the subsurface position (a,b), and the receiver. We combine the two expressions to get

$$R_\mu(a,b) = \iint \mathrm{d}m\mathrm{d}\xi c_\mu[x(\tau(a,b,m,v + \delta v), m, \xi, v), z(\tau(a,b,m,v + \delta v), m, \xi, v)]. \qquad (7)$$

We want to simplify equation 7 (see Appendix A for details.) First, a 2D Fourier transform is applied on c_μ. Then, we exploit the fact that the curvelet is localized near a wedge in the Fourier domain (Figure 1b). This lets us apply the stationary-phase approximation (Bleistein, 1984) and reduce the number of integrals. As explained in Chauris et al. (2002a) and Douma and de Hoop (2007), the key parameters controlling the common-offset migration or demigration schemes are the midpoint $m = (s + r)/2$, the offset $h = (s - r)/2$, the two-way traveltime τ, and the derivative of the traveltime with respect to the midpoint $\partial\tau/\partial m = \partial\tau/\partial s + \partial\tau/\partial r = p_s + p_r$ (Figure 5). The slope values p_s and p_r correspond to the horizontal component of the ray parameter at the surface. The components $(s, r, \tau, p_s + p_r)$ do not depend on the velocity model chosen for migration or demigration. The slope $p_s - p_r$ is not used because each common-offset section is processed independently (Chauris et al., 2002a).

Finally, we further simplify the expression by considering small velocity perturbations. This leads to a simple formulation,

$$R_\mu(a,b) \simeq c_\mu\left(a + \frac{\partial x}{\partial \tau}\frac{\partial \tau}{\partial v}\delta v, b + \frac{\partial z}{\partial \tau}\frac{\partial \tau}{\partial v}\delta v\right). \qquad (8)$$

Equation 8 indicates, from a kinematic point of view, how a curvelet is transformed after demigration and migration in two different velocity models (Figure 6). The quantities in front of δv can be computed by paraxial ray tracing (Farra and Madariaga, 1987). The key element that controls the demigration and migration steps is the direction of the curvelet. Instead of having a double integral as in equation 7, the distortion of a curvelet is a morphing or distortion of the initial curvelet. The final section is obtained as the weighted sum $\Sigma_\mu A_\mu \cdot R_\mu(a,b)$.

ESTIMATING CURVATURE

Equation 8 is obtained by considering the first-order leading term. We extend this formalism and focus on deriving the curvature of the curvelet after demigration/migration, here for the zero-offset case only. By curvature, we mean a parameter controlling the quadratic

approximation of the wavefront after propagation (see Appendix B for a precise definition). Curvelets in the initial migration section have zero curvature. After demigration and migration in the same velocity model, the image remains the same and the final curvature is null.

We derive a quantitative formula expressing the curvature as a function of the velocity perturbation δv. After computations detailed in Appendix B, we get for the zero-offset case

$$\gamma = \left(\frac{\partial^2 \sin \xi_f}{\partial x \partial v} + \frac{\partial^2 \sin \xi_f}{\partial z \partial v} \cdot \tan \xi_i \right) \delta v . \tag{9}$$

The angles ξ_i and ξ_f are associated, respectively, with the curvelet before and after demigration/migration. Equation 9 is valid only for small velocity perturbations. As explained in Appendix B, we compare the formula with the one given by the formalism developed by Červený (2001). For a zero-offset migrated section, a local plane wave is propagated up to the surface in the initial velocity model and is back-propagated in the perturbed model (Figure 6).

As an example of a velocity model and velocity perturbation, we chose a smoothed version of the Marmousi model (Figure 7). The curvatures computed with two different methods are plotted in Figure 8. The two approaches used for the derivation show consistent curvature values.

It is interesting to do the same derivation for the migration part only, again for zero-offset migration. The curvelet propagation is controlled by the initial midpoint position m, the two-way traveltime t, and the initial slope of the curvelet p_μ (Douma and de Hoop, 2005, 2007; Chauris, 2006). In Appendix B, we derive an expression for the expected curvature,

$$\gamma_{\text{mig}} = \frac{\dfrac{\partial \sin \xi_f}{\partial m} + \dfrac{\partial \sin \xi_f}{\partial t} \cdot p_\mu}{\dfrac{\partial x_f}{\partial m} + \dfrac{\partial x_f}{\partial t} \cdot p_\mu}, \tag{10}$$

where x_f is the final position and ξ_f is the final angle. As a result of the denominator and as illustrated in Figure 9, the wavefronts can be largely bent, especially at caustics where the curvature value is infinite in ray theory. These curvature values derived in equation 10 are consistent with those given in Červený (2001) (Figure 9b). Note that the denominator in equation 9 is equal to one (see equations B-2 and B-4).

In this particular velocity model, the curvature induced by a single migration step is at least one order of magnitude larger than the curvature because of the demigration/migration process. This can be understood as follows: The migration, even in a different velocity model, undoes most of the bending of the curvelet. Equation 9 shows that the curvature can be limited to any threshold by selecting a small enough velocity perturbation. This indicates that it is easier to sparsely represent the demigration/migration operator in the curvelet domain than the migration operator itself.

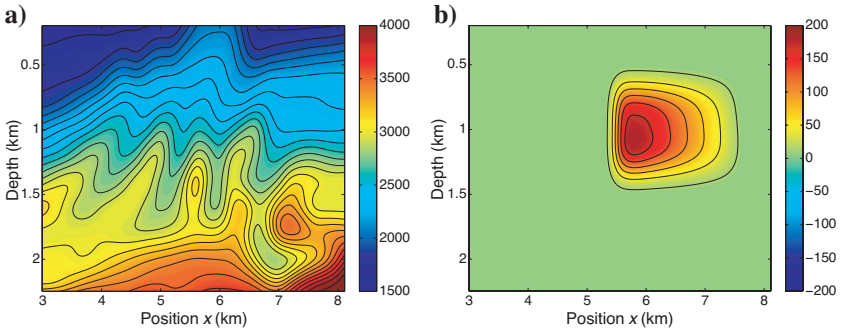


Figure 7. Exact velocity model v_{ref} (a) and velocity perturbation dv (b). The initial velocity model used for migration is defined by $v_{\text{ini}} = v_{\text{ref}} - dv$. Color scale denotes velocity in meters/second.

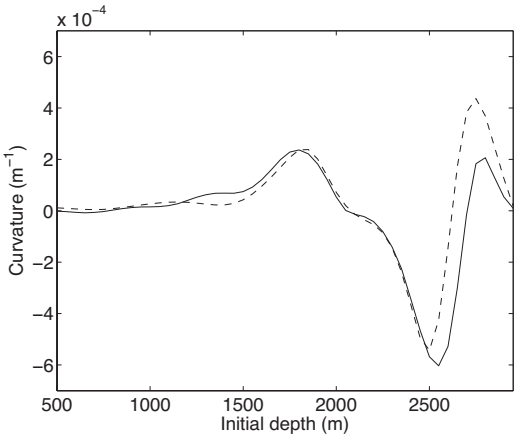


Figure 8. Curvature values after demigration/migration for different initial positions (x_i = 6100 m, $\xi_i = 15°$, and z_i between 500 and 2950 m) computed with equation 9 (solid line) and Červený's (2001) approach (dashed line). The curvature values correspond to a maximum velocity perturbation of 100 m/s.

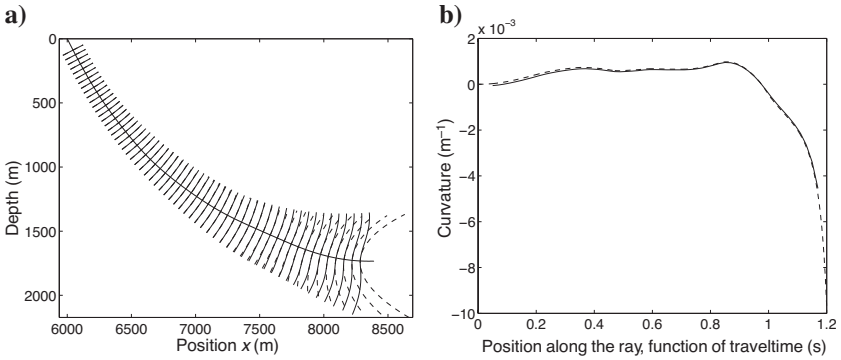


Figure 9. (a) Representation of plane-wave propagation (solid line: wavefront; dashed line: quadratic approximation) and (b) associated curvature values after migration for different propagation times computed with equation 10 (solid line) and the approach proposed in Červený (2001) (dashed line). The curvature values are one order of magnitude larger than for the demigration/migration case (Figure 8).

APPLICATION ON A SYNTHETIC DATA SET

A 2D synthetic data set, with offsets from 0 to 2 km every 50 m, is created with the ray + Born approximation (Bleistein et al., 1987; Thierry et al., 1999), using a smooth version of the Marmousi velocity model (Figure 7a) and the original Marmousi reflectivity (Figure 10a; Versteeg and Grau, 1991). Several common-offset sections are generated and processed independently (Figure 10b). First, we deal only with offset 600 m. The common-offset section is migrated with an initial velocity model $v_{\text{ini}} = v_{\text{ref}} - \delta v$, where v_{ref} is the exact velocity model (Figure 7). This provides an initial migration section (Figure 11). The objective is to predict the image that would be obtained by directly migrating the data in v_{ref}.

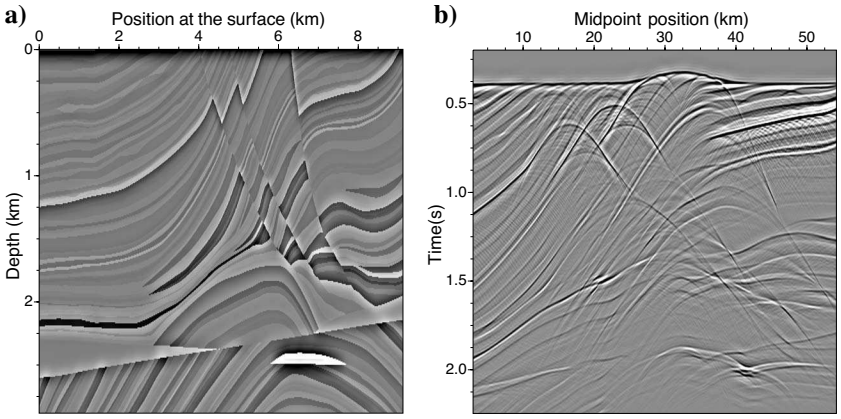


Figure 10. (a) Marmousi reflectivity model. (b) Associated ray + Born common-offset section (offset 600 m), generated in the exact velocity model v_{ref} (Figure 7a).

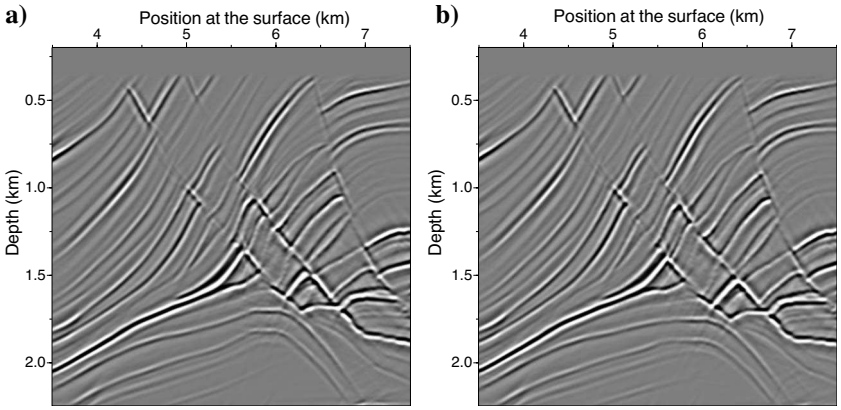


Figure 11. Kirchhoff common-offset section (for offset 600 m) migrated in (a) the initial velocity model v_{ini} and (b) the exact velocity model v_{ref}. The figure on the left is used as an initial image. The objective is to predict the image on the right, knowing v_{ini} and δv.

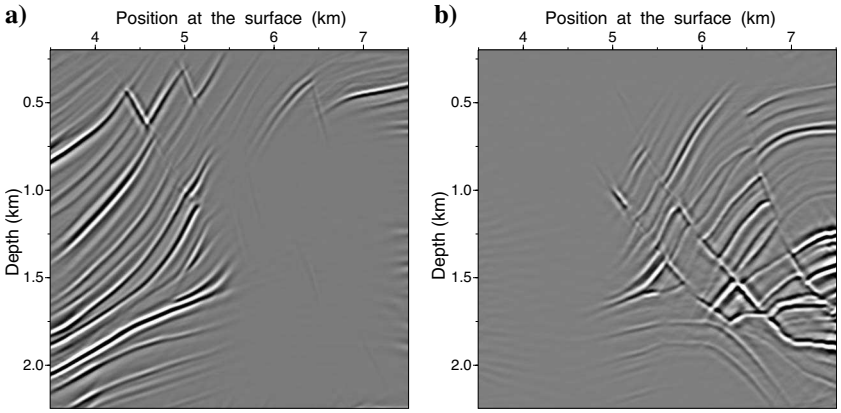


Figure 12. (a) Unchanged and (b) perturbed portions of the initial image (Figure 11a) after demigration/migration. Starting from the initial section, the image on the left is obtained by considering only the curvelet coefficients that are not modified after demigration/migration. On the contrary, the image on the right only uses curvelet coefficients affected by the velocity perturbation.

The initial image is decomposed in the curvelet domain. Each significant coefficient (here, 25% of the total number of the coefficients) is then processed independently. In a direct implementation of equation 8 for each significant curvelet coefficient, one would construct the associated curvelet in the spatial domain, apply the transformation (equation 8), and combine all results to get the final image. For a more efficient implementation, we take advantage of the radial construction of curvelets (Figure 3).

As explained in Appendix C, the curvelet transformation associated with demigration/migration (equation 8) is approximated by a combination of a shift, a rotation, and a dilation. When the computed shift, rotation, or dilation is not a multiple of the discrete value in the digital curvelet implementation, then the Shannon interpolation scheme (Zayed, 1993) is used (Appendix C). The shift, rotation, and dilation values are computed by ray tracing on a smooth velocity model (Farra and Madariaga, 1987). The practical details are given in Appendix C.

For each curvelet, the expected shift, rotation, and stretch are estimated by ray tracing, determining whether the curvelet should be modified. For example, when the velocity perturbation δv is far away, the curvelet coefficient is unchanged and there is no need for interpolation. Otherwise, the coefficient is modified according to the approximation of equation 8 (shift, rotation, and dilation). For the proof of concepts, two images are reconstructed by applying the inverse curvelet transform on the unchanged and perturbed coefficients (Figure 12). The final image is obtained as the sum of these two results (Figure 13a). In practice, a single inverse curvelet transform is applied on all coefficients, modified or not after demigration/migration.

Extracting a vertical section from the seismic migrated image allows us to better compare the prediction to the optimal result (Figure 13b). As a result of the velocity perturbation δv, a depth difference of 60 m can be observed. After prediction, the maximum depth difference is reduced to less than 2 m. Also, the phase of the signal is well predicted.

A direct gain in efficiency can be sought by processing fewer curvelet coefficients (Figure

14). The more coefficients, the more the lower-amplitude events become visible. In Figure 13a, 25% of the coefficients were used. Even with a very small number, the predicted image can be interpreted easily, at least in a tomographic approach where the demigration/migration process would be used to evaluate the quality of different velocity models.

All common-offset sections are processed independently in the same way. After demigration/migration in the curvelet domain, all results are combined and CIGs are extracted at three different positions (Figures 15–17). Once again, the predicted result shows an excellent match with the perfect result, obtained by fully migrating the input data in the exact velocity model. Note that it is not possible to predict the CIG in Figure 16c directly from the one in Figure 16a because of the lateral velocity variations. The full common-offset sections must be considered to extract CIGs later.

DISCUSSION

The derivation of the expected curvature formulas and their associated values (Figures 8 and 9b) indicate that curvelets are more suited for demigration/migration than for migration only. The illustration on synthetic data demonstrates that curvelets can handle a local velocity perturbation of a few 100 m/s in a smooth, heterogeneous model (Figure 7). Even in two different velocity models, the migration step undoes most of the induced curvature that appears in the demigration step. This is valid in heterogeneous models, even in the presence of triplications (Figure 18).

The explicit formulas (equations 9 and 10) give consistent results with the approach proposed by Červený (2001). The advantage of the formulation derived here is its interpretation. In equation 9, the velocity perturbation explicitly controls the maximum curvature. With the use of equation 10, it becomes clear that large curvature values are expected around caustics in case of migration. where the denominator tends to zero in the frame of the ray theory.

In the derivation of equation 8, only the kinematic aspects of migration and demigration were considered. However, the application on synthetic data — in particular, the result in Figure 13b — shows that this approximation is justified: The shape and amplitude of the signal after demigration/migration is well retrieved. Another approximation was used in the implementation of equation 8. For efficiency in relation with the curvelet construction, the curvelet distortion is treated as a combination of shift, rotation, and dilation (Appendix C). Such an approximation is also used in Douma and de Hoop (2005) and Chauris (2006) for migration only. The application proposed here in the case of demigration/migration shows that for small velocity perturbations, shift, rotation, and dilation are enough to reduce the maximum depth error to a few meters and to preserve the shape of the signal (Figure 13b).

The result expressed in equation 8 is a generalization of the result obtained by Adler (2002). He considers only how the positions, referred to as image points, are modified after demigration/migration

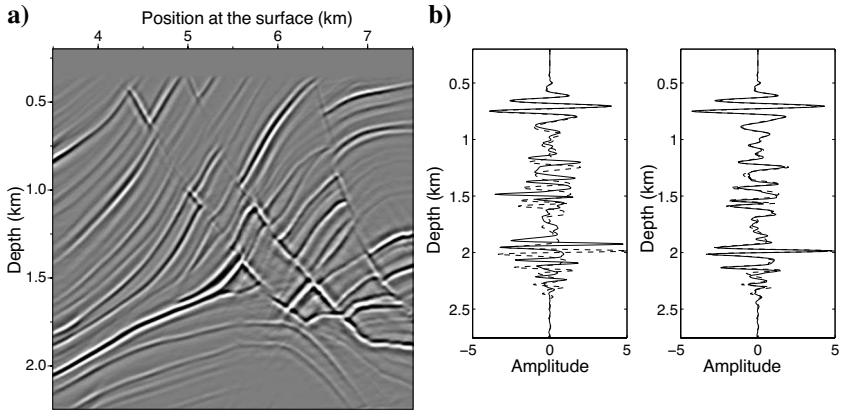


Figure 13. (a) Sum of the two images displayed in Figure 12 that should be compared to the image migrated in the exact velocity model (Figure 11b). (b) Extracted vertical sections for the position $x = 6000$ m. The dashed lines correspond to the image migrated in the exact velocity model v_{ref} and the solid lines to the image migrated in the initial model v_{ini} (left) and the predicted image (right). The maximum depth difference is reduced from 60 m to less than 2 m.

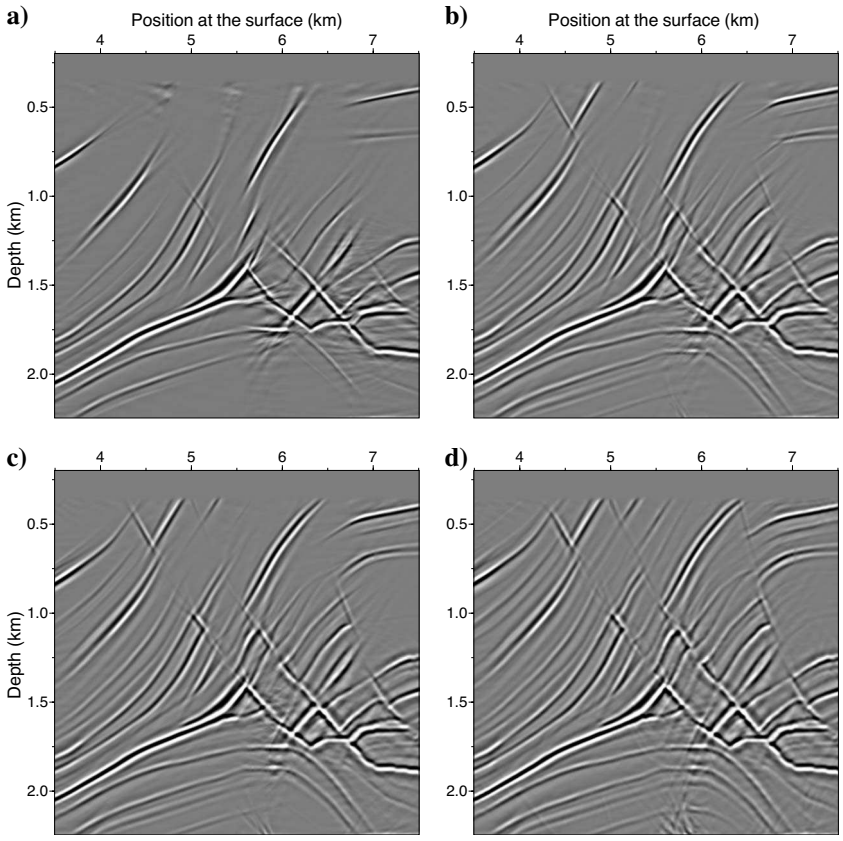


Figure 14. Predicted images obtained with (a) 4%, (b) 12%, (c) 20%, and (d) 40% of the total number of curvelet coefficients.

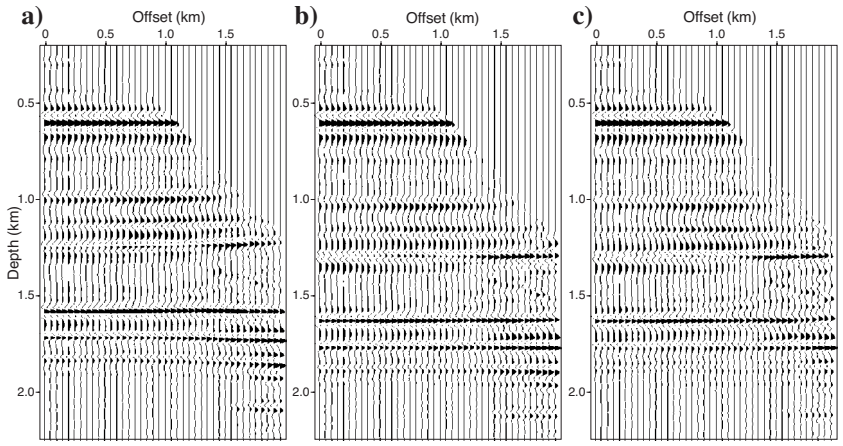


Figure 15. CIGs for position $x = 6000$ m; (a) initial velocity model v_{ini}, (b) the reference model v_{ref}, and (c) after prediction in the curvelet domain. The value v_{ref} corresponds to the exact velocity model.

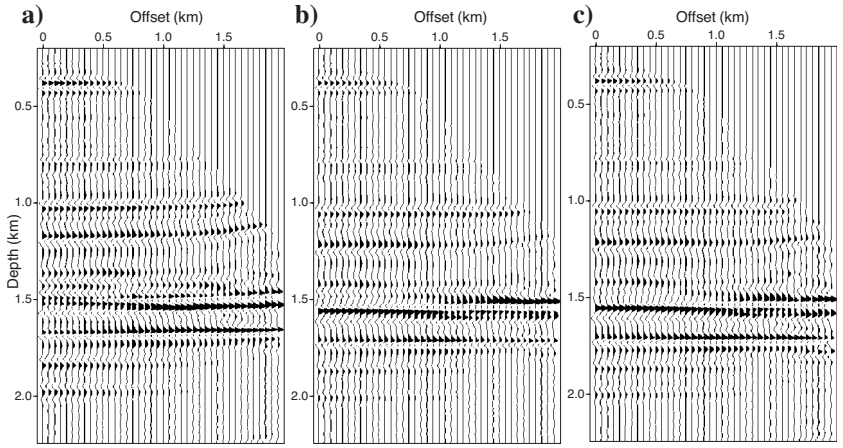


Figure 16. CIGs for $x = 6400$ m; (a) the initial velocity model v_{ini}, (b) the reference model v_{ref}, and (c) after prediction in the curvelet domain.

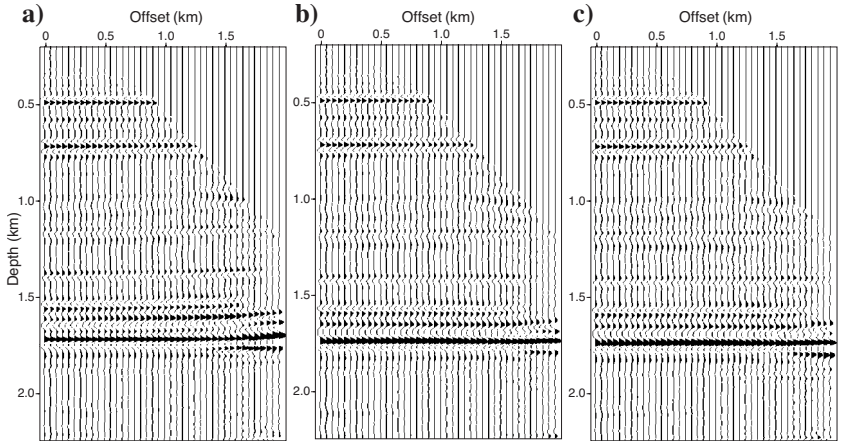


Figure 17. CIGs for $x = 6800$ m; (a) the initial velocity model v_{ini}, (b) the reference model v_{ref}, and (c) after prediction in the curvelet domain.

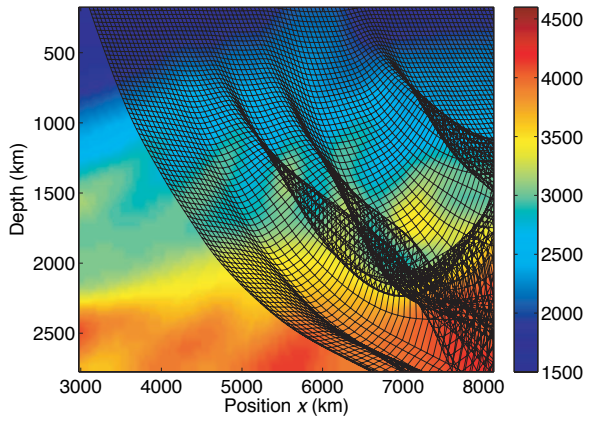


Figure 18. Rayfield superimposed on the exact velocity model v_{ref} from Figure 7a.

in two different velocity models. In our approach, we show how a curvelet is distorted by the demigration/migration operator. Such a curvelet contains the notion of image point as in Adler (2002) but is also associated with a local dip angle and a frequency content. In other words, our derivation shows how the migrated image is distorted locally. Moreover, we derive an expression for the expected curvature.

CONCLUSIONS

As for seismic depth migration with the use of curvelets, the demigration/migration in two different velocity models is reduced to a map migration; energy is propagated only along a single direction related to the curvelet. In the case of demigration/migration, the expected curvature is expressed as a linear function of the velocity perturbation. For small local velocity changes, curvelets sparsify the demigration/migration operator. Curvelets also efficiently decompose seismic images, so they are well suited to estimate the sensitivity of a migrated section with respect to the velocity model.

The analysis is valid for smooth, heterogeneous models. The extension to media with interfaces would enable applications on subsalt imaging. However, more developments are needed to handle rough interfaces or nonsmooth models because the derivations proposed here are based on the classical high-frequency approximation.

ACKNOWLEDGMENTS

The authors would like to thank Shell E&P for partly funding the project and for permission to publish this work. The authors thank Emmanuel Candès (Caltech) for providing them with his curvelet transform. They are grateful to Fons ten Kroode, Frans Kuiper (Shell E&P), Mark Noble, and Pascal Podvin (École des Mines de Paris) for fruitful discussions. They are indebted to the three reviewers for largely improving and clarifying the initial manuscript.

APPENDIX A

DEMIGRATION/MIGRATION OF CURVELETS

To simplify equation 7, we first apply an inverse 2D Fourier transform of c_μ, where

$$c_\mu(x,z) = \frac{1}{2\pi}\iint \mathrm{d}k_x\mathrm{d}k_z\hat{c}_\mu(k_x,k_z)\cdot e^{-ik_x\cdot x}\cdot e^{-ik_z\cdot z}. \tag{A-1}$$

We then change the integration order, leading to

$$R(a,b) = \frac{1}{2\pi}\iint \mathrm{d}k_x\mathrm{d}k_z I, \tag{A-2}$$

where

$$I = \hat{c}_\mu(k_x,k_z)\cdot\iint \mathrm{d}m\mathrm{d}\xi e^{-ik_x\cdot x[\tau(a,b,m,v+\delta v),m,\xi,v]} \cdot e^{-ik_z\cdot z[\tau(a,b,m,v+\delta v),m,\xi,v]}. \tag{A-3}$$

Curvelets are localized by construction near a wedge in the wavenumber domain, meaning that $\hat{c}_\mu(k_x,k_z) = 0$, except when k_x/k_z is close to $\tan\xi_\mu$ (Figure 1b). We thus replace k_x by $k_z\cdot\tan\xi_\mu$ in equation A-3 and obtain

$$I = \hat{c}_\mu(k_x,k_z)\cdot\iint \mathrm{d}m\mathrm{d}\xi e^{-ik_z\cdot\varphi(\tan\xi_\mu,a,b,m,\xi,v,v+\delta v)}, \tag{A-4}$$

where

$$\varphi(\tan\xi_\mu,a,b,m,\xi,v,v+\delta v) = z[\tau(a,b,m,v+\delta v),m,\xi,v] + \tan\xi_\mu\cdot x[\tau(a,b,m,v+\delta v),m,\xi,v]. \tag{A-5}$$

Now that φ does not depend on k_z any more, we can apply twice the stationary phase approximation (Bleistein, 1984) to evaluate I. The derivatives of φ with respect to ξ and m give

$$0 = \frac{\partial\varphi}{\partial\xi} = \frac{\partial z}{\partial\xi} + \tan\xi_\mu\frac{\partial x}{\partial\xi}, \tag{A-6}$$

$$0 = \frac{\partial\varphi}{\partial m} = \frac{\partial z}{\partial\tau}\frac{\partial\tau}{\partial m} + \frac{\partial z}{\partial m} + \tan\xi_\mu\frac{\partial x}{\partial\tau}\frac{\partial\tau}{\partial m} + \tan\xi_\mu\frac{\partial x}{\partial m}. \tag{A-7}$$

Equation A-6 determines the stationary dip $\xi = \xi_\mu$. Equation A-7 provides a specular midpoint position m_μ.

We now replace once more $\tan\xi_\mu$ by k_x/k_z and obtain

$$I \simeq \hat{c}_\mu(k_x,k_z)\cdot e^{-ik_x\cdot x[\tau(a,b,m_\mu,v+\delta v),m_\mu,\xi_\mu,v]} \cdot e^{-ik_z\cdot z[\tau(a,b,m_\mu,v+\delta v),m_\mu,\xi_\mu,v]}, \tag{A-8}$$

where we keep only the terms contributing to kinematic effects. As in Douma and de Hoop (2007) for the migration case, equation A-2

can now be evaluated as an inverse Fourier transform, leading to

$$R(a,b) \simeq c_\mu(x[\tau(a,b,m_\mu,v+\delta v),m_\mu,\xi_\mu,v], z[\tau(a,b,m_\mu,v+\delta v),m_\mu,\xi_\mu,v]). \quad \text{(A-9)}$$

In the following applications, the velocity perturbation is small compared to the background velocity model. We can further simplify equation A-9, yielding

$$R(a,b) \simeq c_\mu\Big(x[\tau(a,b,m_\mu,v),m_\mu,\xi_\mu,v] + \frac{\partial x}{\partial \tau}\frac{\partial \tau}{\partial v}\delta v, z[\tau(a,b,m_\mu,v),m_\mu,\xi_\mu,v] + \frac{\partial z}{\partial \tau}\frac{\partial \tau}{\partial v}\delta v\Big). \quad \text{(A-10)}$$

If $\delta v = 0$, then $R(a,b) = c_\mu(a,b)$ by definition of the migration operator that can be seen as the inverse of the demigration operator. The identity would not be true in the case of limited acquisition; but for a well-illuminated subsurface, it means that $x[\tau(a,b,m_\mu,v),m_\mu,\xi_\mu,v] = a$ and $z[\tau(a,b,m_\mu,v),m_\mu,\xi_\mu,v] = b$. We thus finally obtain equation 8.

APPENDIX B

DERIVATION OF CURVATURE

First approach

The curvature at any point along a 2D curve is defined as the rate of change in angle direction of the contour, as function of the arc length (Farin, 1993). For the demigration/migration process, the central point of the curvelet at position (x_i,z_i) with the angle ξ_i will be mapped at (x_f,z_f) with the angle ξ_f (Figure B-1). Another point $(x_i + \cos\xi_i\delta q, z_i + \sin\xi_i\delta q)$ that belongs to the main part of the curvelet will be mapped with a first-order approximation into a new position $(x_f + \cos\xi_f\delta l, z_f + \sin\xi_f\delta l)$ with the angle $\xi_f + \delta\xi_f$. The curvature value is thus

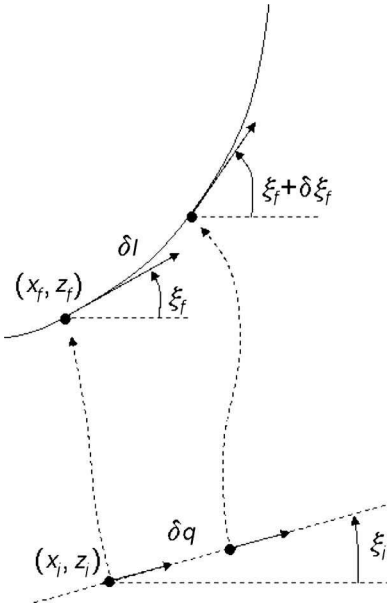


Figure B-1. Definition of the curvature. The two points (x_i,z_i) and $(x_i + \cos\xi_i\delta q, z_i + \sin\xi_i\delta q)$ are mapped after demigration/migration along a curve (solid line). The arc length between the new positions is denoted by δl.

$$\gamma = \frac{\dfrac{\partial \xi_f}{\partial q}}{\dfrac{\partial l}{\partial q}}. \quad \text{(B-1)}$$

By definition of ξ_f, we also have

$$\frac{\partial l}{\partial q} = \frac{1}{\cos\xi_f}\frac{\partial x_f}{\partial q}$$

(Figure B-1) and thus

$$\gamma = \frac{\dfrac{\partial \sin\xi_f}{\partial q}}{\dfrac{\partial x_f}{\partial q}}. \quad \text{(B-2)}$$

In the case of demigration/migration, q indicates that we consider an initial plane wave along direction ξ_i

$$\frac{\partial \cdot}{\partial q} = \frac{\partial \cdot}{\partial x} + \frac{\partial \cdot}{\partial z}\cdot\tan\xi_i. \quad \text{(B-3)}$$

Moreover, $\partial x_f/\partial q$ should be read $\partial x_f/\partial q(v+\delta v)$, which for small velocity perturbations is

$$\frac{\partial x_f}{\partial q} = \frac{\partial x_f}{\partial q}(v) + \frac{\partial^2 x_f}{\partial q\partial v}\delta v = 1 + \frac{\partial^2 x_f}{\partial q\partial v}\delta v \simeq 1. \quad \text{(B-4)}$$

The first term in expression B-4 is equal to one because the same image is obtained if there is no velocity perturbation. Similarly, the numerator can be simplified as the final angle; ξ_f should not change if $\delta v = 0$:

$$\frac{\partial \sin\xi_f}{\partial q} = \frac{\partial \sin\xi_f}{\partial q}(v) + \frac{\partial^2 \sin\xi_f}{\partial q\partial v}\delta v = 0 + \frac{\partial^2 \sin\xi_f}{\partial q\partial v}\delta v. \quad \text{(B-5)}$$

By combining equations B-3–B-5, we obtain the final formula (equation 9). The derivation for the migration case is very similar: q indicates that the derivative has to be done along the initial direction of the curvelet p_μ:

$$\frac{\partial \cdot}{\partial q} = \frac{\partial \cdot}{\partial m} + \frac{\partial \cdot}{\partial t}\cdot p_\mu. \quad \text{(B-6)}$$

By combining equations B-2 and B-6, we obtain the final formula (equation 10).

Second approach

To check the validity of equations 9 and 10, we consider another approach where the curvatures are computed by integrating dynamic ray quantities along a central ray as developed in Červený (2001). For demigration/migration in for 2D zero offset, a first ray is shot

from the subsurface toward the surface in the initial velocity model (Figure 6). Using the final conditions as new initial conditions, a new ray is back-propagated in the perturbed velocity model.

We use the same notations as in Červený (2001). The curvature value γ is obtained by considering a quadratic expansion of the traveltime in ray-centered coordinates (Červený 2001, his equation 4.1.77), where P and Q are integrated numerically according to the dynamic ray-tracing system of Červený's (2001) equation 4.1.76. It follows that the curvature from equation B-1 can be expressed as $\gamma = vP/Q$. The initial values satisfy the conditions for a local plane wave: $P(0) = 0$ and $Q(0) = 1$.

For the migration case and again for zero-offset input data, the derivation is very similar. The only difference is that the integration along the ray is done from the surface downward without the back-propagation step.

APPENDIX C

MODIFIED DEMIGRATION/MIGRATION TRANSFORM

In this appendix, we describe how the demigration/migration operator (equation 8) is implemented in the curvelet domain.

Restriction to a shift, a rotation, and a dilation

The new coordinates (a,b) can be expressed formally as a function of the original positions (x,z) with $a = x + f(x,z)\delta v$ and $b = z + g(x,z)\delta v$, where $f = \partial x/\partial\tau \cdot \partial\tau/\partial v$ and $g = \partial z/\partial\tau \cdot \partial\tau/\partial v$. To implement in the curvelet domain, we exploit the curvelet structure and thus restrict the transformation to a combination of a shift, a rotation, and a dilation. These operations are indeed defined naturally in the radial curvelet construction. For a curvelet centered at (x_i,z_i), we impose the new center to be $(a_i,b_i) = (x_i + f(x_i,z_i)\delta v, z_i + g(x_i,z_i)\delta v)$. We want to find a rotation angle θ and a dilation factor λ such that the combination of the rotation and the dilation, in addition to the shift, mimics the demigration/migration operator. In other words, we want to find optimal parameters (θ,λ) such that

$$\begin{bmatrix} a - a_i \\ b - b_i \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x - x_i \\ z - z_i \end{bmatrix}. \tag{C-1}$$

We exploit the fact that the curvelet is fairly local in the spatial domain. For small $(\delta x, \delta z) = (x - x_i, z - z_i)$, we get

$$\begin{bmatrix} 1 + \dfrac{\partial f}{\partial x}\delta v - \lambda\cos\theta & \dfrac{\partial f}{\partial z}\delta v - \lambda\sin\theta \\ \dfrac{\partial g}{\partial x}\delta v + \lambda\sin\theta & 1 + \dfrac{\partial g}{\partial z}\delta v - \lambda\cos\theta \end{bmatrix} \begin{bmatrix} \delta x \\ \delta z \end{bmatrix} = 0. \tag{C-2}$$

Equation C-2 can be read as $A \cdot [\delta x/\delta z] = 0$. The unknowns θ and λ are obtained by minimizing the norm of the matrix A in a least-squares sense. For small velocity perturbations, it gives

$$\begin{bmatrix} \tan\theta \\ \lambda - 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \dfrac{\partial f}{\partial z} - \dfrac{\partial g}{\partial x} \\ \dfrac{\partial f}{\partial x} + \dfrac{\partial g}{\partial z} \end{bmatrix}\delta v. \tag{C-3}$$

For $\delta v = 0$, we obviously have $\theta = 0$ and $\lambda = 1$.

Practical implementation

Instead of performing the general transform (equation 8), we restrict the demigration/migration to a shift, a rotation, and a dilation. The f and g functions are evaluated for each curvelet in a two-step approach. First, as in Billette and Lambaré (1998) and Chauris et al. (2002b), for a fixed-velocity model, two rays are shot from the initial position (x_i,z_i,ξ_i) up to the surface. With an iterative minimization process, the opening angle α at the starting point is modified to obtain the correct final offset at the surface. The final rays define the specular rays and their associated $(s,r,\tau,p_m = p_s + p_r)$ values (Figure 6).

In a second phase, paraxial quantities and the Fréchet derivatives are computed as in Billette and Lambaré (1998). For a given velocity perturbation δv, the initial conditions will be perturbed by $(\delta x, \delta z, \delta z, \delta\alpha)$, such that (s,r,τ,p_m) remain fixed In other terms,

$$\begin{bmatrix} \delta s \\ \delta r \\ \delta\tau \\ \delta p_m \end{bmatrix} = 0$$

$$= \begin{bmatrix} \dfrac{\partial s}{\partial x} & \dfrac{\partial s}{\partial z} & \dfrac{\partial s}{\partial\xi} & \dfrac{\partial s}{\partial\alpha} \\ \dfrac{\partial r}{\partial x} & \dfrac{\partial r}{\partial z} & \dfrac{\partial r}{\partial\xi} & \dfrac{\partial r}{\partial\alpha} \\ \dfrac{\partial\tau}{\partial x} & \dfrac{\partial\tau}{\partial z} & \dfrac{\partial\tau}{\partial\xi} & \dfrac{\partial\tau}{\partial\alpha} \\ \dfrac{\partial p_m}{\partial x} & \dfrac{\partial p_m}{\partial z} & \dfrac{\partial p_m}{\partial\xi} & \dfrac{\partial p_m}{\partial\alpha} \end{bmatrix} \cdot \begin{bmatrix} \delta x \\ \delta z \\ \delta\xi \\ \delta\alpha \end{bmatrix} + \begin{bmatrix} \dfrac{\partial s}{\partial x} \\ \dfrac{\partial r}{\partial z} \\ \dfrac{\partial\tau}{\partial\xi} \\ \dfrac{\partial p_m}{\partial\alpha} \end{bmatrix} \cdot \delta v. \tag{C-4}$$

A similar approach has been developed by Chauris et al. (2002b). The resolution of equation C-4 provides the quantities $(f,g) = (\delta x, \delta z)$. The derivatives of f and g needed for θ and λ are performed numerically by considering neighboring curvelets.

Shannon interpolation

Using radial curvelet decomposition (Figure 3), we can implement a rotation in the curvelet domain through the Shannon interpolation scheme (see, for example, Zayed, 1993). All filters G_k (equation 3) can de deduced from a reference filter G_1 with $G_k(\theta)$

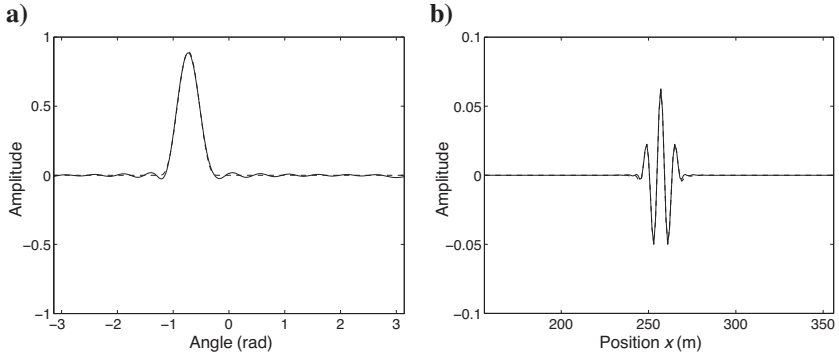


Figure C-1. Interpolated (dashed line) and exact (solid line) filters $G_k(\theta - d\theta)$ for $d\theta = 0.25\Delta\theta$ in (a) the wavenumber domain and (b) after reconstruction in the spatial domain. The interpolated filter is expressed as a weighted summation according to equation C-5.

$= G_1(\theta - \Delta\theta(k-1))$, where $\Delta\theta = \pi/4N$. For any arbitrary value $d\theta$, the shifted filter is approximated by

$$G_k(\theta - d\theta) \simeq \sum_p \beta_{p,k}(d\theta) G_p(\theta), \qquad \text{(C-5)}$$

where $\beta_{p,k} = \text{sinc}(\pi \cdot [d\theta/\Delta\theta - p + k])$ are the Shannon weights.

Figure C-1 illustrates that the approximation is good enough for the application. These weights are the new curvelet coefficients after rotation. If there is an integer p such that $d\theta/\Delta\theta = p$, then $\beta_{p,k} = 0$ except for $p = k$, where $\beta_{k,k} = 1$. In that case, the rotation is a change of coordinates in the curvelet domain. The same approach is implemented for the shift and the dilation in the curvelet domain.

REFERENCES

Adler, F., 2002, Kirchhoff image propagation: Geophysics, **67**, 126–134.

Billette, F., and G. Lambaré, 1998, Velocity macro model estimation from seismic reflection data by stereotomography: Geophysical Journal International, **135**, 671–680.

Bleistein, N., 1984, Mathematical methods for wave phenomena: Academic Press Inc.

Bleistein, N., J. K. Cohen, and F. G. Hagin, 1987, Two and one-half dimensional Born inversion with an arbitrary reference: Geophysics, **52**, 26–36.

Boschetti, F., and L. Moresi, 2001, Interactive inversion in geosciences: Geophysics, **66**, 1226–1234.

Candès, E., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528.

Candès, E., L. Demanet, D. Donoho, and L. Ying, 2006, Fast discrete curvelet transform: SIAM Multiscale Modeling and Simulation, **5**, 861–899.

Candès, E., and D. Donoho, 2004, New tight frames of curvelets and optimal representations of objects with C^2 singularities: Communications on Pure and Applied Mathematics, **57**, 219–266.

Červený, V., 2001, Seismic ray theory: Cambridge University Press.

Chauris, H., 2006, Seismic imaging in the curvelet domain and its implication for the curvelet design: 76th Annual International Meeting, SEG, Expanded Abstracts, 2406–2410.

Chauris, H., M. S. Noble, G. Lambaré, and P. Podvin, 2002a, Migration velocity analysis from locally coherent events in 2D laterally heterogeneous media, Part I, Theoretical aspects: Geophysics, **67**, 1202–1212.

———, 2002b, Migration velocity analysis from locally coherent events in 2D laterally heterogeneous media, Part II, Applications on synthetic and real data: Geophysics, **67**, 1213–1224.

Daubechies, I., 1992, Ten lectures on wavelets: SIAM CBMS-NFS Lecture Notes 61.

Do, M. N., 2001, Directional multiresolution image representations: Ph.D. thesis, Swiss Federal Institute of Technology, Lausanne.

Do, M. N., and M. Vetterli, 2003, Contourlet *in* G. V. Welland, ed., Beyond wavelets: Academic Press Inc., 83–106.

Douma, H., and M. V. de Hoop, 2005, On common-offset prestack time migration with curvelets: 75th Annual International Meeting, SEG, Expanded Abstracts, 2009–2012.

———, 2007, Leading-order seismic imaging using curvelets: Geophysics, **72**, no. 6, S231–S248.

Farin, G., 1993, Curves and surfaces for computer aided geometric design, A practical guide: Academic Press Inc.

Farra, V., and R. Madariaga, 1987, Seismic waveform modeling in heterogeneous media by ray perturbation theory: Journal of Geophysical Research, **92**, 2697–2712.

Fomel, S., 2003, Time-migration velocity analysis by velocity continuation: Geophysics, **68**, 1662–1672.

Gray, S., S. Cheadle, and B. Law, 2000, Depth model building by interactive manual tomography: 70th Annual International Meeting, SEG, Expanded Abstracts, 914–917.

Herrmann, F., 2003, Optimal imaging with curvelets: 73rd Annual International Meeting, SEG, Expanded Abstracts, 997–1000.

Hubral, P., M. Tygel, and J. Schleicher, 1996, Seismic image wave: Geophysical Journal International, **125**, 431–442.

Mallat, S., 1989, A theory for multiresolution signal decomposition, The wavelet representation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**, 674–693.

Meyer, Y., 1993, Wavelets: Algorithms and applications, SIAM.

Murphy, G. E., N. D. Whitmore, and M. A. Thornton, 1993, Interactive prestack depth migration: 63rd Annual International Meeting, SEG, Expanded Abstracts, 899–902.

Nolan, C. J., and W. W. Symes, 1996, Imaging and coherency in complex structure: 66th Annual International Meeting, SEG, Expanded Abstracts, 359–363.

Sava, P., 2003, Prestack residual migration in the frequency domain: Geophysics, **68**, 634–640.

Simoncelli, E. P., W. T. Freeman, E. H. Adelson, and D. J. Heeger, 1992, Shiftable multiscale transform: IEEE Transactions on Information Theory, **38**, 587–607.

Smith, H., 1998, A parametrix construction for wave equation with $C^{1,1}$ coefficients: Annales de l'Institut Fourier, **48**, 797–835.

Stolk, C. C., and W. W. Symes, 2004, Kinematic artifacts in prestack depth migration: Geophysics, **69**, 562–575.

Stolt, R. H., 1996, A prestack residual time migration operator: Geophysics, **61**, 605–607.

ten Kroode, A. P. E., D. J. Smit, and A. R. Verdel, 1998, A microlocal analysis of migration: Wave Motion, **28**, 149–172.

Thierry, P., S. Operto, and G. Lambaré, 1999, Fast 2-D ray + Born migration/inversion in complex media: Geophysics, **64**, 162–181.

Versteeg, R. J., and G. Grau, eds., 1991, The Marmousi experience: Proceedings, 1990 EAEG workshop on Practical Aspects of Seismic Data Inversion.

Zayed, A. I., 1993, Advances in Shannon's sampling theory: CRC Press.

Uniform Discrete Curvelet Transform

Truong T. Nguyen and Hervé Chauris

***Abstract*— An implementation of the discrete curvelet transform is proposed in this work. The transform is based on and has the same order of complexity as the Fast Fourier Transform (FFT). The discrete curvelet functions are defined by a parameterized family of smooth windowed functions that satisfies two conditions: (i)2π periodic, (ii) their squares form a partition of unity. The transform is named the Uniform Discrete Curvelet Transform (UDCT) because the centers of the curvelet functions at each resolution are positioned on a uniform lattice. The forward and inverse transform form a tight frame, in the sense that they are the exact transpose of each other. Generalization to M dimensional version of the UDCT is also presented. The novel discrete transform has several advantages over existing transforms, such as lower redundancy ratio, hierarchical data structure and ease of implementation.**

***Index Terms*— Contourlet, Curvelet, Directional filter bank, Directional decomposition, Multidimensional filter bank, Multiresolution representation, Wavelet**

I. Introduction

During the last two decades, there have been a lot of research activities on new mathematical and computational tools for multiresolution data representation. The wavelet transform, arising from a particular question in geophysics, decomposes a function on the real line to a sum of local wavelike functions at multiple scales. It is shown that the continuous wavelet functions and their discrete counterpart implemented by regular filter banks are optimal representations of one dimensional piece-wise smooth signals [1]. However, the direct extension of a wavelet to two dimensions by the tensor product of two one-dimensional wavelets is no longer optimal for representing a signal that has features along smooth curves [2]. There have been many 2-D transforms with directional basis that can better represent this type of signal. Without being exhaustive, we can list several examples of directional, discrete and non-adaptive transforms such as the steerable pyramid [3], the complex wavelet [4], [5], discrete curvelet [6], shearlet transforms [7], and contourlet transform [8].

Recently, Candès and Donoho constructed the curvelet transform [9], and proved that it is an essentially optimal representation of two-variable functions, which are smooth except at discontinuities along $\boldsymbol{C}^2$ (twice differentiable) curve. The nonlinear approximation of a function f, $f_M^{(c)}$, reconstructed by M largest curvelet coefficients has an asymptotic decay rate of $\|f - f_M^{(c)}\|^2 \leq CM^{-2}(\log_2 M)^3$. This decay rate of the approximation error is a significant theoretical improvement compared to those by wavelets or Fourier coefficients, which are $O(M^{-1})$ and $O(M^{-1/2})$, respectively [1]. The 2-D curvelet transform decomposes a signal into a sum of basis functions that represent local waves with strong directionality. The effective support of these curvelet functions are highly anisotropic at fine scale, following a parabolic scaling rule: $length^2 \sim width$.

It has been proved that curvelet representation for wave propagation is optimally sparse [10]. It is argued that the representation of seismic data by the curvelet transform is also optimally sparse and a series of applications can be found in the area of seismic imaging [11], [12], [13]. The curvelet transform has also found applications in other areas of image processing [14], [15].

In the last few years, several discrete curvelet and curvelet-like transforms have been proposed. They can be divided into discrete transforms based on the Fast Fourier Transform (FFT) [6], [16], or based on filter bank (FB) implementations [8], [17], [18].

The major contribution of this work is a new discrete curvelet transform that uses the ideas of both FFT-based discrete curvelet transform and filter-bank based contourlet transform. The motivation of the new transform is from the difficulties we encountered while trying to use curvelet transform in industrial applications. The new uniform discrete curvelet transform (UDCT) is implemented by the FFT algorithm, but it is designed as a multiresolution filter bank. By this way, we are able to take advantage of the two methods. For example, the new curvelet basis functions have excellent frequency responses, which is not possible with the contourlet transform. The UDCT has significantly lower redundancy compared to other FFT-based discrete curvelet transforms [6], [16], especially in high-dimension version of the transform. This makes the UDCT very practical in industrial applications.

Paper outline. In the next section the notations used in this paper are introduced. Since the UDCT transform is interpreted as a multiresolution FB, this section also includes a short review of some formulae for multidimensional multirate systems. The continuous curvelet transform is reviewed in Section III-A. This presentation closely follows the construction of 'second-generation' curvelet transform in [19]. We discuss the typical examples of the two approaches in discrete curvelet (or curvelet-like) transform, the fast discrete curvelet transform (FDCT) [6] and the contourlet transform [8], in Section III-B. After an analysis of the pros and cons of the two discrete transforms, we construct a discrete curvelet transform that combines the advantages of both methods in the rest of the paper. In Section IV, a parameterized family of smooth windowed functions is defined. These functions will later be used as curvelet functions in the frequency domain. They are all $\mathbf{2}\pi$ periodic and their squares form a partition of unity, which means that the sum of squares of all functions is equal to 1 on the 2-D plane. The set of windows is then used in Section VI to define a multi-resolution multi-directional FB.

Truong T. Nguyen is with the R&D department, Fugro Seismic Imaging, Swanley, UK. Hervé Chauris is with the Centre de Géoscience , Mines ParisTech, UMR Sisyphe 7619, Fontainebleau, 77300 France (email: truong.nguyen@fugro-fsi.com, herve.chauris@mines-paristech.fr).

The discrete transform implemented by the FB is formally defined in Section VI. The proposed transform, namely the uniform discrete curvelet transform, is generalized to M-D in Section VII. The complexity of the transform and various implementation aspects are discussed in Section VIII. We present several numerical experiments using the UDCT in Section IX and conclude the paper in Section X.

II. Preliminary

Notation. In this paper, common multirate signal processing notations as in [20] are used. However, we also try to keep the same notation as in other works about discrete curvelet transforms [6], [8], [21]. Since curvelet decomposition is a two dimensional transform, the functions in this paper are typically two-variable. We use bold letters and numbers to signify vectors, and bold capital letters to signify matrices. For example, $\boldsymbol{I} = \text{diag}\{1, 1\}$ is the 2×2 identity matrix. Bold letter variables signify functions of two or more variables. We denote by t the spatial variable in $\mathbb{R}$ and $\boldsymbol{t} = (t_1, t_2)$ the spatial variable in $\mathbb{R}^2$. In the Fourier domain, $\boldsymbol{\xi} = (\xi_1, \xi_2)$ is the 2-D frequency variable. $\boldsymbol{n} = (n_1, n_2)$ are coordinates of 2-D discrete functions (or signal). Similarly, $\boldsymbol{\omega} = (\omega_1, \omega_2)$ is discrete frequency variable. By an abuse of notation, r and θ are used as the polar coordinate of the $\boldsymbol{\xi}$ or $\boldsymbol{\omega}$ frequency planes. Throughout the paper, bold letter variables are employed frequently to compress mathematical expressions. For example the following expressions are equivalent.

$$\sum_{n_1,n_2=-\infty}^{\infty} w_0(\omega_1 + 2n_1\pi, \omega_2 + 2n_2\pi) = \sum_{\boldsymbol{n}\in\mathbb{Z}^2} w_0(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) \tag{1}$$

Small letters and capital letters denote a discrete filter in the spatial and frequency domains, for example $f(\boldsymbol{n})$ and $F(\boldsymbol{\omega})$. However, the hat symbol denotes the Fourier transform of continuous functions. For example the Fourier transform of $\varphi(\boldsymbol{t})$ is $\hat{\varphi}(\boldsymbol{\xi})$.

In the continuous domain, a curvelet basis function can be indexed by scale j, direction l and position $\boldsymbol{k}$. Therefore, the curvelet index set is $\mu = (j, l, \boldsymbol{k})$. However, the index for the new discrete curvelet transform is $\mu = (i, j, l, \boldsymbol{k})$ where i is the index for symmetric/antisymmetric property of the curvelet functions.

Grid and lattice. A *grid* $\Lambda(\boldsymbol{M})$ is a set of points in a 2-D space generated by a nonsingular matrix $\boldsymbol{M}$, defined by

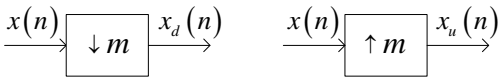
$$\Lambda(\boldsymbol{M}) = \{\boldsymbol{M}\boldsymbol{n}\}, \boldsymbol{n} \in \mathbb{Z}^2 \tag{2}$$

In this paper, we use the word *lattice* to denote a grid in case all elements of $\boldsymbol{M}$ are integer.

Rotation and shearing. A rotation of a 2-D function $f(\boldsymbol{t})$ by an angle θ is $f(\boldsymbol{R}_\theta \boldsymbol{t}) = f(\cos(\theta)t_1 + \sin(\theta)t_2, -\sin(\theta)t_1 + \cos(\theta)t_2)$. When the matrix $\boldsymbol{R}_\theta$ is replaced by an upper or lower 2×2 matrix with diagonal elements equal to 1, the new function is a shearing version of $f(\boldsymbol{t})$.

A. Discrete signal decimation and upsampling, and filter bank

A discrete filter bank is a set of cascading discrete filters, decimation and upsampling blocks (see Fig. 10 as an example). If the elements of a FB are designed in such a way that the output $y(\boldsymbol{n})$ is exactly $x(\boldsymbol{n})$, then the FB corresponds to a discrete transform. Since we will use the family of windows $u_l^2(\omega_1, \omega_2)$ as discrete filters in frequency domain of a FB, it is important to establish the equivalent operation in frequency domain when a discrete signal passes through the decimation and upsampling blocks as in Fig. 2.



(a) (b)

Fig. 2. Two basic operations in multirate signal processing: down and up sampling .

The relationship between a decimated signal $x_d(n)$ and the original signal $x(n)$ in the spatial domain is $x_d(mn) = x(n)$, where m is a natural number. The relationship between $X_d(\omega)$ and $X(\omega)$ is

$$X_d(\omega) = \frac{1}{m} \sum_{k=0}^{m-1} X\left(\frac{\omega - 2k\pi}{m}\right) \tag{3}$$

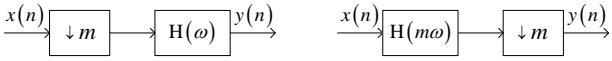
If a signal $x(n)$ is upsampled by m to produce $x_u(n) = x(n/m)$ if n/m is an integer, and $x_u(n) = 0$ otherwise, then in the frequency domain

$$X_u(\omega) = X(m\omega) \tag{4}$$

The two equations (3) and (4) are for 1-D signal. If a 2-D signal $x(\boldsymbol{n})$ is downsampled and then upsampled by 2 in two dimensions (or by $2\boldsymbol{I}$), the resulting signal $x_{du}(\boldsymbol{n}) = x(\boldsymbol{n})$ when $n \in \Lambda(2\boldsymbol{I})$, otherwise $x_{du}(\boldsymbol{n}) = 0$. Following (3) and (4), in frequency domain

$$X_{du}(\boldsymbol{\omega}) = \frac{1}{4} \sum_{\boldsymbol{n}\in\mathcal{S}} X(\boldsymbol{\omega} + \boldsymbol{n}\pi) \tag{5}$$

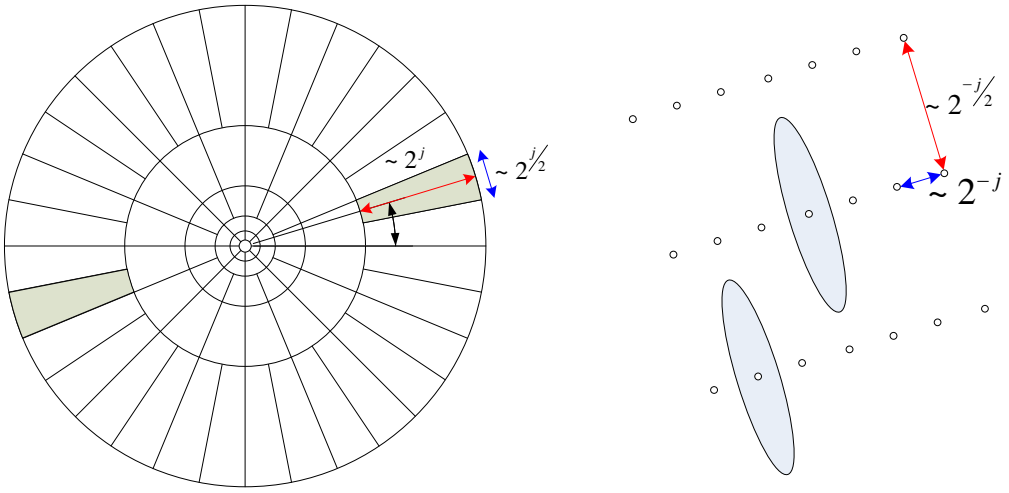
where $\mathcal{S} = \{(0,0), (0,1), (1,0), (1,1)\}$. Visually, $X_{du}(\boldsymbol{\omega})$ is a scaled sum of four copies of $X(\boldsymbol{\omega})$ shifted by $\boldsymbol{n}\pi$, $\boldsymbol{n} \in \mathcal{S}$. Fig. 9(b) and (c) show the support of windows $u_2(\boldsymbol{\omega})$ and its shifted copies by $\boldsymbol{n}\pi$.



(a) (b)

Fig. 3. The two structures are equivalent: the same $x(n)$ will produce the same $y(n)$.

Another result in filter bank theory that will be needed is the equivalent relation between the two structures in Fig. 3. Signal $y(n)$ is obtained from $x(n)$ by decimation by m and filtering by $H(\omega)$. Using Equation (3), it is easy to show that the operation is equivalent to filtering by $H(m\omega)$ followed by decimation by m. This relationship is called 'noble property' in signal processing literature.



(a) (b)

Fig. 1. Curvelet tiling of space and frequency (a) Tiling of the frequency plane, (b) The spatial grid of the curvelet at a given scale and orientation [19].

III. The Curvelet and Contourlet Transforms

A. The second generation continuouos curvelet transform

The first generation of curvelet transform is a ridgelet transform of bandpass filtering images [9]. It is not clear what the basis functions are in this construction of curvelet transform. The recent curvelet transform described in [19], sometimes called 'second-generation curvelet', is considerably simpler and easier to use. The main idea of this construction is to decompose any 2-D function into spaces that are strictly bandpass in frequency domain. The support shapes of these function spaces are concentric wedges (see Fig. 1(a)). The curvelet coefficients for each scale-direction wedges are estimated as the inner product of the given function and the band-limited curvelet function centered on a grid that is inversely proportional to the wedge-shape support of the curvelet in frequency domain (see Fig. 1(b)). Assume that we have two smooth functions W and V, satisfying the 'admissibility' condition

$$\sum_{j=-\infty}^{\infty} W^2(2^{-j}r) = 1, \tag{6}$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1. \tag{7}$$

For each $j > j_0$, the frequency window U_j for the curvelet function is given by

$$U_j(r,\theta) = 2^{-3j/4} W(2^{-j}r) V\left(\frac{2^{\lfloor j/2 \rfloor}\theta}{2\pi}\right) \tag{8}$$

The real-valued curvelet has symmetry window function in frequency domain

$$\hat{\varphi}_j(r,\theta) = U_j(r,\theta) + U_j(r,\theta+\pi) \tag{9}$$

This window is used to define the curvelet function at the first direction of scale j, $\hat{\varphi}_{j,1}(\boldsymbol{\xi}) = \hat{\varphi}_j(r,\theta)$. In the frequency plane, each curvelet direction is indexed by $L = (j,l)$. At each scale j, the number of directions is $2^{\lfloor j/2 \rfloor}$. At scale j, the curvelet direction l is generated by rotating the curvelet window function $\hat{\varphi}_j(r,\theta)$ by a rotation angle $\theta_L = (l-1)\cdot 2\pi \cdot 2^{-\lfloor j/2 \rfloor}$, with $l = 1, 2, \ldots, 2^{\lfloor j/2 \rfloor}$.

The coarse scale curvelet function is defined by polar window $W(r)$ as follows $\hat{\varphi}_0(\boldsymbol{\xi}) = W(r)$. From the definition of $W(r)$ and $V(\theta)$ functions in (6) and (7), it can be shown that the function $\hat{\varphi}_0^2(\boldsymbol{\xi})$ and scaled version of $\hat{\varphi}_{j,l}^2(\boldsymbol{\xi})$ are summed up to one for all $\boldsymbol{\xi}$.

The scale j and angle (or direction) l determine the curvelet in frequency domain, or the shape of the curvelet function. The spatial function also has a translation parameter $\boldsymbol{k} \in \mathbb{Z}^2$. The translation index specifies the amount of translation $t_{\boldsymbol{k}}^{(j,l)}$ that belongs to a grid $\mathcal{M}_{j,l}$. The 2-D curvelet function of variables t_1 and t_2 at scale j and direction θ_L is given by

$$\varphi_{j,l,\boldsymbol{k}} = \varphi_j\left(R_{\theta_L}(\boldsymbol{t} - \boldsymbol{t}_{\boldsymbol{k}}^{(j,l)})\right) \tag{10}$$

where R_{θ_L} is rotation by an angle θ_L. The center of the curvelet function is on a rotated grid $\boldsymbol{t}_{\boldsymbol{k}}^{(j,l)} = R_{\theta_L}^{-1}(k_1 2^{-j}, k_2 2^{-j/2})$. An example of the frequency support of a curvelet and its position in the spatial domain is illustrated in Fig. 1.

Let the index parameter μ denote the set of three parameters $(j,l,\boldsymbol{k})$. The curvelet coefficient is the inner product between an element $f(t_1,t_2) \in L^2(\mathbb{R}^2)$ and a curvelet c_μ

$$c_\mu = \langle f(\boldsymbol{t}), \varphi_\mu(\boldsymbol{t}) \rangle = \int_{\mathbb{R}^2} f(\boldsymbol{t}) \varphi_\mu(\boldsymbol{t}) d\boldsymbol{t} \tag{11}$$

By definition, the curvelet functions are bandpass functions with compact support in frequency domain. Therefore, the convolution of a two-variable function $f(\boldsymbol{t})$ with a curvelet $\varphi_{j,l}(\boldsymbol{t})$ is also a bandpass function. The curvelet coefficients are obtained by sampling $f(\boldsymbol{t}) * \varphi_{j,l}(\boldsymbol{t})$ on a grid $\mathcal{M}_{j,l}$. If this grid is denser than the Shannon sampling ratio associated with the support in frequency domain of $\hat{f}(\boldsymbol{\xi})\hat{\varphi}_{j,l}(\boldsymbol{\xi})$, then one can recover the bandpass function from the curvelet coefficients c_μ. Moreover, because the sum of squares of the scale curvelet functions in frequency domain is equal to 1, the original

function can also be recovered:

$$f(\boldsymbol{t}) = \sum_{\mu \in \mathcal{M}} c_\mu \varphi_\mu(\boldsymbol{t}) \tag{12}$$

B. The fast discrete curvelet transform

Since the theory of the continuous curvelet transform promises an order of magnitude improvement over the wavelet decomposition in many important applications, there have been several implementations of the discrete curvelet transform to operate on digital data.

The Fast Discrete Curvelet Transform (FDCT) described in [6] with source code (CurveLab package) is available for academic use at www.curvelet.org. This software package actually contains two discrete implementations of the curvelet transform. The first is called Digital Curvelet Transform via Unequispaced FFT (USFFT), and the second is called Digital Curvelet Transform via wrapping. In this paper, the FDCT is referred to the wrapping FDCT because this is the version usually used in actual applications [22], [15], [23].

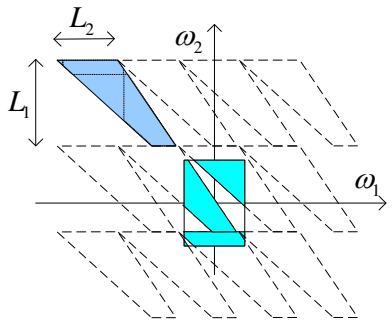


Fig. 5. After multiplication of FFT data with curvelet window, the data on a wedge-shaped support is mapped to a rectangle.

The wrapping FDCT implementation is based on the FFT algorithm. The data flow diagram of the forward and inverse wrapping FDCT are plotted in Fig. 4. The data are first transformed into the frequency domain by forward FFT. The transformed data are then multiplied with a set of window functions. The shape of these windows are defined according to the requirements of the ideal curvelet transform, such as the parabolic scaling rule. The curvelet coefficients are obtained by inverse FFT from windowing data. Since the window functions are zero except on support regions of elongated wedges, the inverse FFT algorithm can operate on regions much smaller than the original data. On the wrapping FDCT, the FFT coefficients on these regions are 'wrapped' or folded into rectangular shape before being applied to inverse FFT algorithm. Fig. 5 shows the mapping of a wedge-shaped region to several other regions, which lie inside of a rectangular.

The FFT-based FDCT transform can be viewed as a straightforward discretization of the continuous curvelet transform. Therefore the FDCT is faithful to the definition of its continuous counterpart. However, the FDCT implementation suffers from several practical problems that makes it difficult to use in industrial applications.

1) The curvelet functions in the frequency domain are designed to cover the whole frequency plane. This is not necessary when processing real-valued data. As a result, the FDCT has high-redundancy ratio, especially in 3-D.
2) In the FDCT, the curvelet functions are defined as the product of concentric square functions and sheared angle functions. However, the construction of the FDCT does not take into account specific aspects of discrete transform. For example, the window functions are not automatically $\mathbf{2}\pi$ periodic. Therefore, the FDCT has some difficulties in handling curvelet band at highest frequency or at junctions (Section 7 in [6]).
3) The basis of FDCT functions are located on non-integer grids. Moreover, these grids are different for each pair of resolution and direction. This will lead to problems when one wants to exploit the inter-band or inter-scale relationship of curvelet coefficients in actual applications [24].
4) Other inconveniences in implementation of the FDCT are: the curvelet basis functions do not have the same norm; the curvelet coefficients organized into subbands with different size; the redundancy ratio is not fixed, but varies within a range, which may lead to a memory allocation problem.

C. The Contourlet transform

Parallel to the development of the curvelet transform, there exists another transform proposed in signal processing literature and known as the 'contourlet' transform [8]. If the curvelet transform is originally defined as a transform for continuous functions, the contourlet transform is inherently discrete, without a continuous version. The contourlet transform is implemented by a multirate filter bank; the discrete contourlet basis functions simply being the 2-D filters that are used in that FB.

If we consider the ideal basis functions, the basis functions of the contourlet transform are similar to the basis functions of a discrete curvelet transform. Given the ideal basis functions assumption, the contourlet transform has also the same properties as the curvelet transform. The main differences between the contourlet transform and the FDCT lie in their implementations. The FDCT is faithful to its continuous counterpart. Its basis functions are strictly band-limited in frequency domain. In the spatial domain, they have infinite support. On the other hand, the contourlet transform is constructed from a cascade of finite impulse response discrete filters. Therefore, the calculation of the contourlet transform requires convolution operation. The contourlet bases are finite in the spatial domain. They do not satisfy strict requirements of the curvelet transform such as being a rotation of the same function, or having a finite support region in frequency domain.

In the first construction of the contourlet transform [8], the contourlet functions were not sharply localized in the frequency domain. In a recent construction of the contourlet transform, this problem has been reduced [25], [26]. In the

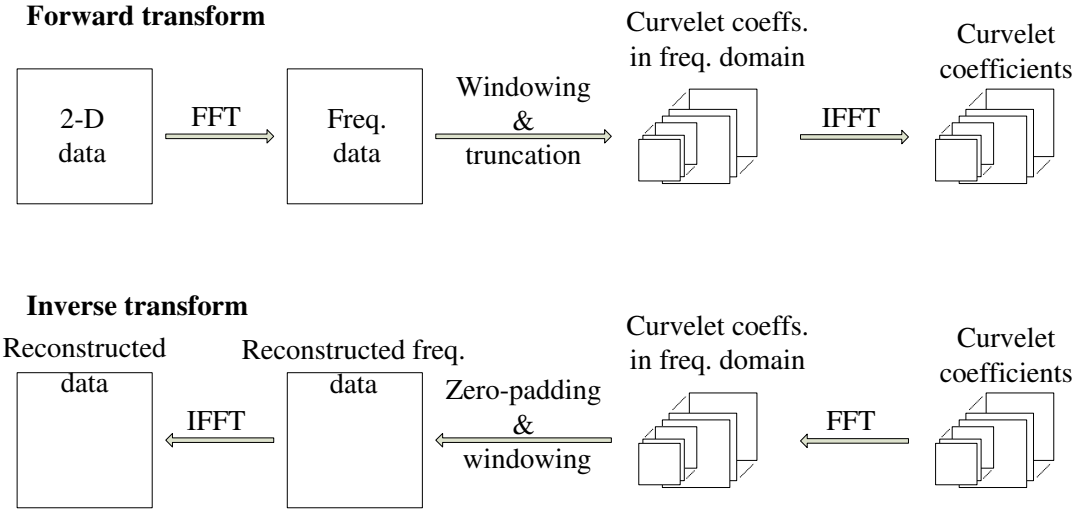


Fig. 4. Data flow structure of the FDCT forward and inverse transforms.

newly designed transform, the actual contourlet functions are quite similar to the curvelet functions in both spatial and frequency domain. Therefore, the two types of discrete transforms should have comparable performances in practical applications.

The contourlet transform, being implemented by a filter bank, is a digital-friendly transform. It has very low redundancy ratio and very fast implementation (faster than the FFT). The problem of this approach is that it is only an approximation of the curvelet transform, because the discrete basis contourlet function does not satisfy the criteria of a curvelet function. According to the construction of the curvelet transform, the basis functions in Fourier domain need to have a compact support, while the contourlet basis functions, created by cascading digital filters, do not have this property.

IV. 2-D Curvelet window definition

Since this section contained rather involved notations, let us make clear our objective. This section aims to construct a family of $2N+1$ smooth 2-D curvelet windows, denoted as $u_l(\boldsymbol{\omega})$, $l = 0, 1, \ldots, 2N$. This family of smooth windows will be used in later sections in defining the curvelet functions in frequency domain for the UDCT. The set of 2-D functions satisfy the following criteria

- All windows functions are $\mathbf{2}\pi$ periodic in ω_1 and ω_2. In the following, we limit the specification of $u_l(\boldsymbol{\omega})$ to the $(-\pi, \pi)^2$ square.
- The first window $u_0(\boldsymbol{\omega})$ has a square-shape support and zero outside $[-\pi/2, \pi/2]^2$ region (Fig. 9(d)). This function is sometimes referred to as the lowpass window.
- The other $2N$ windows have wedge-shaped supports (Fig. 9(a)).
- All $u_l(\boldsymbol{\omega})$ are smooth functions, with compact support. The function has value 1 in an 'essential support' region and smoothly transition to zero outside a compact support region. The parameter η_a and η_b control the width of these transition regions.
- For the wedge-shaped support function, the widest part of the wedge must be smaller than a fractional of 2π. In the case illustrated in Fig. 9, the widest part of the wedge-shaped support of $u_l(\boldsymbol{\omega})$ windows, $l \neq 0$, should be less than or equal to π.
- The set of $u_0^2(\boldsymbol{\omega})$ and $u_l^2(\boldsymbol{\omega}) + u_l^2(-\boldsymbol{\omega})$ form a partition of unity, which means their sum equal to 1.

The construction of 2-D window functions $u_l(\boldsymbol{\omega})$ are divided in to several steps. We first define a function $\beta(t)$ which have a smooth transition from 0 to 1. Based on $\beta(t)$ functions, a set of 2 smooth concentric square windows and a set of $2N$ angle windows are created. The parameter η_a controls the width of the transition regions of concentric square windows, while parameter η_b controls the transition regions of angle windows. The square-shape support $u_0(\boldsymbol{\omega})$ is created from the product of two 1-D functions. The wedge-shaped support $u_l(\boldsymbol{\omega})$ are constructed as the products of a concentric square passband window and the set of $2N$ angle windows.

A. One dimensional projector function.

We define a function $\beta(t)$ that has a smooth transition from 0 to 1 when t goes from -1 to 1 (see Fig. 6(a)). The $\beta(t)$ function satisfy

$$\begin{cases} \beta(t) = 1 & \text{when } 1 \leq t \\ \beta(t) = 0 & \text{when } t \leq -1 \\ \beta^2(t) + \beta^2(-t) = 1 & \text{when } -1 \leq t \leq 1 \end{cases} \tag{13}$$

One can easily find many smooth functions that satisfy (13). In our implementation we chose the following spline function when $-1 \leq t \leq 1$

$$\beta^2(t) = -\frac{5}{32}t^7 + \frac{21}{32}t^5 - \frac{35}{32}t^3 + \frac{35}{32}t^1 + \frac{1}{2} \tag{14}$$

B. Square-support functions.

Based on the β function, we now define two functions of t, which will be used to define the set of lowpass and concentric

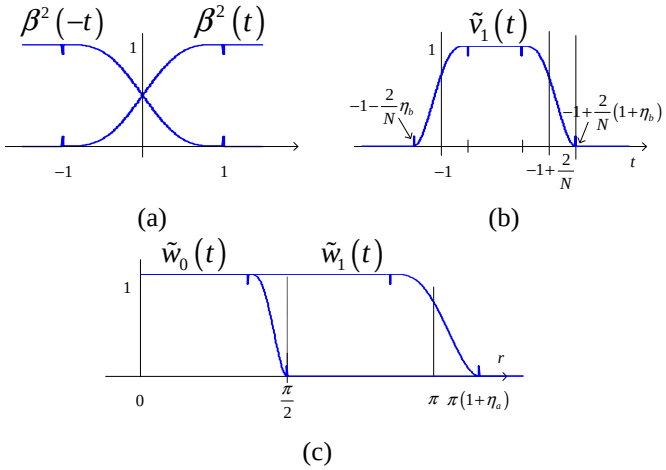


Fig. 6. The functions that are used to define the curvelet functions in frequency domain, (a) $\beta^2(t)$ function (13), (b) $\tilde{v}_1(t)$ function (26) and (c) $\tilde{w}_0(t)$ in and $\tilde{w}_1(t)$ functions in (17) and (15).

bandpass window functions in 2-D plane. The functions $\tilde{w}_1(t)$ are parameterized by $\eta_a < 1/2$

$$\tilde{w}_1(t) = \beta\left(\frac{\pi - |t|}{\pi\eta_a}\right) \tag{15}$$

Essentially, $\tilde{w}_1(t)$ is a window function that has a smooth transition from 1 to 0 when $|t|$ goes from $\pi(1-\eta_a)$ to $\pi(1+\eta_a)$ From definition of $\beta(t)$, it is easy to show that

$$\sum_{n=-\infty}^{\infty} \tilde{w}_1^2(t+2n\pi) = 1 \tag{16}$$

The function $\tilde{w}_0(t)$ is given as a scaling version of $\tilde{w}_1(t)$, so that the support of $\tilde{w}_0(t)$ is $[-\pi/2, \pi/2]$.

$$\tilde{w}_0(t) = \tilde{w}_1\left(2t(1+\eta_a)\right) \tag{17}$$

The two function $\tilde{w}_0(t)$ and $\tilde{w}_1(t)$ are plotted in Fig. 6(c). A low pass function in the frequency plane can be defined from the $\tilde{w}_0(t)$ as follows

$$w_0(\boldsymbol{\omega}) = \tilde{w}_0(\omega_1)\tilde{w}_0(\omega_2) \tag{18}$$

The $w_0(\boldsymbol{\omega})$ is zero outside the square $[-\pi/2, \pi/2]^2$. We now define $w_1(\boldsymbol{\omega})$ based on lowpass window $w_0(\boldsymbol{\omega})$ and $\tilde{w}_1(t)$ in (15)

$$w_1(\boldsymbol{\omega}) = \left(1 - w_0^2(\boldsymbol{\omega})\right)^{1/2} \tilde{w}_1(\omega_1)\tilde{w}_1(\omega_2) \tag{19}$$

Following definition of $w_0(\boldsymbol{\omega})$ and $w_1(\boldsymbol{\omega})$ in (18) and (19),

$$w_0^2(\boldsymbol{\omega}) + w_1^2(\boldsymbol{\omega}) = \tilde{w}_1^2(\omega_1)\tilde{w}_1^2(\omega_2) \tag{20}$$

This is because with $\eta_a < 1/2$, $\tilde{w}_1(\omega_i) = 1$ when $|\omega_i| < \pi/2$ (see Fig. 6(c)). Therefore $\tilde{w}_1^2(\omega_1)\tilde{w}_1^2(\omega_2)w_0^2(\boldsymbol{\omega}) = w_0^2(\boldsymbol{\omega})$.

Because our windows will be used as the discrete curvelet functions in frequency domain, they have to be $\mathbf{2}\pi$ periodic. The function $w_0(\boldsymbol{\omega})$ is periodized as follows

$$u_0(\boldsymbol{\omega}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^2} w_0(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) \tag{21}$$

A 3-D view of $u_0(\boldsymbol{\omega})$ is in Fig. 9(d). On the $[-\pi, \pi]^2$ square, the 2-D function has a square support $[-\pi/2, \pi/2]^2$. Moreover, by the scaling relationship in (17), $u_0(\boldsymbol{\omega}) = 1$ when $|\omega_1|, |\omega_2| < \frac{\pi}{2}\frac{1-\eta_a}{1+\eta_a}$. If $\frac{1-\eta_a}{1+\eta_a} > \frac{1}{2}$, we have the following relationship

$$u_0(2\boldsymbol{\omega})u_0(\boldsymbol{\omega}) = \begin{cases} u_0(2\boldsymbol{\omega}) & \text{when } |\omega_1|, |\omega_2| < \pi/4 \\ 0 & \text{when } \pi/4 \le |\omega_1|, |\omega_2| < \pi \end{cases} \tag{22}$$

The function $u_0(\boldsymbol{\omega})$ corresponds to a lowpass discrete filter, and the function $w_1(\boldsymbol{\omega})$ is a smooth bandpass function with concentric square support. Intuitively, their sum of squares is equal to one.

$$\begin{aligned} u_0^2(\boldsymbol{\omega}) + \sum_{\boldsymbol{n}\in\mathbb{Z}^2} w_1^2(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) & \\ = \sum_{\boldsymbol{n}\in\mathbb{Z}^2} w_0^2(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) + w_1^2(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) & \quad (23) \\ = \sum_{\boldsymbol{n}\in\mathbb{Z}^2} \tilde{w}_1^2(\omega_1 + 2n_1\pi)\tilde{w}_1^2(\omega_2 + 2n_2\pi) & \quad (24) \\ = 1 & \quad (25) \end{aligned}$$

(23) is obtained from (21) and the fact that $w_0(\boldsymbol{\omega} + 2\boldsymbol{n}\pi)$ are not overlapping. The next two steps are results of (20) and (16).

C. Angle polar functions.

A set of polar angle functions is defined in this section. Similar to the angle function in continuous case ($V(t)$ in Equation (7)), their squares are summed up to one. However, since the new angle functions are defined for discrete frequency plane, the rotation relationship between two functions will be replaced by the shearing relationship.

Assume that the angle functions that need to be defined is $2N$, with essential support in the range $(-\pi/4, 3\pi/4)$. First, we need to define N intermediary functions $\tilde{v}_l(t), l = 1, \ldots, N$.

$$\tilde{v}_1(t) = \beta\left(\frac{\frac{2}{N} - 1 - t}{\frac{2}{N}\eta_b}\right)\beta\left(\frac{t+1}{\frac{2}{N}\eta_b}\right) \text{ with } \eta_b \le \frac{1}{2} \tag{26}$$

$$\tilde{v}_l(t) = \tilde{v}_1\left(t - \frac{2(l-1)}{N}\right) \text{ with } l = 2, \ldots, N \tag{27}$$

The definition of $\tilde{v}_1(t)$ specifies that the function, showed in Fig. 6(b), is a smooth window with main support from -1 to $-1 + 2/N$. The width of the transition area is controlled by parameter η_b; The function changes smoothly from 0 to 1 (or from 1 to 0) when t goes from $-1-\eta_b 2/N$ to $-1+\eta_b 2/N$ (or from $-1+(1-\eta_b)2/N$ to $-1+(1+\eta_b)2/N$). The condition $\eta_b \le \frac{1}{2}$ ensures that the set t satisfying $\tilde{v}_1(t) = 1$ is not empty. The other $\tilde{v}_l(t)$ are shifted copies of $\tilde{v}_1(t)$ by $2(l-1)/N$. From the way $\beta(t)$ was defined in (13), one can verify that the sum of squares of $\tilde{v}_l(t)$, $l = 1, \ldots, N$ is also a smooth window with support from $-1 - 2/N\eta_b$ to $1 + 2/N\eta_b$.

In order to convert $\tilde{v}_l(t)$ to polar functions, we define a function $T(\theta)$ that maps the angle value θ in the range $(-\pi/2, \pi/2)$ to value $(-2, 2)$ as follows

$$T(\theta) = \begin{cases} \tan(\theta) & \text{when } -\pi/4 \le \theta \le \pi/4 \\ 2 - \tan(\pi/2 - \theta) & \text{when } \pi/4 < \theta < \pi/2 \\ -2 - \tan(\pi/2 - \theta) & \text{when } -\pi/2 < \theta < -\pi/4 \end{cases} \tag{28}$$

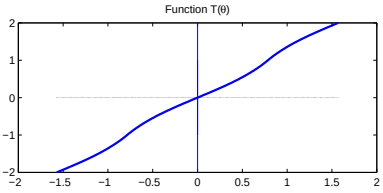


Fig. 7. Function $T(\theta)$ is used to map the 1-D function $\tilde{v}_l(t)$ to polar function $v_l(\theta)$.

From the functions $\tilde{v}_l(t)$, we can define N polar angle functions of θ

$$v_l(\theta) = \tilde{v}_l(T(\theta)) \tag{29}$$

Function $T(\theta)$ is plotted in Fig. 7. This function map the 1-D functions $v_l(t)$ to polar angle functions $v_l(\theta)$, whose essential support and 3-D views are showed in Fig. 8.

It is easy to check that the two functions $T(\theta \pm \pi/4) \pm 1$ are anti-symmetric. This allows us to construct other N polar angle functions by flipping $v_l(\theta), l = 2, \ldots, N-1$ around the value $\pi/4$,

$$v_l(\theta) = v_{2N+1-l}\left(\frac{\pi}{2} - \theta\right), l = N+1, \ldots, 2N \tag{30}$$

The anti-symmetric property of $T(\theta - \pi/4) - 1$ is used to show that the sum of square of $v_l(\theta)$ is still equal to 1 in the overlapping regions of $v_N(\theta)$ and $v_{N+1}(\theta)$. Moreover, the sum of squares of all angle polar functions $v_l(\pm\theta)$ is also equal to one.

$$\sum_{l=1}^{2N} v_l^2(\theta) + v_l^2(-\theta) = 1 \tag{31}$$

D. Discrete curvelet windows.

In a polar coordinate, $2N$ functions $v_l(\theta)$ are depending only on angle coordinate. Since they are two-variable functions on a plane, we can refer to them as function of two variables ω_1 and ω_2, or $v_l(\boldsymbol{\omega})$. Equation (31) can be rewritten

$$\sum_{l=1}^{2N} v_l^2(\boldsymbol{\omega}) + v_l^2(-\boldsymbol{\omega}) = 1 \tag{32}$$

We define the following directional wedge functions in the 2-D plane

$$\tilde{u}_l(\boldsymbol{\omega}) = v_l(\boldsymbol{\omega}) w_1(\boldsymbol{\omega}) \text{ with } l = 1, 2, \ldots, 2N \tag{33}$$

From the two above equations and the fact that $w_1(\boldsymbol{\omega})$ is symmetric, it is obvious that

$$\sum_{l=1}^{2N} \tilde{u}_l^2(\boldsymbol{\omega}) + \tilde{u}_l^2(-\boldsymbol{\omega}) = w_1^2(\boldsymbol{\omega}) \tag{34}$$

$2N$ functions $\tilde{u}_l(\boldsymbol{\omega})$ are wedge-shaped support windows with smooth transition regions. In order for a function to correspond to a discrete filter in frequency domain, it has to be $\mathbf{2}\pi$ periodic in both ω_1 and ω_2. A new set of $\mathbf{2}\pi$-periodic functions $u_l(\boldsymbol{\omega})$ is given

$$u_l(\boldsymbol{\omega}) = \sum_{\boldsymbol{n} \in \mathbb{Z}^2} \tilde{u}_l(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) \tag{35}$$

with $l = 1, 2, \ldots, 2N$. Because $\tilde{u}_l(\boldsymbol{\omega})$ have wedge-shaped compact supports (example of its supports are in Fig. 9(a)), it does not overlap with its shifted copies by $2\boldsymbol{n}\pi$. Thus

$$u_l^2(\boldsymbol{\omega}) = \sum_{\boldsymbol{n} \in \mathbb{Z}^2} \tilde{u}_l^2(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) \tag{36}$$

Following the relation in (34), we have

$$\sum_{l=1}^{2N} u_l^2(\boldsymbol{\omega}) + u_l^2(-\boldsymbol{\omega}) = \sum_{\boldsymbol{n} \in \mathbb{Z}^2} w_1^2(\boldsymbol{\omega} + 2\boldsymbol{n}\pi) \tag{37}$$

Combine with (25), the set of $u_l^2(\boldsymbol{\omega})$ window form a partition of unity

$$u_0^2(\boldsymbol{\omega}) + \sum_{l=1}^{2N} u_l^2(\boldsymbol{\omega}) + u_l^2(-\boldsymbol{\omega}) = 1 \tag{38}$$

The set of $2N+1$ window functions $u_l(\boldsymbol{\omega})$ satisfy all the criteria that have been listed at the beginning of this section.

Example 1. The construction of a set of 2-D windows $u_l(\boldsymbol{\omega})$ with three parameters $2N = 6$, $\eta_a = \eta_b = 0.15$ is illustrated in Figs. 8 and 9.

V. The Uniform Discrete Curvelet Transform as a Filter Bank

The uniform discrete curvelet transform as a FB is defined in this section, using the parameterized family of smooth windowed functions $u_l^2(\omega_1, \omega_2)$, $l = 0, 1, \ldots, 2N$. The UDCT is built up as a simple 2-D FB with one lowpass band and six directional highpass bands. Then it is shown that by cascading the same FB at lower resolution and fixing the number of directional bands following the parabolic scaling rule, a discrete decomposition faithful to the definition of the curvelet transform is created.

A. 2-D filter bank implementation in the frequency domain

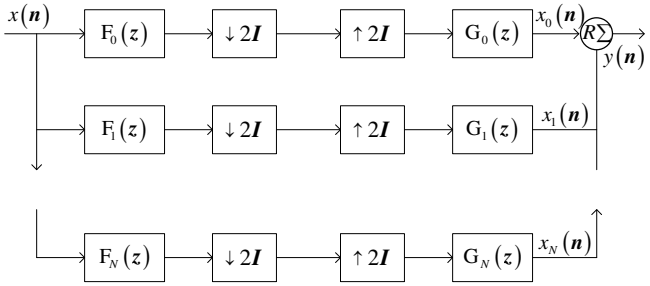


Fig. 10. A 7-band FB constructed from $u_l(\boldsymbol{\omega})$ windows in Example 1.

1) A 2-D filter bank using the set of 7 windows defined in Example 1: In example 1 we have defined a parameterized family of 2-D windows that forms a partition of unity (Equation (38)). Let us define a 7-band FB with the analysis filters $F_l(\boldsymbol{\omega})$ defined from 7 $u_l(\boldsymbol{\omega})$ windows as follows

$$F_0(\boldsymbol{\omega}) = 2u_0(\boldsymbol{\omega}) \tag{39}$$

$$F_l(\boldsymbol{\omega}) = 2\sqrt{2}u_l(\boldsymbol{\omega}), l = 1, \ldots, 6 \tag{40}$$

$$G_l(\boldsymbol{\omega}) = F_l(\boldsymbol{\omega}) \tag{41}$$

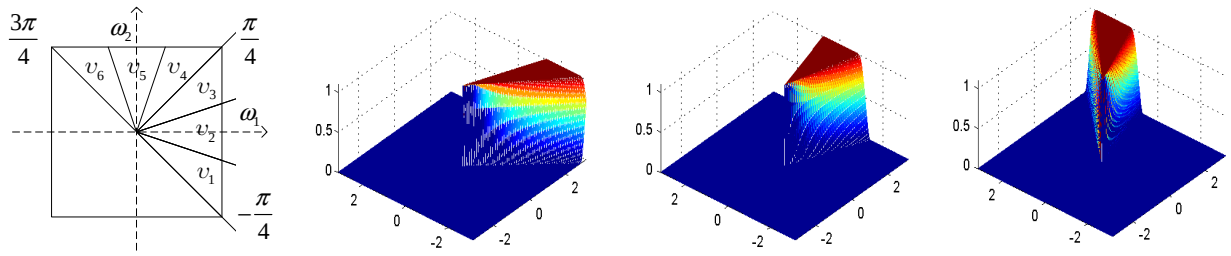


Fig. 8. (a) Essential supports of the angle functions $v_l(\theta)$ as polar functions, and 3-D view of these polar function in the frequency plane, (b) 2-D angle functions $v_1(\boldsymbol{\omega})$, (c) 2-D angle functions $v_2(\boldsymbol{\omega})$, (d) 2-D angle functions $v_3(\boldsymbol{\omega})$.

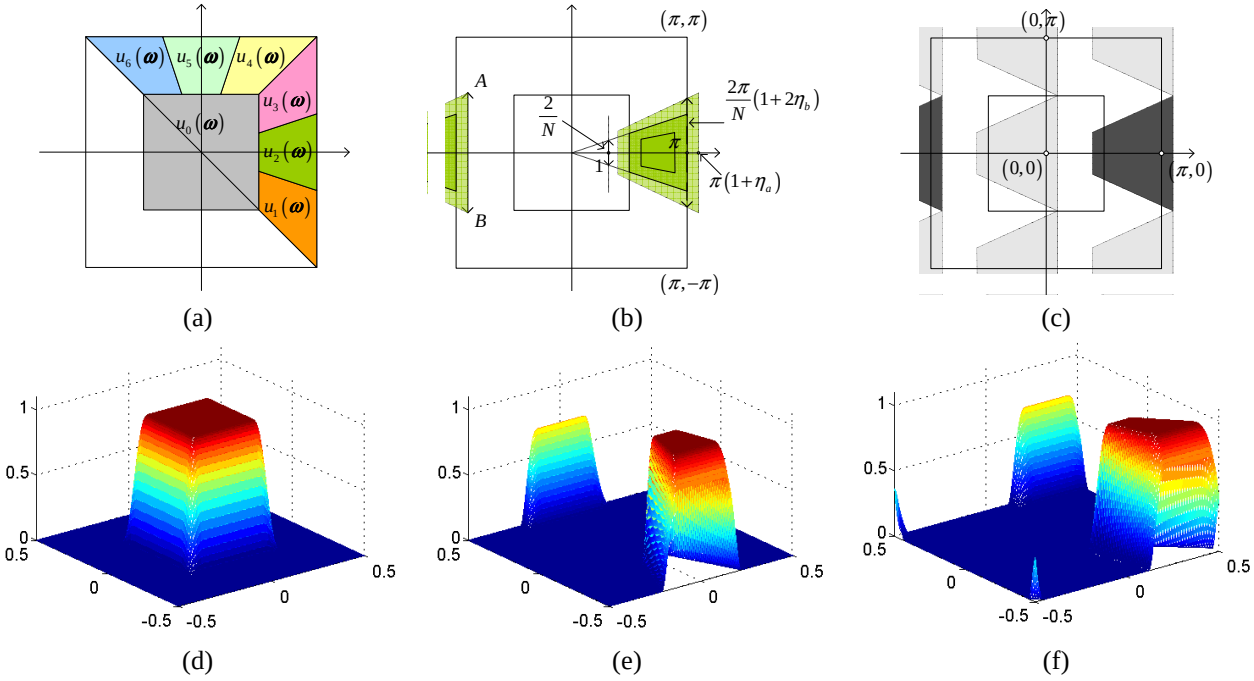


Fig. 9. An example of a family of 7 $u_l(\boldsymbol{\omega})$ windows in (ω_1, ω_2) plane. (a) Regions of essential support of $u_l(\boldsymbol{\omega})$. (b) Support of $u_2(\boldsymbol{\omega})$, distance AB can be estimated from N, η_a and η_b. (c) The support area of $F_2(\boldsymbol{\omega})$ and $G_2(\boldsymbol{\omega})$ (darker area) that are built from function $u_2(\boldsymbol{\omega})$. The lighter shade areas are support regions of $F_l(\boldsymbol{\omega} + \boldsymbol{n}\pi)$, $n \in \{\mathcal{S} \setminus (0,0)\}$. (d), (e), (f) 3-D view of functions $u_0(\boldsymbol{\omega})$, $u_2(\boldsymbol{\omega})$, $u_3(\boldsymbol{\omega})$.

In the spatial domain, the synthesis filters $g_l(\boldsymbol{n})$ are the same as the analysis filters, $g_l(\boldsymbol{n}) = f_l(\boldsymbol{n})$. The decimation ratio for all subbands is $2\boldsymbol{I}$. At the output of the FB, only real values of the reconstructed image are kept. In the following, we will show that the FB illustrated in Fig. 10 implements a discrete transform, or the FB is perfect reconstruction.

Let us consider the decomposition and reconstruction of the FB for a 2-D signal $x(\boldsymbol{n})$. The signal $x(\boldsymbol{n})$ is first filtered by $f_l(\boldsymbol{n})$. The filtered signals are decimated by $2\boldsymbol{I}$ (keeping every other row and column). On the synthesis side, the subband images are first upsampled by $2\boldsymbol{I}$ (alternating every row and column with zero), then convolved with $g_l(\boldsymbol{n})$. By the multirate filter bank theory [20], these signals can be written in frequency domain

$$X_l(\boldsymbol{\omega}) = \frac{1}{4} G_l(\boldsymbol{\omega}) \sum_{\boldsymbol{n} \in \mathcal{S}} X(\boldsymbol{\omega} + \boldsymbol{n}\pi) F_l(\boldsymbol{\omega} + \boldsymbol{n}\pi) \qquad (42)$$

where $\mathcal{S} = \{(0,0),(0,1),(1,0),(1,1)\}$. Finally, the reconstructed output signal is obtained by $y(\boldsymbol{n}) = \text{Real}\left(\sum_{l=0}^{2N}(x_l(\boldsymbol{n})\right)$.

The wedge-shaped support of window function $u_2(\boldsymbol{\omega})$ (or the filter $F_2(\boldsymbol{\omega})$) in Fig. 9(b) has the widest part $AB = \frac{2\pi}{N}(1+2\eta_b)(1+\eta_a)$. Because $\eta_a = \eta_b = 0.15$ and $N = 3$, the distance AB is less than π. Therefore, the support of $F_l(\boldsymbol{\omega} + \boldsymbol{n}\pi)$, $\boldsymbol{n} \in \mathbb{Z}^2$, are not overlapping. Therefore

$$G_l(\boldsymbol{\omega}) \sum_{\boldsymbol{n} \in \mathcal{S}} F_l(\boldsymbol{\omega} + \boldsymbol{n}\pi) = G_l(\boldsymbol{\omega}) F_l(\boldsymbol{\omega}) \qquad (43)$$

Equation (42) can be rewritten

$$X_l(\boldsymbol{\omega}) = \frac{1}{4} X(\boldsymbol{\omega}) F_l(\boldsymbol{\omega}) G_l(\boldsymbol{\omega}) \qquad (44)$$

Since the output $y(\boldsymbol{n})$ is obtained from the real part of a complex signal, its Fourier transform is given

$$Y(\boldsymbol{\omega}) = \frac{1}{2} \sum_{l=0}^{6} \left(X_l(\boldsymbol{\omega}) + X_l^*(-\boldsymbol{\omega})\right) \qquad (45)$$

Combining (44) and (45) and note that $X(\boldsymbol{\omega}) = X^*(-\boldsymbol{\omega})$, we have

$$Y(\boldsymbol{\omega}) = \frac{1}{8}X(\boldsymbol{\omega})\sum_{l=0}^{6}\left(G_l(\boldsymbol{\omega})F_l(\boldsymbol{\omega}) + G_l^*(-\boldsymbol{\omega})F_l^*(-\boldsymbol{\omega})\right) \tag{46}$$

Since $F_l(\boldsymbol{\omega})$ and $G_l(\boldsymbol{\omega})$ are defined from $u_l(\boldsymbol{\omega})$, which are positive functions. Therefore, we have $G_l(\boldsymbol{\omega})F_l(\boldsymbol{\omega}) = 8u_l^2(\boldsymbol{\omega})$, $l = 1, \ldots, 6$. The lowpass function $u_0(\boldsymbol{\omega})$ is symmetric, which leads to $G_0(\boldsymbol{\omega})F_0(\boldsymbol{\omega}) = G_0^*(-\boldsymbol{\omega})F_0^*(-\boldsymbol{\omega}) = 4u_0^2(\boldsymbol{\omega})$. Equation (46) is rewritten

$$Y(\boldsymbol{\omega}) = X(\boldsymbol{\omega})\left(u_0^2(\boldsymbol{\omega}) + \sum_{l=1}^{6}u_l^2(\boldsymbol{\omega}) + u_l^2(-\boldsymbol{\omega})\right) \tag{47}$$

From the equation (38), we have $Y(\boldsymbol{\omega}) = X(\boldsymbol{\omega})$, or $y(\boldsymbol{n}) = x(\boldsymbol{n})$.

2) Relationship between η_a, η_b and the redundancy of the decomposition: In order to reduce the overcomplete ratio of a discrete curvelet decomposition, one would need to decimate the subband after filtering. We limit our choice of decimation ratio to the power of 2. The shape of support of a function $u_l(\omega_1, \omega_2)$ is fully determined by parameters η_a, η_b and N. The distance AB in Fig. 9 is $\frac{2\pi}{N}(1+2\eta_b)(1+\eta_a)$. Our objective is to determine η_a, and η_b so that the distance $AB < \pi/2^n$. Following are set of parameters that satisfy this requirement

$$N = 3 \cdot 2^n, n \geq 0 \quad , \quad \eta_a = \eta_b = 0.15 \tag{48}$$
$$N = 4 \cdot 2^n, n \geq 0 \quad , \quad \eta_a = \eta_b = 0.3 \tag{49}$$
$$N = 5 \cdot 2^n, n \geq 0 \quad , \quad \eta_a = \eta_b = 0.5 \tag{50}$$

From now on we pick the parameters $\eta_a = \eta_b = 0.15$ and N in the form $3 \cdot 2^n$ as in (48) . This set of parameters are interesting because it allows the curvelet window have a regions of transition around the essential support region of the curvelet. It makes the curvelet well-localized in the spatial domain. Moreover, the estimation of the redundancy ratio in (86) shows that the 2-D UDCT with this set of parameters has an acceptable redundancy of 4

3) Two-dimensional filter bank with $2N$ directional band : We have created a 7-band FB based on the set of 7 $u_l(\boldsymbol{\omega})$ windows. In general case, we can set $N = 3 \cdot 2^n, n \geq 0$, $\eta_a = \eta_b = 0.15$ as in (48) and use the $2N+1$ $u_l(\boldsymbol{\omega})$ windows to define a $(2N+1)$-band FB as follows

$$F_0(\boldsymbol{\omega}) = 2u_0(\boldsymbol{\omega}) \tag{51}$$
$$F_l(\boldsymbol{\omega}) = 2^{\frac{n+3}{2}}u_l(\boldsymbol{\omega}), l = 1, \ldots, 2N \tag{52}$$
$$G_l(\boldsymbol{\omega}) = F_l(\boldsymbol{\omega}) \tag{53}$$

Since the two parameters $\eta_a = \eta_b = 0.15$, it can be shown that the widest part of the wedge-shaped supports of $u_l(\boldsymbol{\omega})$ windows are less than $\pi/(N/3)$, or $\pi/2^n$. Therefore, the decimation ratio for the $2N$ directional bands of the $2N$-band FB will be 2^n times higher than the case $2N = 6$. The decimation ratio for the first $3 \cdot 2^n$ directional bands are $\boldsymbol{D}_1^{(N)} = \text{diag}\left\{2, 2^{n+1}\right\}$ and for the last $3 \cdot 2^n$ directional bands are $\boldsymbol{D}_2^{(N)} = \text{diag}\left\{2^{n+1}, 2\right\}$. The decimation ratio for the lowpass band is $\boldsymbol{D}_0^{(N)} = 2\boldsymbol{I}$. Let us rewrite the three decimation ratios used in a $2N$-band UDCT filter bank as follows

$$\boldsymbol{D}_0^{(N)} = 2\boldsymbol{I}, \boldsymbol{D}_1^{(N)} = \text{diag}\left\{2, \frac{2N}{3}\right\}, \boldsymbol{D}_2^{(N)} = \text{diag}\left\{\frac{2N}{3}, 2\right\} \tag{54}$$

Following similar argument as for the case of 7-band FB, one can easily show that the constructed FB is indeed perfect reconstruction. Since the synthesis filters are the same as the analysis filters, the FB is implementing a tight frame decomposition.

B. The uniform discrete curvelet transform as an iterative multiscale filter bank

We have constructed a set of $2N$ 2-D directional filters $F_l(\boldsymbol{\omega}), l = 1, \ldots, 2N$ and a lowpass filter $F_0(\boldsymbol{\omega})$ in such a way that the directional subbands and the lowpass subband can be decimated without aliasing. The defined filters in the frequency domain are real-valued functions and satisfy the perfect reconstruction condition, taking into account the decimation ratio. Similar to the complex wavelet [4], our directional filters have one-sided support in the frequency domain, making the subband coefficients have complex values. In the reconstruction procedure, the final complex components are simply discarded.

In order to obtain a multiresolution decomposition, another FB with the same η_a, η_b can be reiterated at lower band. Since our objective is to create a discrete decomposition faithful to the curvelet transform, the number of directional subbands should follow the parabolic scaling rule. For example, a signal $x(\boldsymbol{n})$ can be decomposed by the UDCT transform into one lowpass band and J highpass band, numbered from 1 (lowest resolution) to J (highest resolution). The UDCT is implemented by cascading J multiresolution FBs that are constructed in the previous subsection. The number of directional band $2N_j$ at resolution j should be proportional to $2^{\lfloor \frac{j}{2} \rfloor}$

However, in practical implementation, the number of directions of UDCT filter bank at each level can be flexible. For instance, the multiresolution FB consist of J level; the UDCT filter bank at level j has $2N_j$ directions. The three parameters η_a, η_b and $2N_j$ are chosen as in (48). The filters at scale j are denoted by $F_l^{(j)}(\boldsymbol{\omega})$, $l = 0, 1, \ldots, 2N_j$. Since we use the same parameter η_a for all J set of parameterized windows, the lowpass windows $F_0^{(j)}(\boldsymbol{\omega})$ are the same for all scale j.

Example 2: A multilevel UDCT decomposition. Let us demonstrate an example of the UDCT decomposition, with configuration as follows: $J = 3$, $2N_1 = 6$, $2N_2 = 12$, $2N_3 = 24$. The 3 level UDCT FB associated with this decomposition is in Fig. 11(a).

Fig. 12 are examples of the zoneplate image decomposed by our discrete curvelet transform with above parameters J and N_j. Fig. 12(c) is the magnitude of the complex value coefficients in the transform domain.

For any multilevel FB, one can find an equivalent one-level FB, by moving the decimation block towards the end (or the start) of the analysis-side (or synthesis-side) FB. The equivalent FB for structure in Fig. 11(a) is in Fig. 11(b), with

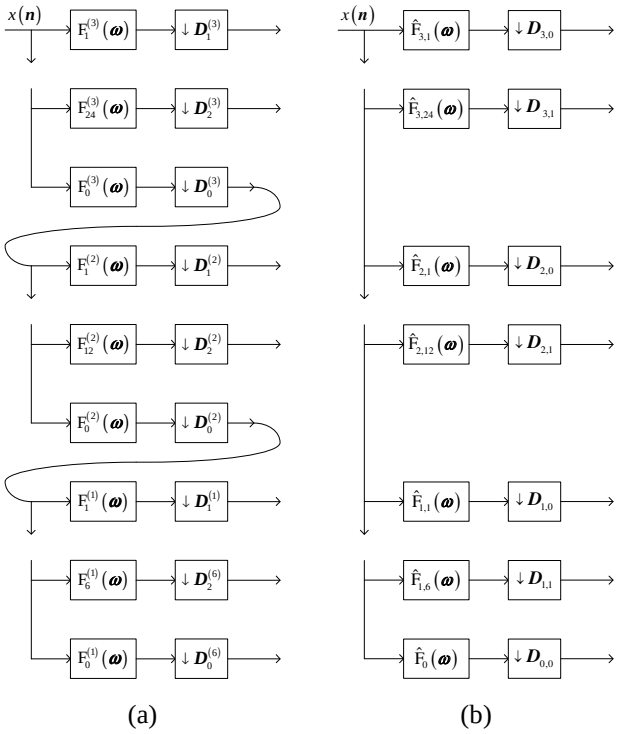


Fig. 11. Discrete Curvelet transform is obtained by cascading multiple level of UDCT filter bank. (a) The three-level UDCT filter bank in Example 2, and (b) Equivalent FB, filters $\hat{F}_{j,l}(\boldsymbol{\omega})$ as in (55), and $\boldsymbol{D}_{j,0}$, $\boldsymbol{D}_{j,1}$, as in (59) and (60)

equivalent filters in scale j and direction l denoted by $\hat{F}_{j,l}(\boldsymbol{\omega})$. By the filter bank theory [20],

$$\hat{F}_{j,l}(\boldsymbol{\omega}) = F_l^{(j)}(2^{J-j}\boldsymbol{\omega}) \prod_{i=0}^{J-j-1} F_0^{(J-i)}(2^i\boldsymbol{\omega}) \qquad (55)$$

Since all $F_0^{(i)}(\boldsymbol{\omega})$, $i = 1, \ldots, J$ are the same and defined from $u_0(\boldsymbol{\omega})$ as in (51), and the product $u_0(\boldsymbol{\omega})u_0(2\boldsymbol{\omega})$ is already estimated in (22), we have

$$\prod_{i=0}^{j} F_0^{(J-i)}(2^i\boldsymbol{\omega}) = \begin{cases} 2^j F_0^{(J-j)}(2^j\boldsymbol{\omega}), & \text{when } |\omega_1|, |\omega_2| < \pi/2^{j+1} \\ 0, & \text{when } \pi/2^{j+1} \le |\omega_1|, |\omega_2| < \pi \end{cases} \qquad (56)$$

Therefore, in the $[-\pi, \pi]^2$ square of the $\boldsymbol{\omega}$ frequency plane, $\hat{F}_{j,l}(\boldsymbol{\omega})$ when $j < J$ has the following value

$$\hat{F}_{j,l}(\boldsymbol{\omega}) = \begin{cases} 2^{J-j-1} F_l^{(j)}(2^{J-j}\boldsymbol{\omega}) F_0^{(j+1)}(2^{J-j-1}\boldsymbol{\omega}), \\ \qquad \text{when } |\omega_1|, |\omega_2| < \pi/2^{J-j} \\ 0, \quad \text{when } \pi/2^{J-j} \le |\omega_1|, |\omega_2| < \pi \end{cases} \qquad (57)$$

where $F_0(\boldsymbol{\omega})$ is defined in (39) and (21). The equivalent lowpass filter is

$$\hat{F}_0(\boldsymbol{\omega}) = \begin{cases} 2^{J-1} F_0(2^{J-1}\boldsymbol{\omega}) & \text{when } |\omega_1|, |\omega_2| < \pi/2^J \\ 0 & \text{when } \pi/2^J \le |\omega_1|, |\omega_2| < \pi \end{cases} \qquad (58)$$

We have determined the equivalent curvelet filters of an iterated UDCT filter bank in (57) and (58) by pushing all the analysis filters to the left and all the decimation blocks to the right on the analysis side. Associated with the equivalent filters are the overall decimation ratio, which are the product of all decimation blocks along the branch of the iterative FB. The decimation ratio for the first half of directional band (mostly horizontal band) is

$$\begin{aligned} \boldsymbol{D}_{j,0} &= \boldsymbol{D}_1^{(j)} \prod_{i=j+1}^{J} \boldsymbol{D}_0^{(i)} \\ &= \text{diag}\left\{2, \frac{2N_j}{3}\right\} \cdot 2^{J-j}, \end{aligned} \qquad (59)$$

and for the second half of directional band (mostly vertical band) is

$$\begin{aligned} \boldsymbol{D}_{j,1} &= \boldsymbol{D}_2^{(j)} \prod_{i=j+1}^{J} \boldsymbol{D}_0^{(i)} \\ &= \text{diag}\left\{\frac{2N_j}{3}, 2\right\} \cdot 2^{J-j}, \end{aligned} \qquad (60)$$

The two above equations are obtained simply by applying (54). The decimation ratio corresponds to the lowpass window $\hat{F}_0(\boldsymbol{\omega})$ in (58) is $\boldsymbol{D}_{0,0} = \text{diag}\left\{2^J, 2^J\right\}$.

Example 2 (continued). In the case of example 2, the equivalent FB in Fig. 11(b) produces the same decomposition as the 3 level FB in Fig. 11(a). The overall decimation ratio for the equivalent FB are

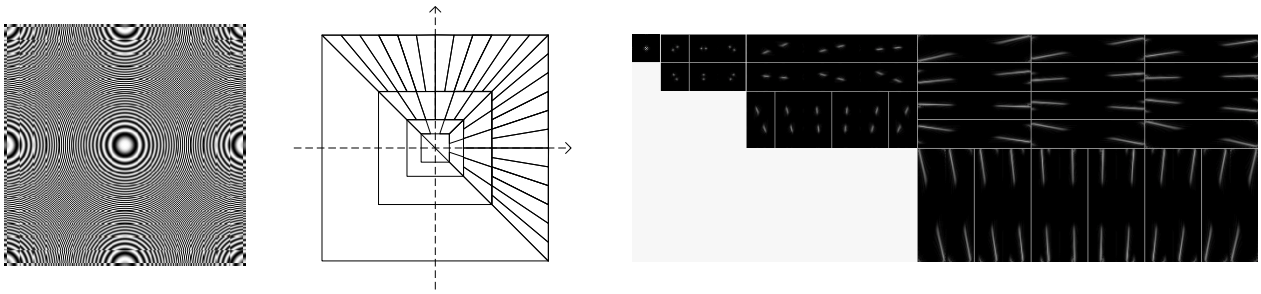
$$\begin{aligned} \boldsymbol{D}_{0,0} &= \text{diag}\left\{8, 8\right\}, \\ \boldsymbol{D}_{1,0} &= \text{diag}\left\{8, 8\right\}, \boldsymbol{D}_{1,1} = \text{diag}\left\{8, 8\right\} \\ \boldsymbol{D}_{2,0} &= \text{diag}\left\{4, 8\right\}, \boldsymbol{D}_{2,1} = \text{diag}\left\{8, 4\right\} \\ \boldsymbol{D}_{3,0} &= \text{diag}\left\{2, 8\right\}, \boldsymbol{D}_{3,1} = \text{diag}\left\{8, 2\right\} \end{aligned} \qquad (61)$$

VI. The uniform discrete curvelet transform

A. The discrete transform implemented by UDCT filter bank

The UDCT has been introduced as a multiresolution FB implemented in the frequency domain by using a set of curvelet windows. The curvelet basis function of the UDCT can be estimated from the FB implementation as in (57) and (58). Since the directional filters used in the construction of UDCT have one-sided support in the frequency domain, they have complex valued coefficients in the spatial domain. Therefore we have to interpret each complex valued subband coefficient produced by UDCT as the projection of input signal to two curvelet basis functions, which are the real and imaginary parts of a complex function. The first function is symmetric and the second function is anti-symmetric. Their centers of symmetry and anti-symmetry are at the same point, which is also considered the position of the two curvelet functions.

The index set for the discrete basis function is $\mu = (i, j, l, \boldsymbol{k}) \in \mathcal{M}$, where $i = 1, 2$ is an indication on symmetric/antisymmetric property of the basis function; parameter $j = 1, \ldots, J$ is the scale index, $l = 1, \ldots, 2N_j$ is the directional index, and $\boldsymbol{k}$ belongs to the lattice $\Lambda\left(\boldsymbol{D}_{j,\left\lfloor \frac{l-1}{N_j} \right\rfloor}\right)$. The frequency domain representation of $\hat{F}_{j,l}(\boldsymbol{\omega})$ in (57) corresponds to a complex function in the spatial domain whose real part is $\varphi_{1,j,l,\mathbf{0}}(\boldsymbol{n})$ and imaginary part is $\varphi_{2,j,l,\mathbf{0}}(\boldsymbol{n})$. The



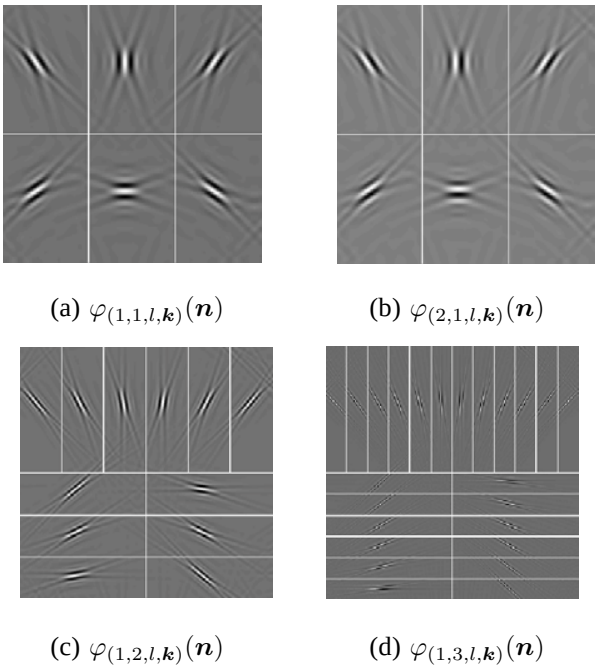
(a) (b) (c)

Fig. 12. An example of the UDCT decomposition (a) Zoneplate image, (b) Essential support of the curvelet function in Fourier domain and (c) Curvelet coefficients.

other functions at different location index $\boldsymbol{k}$ are simply shifted versions of this function on the $\Lambda\left(\boldsymbol{D}_{j,\left\lfloor \frac{l-1}{N_j} \right\rfloor}\right)$ grid.

$$\varphi_{(i,j,l,\boldsymbol{k})}(\boldsymbol{n}) \quad = \quad \varphi_{(i,j,l,\mathbf{0})}(\boldsymbol{n}-\boldsymbol{k}) \tag{62}$$

We extend the index set $\mathcal{M}$ to include index elements $\mu = (1,0,0,\boldsymbol{k})$ for the 'scaling' curvelet that corresponds to the lowpass filter defined in (58). The Fourier transform of $\varphi_{(1,0,0,\mathbf{0})}(\boldsymbol{n})$ is $\hat{F}_0(\boldsymbol{\omega})$. The other elements of these lowpass curvelets corresponds to $\mathbf{k} \in \Lambda(\boldsymbol{D}_{0,0})$ lattice.

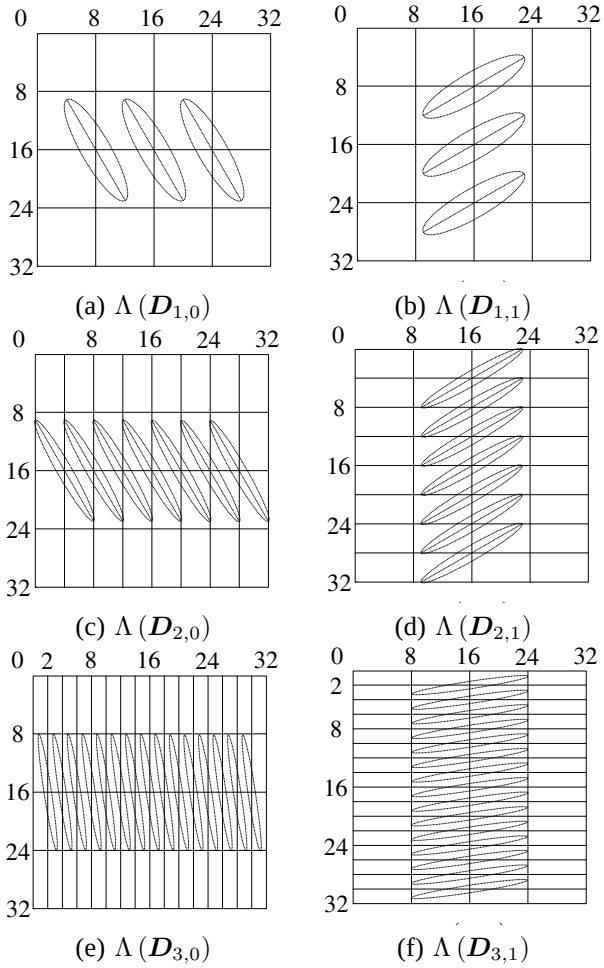


(a) $\varphi_{(1,1,l,\boldsymbol{k})}(\boldsymbol{n})$ (b) $\varphi_{(2,1,l,\boldsymbol{k})}(\boldsymbol{n})$

(c) $\varphi_{(1,2,l,\boldsymbol{k})}(\boldsymbol{n})$ (d) $\varphi_{(1,3,l,\boldsymbol{k})}(\boldsymbol{n})$

Fig. 13. Examples of 2-D UDCT basis functions corresponds to a $6, 12, 24$ configuration, (a) $\varphi_{(1,1,l,\boldsymbol{k})}(\boldsymbol{n})$, $l = 1, \ldots, 2N_1 = 6$, (b) $\varphi_{(2,1,l,\boldsymbol{k})}(\boldsymbol{n})$, $l = 1, \ldots, 2N_1 = 6$ (c) $\varphi_{(1,2,l,\boldsymbol{k})}(\boldsymbol{n})$, $l = 1, \ldots, 2N_2 = 12$, (d) $\varphi_{(1,3,l,\boldsymbol{k})}(\boldsymbol{n})$, $l = 1, \ldots, 2N_3 = 24$.

Example 2 *(continued)*. The decomposition in this case has three level of directional curvelet. The lowest directional scale has six direction, where the corresponding symmetric and anti-symmetric curvelet functions are showed in Fig. 13 (a) and (b). The center of these curvelet functions are situated on the lattices $\Lambda(\boldsymbol{D}_{1,0})$ and $\Lambda(\boldsymbol{D}_{1,1})$ in Fig. 14 (a) and (b). Other lattices for scales $J = 2, 3$ are also included in Fig. 14



(a) $\Lambda(\boldsymbol{D}_{1,0})$ (b) $\Lambda(\boldsymbol{D}_{1,1})$

(c) $\Lambda(\boldsymbol{D}_{2,0})$ (d) $\Lambda(\boldsymbol{D}_{2,1})$

(e) $\Lambda(\boldsymbol{D}_{3,0})$ (f) $\Lambda(\boldsymbol{D}_{3,1})$

Fig. 14. The lattices of UDCT transform, corresponding to UDCT transform described in Example 2. (a) and (b), (c) and (d), (e) and (f) are lattices for the first, second and third directional resolutions, see (61)

Any two dimensional discrete function $x(\boldsymbol{n}) \in \ell^2$ can be expressed as a sum of φ_μ.

$$x(\boldsymbol{n}) \quad = \quad \sum_{\mu \in \mathcal{M}} c_\mu \varphi_\mu(\boldsymbol{n}) \tag{63}$$

where c_μ are estimated by

$$c_\mu \quad = \quad \langle x(\boldsymbol{n}), \varphi_\mu(\boldsymbol{n}) \rangle \tag{64}$$

c_μ are the real and imaginary part of complex coefficients

produced by the UDCT filter bank. Combining the two formulae (63) and (64) and writing out the range of index in $\mathcal{M}$, we have the following self-inversion formula.

$$x(\boldsymbol{n}) = \sum_{\boldsymbol{k}\in\Lambda(\boldsymbol{D}_{0,0})} \left\langle x(\boldsymbol{n}), \varphi_{(1,0,0,\boldsymbol{k})}(\boldsymbol{n})\right\rangle \varphi_{(1,0,0,\boldsymbol{k})}(\boldsymbol{n}) + \\ + \sum_{j=1}^{J}\sum_{l=1}^{2N_j}\sum_{i=1}^{2} \sum_{\boldsymbol{k}\in\Lambda\left(\boldsymbol{D}_{j,\left\lfloor\frac{l-1}{N_j}\right\rfloor}\right)} \left\langle x(\boldsymbol{n}), \varphi_{(i,j,l,\boldsymbol{k})}(\boldsymbol{n})\right\rangle \varphi_{(i,j,l,\boldsymbol{k})}(\boldsymbol{n}) \quad (65)$$

B. The UDCT and other curvelet transforms

The UDCT has some similarities to the wrapping-based FDCT [6], in the sense that both the two transforms employ alias free subsampling in frequency domain. Therefore the curvelet functions of the two discrete transforms have compact support windows in Fourier domain. Both the UDCT and the FDCT are implemented by FFT algorithm. Therefore they have the same data flow structure as illustrated in Fig. 4. One can regard the UDCT as a generalization and parameterization of the FDCT. The practical advantages of the UDCT are detailed in the following list, comparing against the list in subsection III-B

1) The UDCT use the fact that only half of the number of Fourier transform coefficients are needed to reconstruct a real signal. Therefore, the curvelet window functions are designed to cover only approximately half of the frequency plane. This reduces the redundancy ratio of the UDCT compared to the FDCT.
2) In the UDCT, the curvelet window functions are systematically defined and parameterized by η_a and η_b . By varying these parameters, one can generate UDCT with different redundancy ratio and curvelet support shapes.
3) The UDCT is built up using multirate FB theory. The set of window functions is generated from a cascade of multiresolution FBs, each FB is implemented in frequency domain. This makes all the differences between the UDCT and the FDCT. For example, because the FB employed two integer subsampling matrices for all the directional bands, the generated curvelet functions are located on two lattices; one for mostly vertical and one for mostly horizontal curvelet. The discrete curvelet functions at each resolution and scale are the same function shifted to different nodes on one lattice. Moreover, the lattices of lower scales are subset of those at higher scales. Therefore, the UDCT coefficients have a clear parent-children relationship.
4) The UDCT provide several other properties that can be useful in practical applications: i) All the norms of the curvelet functions at each scale are approximately the same. ii) The number of coefficients at each scale are fixed and do not depend on the number of directions. iii) The size of subbands are the same for each scale (subject to a transpose operation).

The UDCT can also be viewed as a complex wavelet decomposition, implemented by filter bank, in the same way as shiftable complex directional pyramid [18] or other complex wavelet transforms [27]. Unlike the contourlet FB, each complex band of the UDCT is shift invariant in the energy sense [28], and should offer better performances in image analysis applications. The countourlet FB can also be redesigned to achieve complete linear shift invariant as in the nonsubsampled contourlet transform [17].

VII. The UDCT in three or more dimensions

One of the advantages of the UDCT is that it can be generalized to M-dimensions without significant difficulty. The M-D UDCT can be constructed following the same steps as the 2-D UDCT. First a parameterized family of M variable window functions are defined. These window functions are used to defined M-D filters for a M-D FB. By cascading iteratively this M-D FB, a M-D UDCT decomposition is obtained. We focus on how to construct a parameterized family of M-D curvelet windows. The details on using this set of windows to construct a M-D FB, and cascading multiples M-D FBs to create M-D DCT transforms are omitted. The 3-D UDCT case is treated as an example of the M-D cases.

A. A set of M-D curvelet window functions

In this section, function of bold letter variable denotes M-D function. The M-D curvelet windows are defined as the product of a concentric square window and a set of angle windows in M-D space. The M-D generalization of concentric square window is straightforward. We reuse the two functions $\tilde{w}_0(t)$ and $\tilde{w}_1(t)$ defined in (15) and (17). Similar to the 2-D case in (18) and (19), the two functions $w_0(\boldsymbol{\omega})$ and $w_1(\boldsymbol{\omega})$ in M-D space are given

$$w_0(\boldsymbol{\omega}) = \prod_{k=1}^{M} \tilde{w}_0(\omega_k) \quad (66)$$

$$w_1(\boldsymbol{\omega}) = \left(1 - w_0^2(\boldsymbol{\omega})\right)^{1/2} \prod_{k=1}^{M} \tilde{w}_1(\omega_k) \quad (67)$$

Let us consider the generalization of angle functions in M dimensions. We begin by assuming that we have a set of polar angle function $v_l(\theta)$, $l = 1..,N$, $N = 3\cdot 2^n$ as defined in Section IV-C. The main support of the sum of square of these functions is an infinite pyramid shape region satisfying $\omega_1 > |\omega_2|$; On a polar coordinate, its values depend only on the angle coordinate. Three functions $v_1(\theta)$, $v_2(\theta)$ and $v_3(\theta)$ are illustrated in Fig. 8(b, c, d), in the case $N = 3$. Instead of consider v_l as a function of angle θ, we can write it as function of (ω_1, ω_2). Building up from 2-D angle function $v_l(\omega_1, \omega_2)$, we can define a set of $M \times N^{M-1}$ polar function on M-D space $\tilde{v}_{\boldsymbol{l}}(\boldsymbol{\omega})$, index by $\boldsymbol{l} = (l_1, .., l_M)$, where l_1 is the index of the hyperpyramid that the angle function belongs to, $1 \le l_1 \le M$; the rest are $M-1$ direction indices of the angle function within that hyperpyramid, $1 \le l_2, \ldots, l_M \le N$. Each function is defined from $M-1$ functions v_l as follows

$$\tilde{v}_{l_1\ldots l_M}(\omega_1, \ldots, \omega_M) = \\ = v_{l_2}(\omega_{l_1}, \omega_{k_1}) v_{l_3}(\omega_{l_1}, \omega_{k_2}) \ldots v_{l_M}(\omega_{l_1}, \omega_{k_{M-1}}) \quad (68)$$

where $1 \le k_1 < k_2 \ldots < k_{M-1} \le M$ and all $k_m \ne l_1$, $1 \le m \le M-1$.

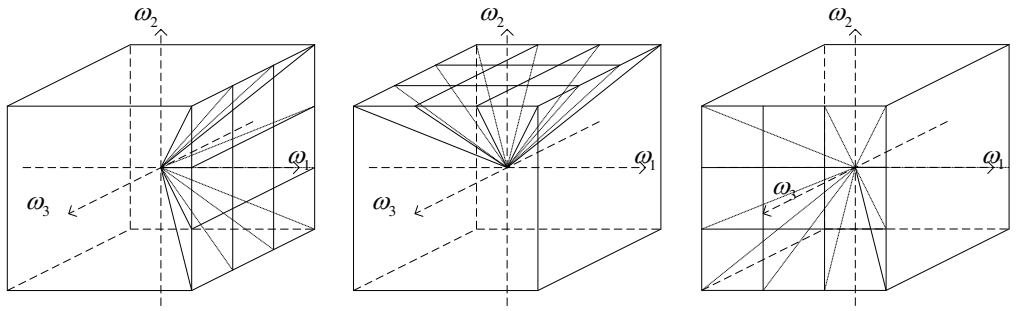


Fig. 15. Construct the angle functions in 3-D space, (a) 3-D angle functions in the first pyramid, (b) 3-D angle functions in the second pyramid, (c) 3-D angle functions in the third pyramid.

For each l_1, N^{M-1} function $\tilde{v}_{l_1,\ldots,l_M}(\pm\boldsymbol{\omega})$ covers a hyper pyramid in l_1 direction of M dimensional space. Unlike the 2-D case, the sum of square of these functions and their mirror through the origin are not equal to 1 in the regions where three or more $\tilde{v}_{\boldsymbol{l}}(\pm\boldsymbol{\omega})$ functions are overlapping. Denote their sum of square as

$$V^2(\boldsymbol{\omega}) = \sum_{l_1=1}^{M}\sum_{l_2=1}^{N}\cdots\sum_{l_M=1}^{N}\tilde{v}^2_{l_1\ldots l_M}(\boldsymbol{\omega}) + \tilde{v}^2_{l_1\ldots l_M}(-\boldsymbol{\omega}) \quad (69)$$

We can define a set of normalized angle functions $v_{l_1\ldots l_M}(\boldsymbol{\omega})$ as follows

$$v_{l_1\ldots l_M}(\boldsymbol{\omega}) = \frac{\tilde{v}_{l_1\ldots l_M}(\boldsymbol{\omega})}{V(\boldsymbol{\omega})} \quad (70)$$

Because of the normalization, the sum of square of all $v_{l_1\ldots l_M}(\boldsymbol{\omega})$ is equal to 1.

A set of M-D window functions that are $\mathbf{2}\pi$ periodic in M dimensional space can be constructed from $w_0(\boldsymbol{\omega})$, $w_1(\boldsymbol{\omega})$ and $v_{l_1\ldots l_M}(\boldsymbol{\omega})$ as follows

$$u_0(\boldsymbol{\omega}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^M} w_0(\boldsymbol{\omega}+2\boldsymbol{n}\pi) \quad (71)$$

$$u_{l_1\ldots l_M}(\boldsymbol{\omega}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^M} v_{l_1\ldots l_M}(\boldsymbol{\omega}+2\boldsymbol{n}\pi)w_1(\boldsymbol{\omega}+2\boldsymbol{n}\pi) \quad (72)$$

Similar to the 2-D case (38), these functions also form a partition of unity in M-D space

$$u_0^2(\boldsymbol{\omega}) + \sum_{l_1=1}^{M}\sum_{l_2=1}^{N}\cdots\sum_{l_M=1}^{N} u^2_{l_1\ldots l_M}(\boldsymbol{\omega}) + u^2_{l_1\ldots l_M}(-\boldsymbol{\omega}) = 1 \quad (73)$$

Example 3, *UDCT in three dimensions*. For illustration, the 3-D UDCT is considered as a particular case of M-D UDCT. In 3-D case, the three-variable angle functions $\tilde{v}_{l_1l_2l_3}(\omega_1,\omega_2,\omega_3)$, $l_2, l_3 = 1,\ldots,N$, correspond to M-D functions in (68), are

$$\tilde{v}_{1l_2l_3}(\boldsymbol{\omega}) = v_{l_2}(\omega_1,\omega_2)v_{l_3}(\omega_1,\omega_3) \quad (74)$$

$$\tilde{v}_{2l_2l_3}(\boldsymbol{\omega}) = v_{l_2}(\omega_2,\omega_1)v_{l_3}(\omega_2,\omega_3) \quad (75)$$

$$\tilde{v}_{3l_2l_3}(\boldsymbol{\omega}) = v_{l_2}(\omega_3,\omega_1)v_{l_3}(\omega_3,\omega_2) \quad (76)$$

These functions cover the three infinite pyramids in Fig. 15(a, b, c). Following the normalization in (69) and (70), one can define the set of $3\times N^2$ $v_{l_1l_2l_3}(\boldsymbol{\omega})$ functions in 3-D space whose sum of square is equal to one:

$$v_{l_1l_2l_3}(\boldsymbol{\omega}) = \frac{\tilde{v}_{l_1l_2l_3}(\boldsymbol{\omega})}{V(\boldsymbol{\omega})} \quad (77)$$

Similar to (71) and (72), a set of 3-D curvelet window functions can be constructed from $w_0(\boldsymbol{\omega})$, $w_0(\boldsymbol{\omega})$ and $v_{l_1l_2l_3}(\boldsymbol{\omega})$ as follows

$$u_0(\boldsymbol{\omega}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^3} w_0(\boldsymbol{\omega}+2\boldsymbol{n}\pi) \quad (78)$$

$$u_{l_1l_2l_3}(\boldsymbol{\omega}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^3} v_{l_1l_2l_3}(\boldsymbol{\omega}+2\boldsymbol{n}\pi)w_1(\boldsymbol{\omega}+2\boldsymbol{n}\pi) \quad (79)$$

Similar to the 2-D case in (38), these functions also form a partition of unity in 3-D space

$$u_0^2(\boldsymbol{\omega}) + \sum_{l_1=1}^{3}\sum_{l_2=1}^{N}\sum_{l_3=1}^{N} u^2_{l_1l_2l_3}(\boldsymbol{\omega}) + u^2_{l_1l_2l_3}(-\boldsymbol{\omega}) = 1 \quad (80)$$

B. Uniform discrete curvelet transform in M dimensions

From the set of M-D windows $u_0(\boldsymbol{\omega})$ and $u_{\boldsymbol{l}}(\boldsymbol{\omega})$, we can define a M-D FB with analysis filters

$$F_0(\boldsymbol{\omega}) = 2^{\frac{M}{2}}u_0(\boldsymbol{\omega}) \quad (81)$$

$$F_{l_1\ldots l_M}(\boldsymbol{\omega}) = 2^{\frac{(M-1)(n+1)}{2}+1}u_l(\boldsymbol{\omega}), \quad (82)$$

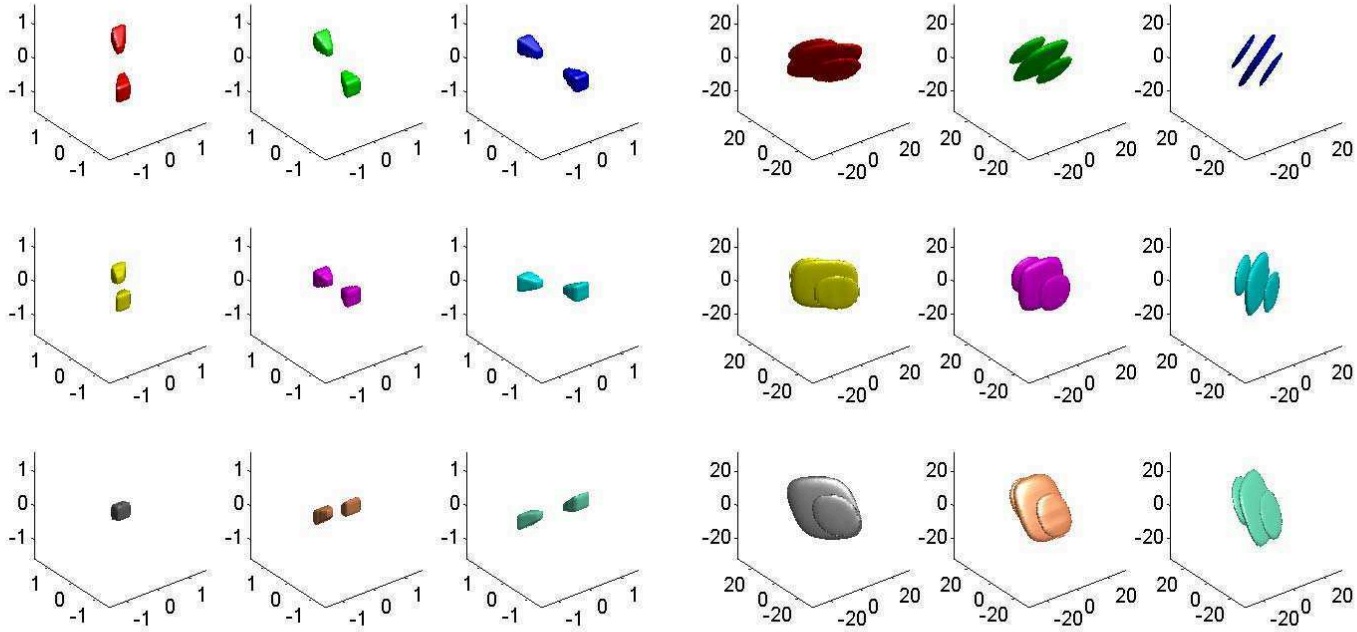
The decimation matrices for M dimensional signal are matrices of size $M\times M$. The decimation matrix for the lowpass branch $\boldsymbol{D}_0 = \mathrm{diag}\{2,\ldots,2\}$; The decimation matrices for the M groups of directional branches index by $l_1l_2\ldots l_M$ are $\boldsymbol{D}_{1l_2\ldots l_M} = \mathrm{diag}\{2,2^{n+1},\ldots,2^{n+1}\}$, $\boldsymbol{D}_{2l_2\ldots l_M} = \mathrm{diag}\{2^{n+1},2,\ldots 2^{n+1}\}$, and continue until $\boldsymbol{D}_{Ml_2\ldots l_M} = \mathrm{diag}\{2^{n+1},\ldots,2^{n+1},2\}$. The synthesis side filters are the same as the corresponding analysis filters in all three dimensions.

Example 3, *(continued)*. In 3-D case, the filters are defined from the curvelet windows in (81) and (82) as follows

$$F_0(\boldsymbol{\omega}) = 2\sqrt{2}u_0(\boldsymbol{\omega}) \quad (83)$$

$$F_{l_1l_2l_3}(\boldsymbol{\omega}) = 2^{n+2}u_{l_1l_2l_3}(\boldsymbol{\omega}), \quad (84)$$

The M-D UDCT is constructed by cascading multiples of the M-D FB described above, with the condition that number



(a) (b)

Fig. 16. Isosurface of 9 3-D uniform curvelet in (a) the frequency domain and (b) in the spatial domain.

of directions at each level follows the rule of the curvelet transform. The derivation of the result M-D curvelet function is straightforward and will not be included here. The index set for the curvelet function is similar to the 2-D case, except that the directional index $\boldsymbol{l}$ and position index $\boldsymbol{k}$ are now vectors with M components. The spatial M-D curvelet function has value oscillating along the direction defined by $\boldsymbol{l}$ index. The isosurfaces, on which the curvelet function has constant value, are approximately perpendicular to the direction $\boldsymbol{l}$. In order to show the 3-D wedge-shaped support in the frequency domain as well as the directionality in the spatial domain of 3-D curvelet functions, the isosurfaces of 9 curvelet functions are plotted in Fig. 16.

The structure of the M-D discrete uniform curvelet transform is similar to the one in Fig. 4. The M-D data are first transformed to frequency domain. Then the data are multiplied with the set of directional windows defined above. The data are then wrapped into smaller M-D hypercube to reduce redundancy before being transformed back to spatial domain.

We estimate the redundancy ratio for the UDCT in the general M dimensional case. Like other multiscale FB-based transform, the overcomplete ratio of the UDCT depends on number of scale J, but it does not depend on the number of directions at each scale, or N_j. Let us consider the M-D data block being applied to the UDCT with 3 direction on each dimension (or $N_j = 3$). The FB to implement UDCT has $M \times 3^{M-1}$ complex directional bands with decimation ratio 2^M. Therefore the number of real-valued coefficients is $M \times \left(\frac{3}{2}\right)^{M-1}$. The lowpass band also has decimation ratio 2^M. Assuming the transform is reiterated indefinitely, the redundancy is bounded by

$$\text{Red. UDCT MD} < M \times \left(\frac{3}{2}\right)^{M-1} \times \frac{1}{1-\left(\frac{1}{2}\right)^M} \qquad (85)$$

or

$$\text{Red. UDCT MD} < 2 \times M \times \frac{3^{M-1}}{2^M - 1} \qquad (86)$$

Therefore, the 2-D UDCT has a redundancy ratio less than 4 and the 3-D UDCT has a redundancy ratio less than $54/7$.

C. 3-D curvelet comparison

The 3-D FDCT is described in [29]. This implementation is a straightforward generalization of the FDCT. The 3-D curvelet windows are defined by multiplying two shearing angular windows with concentric square cubes. All window functions are then normalized to ensure that their sum is equal to one.

The contourlet transform also has a 3-D implementation in [30]. The transform is called the surfacelet transform. The structure of this transform is similar to the contourlet transform and is also implemented by a filter bank. The new feature in this generalization is a 3-D directional filter bank. Unlike the biorthogonal 2-D directional filter bank, this 3-D FB is 3 times overcomplete.

The three 3-D discrete curvelet transforms are used to decompose a synthetic seismic data cube of size $64 \times 64 \times 64$. All three decompositions have one low resolution and two directional resolutions. The overcomplete ratios of the three 3-D discrete transforms are displayed in Table I.

We believe that the overcomplete ratio is a key factor that distinguishes different curvelet implementations. The 3-D uniform curvelet has a redundancy twice higher compared to the

Name	3-D FDCT	Surfacelet	3-D UDCT
Configuration	[24 96]	[48 192]	[27 108]
Number of coeffs.	7.0e+6	8.9e+5	2.0e+6
Overcomplete ratio	≈ 27.0	≈ 3.4	≈ 7.6

TABLE I
NUMBER OF COEFFICIENTS AND REDUNDANCY RATIO FOR A DECOMPOSITION OF $64 \times 64 \times 64$ DATA.

surfacelet transform. Both of them have acceptable redundancy ratio. This is not the case for the current implementation of the FDCT transform.

VIII. IMPLEMENTATIONAL ASPECTS OF THE 2-D AND 3-D UDCT

A software package to implement the 2-D and 3-D UDCT in Matlab and Fortran has been developed at the Geoscience center, Mines ParisTech. Following are several practical aspects of the UDCT implementation.

A. Implementation, data flow and complexity of the UDCT transform

In Section V, the UDCT is constructed as multiresolution FB for the purpose of showing that the UDCT is indeed a forward and inverse discrete transform. Using FB theory, we can estimate the curvelet basis functions and establish a clear forward and inverse transform as expressed in (64) and (63).

In practice, the UDCT implementation does not need to follow a multiple level of structure as in Fig. 11(a); It is implemented directly like the structure in Fig. 11(b). All of the curvelet windows (or curvelet functions in frequency domain) are estimated at once based on the configuration (number of resolutions and number of direction for each resolution) of the transform, following (57) and (58). The set of $1+\sum_{j=1}^{J} 2N_j$ curvelet windows (or filters in frequency domain) of J resolutions are multiplied with the Fourier transform of data. The new set of windowed data is then wrapped to a rectangular according to decimation rule in the frequency domain, as in Equation (3). The decimation in frequency domain is done by shifting and adding the windowed data based on the overall decimation ratio. For example, the windowed data in Fig. 17 is decimated by 4 in the row direction, and 8 in the column direction. The frequency data are partitioned into 32 smaller matrices, 4 in the row and 8 in the column directions. Those that are not zero matrices are added together. The curvelet coefficients for each resolution and each direction are recovered by inverse FFT on corresponding smaller frequency data.

The implementation of the UDCT by pre-computing all the curvelet windows has several advantages. Once all the windows are computed, the actual forward and inverse UDCT computations are straightforward. The complexity of the transform is equal to the redundancy ratio multiplied with the complexity of FFT on data, because the complexity of window multiplication and decimation, as in Fig. 4, is negligible.

Another important advantage of the UDCT is that the computational and storage of the set of $1+\sum_{j=1}^{J} 2N_j$ curvelet windows is not a significant part of the transform. In fact, because of the inherent symmetry in each set of $2N_j$ windows in each scale, only $N_j/2$ windows need to be computed. The other $3N_j/2$ windows are obtained from the pre-computed windows by flipping and rotating operations. More over, all window functions are zero outside a region of small support. By using sparse matrix format, the memory needed for storing $1+\sum_{j=1}^{J} 2N_j$ windows is only about a fraction of the data storage. Efficient memory handling of the UDCT is very useful in processing 3-D transform, when the size of the data can be in the order of the system memory.

B. Data size and parallel structure

An implicit assumption in all implementations of the curvelet transform is that the input data are square (for 2-D) or cubic (for 3-D), and its size are power of 2 numbers. However, in industrial applications, the typical data can be of any size. The current implementation of UDCT can accommodate data of different size by indexing the directional curvelet depending on its dimension. At each scale, the UDCT can have different numbers of directions on each dimension, which means there are M values of N_j, denoted by $N_j^{(m)}$. The index for the curvelet direction $\boldsymbol{l}$ has M elements, similar to index $\boldsymbol{l}$ in (68); The first element $l_1, 1 \le l_1 \le M$ corresponds to the dimension (or the pyramid that the curvelet direction belongs to) and the rest are directions within that pyramid. For example, in the 2-D case, the directional curvelets at a certain scale will be divided into 'mostly vertical' and 'mostly horizontal' curvelets. A 2-D image that has twice the number of columns as rows, can be decomposed by a UDCT transform having twice the number of 'mostly vertical' curvelet functions compared to 'mostly horizontal' curvelet functions.

The data in an actual application of a discrete curvelet transform can also be very large. It can even be larger than the size of the memory available on the computer. The solution is to divide the data into multiple blocks to process in sequence or in parallel. A method of dividing data using overlapping tapered windows is proposed in [31] to be used with the FDCT transform. The same approach can also be used with the UDCT transform.

C. Border extension

The UDCT and FDCT are based on the FFT transform. The use of FFT transform to implement the convolution operation in the filter bank structure of the UDCT has one potential problem: it assumes that the image is periodically extended across its boundaries. Therefore, the discrete curvelet basis function is periodically extended around the opposite border of the finite size discrete 2-D data (curvelet 1 in Fig. 18). This may lead to potential border artifact in practical application of the transform. For example, a significant curvelet coefficient corresponding to a curvelet function located near the border of a 2-D seismic section may actually corresponds to an event near the opposite border.

A proposed solution to this problem is to replace the FFT with the Discrete Cosine Transform (DCT) as in [32].

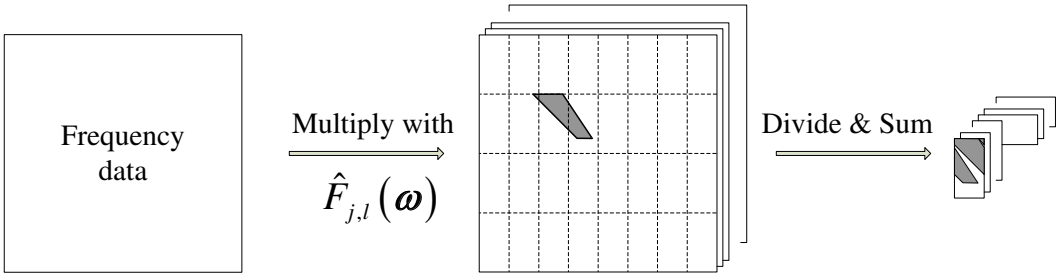


Fig. 17. Decimation in frequency domain of the UDCT. At each scale there are $2N_j$ windows, the first and second N_j data matrices are divided by $\boldsymbol{D}_{j,0}$ in (59) and $\boldsymbol{D}_{j,1}$ (60) , respectively.

Because with the DCT, one can assume symmetric extension at the border of the transformed images, the curvelet functions in [32] is mirror-extended around the border of the finite size processed data (curvelet 2 in Fig. 18). However, this approach has not completely solved the potential border artifact problem. By definition, the curvelet function is highly anisotropic at high frequency, with one dominant direction. Because of the use of the DCT, the curvelet basis near the border is also mirrored around the border edges. Thus they have two dominant directions and may not be considered as a 'true' curvelet. The potential artifact problem is that a significant curvelet coefficient belonging to a subband of a certain dominant direction may actually correspond to events at another direction that happens to mirror the direction of the subband.

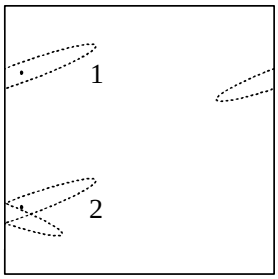


Fig. 18. Support of curvelet functions near borders of size-limited images: (1) is the case of periodic extension and (2) is the case of symmetric extension.

A practical general solution is to multiply the actual data with tapered windows. The processed data usually do not have the size optimal for the FFT. For the UDCT implementation, the data are first extended to suitable size by repeating the first and last rows (or columns). It is then multiplied with a smooth window similar to $w_0(\boldsymbol{\omega})$ in (19). The window function has a smooth transition area from 0 to 1 that starts from value 0 at the border of the extended data to reach value 1 at the border of the actual data (before being extended). This procedure leads to a slightly higher redundancy ratio. But it also removes the border artifact problem by guaranteeing two aspects. First, there is no significant feature near the border of the processed data. Second, there is no abrupt change at the border of the actual data.

IX. Numerical results

In this Section, we presents several simple applications of the UDCT transform. The parameters η_a and η_b receive values as in (48). The UDCT has also been used in image denoising and texture classification [33].

A. Application of the UDCT to image inpainting

Here we demonstrate the use of UDCT in image inpainting problem. We use the image inpainting algorithm in [34] with publicly available Matlab implementation. The original code employs the FDCT dictionary, which is replaced by the UDCT dictionary. The testing experiment assumes that 80% of number of pixels in Lena image are randomly missing. The inpainted result is nearly identical as the one produced by FDCT dictionnary. The PSNR of the inpainted images by UDCT and FDCT dictionaries are 26.44 dB and 26.56 dB, respectively.

(a) (b)

Fig. 19. Image inpainting application of the UDCT, (a) Lena image with 80% number of pixels missing PSNR=6.28 dB, and (b) inpainted Lena image by algorithm in [34] with UDCT dictionary, PSNR=26.44 dB.

B. Image and Video denoising by simple thresholding

1) Image denoising: The wavelet-type transform is widely used for image denoising. Denoising by simple thresholding also provide a simple and effective indication of the peformance of a discrete tranform. We compared the denoising performance of four directional image representations: the contourlet transform, the UDCT, the FDCT and the nonsubsampled contourlet transform (NSCT). The Lena testing image is contaminated with Gaussian white noise. The noisy image is transformed by a five-level UDCT transform. The number of subbands at five resolutions are 6, 12, 12, 24, 24. The UDCT

subband coefficients are hard-thresholded at 3 times the noise variances in the subband. For other three transforms we use the same experiment set up as in [17]. The results shows that the denoising performances are directly related to the redundancy ratio. The UDCT results are better than the contourlet transform but lower than the FDCT. Therefore, we can position the UDCT as a tranform that combine the advantages of the two decompositions: practicality and performances.

Noise variance	10	20	30	40	50
Noisy	28.13	22.13	18.63	16.13	14.2
Contourlet (1.33)	31.9	28.34	27.1	25.84	24.87
UDCT (4)	33.11	30.13	28.31	27.12	26.14
FDCT (7.6)	34.17	31.52	30.01	28.84	27.78
NSCT(53)	34.69	32.03	30.35	29.1	28.1

TABLE II

DENOISING RESULTS ON LENA IMAGE, THE NUMBERS ON THE LEFT ARE REDUNDANT RATIO OF THE TRANSFORMS.

2) Video denoising: Using 3-D curvelet transform to denoise video data should provide even better performances than in 2-D image denoising, because it exploit temporal correlation between different frames. We test the UDCT decomposition in denoising video data contaminated by white Gaussian noise.

The experiment setups are the same as in [30]. The video data decomposed by the UDCT transform into one lowpass and four directional scale. Number of directional subbands from coarse to fine scale are 27, 27, 108 and 432. The hard threshold is set at three times the noise variances, multiplied with the norm of the curvelet function. As expected, the result in Table III shows that the performance of the UDCT is better than the surfacelet transform. The improvement can be explained by higher redundancy ratio and better frequency characteristics of the UDCT curvelet, compared to the filters used in the surfacelet FB. We do not include the FDCT in the experiment because with the current implementation, the redundancy is too high and can not run on our workstation.

Noise variance	30	40	50
Surfacelets [30] (3.4)	25.86	24.72	23.88
3-D UDCT (7.6)	26.63	25.1	23.93

TABLE III

DENOISING RESULT ON MOBILE DATA CUBE OF SIZE $192 \times 192 \times 192$ AS IN [30].

X. CONCLUSION

We have presented a novel discrete curvelet transform, which is called Uniform Discrete Curvelet Transform. It can be considered an 'engineering' approach to implement the curvelet transform. It has several advantages over existing transforms, such as lower redundancy ratio, hierarchical data structure and ease of implementation. Applications of the transform has been reported in research as well as in industrial applications.

XI. ACKNOWLEDGEMENTS

The authors would like to thank Shell E&P for partly funding the project and for permission to publish this work. They would also like to thank Fons ten Kroode (Shell E&P), Mark Noble, Pascal Podvin (Mines ParisTech) and Richard Dyer (Fugro Seismic Imaging) for fruitful discussions.

REFERENCES

[1] S. Mallat, *A Wavelet tour of signal processing*, 2nd ed. Academic Press, 1999.
[2] M. Vetterli, "Wavelets, approximation and compression," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 59–73, 2001.
[3] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transform," *IEEE Transaction on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar 1992.
[4] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.
[5] ——, "Image processing with complex wavelets," *Phil. Trans. Royal Society London A*, vol. 357, no. 1760, pp. 2543–2560, Sept 1999.
[6] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861–899, 2006.
[7] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis*, vol. 25, pp. 25 – 46, July 2008. [Online]. Available: http://www.shearlet.org/
[8] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, pp. 2107– 2116, Dec. 2005.
[9] E. J. Candès and D. L. Donoho, "Curvelets: A surprisingly effective nonadaptive representation for objects with edges," in *Curve and Surface Fitting: Saint-Malo 99*, A. C. et al., Ed. Nashville, TN: Vanderbilt University Press, 2000, pp. 105–120.
[10] E. J. Candès and L. Demanet, "The curvelet representation of wave propagators is optimally sparse," *Communications on Pure and Applied Mathematics*, vol. 58, no. 11, pp. 1472–1528, 2005.
[11] F. Herrmann, "Optimal imaging with curvelet," in *70th SEG Annual Meeting and Exposition, Expanded abstracts*, 2003.
[12] H. Douma and M. V. de Hoop, "Leading-order seismic imaging using curvelets," *Geophysics*, vol. 72, no. 6, pp. S231 – S248, 2007.
[13] H. Chauris and T. T. Nguyen, "Seismic demigration/migration in the curvelet domain," *Geophysics*, vol. 73, no. 2, pp. S35 – S46, 2008.
[14] J.-L. Starck, E. J. Candès, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Transaction on Image Processing*, vol. 12, no. 6, pp. 706–717, June 2003.
[15] J. Ma and G. Plonka, "Combined curvelet shrinkage and nonlinear anisotropic diffusion," *IEEE Transaction in Image Processing*, vol. 16, pp. 2198–2206, Sep. 2007.
[16] H. Chauris, "Seismic imaging in the curvelet domain and its implications for the curvelet design," in *73th SEG Annual Meeting and Exposition, Expanded abstracts*, 2006.
[17] A. L. Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089–3101, 2006.
[18] T. T. Nguyen and S. Oraintara, "The shiftable complex directional pyramid, part 1: Theoretical aspects," *IEEE Transaction on Signal Processing*, vol. 56, no. 10, pp. 4651–4660, Oct 2008.
[19] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with C^2 singularities," *Communication on Pure and Applied Mathematics*, vol. 57, pp. 219–266, 2004.
[20] P. Vaidyanathan, *Multirate Systems and Filter Banks*. Prentice-Hall,Englewood Cliffs, NJ, 1993.
[21] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transaction on Image Processing*, vol. 11, no. 6, pp. 670–684, June 2002.
[22] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Journal on Applied and Computational Harmonic Analysis*, vol. 19, pp. 340–358, Nov. 2005.
[23] F. J. Herrmann, U. Boeniger, and D. J. Verschuur, "Nonlinear primary-multiple separation with directional curvelet frames," *Geophysical Journal International*, vol. 170, pp. 781–799, 2007.

[24] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transaction on Image Processing*, vol. 12, no. 11, pp. 1338–1351, Nov 2003.

[25] T. T. Nguyen and S. Oraintara, "On the aliasing effect of the contourlet filter banks," in *Proc. of the 14th European Signal Processing Conference (EUSIPCO 2006)*, Florence, Sept. 2006.

[26] Y. Lu and M. N. Do, "A new contourlet transform with sharp frequency localization," in *Proc. of International Conference on Image Processing (ICIP 06)*, Oct. 2006, pp. 1629–1632.

[27] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov 2005.

[28] F. C. A. Fernandes, R. L. C. van Spaendonck, and C. S. Burrus, "A new framework for complex wavelet transforms," *IEEE Transaction on Signal Processing*, vol. 51, no. 7, pp. 1825–1837, Jul 2003.

[29] L. Ying, L. Demanet, and E. J. Candès, "3D discrete curvelet transforms," in *Proc. of SPIE Conference on Wavelet Applications in Signal and Image Processing XI*, July 2005.

[30] Y. Lu and M. N. Do, "Multidimensional directional filter banks and surfacelets," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 918–931, 2007.

[31] D. Thomson, G. Hennenfent, H. Modzelewski, and F. Herrmann, "A parallel windowed fast discrete curvelet transform applied to seismic processing," in *73th SEG Annual Meeting and Exposition, Expanded abstracts*, 2006.

[32] L. Demanet and L. Ying, "Curvelets and wave atoms for mirror-extended images," in *Proc. of SPIE Conference on Wavelet Applications in Signal and Image Processing XII*, Aug. 2007.

[33] A. Vo, S. Oraintara, and T. T. Nguyen, "Statistical image modeling using distribution of relative phases in the complex wavelet domain," in *Proceedings of The 16th European Signal Processing Conference (EUSIPCO'08)*, Aug 2008.

[34] M. J. Fadili, J. L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *IEEE Transaction on Image Processing*, vol. 52, no. 1, pp. 64–79, 2009.

Seismic wave-equation demigration/migration

Hervé Chauris[*†] *and Mondher Benjemaa*[*‡]

ABSTRACT

Reverse-time migration is a well-known method based on single-scattering approximation and designed to obtain seismic images in the case of a complex sub-surface. It can however be a very time consuming task as the number of computations is directly proportional to the number of processed sources. In the context of velocity model building, iterative approaches require to derive a series of migrated sections for different velocity models. We propose to replace the summation over sources by a summation over depth-offsets or time delays defined in the sub-surface. For that, we provide a new relationship between two migrated sections obtained for two different velocity models: starting from one of the two images, we show how the second section can be correctly and efficiently obtained. In practice, for each time-delay, we compute a generalized source term by extending the concept of exploding reflector to non-zero offset. The final migrated section is obtained by solving the same wave equation in the perturbed model with the modified source term. We illustrate the methodology on 2D synthetic data sets, in particular when the initial and perturbed velocity models largely differ.

INTRODUCTION

The derivation of focused seismic migrated images strongly relies on the determination of an accurate velocity model to correctly predict the propagation of the wavefield. Migration Velocity Analysis (MVA) techniques require a number of migrated results for different velocity models (Al-Yahya, 1989). The quality of the velocity model and the associated stack section is iteratively improved by analyzing the coherency between different migrated common offset sections. The MVA principle has been studied and modified, among others by Lee and Zhang (1992); Jin and Madariaga (1994); Liu and Bleistein (1995); Jervis et al. (1996). In this context, Symes and Kern (1994); Symes (1998); Chauris et al. (2002); Mulder and ten Kroode (2002) made attempts to derive automatic schemes. Even for travel time tomography, an alternative method to MVA, events are usually picked on migrated sections (Lailly and Sinoquet, 1996). On the interpretation side, it would be very convenient to derive in an interactive manner how a given migrated image get distorted when the velocity model is modified (Murphy et al., 1993; Gray et al., 2000; Boschetti and Moresi, 2001). For example, the interpreters may question the position of a given reflector or may be uncertain about the uniqueness of the velocity model. In other words, there is a need for a quantitative evaluation of image sensitivity with respect to the velocity model used for migration.

We address this issue with the concept of demigration/migration developed in Chauris and Nguyen (2008): instead of migrating the full data set with different velocity models, a first migration is performed with an initial velocity model and then other migrated sections

are derived from this initial migrated image. The modeling or demigration in the initial velocity model is combined with the migration in a different velocity model. This provides the sensitivity of a migration result with respect to model parameters. In Chauris and Nguyen (2008), the approach was restricted to smooth velocity models with the use of ray theory for predicting the new images.

More realistic applications require to migrate data in models with discontinuities where the high frequency approximation is not valid anymore, typically for sub-salt or sub-basalt imaging (Albertin et al., 2001). In the context of rough interfaces and large steep dips, only the two-way reverse-time migration may correctly image the subsurface (Baysal et al., 1983; Farmer et al., 2006). This is not the case with the integral methods (Kirchhoff migration) nor with the one-wave wave equation migration. We investigate how the demigration/migration concept can be combined with reverse-time migration. In essence, the approach is similar to the one proposed by Wang et al. (2006) where they perform a post-stack demigration followed by a series of post-stack migrations in order to determine the optimal velocity model. We generalize their approach by considering *pre-stack* modeling and migration schemes. The pre-stack aspect is necessarily for imaging in complex models.

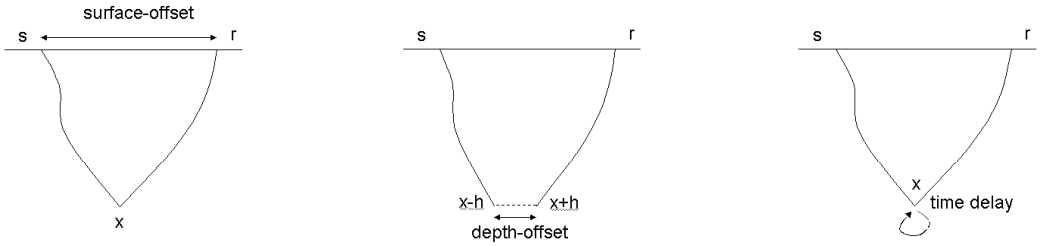
The outline of the paper is the following: we first establish the relationship between two migrated sections obtained in two different velocity models. For that purpose, we introduce a non-standard definition for Common Image Gathers (CIGs) as proposed by Sava and Fomel (2006). We then illustrate our methodology through two applications on 2D synthetic data sets. The results on the first data set allow to better understand the limitations of the predicted migrated sections. In the second application, we discuss the efficiency of the new approach.

DEMIGRATION/MIGRATION

The main idea is the following: for a given velocity model $c_0(x)$, where x denotes the spatial coordinates, we build the associated reverse-time migrated image $V_0(x,h)$. The offset h is not defined at the surface (Figure 1) but around the image point (Sava and Fomel, 2005; Shen and Symes, 2008). The exact definition we adopt here has been proposed by Sava and Fomel (2006) and is given below. For a different velocity model $c_p(x)$, one could perform a new migration and obtain $V_p(x,h)$. The alternative approach investigated here consists of studying the relationship between V_0 and V_p, for two given velocity models c_0 and c_p. The link between V_0 and V_p can be understood as a demigration (or modeling) in the initial velocity model c_0 followed by a migration in a perturbed model c_p. For identical models, this would provide the resolution operator. Studying the link between V_0 and V_p is attractive as the two sections are defined in the same domain. Demigration/migration can also be understood as velocity continuation (Hubral et al., 1996; Fomel, 2003) or image propagation (Adler, 2002).

Theory

Under the Born approximation where multiple scattering is not modeled, the data $D(s,r,\omega)$ linearly depends on the reflectivity $V_0(x,h)$. In 2D, h is a scalar and in 3D, h is a vector with two components. The coordinates (s,r) respectively correspond to the source and



(a) (b) (c)

Figure 1: Definitions of the surface-offset (a), the depth-offset (b) and the time delay (c).

the receiver positions, and ω to the angular frequency. In the original work by Sava and Fomel (2005); Shen and Symes (2008), the depth-offset reflectivity was defined up to some amplitude terms (Figure 1(b)) as

$$V_0(x,h) = -\iiint \mathrm{d}s\,\mathrm{d}r\,\mathrm{d}\omega\,\omega^2 S(\omega) G_0(s,x-h,\omega) D(s,r,\omega) G_0(r,x+h,\omega), \tag{1}$$

where G_0 is the Green's function computed in the reference velocity model c_0 and $S(\omega)$ the source term. The construction of reflectivity V_0 should be read as the correlation between the forward field from the source and the backward field from the receivers (Tarantola, 1986). Before correlation, a spatial shift is introduced to obtain a reflectivity function of offset h. For the correct velocity model c_0, most of the energy collapses at $h = 0$ (Shen and Symes, 2008). For the simplest acoustic wave equation, G_0 is the solution of the equation

$$\left(\frac{\omega^2}{c_0^2(x)} + \Delta\right) G_0(s,x,\omega) = \delta(x-s). \tag{2}$$

In this paper, we define another reflectivity W in order to obtain later a simpler relationship between W_0 and W_p computed in c_0 and c_p.

$$W_0(x,h) \;=\; -\iiint \mathrm{d}s\,\mathrm{d}r\,\mathrm{d}\omega\,\omega^2 S(\omega) G_0(s,x,\omega) D(s,r,\omega) G_0(r,x,\omega) F(s,r,x,h,\omega), \tag{3}$$

where F is a function to be defined. We impose $F(s,r,x,h=0,\omega) = 1$ to guarantee the same stack sections for $h = 0$ in equations 1 and 3. The simplest definition for F is $F = e^{i\omega h/\bar{c}} = e^{i\omega\tau}$ where $\bar{c}$ is a given constant value, similar to a velocity. Typically, $\bar{c}$ is the mean estimated velocity model. The new variable $\tau = h/\bar{c}$ is expressed in seconds (Figure 1(c)). The F term can be interpreted as a time delay to be added before correlation (Sava and Fomel, 2006). It is known that migrated images focus at a different time delay when the velocity model is inaccurate (de Vries and Berkhout, 1984; Faye and Jeannot, 1986). The new migrated section W_p is obtained by migrating the data in the perturbed velocity model c_p. It is thus associated to the Green's function G_p in c_p. As for equation 3, the perturbed migrated section is given by

$$W_p(x,h) \quad = \quad -\iiint \mathrm{d}s\,\mathrm{d}r\,\mathrm{d}\omega\,\omega^2 S(\omega) e^{i\omega h/\bar{c}} G_p(s,x,\omega) D(s,r,\omega) G_p(r,x,\omega), \tag{4}$$

where the Green's function G_p is the solution of the equation

$$\left(\frac{\omega^2}{c_p^2(x)} + \Delta\right) G_p(s,x,\omega) = \delta(x-s). \tag{5}$$

For the definition of V_0 and W_0, the selection of a single offset or time delay h corresponds to a full stack, i.e. to the contribution of all sources and receivers. By comparison, the migration of a single *surface*-offset is only a partial migration result. Note that in 2D or 3D, h is a scalar in equation 4 whereas h would be a vector in 3D for equation 1. In the remaining part, we only deal with the two-dimensional case.

Implementation

The objective is to find an efficient implementation for the computation of W_p from W_0. We summarize the key equations and their relative cost in terms of number of calculations. More details are explained in Appendix–A. Once W_p is expressed in function of W_0, we simplify the result for an efficient implementation. From W_0, we include all offsets and generate a new source term L_0 at the surface in the velocity model c_0. This source is then used to obtain W_p in c_p. More precisely, the implementation consists of three steps:

- We first compute a reference migrated image using the velocity model c_0:

$$W_0(x,h) = -\int \mathrm{d}s \int \mathrm{d}\omega\,\omega^2 S(\omega) e^{i\omega h/\bar{c}} G_0(s,x,\omega) \int \mathrm{d}r D(s,r,\omega) G_0(r,x,\omega). \tag{6}$$

  As for the reverse time migration (Tarantola, 1986), $W_0(x,h)$ can be obtained by propagating the wavefield from the source $-\omega^2 S(\omega) G_0(s,x,\omega)$, back-propagating the residuals $D(s,r,\omega) G_0(r,x,\omega)$, including a correction term $e^{i\omega h/\bar{c}}$, and then correlating the two wavefields for all offset h. This operation has to be repeated for all sources. In practice, the cost of the correlation is negligible compared to the cost of the modeling. In particular, the integration over ω can be limited to the support of the source function $S(\omega)$. The total cost is thus $2N_sC$, where N_s is the number of sources and C the cost of a (forward or backward) modeling.

- For each offset h, we compute a new source term $L_0(x,h,\omega)$ where x is located at the surface (otherwise $L_0 = 0$). After simplifications (equations A-5–A-12), the source term reads

$$L_0(x,h,\omega) = \omega^2 \hat{S}(\omega) \iint \mathrm{d}y\mathrm{d}k\, e^{-i\omega(k-h)/\bar{c}} W_0(y,k) G_0^*(x,y,2\omega) \tag{7}$$

  for x at the surface. An efficient computation can be obtained in a similar approach as for the exploding reflector concept (Lowenthal et al., 1976), where each point x

in the sub-surface is activated as a source, with a weight $\int \mathrm{d}k\, e^{-i\omega(k-h)/\bar{c}} W_0(y,k)$. The resulting signal is recorded at the surface, after half-time propagation due to the factor 2 in equation 7. This has to be repeated for all offsets h. The total cost is thus $N_h C/2$, where N_h is the number of depth-offsets.

- Finally, for all offsets h, we need to solve the wave equation

$$\left(\frac{\omega^2}{c_p^2(x)} + \Delta\right) Q_p(x,h,\omega) = L_0(x,h,\omega) \tag{8}$$

We then obtain $W_p(x,h)$ by solving

$$W_p(x,h) \simeq \int \mathrm{d}\omega\, Q_p(x,h,\omega) = Q_p(x,h,t=0) \tag{9}$$

This has to be repeated for all offsets, the associated CPU cost is thus $N_h C/2$, as we integrate for half of the time only (until $t=0$). Equation 9 is not an exact result due to the approximation detailed in Appendix-A.

The algorithm is summarized here:

```
FOR EACH SOURCE
    FORWARD MODELING IN THE INITIAL MODEL
    BACKWARD MODELING OF RESIDUALS IN THE INITIAL MODEL
    CORRELATION OF THE TWO FIELDS WITH A TIME-DELAY
END
```

This provides the initial migrated section, function of h.

```
FOR EACH TIME-DEALY
    DEMIGRATION IN THE INITIAL MODEL TO COMPUTE THE SOURCE TERM
END
```

The generalized source term is obtained for each time-delay by using the exploding reflector concept.

```
FOR EACH TIME-DELAY
    SOLVING THE WAVE EQUATION IN THE PERTURBED MODEL
    USING THE GENERALIZED SOURCE
END
```

The predicted migrated section is derived in the perturbed velocity model.

In a direct approach, W_p would be obtained for the same cost as W_0 (i.e. $2N_sC$). The gain to get an additional W_p is thus $2N_s/N_h$. Typically, $N_s = 10^3$ and $N_h = 10^2$. It means that we can get additional migrated sections for a very limited cost once the generalized source term L_0 has been obtained. Note that there is currently no limitation on the perturbed velocity model. We test the validity of this approach by comparing the demigration/migration with the classical migration on different 2D synthetic data sets.

Wave-equation demigration/migration

APPLICATIONS

We shall distinguish between the exact, initial and perturbed velocity models. The exact model is the one used to generate the shot gathers using a time domain finite difference code of order 4 in time and 8 in space (Noble, 1992). The initial velocity model c_0 is used to derive the initial migrated section W_0. This result is considered as the starting point for the demigration/migration. The tested or perturbed velocity model c_p is another model used to generate the reflectivity W_p and may not be close to the exact velocity model. In a real application, c_p would for instance be obtained by tomography or defined by the interpreter to test the sensitivity of the migrated image with respect to the velocity model used for migration.

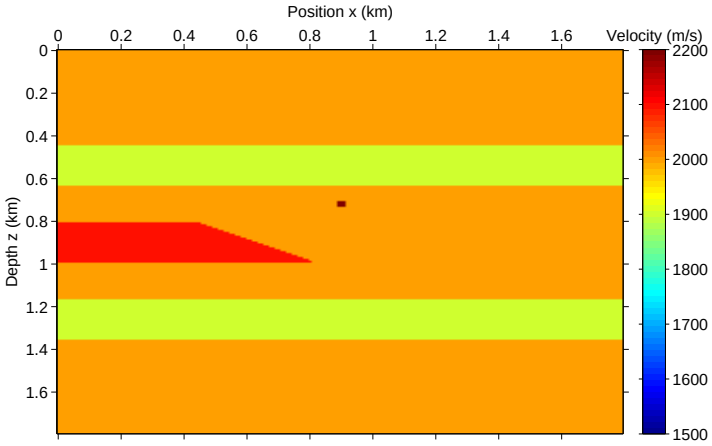
Simple synthetic data set

The first velocity model consists of three horizontal layers, including a wedge, and a diffraction point. Both $c_0 = 1500$ m/s and $c_p = 2000$ m/s are homogeneous. The c_p model can be considered for migration as a smooth version of the exact velocity model (Figure 2).

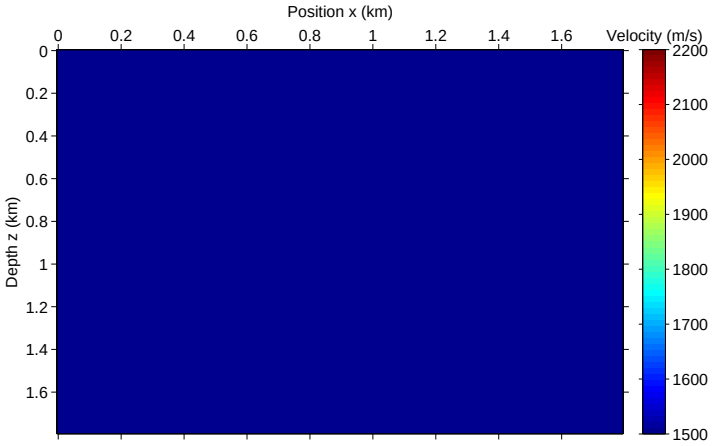
More than 200 different offset sections are generated for $W_0(x, h)$ in the initial velocity model. The zero-offset section shows the main structures but at a wrong depth and/or not perfectly focused (Figure 3). These sections are used to compute the predicted image W_p which is compared to the result obtained by directly migrating the shot gathers in c_p (Figure 4). Energy is there focused at the right position. More precisely, extracted logs at the central position of the model $x = 0.9$ km show a nice agreement in terms of kinematics between the two sections (Figure 5). The stacking power increases as c_p is close to the exact model (Figure 5(a)). One can however observe more spurious oscillations (Figure 5(b), around 1.5 km) and a lower resolution, with significant frequencies reduced from 21 to 18 Hz (Figure 5(c)). These effects are due to the simplification applied to modify the relationship between W_0 and W_p.

The generalized source term $L_0(x, h, \omega)$ for x at the surface is displayed in Figure 6(a) for $h = 0$ and compared to the *surface* zero-offset section. In this case, where only one shot over ten was computed, most of the traces are empty (Figure 6(b)). This was made on purpose to prove that L_0 is different from a surface zero-offset section. The associated migrated result suffers from the classical migration artifacts (Figure 7(a)). These artifacts are no longer present in the W_p result when signals from different surface-offsets are combined. The derivation of $W_p(x, h = 0)$ only from $W_0(x, h = 0)$ can be obtained by artificially setting $W_0(y, k) = W_0(y, 0) \cdot \delta(k)$ in equation 7. This does not provide the correct answer either (Figure 7(b)), meaning that information from different offsets should be combined to deduce W_p from W_0, according to equations 7 and 8.

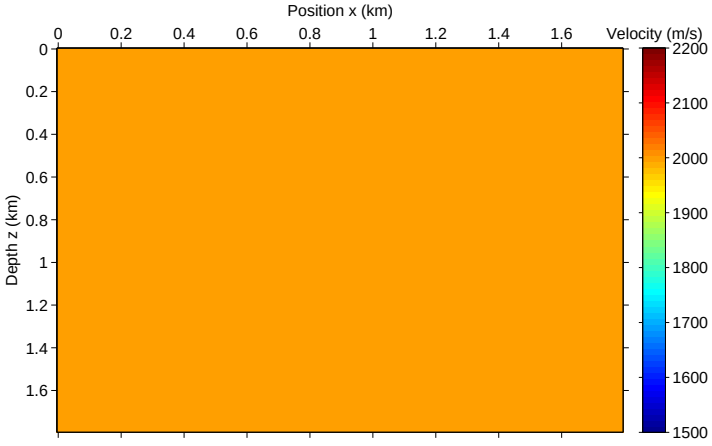
Finally, CIGs were generated for $x = 0.9$ km and displayed function of $\tau = h/\bar{c}$. In Figure 8, we compare CIGs obtained either by migration or by demigration/migration, in c_0 and c_p. The unusual linear trend is directly related to the definition of CIGs and is also observed in Sava and Fomel (2006). From this example, we conclude that the approximation used to derive the final equation 9 is robust in terms of kinematics, but leads to wrong amplitudes. Beyond these limitations, a new migrated section for a single h can be obtained for half a modeling cost, once the generalized source is computed.



(a)



(b)



(c)

Figure 2: Exact (a), initial c_0 (b) and perturbed c_p (c) velocity models.

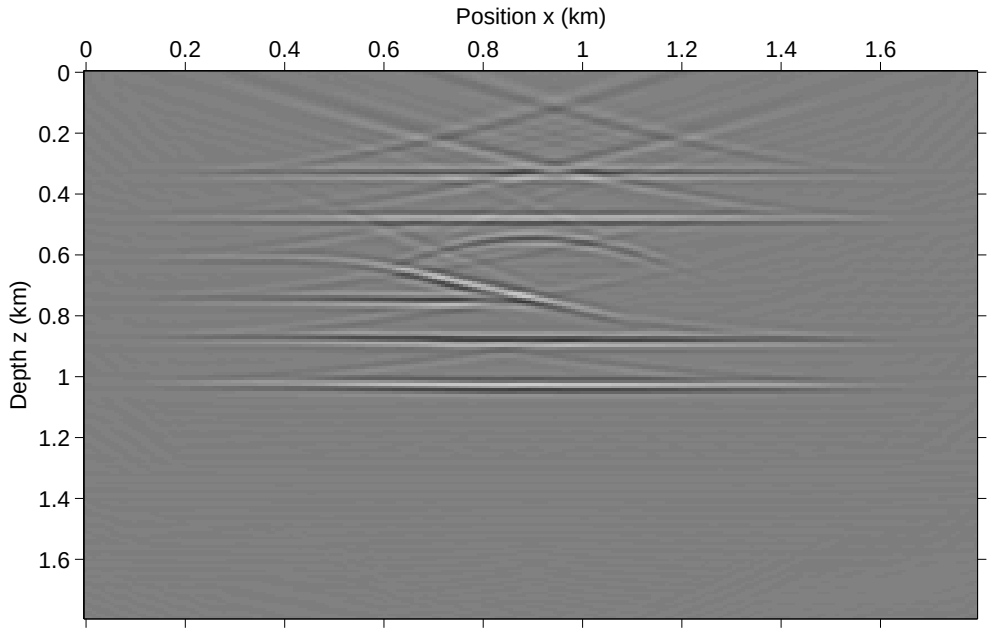


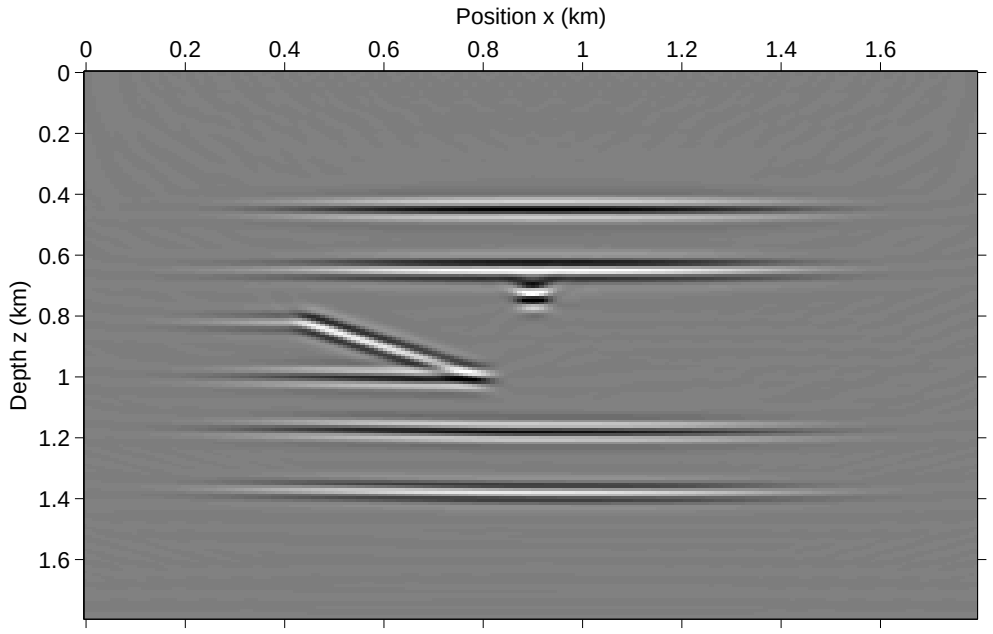
Figure 3: Image after migration in the initial velocity model c_0 for $h = 0$.

Salt-dome synthetic data set

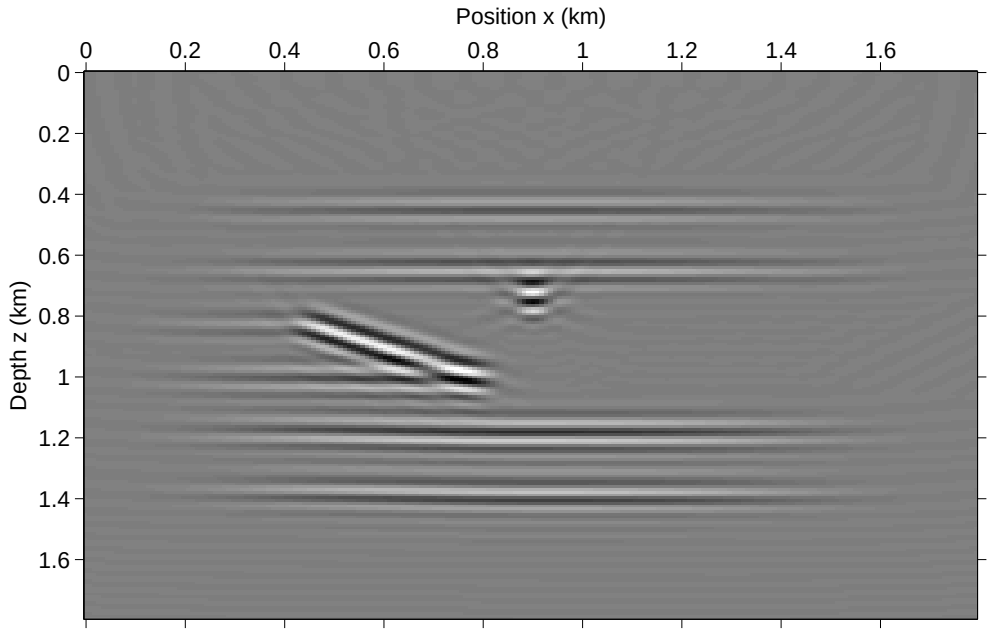
The same approach is carried on a more complex model. The exact model (Figure 9) is composed of a salt body extracted from the BP velocity model (Billette and Brandsberg-Dahl, 2005) superimposed to a velocity model with a dense reflectivity. The velocity within the salt is 3500 m/s. The model is discretized on a grid of 320 by 1200 points with a vertical and horizontal distance of 9 m and a maximal frequency of 30 Hz. The initial velocity model is homogeneous at 1500 m/s, far from the exact model. Two models c_{p1} and c_{p2} are tested, as it would be the case in an iterative tomographic approach. For the two models, the velocity outside the salt is a smooth version of the exact velocity model. For the first model, salt velocity is filled up from top salt to bottom model. In the second test, we have used the exact salt dome.

On the associated initial sections W_0, one could recognize the top salt, but deeper structures would be more difficult to be interpreted (Figure 10). Classical artifacts in the very shallow part could be removed by a Laplacian filter (Zhang and Sun, 2009). In the c_{p2} model, energy is nicely focused. The predicted W_{p1} and W_{p2} sections are similar except near the salt body (Figure 11). In both cases, bottom salt can be recognized. Only in W_{p2}, the sediment below salt and close to the salt flank is correctly imaged.

The L_0 source term is displayed in Figure 12(a) for $h = 0$. Once L_0 has been computed, the derivation of W_p becomes very efficient. If $W_{p2}(x, h = 0)$ is derived from $W_0(x, h = 0)$ as for the previous case, then the result suffers from artifacts (Figure 12(b)), for example around locations $x = 3$ km and $z = 2$ km or $x = 8.5$ km and $z = 1$ km.

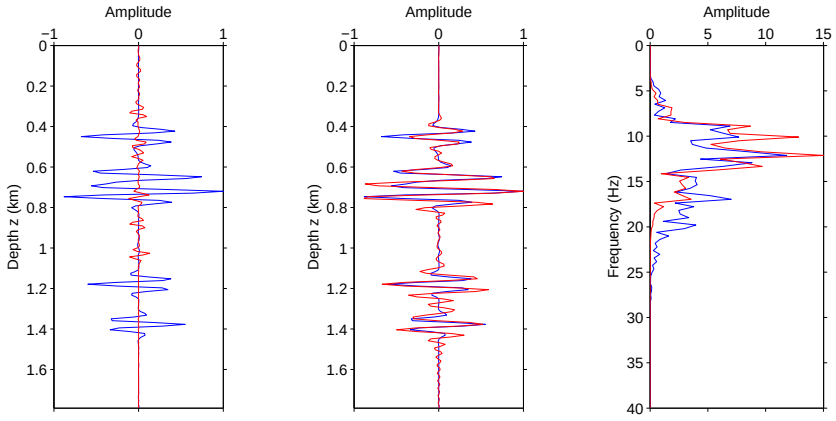


(a)



(b)

Figure 4: Images for $h = 0$ after migration (a) and demigration/migration (b) in the perturbed velocity model c_p.



(a) (b) (c)

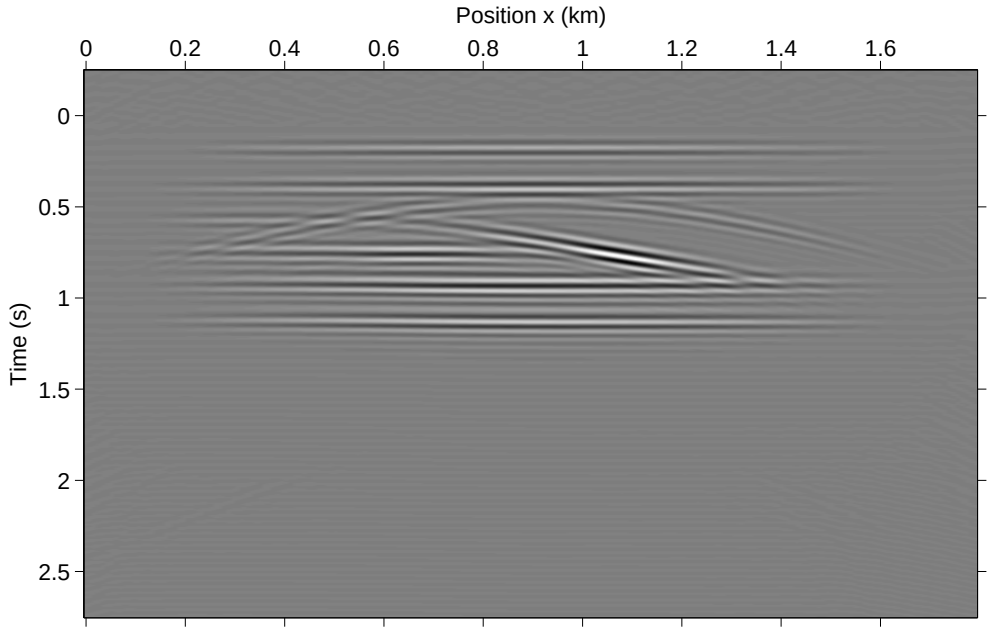
Figure 5: Comparisons between the migrated sections in the c_0 (red line) and in the c_p (blue line) models (a), the migration (blue line) and demigration/migration (red line) results in c_p (b) and the equivalent of (b) in the Fourier domain (c). All traces are selected for $x = 0.9$ km.

DISCUSSION

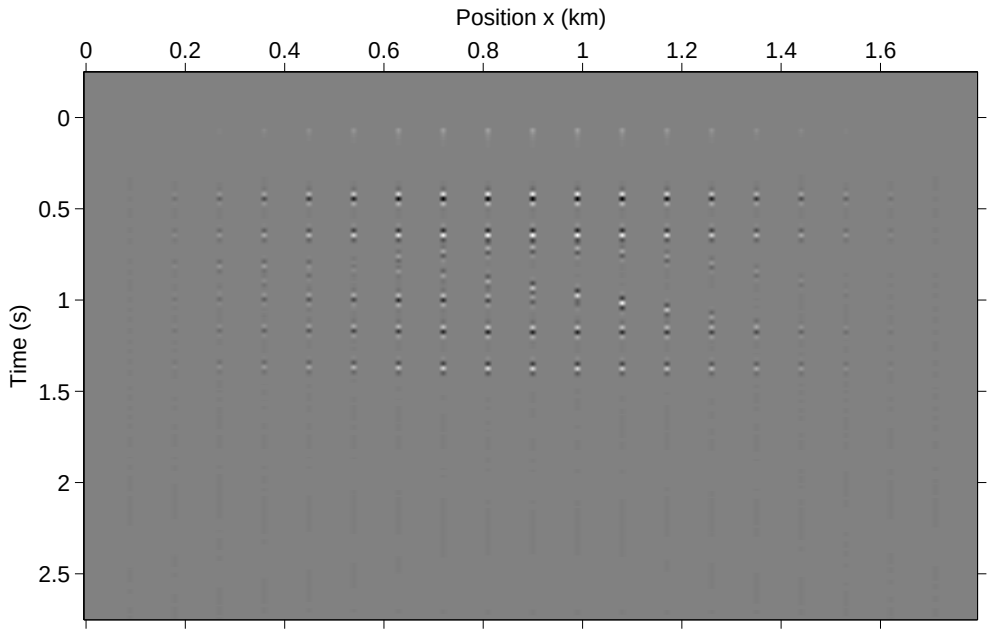
In order to obtain an efficient strategy to derive the predicted migration section W_p from the reference section W_0, we have used an alternative definition for CIGs (Sava and Fomel, 2006) and provided an approximation of the relationship between the migrated images in different velocity models. As a result, the predicted sections after demigration/migration are not well calibrated in terms of amplitude and one can observe additional oscillations. However, the kinematic is well predicted.

The derivation of W_p is first obtained by generating a new source term L_0. With the exploding reflector concept, the associated CPU-cost is proportional to the number of depth offsets, and not to the number of sources. More precisely, the CPU-cost for getting L_0 is $N_h C/2$ versus $2N_s C$ for the classical approach. In the salt dome example, $N_s = 120$ (1 shot over 10 is selected) and $N_h = 83$. It thus goes 5.8 times faster for the derivation of L_0. After that step, the cost of the derivation of W_p for a single h is only $C/2$. A single h corresponds to a full stack with a possible time delay, and not to a single *surface*-offset section. The number of shots mainly depends on the size of the investigated zone, whereas the total number of depth offsets needed to construct W_p from W_0 depends on how far c_0 is from the exact velocity model. If c_0 is indeed close to the exact model, then most of the energy is localized around $h = 0$ and the number of offsets could thus be reduced. In essence, the proposed approach is similar to an offset-domain reverse time migration, where the summation over shots is replaced by a summation over offsets. For any homogeneous velocity models c_0, one could analytically solve the wave equation for different source terms to obtain W_0 and then L_0. Only the second step for deriving W_p requires to numerically solve the wave equation, except for very simple c_p models. This could be a way to further speed-up the derivation of L_0.

With the original definition of CIGs (equation 1), where a space-shift imaging condi-

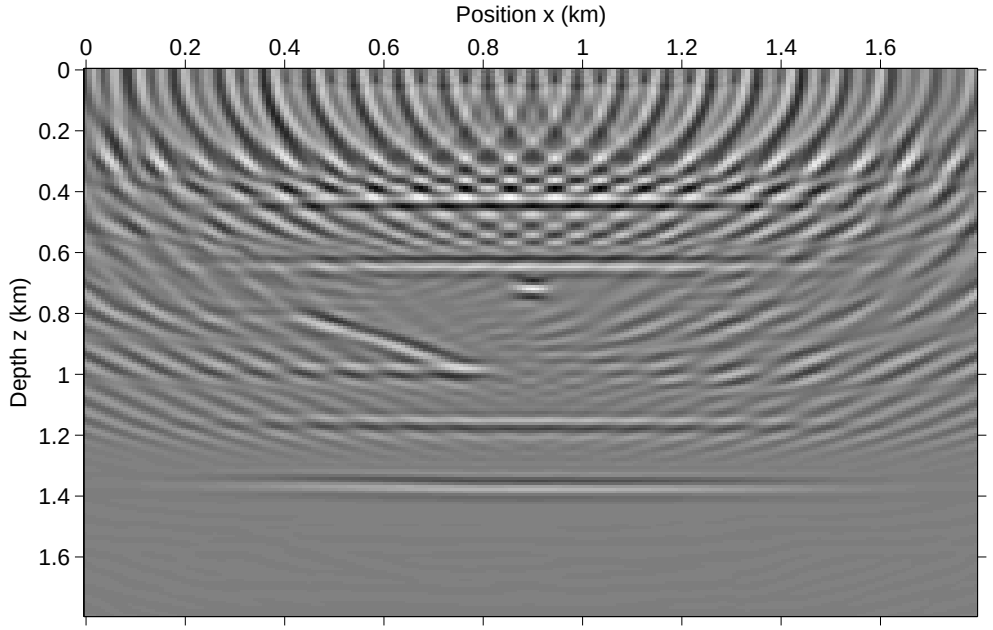


(a)

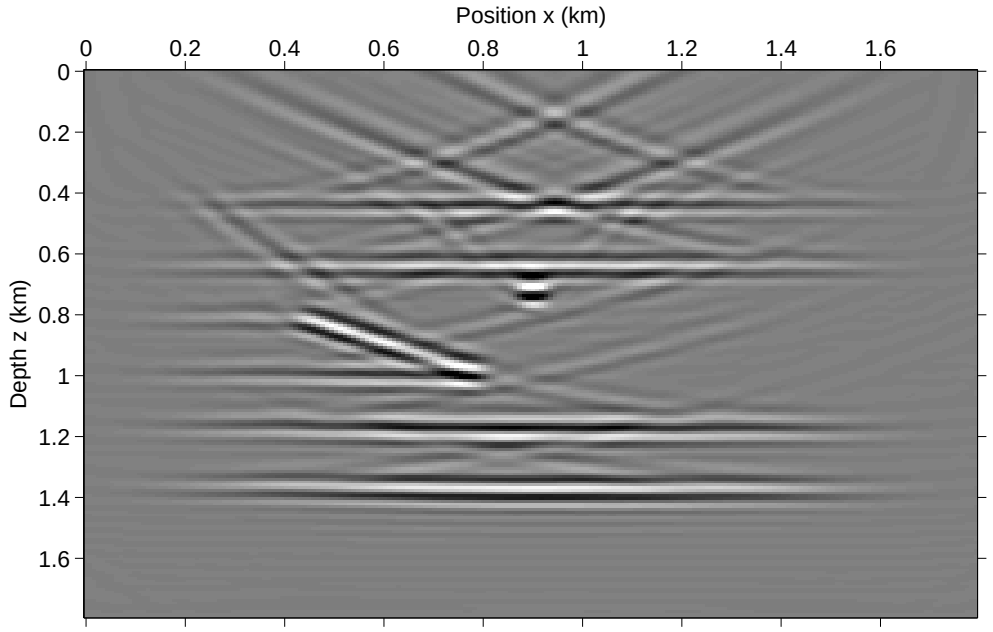


(b)

Figure 6: Generalized source for zero time delay (a) and *surface* zero-offset section for the same number of input shots (b).



(a)



(b)

Figure 7: Migration of the surface zero-offset data from Figure 6(b) (a) and after demigration/migration using only the zero time delay section (b) which should be compared to Figure 4(b).

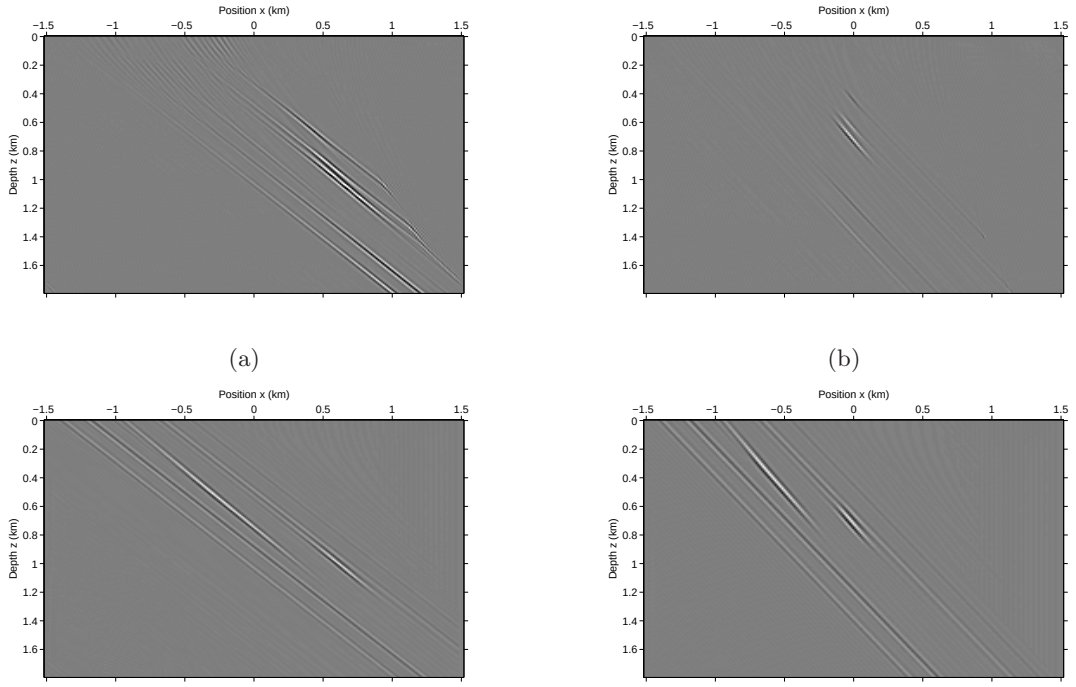


Figure 8: Common Image Gathers after migration (a and b) and after demigration/migration (c and d), in the initial velocity model c_0 (a and c) and perturbed velocity model c_p (b and d).

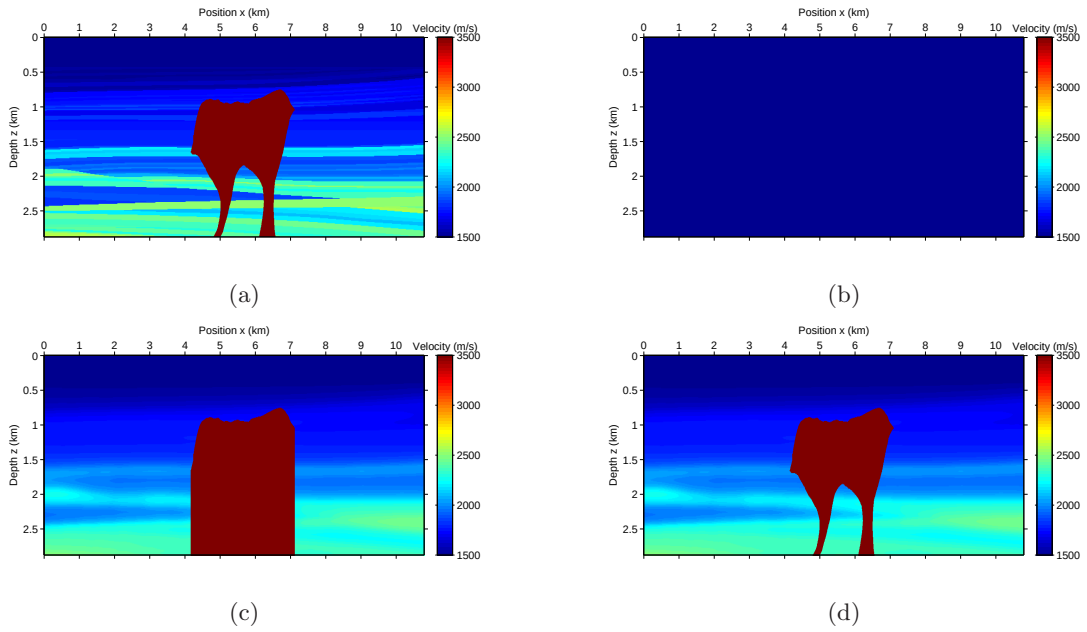
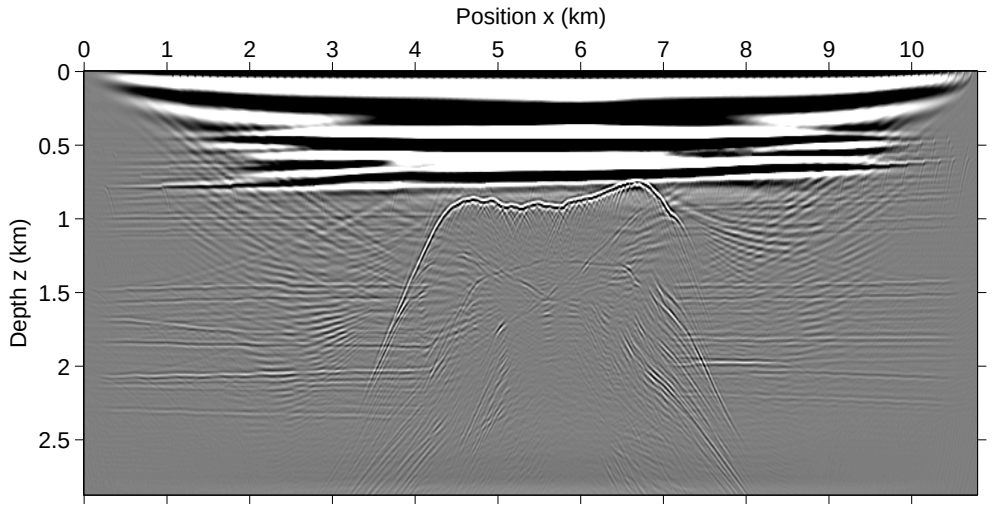
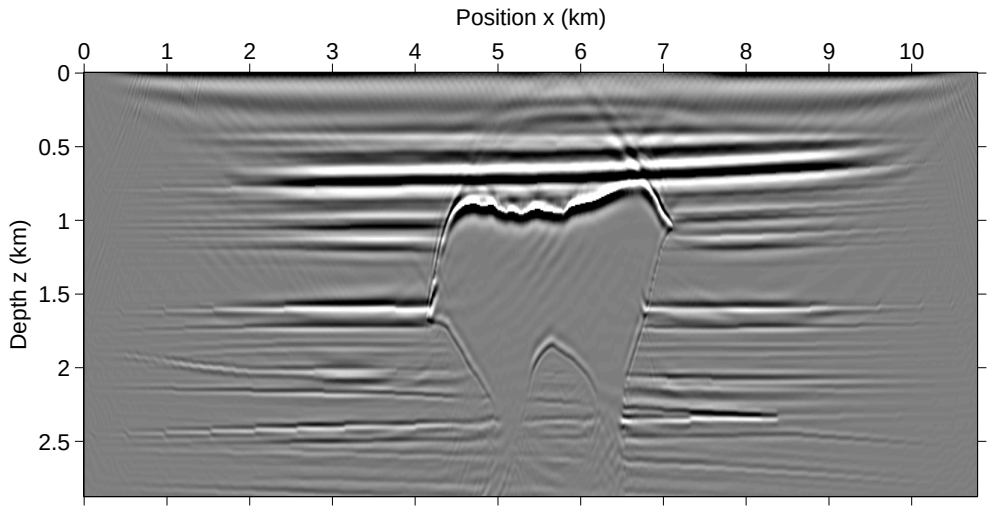


Figure 9: Exact (a), initial c_0 (b), intermediate c_{p1} (c) and perturbed c_{p2} (d) velocity models.

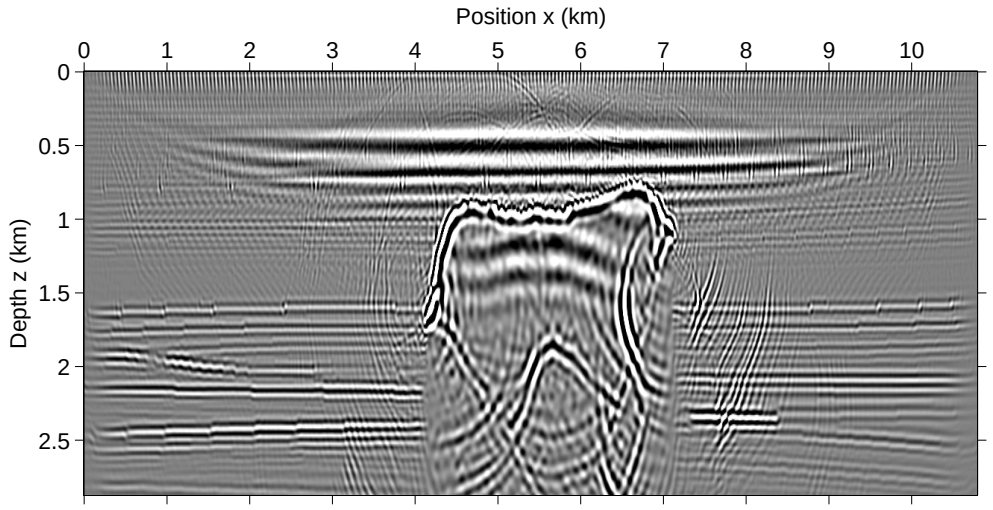


(a)

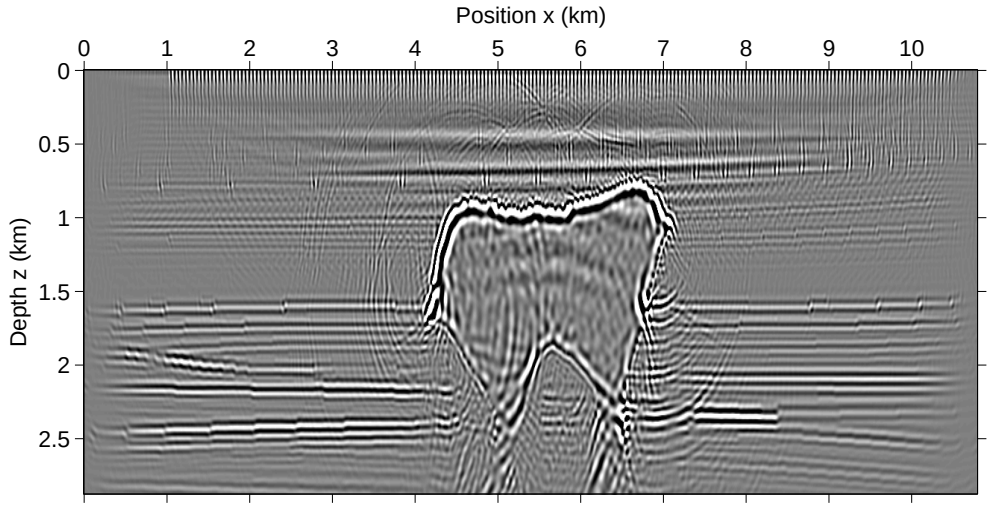


(b)

Figure 10: Migration in the initial c_0 (a) and c_{p2} (b) velocity models for $h = 0$.

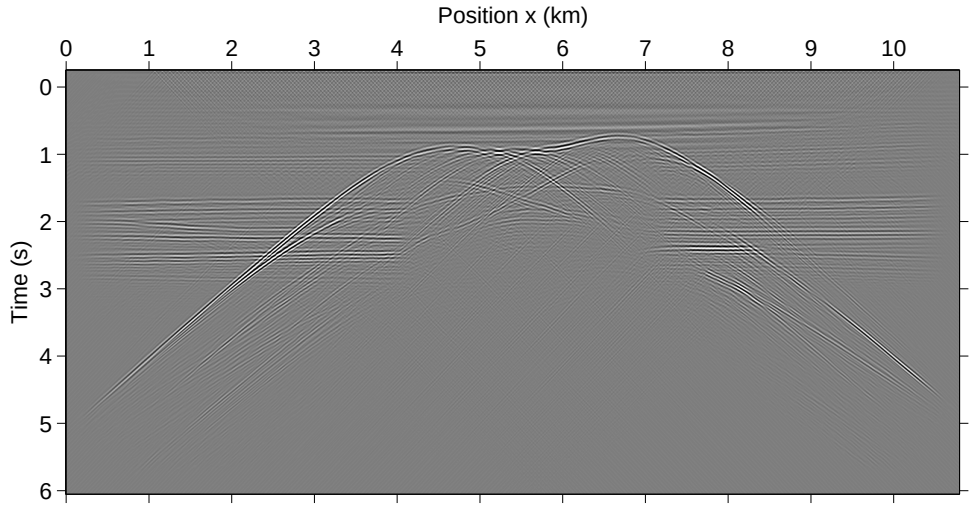


(a)

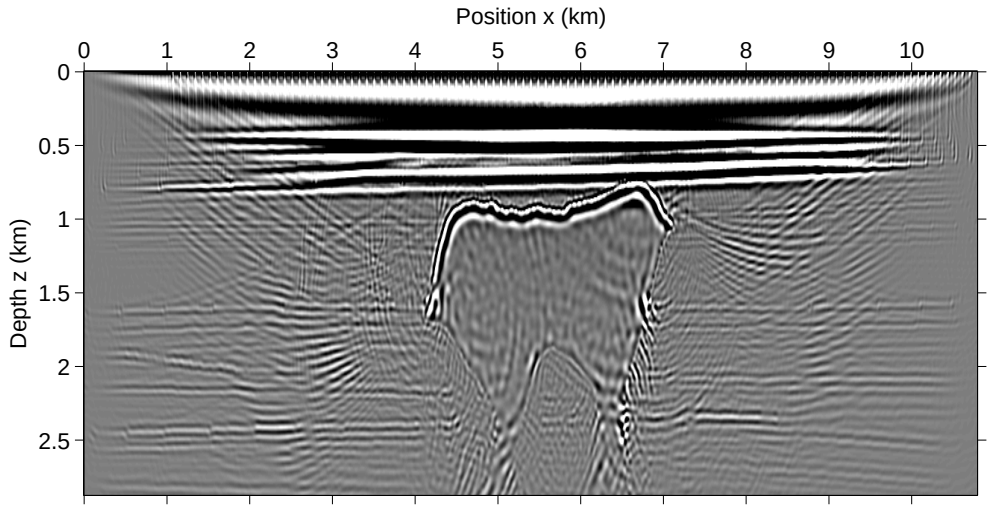


(b)

Figure 11: Demigration/migration in the intermediate c_{p1} (a) and perturbed c_{p2} (b) velocity models, when all offsets h are used in W_0.



(a)



(b)

Figure 12: The new source for zero time delay (a), and the result of demigration/migration in the perturbed velocity model c_{p2} only using the zero time delay section $W_0(x, h = 0)$ (b).

tion is applied, a similar methodology cannot be directly transposed. The equivalent of equation A-4 would not have a squared factor due to the dissymmetry between sources and receivers: either $x - h$ at the shot side and $x + h$ at the receiver side in the Green's functions. Equation A-4 cannot be simplified in the same way needed to obtain an efficient implementation.

In Xie and Yang (2008), the authors are also interested in the dependency of a migrated result with respect to the velocity model used for migration. But the objective is rather different: they derive a relationship between the residual moveout and the unknown velocity model in order to design a strategy for velocity estimation. In this paper, we concentrate on an efficient prediction of a new migrated section that should be valid in complex models.

In the classical approach proposed by Shen and Symes (2008), the 2D and 3D formalisms do not differ: the index h is either a scalar or a vector. Here, $h/\bar{c}$ is a time-delay and there is no direct extension to a "vector" time-delay. For example in Sava and Fomel (2006), $h/\bar{c}$ remains a scalar in 3D. That could be an issue as the input data $D(s, r, t)$ is defined in a 5D space with two coordinates for the sources and the receivers, whereas the reflectivity $W_0(x, h)$ is a 4D function and thus does not a priori contain the same information as the input data. Beyond this aspect, the implementation is certainly more evolved in 3D. In particular, the full wavefield in 2D can be stored in memory. This is usually not the case in 3D, even for a single shot.

For the derivation, we have used the constant density acoustic wave equation. It appears that one could use the variable density or the elastic wave equation. In fact, the same solver for different velocity models and source terms has to be used to generate migration section W_0, the source term L_0 and the demigration/migration result W_p. In the current stage, more work is needed to predict accurate amplitudes. For example, Figures 8a and c are very similar in terms of kinematics, but not for the amplitudes. After prediction, energy is not only concentrated around $h = 0$. This aspect is important for subsequent automatic velocity analysis as proposed by Shen and Symes (2008), where for the exact velocity model, most of the energy is concentrated around $h = 0$.

CONCLUSIONS

Starting from an initial migrated section and modifying the offset definition, we have been able to derive a simple relationship between migrated sections obtained in different velocity models. Once a generalized source term has been computed for all depth-offsets h, the derivation of a new migrated image for a single offset almost becomes interactive. A single h value corresponds to a stack of all data with a possible time delay. The method can be used to test the sensitivity of a migration result with respect to some velocity parameters such as the salt velocity or the position of the salt interface. This work should be understood as a first step for efficient MVA in complex media.

ACKNOWLEDGMENTS

The authors would like to thank Shell E&P for partly funding the project and for permission to publish this work. They are grateful to Fons ten Kroode, Eric Duvenek (Shell E&P), Mark Noble and Alexandrine Gesret (Mines Paristech) for fruitful discussions.

APPENDIX A

RELATIONSHIPS BETWEEN REFLECTIVITY SECTIONS

General formulation

The purpose is to derive the perturbed migrated section $W_p(x,h)$ from the original reflectivity $W_0(x,h)$. Up to some amplitude terms, the adjoint of equation 3 provides the expression of the data $D(s,r,\omega)$ function of the reflectivity $W_0(x,h)$.

$$D(s,r,\omega) \quad = \quad -\omega^2 S^*(\omega) \iint \mathrm{d}x\, \mathrm{d}h\, e^{-i\omega h/\bar{c}} G_0^*(s,x,\omega) W_0(x,h) G_0^*(r,x,\omega). \qquad \text{(A-1)}$$

Combining equations 4 and A-1, it is possible to express the perturbed reflectivity W_p function of the original reflectivity W_0.

$$W_p(x,h) = \int \mathrm{d}\omega\, Q_p(x,h,\omega) = Q_p(x,h,t=0), \qquad \text{(A-2)}$$

with

$$\begin{aligned} Q_p(x,h,\omega) \quad = \quad & \omega^4 |S(\omega)|^2 e^{i\omega h/\bar{c}}, \\ & \iint \mathrm{d}y\, \mathrm{d}k\, e^{-i\omega k/c} W_0(y,k) \\ & \int \mathrm{d}s\, G_p(s,x,\omega) G_0^*(s,y,\omega) \int \mathrm{d}r\, G_p(r,x,\omega) G_0^*(r,y,\omega) \end{aligned} \qquad \text{(A-3)}$$

For sources and receivers at the same positions at the surface, the integral over s is equal to the integral over r. We thus have

$$Q_p(x,h,\omega) = e^{i\omega h/\bar{c}} \iint \mathrm{d}y\, \mathrm{d}k\, e^{-i\omega k/c} W_0(y,k) \left[\omega^2 |S(\omega)| \int \mathrm{d}s\, G_p(s,x,\omega) G_0^*(s,y,\omega) \right]^2 \qquad \text{(A-4)}$$

We now investigate a possibility to exploit the relationship (A-4) in a practical way.

Simplification

Equations A-2 and A-4 provide a linear relationship between the perturbed reflectivity W_p and the original reflectivity W_0. However, equation A-4 is not linear with respect to the Green's function due to the square. From a very practical point of view, we shall approximate the P expression defined as

$$P(x,y,\omega) = \left[\omega^2 |S(\omega)| \int \mathrm{d}s\, G_p(s,x,\omega) G_0^*(s,y,\omega) \right]^2 \qquad \text{(A-5)}$$

We temporarily replace the Green's functions G_p and G_0 by their expressions in the high frequency approximation, where $G_p(s,x,\omega) = A_p(s,x)e^{i\omega\tau_p(s,x)}$ and $G_0(s,x,\omega) = A_0(s,x)e^{i\omega\tau_0(s,x)}$. The A_p and A_0 terms are amplitude and τ_p and τ_0 travel times. The expression for equation A-5 becomes

$$P(x,y,\omega) \simeq \left[\omega^2|S(\omega)|\int \mathrm{d}s\, A_p(s,x)A_0^*(s,y)e^{i\omega[\tau_p(s,x)-\tau_0(s,x)]}\right]^2 \tag{A-6}$$

Considering large angular frequency ω and the fact that the amplitude and travel time terms are smooth terms, we can apply the stationary phase approximation

$$\int \mathrm{d}s f(s)e^{i\omega g(s)} \simeq \sqrt{\frac{2\pi}{\omega|g''(s^*)|}}f(s^*)e^{i\omega g(s^*)} \tag{A-7}$$

where $g'(s^*) = 0$. The specular source position s^* is only used as an intermediate step. Neglecting the amplitude terms, we get

$$\begin{aligned} P(x,y,\omega) &\simeq \left[\omega^{3/2}|S(\omega)|e^{i\omega[\tau_p(s^*,x)-\tau_0(s^*,x)]}\right]^2 &\text{(A-8)}\\ &= \omega^3|S(\omega)|^2e^{2i\omega[\tau_p(s^*,x)-\tau_0(s^*,x)]} &\text{(A-9)}\end{aligned}$$

Again, the stationary phase approximation yields

$$\begin{aligned} P(x,y,\omega) &\simeq \omega^2\omega^{3/2}|S(\omega)|^2\int \mathrm{d}s\, G_p(s,x,2\omega)G_0^*(s,y,2\omega) &\text{(A-10)}\\ &= \omega^2\hat{S}(\omega)\int \mathrm{d}s\, G_p(s,x,2\omega)G_0^*(s,y,2\omega) &\text{(A-11)}\end{aligned}$$

with the modified source term

$$\hat{S}(\omega) = \omega^{3/2}|S(\omega)|^2 \tag{A-12}$$

Note that we have used twice the stationary phase approximation and the high frequency approximation during the proof, but not in the final formula. In the 1-D case, there is a single source and thus there is no need for the stationary phase approximation. We thus have $\hat{S}_{1D}(\omega) = |S(\omega)|^2$.

Modified relationship

Following the previous approximation, the modified expression of the reflectivity W_p is

$$W_p(x,h) = \int \mathrm{d}\omega\, Q_p(x,h,\omega) = Q_p(x,h,t=0), \tag{A-13}$$

with

$$Q_p(x,h,\omega) = e^{i\omega h/\bar{c}} \iint \mathrm{d}y\,\mathrm{d}k\, e^{-i\omega k/c} W_0(y,k)\,\omega^2 \hat{S}(\omega) \int \mathrm{d}s\, G_p(s,x,2\omega) G_0^*(s,y,2\omega) \quad \text{(A-14)}$$

To obtain an efficient way for the computation of Q_p, one can observe that Q_p is solution of the same acoustic wave equation as in equation 2, with a particular source term. Indeed we have

$$\left(\frac{\omega^2}{c_p^2(x)} + \Delta\right) Q_p(x,h,\omega) = L_0(x,h,\omega) \quad \text{(A-15)}$$

with

$$L_0(x,h,\omega) = \omega^2 \hat{S}(\omega) e^{i\omega h/\bar{c}} \iint \mathrm{d}y\mathrm{d}k\, e^{-i\omega k/\bar{c}} W_0(y,k) \int \mathrm{d}s\, \delta(x-s) G_0^*(s,y,2\omega) \quad \text{(A-16)}$$

if x belongs to the surface, and $L_0(x,h,\omega) = 0$ otherwise.

REFERENCES

Adler, F., 2002, Kirchhoff image propagation: Geophysics, **67**, 126–134.

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

Albertin, U., M. Woodward, J. Kapoor, W. Chang, S. Charles, D. Nichols, P. Kitchenside, and W. Mao, 2001, Depth imaging examples and methodology in the Gulf of Mexico: The Leading Edge, **20**, 498–506.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 15141524.

Billette, F. and S. Brandsberg-Dahl, 2005, The 2004 BP velocity benchmark.: 67th Annual International Meeting, EAGE, Expanded Abstracts, B035.

Boschetti, F. and L. Moresi, 2001, Interactive inversion in Geosciences: Geophysics, **66**, 1226–1234.

Chauris, H. and T. T. Nguyen, 2008, Seismic demigration/migration in the curvelet domain: Geophysics, **73**, S35–S46.

Chauris, H., M. S. Noble, G. Lambaré, and P. Podvin, 2002, Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Part I: Theoretical aspects: Geophysics, **67**, 1202–1212.

de Vries, D. and A. J. Berkhout, 1984, Influence of velocity errors on the focusing aspects of migration: Geophysical Prospecting, **32**, 629–648.

Farmer, P. A., I. F. Jones, H. Zhou, R. I. Bloor, and M. C. Goodwin, 2006, Application of reverse time migration to complex imaging problems: First Break, **24**, 65–74.

Faye, J. P. and J. P. Jeannot, 1986, Prestack migration velocities from focusing depth analysis: 56th Annual International Meeting, SEG, Expanded Abstracts, 438–440.

Fomel, S., 2003, Time-migration velocity analysis by velocity continuation: Geophysics, **68**, 1662–1672.

Gray, S., S. Cheadle, and B. Law, 2000, Depth model building by interactive manual tomography: 70th Annual International Meeting, SEG, Expanded Abstracts, 914–917.

Hubral, P., M. Tygel, and J. Schleicher, 1996, Seismic image wave: Geophysical Journal International, **125**, 431–442.

Jervis, M., M. Sen, and P. Stoffa, 1996, Prestack migration velocity estimation using nonlinear methods: Geophysics, **61**, 138–150.

Jin, S. and R. Madariaga, 1994, Nonlinear velocity inversion by a two-step Monte Carlo method: Geophysics, **59**, 577–590.

Lailly, P. and D. Sinoquet, 1996, Smooth velocity models in reflection tomography for imaging complex geological structures: Geophysical Journal International., **124**, 349–362.

Lee, W. B. and L. Zhang, 1992, Residual shot profile migration: Geophysics, **57**, 815–822.

Liu, Z. and N. Bleistein, 1995, Migration velocity analysis: Theory and an iterative algorithm: Geophysics, **60**, 142–153.

Lowenthal, D., L. Lu, R. Roberson, and J. Sherwood, 1976, The wave equation applied to migration: Geophysical Prospecting, 380–399.

Mulder, W. and A. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191.

Murphy, G. E., N. D. Whitmore, and M. A. Thornton, 1993, Interactive prestack depth migration: 63rd Annual International Meeting, SEG, Expanded Abstracts, 899–902.

Noble, M., 1992, Inversion non linéaire de données de prospection pétrolière: PhD thesis, Université Paris 7, in French.

Sava, P. and S. Fomel, 2005, Coordinate-independent angle-gathers for wave equation migration: 65th Annual International Meeting, SEG, Expanded Abstracts, 2052–2055.

———, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Symes, W. and M. Kern, 1994, Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples: Geophysical Prospecting, **42**, 565–614.

Symes, W. W., 1998, High frequency asymptotics, differential semblance and velocity estimation: 68^{th} Annual SEG Meeting and Exposition, Expanded Abstracts, 1616–1619, Soc. Expl. Geophys.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

Wang, B., D. Wheaton, F. Audebert, and V. Dirks, 2006, Separation of focusing and positioning effects using wave equation based focusing analysis and post-stack modeling: 66th Annual International Meeting, SEG, Expanded Abstracts, 2445–2449.

Xie, X.-B. and H. Yang, 2008, The finite-frequency sensitivity kernel for migration residual moveout and its applications in migration velocity analysis: Geophysics, **73**, S241–S249.

Zhang, Y. and J. Sun, 2009, Practical issues in reverse time migration: true amplitude gathers, noise removal and harmonic source encoding: First Break, **27**, 53–59.