MULTISOURCE LEAST-SQUARES MIGRATION AND PRISM WAVE REVERSE TIME MIGRATION

by

Wei Dai

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Geophysics

Department of Geology and Geophysics

The University of Utah

December 2012

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Wei Dai**

has been approved by the following supervisory committee members:

Gerard T. Schuster	, Chair	**10/31/2012** Date Approved
Ron L. Bruhn	, Member	**10/31/2012** Date Approved
Michael S. Throne	, Member	**10/31/2012** Date Approved
Sergio Chavez-Perez	, Member	**10/31/2012** Date Approved
Sherif M. Hanafy	, Member	**10/31/2012** Date Approved

and by **D. Kip Solomon**, Chair of the Department of **Geology and Geophysics**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Least-squares migration has been shown to be able to produce high quality migration images, but its computational cost is considered to be too high for practical imaging. In this dissertation, a multisource least-squares migration algorithm (MLSM) is proposed to increase the computational efficiency by utilizing the blended sources processing technique. The MLSM algorithm is implemented with both the Kirchhoff migration and reverse time migration methods. In the last chapter, a new method is proposed to migrate prism waves separately to illuminate vertical reflectors such as salt flanks. Its advantage over standard RTM method is that it does not require modifying the migration velocity model.

There are three main chapters in this dissertation.

In Chapter 2, the MLSM algorithm is implemented with Kirchhoff migration and random time-shift encoding functions. Numerical results with Kirchhoff least-squares migration on the 2D SEG/EAGE salt model show that an accurate image is obtained by migrating a supergather of 320 phase-encoded shots. When the encoding functions are the same for every iteration, the I/O cost of MLSM is reduced by 320 times. Empirical results show that the crosstalk noise introduced by blended sources is more effectively reduced when the encoding functions are changed at every iteration. The analysis of the signal-to-noise ratio (SNR) suggests that an acceptable number of iterations are needed to enhance the SNR to an acceptable level. The benefit is that Kirchhoff MLSM is a few times faster than standard LSM, and produces much more resolved images than standard Kirchhoff migration.

In Chapter 3, the MLSM algorithm is implemented with the reverse time migration method and a new parameterization, where the migration image of each shot gather is updated separately and an ensemble of prestack images is produced along with common image gathers. The merits of prestack plane-wave LSRTM are the following: (1) plane-wave prestack LSRTM can sometimes offer stable convergence even when the migration velocity has bulk errors of up to 5%; (2) to significantly reduce computation cost, linear phase-shift encoding is applied to hundreds of shot gathers to produce dozens of plane waves. Unlike phase-shift encoding with random time shifts applied to each shot gather, plane-wave encoding can be effectively applied to data with a marine streamer geometry; (3) plane-wave

prestack LSRTM can provide higher quality images than standard RTM. Numerical tests on the Marmousi2 model and a marine field dataset are performed to illustrate the benefits of plane-wave least-squares reverse time migration.

In Chapter 4, I present a new reverse time migration method for imaging salt flanks with prism wave reflections. It consists of four steps: (1) migrating the seismic data with conventional RTM to give the RTM image; (2) using the RTM image as a reflectivity model to simulate source-side reflections with the Born approximation; (3) zero-lag correlation of the source-side reflection wavefields and receiver-side wavefields to produce the prism wave migration image; and (4) repeating steps 2 and 3 for the receiver-side reflections. An advantage of this method over standard RTM is that there is no need to pick the horizontal reflectors prior to migration of the prism waves. It also separately images the vertical structures at a different step to reduce crosstalk interference. The empirical results with salt model data suggest that prism wave migration can be an effective method for salt flank delineation in the absence of diving waves.

To my family.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Gerard T. Schuster, for his guidance, support and encouragement throughout my study at the University of Utah. He taught me the in-depth understanding of geophysics theories as well as the methodology of scientific research. His passion and enthusiasm for research always encourage and inspire me during my program of study. I also wish thank other members of my advisory committee, Dr. Ron Bruhn, Dr. Michael S. Thorne, Dr. Sherif Hanafy and Dr. Sergio Chavez-Perez for reading and correcting this dissertation and the advisory of my research.

I also thank all of my UTAM colleagues for their help and discussions. The help from Dr. Chaiwoot Boonyasiriwat, Dr. Weiping Cao, and Dr. Yanwei Xue is highly appreciated. I also enjoyed and benefited from the fruitful discussion with Naoshi Aoki, Samuel Brown, Shuqian Dong, Simin Huang, Shengdong Liu, Qiong Wu, Xin Wang and Ge Zhan. I also appreciate the efforts from all of my colleagues to make UTAM a great place for research and study.

I am very grateful for the financial support from the UTAM sponsors. I thank WesternGeco, Tullow Oil and Total for offering internships, from which I obtained industrial experiences. I'm in debt to Dr. John Doherty at Tullow Oil, Dr. Paul J. Fowler and Dr. Alfonso Gonzalez at WesternGeco, Dr. Paul Williamson, Dr. Fuchun Gao and Dr. Andreas Atle at Total for their guidance.

The support from computation facilities at the Center of High Performance Computation (CHPC) at the University of Utah is also highly appreciated.

CHAPTER 1

INTRODUCTION AND OVERVIEW

Least-squares migration (Nemeth et al., 1999) has been shown to have the following advantages: (1) it can reduce migration artifacts from a limited recording aperture and/or coarse source and receiver sampling; (2) it can balance the amplitudes of the reflectors; and (3) it can improve the resolution of the migration images. However, least-squares migration is usually considered to be too expensive for practical use. In this dissertation, I propose a new algorithm to combine the blended sources processing technique with least-squares migration to increase its computational efficiency. The proposed algorithm is first implemented with the Kirchhoff migration method in Chapter 2, and then with the reverse time migration method in Chapter 3. To adapt this method for data recorded with a marinestreamer geometry, plane-wave encoding can be used instead of random time-shift encoding. In the following chapters, the multisource least-squares migration algorithm is tested with synthetic and real data examples to illustrate its advantages.

When the horizontal reflectors are embedded in the migration velocity, reverse time migration can migrate prism waves correctly to illuminate the vertical reflectors such as salt flanks. In Chapter 4, I propose a new method to migrate prism waves separately and reduce the interference between primary reflections and prism waves. It also avoids the step of modifying the migration velocity or delineating the reference reflector boundaries.

1.1 Chapter 2: Least-squares Migration of Multisource Data with a Deblurring Filter

In this chapter, I propose to use a summation of phase encoded shot gathers as input data to reduce the computational burden of least-squares migration. The blended data is similar to that used in the blended sources method (Romero et al., 2000), but my proposed scheme of multisource least-squares migration (MLSM) aims to improve the image quality while reducing crosstalk noise. During the inversion, a deblurring filter is used as a

preconditioner (Hu and Schuster, 2000; Guitton, 2004; Aoki and Schuster, 2009) to speed up the convergence.

With blended sources processing, many conventionally acquired shot gathers are phase-encoded and blended together to form supergathers to reduce the computational cost and I/O burden of migration. However, blended sources processing introduces crosstalk noise, which needs to be removed from the final migration images. Simultaneous sources acquisition shares some common ground with blended sources, as it reduces the acquisition cost, but also introduces crosstalk noise. In this chapter, a multisource least-squares migration algorithm is proposed to combine the strengths of least squares migration and blended sources processing to produce high quality images with low computational cost. The least-squares migration improves the image quality by suppressing migration artifacts, balancing reflector amplitudes and enhancing image resolution, and blended sources processing increases the computational efficiency. During the iterations of least-squares migration, the crosstalk noise introduced by blended sources is effectively reduced. The MLSM algorithm can be implemented with any migration method and the gain in efficiency depends on the migration method. My goal is to test the effectiveness of the MLSM algorithm with a Kirchhoff migration method.

1.2 Chapter 3: Plane-wave Least-squares Reverse Time Migration

The original implementation of least-squares migration was with Kirchhoff migration (Nemeth et al., 1999; Duquet et al., 2000), but was later developed for phase shift migration algorithms (Kaplan et al., 2010; Huang and Schuster, 2012). When least-squares migration is implemented with the reverse time migration method (Tang and Biondi, 2009; Dai and Schuster, 2010; Dai et al., 2010; Wong et al., 2011; Dai et al., 2012), it can reduce not only the acquisition footprint but also the artifacts in the RTM image, while enhancing the image resolution. In addition, Romero et al. (2000); Krebs et al. (2009); Tang and Biondi (2009); Schuster et al. (2011); Dai et al. (2011, 2012) employed a phase-encoding multisource approach to increase the computational efficiency by more than an order-of-magnitude compared to conventional LSRTM.

One significant problem with random encoding LSRTM is that it requires all the encoded shot gathers to share the same receivers (fixed spread geometry). Therefore, it is not applicable to marine streamer data which are recorded by a towed receiver array (Routh et al., 2011; Huang and Schuster, 2012). To remedy this problem, I devise a plane-wave LSRTM method that can be applied to both land and marine datasets (An alternative

remedy is to use frequency selection encoding, as proposed by Huang and Schuster (2012)).

Another drawback of the multisource least-squares reverse time migration algorithm is that its convergence is sensitive to the accuracy of the velocity model. When the velocity model contains large bulk errors, the migration images from different shots are inconsistent with each other, so the stacking process become less effective in reducing crosstalk noise and the resolution of the final image is spoiled. In addition, when many shots are blended together, it is difficult to separate them to produce common image gathers as quality control tools.

This problem is now remedied by incorporating a regularization term into the LSRTM method that penalizes misfits between the images in the plane-wave domain. In this way the defocusing due to velocity errors is reduced. The formulation is similar to differential semblance optimization (Symes and Carazzone, 1991) which inverts for the velocity model, but in this chapter only the reflectivity image is produced. In contrast to a stacked image, the prestack image ensemble accommodates more unknowns to allow for better fitting of the observed data, and so the convergence of least-squares migration is improved (see Appendix A).

In summary, I present a plane-wave prestack least-squares migration method where the migration image of each shot is updated separately and an ensemble of prestack images is produced with common image gathers. The advantage over conventional LSRTM where all the shot gathers are explained by a single migration image is that it is relatively less sensitive to bulk errors in the migration velocity. The plane-wave encoding technique can significantly reduce the computational and input/output (I/O) cost. In contrast to conventional multisource least-squares migration with phase-encoded supergathers, it can be applied to marine data.

1.3 Chapter 4: Reverse Time Migration of Prism Waves for Salt Flank Delineation

Vertical structures such as salt flanks are usually not illuminated by primary reflections and so cannot be well imaged by conventional migration methods (Hale et al., 1992). If on the other hand strong diving waves are present, they can be reflected from the salt flank, recorded on the surface, and migrated by a two-way migration method, such as Kirchhoff migration (Ratcliff et al., 1991, 1992) or reverse time migration (RTM) (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983). Even a one-way migration method can be modified (Hale et al., 1992) to incorporate diving waves for salt flank imaging.

If the diving wave is not extant due to the absence of a strong velocity gradient or a

limited recording aperture, prism waves can be migrated to illuminate vertical reflectors. With reverse time migration, prism waves can be migrated by embedding the subhorizontal reflection boundaries in the velocity model (Jones et al., 2007). However, incorporating the sharp boundaries into the velocity model is not trivial, and the complex migration velocity will excite complex wavefields that lead to artifacts in the RTM images (Liu et al., 2011). Another problem is that prism waves are doubly scattered waves, which are usually weaker than primaries, so that the contribution from the prism waves might be weak. In this chapter, I propose a new RTM method for migrating the prism waves separately from the other reflectors by utilizing the migration image from conventional RTM. The advantages of this approach over conventional RTM are as follows: (1) It does not require modifying the migration velocity as conventional RTM does; (2) It separately images different structures at different steps and reduces the artifacts from crosstalk of different phases. The disadvantage of the proposed method is that its computational cost is twice that of conventional RTM.

1.4 Technical Contributions in this Dissertation

Chapter 2 presents the novel technique of multisource least-squares migration to efficiently produce high quality reflectivity images. This algorithm is implemented with the Kirchhoff migration method and tested with 320 synthetic shot gathers associated with the 2D SEG/EAGE salt model. An accurate image is obtained by migrating a supergather composite of all these 320 shot gathers after 60 iterations. Compared to the conventional Kirchhoff migration image, the I/O cost of MLSM with static encoding is reduced by 320 times. The MLSM image is much more resolved than conventional Kirchhoff migration image, because the migration artifacts are suppressed, the reflector amplitudes are balanced, the image resolution is enhanced and the crosstalk noise is reduced. Two types of encoding strategies are proposed: static encoding and dynamic encoding. Their performance in crosstalk-noise reduction is studied with the measurements of signal-to-noise ratio of migration images. For the 2D SEG/EAGE salt model example, The MLSM algorithm with static encoding enjoys lower I/O cost compared to the MLSM with dynamic encoding, but the empirical results show that the MLSM with dynamic encoding, on the other hand, is more effective in reducing crosstalk noise introduced by blended sources.

In Chapter 3, the multisource least-squares migration algorithm is implemented with the reverse time migration method and the blended sources processing technique can increase the computational efficiency significantly. The random encoding strategy is not applicable for marine-streamer data, so a plane-wave encoding method is adopted. When an ensemble

of prestack images is incorporated into the inversion, prestack plane-wave least-squares reverse time migration shows the following advantages: (1) stable convergence even with velocity errors up to 5% in my example and (2) the common image gathers are available for quality control and migration velocity analysis. I conclude that the least-squares reverse time migration in the plane-wave domain can be an efficient method improving the quality of RTM images and producing common image gathers for MVA.

In Chapter 4, I proposed a new method for migrating prism waves by RTM. There are two steps to the method: (1) Conventional RTM is applied to the data to estimate the geometry of the horizontal reflectors near the salt flank; (2) Prism wave RTM is applied to the data again, except the prism imaging condition is used rather than the conventional one. A high quality image is obtained by summation of two partial migration images: one from conventional RTM and the other from the migration of the prism waves. The empirical results suggest that the proposed method can migrate the prism waves correctly to delineate salt flanks and improve the image quality.

CHAPTER 2

LEAST-SQUARES MIGRATION OF MULTISOURCE DATA WITH A DEBLURRING FILTER

Least-squares migration (LSM) has been shown to be able to produce high quality migration images, but its computational cost is considered to be too high for practical imaging. In this chapter, a multisource least-squares migration algorithm (MLSM) is proposed to increase the computational efficiency by utilizing the blended sources processing technique. To expedite convergence, a multisource deblurring filter is used as a preconditioner to reduce the data residual. This MLSM algorithm is applicable with Kirchhoff migration, wave-equation migration or reverse time migration, and the gain in computational efficiency depends on the choice of migration method. Numerical results with Kirchhoff least-squares migration on the 2D SEG/EAGE salt model show that an accurate image is obtained by migrating a supergather of 320 phase-encoded shots. When the encoding functions are the same for every iteration, the I/O cost of MLSM is reduced by 320 times. Empirical results show that the crosstalk noise introduced by blended sources is more effectively reduced when the encoding functions are changed at every iteration. The analysis of signal-to-noise ratio (SNR) suggests that not too many iterations are needed to enhance the SNR to an acceptable level. Therefore, when implemented with wave-equation migration or reverse time migration methods, the MLSM algorithm can be more efficient than the conventional migration method.

2.1 Introduction

Conventional migration (Claerbout, 1971) computes the reflectivity image by applying the adjoint operator to the data. Migration can also be interpreted as the first iteration of iterative inversion, where the Hessian of the misfit functional is approximated as a diagonal matrix. This approximation is violated when the data are incomplete (Nemeth et al., 1999)

and the migration image will be obscured by migration artifacts.

It has been shown that least-squares migration (LSM) (Nemeth et al., 1999; Duquet et al., 2000) can improve the resolution of the migration image and suppress migration artifacts. However, one of the drawbacks of least-squares migration is its high computational cost. In this chapter, I propose to use a summation of phase encoded shot gathers as input data to reduce the computational burden of least-squares migration. The blended data are similar to that used in the blended sources method (Romero et al., 2000), but my proposed scheme of multisource least-squares migration (MLSM) aims to improve the image quality while reducing crosstalk noise. During the inversion, a deblurring filter is used as a preconditioner (Hu and Schuster, 2000; Guitton, 2004; Aoki and Schuster, 2009) to speed up the convergence.

2.1.1 Blended Sources Processing

In blended sources processing, many conventionally acquired shot gathers are phase-encoded and blended together to form supergathers to reduce the computational cost and I/O burden of migration. Romero et al. (2000) first explored this idea with the wave-equation migration of synthetic data associated with the Marmousi model. They produced acceptable images with less cost than the conventional method. The limitation of their approach was that the blended sources images were always no better in quality than the corresponding conventional images, because the blended sources introduced unacceptable crosstalk noise into the final migration section. Krebs et al. (2009) presented their full waveform inversion result with blended sources encoded by random encoding functions. Their computational efficiency was increased by a factor of 50 compared to standard full waveform inversion and their method has been mostly tested for a fixed-spread acquisition geometry. The extension of blended sources processing to marine acquisition is a topic of current research.

2.1.2 Crosstalk Noise Reduction

As mentioned earlier, blended sources processing introduces crosstalk noise, which needs to be removed from the final migration images. Simultaneous sources acquisition shares some common ground with blended sources, as it reduces the acquisition cost, but introduces crosstalk noise also. The simplest approach for reducing crosstalk noise in multisource data is to use standard migration and stacking procedures. Lynn et al. (1987) showed that coherent noise in multisource data (with several shot gathers per supergather) can effectively be suppressed by weighted stacking. Hampson et al. (2008) reported their synthetic and

field data examples and showed that for 2D cases and two shot gathers per supergather, simple stacking was effective enough, but for their 3D example, they found that the shot separation technique was necessary, due to the strong reflections from the shallow water bottom. Fromyr et al. (2008) achieved similar image quality with two-source shooting as compared to conventional acquisition in their wide azimuth experiment. With careful survey design, a suitable marine environment and a small number of multiple sources, simple stacking alone might be sufficient for quality imaging. To assist in this design process, Schuster et al. (2011) provide rigorous formulas for predicting the level of crosstalk noise as a function of the encoding parameters.

2.1.3 Scope of This Chapter

In this chapter, a multisource least-squares migration algorithm is proposed to combine the strengths of least squares migration and blended sources processing to produce high quality images with low computational cost. The least-squares migration improves the image quality by suppressing migration artifacts, balancing reflector amplitudes and enhancing image resolution, and blended sources processing increases the computational efficiency. During the iterations of least-squares migration, the crosstalk noise introduced by blended sources is effectively reduced. The MLSM algorithm can be implemented with any migration method and the gain in efficiency depends on the migration method. My goal is to test the effectiveness of the MLSM algorithm with a Kirchhoff migration method.

2.2 Theory

For a fixed-spread acquisition, the phase-encoded multisource data (i.e. supergathers) can be represented as

$$\mathbf{d} = \sum_{i=1}^{S} \mathbf{P}_i \mathbf{d}_i, \tag{2.1}$$

where S is the number of multiple shots and matrix $\mathbf{P}_i$ represents the phase-encoding functions (in this study, the encoding functions involve random source time delay). All the $\mathbf{P}_i$ are chosen to be unitary so that $\mathbf{P}_i^T \mathbf{P}_i$ is equal to the identity matrix.

In equation 2.1, I define $\mathbf{d}$ as a supergather, which is the summation of shot gathers, each with shot excitation time shifted by a random time shift with a standard deviation greater than the source period. It is shown in Schuster et al. (2011) that the combination of random polarity changes, random time shifts and random shot locations is more effective

at reducing crosstalk noise than the use of any of the three encoding functions alone. I assume that the i-th CSG $\mathbf{d}_i$ and the reflectivity model $\mathbf{m}$ are related by

$$\mathbf{d}_i = \mathbf{L}_i \mathbf{m}, \tag{2.2}$$

where $\mathbf{L}_i$ is the linear forward modeling operator associated with the i-th shot. This operator can represent either a Kirchhoff or a wave-equation modeling method (Mulder and Plessix, 2004). Plugging equation 2.2 into 2.1, I get

$$\mathbf{d} = \sum_{i=1}^{S} \mathbf{P}_i \mathbf{L}_i \mathbf{m} = \mathbf{L}\mathbf{m}, \tag{2.3}$$

where the supergather modeling operator is defined as

$$\mathbf{L} = \sum_{i=1}^{S} \mathbf{P}_i \mathbf{L}_i. \tag{2.4}$$

2.2.1 Multisource Migration

From equation 2.4, the supergather migration operator is defined as the adjoint of the supergather modeling operator,

$$\mathbf{L}^T = \sum_{i=1}^{S} \mathbf{L}_i^T \mathbf{P}_i^T, \tag{2.5}$$

so that the supergather migration image is

$$\begin{aligned}
\mathbf{m}_{mig} &= \mathbf{L}^T \mathbf{d} = \mathbf{L}^T \sum_{i=1}^{S} \mathbf{P}_i \mathbf{L}_i \mathbf{m} \\
&= \sum_{j=1}^{S} \mathbf{L}_j^T \mathbf{P}_j^T \sum_{i=1}^{S} \mathbf{P}_i \mathbf{L}_i \mathbf{m} \\
&= \sum_{i=1}^{S} \sum_{j=1}^{S} \mathbf{L}_j^T \mathbf{P}_j^T \mathbf{P}_i \mathbf{L}_i \mathbf{m} \\
&= \overbrace{\sum_{i=1}^{S} \mathbf{L}_i^T \mathbf{L}_i \mathbf{m}}^{standard\ migration\ image} + \overbrace{\sum_{j \neq i} \sum_{i=1}^{S} \mathbf{L}_j^T \mathbf{P}_j^T \mathbf{P}_i \mathbf{L}_i \mathbf{m}}^{crosstalk},
\end{aligned} \tag{2.6}$$

consisting of two terms: the first term is the standard migration image and the second term is the crosstalk noise introduced by multisource blending of shot gathers. The magnitude of the crosstalk term for a variety of different phase encoding functions is derived in Schuster et al. (2011).

2.2.2 Multisource Least-squares Migration (MLSM)

In order to suppress crosstalk noise to an acceptable level when the number of multiple sources S is large, I solve equation 2.3 in the least-squares sense (Dai and Schuster, 2009; Dai et al., 2009). That is, define the objective function as

$$f(\mathbf{m}) = \frac{1}{2}||\mathbf{d} - \mathbf{Lm}||^2 + \frac{1}{2}\lambda||\mathbf{m} - \mathbf{m}_{apr}||^2, \tag{2.7}$$

so that, an optimal $\mathbf{m}$ is sought to minimize the objective function in equation 2.7. In equation 2.7, Tikhonov regularization (Tikhonov and Arsenin, 1977) is used and λ is the regularization parameter, determined by a trial and error method. Smoothness constraints in the form of second-order derivatives of the model function can expedite convergence (Kühl and Sacchi, 2003) and partly overcome the problems associated with errors in the velocity model.

With the assumption that nothing is known about $\mathbf{m}$, $\mathbf{m}_{apr}$ is set to be equal to zero. The model $\mathbf{m}$ that minimizes equation 2.7 can be found by a gradient type optimization method

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha\mathbf{F}(\mathbf{L}^T(\mathbf{Lm}^{(k)} - \mathbf{d}) + \lambda\mathbf{m}^{(k)}), \tag{2.8}$$

where $\mathbf{L}^T(\mathbf{Lm}^{(k)} - \mathbf{d}) + \lambda\mathbf{m}^{(k)}$ is the gradient, $\mathbf{F}$ is a preconditioning matrix and α is the step length. As both the forward modeling and migration operators are linear and adjoint to each other, the analytical step length formula can be used. Alternatively, in order to improve the robustness of the MLSM algorithm, a quadratic line search method is carried out with the current model and two trial models. In this study, I use the conjugate gradient (CG) method, which generally converges faster than the steepest decent method. Moreover, static encoding is used where the encoding functions are the same for every iteration to reduce the I/O cost. Boonyasiriwat and Schuster (2010) show that dynamic encoding (encoding functions are changed at every iteration) is more effective in 3D multisource full waveform inversion and so dynamic encoding results are presented as well. To ensure the convergence of MLSM, the migration velocity should be close to the true velocity model.

2.2.3 Numerical Implementation

The numerical scheme in equation 2.8 is applicable to any migration method and its associated forward modeling (demigration) operator. Each type of migration method, e.g. Kirchhoff migration, one-way wave-equation migration or reverse time migration, can be implemented in the mode of least-squares migration (Nemeth et al., 1999; Duquet et al., 2000; Kaplan et al., 2010; Dai and Schuster, 2010; Dai et al., 2010). When combined with

blended sources processing, each specific implementation will bear different advantages. The computational cost of the one-way wave-equation migration or reverse time migration method is reduced by S times when S sources are blended together. In addition, the I/O cost is reduced by S times with static encoding method. On the other hand, the computational cost of Kirchhoff migration is relatively low, but it cannot be further reduced with blended sources processing because the Kirchhoff migration operation of $\mathbf{L}_i^T\mathbf{P}_i^T$ in equation 2.5 must be applied separately to the supergather for $i = 1, 2, ..., S$. However, the I/O cost is reduced by inputting only a supergather so this will reduce the overall run time of Kirchhoff least-squares migration.

In this chapter, the multisource least-squares migration algorithm is implemented with Kirchhoff migration and tested on synthetic blended sources data. To expedite convergence, a deblurring filter (see Appendix A) is used as a preconditioner (Hu and Schuster, 2000; Guitton, 2004; Aoki and Schuster, 2009) that can reduce the migration artifacts related to Kirchhoff migration (frowns and smiles) and compensate for the energy loss from geometric spreading; and therefore, speed up the convergence. Numerical simulations are conducted to validate these statements.

2.2.4 Signal-to-noise Ratio Analysis

It is desirable to estimate the relationship between the signal-to-noise ratio (SNR, defined in Appendix B) enhancement and the number of shot gathers (S) for iterative least-squares migration of supergathers. While it affords no simple analytical expression for the dependence of SNR on the number of iterations of least-squares migration, I focus on how the SNR is reduced by iterative stacking (multiple migrations of all shots) of Romero et al. (2000), where all the shots in a survey are phase-encoded and blended together to form a supergather before migration (Figure 2.1). Here, I assume the data are noise free and the noise is defined to be the crosstalk noise only. In other words, in equation 2.6 the 'standard migration image' term is assumed to be noise free whereas the 'crosstalk' term is responsible for all the noise. For convenience, the terms in equation 2.6 are regrouped here as follows,

$$\mathbf{m}_{mig} = \sum_{i=1}^{S} \mathbf{m}_{mig,i} = \sum_{i=1}^{S} (\overbrace{\mathbf{L}_i^T\mathbf{L}_i\mathbf{m}}^{signal} + \overbrace{\sum_{j\neq i} \mathbf{L}_j^T\mathbf{P}_j^T\mathbf{P}_i\mathbf{L}_i\mathbf{m}}^{noise}). \tag{2.9}$$

In equation 2.9, I further assume the signal term and $S-1$ noise terms in the parentheses are of comparable energy, and that those $S-1$ noise terms are incoherent. Consequently the SNR is roughly $1/\sqrt{S-1}$ for $\mathbf{m}_{mig,i}$, the image associated with i-th sources . After

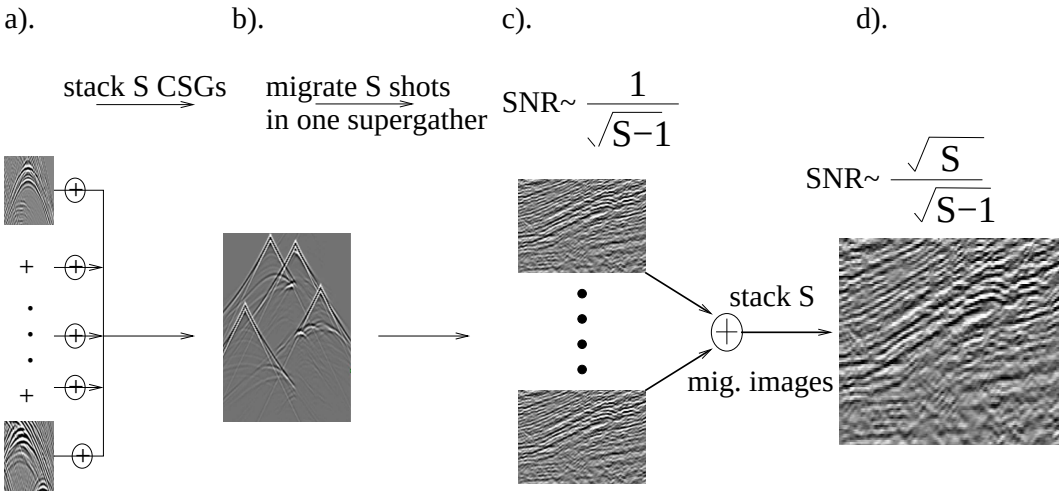

Figure 2.1. Blending and migration of a supergather: (a) time-shifted shot gathers, (b) blended supergather created by blending S time-shifted shot gathers, (c) migration images after migrating the supergather for each shot position with SNR approximately $\frac{1}{\sqrt{S-1}}$, (d) final image after summing S migration images. The final SNR is $\frac{\sqrt{S}}{\sqrt{S-1}}$.

summation over all the S sources, the SNR of $\mathbf{m}_{mig}$ is $\sqrt{S}/\sqrt{S-1}$, assuming the signal term from all the S sources are coherent.

This SNR analysis is summarized in Figure 2.1. Here, S shots in Figure 2.1(a) are encoded and stacked together to form a supergather, which is noise free, in Figure 2.1(b). The supergather is then migrated S times—once for each of the S source locations—to produce S images as shown in Figure 2.1(c). Every image contains one signal image from a correctly decoded and migrated shot and $S-1$ noisy images from the rest $S-1$ shots being migrated with wrong source locations and wrong time shifts. As analyzed before, every image in Figure 2.1(c) has a SNR approximately $1/\sqrt{S-1}$. After stacking all the S images together in Figure 2.1(d), the SNR becomes $\sqrt{S}/\sqrt{S-1}$.

Here the key assumptions are:

(1) The correctly decoded and migrated shots from all the S images give coherent signal, which will constructively stack after stacking. In addition, geometrical spreading effects can be ignored;

(2) The incorrectly decoded and migrated shots generate random noise with the same strength due to random encoding, which will destructively stack after stacking;

(3) The crosstalk noise from each migration at each iteration is uncorrelated.

$$\text{SNR} \approx \sqrt{SI}/\sqrt{S-1}. \tag{2.10}$$

In the case that there are N supergathers in the survey, the SNR is proportional to

$$\begin{aligned} \text{SNR} &\approx \sqrt{NSI}/\sqrt{S-1} \\ &\approx \sqrt{NI}, \text{ when } S \gg 1 \end{aligned} \tag{2.11}$$

where N is the number of supergathers and I is the number of iterations. The total number of shots is $N \times S$. When S is equal to 1 for the conventional sources situation, there will be no crosstalk noise. Since I assume there is no noise in the original shot gathers, the SNR of the migration image is infinity, and when S is much greater than 1, the SNR is independent of S. Equations 2.10 and 2.11 will be validated with numerical examples for $S \gg 1$. In the case of iterative least-squares migration, the crosstalk noise in the gradient or conjugate direction from each iteration is correlated with each other for static encoding; moreover, after being scaled with the step length, the variance of the crosstalk noise would be different for every iteration, where early iterations receive large weight. Therefore, I expect the SNR enhancement to be less than the prediction from equation 2.11, where crosstalk noise is assumed to be of comparable energy.

When S is small, e.g. $S = 2$, the SNR of a conventional Kirchhoff migration image is often large enough, because N is large in this case. Several studies (Beasley, 2008; Hampson et al., 2008; Berkhout, 2008; Fromyr et al., 2008) have shown that conventional stacking and migration of simultaneously acquired supergathers can effectively suppress the interference of reflections from different sources, i.e., crosstalk. However, if S is large, the crosstalk noise is intolerable due to the decrease of the number of supergathers (N). In the next section, multisource least-squares migration is applied to synthetic multishot supergathers to suppress the crosstalk and improve the SNR.

2.3 Numerical Results

The multisource least-squares migration algorithm wth Kirchhoff modeling and adjoint operators is tested on synthetic data generated by a Born modeling method for the 2D SEG/EAGE salt model. Figure 2.2 shows the reflectivity model calculated from the velocity model using vertical rays and constant density assumptions; the true velocity model is used for migration. The ocean bottom reflector is muted in order to better illustrate the deep structure, and 320 sources and 320 receivers are deployed on the surface with the same sampling interval of 18.3 m. The modeling parameters are listed in Table 2.1 (see Appendix A for the meaning of deblurring filter parameters), where the deblurring filter is only applied at the first two iterations[4] to provide a good initial model for the inversion. The regularization parameter is chosen based on a trial and error method and is reduced by half after each iteration. Regularization is important for attenuating crosstalk noise and high-frequency noise associated with the deblurring filter.

2.3.1 Conventional Sources Least-squares Migration and Deblurring

Figure 2.3a shows the 2D prestack Kirchhoff migration image (color scale boosted to show deep structures) for a conventional acquisition geometry of 320 individual shots with 320 receivers per shot. To reduce the artifacts, a non-stationary preconditioner (also denoted as a deblurring filter in Aoki and Schuster (2009)) is applied to the Kirchhoff migration image to give the result shown in Figure 2.3b. It is referred to as the deblurred image.

Comparison of the deblurred image and non-deblurred images shows that the deblurred image has a more balanced reflectivity amplitude, which means that amplitude weakening

[4]The standard preconditioner (Nemeth et al., 1999; Plessix and Mulder, 2004) of inverse geometric spreading is used at every iteration

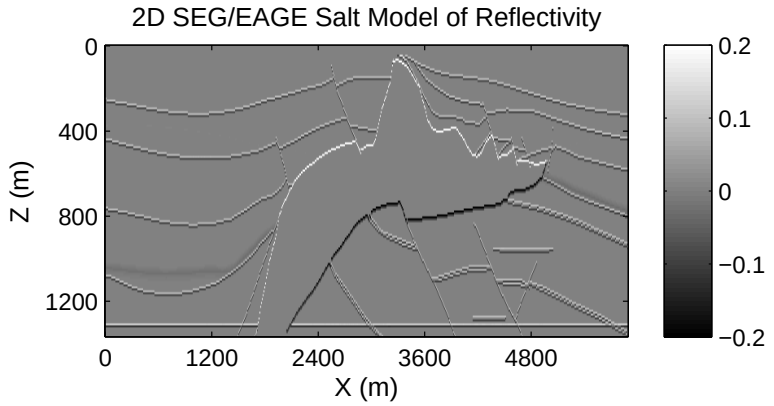

Figure 2.2. 2D SEG/EAGE salt model (reflectivity).

Table 2.1. Modeling parameters for conventional sources simulation.

Model size	645×150	Src Wavelet	Ricker
Grid interval	9.14 m	Peak freq.	50 Hz
Src number	320	Src interval	18.3 m
Rec. number	320	Rec. interval	18.3 m
Trace length	4 sec	Regularization	1.0e-7
Subsec. size	5×5 pts	Filter size	3×3 pts

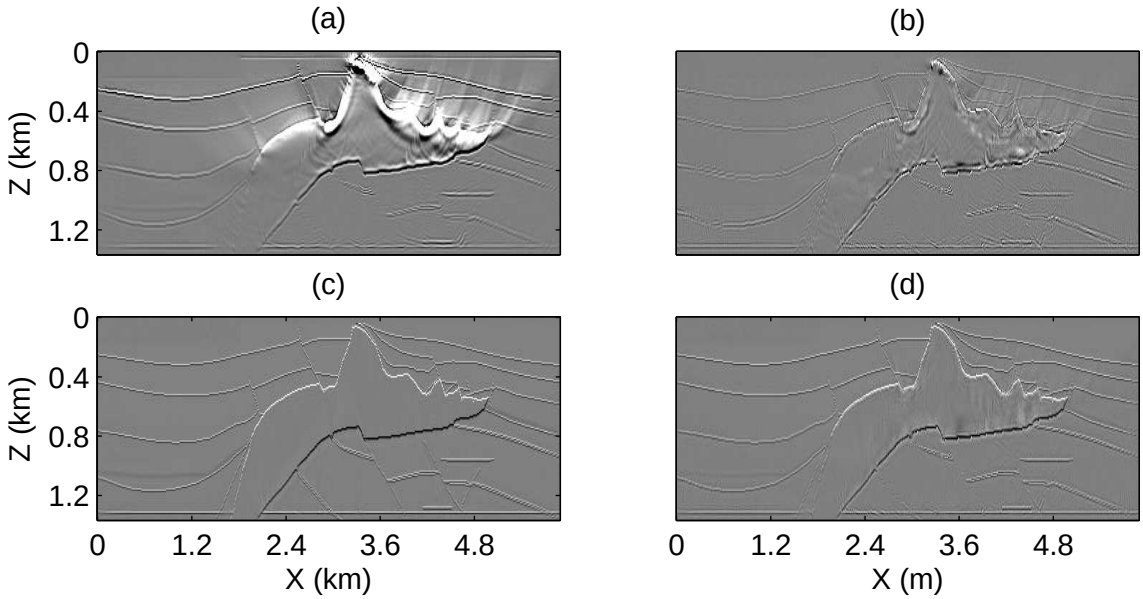

Figure 2.3. Migration images obtained with different methods: (a) Kirchhoff migration image for conventional sources data, (b) KM image after deblurring (deblurred image), (c) Least-squares migration image after 30 CG iterations, (d) Preconditioned least-squares migration image after 30 DCG iterations.

due to geometric spreading is compensated. The migration artifacts are also suppressed in the deblurred image. However, the deblurring filter also introduces some high-frequency noise into the deblurred image because it only approximates the inverse Hessian (see Appendix A for details). In the end, the filter is effective for deblurring Kirchhoff migration images, but it comes with the price of adding high-frequency noise. A more effective deblurring filter (Yu et al., 2006) can be used but comes with added computation cost.

To summarize the overall effect of the deblurring filter, Figure 2.4 depicts the convergence curves for both standard (CG) and deblurred LSM (it is referred to as DCG). Here, the CG result after one iteration is equivalent to the Kirchhoff migration image, and the first iteration result of DCG represents the deblurred image. I can see that the deblurring filter reduces the data residual by 52%, in spite of the high-frequency noise it introduced. It is not used after a few iterations and allows the least-squares migration to reduce the remaining noise. Figure 2.3c shows the conventional sources least-squares migration image after 30 CG iterations[5], which is almost identical to the original model. It demonstrates that least-squares migration can sometimes produce images of higher quality and resolution compared to Kirchhoff migration (Nemeth et al., 1999), if the migration velocity is a somewhat accurate rendering of the actual smoothed velocity.

2.3.2 Multisource Least-squares Migration

To simulate multisource data, conventional sources data are encoded and blended together to form a small number of supergathers. The 320 shot gathers are separated into different clusters of supergathers, where each supergather in a cluster is formed by stacking a unique set of shot gathers together to form the following data sets: thirty-two 10-shot supergathers, sixteen 20-shot supergathers, eight 40-shot supergathers, four 80-shot supergathers, two 160-shot supergathers and one 320-shot supergather. Each shot gather has a random time shift applied to it with a standard deviation equal to about seven times the dominant period of the source wavelet. All the random time shifts are generated by a random number generator that honors a uniform probability distribution. Figure 2.5 shows the Kirchhoff migration images from all the experiments. Consistent with equation 2.11, these results show that decreasing the number of supergather leads to increasing levels of crosstalk.

To further validate equation 2.11, I adopt the iterative stacking approach (multiple migrations of all shots) in Romero et al. (2000), where all the 320 shots are encoded and

[5]DCG produces basically the same result after so many iterations, shown in Figure 2.3d.

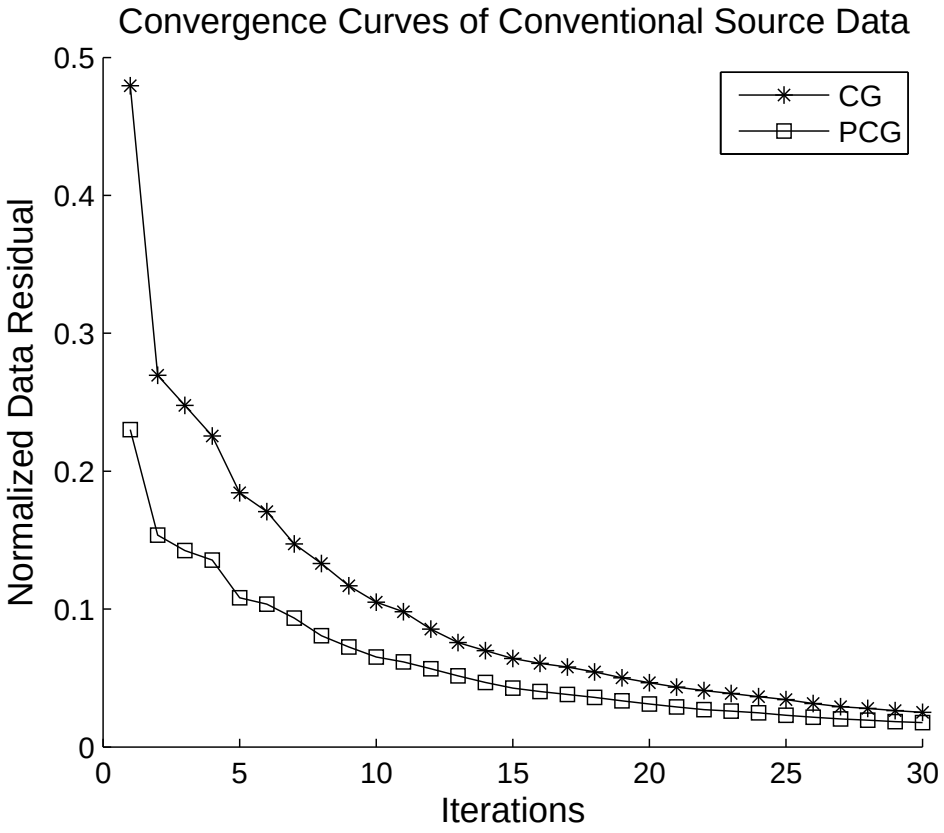

Figure 2.4. Normalized data residual plotted against iteration number. The line with stars indicates the convergence of the conjugate gradient method and the line with squares shows the convergence when the deblurring filter is used as a preconditioner.

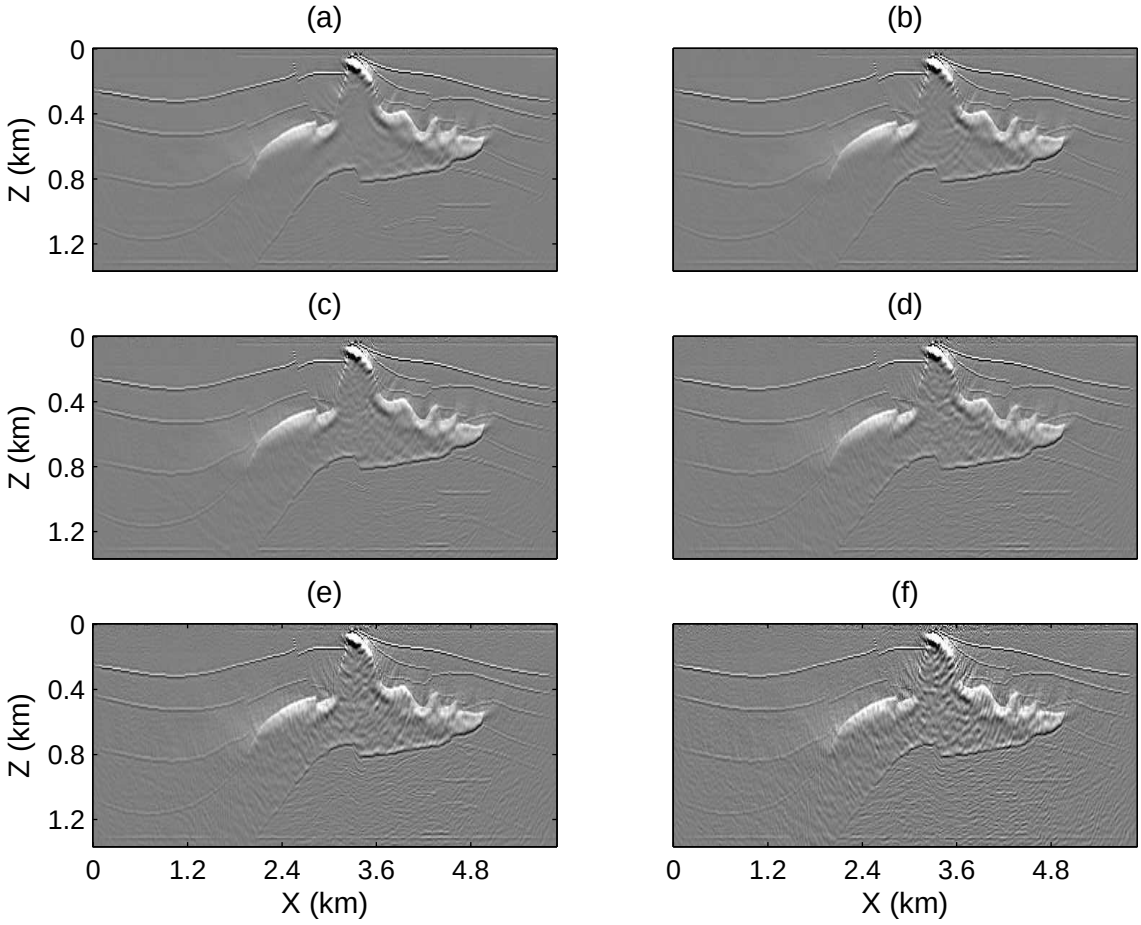

Figure 2.5. Kirchhoff migration images obtained from the following clusters of supergathers, (a) thirty-two 10-shot supergathers, (b) sixteen 20-shot supergathers, (c) eight 40-shots supergathers, (d) four 80-shot supergathers, (e) two 160-shot supergathers and (f) one 320-shot supergather. Here, all shot gathers consisted of 320 traces, and each supergather in a cluster was formed from a unique set of shot gathers.

blended together and migrated with different encoding functions for many iterations. The migration images from different iterations are then stacked together to improve the SNR. Numerically, I use the formula

$$\text{SNR} = \frac{||m_{ref}||}{||m^{(k)} - m_{ref}||}, \tag{2.12}$$

for the SNR calculation, where m_{ref} is the reference migration image for conventional sources (Figure 2.3a) and $m^{(k)}$ is the stacked image after k iterations (k-fold). According to equation 2.11, the SNR is proportional to $\sqrt{I}$, I being the number of iterations. The numerical results in Figure 2.6 largely agree with the prediction, where the measured SNR is normalized by the SNR of the 1st iteration to compare with the $\sqrt{I}$ curve. Figure 2.7 shows (a) the Kirchhoff migration image of a 320-shot supergather with only 1 stack ($I = 1$); (b) the 5-fold stacked image ($I = 5$); (c) the 10-fold stacked image ($I = 10$); and (d) the 20-fold stacked image ($I = 20$). These numerical results suggest that iterative stacking is very effective in suppressing random crosstalk noise. The iterative stacking method is applicable to marine data with either wave-equation or reverse time migration, where the supergathers are not explicitly formed, but instead, the back-propagated wavefields are superimposed together. However, without least-squares migration, these migration artifacts will persist in the images.

For the case where S is much greater than 1, Figure 2.8 presents the migration images where the input data consist of only one supergather ($N = 1$, in equation 2.11) but there are different numbers of shot gathers in the supergather: (a) S=40, (b) S=80 and (c) S=160. These results along with Figure 2.5f demonstrate that the SNRs of these migration images are mostly independent of the number of shot gathers in the supergather. At first glance, this result appears contradictory to intuition because the migration of a 160-shot supergather might be expected to yield a less noisy image than a 80-shot supergather. However, the 160-shot supergather has a higher crosstalk noise level (by a factor of $\sqrt{2}$) than the 80-shot supergather, which cancels the $\sqrt{2}$ SNR enhancement in migrating a 160-shot supergather. The key point here is that increasing the number of unique supergathers is more effective at SNR enhancement than increasing the number of unique shot gathers per supergather.

According to equation 2.11, even a single 320-shot supergather can be used to get an accurate image if the number of iterations is large enough. To verify this prediction, Figure 2.9a-c shows the MLSM images after iteratively migrating a 320-shot supergather; here, the deblurring filter is applied to stabilize and speed up the convergence. It is clear that the image quality increases with the number of iterations. After 60 iterations, the MLSM image is of high quality and mostly free of migration artifacts and crosstalk. It indicates

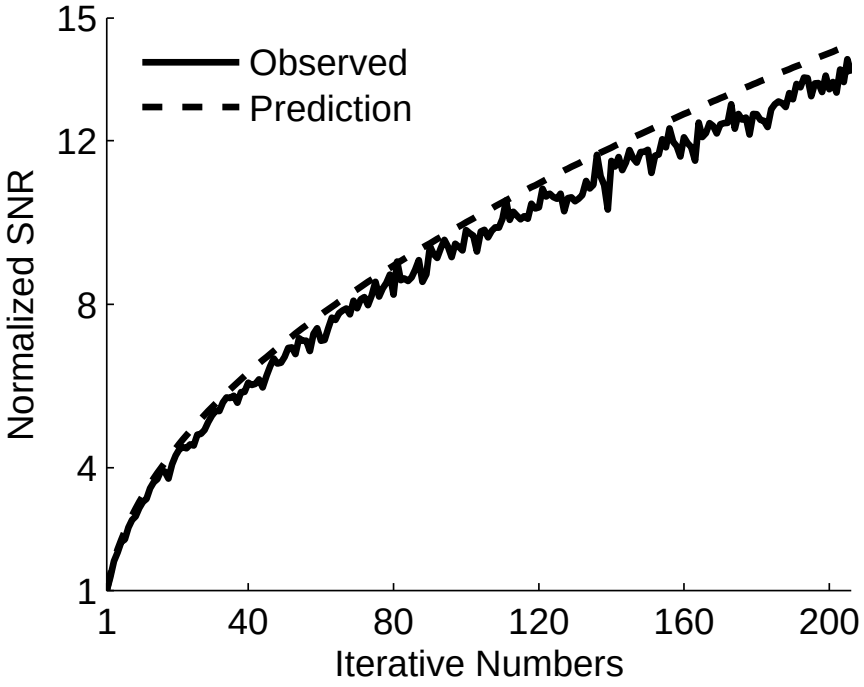

Figure 2.6. The predicted and measured signal-to-noise ratios of iterative stacking method are plotted against iteration number as dashed and solid lines. The measurements have been normalized by the 1st iteration result.

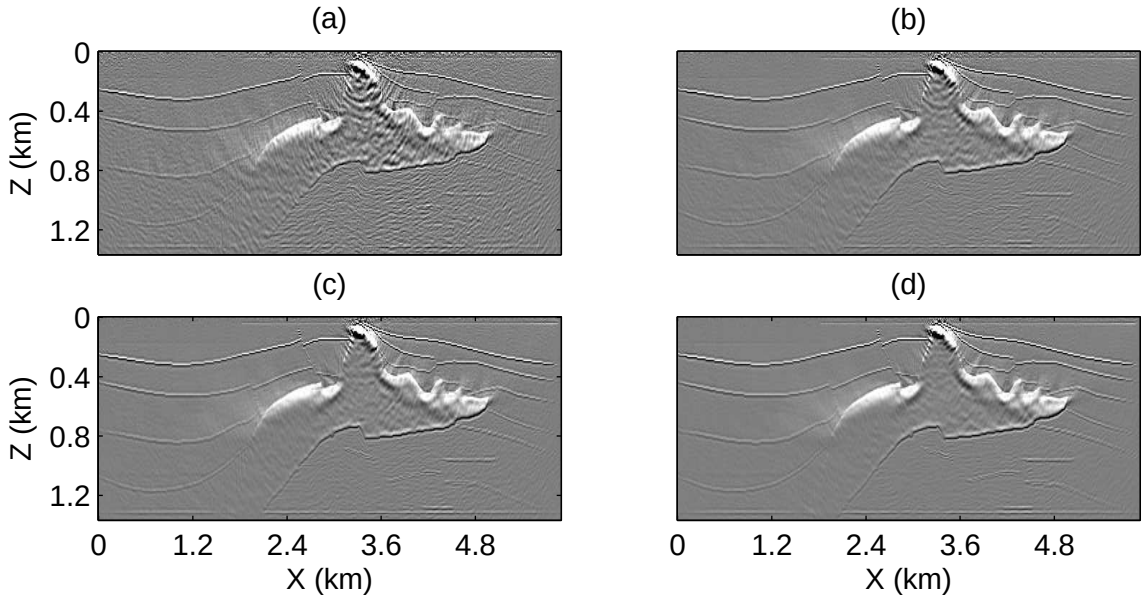

Figure 2.7. Stacked images for iterative stacking after (a) 1 iteration; (b) 5 iterations; (c) 10 iterations; (d) 20 iterations.

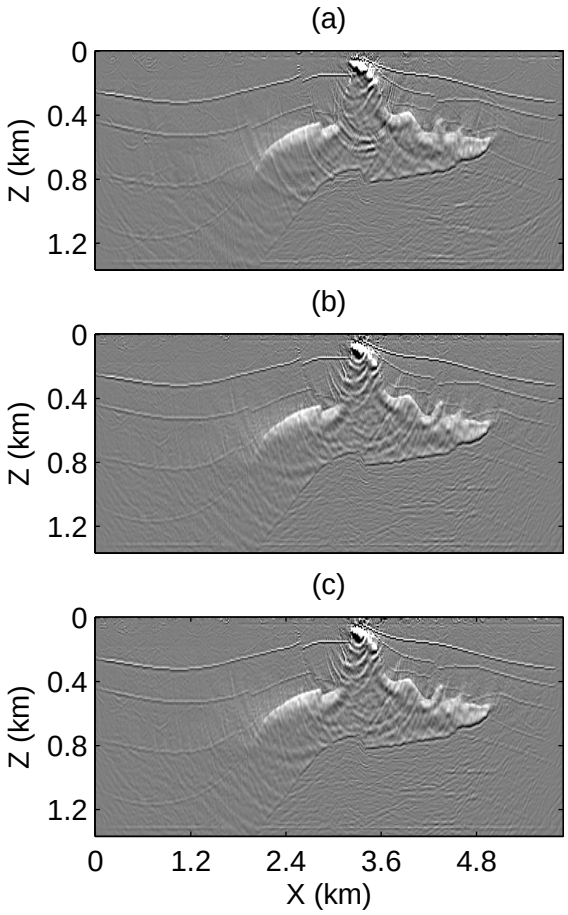

Figure 2.8. Kirchhoff migration images obtained from the following supergathers (a) one 40-shot supergather, (b) one 80-shot supergather and (c) one 160-shot supergather.

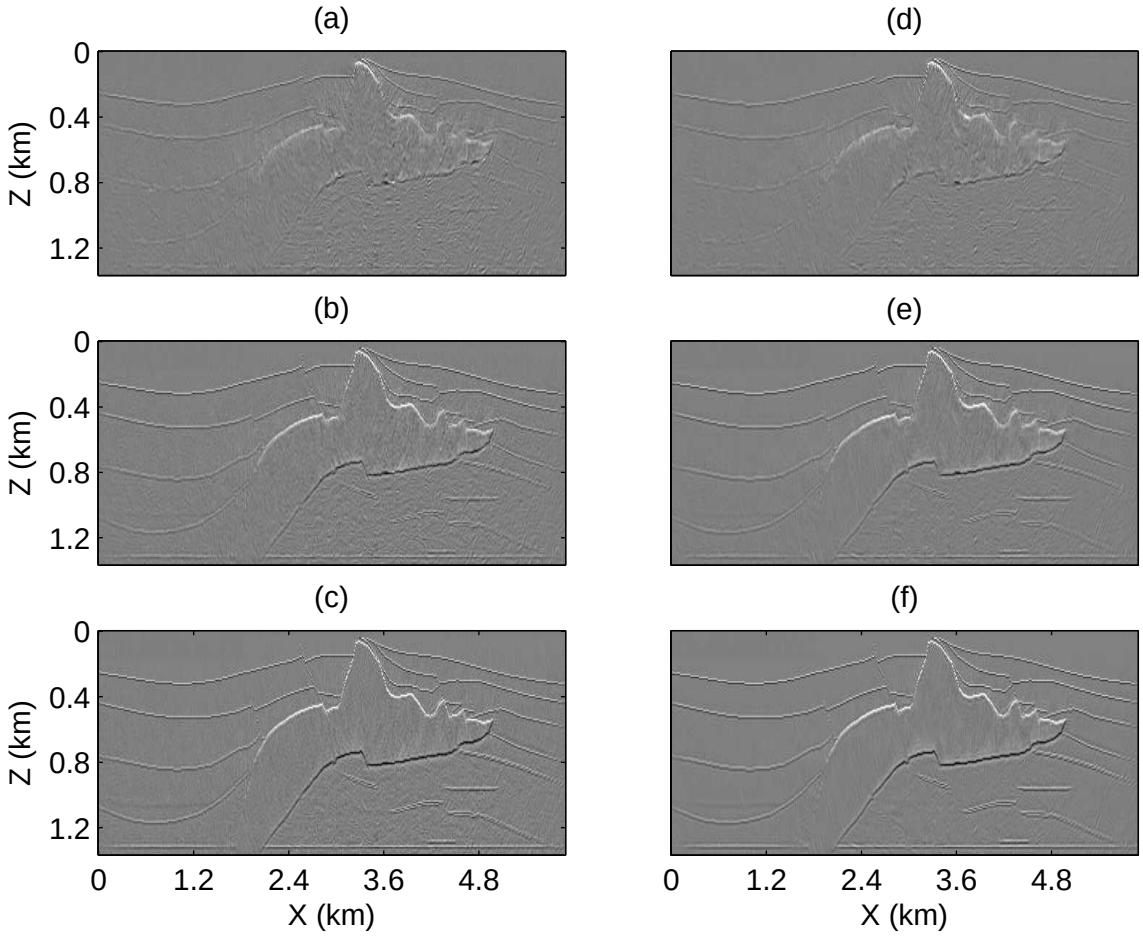

Figure 2.9. Least-squares migration images of a 320-shot supergathers after (a) 10, (b) 30, (c) 60 iterations with static encoding or (d) 10, (e) 30, (f) 60 iterations with dynamic encoding.

that MLSM can accurately estimate the model even when hundreds of shots are blended together in the processing step, and it does not require too many iterations. Moreover, since the SNR is proportional to $\sqrt{N}$, N being the number of supergathers, the crosstalk noise will be more effectively suppressed when there is more than one supergather.

2.3.3 Dynamic Encoding vs Static Encoding

Following Krebs et al. (2009) and Boonyasiriwat and Schuster (2010), a different time-shift encoding of the shot gathers at each iteration can be used for MLSM; I call this dynamic encoding compared to static encoding where a shot gather has the same time shift for any iteration. To compare the effectiveness of the dynamic encoding method relative to static encoding, the MLSM of one 320-shot gather is computed with dynamic encoding. Figure 2.9e-f shows the migration images after 10, 30, and 60 iterations. Compared to Figure 2.9a-c, the MLSM result is improved, which indicates that dynamic encoding is more effective than static encoding in reducing crosstalk.

To quantitatively show the image quality improvement due to dynamic encoding, the SNR is calculated for the MLSM images and compared to the SNR of the statically encoded images in Figure 2.10. For each iteration, the corresponding conventional sources least-squares migration image is used as the reference signal. Here, I assume that the convergence rate is the same for conventional sources and multisource least-squares migration. Results clearly show that the dynamic encoding helps suppress the crosstalk and produce images with higher SNR compared to static encoding. With dynamic encoding, the assumption that the crosstalk noise at every iteration is uncorrelated with the crosstalk at previous iterations is closer to the ideal case compared to static encoding. The drawback is that now I supergathers with different encoding functions are required at input, so that the I/O cost will increase and approach that of conventional migration for a large number of iterations (I).

However, the numerical results show that MLSM algorithm is less efficient in reducing crosstalk than the iterative stacking method as shown in Figure 2.10. The SNR of the 60-iteration MLSM image with dynamic encoding (Figure 2.9f) is comparable to the SNR of the 20-fold stacked image (Figure 2.7d: Note that the migration artifacts in this image are considered as signal in the SNR calculation). One possible explanation is that during the iterations of MLSM the gradients or conjugate directions are computed from different residual data and scaled by different step lengths to make different contributions to the MLSM image and cause the SNR enhancement of MLSM to be suboptimal. Therefore, in real applications, many supergathers (N) should be used. According to equation 2.11, more

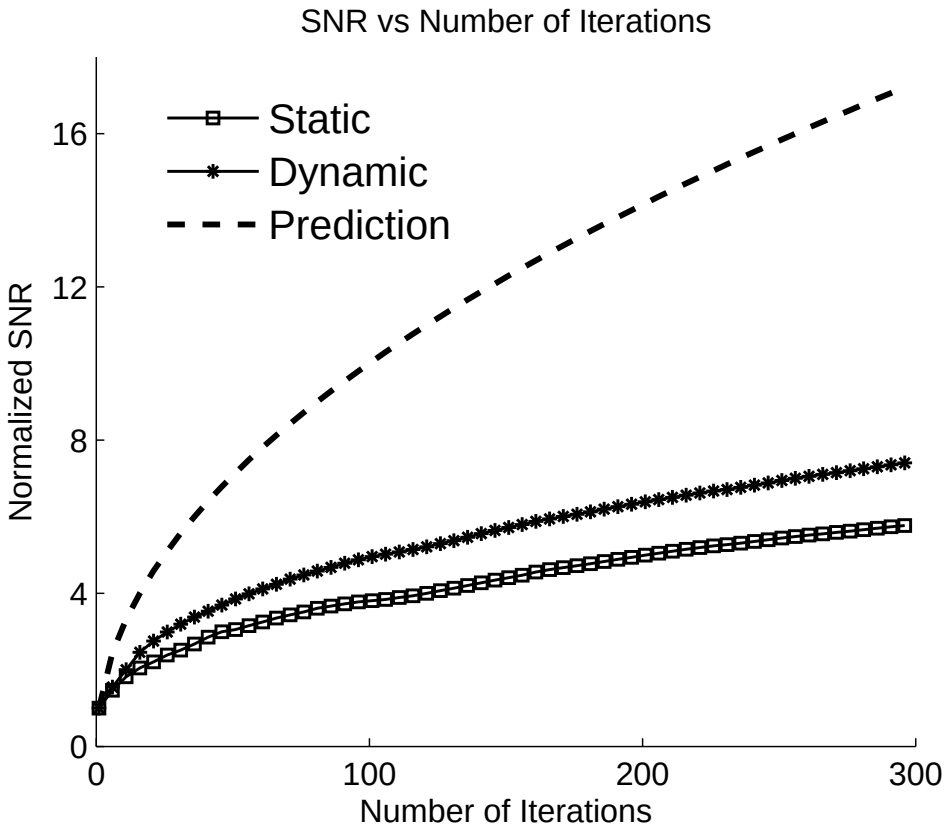

Figure 2.10. The solid line with squares shows the measured SNR for images of one 320-shot supergather with static encoding; the solid line with stars shows the results with dynamic encoding. Here the measured SNR is normalized by the first iteration result. The dashed line indicates the prediction from equation 2.11.

supergathers will greatly improve the SNR of final images, which is evident in examining the change from (f) to (a) in Figure 2.5.

When the processing technique for blending sources is used in full waveform inversion, the SNR of the inverted result is expected to behave in a manner similar to that of MLSM, but analysis is difficult because full waveform inversion is a highly non-linear process.

2.3.4 Computational Cost

Each iteration of iterative LSM costs about two migrations, so the cost of iterative LSM is about $2I$ times that of standard migration. Assuming an ideal land acquisition geometry where the geophones are fixed and S shot gathers are recorded, the total computational cost in computing the migration image is $Cost^{conv} \approx S\alpha$ for conventional prestack migration, where α is the cost of one wave-equation migration. In comparison, if N supergathers are migrated, then the cost[5] per iteration of LSM is only $2N\alpha$. If I iterations are needed then the total cost of LSM is $Cost^{multi} \approx 2N\alpha I$. Therefore I conclude that the cost of MLSM can be less than standard migration if

$$2NI < S. \tag{2.13}$$

In the empirical results, a high quality image is obtained after 60 iterations for a 320-shot supergather, which translates to about 2.7 times speedup if the numerical tests are performed with wave-equation migration or reverse time migration. Meanwhile, the image is free of migration artifacts and with balanced amplitudes (Figure 2.9c).

Another important saving is the reduction of I/O cost. For Kirchhoff migration, the I/O cost can be the dominant factor for the run time. By statically encoding S shots into a supergather, the I/O cost is reduced to $1/S$ of the original cost, which allows significant saving in run time of MLSM. For dynamic encoding, if I iterations are needed, I supergathers with different encoding functions are required at input, so that the I/O cost is reduced to I/S of the original cost. Therefore, MLSM with dynamic encoding does not enjoy a large I/O cost reduction if the number of iterations is large. An optional strategy is to periodically stop the iterations in static iterative LSM and restart them at the stopping model but with a new encoding function in the supergathers. In the above calculation, the cost of computation and I/O of preprocessing step is not considered.

[5]This assumes a wave-equation migration method.

2.4 Conclusions

A multisource least-squares migration algorithm is proposed to efficiently produce high quality images. This algorithm is implemented with Kirchhoff migration method and tested with 320 synthetic shot gathers for the 2D SEG/EAGE salt model. An accurate image is obtained by migrating a supergather composite of all these 320 shot gathers after 60 iterations. Compared to conventional Kirchhoff migration image, the I/O cost of MLSM with static encoding is reduced by 320 times. The MLSM image is much more resolved than conventional Kirchhoff migration image, because the migration artifacts are suppressed, the reflector amplitudes are balanced, the image resolution is enhanced and the crosstalk noise is reduced. According to the signal-to-noise ratio analysis, an acceptable number of iterations are needed to achieve high enough SNR. This suggests that high quality images can be produced with less cost than conventional migration method, if the MLSM algorithm is implemented with the wave-equation migration method.

Two encoding strategies are discussed in this chapter. The MLSM algorithm with static encoding enjoys lower I/O cost compared to the MLSM with dynamic encoding, but the empirical results show that the MLSM with dynamic encoding, on the other hand, is more effective in reducing crosstalk noise introduced by blended sources. Compared to the iterative stacking method, the MLSM algorithm improves the image quality by suppressing the migration artifacts, balancing the reflector amplitudes and enhancing the image resolution, although the MLSM algorithm requires more iterations to reduce crosstalk than the iterative stacking method. For example, the measured SNR of the 60-iteration MLSM image with dynamic encoding is comparable with the SNR of the 20-fold stacked image.

Future research is needed to address following questions. Firstly, the MLSM has only been tested with fixed-spread acquisition geometry. The extension to marine acquisition will be significant. Secondly, the least-square migration seeks a model that optimally fits the data. This process is sensitive to the velocity model, and it is important to reduce this sensitivity for real applications. A third interesting research topic is to look for model dependent efficient encoding functions.

CHAPTER 3

PLANE-WAVE LEAST-SQUARES REVERSE TIME MIGRATION

A plane-wave least-squares reverse time migration (LSRTM) is formulated with a new parameterization, where the migration image of each shot gather is updated separately and an ensemble of prestack images is produced along with common image gathers. The merits of plane-wave LSRTM are the following: (1) plane-wave prestack LSRTM can sometimes offer stable convergence even when the migration velocity has bulk errors of up to 5%; (2) to significantly reduce computation cost, linear phase-shift encoding is applied to hundreds of shot gathers to produce dozens of plane waves. Unlike phase-shift encoding with random time shifts applied to each shot gather, plane-wave encoding can be effectively applied to data with a marine streamer geometry; (3) plane-wave prestack LSRTM can provide higher quality images than standard RTM. Numerical tests on the Marmousi2 model and a marine field dataset are performed to illustrate the benefits of plane-wave least-squares reverse time migration. Empirical results show that LSRTM in the plane-wave domain, compared to standard reverse time migration, produces images efficiently with fewer artifacts and better spatial resolution. Moreover, the prestack image ensemble accommodates more unknowns to makes it more robust than conventional least-squares migration in the presence of migration velocity errors.

3.1 Introduction

The least-squares migration method (Lailly, 1984; Cole and Karrenbach, 1992; Schuster, 1993; Nemeth et al., 1999; Duquet et al., 2000) has been shown to sometimes produce migration images with better quality than those computed by conventional migration. Its original implementation was with Kirchhoff migration (Nemeth et al., 1999; Duquet et al., 2000), but was later developed for phase shift migration algorithms (Kaplan et al., 2010; Huang and Schuster, 2012). When least-squares migration is implemented with the reverse time migration method (Tang and Biondi, 2009; Dai and Schuster, 2010; Dai et al., 2010;

Wong et al., 2011; Dai et al., 2012), it can reduce not only the acquisition footprint but also the artifacts in the RTM image, while enhancing the image resolution. In addition, Romero et al. (2000); Krebs et al. (2009); Tang and Biondi (2009); Schuster et al. (2011); Dai et al. (2011, 2012) employed a phase-encoding multisource approach to increase the computational efficiency by more than an order-of-magnitude compared to conventional LSRTM.

For iterative phase-encoded multisource migration, many shot gathers are encoded with random encoding functions and blended together to form a supergather. One supergather can be modeled and migrated with one finite-difference solution to the wave equation for multiple sources and so provide a high computational efficiency compared to standard LSM. With increasing iteration number, the crosstalk between different shots will be increasingly suppressed. Consequently, the computational cost of LSRTM is reduced to a level comparable to conventional reverse time migration or even lower, depending on the acquisition geometry.

There are two significant problems with LSRTM. The problems and my proposed solutions are now presented.

1. The standard multisource LSRTM combined with the random encoding method is that it requires all the encoded shot gathers to share the same receivers (fixed spread geometry). Therefore, it is not applicable to marine streamer data which are recorded by a towed receiver array (Routh et al., 2011; Huang and Schuster, 2012). To remedy this problem, I devise a plane-wave LSRTM method that can be applied to both land and marine datasets.[1] The encoded source represents a physically realizable planar or line source on the surface, given that the sampling of the shot location is dense, regular, and continuous (Liu et al., 2006). Hence, the blending process with linear phase encoding is identical to a tau-p transformation that is used to transform shot-domain data to plane waves for plane-wave migration (Zhang et al., 2005). Liu et al. (2006) described the relationship between linear time-shift encoding and a plane-wave transformation. They also reported the existence of crosstalk when the sampling of shots is too coarse, and proposed to stack over many different encoding functions (different surface shooting angles) to reduce the crosstalk, which is similar to the crosstalk reduction procedure for random phase encoding. In this way, Vigh

[1] An alternative remedy is to use frequency selection encoding, as proposed by Huang and Schuster (2012).

and Starr (2008) implemented full waveform inversion in the plane-wave domain and achieved significant computational savings.

2. Another drawback of multisource least-squares reverse time migration algorithm is that its convergence is sensitive to the accuracy of the velocity model. When the velocity model contains large bulk errors, the migration images from different shots are inconsistent with each other, so the stacking process become less effective in reducing crosstalk noise and the resolution of the final image is spoiled. In addition, when many shots are blended together, it is difficult to separate them to produce common image gathers as quality control tools. This problem is now remedied by incorporating a regularization term into the LSRTM method that penalizes misfits between the images in the plane-wave domain. In this way the defocusing due to velocity errors is reduced. The formulation is similar to differential semblance optimization (Symes and Carazzone, 1991) which inverted for the velocity model, but in this chapter only the reflectivity image is produced. In contrast to a stacked image, the prestack image ensemble accommodates more unknowns to allow for better fitting of the observed data, and so the convergence of least-squares migration is improved (see Appendix C).

In summary, I present a plane-wave prestack least-squares migration method where the migration image of each shot is updated separately and an ensemble of prestack images is produced with common image gathers. The advantage over conventional LSRTM where all the shot gathers are explained by a single migration image is that it is relatively less sensitive to bulk errors in the migration velocity. The plane-wave encoding technique can significantly reduce the computational and input/output (I/O) cost. In contrast to conventional multisource least-squares migration with phase-encoded supergathers, it can be applied to marine data.

This chapter is organized into four sections. The first one is this introduction, which is followed by the theory of LSRTM. The synthetic and field data examples are then presented in the numerical results section, and this is followed by the summary.

3.2 Theory

The theory of least-squares reverse time migration is well established (Symes and Carazzone, 1991; Mulder and Plessix, 2004; Dai et al., 2012). In this section, I will first review the theory of LSRTM assuming the constant density acoustic wave equation,

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \bigtriangledown^2 p(\mathbf{x},t;\mathbf{x}_s) = s(t;\mathbf{x}_s), \tag{3.1}$$

where $c(\mathbf{x})$ is the velocity distribution, and $p(\mathbf{x},t;\mathbf{x}_s)$ is the pressure field associated with the source term $s(t;\mathbf{x}_s)$. A perturbation in the velocity model $c(\mathbf{x}) \rightarrow c(\mathbf{x}) + \delta c(\mathbf{x})$ will generate a wavefield $p(\mathbf{x},t;\mathbf{x}_s) \rightarrow p(\mathbf{x},t;\mathbf{x}_s) + \delta p(\mathbf{x},t;\mathbf{x}_s)$, which obeys the equation

$$\frac{1}{(c(\mathbf{x}) + \delta c(\mathbf{x}))^2}\frac{\partial^2 (p(\mathbf{x},t;\mathbf{x}_s) + \delta p(\mathbf{x},t;\mathbf{x}_s))}{\partial t^2} - \bigtriangledown^2 (p(\mathbf{x},t;\mathbf{x}_s) + \delta p(\mathbf{x},t;\mathbf{x}_s)) = s(t;\mathbf{x}_s). \tag{3.2}$$

Expanding the velocity term according to

$$\frac{1}{(c(\mathbf{x}) + \delta c(\mathbf{x}))^2} \approx \frac{1}{c(\mathbf{x})^2} - \frac{2\delta c(\mathbf{x})}{c(\mathbf{x})^3}, \tag{3.3}$$

and subtracting equation 3.1 from equation 3.2 yields the wave equation for the wavefield perturbation $\delta p(\mathbf{x},t;\mathbf{x}_s)$

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2 \delta p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \bigtriangledown^2 \delta p(\mathbf{x},t;\mathbf{x}_s) = \frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\frac{2\delta c(\mathbf{x})}{c(\mathbf{x})^3} + O(\delta c(\mathbf{x})^2). \tag{3.4}$$

Neglecting the higher order terms and defining the reflectivity model as $m(\mathbf{x}) = \frac{2\delta c(\mathbf{x})}{c(\mathbf{x})}$, the above equation becomes

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2 \delta p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \bigtriangledown^2 \delta p(\mathbf{x},t;\mathbf{x}_s) = m(\mathbf{x}) \bigtriangledown^2 p(\mathbf{x},t;\mathbf{x}_s). \tag{3.5}$$

Equations 3.1 and 3.5 will be used to derive the Born modeling operator. Numerically, the calculation of the reflection data $\delta p(\mathbf{x},t;\mathbf{x}_s)$ requires two finite-difference simulations: one to solve equation 3.1 to obtain the wavefield $p(\mathbf{x},t;\mathbf{x}_s)$, and one to solve equation 3.5 for the reflection data $\delta p(\mathbf{x},t;\mathbf{x}_s)$. The wavefield $\delta p(\mathbf{x},t;\mathbf{x}_s)$ will be recorded at the receiver position $\mathbf{x}_g$ to give the shot gather $d(\mathbf{x}_g,t;\mathbf{x}_s)$. By the adjoint state method (Plessix, 2006), the migration operation of a shot gather $d(\mathbf{x}_g,t;\mathbf{x}_s)$ requires two finite-difference simulations, one for the source-side wavefield and one for the receiver-side wavefield:

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \bigtriangledown^2 p(\mathbf{x},t;\mathbf{x}_s) = s(t;\mathbf{x}_s), \tag{3.6}$$

$$\frac{1}{c(\mathbf{x})^2}\frac{\partial^2 q(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \bigtriangledown^2 q(\mathbf{x},t;\mathbf{x}_s) = d(\mathbf{x}_g,t;\mathbf{x}_s), \tag{3.7}$$

where $q(\mathbf{x},t;\mathbf{x}_s)$ is the receiver-side wavefield. Note that the source-side wavefield $p(\mathbf{x},t;\mathbf{x}_s)$ propagates forward in time but the receiver-side wavefield $q(\mathbf{x},t;\mathbf{x}_s)$ propagates backward

in time. The migration image associated with the shot at $\mathbf{x}_s$ is produced by applying the imaging condition

$$m(\mathbf{x};\mathbf{x}_s) = \sum_t \bigtriangledown^2 p(\mathbf{x},t;\mathbf{x}_s)\cdot q(\mathbf{x},t;\mathbf{x}_s). \tag{3.8}$$

To simplify the formulas, matrix-vector notation will be used to represent the Born modeling operator

$$\mathbf{d}_i = \mathbf{L}_i\mathbf{m}, \tag{3.9}$$

where $\mathbf{d}_i$ is the reflection data vector for the *ith* shot, $\mathbf{m}$ is a reflectivity model, and $\mathbf{L}_i$ represents the Born modeling operator associated with the *ith* shot. Similarly, the reverse time migration operator can be expressed as

$$\mathbf{m}_{mig,i} = \mathbf{L}_i^T\mathbf{d}_i. \tag{3.10}$$

with $\mathbf{m}_{mig,i}$ indicating the migration image for the *ith* shot and $\mathbf{L}_i^T$ representing the migration operator associated with the *ith* shot.

3.2.1 Least-squares Migration

For conventional least-squares migration (Nemeth et al., 1999), a reflectivity model $\mathbf{m}$ is assumed to be independent of the shot position. For a dataset with N_s shots, the modeling process can be expressed as

$$\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_2\\ \cdot\\ \mathbf{d}_{N_s}\end{bmatrix} = \begin{bmatrix}\mathbf{L}_1\\ \mathbf{L}_2\\ \cdot\\ \mathbf{L}_{N_s}\end{bmatrix}[\mathbf{m}], \tag{3.11}$$

and similarly for the migration

$$\mathbf{m}_{mig} = \begin{bmatrix}\mathbf{L}_1^T & \mathbf{L}_2^T & \cdot & \mathbf{L}_{N_s}^T\end{bmatrix}\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_2\\ \cdot\\ \mathbf{d}_{N_s}\end{bmatrix} = \sum_{i=1}^{Ns}\mathbf{L}_i^T\mathbf{d}_i, \tag{3.12}$$

where the final image is the stack of migration images from all of the individual shots.

For conventional least-squares migration, a reflectivity model $\mathbf{m}$ is sought that minimizes the misfit functional

$$f(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{N_s}||\mathbf{L}_i\mathbf{m}-\mathbf{d}_i||^2 + \frac{\gamma}{2}||\mathbf{m}||^2, \tag{3.13}$$

where γ is the damping coefficient, and $\mathbf{m}$ is defined as the stacked migration image. This method will also be referred as least-squares migration with the stacked image. In Dai et al. (2012), the stacked migration image is computed from a blend of phase encoded shot gathers, also known as a supergather. When the migration velocity is not accurate, the

prestack images are not exactly the same from different shots, and the stacked image can become blurred and convergence stalls.

In order to improve the robustness of LSRTM, I define the ensemble of prestack images as a function of the shot position: $\underline{\mathbf{m}} \rightarrow \underline{\mathbf{m}}(\mathbf{x}, \mathbf{x}_s)$, or in matrix-vector notation

$$\underline{\mathbf{m}} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ . \\ \mathbf{m}_{N_s} \end{bmatrix}, \tag{3.14}$$

where $\mathbf{m}_i$ is the migration image for the ith shot. Now for a dataset with N_s shots, the forward modeling operation is

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ . \\ \mathbf{d}_{N_s} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & . & \\ & & & \mathbf{L}_{N_s} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ . \\ \mathbf{m}_{N_s} \end{bmatrix}, \tag{3.15}$$

and the migration operation is

$$\begin{bmatrix} \mathbf{m}_{mig,1} \\ \mathbf{m}_{mig,2} \\ . \\ \mathbf{m}_{mig,N_s} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1^T & & & \\ & \mathbf{L}_2^T & & \\ & & . & \\ & & & \mathbf{L}_{N_s}^T \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ . \\ \mathbf{d}_{N_s} \end{bmatrix}. \tag{3.16}$$

For simplicity, I define $\underline{\mathfrak{L}}$ as the forward modeling operator for all the shot gathers

$$\underline{\mathfrak{L}} = \begin{bmatrix} \mathbf{L}_1 & & & \\ & \mathbf{L}_2 & & \\ & & . & \\ & & & \mathbf{L}_{N_s} \end{bmatrix} \tag{3.17}$$

and equation 3.15 and 3.16 are rewritten in compact form

$$\mathbf{d} = \underline{\mathfrak{L}}\underline{\mathbf{m}} \tag{3.18}$$

and

$$\underline{\mathbf{m}}_{mig} = \underline{\mathfrak{L}}^T \mathbf{d}. \tag{3.19}$$

Therefore, the misfit functional with the ensemble of prestack images is defined as

$$\begin{aligned} f(\underline{\mathbf{m}}) &= \frac{1}{2}\sum_{i=1}^{N_s} ||\mathbf{L}_i\mathbf{m}_i - \mathbf{d}_i||^2 + \frac{\gamma}{2}\sum_{i=1}^{N_s} ||\mathbf{m}_i||^2 \\ &= \frac{1}{2}||\underline{\mathfrak{L}}\underline{\mathbf{m}} - \mathbf{d}||^2 + \frac{\gamma}{2}||\underline{\mathbf{m}}||^2. \end{aligned} \tag{3.20}$$

A preconditioned conjugate gradient algorithm

$$\begin{aligned}
\mathbf{g}^{(k+1)} &= \underline{\mathfrak{L}}^T[\underline{\mathfrak{L}}\underline{\mathfrak{m}}^{(k)} - \mathbf{d}] + \gamma\underline{\mathfrak{m}}^{(k)} \\
\beta &= \frac{\mathbf{g}^{(k+1)}\mathbf{P}\mathbf{g}^{(k+1)}}{\mathbf{g}^{(k)}\mathbf{P}\mathbf{g}^{(k)}} \\
\mathbf{z}^{(k+1)} &= \mathbf{P}\mathbf{g}^{(k+1)} + \beta\mathbf{z}^{(k)} \\
\alpha &= \frac{[\mathbf{z}^{(k+1)}]^T\mathbf{g}^{(k+1)}}{[\mathbf{L}\mathbf{z}^{(k+1)}]^T\mathbf{L}\mathbf{z}^{(k+1)} + \lambda||\mathbf{z}^{(k+1)}||^2} \\
\underline{\mathfrak{m}}^{(k+1)} &= \underline{\mathfrak{m}}^{(k)} - \alpha\mathbf{z}^{(k+1)},
\end{aligned} \tag{3.21}$$

can be implemented to find the solution $\underline{\mathfrak{m}}$ that minimizes the misfit in equation 3.20. In above equation, $\mathbf{P}$ is the matrix representing the illumination compensation preconditioner (Plessix and Mulder, 2004).

The ensemble of prestack images $\underline{\mathfrak{m}}$ contains many more unknowns than the stacked image, and therefore provides more freedom to fit the observed data (see Appendix C). Another advantage is that the common image gathers can be extracted from the prestack images as an indication of the image quality.

3.2.2 Plane-wave Prestack LSRTM

In Dai et al. (2012), the multisource technique is implemented with random time shifts and random source polarity encoding functions to greatly reduce the computational cost. By stacking images from different supergathers and gradients from different iterations, the coherent signal is enhanced while the crosstalk noise is reduced. In this report, the shot-domain data are encoded with linear time-shift encoding functions and transformed into plane waves. Assuming a 2D survey geometry, the encoding process can be expressed as:

$$d(x_g, t; p) = \sum_{x_s} d(x_g, t; x_s) * \delta(t - p \cdot x_s), \tag{3.22}$$

where the shot-domain data $d(x_g, t; x_s)$ are encoded with a time shift function $\delta(t - p \cdot x_s)$ and stacked together. As illustrated by Figure 3.1, the time shift $p \cdot x_s$ is a linear function of source position x_s, and p is the ray parameter defined as

$$p = \frac{sin\theta}{v}, \tag{3.23}$$

where θ is the surface shooting angle and v is the velocity at the surface. For the 3D case, the plane-wave is computed from linear combination of surface sources along the x-direction or both the x- and y-directions to form planar sources (see Zhang et al. (2005) and Duquet and Lailly (2006) for details). Since the plane waves are coherent signals, the migration

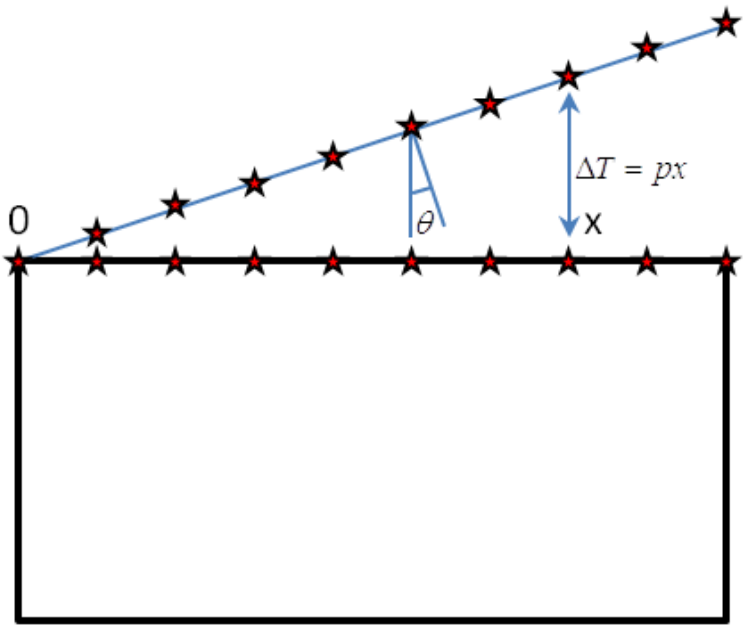

Figure 3.1. The diagram of plane wave encoding (reproduced from Zhang et al. (2005)), where the time shift is linear function to the source location x and the slope is the ray parameter p.

image of one plane wave does not contain crosstalk noise seen in the migration image of a phase-encoded supergather. Instead, it contains aliasing artifacts, which can be reduced by stacking images from many different angles. Zhang et al. (2005) provided an estimate of how many angles are needed as a function of recording aperture, velocity model, and estimated dipping angle range of the reflectors.

In the plane-wave domain, the prestack image ensemble is a function of the ray parameter

$$\underline{\mathfrak{m}} \rightarrow \underline{\mathfrak{m}}(\mathbf{x}, p), \tag{3.24}$$

or in matrix-vector notation

$$\underline{\mathfrak{m}} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ . \\ \mathbf{m}_{N_p} \end{bmatrix}, \tag{3.25}$$

assuming there are N_p plane-wave gathers. In above equation, $\mathbf{m}_i$ is the image associated with the *ith* plane-wave gather. A new misfit functional is defined in the plane-wave domain as

$$\begin{aligned} f(\underline{\mathfrak{m}}) &= \frac{1}{2}\sum_{i=1}^{N_p} ||\mathbf{L}_i\mathbf{m}_i - \mathbf{d}_i||^2 + \frac{\gamma}{2}||\mathbf{C}\underline{\mathfrak{m}}||^2, \\ &= \frac{1}{2}||\underline{\mathfrak{L}}\underline{\mathfrak{m}} - \mathbf{d}||^2 + \frac{\gamma}{2}||\mathbf{C}\underline{\mathfrak{m}}||^2, \end{aligned} \tag{3.26}$$

Note that $\mathbf{d}_i$ represents the *ith* plane-wave gather and $\mathbf{L}_i$ is the forward modeling operator associated with it. In this chapter, I choose a regularization term that penalizes the difference of migration images computed with slightly different incidence angles, and it is defined as

$$||\mathbf{C}\underline{\mathfrak{m}}||^2 = \sum_{i=1}^{N_p-1} ||\mathbf{m}_{i+1} - \mathbf{m}_i||^2 \tag{3.27}$$

γ is the damping coefficient and is chosen empirically. Then a preconditioned conjugate gradient scheme similar to equation 3.21 can be implemented:

$$\begin{aligned} \mathbf{g}^{(k+1)} &= \underline{\mathfrak{L}}^T[\underline{\mathfrak{L}}\underline{\mathfrak{m}}^{(k)} - \mathbf{d}] + \gamma\mathbf{C}^T\mathbf{C}\underline{\mathfrak{m}}^{(k)} \\ \beta &= \frac{\mathbf{g}^{(k+1)}\mathbf{P}\mathbf{g}^{(k+1)}}{\mathbf{g}^{(k)}\mathbf{P}\mathbf{g}^{(k)}} \\ \mathbf{z}^{(k+1)} &= \mathbf{P}\mathbf{g}^{(k+1)} + \beta\mathbf{z}^{(k)} \\ \alpha &= \frac{[\mathbf{z}^{(k+1)}]^T\mathbf{g}^{(k+1)}}{[\mathbf{L}\mathbf{z}^{(k+1)}]^T\mathbf{L}\mathbf{z}^{(k+1)} + \lambda||\mathbf{C}\mathbf{z}^{(k+1)}||^2} \\ \underline{\mathfrak{m}}^{(k+1)} &= \underline{\mathfrak{m}}^{(k)} - \alpha\mathbf{z}^{(k+1)}, \end{aligned} \tag{3.28}$$

to find the LSRTM prestack images. In the next section, unless otherwise denoted, all LSRTM images are produced with the proposed new method.

3.3 Numerical Results

The plane-wave prestack LSRTM algorithm is first tested with the Marmousi2 dataset with a fixed spread acquisition geometry, and then tested with a 2D field dataset with a marine streamer geometry. The numerical scheme in equation 3.28 is implemented with a 2-8 finite-difference method.

3.3.1 Synthetic Example: Marmousi2 Data with a Fixed Spread Survey

The plane-wave LSRTM is applied to a synthetic dataset for the Marmousi2 model. The original model is modified to be the size of 8 km $\times$ 3.5 km with a 10 m grid interval (Figure 3.2(a)). Synthetic data are generated with a fixed spread geometry where 801 shots are excited with a 10 m offset interval at the depth of 10 m. Each shot is recorded with 801 receivers with a 10 m receiver interval. A Ricker wavelet with a 20-Hz peak frequency is used as the source wavelet, and the record length is 8 sec in time. The true data are generated by a pseudo-spectral modeling method (Kosloff and Baysal, 1982). Conventional shot-domain RTM is first applied to the dataset with a smooth migration velocity (Figure 3.2(b)) to give the image shown in Figure 3.3. In the shallow part, there are strong artifacts associated with the source positions and the ocean bottom even after high-pass filtering. The amplitudes of the deep reflectors are very weak.

3.3.2 Plane-wave Transformation

The 801 shot gathers are encoded with linear time-shift encoding (equation 3.22) to form 31 plane-wave gathers with ray parameters (p) ranging from -333 μs/m to 333 μs/m with an even sampling in p. Figure 3.4 shows a plane-wave gather with p=22.2 $\mu s/m$. It is obvious that the direct wave is well separated from the reflections and can be easily muted.

3.3.3 LSRTM of One Plane-wave Gather

Figure 3.5(a) shows the migration image of one plane-wave gather with a zero-degree shooting angle at the surface. It contains strong aliasing artifacts. In this image, some of the steeply dipping reflectors are not illuminated by the zero-degree plane wave. In addition, the deep reflectors are poorly imaged because the planar wavefront is disrupted by shallow structures before it propagates to the deep part and the reflections from deep reflectors are rendered weak and incoherent. Applying the LSRTM method to this plane-wave gather gives the image in Figure 3.5(b) after 30 iterations. Compared to Figures 3.3 and 3.5(a),

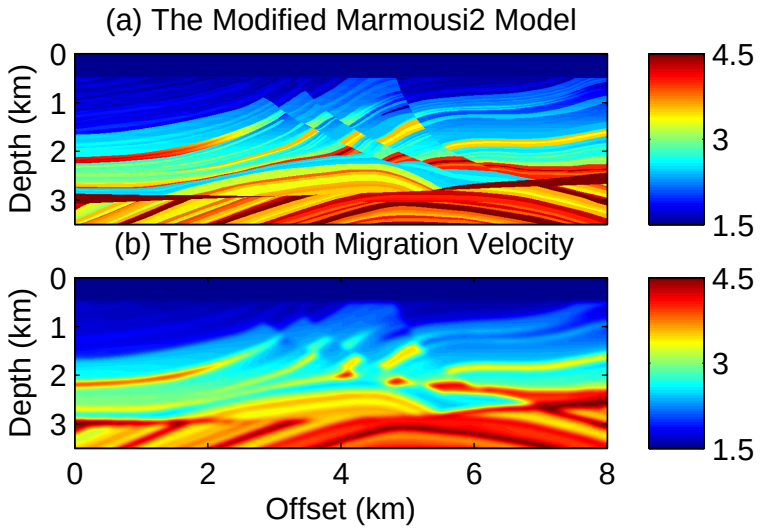

Figure 3.2. The velocity models: (a) modified Marmousi2 model and (b) the smooth migration velocity model. The migration velocity is smoothed by a triangle smoothing filter with a window length of 100 m to get rid of the fine scale structures.

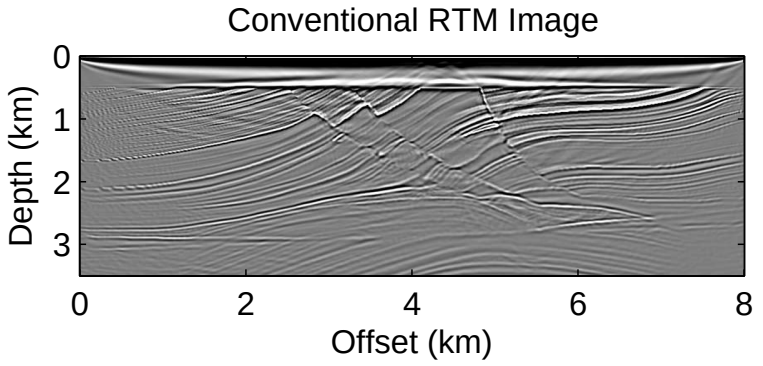

Figure 3.3. The conventional shot-domain RTM image for the Marmousi2 model.

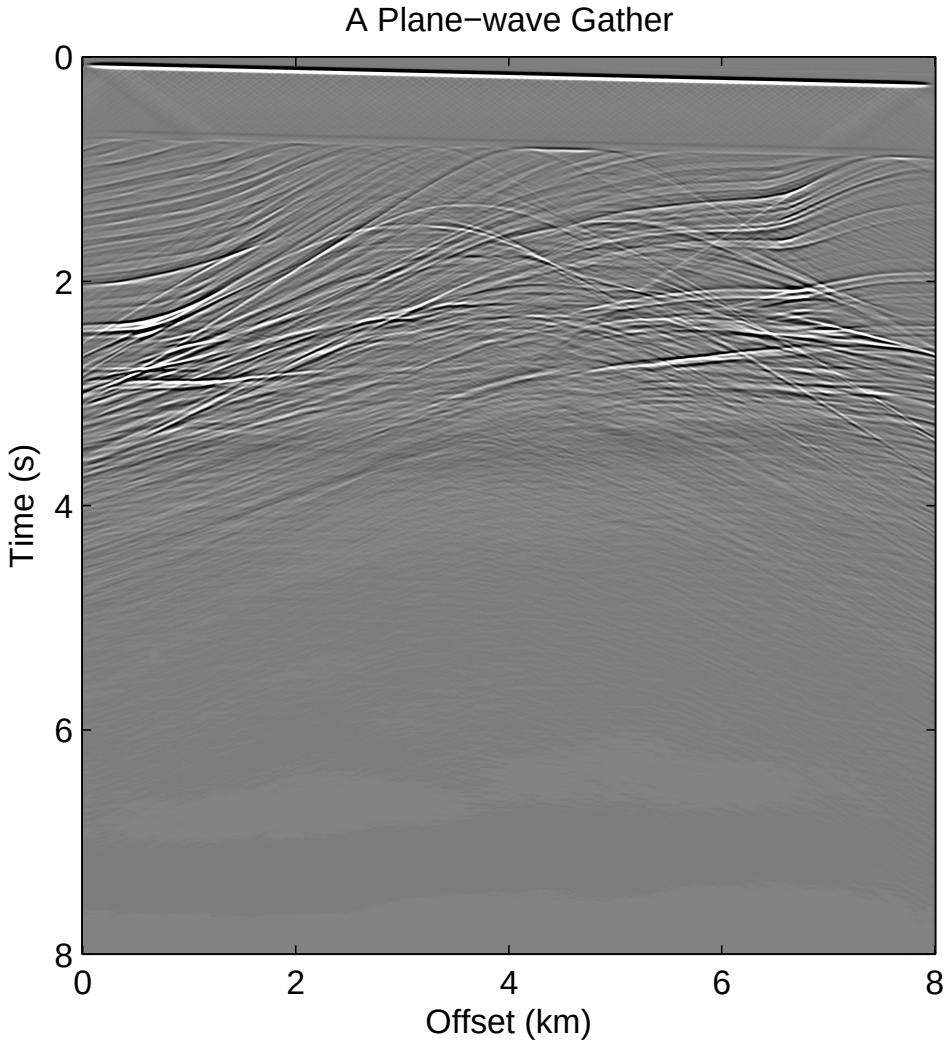

Figure 3.4. A plane-wave gather with p=22.2 μs/m for the Marmousi2 model.

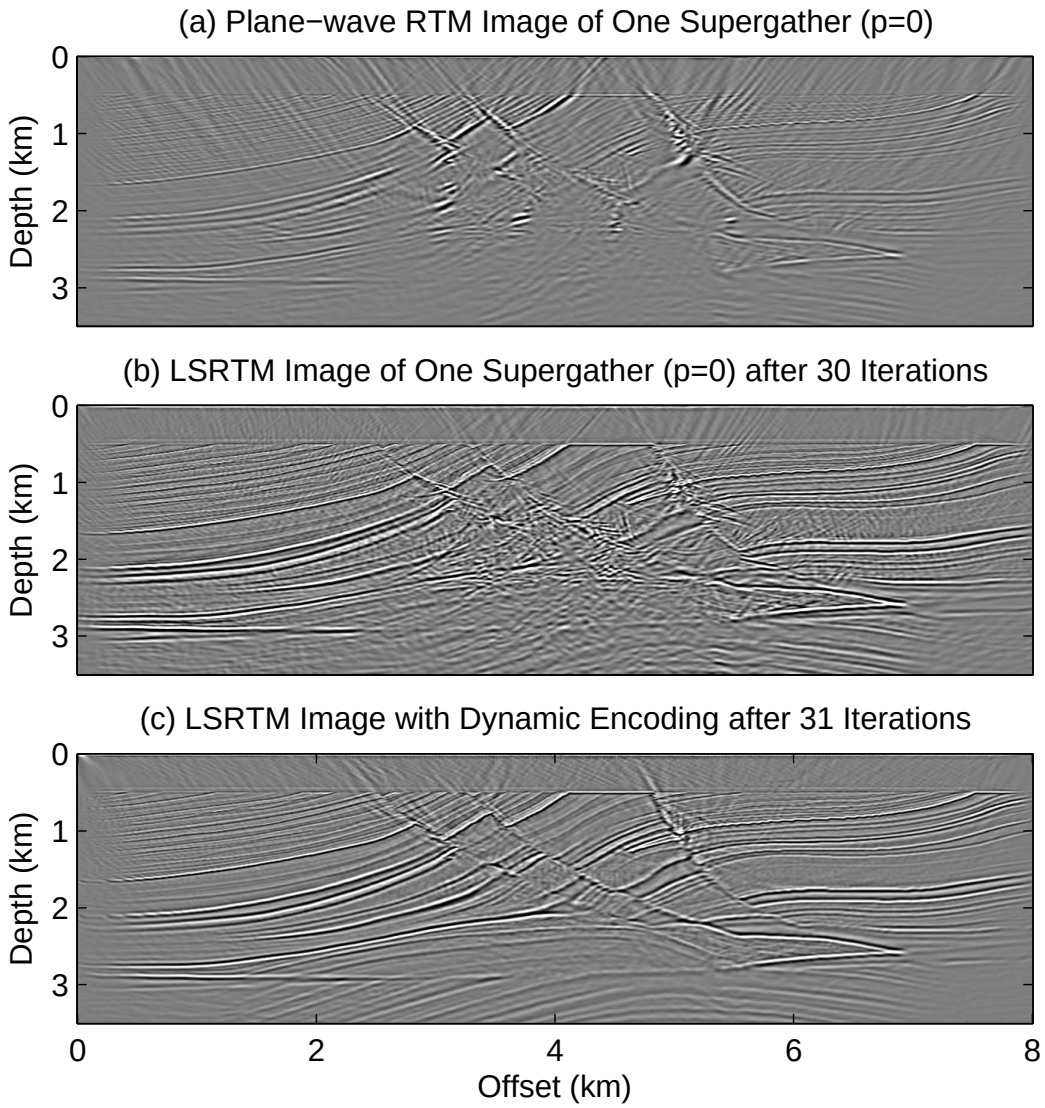

Figure 3.5. The plane-wave RTM image of the Marmousi2 model with only one angle ($p = 0$); (b) The plane-wave LSRTM image of the marmousi2 Model with only one angle ($p = 0$) after 30 iterations; and (c) The plane-wave LSRTM image of the Marmousi2 model with only one angle per iteration. The angle is dynamically changed at every iteration.

the aliasing artifacts are largely suppressed in the LSRTM image and those steeply dipping structures are well illuminated. The LSRTM image also shows enhanced resolution in the shallow part. However, this image contains strong high-frequency noise in the deep portion of the model.

3.3.4 Dynamic Plane-wave LSRTM

To suppress the noise in Figure 3.5(b), the dynamic encoding approach (Krebs et al., 2009; Schuster et al., 2011) is employed. The plane-wave LSRTM is still applied to just one sugergather at each iteration, but the encoding function (in this case, the ray parameter) is changed at every iteration. This time, 31 iterations are used to fully cover the angle range, and the result is shown in Figure 3.5(c). In this example, the ray parameter p ranges from -333 μs/m to 333 μs/m sequentially with a unique p for each iteration. With dynamic encoding, the image has a much higher signal-to-noise ratio (S/N) compared to Figure 3.5(b), and in the deep part, most of the structures are well imaged. The LSRTM with dynamic encoding is very efficient and has the potential to produce high quality images, but there are no image gathers available for quality control or migration velocity analysis.

3.3.5 Plane-wave Prestack LSRTM

When all 31 plane-wave gathers are migrated by conventional RTM and the migration images are stacked together, the final image is of higher quality than Figure 3.5(a) due to the stacking process. Figure 3.6(a) shows the stacked plane-wave RTM image after high-pass filtering. Compared to shot-domain RTM image in Figure 3.3, there are fewer artifacts in the shallow part and the reflector amplitudes are more balanced in the deep part. There are also some migration artifacts that are related to the strong diffractors.

In the next step, the 31 plane-wave gathers are migrated with the numerical scheme in equation 3.28 and Figure 3.6(b) shows the plane-wave LSRTM image after 30 iterations, which is of much higher quality when compared to Figures 3.3 and 3.6(a) in terms of resolution. Also, the migration artifacts are less noticeable in the LSRTM image, and the reflector amplitudes are nearly the same from shallow to deep depths. With the prestack images, the common image gathers (CIGs) can be extracted from the plane-wave RTM and LSRTM images.

Figure 3.7 shows the common image gathers from the RTM image, where it is obvious that the migration velocity is accurate since all the image gathers are flat. Strong low-wavenumber artifacts can be seen, which are expected to be the RTM artifacts (Liu et al., 2011). There is some coherent noise present in the gathers, which can be related to internal

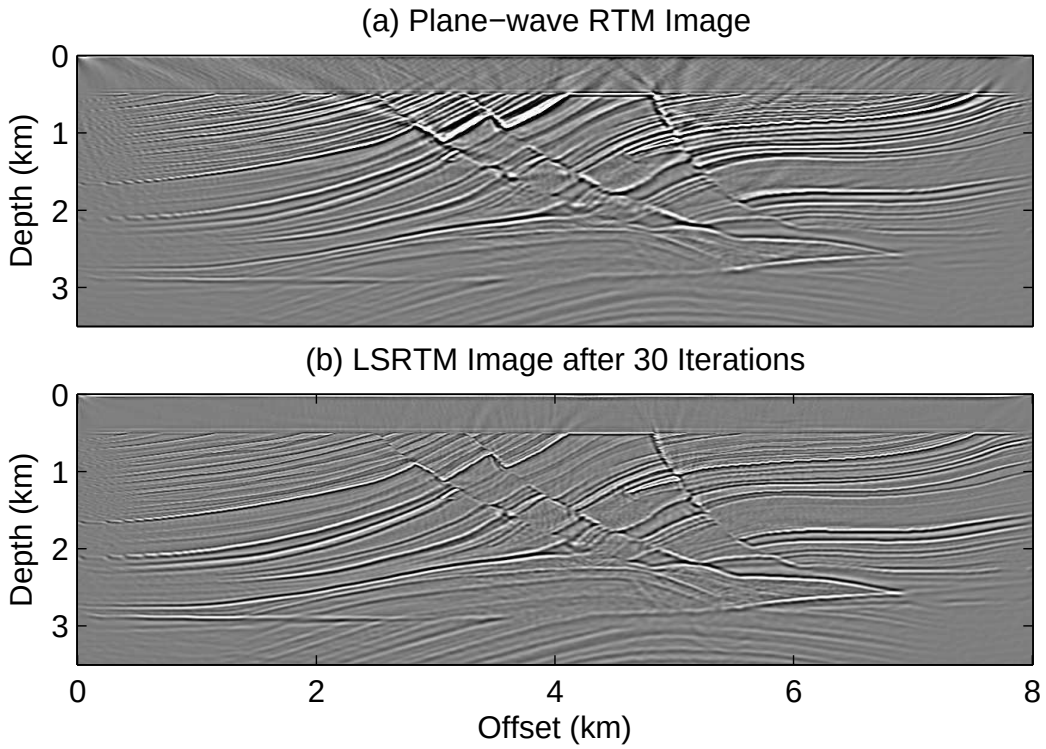

Figure 3.6. Comparison of images: (a) plane-wave RTM image of the Marmousi2 model and (b) plane-wave LSRTM image of the Marmousi2 model after 30 iterations. All the 31 plane-wave gathers are used.

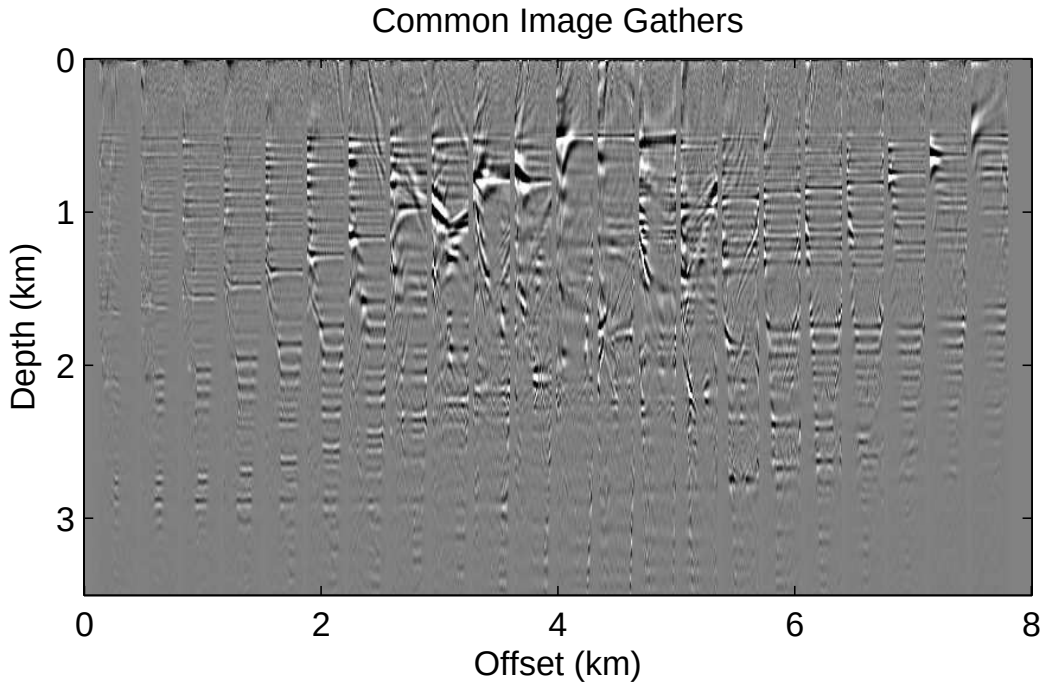

Figure 3.7. The common image gathers extracted from the plane-wave RTM image.

multiples. With an increasing number of iterations, the quality of CIGs is improved, as shown in Figure 3.8. The low-wavenumber artifacts are mostly removed in the CIGs from the LSRTM image. However, in the complex region of the model (faults), some steeply-dipping coherent noise is preserved during the iterations of LSM.

3.3.6 Computational and I/O Cost

Table 3.1 summaries the computational cost and the I/O cost of each of the migration methods. All the calculations are scaled to the conventional RTM method. The conventional LSRTM in shot-domain is the most expensive method with computational cost 60, assuming each iteration takes twice the computation of one RTM operation and 30 iterations are needed. The computational cost of standard plane-wave RTM is $\frac{31}{801} = 0.04$ and the cost of LSRTM for 30 iterations is $\frac{31*2*30}{801} = 2.32$ times that of standard RTM. The dynamic LSRTM is very efficient and its computational cost is only $\frac{31*2}{801} = 0.08$ of standard RTM. In terms of I/O cost, there are only 31 plane-wave gathers, so the plane-wave RTM and LSRTM methods only have $\frac{31}{801} = 4\%$ of the I/O cost of conventional RTM for the Marmousi2 model example, if the plane-wave gathers can be stored in the memory.

3.3.7 Sensitivity to Velocity Error

To test how the proposed method performs in the presence of migration velocity errors, the migration velocity model in Figure 3.2(b) is scaled by 0.95 to introduce a 5% velocity error, and the above 31 plane-wave gathers are migrated with the wrong velocity. If the inversion is computed by conventional LSM, the convergence stalls at high data residual (dashed line with stars in Figure 3.9), but inverting prestack images separately (see equation 3.28) provides a robust convergence (solid line with circles in Figure 3.9). Figure 3.10 shows (a) the RTM image, (b) the conventional plane-wave LSRTM image after 30 iterations, and (c) the plane-wave prestack LSRTM image (30 iterations also). It is obvious that the plane-wave LSRTM image contains many fewer artifacts present in the RTM image. The result obtained with the new method (Figure 3.10(c)) contains fewer high frequency artifacts and shows better continuity for many reflectors compared to the result obtained by conventional method.

Figure 3.11 shows the CIGs extracted from the RTM image, where all the events are curved upwards indicating that migration velocity is too low. The CIGs extracted from the plane-wave LSRTM image (Figure 3.12) are more continuous and show better S/N. There are no CIGs available from conventional least-squares migration.

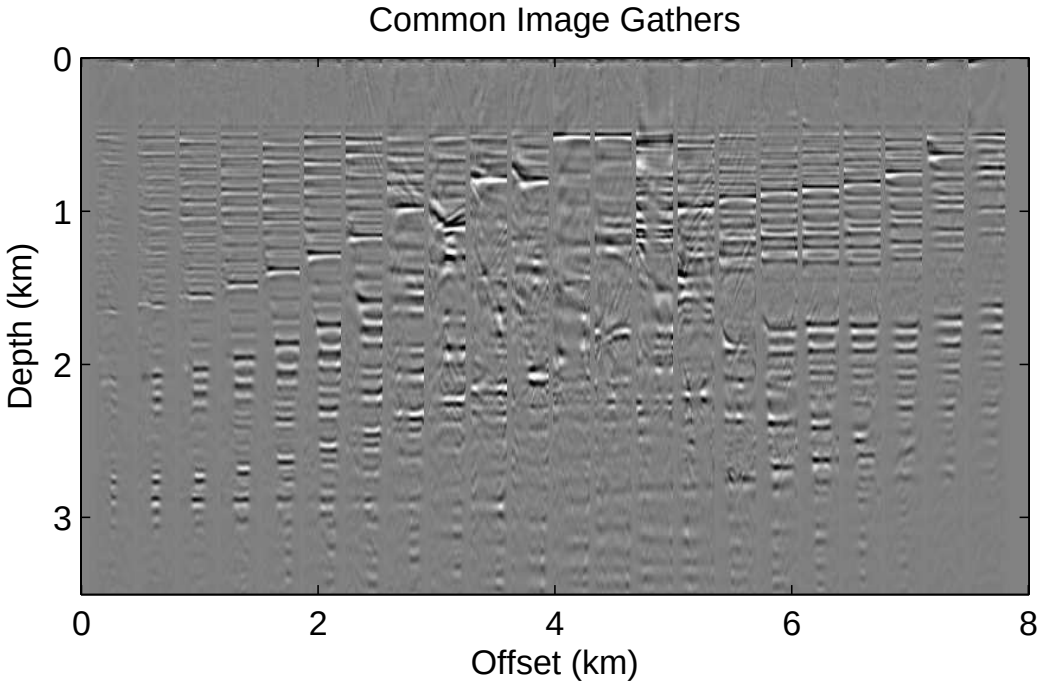

Figure 3.8. The common image gathers extracted from the plane-wave LSRTM image after 30 iterations.

Table 3.1. LSRTM and RTM computational cost, I/O expense, image quality and sensitivity to errors in the migration velocity for the example of Marmousi2 model with a fixed spread acquisition geometry.

	RTM	Shot-domain LSRTM	Plane-wave RTM	Plane-wave LSRTM	Dynamic LSRTM
Computation Cost	1	60	0.04	2.32	0.08
I/O Cost	1	1	0.04	0.04	0.04
Image Quality	good	highest	better	highest	better
Sensitivity to Velocity Errors	good	fair	good	good	good

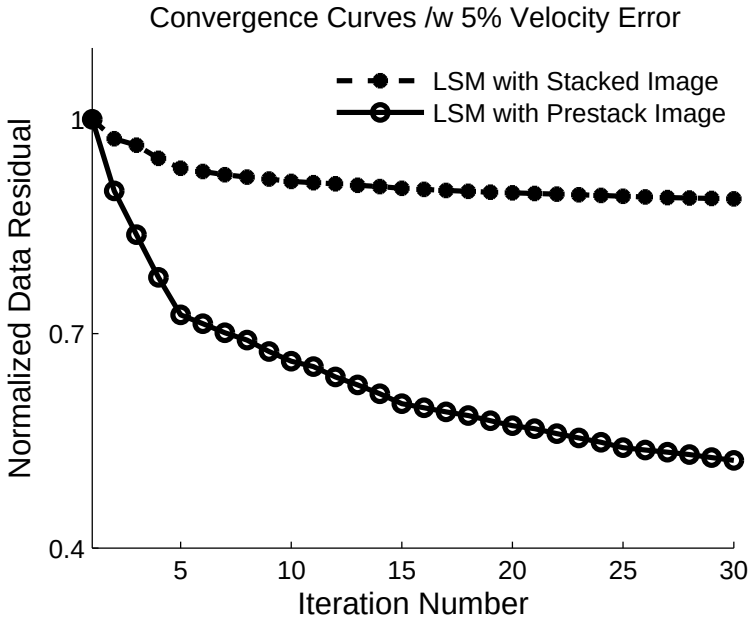

Figure 3.9. The convergence curves for LSRTM with the stacked image and the prestack image. It is clear that the convergence of LSRTM is improved when more unknowns are incorporated into the inversion and the migration velocity contains 5% error.

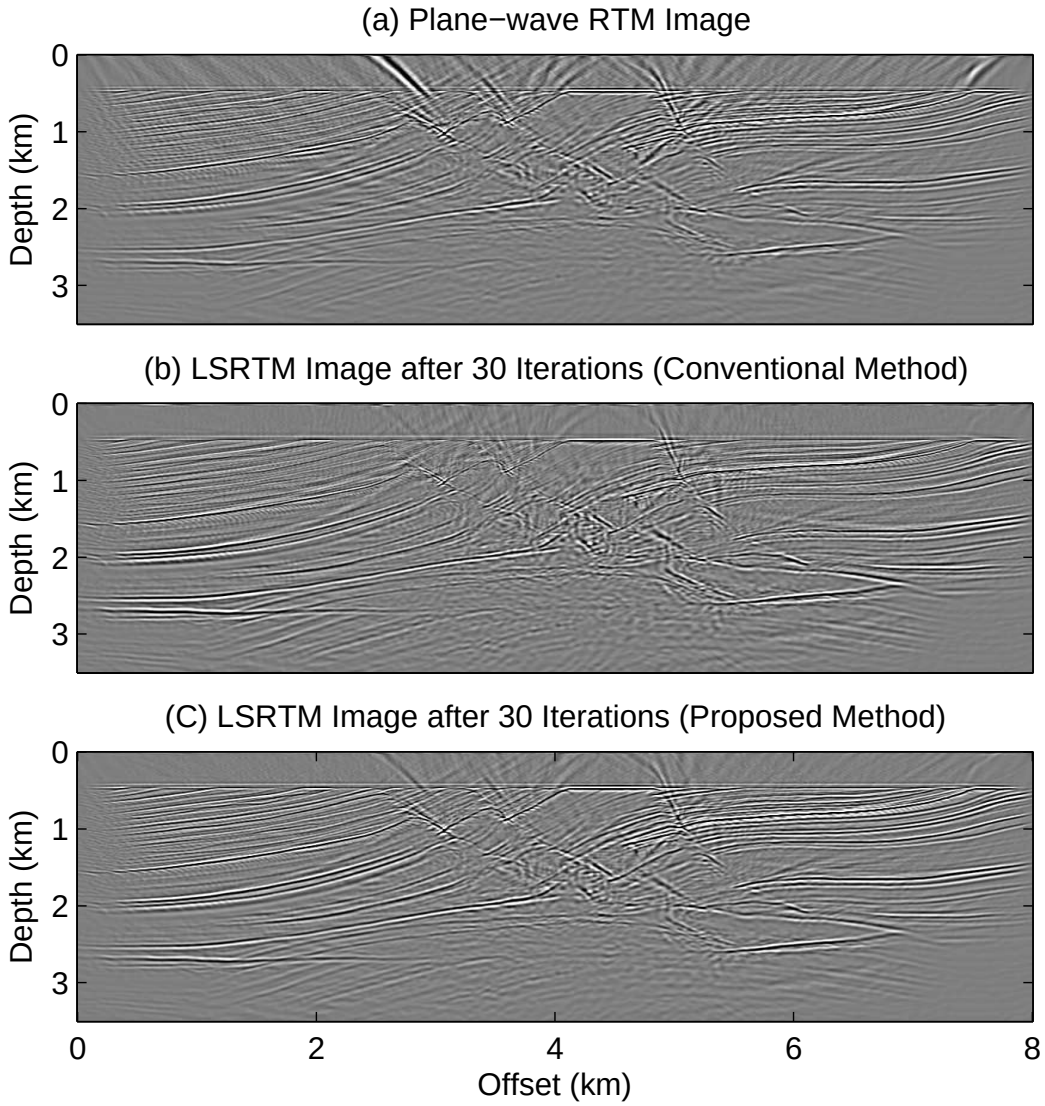

Figure 3.10. Migration images with wrong velocity: (a) the plane-wave RTM image of the Marmousi2 model, (b) image obtained by plane-wave LSRTM with the stacked image after 30 iterations, and (c) image obtained by plane-wave LSRTM with the prestack image after 30 iterations. All the 31 plane-wave gathers are migrated with 5% velocity error.

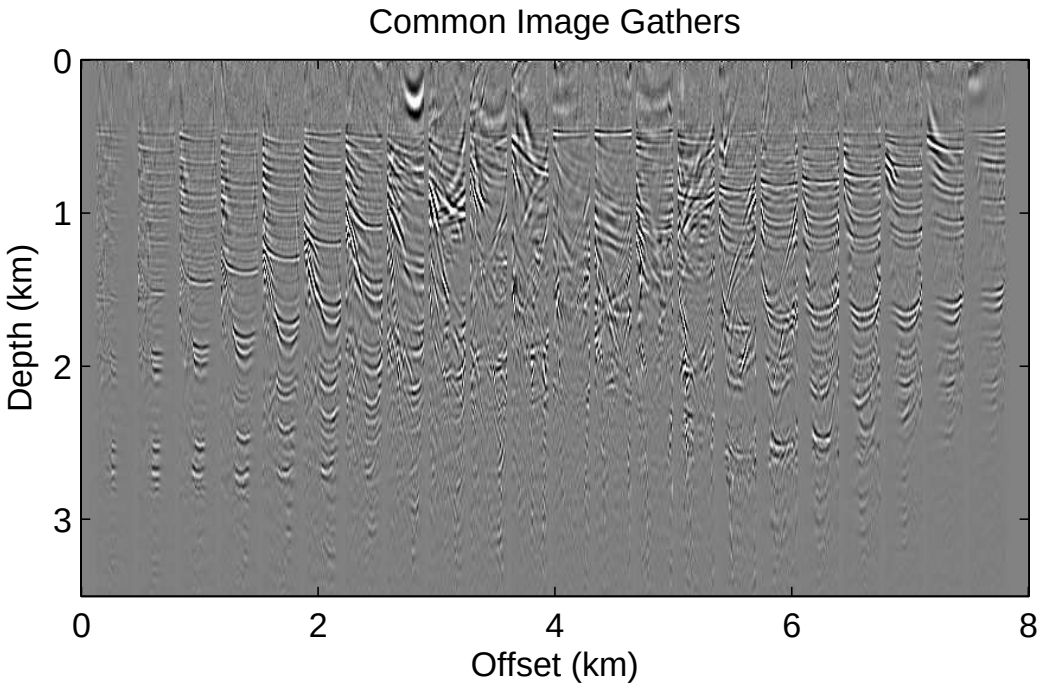

Figure 3.11. The common image gathers extracted from the plane-wave RTM image obtained with 5% velocity error.

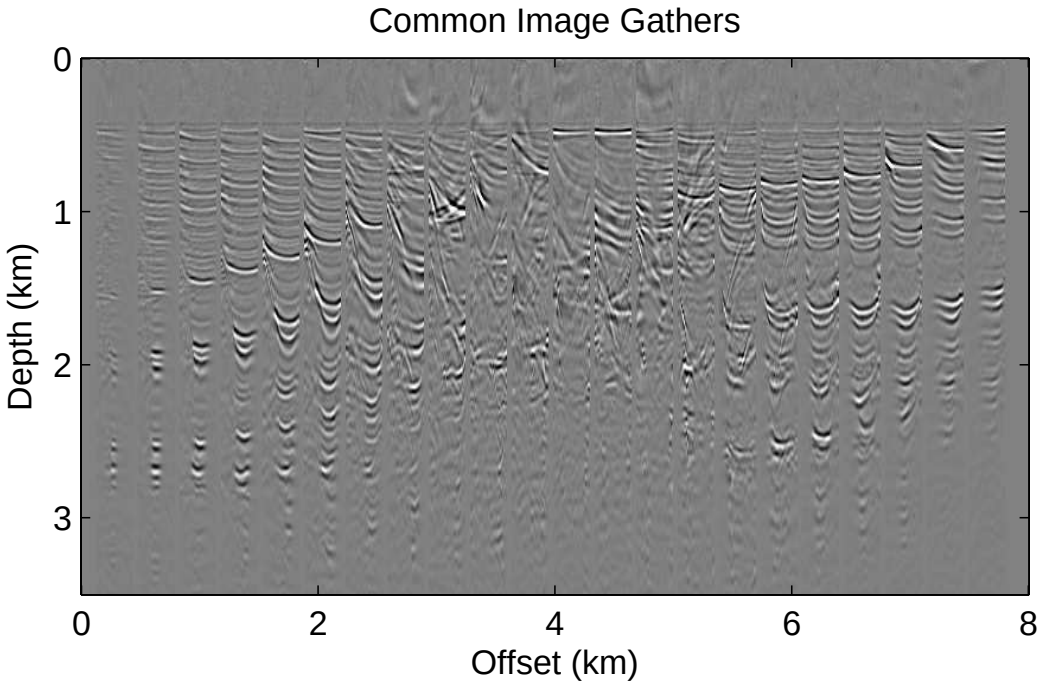

Figure 3.12. The common image gathers extracted from the plane-wave LSRTM image after 30 iterations when the migration velocity contains 5% error.

3.3.8 Field Data Example

The proposed methods are tested on a 2D marine dataset. There are 515 shots with a shot interval of 37.5 m, and each shot is recorded by a 6 km long cable with 480 receivers and a 12.5 m interval. The nearest offset is 198 m. These 515 CSGs are transformed into common midpoint profiles (CMPs), and 2D spline interpolation is used to fill in the near-offset trace gap after normal moveout. The interpolated data are then transformed into common receiver gathers with a split-spread acquisition geometry using reciprocity (Vigh and Starr, 2008). In the CRG domain, each trace is multiplied by $\sqrt{i/\omega}$ in the frequency domain and then scaled by $\sqrt{t}$ in the time domain to correct for 3D geometrical spreading (Zhou et al., 1997). A tau-p transform is applied to each CRG to generate 31 plane-wave gathers with ray parameters (p) ranging from -333μs/m to 333 μs/m with an even sampling in p. The plane-wave gathers are filtered with a Wiener filter to transform the original wavelet to a Ricker wavelet with a 25 Hz peak frequency. The original wavelet is estimated by stacking traces with a strong water bottom reflection, and windowing the water-bottom reflection event. Figure 3.13 shows the plane-wave gather with a surface shooting angle of zero ($p = 0$). The migration velocity (Figure 3.14) is obtained by waveform inversion (Boonyasiriwat et al., 2010).

3.3.9 Shot-domain RTM

The dataset is first migrated with conventional shot-domain RTM method after preprocessing, and the image is shown in Figure 3.15(a), which contains strong artifacts near the shallow reflectors, which are caused by head waves and diving waves (Liu et al., 2011). In the bottom right corner of the image, there are low-frequency horizontal stripes. In the zoom views (Figures 3.16(a) and 3.17(a)), the RTM image shows double-dipping near-vertical artifacts.

3.3.10 Plane-wave RTM

Figure 3.15(b) shows the plane-wave RTM image, which is of higher quality compared to shot-domain RTM image. The artifacts in the shallow part are eliminated in the plane-wave RTM image because head and diving waves do not strongly appear in the plane-wave gathers (Figure 3.13). Similarly, the low-frequency horizontal stripes in the bottom right corner of the conventional RTM image are removed in the plane-wave RTM image. The zoom views (Figures 3.16(b) and 3.17(b)) also show better continuity for the reflectors and fewer artifacts compared to the same zoom areas in Figures 3.16(a) and 3.17(a).

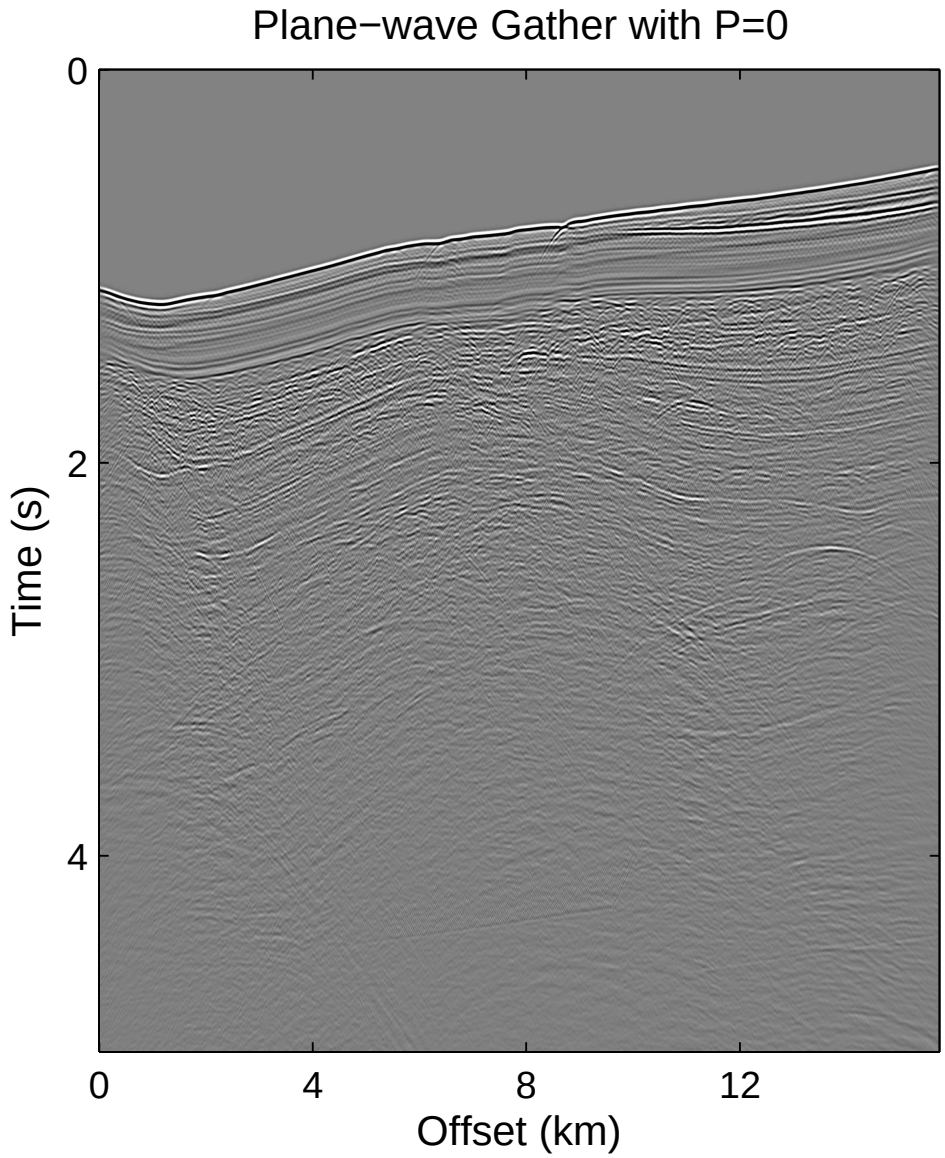

Figure 3.13. A plane-wave gather with zero surface shooting angle (p=0 μs/m).

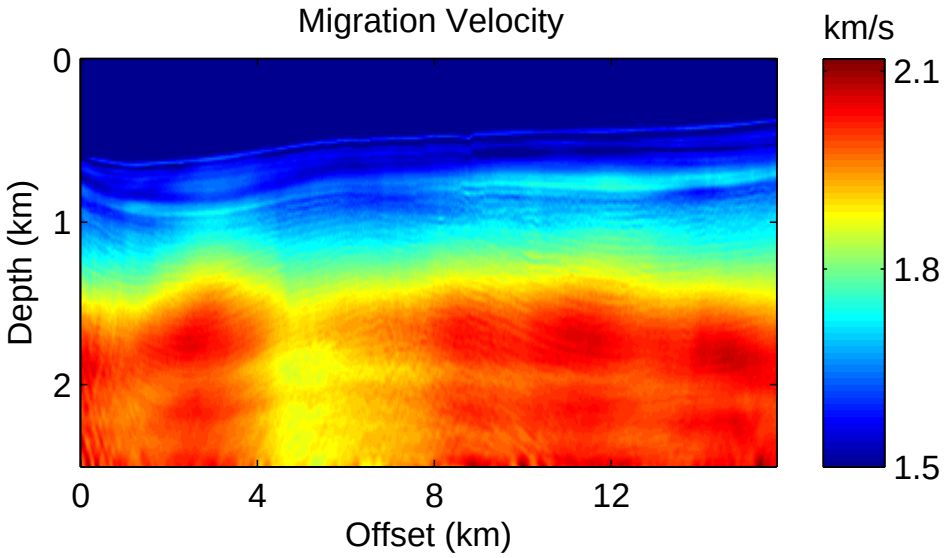

Figure 3.14. The migration velocity model for the field data test, obtained by full waveform inversion.

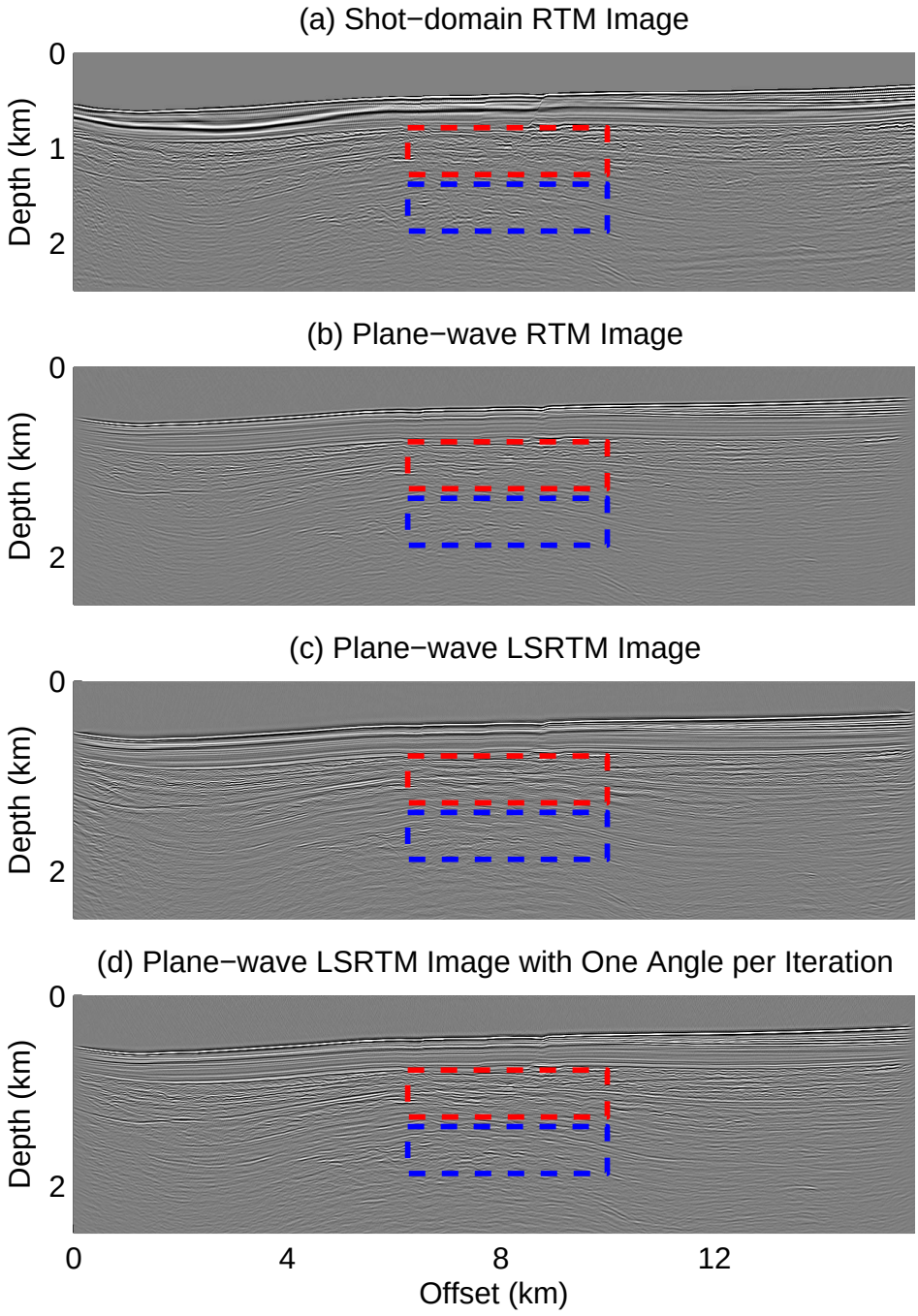

Figure 3.15. The migration images obtained by: (a) conventional shot-domain reverse time migration, (b) plane-wave reverse time migration, (c) plane-wave least-squares reverse time migration and (d) plane-wave LSRTM with dynamic encoding. The blue and red boxes indicate the areas for zoom view.

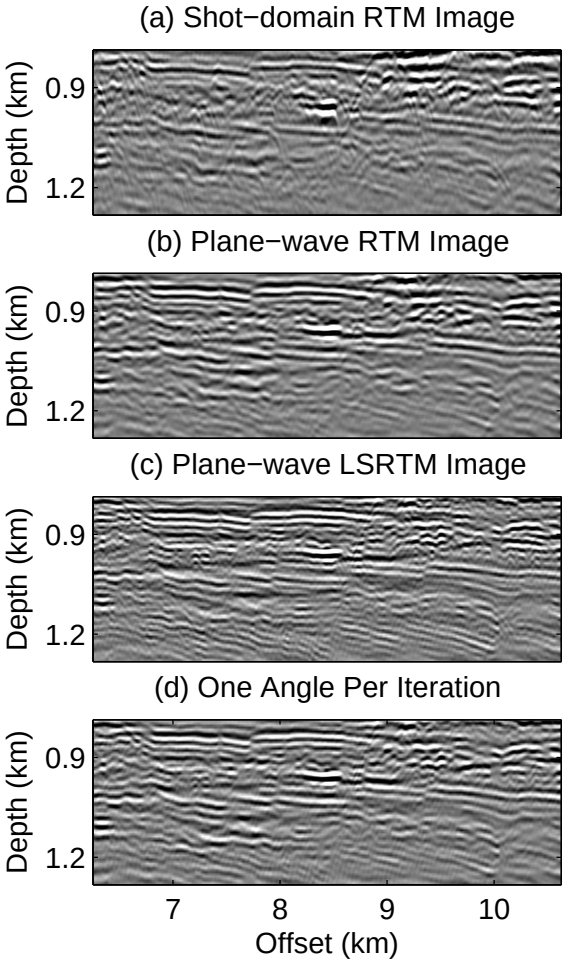

Figure 3.16. The zoom views of the red boxes: (a) conventional shot-domain RTM, (b) the plane-wave RTM, (c) the plane-wave LSRTM and (d) the plane-wave LSRTM images with dynamic encoding.

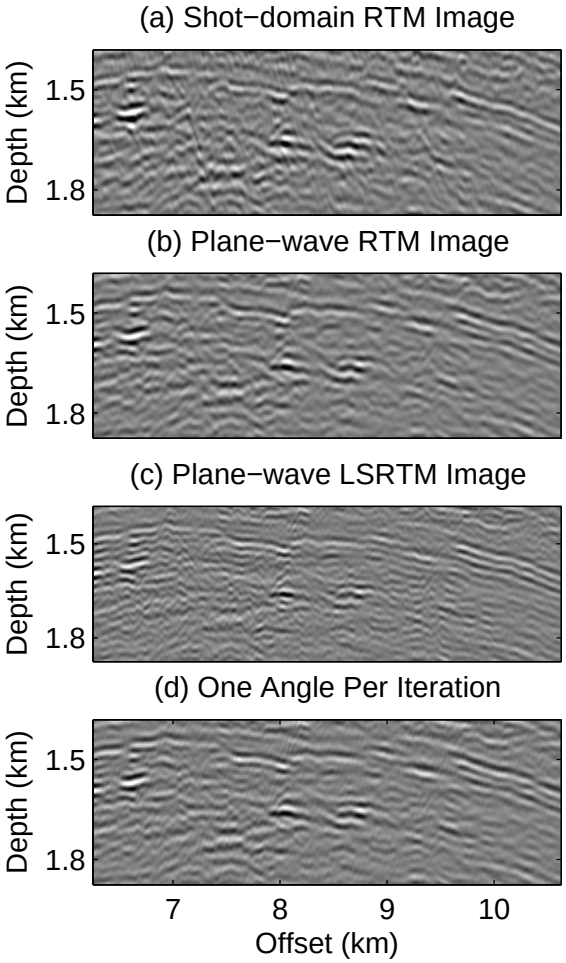

Figure 3.17. The zoom views of the blue boxes: (a) conventional shot-domain RTM, (b) the plane-wave RTM, (c) the plane-wave LSRTM and (d) the plane-wave LSRTM images with dynamic encoding.

3.3.11 Plane-wave Prestack LSRTM

In order to further improve the quality of the image, the plane-wave LSRTM algorithm (equation 3.28) is applied to these 31 plane-wave gathers and the image after 30 iterations is shown in Figure 3.15(c). Compared to the plane-wave RTM image, the LSRTM image contains fewer artifacts and shows better resolution. In the zoom view of the red boxes (Figure 3.16), the horizontal reflectors in the LSRTM image are of higher resolution and are characterized by better balancing of amplitudes, which provides better delineation of the normal faults compared to the RTM image. In the deep part of the section (blue boxes, Figure 3.17), the LSRTM image shows similar advantages to better illuminate the faults, in spite of the fact that these reflectors become more wiggly in the LSRTM image.

The common image gathers can be extracted from the prestack images and they suggest that the migration velocity is not very accurate in the deep part and the reflectors are undermigrated (Figure 3.18 and 3.19). In this example, the LSRTM algorithm can only marginally improve the quality of the CIGs. The vertical artifacts at both edges of any one CIG are removed, and some of the reflector amplitudes are enhanced, so that they appear to be more continuous. In spite of the errors in the velocity model, the convergence of plane-wave LSRTM is stable and robust (solid line with squares in Figure 3.20). In this example, the plane-wave LSRTM still shows better convergence than the conventional LSM.

3.3.12 Dynamic Plane-wave LSRTM

When high computational efficiency is in demand, LSRTM can be performed with the dynamic encoding approach (Krebs et al., 2009; Schuster et al., 2011), where one plane-wave gather is used for each iteration and the ray parameter p (corresponding to surface shooting angle) is dynamically changed from one iteration to another. Figure 3.15(d) shows the LSRTM image with dynamic encoding after 31 iterations. It has resolution comparable to Figure 3.15(c) but contains more noticeable artifacts (see Figure 3.16(d) and 3.17(d)).

3.3.13 Computational and I/O Cost

The computational cost of the plane-wave RTM is about $\frac{31\times3}{515} \approx \frac{1}{5}$ of that for conventional shot-domain RTM. Each iteration of the LSRTM is assumed to cost twice that of the RTM method. So, for 30 iterations, the computational cost of the plane-wave LSRTM is about 12 times that of the conventional RTM. The computational cost of the dynamic plane-wave LSRTM is only about 40% of that for conventional RTM. The drawback is that CIGs are not available for velocity analysis and the convergence is lessened because at every iteration the problem is redefined with a new encoding function. For the I/O

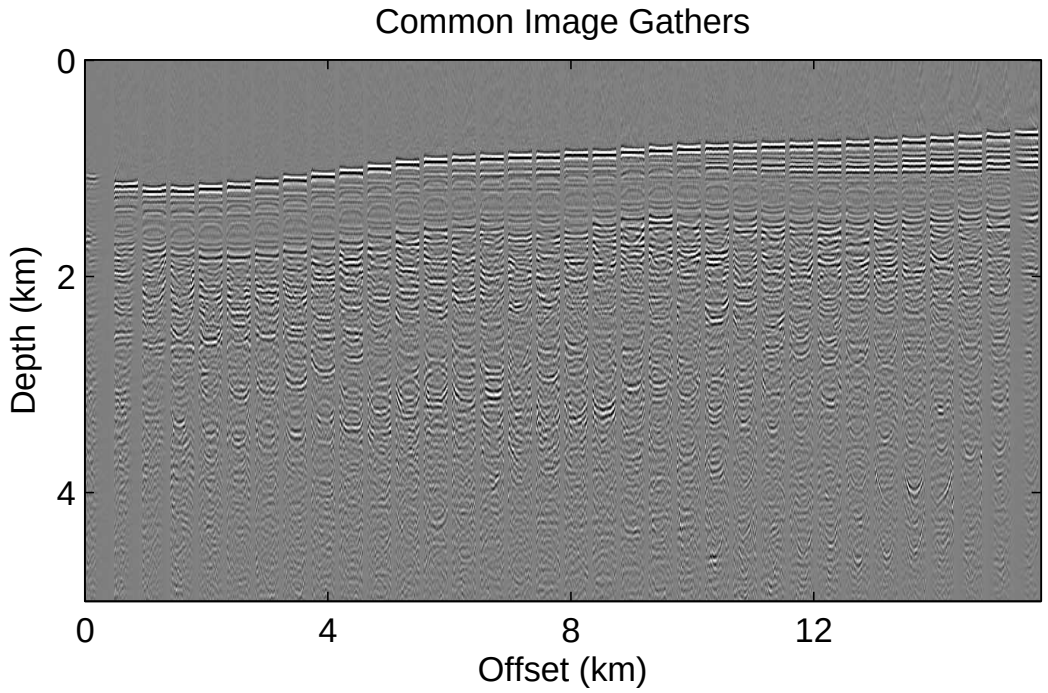

Figure 3.18. The common image gathers extracted from the plane-wave RTM image of the field data.

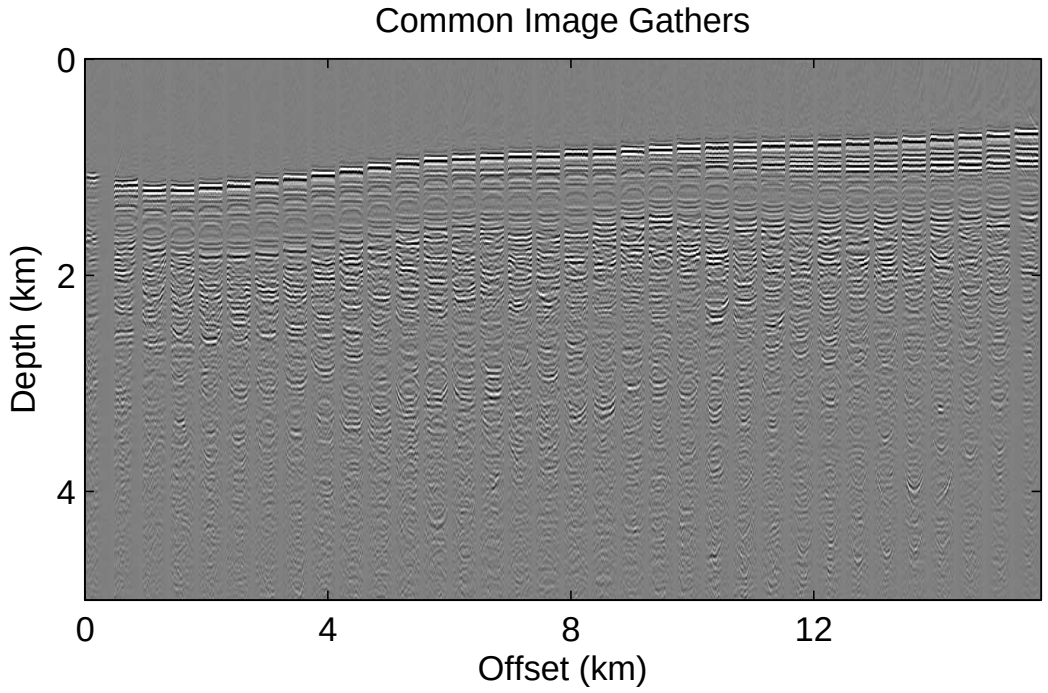

Figure 3.19. The common image gathers extracted from the plane-wave LSRTM image after 30 iterations for the field data test.

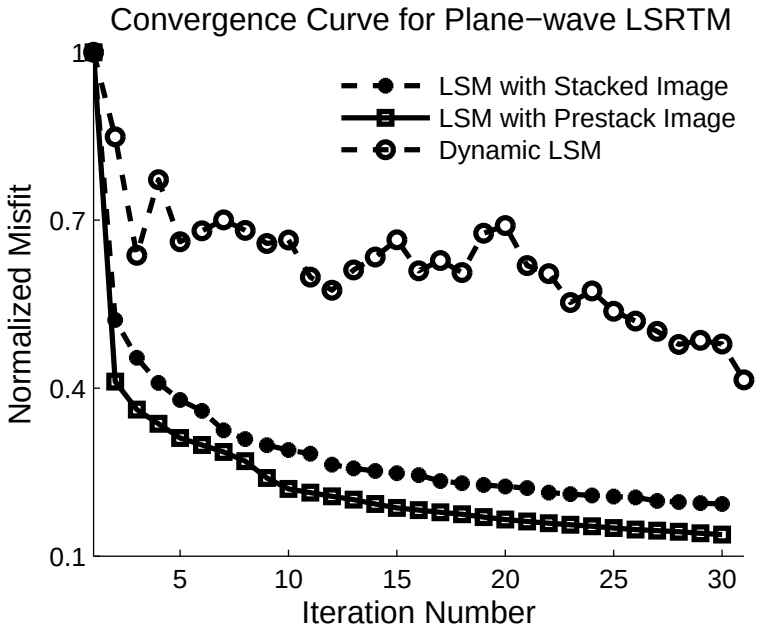

Figure 3.20. The misfit vs iteration number curve for plane-wave LSRTM shows fast and stable convergence even when the velocity is not completely accurate.

costs, conventional RTM inputs 515 shots with 480 traces each, and plane-wave migration only needs to read 31 plane-wave gathers with 1260 traces each. Hence, the I/O cost of plane-wave migration and plane-wave least-squares migration is $\frac{31*1260}{515*480} = 0.15$ that of conventional RTM, if all the data can be stored in the physical memory, so it might be more suitable for GPU calculations. Table 3.2 shows the comparison of different methods in terms of computational and I/O cost.

3.4 Discussion and Conclusion

We propose plane-wave LSRTM as an efficient alternative to RTM and, unlike random phase-encoded LSRTM, a viable method for marine data. To reduce the computational cost, the original shot-domain data are transformed into dozens of plane-wave gathers. The advantages include (1) stable convergence even with velocity errors up to 5% in my example and (2) the common image gathers are available for quality control and migration velocity analysis. The drawback is that extra memory is needed to store all the prestack LSRTM images at any one iteration. Numerical tests on the Marmousi2 model and a 2D marine dataset show that plane-wave prestack LSRTM can produce images with fewer migration artifacts, and higher resolution compared to a corresponding RTM image. The method shows good convergence even when the velocity model is not accurate. I conclude that the least-squares reverse time migration in the plane-wave domain can be an efficient method to improve the quality of RTM images and produce common image gathers for MVA.

Table 3.2. LSRTM and RTM computational cost, I/O expense, image quality and sensitivity to errors in the migration velocity for the field data example with a marine streamer acquisition geometry.

	RTM	Shot-domain LSRTM	Plane-wave RTM	Plane-wave LSRTM	Dynamic LSRTM
Computation Cost	1	60	0.2	12	0.4
I/O Cost	1	1	0.15	0.15	0.15
Image Quality	good	highest	better	highest	better
Sensitivity to Velocity Errors	good	fair	good	good	good

CHAPTER 4

REVERSE TIME MIGRATION OF PRISM WAVES FOR SALT FLANK DELINEATION

In this chapter, I present a new reverse time migration method for imaging salt flanks with prism wave reflections. It consists of four steps: (1) migrating the seismic data with conventional RTM to give the RTM image; (2) using the RTM image as a reflectivity model to simulate source-side reflections with the Born approximation; (3) zero-lag correlation of the source-side reflection wavefields and receiver-side wavefields to produce the prism wave migration image; and (4) repeating steps 2 and 3 for the receiver-side reflections. An advantage of this method is that there is no need to pick the horizontal reflectors prior to migration of the prism waves. It also separately images the vertical structures at a different step to reduce crosstalk interference. The disadvantage of prism wave migration algorithm is that its computational cost is twice that of conventional RTM. The empirical results with a salt model suggest that prism wave migration can be an effective method for salt flank delineation in the absence of diving waves.

4.1 Introduction

Vertical structures such as salt flanks are usually not illuminated by primary reflections and so cannot be well imaged by conventional migration methods (Hale et al., 1992). If on the other hand strong diving waves are present, they can be reflected from the salt flank, recorded on the surface, and migrated by a two-way migration method, such as Kirchhoff migration (Ratcliff et al., 1991, 1992) or reverse time migration (RTM) (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983). Even a one-way migration method can be modified (Hale et al., 1992) to incorporate diving waves for salt flank imaging.

If the diving wave is not extant due to the absence of a strong velocity gradient or a limited recording aperture, prism waves can be migrated to illuminate vertical reflectors. A prism wave is defined to be a doubly scattered wave from, typically, a vertical reflector, as illustrated by the ray diagram in Figure 4.1(a). Cavalca and Lailly (2005) studied

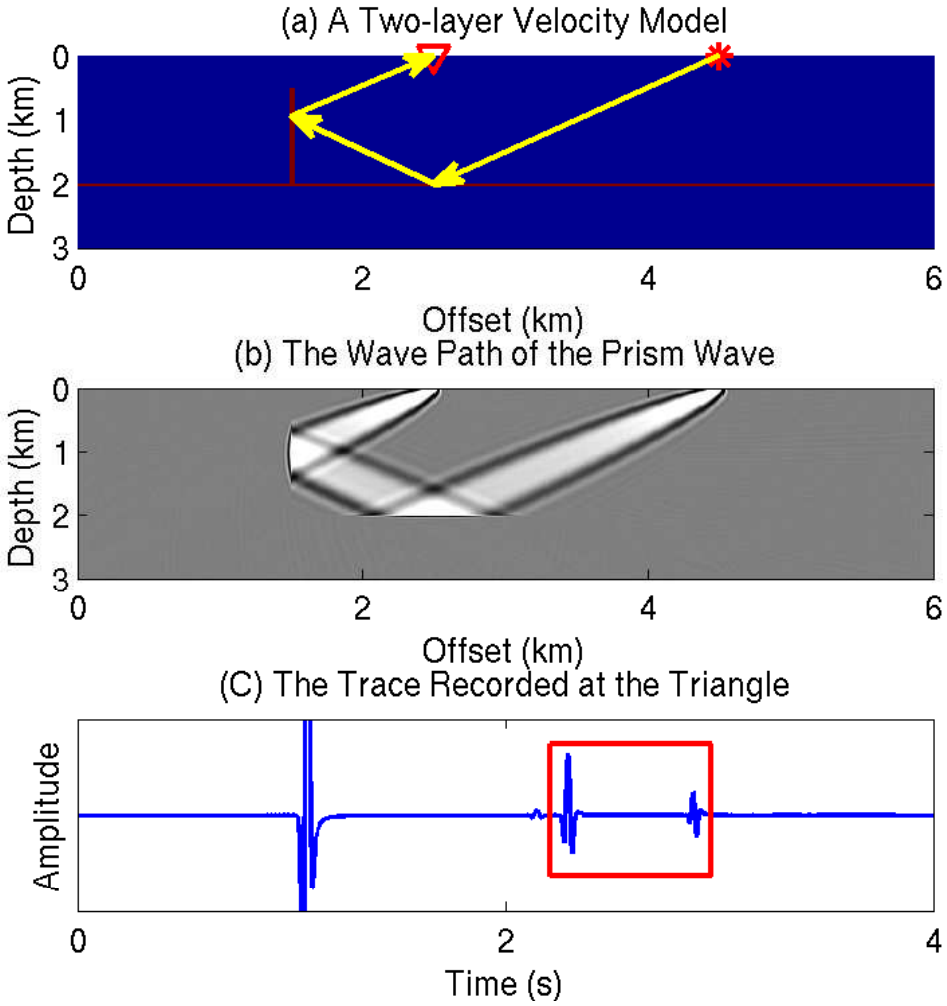

Figure 4.1. Diagrams of a prism wave: (a) a velocity model with a horizontal reflector and a vertical reflector. The yellow arrows indicate the ray path for a prism wave from the source at the star to the receiver at the triangle; (b) the wave path of the prism wave with a 20-Hz Ricker wavelet; and (c) the trace recorded at the triangle. The two arrivals in the red window are the reflections from the horizontal reflector and the prism wave in panel (b).

the kinematics of prism waves and explored the possibility of incorporating the prism waves in traveltime inversion for salt flank locations. To incorporate amplitudes in the imaging, Marmalyevskyy et al. (2005) migrated the prism waves by a Kirchhoff-based method for salt flank delineation with subhorizontal reflection boundaries specified from the previous migration images. An iterative method was proposed by Malcolm et al. (2009) to progressively incorporate migration of prism waves and multiples with a modified one-way wave equation migration method, where each phase was isolated by a data fitting process. At each step, different partial images were computed to illuminate different structures, e.g., the prism waves for salt flanks. They later tested their method on North sea field data with the introduction of a regularization term for the inversion (Malcolm et al., 2011).

With reverse time migration, the migration of the prism waves can be accommodated in the process by embedding the subhorizontal reflection boundaries in the velocity model (Jones et al., 2007). However, incorporating the sharp boundaries into the velocity model is not trivial, and the complex migration velocity will excite complex wavefields that lead to artifacts in the RTM images (Liu et al., 2011). Another problem is that prism waves are doubly scattered waves, which are usually weaker than primaries, so that the contribution from the prism waves might be weak. In this chapter, I propose a new RTM method for migrating the prism waves separately from the other reflectors by utilizing the migration image from conventional RTM. The advantages of this approach over conventional RTM are as follows: (1) It does not require modifying the migration velocity as conventional RTM does; (2) It separately images different structures at different steps and reduces the artifacts from crosstalk of different phases. The disadvantage of the proposed method is that its computational cost is twice that of conventional RTM.

This chapter is organized into four sections. The first one is this introduction, which is followed by the theory section. In the numerical results section, the synthetic examples of a simple model and a salt model are presented. A summary will be provided in the end.

4.2 Theory

In the frequency domain, reverse time migration of a shot gather $d(\mathbf{x}_g|\mathbf{x}_s)$ can be expressed as

$$m_{mig}(\mathbf{x}|\mathbf{x}_s) = \sum_{\omega}\sum_{g} \omega^2 W^*(\omega) G^*(\mathbf{x}|\mathbf{x}_s) G^*(\mathbf{x}|\mathbf{x}_g) d(\mathbf{x}_g|\mathbf{x}_s), \tag{4.1}$$

where $m_{mig}(\mathbf{x}|\mathbf{x}_s)$ is the migration image of the shot at $\mathbf{x}_s$, $W(\omega)$ is the source spectrum, $\mathbf{x}_g$ indicates the receiver location, $G(\mathbf{x}|\mathbf{x}_s)$ is the Green's function from a source at $\mathbf{x}_s$ to

$\mathbf{x}$; This Green's function is computed by a finite-difference solution to the wave equation. The $*$ indicates complex conjugate. For simplicity, the angular frequency ω is silent in the Green's function G and data function d.

For the velocity model in Figure 4.1(a), referred to as the L model, the recorded data contain prism waves. The yellow arrows in Figure 4.1(a) indicate the ray path for a prism wave excited at $(x, z) = (4.5, 0)\ km$ and recorded at $(x, z) = (2.5, 0)\ km$, and Figure 4.1(b) depicts the wavepath (Luo and Schuster, 1991) of the prism wave generated by a source with a 20-Hz Ricker wavelet. The recorded trace is plotted in Figure 4.1(c) with a red window outlining the reflection from the horizontal reflector and the prism wave. For simplicity, I mute the direct wave and diffractions from the trace to keep only the part in the red window

$$d(\mathbf{x}_g|\mathbf{x}_s) = d_1(\mathbf{x}_g|\mathbf{x}_s) + d_2(\mathbf{x}_g|\mathbf{x}_s), \tag{4.2}$$

where $d_1(\mathbf{x}_g|\mathbf{x}_s)$ and $d_2(\mathbf{x}_g|\mathbf{x}_s)$ denote the first-order scattering reflection wave and the doubly scattered prism wave, respectively. When the horizontal reflector is extracted from the migration images and embedded in the migration velocity model (Figure 4.3(a)), conventional RTM can correctly migrate the prism waves to image the vertical reflector (Jones et al., 2007). In this case, the Green's function calculated with the migration velocity in Figure 4.3(a) contains two arrivals: a direct wave arrival and a reflection from the horizontal reflector as shown in Figure 4.2. Therefore, the Green's functions in equation 4.1 can be decomposed into two parts:

$$G(\mathbf{x}|\mathbf{x}_s) = G_o(\mathbf{x}|\mathbf{x}_s) + G_1(\mathbf{x}|\mathbf{x}_s), \tag{4.3}$$

and

$$G(\mathbf{x}|\mathbf{x}_g) = G_o(\mathbf{x}|\mathbf{x}_g) + G_1(\mathbf{x}|\mathbf{x}_g), \tag{4.4}$$

where G_o and G_1 denote the direct and the reflected waves, respectively. Note that in this case G_o is a downgoing wave and G_1 is an upgoing wave.

When the data in the red window of Figure 4.1(c) are migrated with the velocity model in Figure 4.3(a), the migration image is shown in Figure 4.3(b), and is mathematically described by

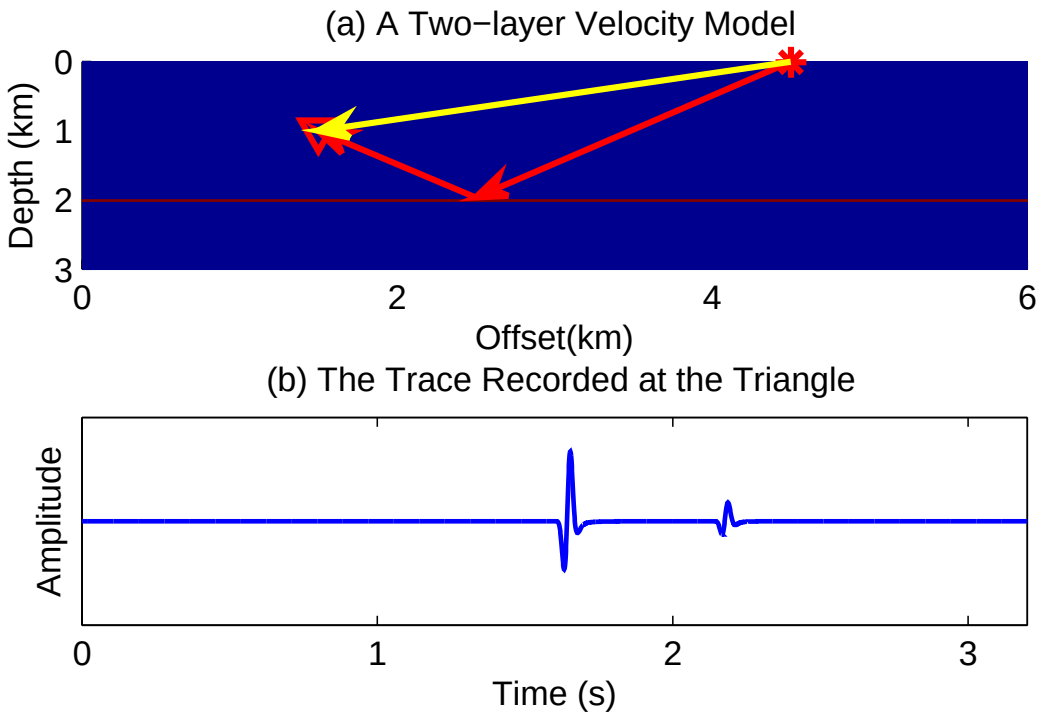

Figure 4.2. Ray diagrams for the Green's functions: (a) a two-layer velocity model. The star and triangle indicate the source and receiver locations. The yellow arrow is the ray path for the direct wave and the red arrows show the ray path for the reflected wave. (b) The trace recorded at the triangle. It is simulated with a 20-Hz Ricker wavelet.

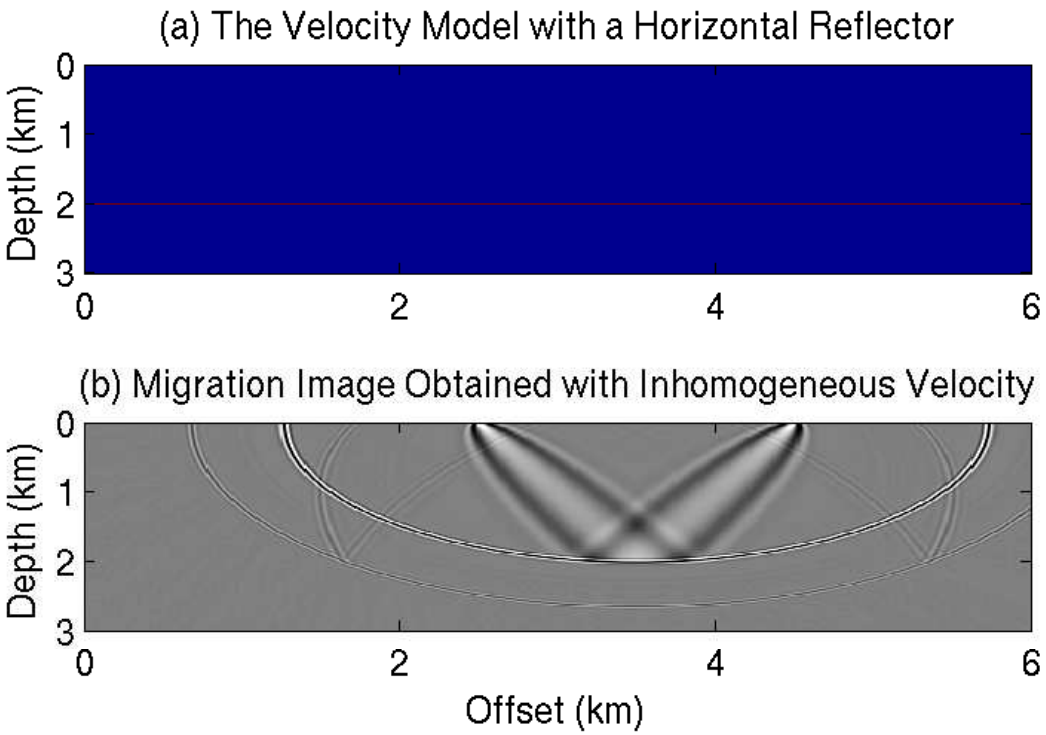

Figure 4.3. When the data are migrated with (a) the homogeneous velocity (2 km/s) with a horizontal reflector embedded (2.5 km/s); (b) the migration image of the data within the red window in Figure 4.1(a) with the velocity model in panel (a).

$$
\begin{aligned}
& m_{mig}(\mathbf{x}|\mathbf{x}_s) \\
= & \sum_{\omega} \omega^2 W^*(\omega)[G_o^*(\mathbf{x}|\mathbf{x}_s) + G_1^*(\mathbf{x}|\mathbf{x}_s)][G_o^*(\mathbf{x}|\mathbf{x}_g) + G_1^*(\mathbf{x}|\mathbf{x}_g)][d_1(\mathbf{x}_g|\mathbf{x}_s) + d_2(\mathbf{x}_g|\mathbf{x}_s)] \\
= & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) G_o^*(\mathbf{x}|\mathbf{x}_g) d_1(\mathbf{x}_g|\mathbf{x}_s)}^{First\ Ellipse \sim O(r)} && (4.5) \\
+ & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s)}^{Second\ Ellipse \sim O(r^2)} && (4.6) \\
+ & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_1^*(\mathbf{x}|\mathbf{x}_s) G_o^*(\mathbf{x}|\mathbf{x}_g) d_1(\mathbf{x}_g|\mathbf{x}_s)}^{Left\ Rabbit\ Ear \sim O(r^2)} && (4.7) \\
+ & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) G_1^*(\mathbf{x}|\mathbf{x}_g) d_1(\mathbf{x}_g|\mathbf{x}_s)}^{Right\ Rabbit\ Ear \sim O(r^2)} && (4.8) \\
+ & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_1^*(\mathbf{x}|\mathbf{x}_s) G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s)}^{First\ Prism\ Wave\ Kernel \sim O(r^3)} && (4.9) \\
+ & \overbrace{\sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) G_1^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s)}^{Second\ Prism\ Wave\ Kernel \sim O(r^3)} && (4.10) \\
+ & other\ terms. && (4.11)
\end{aligned}
$$

Note that the summation over the receiver g is omitted because there is only one trace in this example. With the assumption that the reflection coefficient is the angle-independent value r, the amplitude of the direct wave Green's function G_o is on the order of $O(1)$ and the amplitude of the reflection wave G_1 is on the order of $O(r)$. Similarly, d_1 is with strength of $O(r)$. The prism wave d_2 is a doubly scattered wave and its amplitude is on the order $O(r^2)$. As an example, the first prism wave term in equation 4.9 has $O(r^3)$ because it is a product of the d_2 term with amplitude $O(r^2)$ and the migration kernel $G_1 \times G_o$ with strength $O(r)$. With these assumptions, the amplitude of each term in the above equation can be expressed in terms of r as shown in the labels.

Figure 4.3(b) shows two ellipses. The first one corresponds to the migration kernel in equation 4.5 with the strongest amplitude $O(r)$. When the prism wave is migrated as a primary wave (the term in equation 4.6), it shows up as the second ellipse in Figure 4.3(b) with an amplitude $O(r^2)$. This ellipse is an artifact. The migration kernels in equations 4.7 and 4.8 correspond to these two "rabbit ears" with the strength $O(r^2)$. Equations 4.9 and 4.10 contain the migration kernels for the prism waves corresponding to these near-vertical curves in Figure 4.3(b) and their amplitudes are on the order of $O(r^3)$, which are much

weaker than other kernels, so in the migration image, the vertical reflector is of weaker amplitude compared to the horizontal ones.

4.2.1 Prism Wave Reverse Time Migration

If the migration kernels in equations 4.9 and 4.10 can be computed directly, the prism waves can be directly migrated without crosstalk interference. In the following section, frequency domain formulas are used for mathematical simplicity, but the numerical calculation is actually computed in the time domain by a finite-difference solution to the space-time acoustic wave equation. Given a smooth migration velocity (homogeneous velocity in this example) and a migration image of the horizontal reflector, the Green's function for the reflected wave can be computed with the Born approximation (Stolt and Benson, 1986)

$$G_1(\mathbf{x}|\mathbf{x}_s) = \int_{\mathbf{x}'} \omega^2 m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}_s) G_o(\mathbf{x}'|\mathbf{x}) d\mathbf{x}', \tag{4.12}$$

where $m_1(\mathbf{x}')$ is the reflectivity model representing the horizontal reflector, and the Green's function G_o is calculated using the migration velocity. Plugging equation 4.12 into equation 4.9, I get

$$\begin{aligned} m_{mig}(\mathbf{x}|\mathbf{x}_s) &= \sum_{\omega} \omega^2 W^*(\omega) \int_{\mathbf{x}'} \omega^2 m_1(\mathbf{x}') G_o^*(\mathbf{x}'|\mathbf{x}_s) G_o^*(\mathbf{x}'|\mathbf{x}) d\mathbf{x}' G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s) \\ &= \sum_{\omega} \omega^2 \int_{\mathbf{x}'} \omega^2 W^*(\omega) G_o^*(\mathbf{x}'|\mathbf{x}_s) m_1(\mathbf{x}') G_o^*(\mathbf{x}'|\mathbf{x}) d\mathbf{x}' G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s) \\ &= \sum_{\omega} \omega^2 [P_1(\mathbf{x}|\mathbf{x}_s)]^* [Q_o(\mathbf{x}|\mathbf{x}_s)], \end{aligned} \tag{4.13}$$

with

$$\begin{aligned} P_1(\mathbf{x}|\mathbf{x}_s) &= \int_{\mathbf{x}'} \omega^2 W(\omega) G_o(\mathbf{x}'|\mathbf{x}_s) m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}) d\mathbf{x}'; \\ Q_o(\mathbf{x}|\mathbf{x}_s) &= G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s). \end{aligned} \tag{4.14}$$

Numerically, $P_1(\mathbf{x}|\mathbf{x}_s)$ are computed with two finite-difference simulations in the time domain to solve the following two equations

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x})) P_o(\mathbf{x}) = W(\omega)\delta(\mathbf{x} - \mathbf{x}_s); \tag{4.15}$$

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x})) P_1(\mathbf{x}) = \omega^2 m_1(\mathbf{x}) P_o(\mathbf{x}), \tag{4.16}$$

where the slowness $s_o(\mathbf{x})$ is the reciprocal of the migration velocity model. The receiver-side wavefield $Q_o(\mathbf{x}|\mathbf{x}_s)$ can be computed by solving

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x})) Q_o(\mathbf{x}|\mathbf{x}_s) = d_2(\mathbf{x}_g|\mathbf{x}_s)\delta(\mathbf{x} - \mathbf{x}_g). \tag{4.17}$$

Note the wavefield propagates backward in time when solving the above equation in the time domain with the finite-difference method. When there is more than one trace in

the shot gather, all the traces act as source wavelets of point sources at their respective recording locations, which implies a summation over the receiver g. In summary, prism wave migration requires three finite-difference simulations (equations 4.15, 4.16, and 4.17) to calculate the image corresponding to the term in equation 4.9.

Figure 4.4 illustrates the process of prism wave migration with equation 4.9: (1) The source wavefield $P_o(\mathbf{x})$ propagates downward starting from the source location; (2) $P_o(\mathbf{x})$ is reflected at the horizontal reflector and becomes the reflected wavefield $P_1(\mathbf{x})$; (3) The receiver wavefield $Q_o(\mathbf{x})$ propagates downward from the receiver; (4) The product of $P_1(\mathbf{x})$ and $Q_o(\mathbf{x})$ is the migration image (the vertical curve in Figure 4.4 is part of the prism wave migration kernel and computed by equation 4.13).

Similarly, the term in equation 4.10 can be computed by

$$\begin{aligned} m_{mig}(\mathbf{x}|\mathbf{x}_s) &= \sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) \int_{\mathbf{x}'} \omega^2 m_1(\mathbf{x}') G_o^*(\mathbf{x}'|\mathbf{x}_g) G_o^*(\mathbf{x}'|\mathbf{x}) d\mathbf{x}' d_2(\mathbf{x}_g|\mathbf{x}_s) \\ &= \sum_{\omega} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) \int_{\mathbf{x}'} \omega^2 d_2(\mathbf{x}_g|\mathbf{x}_s) G_o^*(\mathbf{x}'|\mathbf{x}_g) m_1(\mathbf{x}') G_o^*(\mathbf{x}'|\mathbf{x}) d\mathbf{x}' \\ &= \sum_{\omega} \omega^2 [P_o(\mathbf{x}|\mathbf{x}_s)]^* [Q_1(\mathbf{x}|\mathbf{x}_s)], \end{aligned} \tag{4.18}$$

with $Q_1(\mathbf{x}|\mathbf{x}_s)$ computed by a finite-difference solution to

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x})) Q_1(\mathbf{x}|\mathbf{x}_s) = \omega^2 m_1(\mathbf{x}) Q_o(\mathbf{x}|\mathbf{x}_s), \tag{4.19}$$

using the time reversed traces as source wavelets in equation 4.17.

Therefore, the migration image of the prism wave is the sum of the two terms from equations 4.13 and 4.18,

$$m_{mig}(\mathbf{x}|\mathbf{x}_s) = \omega^2 [P_1(\mathbf{x}|\mathbf{x}_s)]^* [Q_o(\mathbf{x}|\mathbf{x}_s)] + \omega^2 [P_o(\mathbf{x}|\mathbf{x}_s)]^* [Q_1(\mathbf{x}|\mathbf{x}_s)], \tag{4.20}$$

and it requires four finite-difference simulations in total. Compared to conventional RTM, its computational cost is doubled. The advantages of this approach are as follows: (1) It avoids modifying the migration velocity as in conventional RTM of prism waves; (2) vertical structures are imaged in a separate step and reduces the crosstalk interference between different phases.

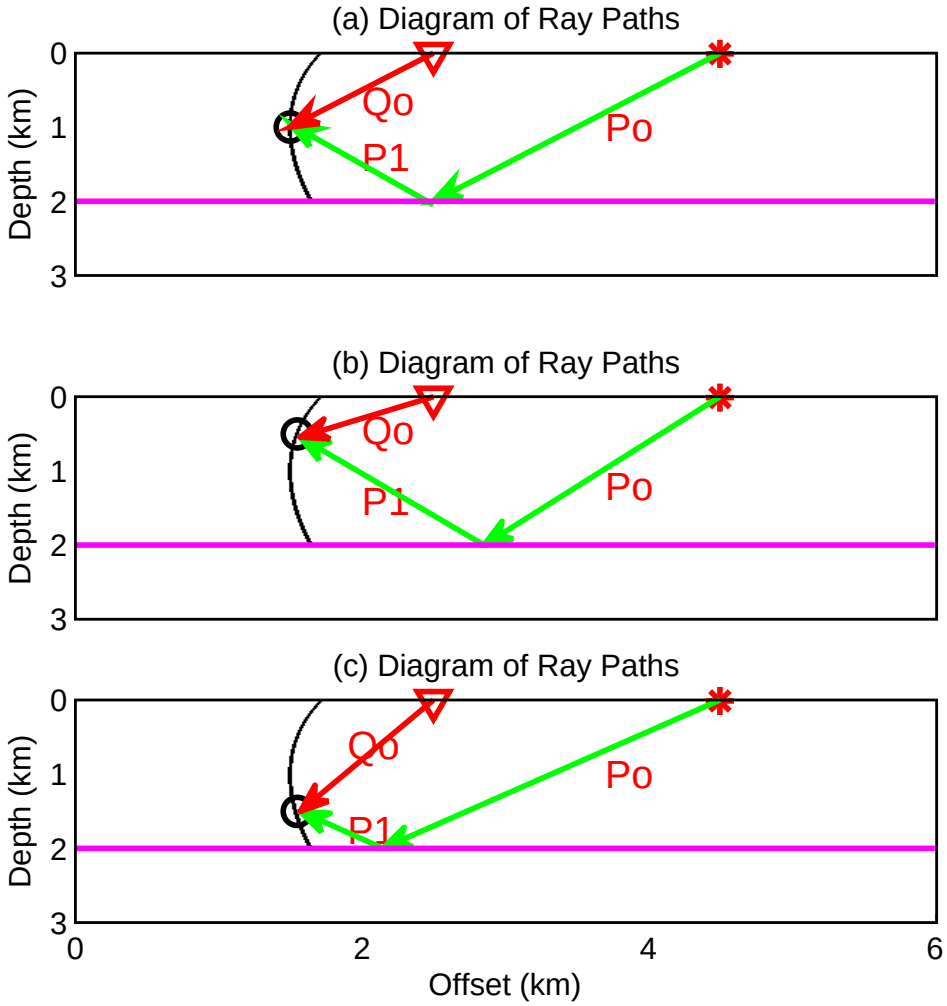

Figure 4.4. Diagrams of the ray paths illuminating the process of prism wave migration: (a) source and receiver wavefields correlate at the correct image point. Panels (b) and (c) show the ray paths to two image points that are above and below the right location. The black vertical curve plots part of the prism wave migration kernel. The circles along the curve show the locations of trial image points.

4.2.2 Geometric Interpretation of the Migration Kernel for Prism Wave

When the migration velocity is a homogeneous model with velocity $c = 2\ km/s$, the reverse time migration kernel plots as an ellipse for fixed source and receiver locations, where the ellipse is defined by the formula

$$\frac{|\mathbf{x}-\mathbf{x}_s|}{c}+\frac{|\mathbf{x}-\mathbf{x}_g|}{c}=\tau_{sg}. \tag{4.21}$$

Here τ_{sg} represents the travel time of a reflection arrival from the source at $\mathbf{x}_s$ to the receiver at $\mathbf{x}_g$. If a prism wave reflects off the horizontal reflector first and then reflects from the vertical reflector (Figure 4.1(a)), and the depth of the horizontal reflector is known as z_h, the migration traveltime equation corresponding to this prism wave can be defined as

$$\frac{|\mathbf{x}'-\mathbf{x}_s|}{c}+\frac{|\mathbf{x}-\mathbf{x}_g|}{c}=\tau'_{sg}. \tag{4.22}$$

In the above equation, τ'_{sg} is the travel time of the prism wave from $\mathbf{x}_s$ to $\mathbf{x}_g$. For $\mathbf{x}=(x,z)$ above the horizontal reflector, $\mathbf{x}'=(x,2z_h-z)$ is the mirror image of $\mathbf{x}$ with respect to the horizontal reflector. For any $\mathbf{x}$ below the horizontal reflector, according to Huygens principle, the Green's function $G_1(\mathbf{x}|\mathbf{x}_s)$ has an arrival time similar to that of the direct wave $G_o(\mathbf{x}|\mathbf{x}_s)$, with an additional amplification caused by ω^2 in equation 4.12. Therefore, below the horizontal reflector, the migration kernel plots as the ellipse in model space defined by

$$\frac{|\mathbf{x}-\mathbf{x}_s|}{c}+\frac{|\mathbf{x}-\mathbf{x}_g|}{c}=\tau'_{sg}. \tag{4.23}$$

This ellipse is an artifact and can be removed by up-down dip filtering applied to the traces associated with G_1 (Zhan and Schuster, 2012).

Figure 4.5(a) depicts the migration kernel corresponding to the ray path in Figure 4.1(a) and equation 4.13. Figure 4.5(b) plots the curves defined by equations 4.22 and 4.23, which are in excellent agreement with those associated with the migration kernel in Figure 4.5(a). In fact, equation 4.22 illustrates the basis of Kirchhoff migration of prism waves (Marmalyevskyy et al., 2005). Similarly, the migration kernel of equation 4.18 is plotted in Figure 4.6(b), which corresponds to the ray path in Figure 4.6(a). By symmetry considerations, it is obvious that a vertical reflector placed on the right side can also fit the observed prism wave. This kernel is plotted in Figure 4.6(c).

4.3 Numerical results

In this section, prism wave RTM is first tested with the simple L model in Figure 4.1(a), and then tested with a salt model to illustrate its advantages over standard RTM.

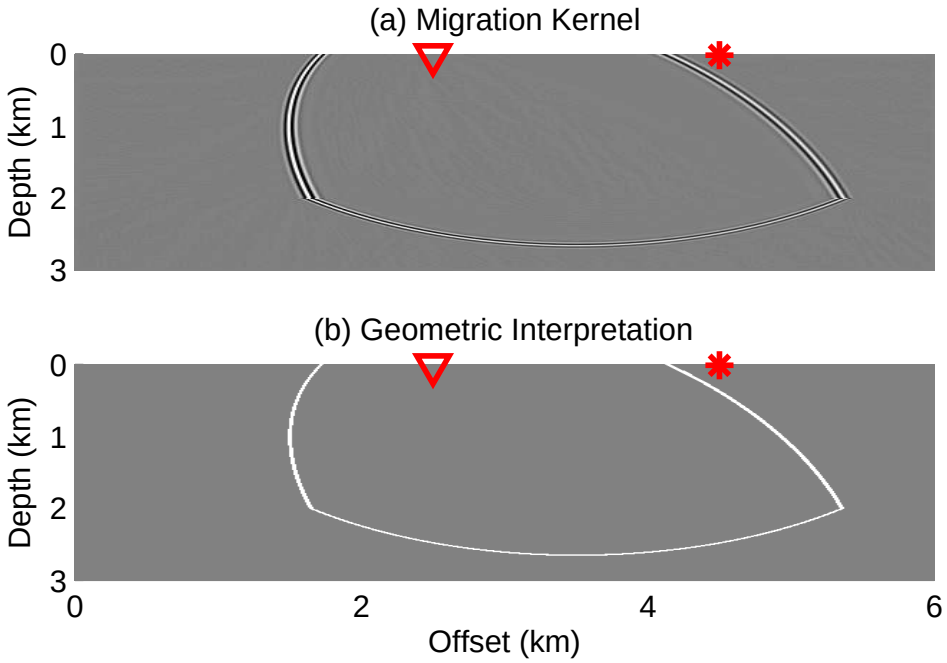

Figure 4.5. Migration kernels of prism waves: (a) the migration kernel of the prism wave corresponding to the term in equation 4.9 in the case the vertical reflector is on the left side. (b) The outline of the migration kernel in panel (a) according to the geometric interpretation. The star and triangle indicate the source and receiver locations respectively.

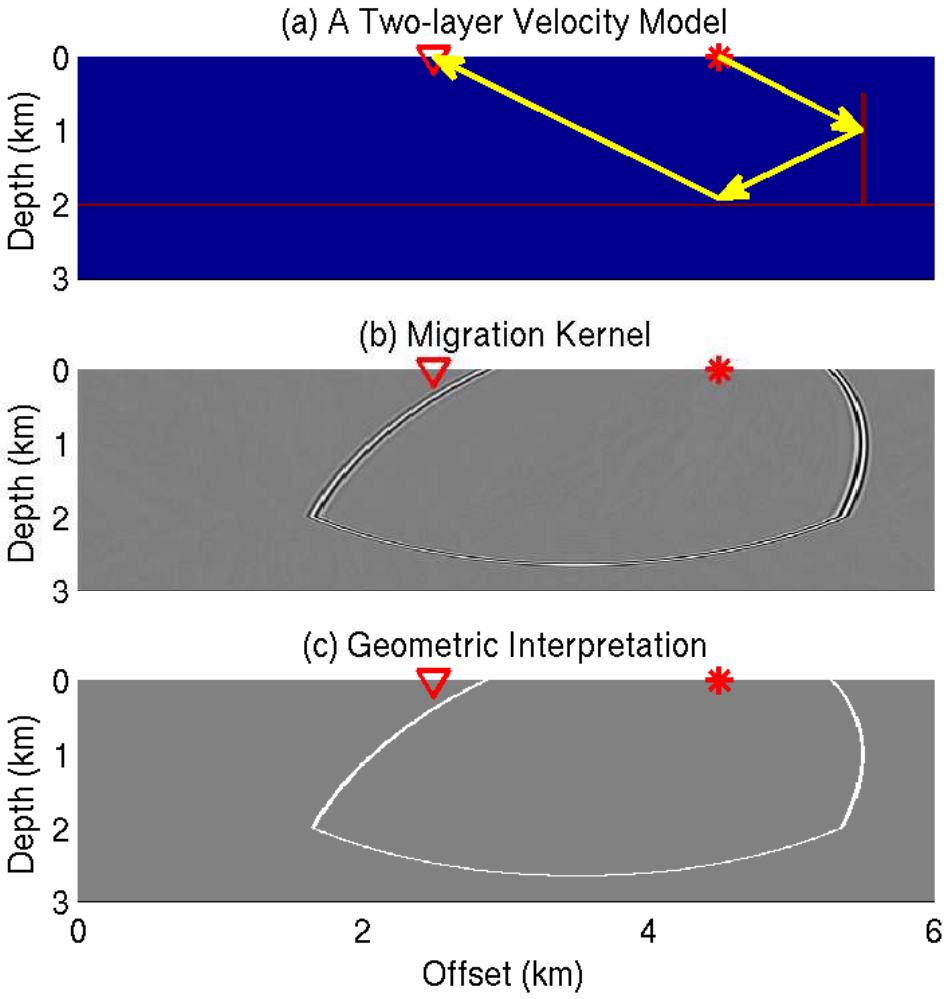

Figure 4.6. Migration kernels of prism waves: (a) the ray path for the prism wave with a vertical reflector on the right side; (b) the migration kernel of the prism wave corresponding to the term in equation 4.10 in the case the vertical reflector is on the right side; and (c) the outline of the migration kernel in panel (b) according to the geometric interpretation. The star and triangle indicate the source and receiver location respectively.

4.3.1 Simple L Model

The L model in Figure 4.1(a) is sampled with a grid size of 301×601 points, and a grid interval of 10 m. There are 31 shots evenly distributed along the x-axis at a 10 m depth with a 200 m shot interval. Every shot is recorded with the same 601 receivers at a 10 m depth with a 10 m receiver interval for a fixed spread acquisition geometry. Figure 4.7 plots a shot gather with the source at $x = 4.6\ km$, which clearly shows the direct waves, the diffractions from the top of the vertical reflector, the reflections off the horizontal reflector, and the prism waves are marked by a yellow arrow.

The 31 shot gathers are first migrated with a homogeneous velocity model ($2\ km/s$) by a conventional RTM method, and the image is shown in Figure 4.8(a), where only the horizontal reflector is visible. Then the proposed prism wave RTM algorithm is applied to the same 31 shot gathers with the same homogeneous velocity and the RTM image (Figure 4.8(a)) to give the image in Figure 4.8(b), which clearly depicts the vertical reflector. The horizontal reflector acts as the location of secondary sources during the migration process, which appears in the Figure 4.8(b) image.

4.3.2 Salt Model

Prism wave RTM can be used to delineate the vertical boundaries of a salt flank. In the velocity model shown in Figure 4.9(a), an irregular salt body is placed along the left boundary. The model size is 601×601 points with a 10 m grid interval. The seismic survey contains 301 shots fired at a depth of 10 m with an even x-sampling of 20 m. Every shot is recorded with 601 receivers at a 10 m depth and a 10 m receiver interval along the x-axis. In this case, the velocity gradient is not strong enough to generate diving waves for the short recording aperture of a 6 km long receiver array. Figure 4.10 shows a shot gather with the source position at $x = 4\ km$, where the prism waves are marked by the yellow arrows.

The 301 shot gathers are migrated with the smooth migration velocity in Figure 4.9(b) by a conventional RTM method, and the result is shown in Figure 4.11(a). This image clearly illuminates the subhorizontal reflectors, but only a few diffractors are visible along the salt flank. If the subhorizontal reflectors are picked from the RTM image and embedded in the velocity model (Figure 4.12(a)), the conventional RTM method can correctly migrate the prism waves to illuminate the steeply dipping salt flank shown in Figure 4.12(b). One problem is that the sharp boundaries in the velocity model cause the wavefield to be complex, e.g., internal multiples, and produce artifacts in the RTM image (Figure 4.12(b)). Another problem is that the subvertical reflectors are of weaker amplitudes compared to the horizontal ones.

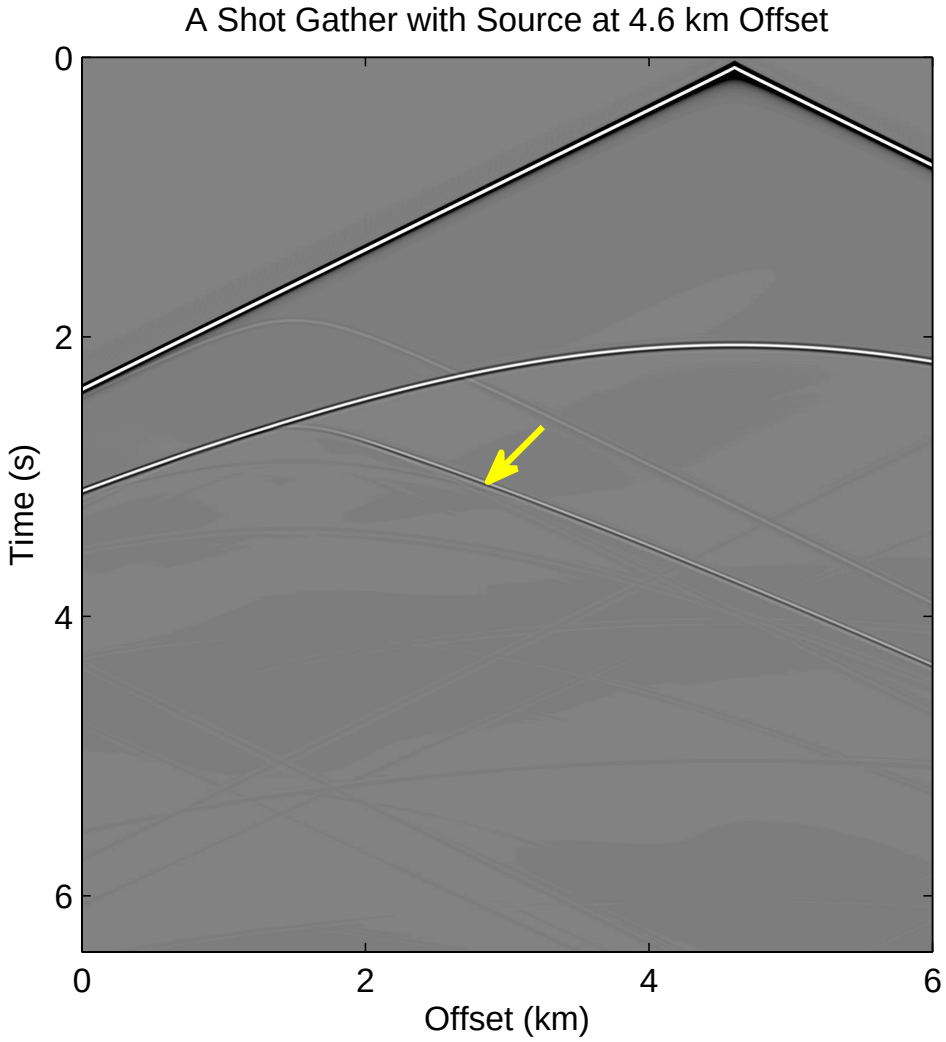

Figure 4.7. A shot gather with the source at $x = 4.6\ km$. The shot gather contains the direct wave, the reflection off the horizontal reflector, and the diffraction from the top of the vertical reflector. The yellow arrow points out the prism wave.

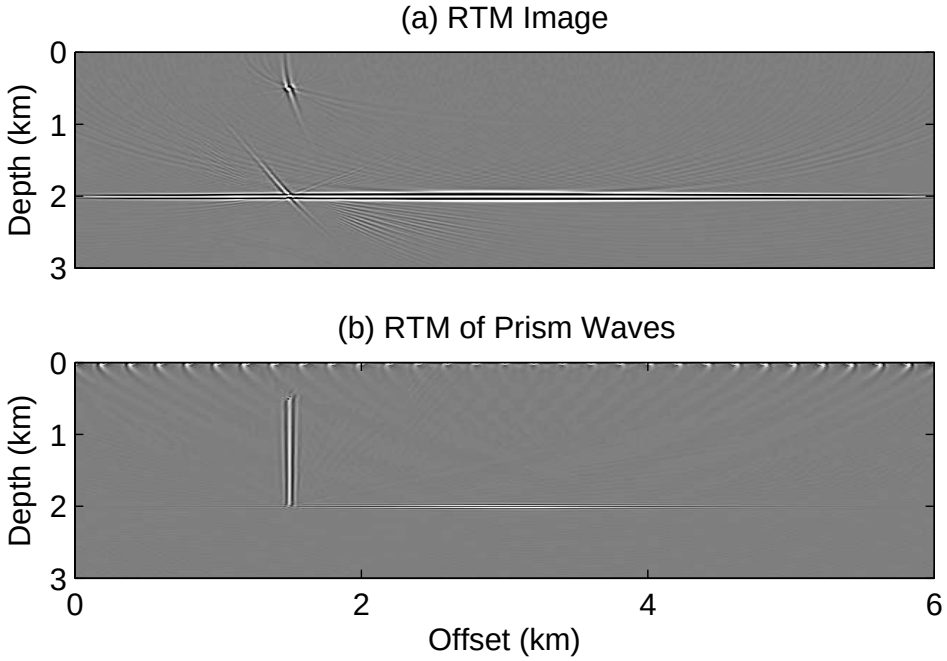

Figure 4.8. Comparison of migration images: (a) the RTM image obtained with a homogeneous velocity model. The vertical reflector is not illuminated. (b) The RTM image of the prism waves with homogeneous velocity and the reflectivity image in panel (a). The vertical reflector is well imaged.

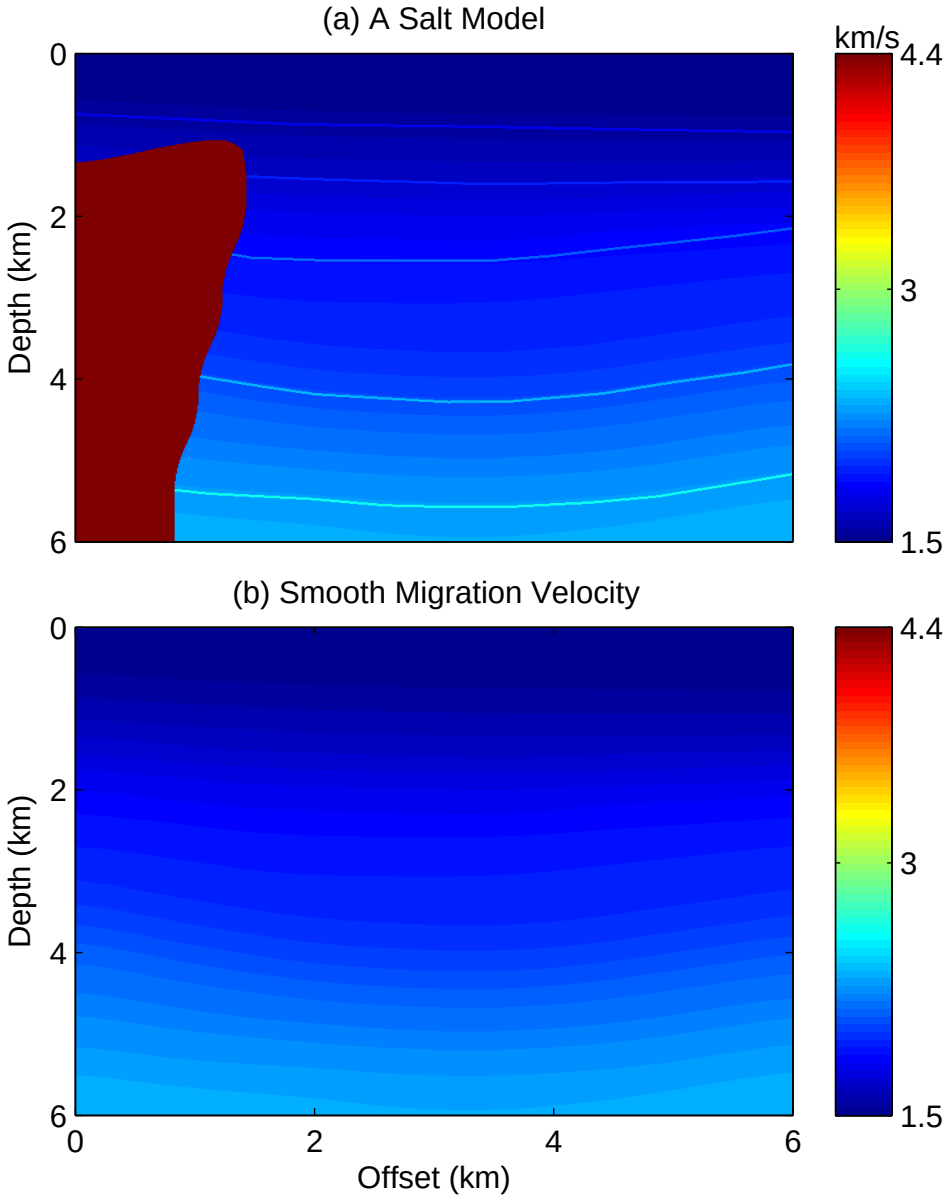

Figure 4.9. The velocity models: (a) a velocity model with a salt body on the left side; (b) the smooth migration velocity model without the salt body.

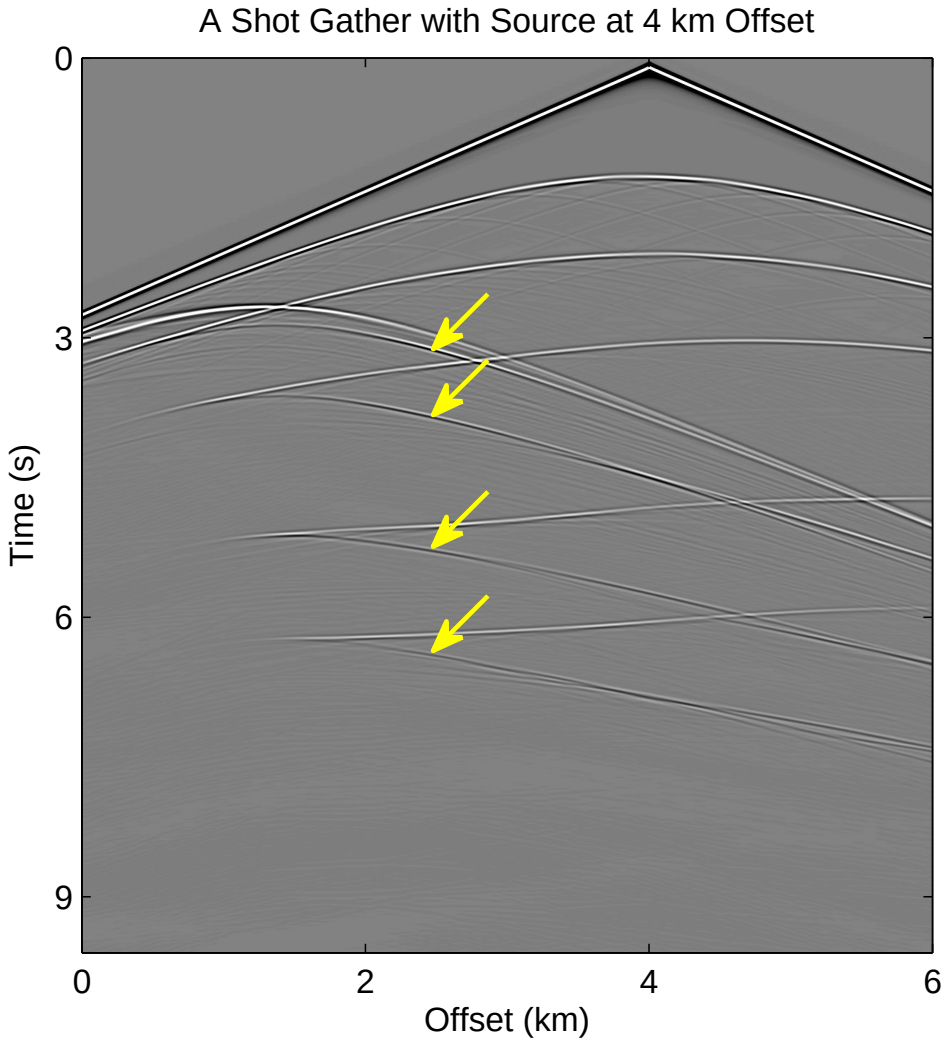

Figure 4.10. A shot gather with the source at $x = 4\ km$. The yellow arrows point out the prism waves.

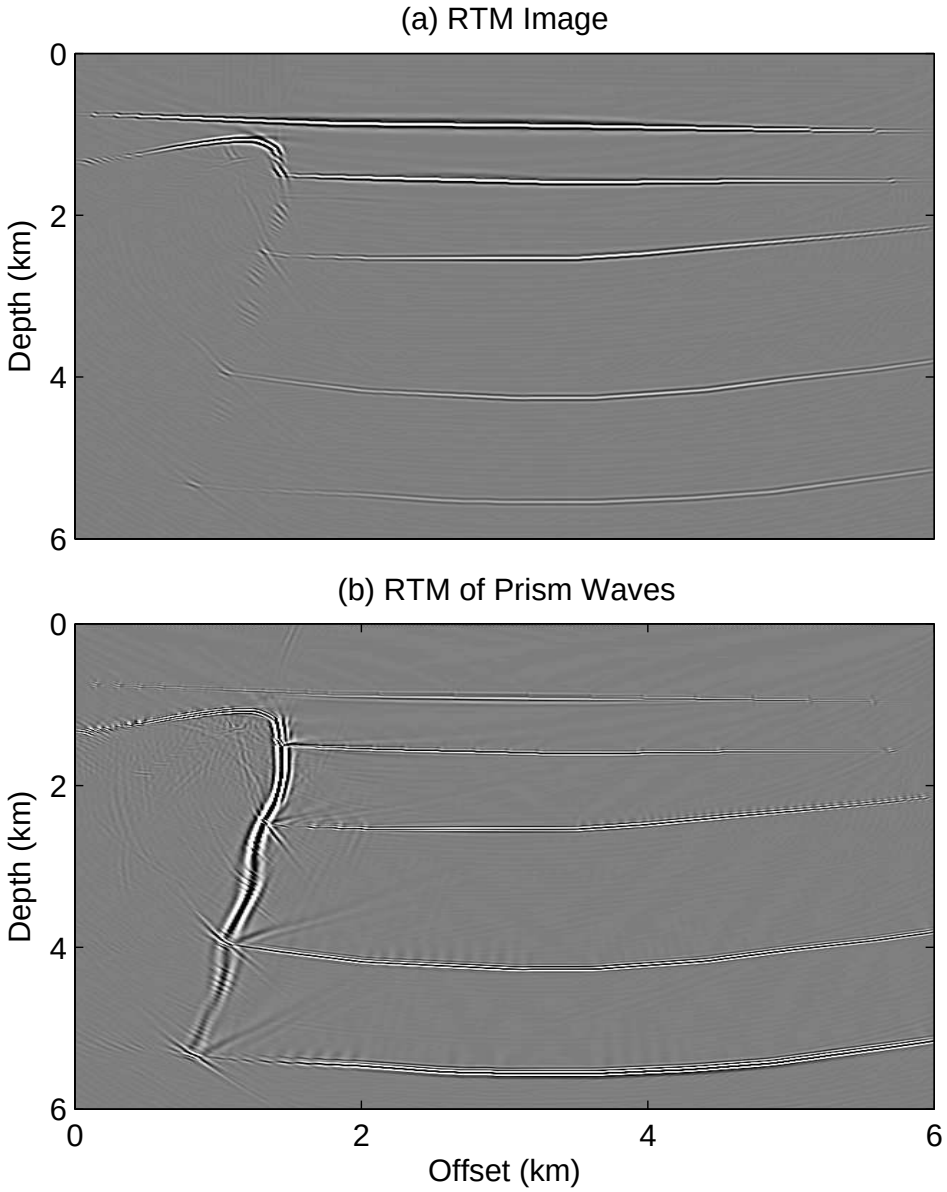

Figure 4.11. Comparison of migration images: (a) the RTM image obtained with the smooth migration velocity model. Along the salt boundary, only a few diffractors are visible. (b) The RTM image of the prism waves with the same velocity model. The irregular salt boundary is well imaged.

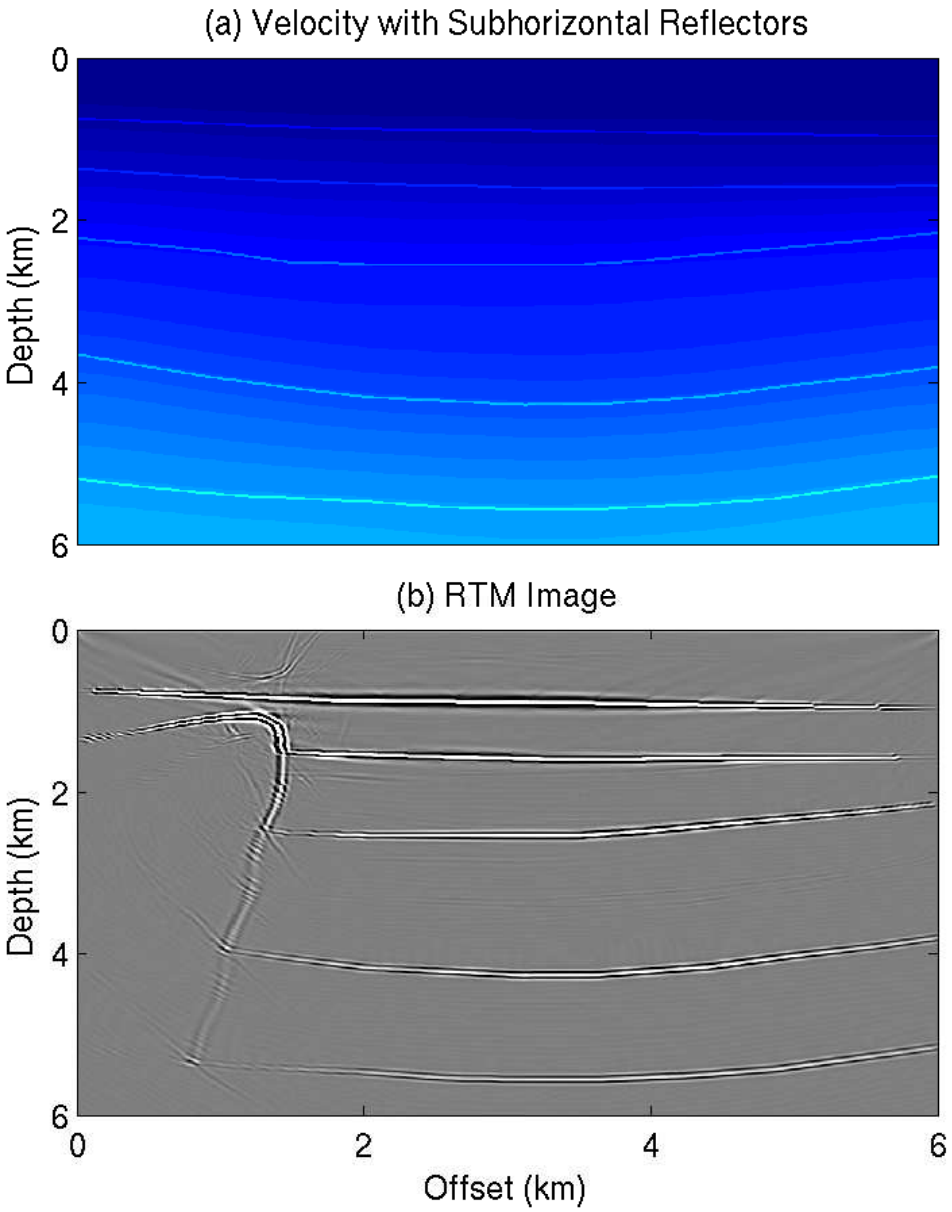

Figure 4.12. Conventional method results: (a) the velocity model with subhorizontal reflectors embedded; (b) the RTM image obtained with the velocity model in panel (a). The irregular salt boundary is well imaged.

The prism wave migration method uses the smooth migration velocity (Figure 4.9(b)) and the conventional RTM image (Figure 4.11(a)) to image the salt flank so that modification of the migration velocity is avoided. Figure 4.11(b) shows the prism wave migration image, where the salt flank is clearly imaged with strong amplitudes. However, this image contains some strong artifacts associated with those in Figure 4.11(a).

To further improve the image quality, I apply a dip filter to Figure 4.11(a) to keep only the subhorizontal reflectors, and the result is shown in Figure 4.13(a). Then, the proposed method is applied with the filtered image and the smooth velocity model to migrate the prism waves to produce the image in Figure 4.13(b), which contains fewer artifacts compared to Figure 4.11(b). Figure 4.14(a) shows the image in Figure 4.13(b) after dip filtering to keep only the subvertical reflectors. The final image is produced by summation of the migration images in Figures 4.13(a) and 4.14(a) to give Figure 4.14(b), which is the migration image with the best quality.

4.4 Discussion and Conclusion

In this chapter, I proposed a new method for migrating prism waves by RTM. There are two steps to the method: (1) Conventional RTM is applied to the data to estimate the geometry of the horizontal reflectors near the salt flank; (2) Prism wave RTM is applied to the data again, except the prism imaging condition is used rather than the conventional one. Dip filtering can also be applied to the images to reduce noise. For the simple L model, the vertical reflector is not visible in the conventional RTM image, but it is well imaged by migration of the prism waves with a homogeneous velocity model. In the example of the salt model, the salt flank can be imaged by embedding the horizontal reflectors in the velocity model, which is not trivial, but the best image is obtained by summation of two dip filtered partial migration images: one from conventional RTM and the other from the migration of the prism waves. The disadvantage of prism wave RTM is that its computational cost is twice that of conventional RTM. The empirical results suggest that the proposed method can migrate the prism waves correctly to delineate salt flanks and improve the image quality with the help of dip filtering.

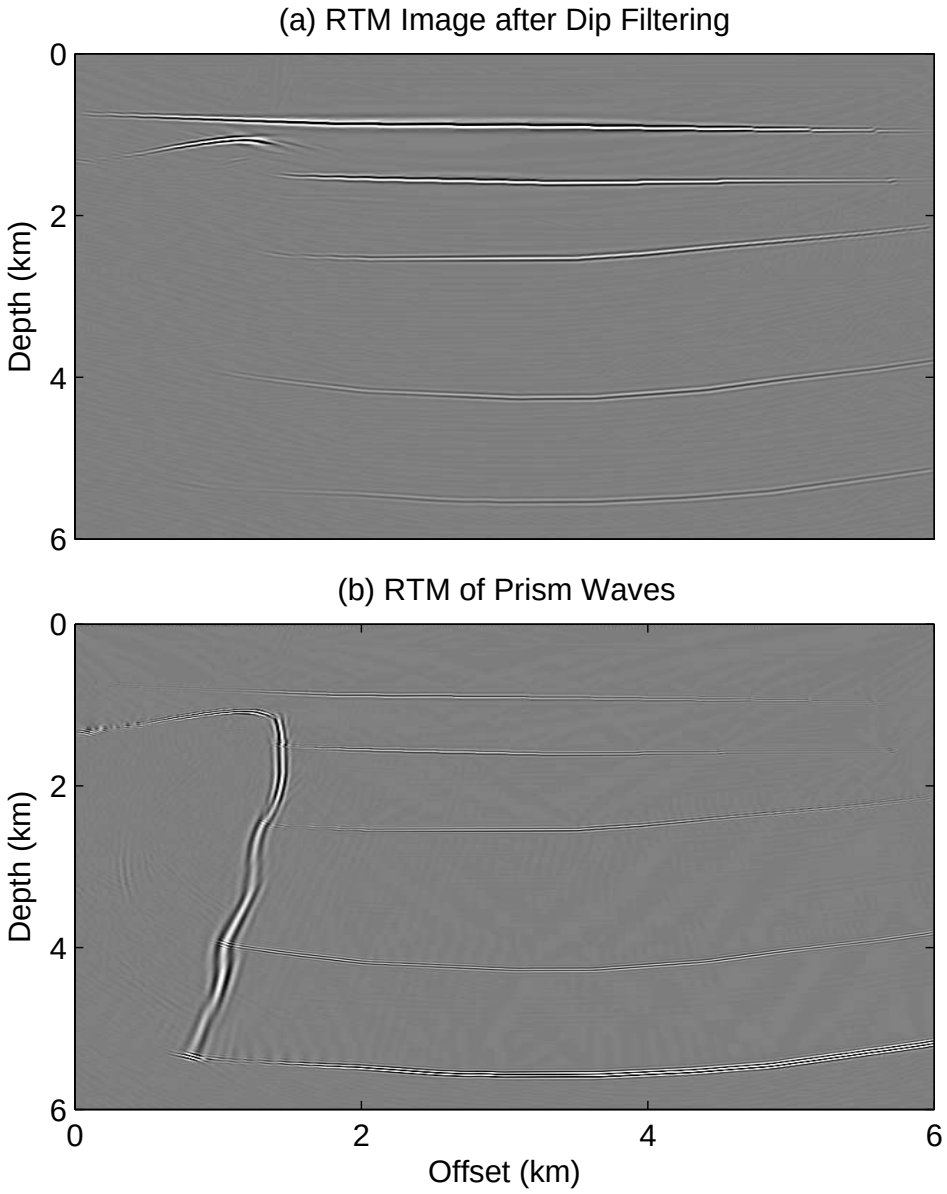

Figure 4.13. Migration image with dip filtering: (a) the RTM image obtained with the smooth migration velocity model after dip filtering to keep subhorizontal reflectors only; (b) the RTM image of the prism waves.

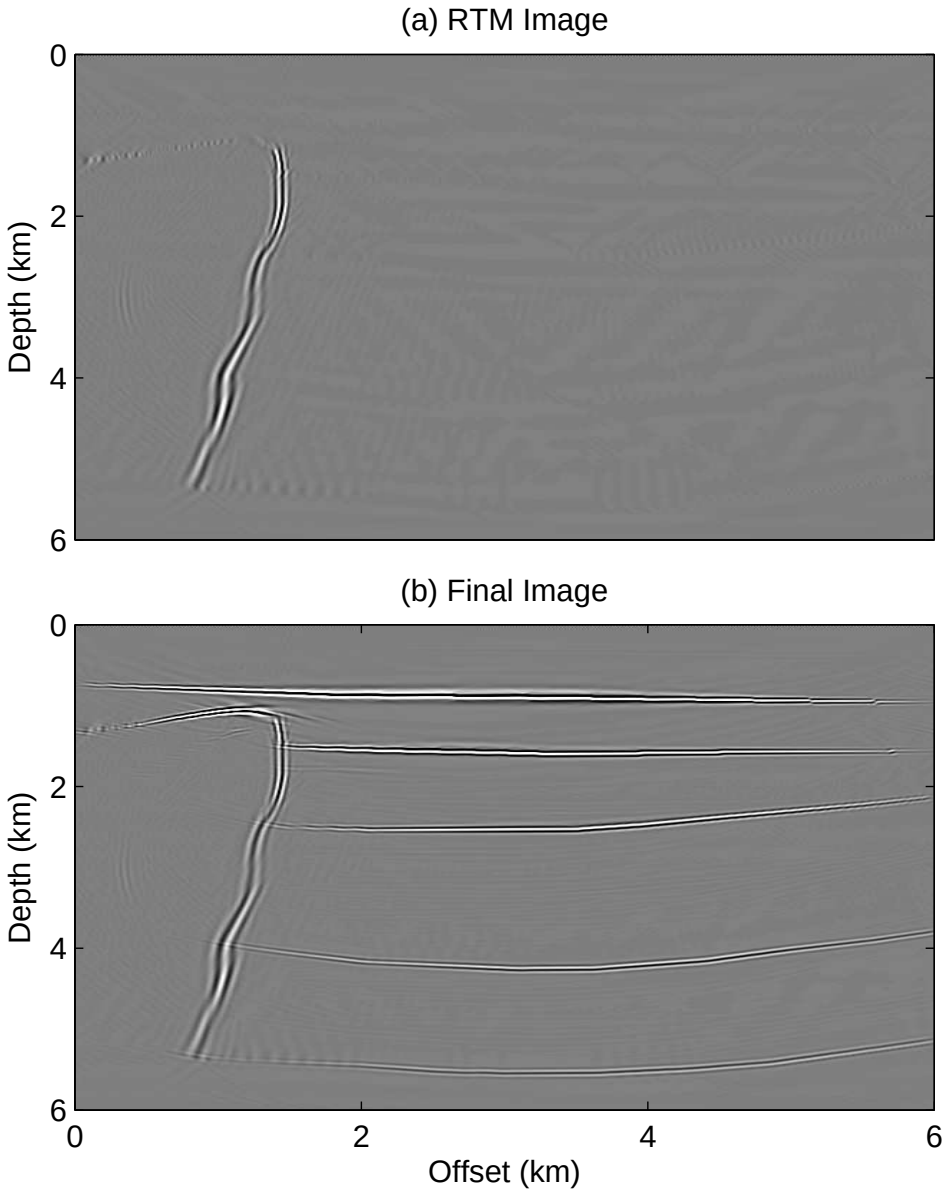

Figure 4.14. Vertial partial image and the final result: (a) the RTM image of the prism waves after dip filtering for subvertical reflectors only; (b) the sum of two partial images: one from conventional RTM and one from migration of the prism waves.

APPENDIX A

DEBLURRING FILTER

Following Aoki and Schuster (2009), I use a grid model with an even distribution of isolated point scatterers $\mathbf{m}_{ref}$ as my reference model. According to equation (6), I get

$$\mathbf{m}_{mig\_ref} = \mathbf{L}^T\mathbf{L}\mathbf{m}_{ref} = \mathbf{L}^T\mathbf{d}, \tag{A.1}$$

where $\mathbf{L}$ is the linear diffraction stack operator, which only depends on the background velocity $\mathbf{v}_o$ and the source receiver configurations. Here a column of the $\mathbf{L}^T\mathbf{L}$ matrix represents a migration Green's function (Schuster and Hu, 2000). Then, as shown in Figure A.1 I divide $\mathbf{m}_{ref}$ into somewhat large subsections centered around each point scatterer. In each subsection, I define a small-sized filter $\mathbf{f}_i$, such that

$$[\mathbf{m}_{mig\_ref}]_i * \mathbf{f}_i = [\mathbf{m}_{ref}]_i. \tag{A.2}$$

where i indicates the ith subsection and the notation $[\ ]_i$ denotes the model in the ith subsection. It is very important to choose a proper size for $[\mathbf{m}_{ref}]_i$ as it has to be big enough to cover the main part of the migration butterflies (Schuster and Hu, 2000). In each subsection, the reference model $[\mathbf{m}_{ref}]_i$ only contains a point scatterer. Thus, $[\mathbf{m}_{mig\_ref}]_i$ represents a migration Green's function, but truncated by the subsection and $\mathbf{f}_i$ is a local filter, which approximates the inverse of the Hessian within the subsection. After solving for $\mathbf{f}_i$ by a least-squares method, I apply $\mathbf{f}_i$ to the ith subsection of the original migration image obtained from the field data, and construct another image $\mathbf{m}_{mf}$. Near the boundaries between subsections, linear interpolation of nearby local filters is computed to make a smoothly varying image. This process can be expressed as

$$\mathbf{m}_{mf} = \mathbf{m}_{mig} * \mathbf{f}. \tag{A.3}$$

Here, $\mathbf{f}$ represents a bank of stationary filters (each filter is constant within its corresponding subsection). We can rewrite equation A.3 in matrix notation

$$\mathbf{m}_{mf} = \mathbf{F}\mathbf{m}_{mig}. \tag{A.4}$$

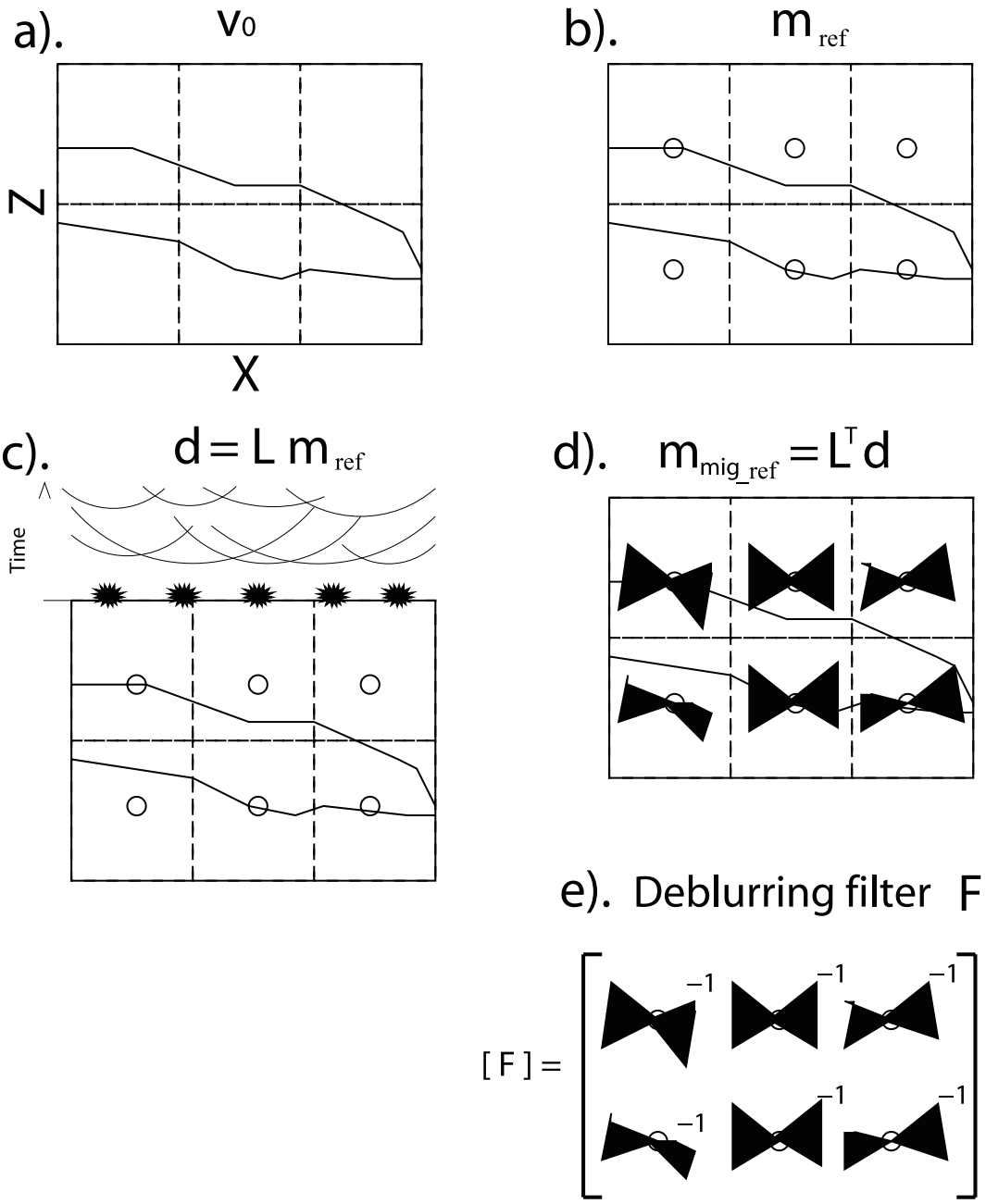

Figure A.1. Steps for computing the deblurring filter. Step (a) Define smooth velocity model with point scatterers denoted as circles in (b). Generate multisource data in (c), migrate the multisource data and get an image shown in (d). Step (e), in each subsection, compute a local filter according to $[\mathbf{m}_{mig\_ref}]_i * \mathbf{f}_i = [\mathbf{m}_{ref}]_i$ and combine all the local filters into the deblurring filter $\mathbf{F}$.

Since $\mathbf{m}_{mf}$ is an approximation of $\mathbf{m}$, and

$$\mathbf{m} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{m}_{mig}, \tag{A.5}$$

then the computed $\mathbf{f}_i$ in each subsection can be formed as the approximated preconditioner matrix

$$\mathbf{F} \approx (\mathbf{L}^T\mathbf{L})^{-1}. \tag{A.6}$$

We can improve the standard migration image by applying $\mathbf{F}$ to it, or, I can use $\mathbf{F}$ as a preconditioner in an iterative LSM solution to speed up convergence.

There are limitations associated with the deblurring filter.

1. The subsection needs to be big enough to cover the main part of migration artifacts. It also has to be large in order to avoid the interface between neighboring sections.

2. The migration Green's function is constant within a subsection, so that I can keep the filter constant with the subsection. To honor these two approximations, the velocity model needs to be smooth, so that the variation in the migration Green's function is smooth; hence, I usually use a high-frequency Ricker source wavelet, which makes the migration artifacts smaller.

APPENDIX B

SIGNAL-TO-NOISE RATIO

Consider an observed trace R_t, consisting of a signal trace S_t and zero-mean independent and identically-distributed[6] noise n_t of variance σ^2, as in

$$R_t = S_t + n_t, \quad t = 1, \ldots, T.$$

When M such observed traces are drawn and stacked, I get

$$\begin{aligned} \breve{R}_t &\triangleq \sum_{m=1}^{M} R_t^{(m)} \\ &= \sum_{m=1}^{M} [S_t + n_t^{(m)}] \\ &= MS_t + \sum_{m=1}^{M} n_t^{(m)}, \end{aligned} \tag{B.1}$$

where $R_t^{(m)}$ denotes the mth random realization of the signal trace S_t. ($n_t^{(m)}$'s are still i.i.d.) The signal and the noise part of the stacked trace $\breve{R}_t$ are denoted by

$$\breve{S}_t \triangleq MS_t \quad \text{and} \tag{B.2}$$

$$\breve{n}_t \triangleq \sum_{m=1}^{M} n_t^{(m)} \tag{B.3}$$

respectively. Note that the root mean squared (rms) amplitude of the stacked signal $\breve{S}_t$ is

$$\begin{aligned} A_M &\triangleq \sqrt{\sum_{t=1}^{T} \breve{S}_t^{\,2}/T} \\ &= M\sqrt{\sum_{t=1}^{T} S_t^{\,2}/T} \\ &= MA_1, \end{aligned} \tag{B.4}$$

[6] A sequence of random variables is independent and identically distributed (i.i.d.) if each random variable has the same probability distribution as the others and all are mutually independent.

where $A_1 = \sqrt{\sum_{t=1}^{T} S_t^2/T}$ is the rms amplitude of the signal trace S_t and the second equality follows from equation B.2; and A_M is defined as the rms amplitude of the M-fold stacked signal $\breve{S}_t$, growing in proportion to M, according to equation B.4. The rms amplitude of the stacked noise $\breve{n}_t$, σ_M, is defined as

$$\begin{aligned}
\sigma_M &\triangleq \sqrt{\sum_{t=1}^{T} < \breve{n}_t{}^2 > /T} \\
&= \sqrt{< \breve{n}_t{}^2 >} \\
&= \sqrt{< [\sum_{m=1}^{M} n_t^{(m)}]^2 >} \\
&= \sqrt{< \sum_{m=1}^{M} n_t^{(m)^2} >} \\
&= \sqrt{M}\sigma, \qquad \text{(B.5)}
\end{aligned}$$

where $<>$ denotes expectation, the second equality follows because n_t's are identically-distributed, the third equality follows from equation B.3, the fourth equality follows because $n_t^{(m)}$'s are zero-mean and independent, and the last equality follows because $n_t^{(m)}$'s are identically-distributed with variance σ^2. Equation B.5 shows that σ_M grows in proportion to $\sqrt{M}$.

Finally, The SNR of $\breve{R}_t$ is defined as the ratio of rms amplitude of signal over that of noise (Papoulis, 1991),

$$\begin{aligned}
\text{SNR} &\triangleq \frac{A_M}{\sigma_M} \\
&= \frac{MA_1}{\sqrt{M}\sigma} \\
&= \sqrt{M}A_1/\sigma, \qquad \text{(B.6)}
\end{aligned}$$

which exhibits a $\sqrt{M}$ enhancement.

APPENDIX C

LEAST-SQUARES MIGRATION WITH PRESTACK IMAGE

In least-squares migration, the goal is to solve the over-determined system of equations

$$\mathbf{d} = \mathbf{L}\mathbf{m}, \tag{C.1}$$

where $\mathbf{d}$ is the data vector, $\mathbf{L}$ matrix represents the forward modeling operator, and $\mathbf{m}$ is the model vector, and the corresponding normal equation is

$$\mathbf{L}^T\mathbf{d} = \mathbf{L}^T\mathbf{L}\mathbf{m}. \tag{C.2}$$

The direct least-squares solution is

$$\mathbf{m} = [\mathbf{L}^T\mathbf{L}]^{-1}\mathbf{L}^T\mathbf{d}. \tag{C.3}$$

Assuming a dataset with three shots, each of dimension $N_g \times N_t$, the total length of the data vector is $3N_g \times N_s$. If the model vector is of the size $N_x \times N_z$, the dimension of equation C.1 will be

$$[\mathbf{d}]^{3N_gN_s} = [\mathbf{L}]^{3N_gN_s \times N_xN_z}\,[\mathbf{m}]^{N_xN_z}\,. \tag{C.4}$$

For example, if the three shots are $\mathbf{d}_1$,$\mathbf{d}_2$, and $\mathbf{d}_3$, each with the length of $N_g \times N_t$, the above equation can be rewritten as

$$\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_2\\ \mathbf{d}_3\end{bmatrix} = \begin{bmatrix}\mathbf{L}_1\\ \mathbf{L}_2\\ \mathbf{L}_3\end{bmatrix}[\mathbf{m}]\,, \tag{C.5}$$

where $\mathbf{L}_1$,$\mathbf{L}_2$, and $\mathbf{L}_3$ are the forward modeling operator associated with each shot respectively. Here, $\mathbf{d}_i$ denotes the shot gather for the *ith* shot.

When the least-squares migration is performed with a stacked image as shown in equation C.5, the answer $\mathbf{m}$ in equation C.3 is the solution to the whole problem. In this dissertation, I propose to introduce an ensemble of prestack images

$$\underline{\mathbb{m}} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix} \tag{C.6}$$

into the inversion scheme, so that the system of equations becomes

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{d}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1\mathbf{m}_1 \\ \mathbf{L}_2\mathbf{m}_2 \\ \mathbf{L}_3\mathbf{m}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1 & & \\ & \mathbf{L}_2 & \\ & & \mathbf{L}_3 \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}, \tag{C.7}$$

where $\mathbf{m}_1$, $\mathbf{m}_2$, and $\mathbf{m}_3$ are the migration image associated for each shot respectively, each with the size of $N_x \times N_z$. The direct solution to the equation C.7 is

$$\begin{aligned} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix} &= \begin{bmatrix} \mathbf{L}_1^T\mathbf{L}_1 & & \\ & \mathbf{L}_2^T\mathbf{L}_2 & \\ & & \mathbf{L}_3^T\mathbf{L}_3 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{L}_1^T\mathbf{d}_1 \\ \mathbf{L}_2^T\mathbf{d}_2 \\ \mathbf{L}_3^T\mathbf{d}_3 \end{bmatrix} \\ = \begin{bmatrix} [\mathbf{L}_1^T\mathbf{L}_1]^{-1} & & \\ & [\mathbf{L}_2^T\mathbf{L}_2]^{-1} & \\ & & [\mathbf{L}_3^T\mathbf{L}_3]^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{L}_1^T\mathbf{d}_1 \\ \mathbf{L}_2^T\mathbf{d}_2 \\ \mathbf{L}_3^T\mathbf{d}_3 \end{bmatrix} &= \begin{bmatrix} [\mathbf{L}_1^T\mathbf{L}_1]^{-1}\mathbf{L}_1^T\mathbf{d}_1 \\ [\mathbf{L}_2^T\mathbf{L}_2]^{-1}\mathbf{L}_2^T\mathbf{d}_2 \\ [\mathbf{L}_3^T\mathbf{L}_3]^{-1}\mathbf{L}_3^T\mathbf{d}_3 \end{bmatrix}. \end{aligned} \tag{C.8}$$

It is clear that the solution $\mathbf{m}_1$, $\mathbf{m}_2$, and $\mathbf{m}_3$ are independent of each other. By introducing the prestack image into the inversion, I solve three small problems instead of one big problem, thus make it possible to find stable solution when the equations are not consistent with each other in the case of wrong migration velocity.

APPENDIX D

MATHEMATICAL DERIVATION WITH ADJOINT STATE METHOD

In Chapter 4, the physical meaning of prism wave migration was explained with a simple geometrical interpretation. From the mathematical point of view, the migration of prism waves can be thought of as the adjoint operation of modeling a prism wave. To show this, I will derive the forward modeling operator of a prism wave and apply the adjoint state method to derive its corresponding migration operator. Given a background slowness model $s_o(\mathbf{x})$ and a reflectivity model $m_1(\mathbf{x})$, the reflection data for a shot at $\mathbf{x}_s$ can be modeled with the Born approximation using the following equations (Dai et al., 2012)

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x}))P_o(\mathbf{x}) = W(\omega)\delta(\mathbf{x}-\mathbf{x}_s), \tag{D.1}$$

$$(\bigtriangledown^2 + \omega^2 s_o^2(\mathbf{x}))P_1(\mathbf{x}) = \omega^2 m_1(\mathbf{x})P_o(\mathbf{x}). \tag{D.2}$$

By introducing a perturbation to the slowness model $s_o \rightarrow s_o + \delta s$, the wavefields become $P_o \rightarrow P_o + \delta P_o$, $P_1 \rightarrow P_1 + \delta P_1$. Expanding the slowness term as

$$(s_o + \delta s)^2 \approx s_o^2 + 2 s_o \delta s, \tag{D.3}$$

equations D.1 and D.2 become

$$(\bigtriangledown^2 + \omega^2 s_o^2 + 2\omega^2 s_o \delta s)(P_o(\mathbf{x}) + \delta P_o(\mathbf{x})) = W(\omega)\delta(\mathbf{x}-\mathbf{x}_s), \tag{D.4}$$

$$(\bigtriangledown^2 + \omega^2 s_o^2 + 2\omega^2 s_o \delta s)(P_1(\mathbf{x}) + \delta P_1(\mathbf{x})) = \omega^2 m_1(\mathbf{x})(P_o(\mathbf{x}) + \delta P_o(\mathbf{x})). \tag{D.5}$$

Assuming $m_2(\mathbf{x}) = -2 s_o(\mathbf{x})\delta s(\mathbf{x})$, and subtracting equation D.1 from equation D.4, I get

$$(\bigtriangledown^2 + \omega^2 s_o^2)\delta P_o(\mathbf{x}) = \omega^2 m_2(\mathbf{x}) P_o(\mathbf{x}). \tag{D.6}$$

Similarly, subtracting equation D.2 from equation D.5, I get

$$(\bigtriangledown^2 + \omega^2 s_o^2)\delta P_1(\mathbf{x}) = \omega^2 m_1(\mathbf{x})\delta P_o(\mathbf{x}) + \omega^2 m_2(\mathbf{x}) P_1(\mathbf{x}), \tag{D.7}$$

where the higher order terms are neglected. Equation D.7 represents the modeling operator for the prism wave $\delta P_1(\mathbf{x})$, which requires solving equations D.1, D.2, and D.6. Calculation of the prism wave needs four finite-difference simulations.

The above equations can be expressed with Green's functions G_o calculated with the slowness s_o, so equation D.2 becomes

$$P_1(\mathbf{x}|\mathbf{x}_s) = \int_{\mathbf{x}'} \omega^2 W(\omega) G_o(\mathbf{x}'|\mathbf{x}_s) m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}) d\mathbf{x}', \tag{D.8}$$

and equation D.6 becomes

$$\delta P_o(\mathbf{x}|\mathbf{x}_s) = \int_{\mathbf{x}''} \omega^2 W(\omega) G_o(\mathbf{x}''|\mathbf{x}_s) m_2(\mathbf{x}'') G_o(\mathbf{x}''|\mathbf{x}) d\mathbf{x}'', \tag{D.9}$$

where $\mathbf{x}'$ and $\mathbf{x}''$ are dummy variables. Thus, the modeling operator of the doubly scattered prism wave can be expressed as

$$\begin{aligned}
\delta P_1(\mathbf{x}|\mathbf{x}_s) &= \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_1(\mathbf{x}''') \delta P_o(\mathbf{x}'''|\mathbf{x}_s) d\mathbf{x}''' \\
+ & \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_2(\mathbf{x}''') P_1(\mathbf{x}'''|\mathbf{x}_s) d\mathbf{x}''' \\
= & \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_1(\mathbf{x}''') \int_{\mathbf{x}''} \omega^2 W(\omega) G_o(\mathbf{x}''|\mathbf{x}_s) m_2(\mathbf{x}'') G_o(\mathbf{x}''|\mathbf{x}''') d\mathbf{x}'' d\mathbf{x}''' \\
+ & \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_2(\mathbf{x}''') \int_{\mathbf{x}'} \omega^2 W(\omega) G_o(\mathbf{x}'|\mathbf{x}_s) m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}''') d\mathbf{x}' d\mathbf{x}'''
\end{aligned} \tag{D.10}$$

If I switch the order of integration for the first term, the above equation becomes

$$\begin{aligned}
& \delta P_1(\mathbf{x}|\mathbf{x}_s) \\
= & \int_{\mathbf{x}''} \omega^2 W(\omega) G_o(\mathbf{x}''|s) m_2(\mathbf{x}'') \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_1(\mathbf{x}''') G_o(\mathbf{x}''|\mathbf{x}''') d\mathbf{x}''' d\mathbf{x}'' \\
+ & \int_{\mathbf{x}'''} \omega^2 W(\omega) G_o(\mathbf{x}'''|\mathbf{x}) m_2(\mathbf{x}''') \int_{\mathbf{x}'} \omega^2 G_o(\mathbf{x}'|s) m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}''') d\mathbf{x}' d\mathbf{x}''' \\
= & \int_{\mathbf{x}''} \omega^2 W(\omega) G_o(\mathbf{x}''|\mathbf{x}_s) m_2(\mathbf{x}'') G_1(\mathbf{x}|\mathbf{x}'') d\mathbf{x}'' \\
+ & \int_{\mathbf{x}'''} \omega^2 W(\omega) G_o(\mathbf{x}'''|\mathbf{x}) m_2(\mathbf{x}''') G_1(\mathbf{x}'''|s) d\mathbf{x}''',
\end{aligned} \tag{D.11}$$

where G_1 represents the Green's function for reflection wave:

$$G_1(\mathbf{x}|\mathbf{x}'') = \int_{\mathbf{x}'''} \omega^2 G_o(\mathbf{x}'''|\mathbf{x}) m_1(\mathbf{x}''') G_o(\mathbf{x}''|\mathbf{x}''') d\mathbf{x}''', \tag{D.12}$$

$$G_1(\mathbf{x}'''|\mathbf{x}_s) = \int_{\mathbf{x}'} \omega^2 G_o(\mathbf{x}'|\mathbf{x}_s) m_1(\mathbf{x}') G_o(\mathbf{x}'|\mathbf{x}''') d\mathbf{x}'. \tag{D.13}$$

When the wavefield δP_1 is recorded at the receiver location $\mathbf{x}_g$, the shot gather $d_2(\mathbf{x}_g|\mathbf{x}_s)$ of the prism wave can be expressed as

$$\begin{aligned}
d_2(\mathbf{x}_g|\mathbf{x}_s) &= \int_{\mathbf{x}} \omega^2 W(\omega) G_o(\mathbf{x}|\mathbf{x}_s) m_2(\mathbf{x}) G_1(\mathbf{x}|\mathbf{x}_g) d\mathbf{x} \\
&+ \int_{\mathbf{x}} \omega^2 W(\omega) G_1(\mathbf{x}|\mathbf{x}_s) m_2(\mathbf{x}) G_1(\mathbf{x}|\mathbf{x}_g) d\mathbf{x},
\end{aligned} \tag{D.14}$$

Equation D.14 is the forward modeling operator for the prism wave. By simply applying adjoint of the forward modeling (Plessix, 2006), the migration image of the shot gather $d_2(\mathbf{x}_g|\mathbf{x}_s)$ can be shown to be

$$\begin{aligned} m_{mig}(\mathbf{x}|\mathbf{x}_s) &= \sum_{\omega}\sum_{g} \omega^2 W^*(\omega) G_o^*(\mathbf{x}|\mathbf{x}_s) G_1^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s) \\ &+ \sum_{\omega}\sum_{g} \omega^2 W^*(\omega) G_1^*(\mathbf{x}|\mathbf{x}_s) G_o^*(\mathbf{x}|\mathbf{x}_g) d_2(\mathbf{x}_g|\mathbf{x}_s), \end{aligned} \tag{D.15}$$

which are exactly the terms in equations 4.9 and 4.10. The computation of these terms is described in the text.

REFERENCES

[1] Naoshi Aoki and Gerard T. Schuster. Fast least-squares migration with a deblurring filter. *Geophysics*, 74(6):WCA83–WCA93, 2009.

[2] Edip Baysal, Dan D. Kosloff, and John W. C. Sherwood. Reverse time migration. *Geophysics*, 48(11):1514–1524, 1983.

[3] Craig J. Beasley. A new look at marine simultaneous sources. *The Leading Edge*, 27(7):914–917, 2008.

[4] A. J. Guus Berkhout. Changing the mindset in seismic data acquisition. *The Leading Edge*, 27(7):924–938, 2008.

[5] Chaiwoot Boonyasiriwat and Gerard T. Schuster. 3d multisource full-waveform inverions using quasi-Monte Carlo phase encoding. *SEG Technical Program Expanded Abstracts, submitted*, 2010.

[6] Chaiwoot Boonyasiriwat, Gerard T. Schuster, Paul Valasek, and Weiping Cao. Applications of multiscale waveform inversion to marine data using a flooding technique and dynamic early-arrival windows. *Geophysics*, 75(6):R129–R136, 2010.

[7] Maud Cavalca and Patrick Lailly. Prismatic reflections for the delineation of salt bodies. *SEG Technical Program Expanded Abstracts*, pages 2350–2354, 2005.

[8] Jon F. Claerbout. Toward a unified theory of reflector mapping. *Geophysics*, 36(3):467–481, 1971.

[9] S. Cole and M. Karrenbach. Least squares kirchhoff migration. *SEP-75*, pages 101–110, 1992.

[10] Wei Dai, C. Boonyasiriwat, and Gerard T. Schuster. 3d multi-source least-squares reverse time migration. *SEG Technical Program Expanded Abstracts*, 29(1):3120–3124, 2010.

[11] Wei Dai, Paul Fowler, and Gerard T. Schuster. Multisource least-squares reverse time migration. *Geophysical Prospecting*, 60:681–695, 2012.

[12] Wei Dai and Gerard T. Schuster. Least-squares migration of simultaneous sources data with a deblurring filter. *SEG Technical Program Expanded Abstracts*, 28(1):2990–2994, 2009.

[13] Wei Dai and Gerard T. Schuster. Multi-source wave equation least-squares migration with a deblurring filter. *EAGE Technical Program Expanded Abstracts*, page 276, 2010.

[14] Wei Dai, Xin Wang, and Gerard T. Schuster. Least-squares migration of multisource data with a deblurring filter. *Geophysics*, 76(5):R135–R146, 2011.

[15] Wei Dai, Ge Zhan, Xin Wang, and Gerard T. Schuster. Multi-source least squares migration, Waveform inversion, and MVA analysis. *SEG Invited Workshop Talk*, 2009.

[16] Bertrand Duquet and Patrick Lailly. Efficient 3d wave-equation migration using virtual planar sources. *Geophysics*, 71(5):S185–S197, 2006.

[17] Bertrand Duquet, Kurt J. Marfurt, and Joe A. Dellinger. Kirchhoff modeling, inversion for reflectivity, and subsurface illumination. *Geophysics*, 65(4):1195–1209, 2000.

[18] Eivind Fromyr, Guillaume Cambois, Ruth Loyd, and Jack Kinkead. Flam — a simultaneous source wide azimuth test. *SEG Technical Program Expanded Abstracts*, 27(1):2821–2825, 2008.

[19] Antoine Guitton. Amplitude and kinematic corrections of migrated images for nonunitary imaging operators. *Geophysics*, 69(4):1017–1024, 2004.

[20] Dave Hale, N. Ross Hill, and Joe Stefani. imaging salt with turning seismic waves. *Geophysics*, 57(11):1453–1462, 1992.

[21] Gary Hampson, Joe Stefani, and Fred Herkenhoff. Acquisition using simultaneous sources. *The Leading Edge*, 27(7):918–923, 2008.

[22] Jianxing Hu and Gerard T. Schuster. Prestack migration deconvolution. *SEG Technical Program Expanded Abstracts*, 19(1):984–987, 2000.

[23] Yunsong Huang and Gerard T. Schuster. Multisource least-squares migration of marine streamer data and land data with frequency-division encoding. *Geophysical Prospecting*, 60(4):663–680, 2012.

[24] I. F. Jones, M. C. Goodwin, I. D. Berranger, H. Zhou, and P. A. Farmer. Application of anisotropic 3D reverse time migration to complex North Sea imaging. *SEG Technical Program Expanded Abstracts*, pages 2140–2144, 2007.

[25] Sam T. Kaplan, Partha S. Routh, and Mauricio D. Sacchi. Derivation of forward and adjoint operators for least-squares shot-profile split-step migration. *Geophysics*, 75(6):S225–S235, 2010.

[26] Dan D. Kosloff and Edip Baysal. Forward modeling by a Fourier method. *Geophysics*, 47(10):1402–1412, 1982.

[27] Jerome R. Krebs, John E. Anderson, David Hinkley, Ramesh Neelamani, Sunwoong Lee, Anatoly Baumstein, and Martin-Daniel Lacasse. Fast full-wavefield seismic inversion using encoded sources. *Geophysics*, 74(6):WCC177–WCC188, 2009.

[28] Henning Kühl and Mauricio D. Sacchi. Least-squares wave-equation migration for AVP/AVA inversion. *Geophysics*, 68(1):262–273, 2003.

[29] P. Lailly. The seismic inverse problem as a sequence of before stack migrations. *Conference on Inverse Scattering*, 1984.

[30] Faqi Liu, Douglas W. Hanson, Norman D. Whitmore, Richard S. Day, and Robert H. Stolt. Toward a unified analysis for source plane-wave migration. *Geophysics*, 71(4):S129–S139, 2006.

[31] Faqi Liu, Guanquan Zhang, Scott A. Morton, and Jacques P. Leveille. An effective imaging condition for reverse-time migration using wavefield decomposition. *Geophysics*, 76(1):S29–S39, 2011.

[32] Y. Luo and G. T. Schuster. Wave-equation traveltime inversion. *Geophysics*, 56(5):645–653, 1991.

[33] Walt Lynn, Mark Doyle, Ken Larner, and Richard Marschall. Experimental investigation of interference from other seismic crews. *Geophysics*, 52(11):1501–1524, 1987.

[34] Alison E. Malcolm, Maarten V. de Hoop, and Bjorn Ursin. Recursive imaging with multiply scattered waves using partial image regularization: A North Sea cast study. *Geophysics*, 76(2):B33–B42, 2011.

[35] Alison E. Malcolm, Bjorn Ursin, and Maarten V. de Hoop. Seismic imaging and illumination with internal multiples. *Geophysical Journal International*, 176:847–864, 2009.

[36] Naum Marmalyevskyy, Yury Roganov, Zinovy Gornyak, Alex Kostyukevych, and Viktor Mershchiy. Migration of duplex waves. *SEG Technical Program Expanded Abstracts*, pages 2025–2029, 2005.

[37] G. A. McMechan. Migration by extrapolation of time-dependent boundary values. *Geophysical Prospecting*, 31:413–420, 1983.

[38] W. A. Mulder and R.-E. Plessix. A comparison between one-way and two-way wave-equation migration. *Geophysics*, 69(6):1491–1504, 2004.

[39] Tamas Nemeth, Chengjun Wu, and Gerard T. Schuster. Least-squares migration of incomplete reflection data. *Geophysics*, 64(1):208–221, 1999.

[40] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1991.

[41] R.-E. Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(3):495–503, 2006.

[42] R.-E. Plessix and W. A. Mulder. Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157(3):975–987, 2004.

[43] D. W. Ratcliff, S. H. Gray, and N. D. Whitmore. Seismic imaging of salt structures in the Gulf of Mexico. *SEG Technical Program Expanded Abstracts*, pages 1164–1165, 1991.

[44] D. W. Ratcliff, S. H. Gray, and N. D. Whitmore. Seismic imaging of salt structures in the gulf of mexico. *The Leading Edge*, 11(4):15–31, 1992.

[45] Louis A. Romero, Dennis C. Ghiglia, Curtis C. Ober, and Scott A. Morton. Phase encoding of shot records in prestack migration. *Geophysics*, 65(2):426–436, 2000.

[46] P.S. Routh, J.R. Krebs, S. Lazaratos, A.I. Baumstein, I. Chikichev, S. Lee, N. Downey, D. Hinkley, and J.E. Anderson. Full-wavefield inversion of marine streamer data with the encoded simultaneous source method. *EAGE Technical Program Expanded Abstracts*, 2011.

[47] Gerard T. Schuster. Least-squares cross-well migration. *SEG Technical Program Expanded Abstracts*, 12(1):110–113, 1993.

[48] Gerard T. Schuster and Jianxing Hu. Green's function for migration: Continuous recording geometry. *Geophysics*, 65(1):167–175, 2000.

[49] Gerard T. Schuster, Xin Wang, Yunsong Huang, Wei Dai, and Chaiwoot Boonyasiriwat. Theory of multisource crosstalk reduction by phase-encoded statics. *Geophysical Journal International*, 184:1289–1303, 2011.

[50] Robert H. Stolt and Alvin K. Benson. *Seismic migration: Theory and practice*. Geophysical press, 1986.

[51] W. W. Symes and J. J. Carazzone. Velocity inversion by differential semblance optimization. *Geophysics*, 56(5):654–663, 1991.

[52] Yaxun Tang and Biondo Biondi. Least-squares migration/inversion of blended data. *SEG Technical Program Expanded Abstracts*, 28(1):2859–2863, 2009.

[53] A. Tikhonov and V. Arsenin. *Solutions of ill-posed problems*. V.H. Winston and Sons, 1977.

[54] Denes Vigh and E. William Starr. 3D prestack plane-wave, full-waveform inversion. *Geophysics*, 73(5):VE135–VE144, 2008.

[55] N. D. Whitmore. Iterative depth migration by backward time propagation. *SEG Technical Program Expanded Abstracts*, 2(1):382–385, 1983.

[56] Mandy Wong, Shuki Ronen, and Biondo Biondi. Least-squares reverse time migration/inversion for ocean bottom data: A case study. *SEG Technical Program Expanded Abstracts*, 30(1):2369–2373, 2011.

[57] Jianhua Yu, Jianxing Hu, Gerard T. Schuster, and Robert Estill. Prestack migration deconvolution. *Geophysics*, 71(2):S53–S62, 2006.

[58] Ge Zhan and Gerard T. Schuster. Mitigation of artifacts in rtm with migration kernel decomposition. *EAGE Technical Program Expanded Abstracts*, page X048, 2012.

[59] Yu Zhang, James Sun, Carl Notfors, Samuel H. Gray, Leon Chernis, and Jerry Young. Delayed-shot 3d depth migration. *Geophysics*, 70(5):E21–E28, 2005.

[60] Changxi Zhou, Gerard T. Schuster, Sia Hassanzadeh, and Jerry M. Harris. Elastic wave equation traveltime and waveform inversion of crosswell data. *Geophysics*, 62(3):853–868, 1997.