IOP Publishing Journal of Geophysics and Engineering
J. Geophys. Eng. **13** (2016) 207–219 doi:10.1088/1742-2132/13/3/207

Amplitude-preserving nonlinear adaptive multiple attenuation using the high-order sparse Radon transform

Yaru Xue[1], Jing Yang[1], Jitao Ma[1] and Yangkang Chen[2]

[1] State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Fuxue Road 18th, Beijing 102200, People's Republic of China
[2] Bureau of Economic Geology, John A and Katherine G Jackson School of Geosciences, The University of Texas at Austin, University Station, Austin, TX 78713-8924, USA

E-mail: xueyaru@cup.edu.cn and ykchen@utexas.edu

Received 1 September 2015, revised 15 January 2016
Accepted for publication 1 February 2016
Published 5 April 2016

Abstract

The Radon transform is widely used for multiple elimination. Since the Radon transform is not an orthogonal transform, it cannot preserve the amplitude of primary reflections well. The prediction and adaptive subtraction method is another widely used approach for multiple attenuation, which demands that the primaries are orthogonal with the multiples. However, the orthogonality assumption is not true for non-stationary field seismic data. In this paper, the high-order sparse Radon transform (HOSRT) method is introduced to protect the amplitude variation with offset information during the multiple subtraction procedures. The HOSRT incorporates the high-resolution Radon transform with the orthogonal polynomial transform. Because the Radon transform contains the trajectory information of seismic events and the orthogonal polynomial transform contains the amplitude variation information of seismic events, their combination constructs an overcomplete transform and obtains the benefits of both the high-resolution property of the Radon transform and the amplitude preservation of the orthogonal polynomial transform. A fast nonlinear filter is adopted in the adaptive subtraction step in order to avoid the orthogonality assumption that is used in traditional adaptive subtraction methods. The application of the proposed approach to synthetic and field data examples shows that the proposed method can improve the separation performance by preserving more useful energy.

Keywords: multiple attenuation, Radon transform, orthogonal polynomial transform, high-order sparse Radon transform, overcomplete dictionary

(Some figures may appear in colour only in the online journal)

1. Introduction

Multiple attenuation is a key step in seismic data processing. There are two main kinds of demultiple methods: the filtering method based on the move-out difference between primaries and multiples and the wave-equation based prediction method (Weglein 1999). The latter consists of two steps: the prediction step and the separation step. The adaptive separation of primaries and multiples is crucial in practice because the initially predicted multiple is not perfectly matched with the true multiple. Many studies have been performed to devise effective and robust adaptive subtraction methods. In Wang (2003b), an expanded multiple multichannel matching filter is proposed, which exploits more local time and phase information to match the multiples. Regularized optimization is also adopted for non-stationary matching filtering (Fomel 2009), which is better suited to real non-stationary seismic data. In Kabir and Abma (2003), the multiples are decomposed into coherent and incoherent components to simulate the diffraction multiples. The above methods are all realized in the

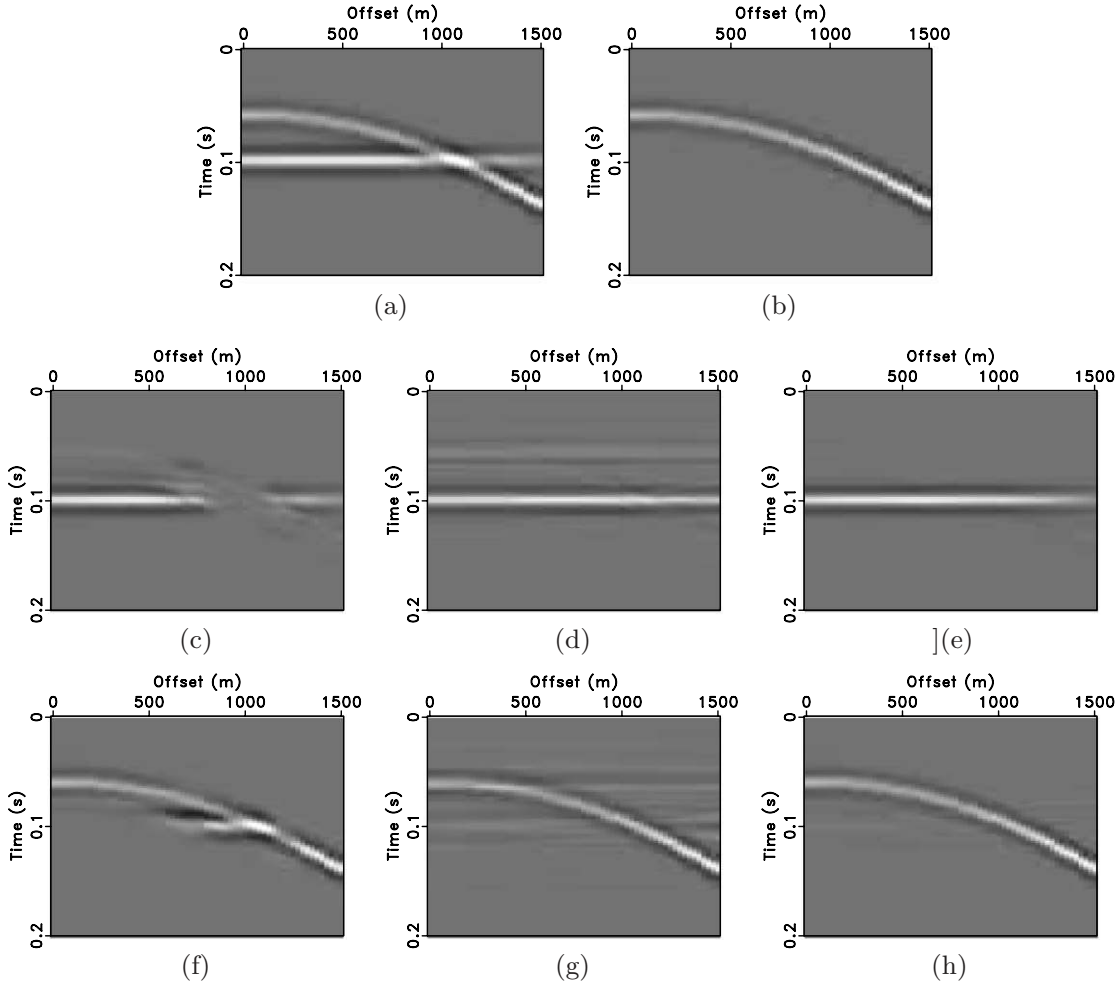

Figure 1. Demultiple comparison for the first synthetic example. (a) Synthetic data. (b) Predicted multiple. (c) Demultiple result using adaptive subtraction in the time–space domain. (d) Demultiple result using adaptive subtraction in the Radon transform domain. (e) Demultiple result using adaptive subtraction in the HOSRT domain. (f) Removed multiple in the time–space domain (corresponding to (c)). (g) Removed multiple using the Radon transform (corresponding to (d)). (h) Removed multiple using the HOSRT (corresponding to (e)).

least-squares (LS) sense. In Lu (2003), an independent component analysis approach is proposed that does not need the orthogonality assumption of the multiples and the primaries. In addition to matching filtering, there are other suggested methods to separate the primaries and multiples. The robust curvelet transform involves subtracting multiples based on the difference of the primaries and multiples in the multiscale and multidirection characteristics (Hermann *et al* 2007). Similar to the expanded multichannel subtraction, the complex-valued curvelet transform is also designed to correct the mismatch between the predicted multiples and the true multiples in the transform domain (Neelamani *et al* 2010).

Compared with wave-equation based methods, filtering methods, such as those based on the Radon transform, are widely used in practical seismic data processing due to their convenience (Foster and Mosher 1992). The demultiple performance using the Radon transform depends on the resolution of multiples and primaries in the Radon domain. Since the Radon transform is not an orthogonal transform, it is often posed as an inversion problem. The damped LS method is an efficient approach for solving the ill-posed inversion problem, but the resolution is not satisfied (Hampson 1986). The high-resolution Radon transform was proposed recently and has found successful applications in data reconstruction, velocity analysis, and multiple attenuation (Thorson and Claerbout 1985, Sacchi and Ulrych 1995). However, the high-resolution Radon transform brings another problem: damaging the amplitude variation with offset (AVO) information while the sparsity priority is too strong. The amplitude-loss problem is significant for seismic data processing and interpretation, in particular when considering pre-stack inversion applications, thus an amplitude-preserving Radon transform is strongly required.

The orthogonal polynomial transform is an efficient way to include the AVO information, and its first few projecting coefficients represent the amplitude variation for the aligned events (Johansen *et al* 1995). When the events are not aligned, the coefficients will smear to high-order coefficients and lose their AVO interpretation. Thus, this orthogonal polynomial transform based method is limited to the exactly normal moveout (NMO)-corrected data.

The high-order sparse Radon transform (HOSRT) embeds the orthogonal polynomial transform in the sparse Radon transform (Xue *et al* 2014). In this method, the Radon

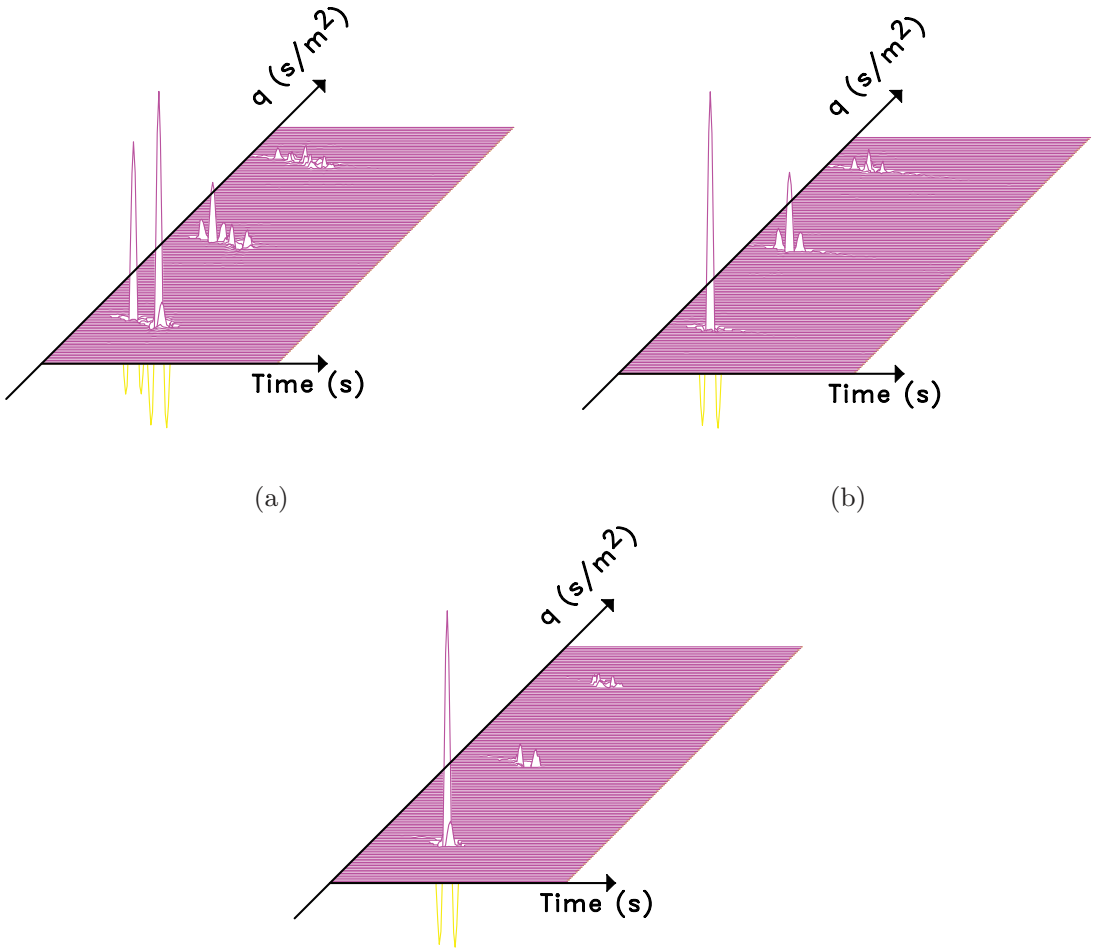

Figure 2. Model space comparison. (a) HOSRT domain of data. (b) HOSRT domain of multiples. (c) HOSRT domain of the adaptively subtracted result. The three well separated regions along the q dimension denote three model spaces ($\mathbf{m}_0$, $\mathbf{m}_1$, and $\mathbf{m}_2$) that have different physical meanings, which are clarified in the main text.

transform describes the trajectory of events and the orthogonal polynomial transform describes the amplitude variation information. The combination of Radon transform and orthogonal polynomial transform constructs an overcomplete transform (Donoho *et al* 2006). The overcomplete dictionary is a signal processing technology. The term *over* means that the dimension of transform bases is much larger than the signal dimension. The overcomplete dictionary is known for its sparse representation for coherent signal (Donoho *et al* 2006). The HOSRT falls into the category of the overcomplete dictionary. We will show that the primaries and multiples can be separated well using the HOSRT method. In the HOSRT domain, we use the nonlinear Butterworth filter to achieve the adaptive subtraction (Wang 2003a), which can help obtain even better separation between primaries and multiples.

The paper is organized as follows: we first introduce the basic theory of the HOSRT, and then we design the adaptive Butterworth filter according to the energy distribution of primaries and multiples in the HOSRT domain to separate the multiples and primaries. Finally we use both synthetic and field data examples to show the advantages of the HOSRT over the traditional approaches.

2. Method

2.1. The high-order sparse Radon transform

The HOSRT has been successfully used in data reconstruction (Xue *et al* 2014). Here we explain it from the viewpoint of an overcomplete transform. The parabolic Radon transform can be expressed as

$$d(t,x) = \sum_{q} m(\tau = t - qx^2, q), \tag{1}$$

where $d(t, x)$ are the seismic data and $m(\tau, q)$ are the model space coefficients. The variables t, x, τ, and q designate the time, offset, intercept time, and residual move-out, respectively. Equation (1) indicates that the seismic data can be represented with a linear combination of the events which have the same amplitude along the parabolic path. The limited aperture results in smearing. The sparse Radon transform reduces the smearing by iteratively re-weighting the inversion matrix (Sacchi and Ulrych 1995). If the q parameter is focused into a point, the offset aperture can be extended infinitely, but the AVO information will be damaged. The reason for amplitude

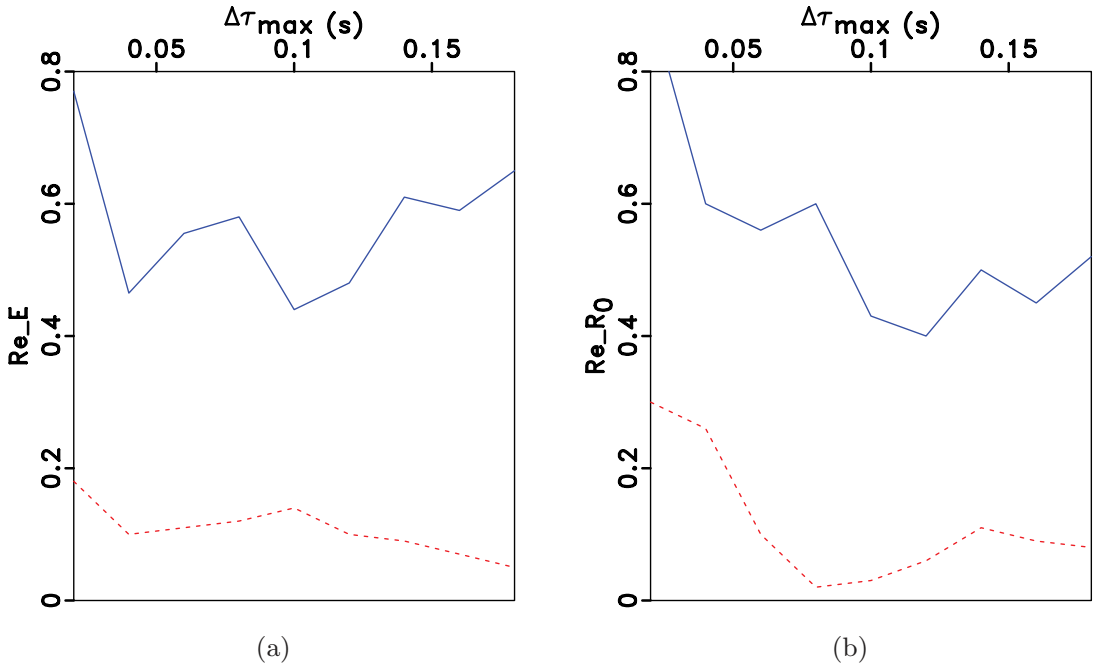

Figure 3. Comparison of relative errors. (a) Primary energy relative errors *RE_E*. (b) Zero-offset primary energy relative errors *Re_R*0 (the dashed line corresponds to the HOSRT and the solid line corresponds to the Radon transform). The measurement shown in (a) is to measure the energy preservation ability of the transform. The measurement shown in (b) is to measure the preservation ability for the zero-offset trace of the transform, which can greatly affect the AVO interpretation accuracy. Both measurements are used to demonstrate the better amplitude preservation ability of the HOSRT.

damage is that the basis function is the constant amplitude events, which is not true for real non-stationary seismic data. Thus, the sparse Radon transform is not adequate to sparsely represent the AVO phenomena.

Considering the lateral continuity of seismic events, seismic data $d(t, x)$ can be fitted with a linear combination of orthogonal polynomial (Johansen *et al* 1995), i.e.

$$d(t,x) = \sum_{j=0}^{N} c_j(t) p_j(x), \tag{2}$$

where $\{p_j(x), j = 0, 1, \cdots, N\}$ are a complete set of unit orthogonal polynomials from the offset coordinate x with $N + 1$ samples. Each $p_j(x)$ has the general form of a polynomial of degree j, $c_j(t)$ is the jth-order fitting coefficient, all of the coefficients finally constitute the orthogonal polynomial spectrum. Usually, the AVO can be represented by the first few orthogonal polynomial coefficients: the zeroth-order components $c_0(t)$ denote the stack along the horizontal direction, the first-order components $c_1(t)$ denote the mean gradient of amplitude variations, the second-order components $c_2(t)$ denote the curvature of amplitude variation that describes the AVO information, and the other higher-order components include amplitude variation details and noise.

Comparing the Radon transform and orthogonal polynomial transform, the basis function of the Radon transform has constant amplitude events along different trajectories, while the basis function of the orthogonal polynomial transform has variable amplitude events without including the trajectory information. To separate the different seismic events with amplitude variation, the above two types of transforms are integrated to construct the high-order Radon transform, which includes the basis function of amplitude variation and the trajectory function:

$$d(t,x) = \sum_{q} \sum_{j} m_j(\tau = t - qx^2, q) p_j(x). \tag{3}$$

Equation (3) indicates that the high-order Radon transform simulates events with a linear superposition of a set of orthogonal polynomials along different trajectories, weighted by their Radon gathers, respectively. In this paper, the second-order Radon transform is adopted, these spectra m_0, m_1, m_2 describe the stack, gradient, and curvature of the amplitude variation, respectively. Equation (3) can be expressed in a matrix-vector form:

$$\mathbf{d} = \mathbf{L}_0\mathbf{m}_0 + \mathbf{L}_1\mathbf{m}_1 + \mathbf{L}_2\mathbf{m}_2 = \begin{pmatrix} \mathbf{L}_0 & \mathbf{L}_1 & \mathbf{L}_2 \end{pmatrix} \begin{pmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix} = \mathbf{L}\mathbf{m}, \tag{4}$$

where $\mathbf{L}_0, \mathbf{L}_1, \mathbf{L}_2$ denote the summation, mean gradient and curvature operator, respectively. Thus, the AVO characterizing capability is incorporated in the forward Radon operator, which matches the real case well.

The above high-order Radon transform is composed of two transforms with different advantages and can overcome the drawbacks of each transform. The Radon transform can depict the trajectory of seismic events but cannot include the amplitude information along the seismic events, while the orthogonal polynomial transform is a complete data transform and can exactly represent the seismic data with detailed amplitude information, but velocity information is not included in this transform. Because the Radon transform is an underdetermined problem, the LS solution recovers data with some

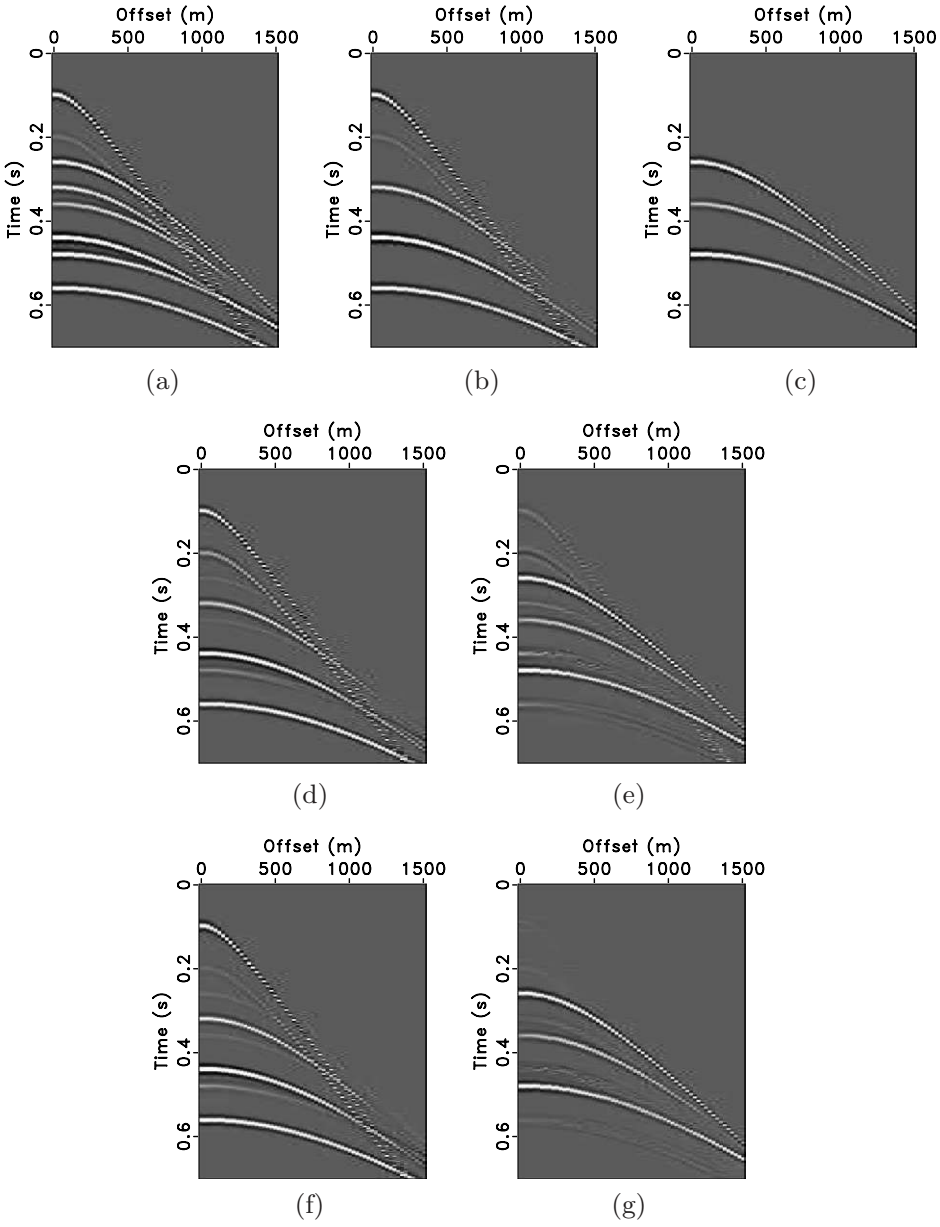

Figure 4. Demultiple comparison for the second synthetic example. (a) Synthetic data. (b) True primaries. (c) True multiples. (d) Demultiple result using adaptive subtraction in the Radon transform domain. (e) Removed multiple using the Radon transform (corresponding to (d)). (f) Demultiple result using adaptive subtraction in the HOSRT domain. (g) Removed multiple using the HOSRT (corresponding to (f)).

signal damage. The compound high-order Radon transform constructs an over underdetermined problem, with more orthogonal polynomial coefficients involved.

The LS solution dilutes across all the basis functions, which brings difficulties in distinguishing between primaries and multiples. The sparse solution of the high-order Radon transform falls into the frame of the overcomplete dictionary representation, which seeks the sparse solution by selecting the best approximation basis among the basis family. The sparse result of $\mathbf{m}$ follows a nonlinear optimization problem

$$\min_{\mathbf{m}} \|\mathbf{d} - \mathbf{L}\mathbf{m}\|_2^2 + \lambda \|\mathbf{m}\|_1, \tag{5}$$

where $\|\cdot\|_2$ and $\|\cdot\|_2$ denote the L_2 and L_1 norms of the input vectors, respectively. λ denotes a controlling factor which balances the weight of the LS misfit and model sparsity constraint.

The matching pursuit and the basis pursuit algorithms are widely used to obtain the sparse solution (Mallat and Zhang 1993, Chen *et al* 2001). Here, we use the matching pursuit method. This method first chooses the best approximate velocity parameters and then transforms the data to the orthogonal polynomial domain. After several iterations, the sparse solution can match all the data. More algorithm details can be founds in Xue *et al* (2014). This solution separates primaries and multiples and forms the sparse representation of the signal. The sparsity not only leads to a low-dimensional

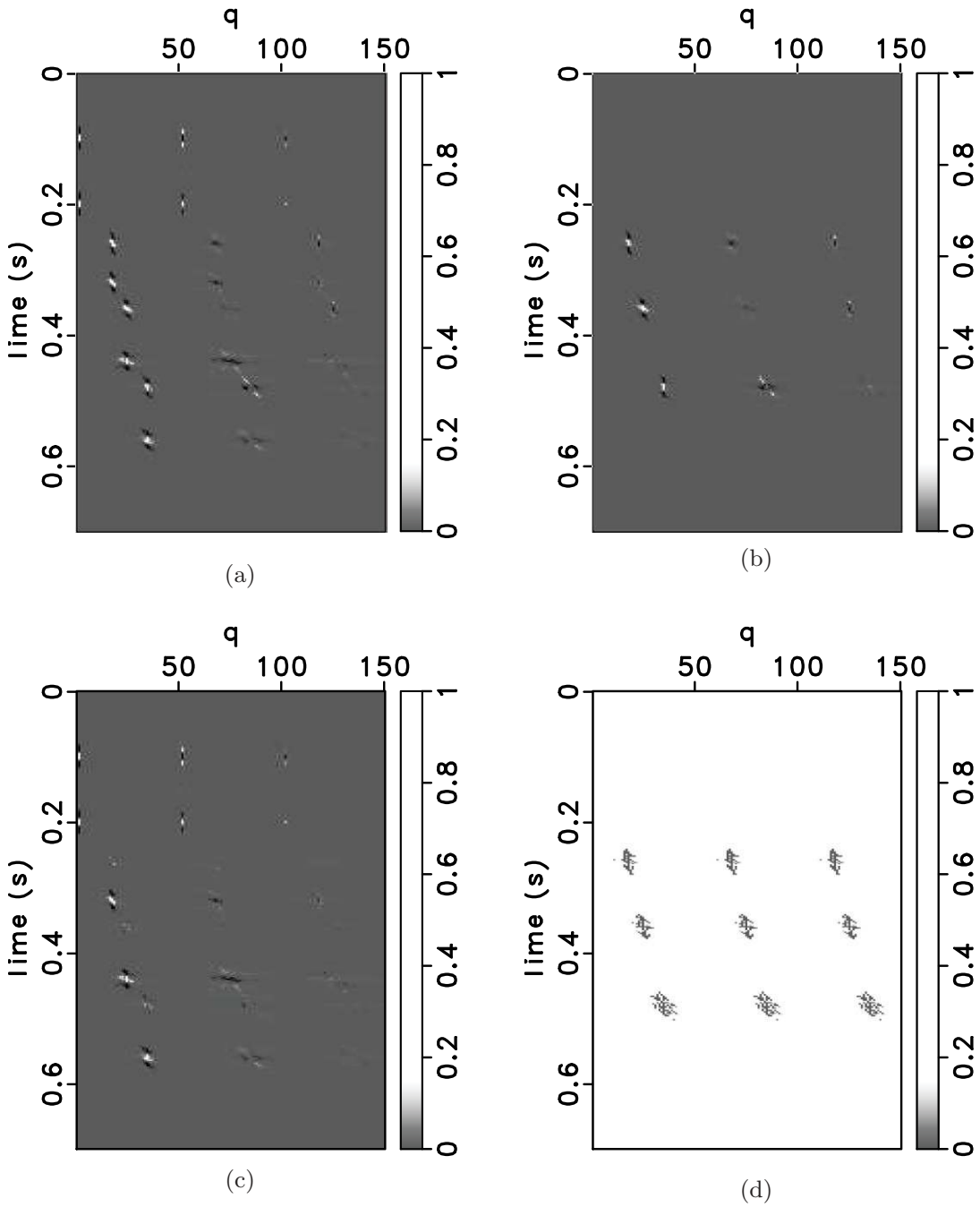

Figure 5. Model space comparison of the second synthetic example. (a) HOSRT domain of data. (b) HOSRT domain of multiples. (c) HOSRT domain of the adaptively subtracted result. (d) The nonlinear filter (f in equation (6)).

signal but reduces the sensitivity to errors in the predicted signal (Hermann *et al* 2007).

2.2. Nonlinear adaptive subtraction

Adaptive subtraction is a key step in multiple attenuation. The widely used LS subtraction is based on the assumption that multiples and primaries are orthogonal to each other, which cannot be satisfied in real case, even in the Radon domain. The nonlinear Butterworth filter is a fast algorithm and can obtain better performance in multiple prediction (Wang 2003a). The filter function can be expressed as:

$$f = \frac{1}{\sqrt{1 + \left(\eta \frac{M}{D}\right)^{2\alpha}}}, \tag{6}$$

where M denotes the predicted multiples, D denotes the data including both primaries and multiples, α is an adjustable parameter to control the smoothness of the filter, and η is the eliminating coefficient. In the HOSRT domain, the ratio between the multiples and data is replaced by the energy of the predicted multiples and the original data, namely

$$\begin{aligned} M &= M^2_{\mathbf{m}_0} + M^2_{\mathbf{m}_1} + M^2_{\mathbf{m}_2}, \\ D &= M^2_{\mathbf{d}_0} + M^2_{\mathbf{d}_1} + M^2_{\mathbf{d}_2}, \end{aligned} \tag{7}$$

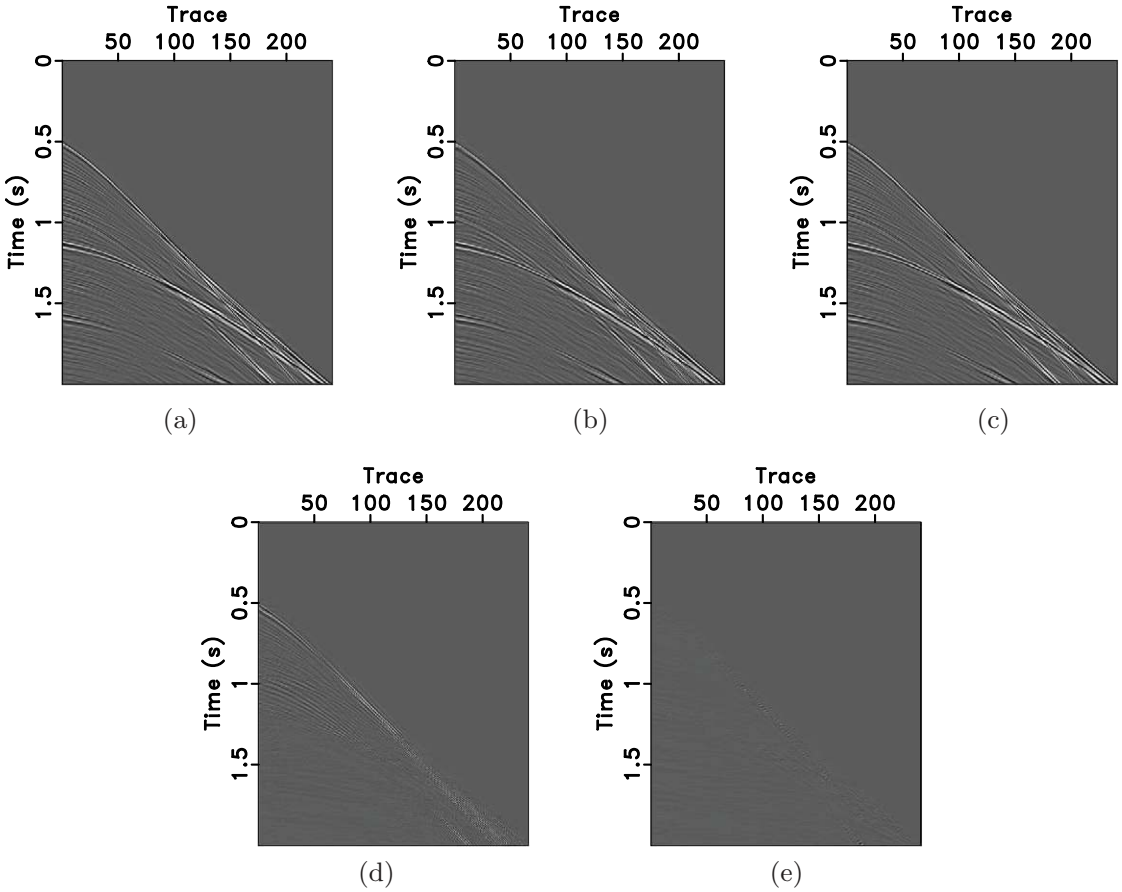

Figure 6. Reconstruction comparison between the sparse Radon transform and the HOSRT. (a) Original common shot gather. (b) Reconstructed data using the sparse Radon transform. (c) Reconstructed data using the HOSRT. (d) Reconstruction residual using the sparse Radon transform (magnified by a factor of two). (e) Reconstruction residual using the HOSRT (magnified by a factor of two).

where $M_{\mathbf{m}_i}$ and $M_{\mathbf{d}_i}$ represent the HOSRT parameters of the predicted multiples and data. The multiples and primaries are separated to some extent by the HOSRT transform. When the local energy of the multiples' model is stronger than that of the primaries' model, the filter achieves an attenuating gain close to 0 to eliminate the multiples, while the primaries' energies are stronger, the filter passes all the data and the primaries are preserved well. Thus, in the HOSRT domain, most primaries are kept well without an orthogonality requirement between the primaries and multiples as required in the LS subtraction method. In addition, this nonlinear adaptive subtraction method can be realized conveniently and efficiently.

The nonlinear adaptive multiple subtraction algorithm in the HOSRT domain can be realized using the following workflow:

1. Wave-equation based multiple prediction.
2. Transform the data and the predicted multiples into the HOSRT domain, respectively, according to equation (5).
3. Filter the multiples out from the data in the HOSRT domain according to equation (6).
4. Inverse transform the primaries from the HOSRT domain to the time–space domain.

The primaries reconstructed from the adaptively subtracted high-order Radon domain can recover the amplitude variation information of the seismic events. This improved amplitude information from the HOSRT method can help us obtain improved common offset profiles with better signal preservation.

3. Examples

3.1. Synthetic data example

The first synthetic example compares the demultiple performance of adaptive subtraction in the time–space domain, the sparse Radon domain, and the HOSRT domain, respectively. For the time–space domain adaptive subtraction approach, we use the classic LS-based method, as used in Verschuur *et al* (1992). Figure 1(a) shows the synthetic data with one multiple event and one primary event. Both the primary event and the multiple event have amplitude variation. Figure 1(b) shows the predicted multiple. Figure 1(c) shows the subtraction result in the time–space domain. The adaptively removed multiple is shown in figure 1(f). It is obviously that the primaries overlapping the multiples are also subtracted in the time–space domain. The adaptive subtraction result in the sparse Radon domain is shown in figure 1(d) and the removed multiple is shown in figure 1(g). It is clear that the primary variations are left in the removed multiple, indicating a loss of useful energy. Figures 1(e) and (h) show the subtraction result and the removed multiple using the HOSRT, respectively. It is

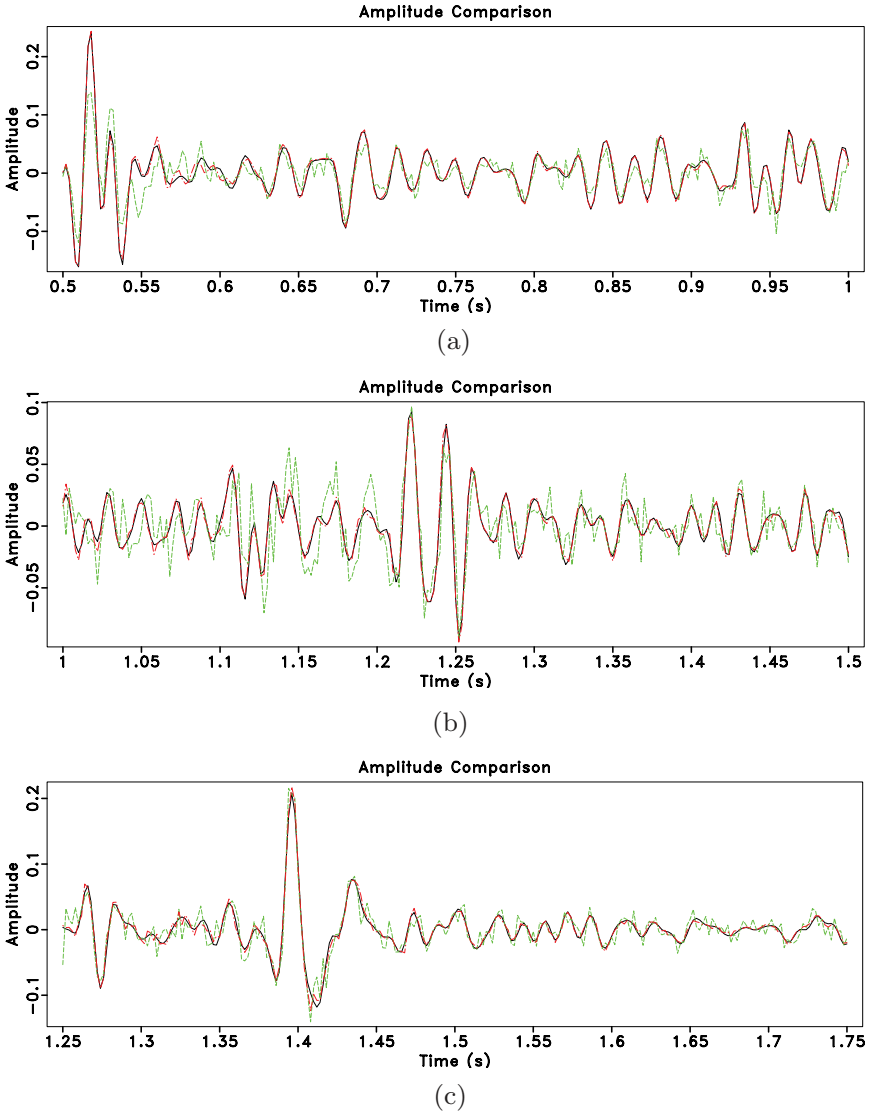

Figure 7. Comparison of the amplitude of different traces of the original data, and the reconstructed data using two methods. (a) Comparison of the first trace between 0.5 s and 1 s. (b) Comparison of the 50th trace between 1 s and 1.5 s. (c) Comparison of the 100th trace between 1.25 s and 1.75 s. The black solid line denotes the original trace (true trace). The green dashed line corresponds to the sparse Radon transform. The red dot–dashed line corresponds to the HOSRT.

obvious that the primary is almost recovered. The comparison of the removed multiples shows that the HOSRT can preserve AVO information well and barely damages the primary event. Figure 2 demonstrates the successful performance of HOSRT in obtaining a sparse model domain. Figure 2(a) is the model space of the HOSRT for the synthetic data. There are zeroth-order, first-order, and second-order parameters in figure 2(a). The model space of HOSRT for the predicted multiple is shown in figure 2(b). Figure 2(c) is the model space of HOSRT for the result after the adaptive subtraction. Figure 2 shows that the energy of multiple can be subtracted clearly in any order, and the energy of primary in different order Radon gathers will be helpful to preserve the AVO characters of the primary reflections.

We also show the primary reflection recovery error for the changing far offsets. Let $\Delta\tau_0$ be the travel time difference between the primary and interfering multiple at the minimum offset and $\Delta\tau_{\max}$ be the difference at maximum offset. The primary energy relative error RE_E and the zero-offset primary energy relative error RE_R_0 can measure the AVO-preserving performance of the Radon transform and the HOSRT. Their definitions are

$$RE\_E = \frac{(P^{\text{true}} - P)^2}{(P^{\text{true}})^2}, \tag{8}$$

$$RE\_R_0 = \frac{(R_0^{\text{true}} - R_0)^2}{(R_0^{\text{true}})^2}, \tag{9}$$

where P^{true} denotes the primary energy and P denotes the primary energy after adaptive subtraction. R_0^{true} denotes the zero-offset primary energy and R_0 denotes the zero-offset primary energy after adaptive subtraction. Figure 3 shows the curves of the relative errors, which change with the far offset. The relative errors are demonstrated with $\Delta\tau_0 = 0.02$ s. In a more straightforward way, we changed the maximum offset with a different

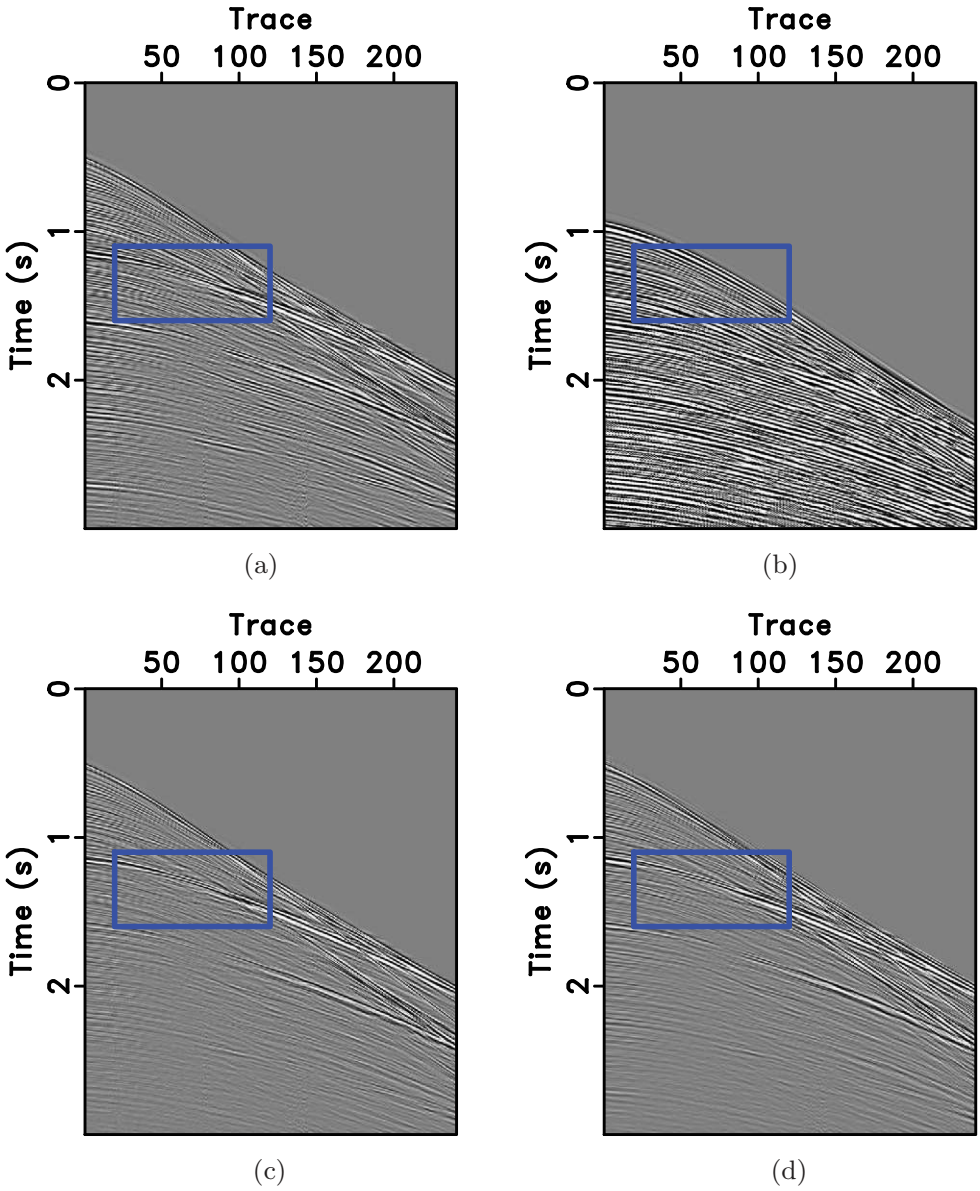

Figure 8. Common shot gather comparison for the marine data example. (a) A shot gather. (b)The predicted multiple. (c) Adaptive subtraction result in the time–space domain. (d) Adaptive subtraction result in the HOSRT domain.

maximum $\Delta\tau_{\max}$, and we performed several experiments for the adaptive subtraction using different Radon transforms. We then calculated P and R_0, which denote the total and zero-offset primary energy, respectively. We measured the amplitude preservation abilities of different Radon transforms by utilizing equations (8) and (9). Figure 3(a) shows the primary energy relative errors based on the adaptive subtraction. Figure 3(b) shows the zero-offset primary energy relative errors. The solid line corresponds to the relative errors of the sparse Radon transform, and the dashed line corresponds to the relative errors of the HOSRT. The comparison illustrates that the HOSRT method has fewer relative errors and preserves more primary energy and information of the zero-offset trace.

Then, we use a relatively more complicated synthetic example to compare the performance between the Radon transform and the HOSRT. Figure 4(a) shows the simulated synthetic data which contain multiples. Figures 4(b) and (c) show the true primaries and true multiples, respectively, used for the simulation. Figures 4(d) and (e) show the demultiple result and corresponding removed multiples using the Radon transform. Figures 4(f) and (g) show the demultiple result and corresponding removed multiples using the HOSRT. Although not as perfect as the first synthetic example, this example also shows an obvious superior performance using the proposed approach over the traditional Radon transform based approach. The removed multiples using the proposed approach, as shown in figure 4(g), are very close to the true multiples as shown in figure 4(c), while the removed multiples using the Radon transform method show some artifacts. The demultiple result using the proposed approach is also much closer to the true primaries' model. In order to show what the nonlinear filter (the muting function) looks like, we plot the model domains of the original data, thetrue multiples, the

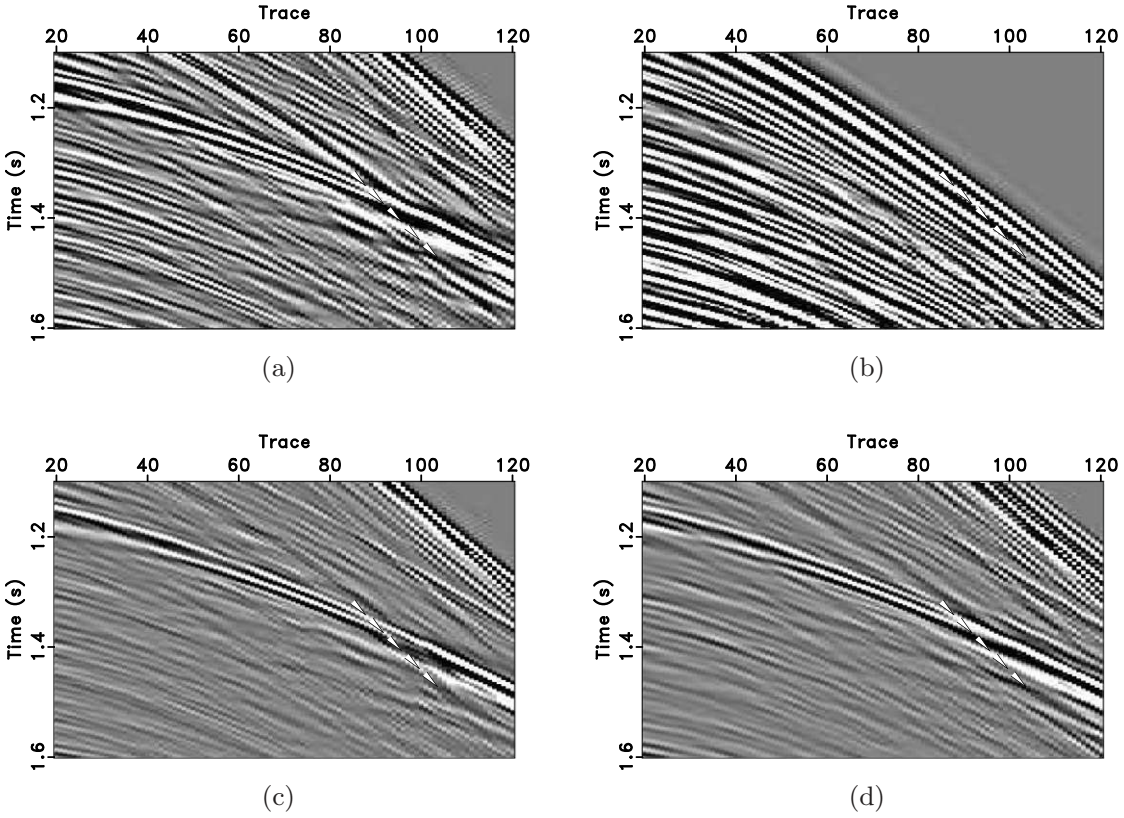

Figure 9. Zoomed common shot gather comparison for the marine data example. (a) Zoomed shot gather. (b) Zoomed predicted multiple. (c) Zoomed result using adaptive subtraction in the time–space domain. (d) Zoomed result using adaptive subtraction in the HOSRT domain.

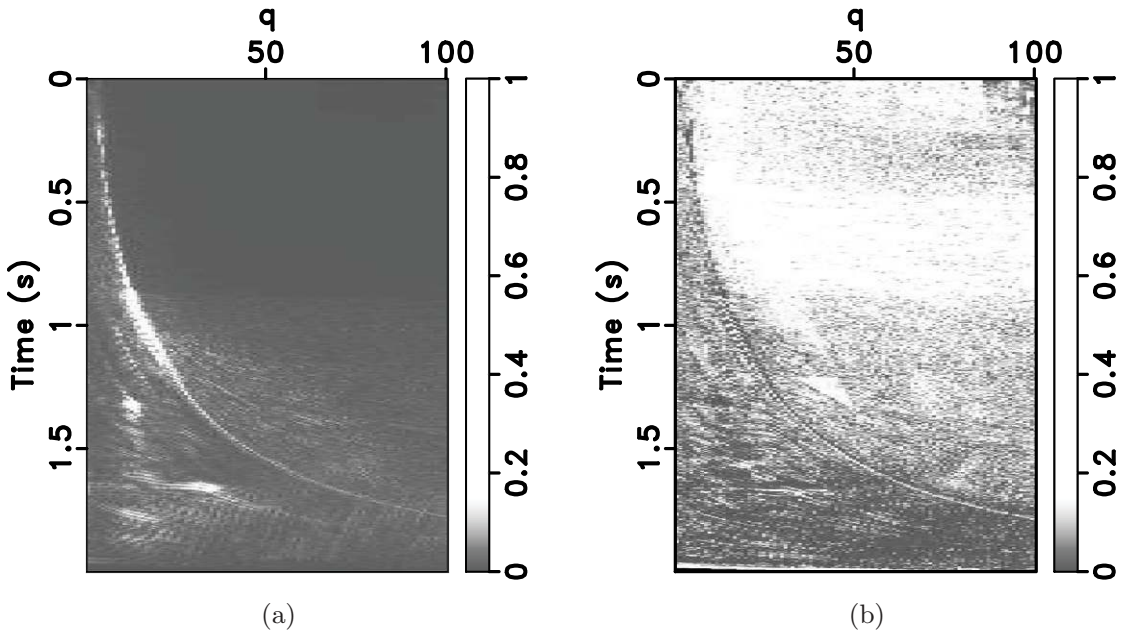

Figure 10. (a) Data in the HOSRT domain. (b) The filter function (the mute function).

demultiple result, and the nonlinear filter used to obtain the demultiple result, in figure 5. It is clear that the nonlinear filter accurately captures the distribution of multiples in the sparse domain, and successfully eliminates the spikes that correspond to multiples by applying a masking operator with very small values.

4. The marine data example

The group of field data experiments contains three different experiments. The first experiment shows the reconstruction comparison (figure 6). Figure 6(a) shows a selected shot gather. Figures 6(b) and (c) demonstrate the reconstruction

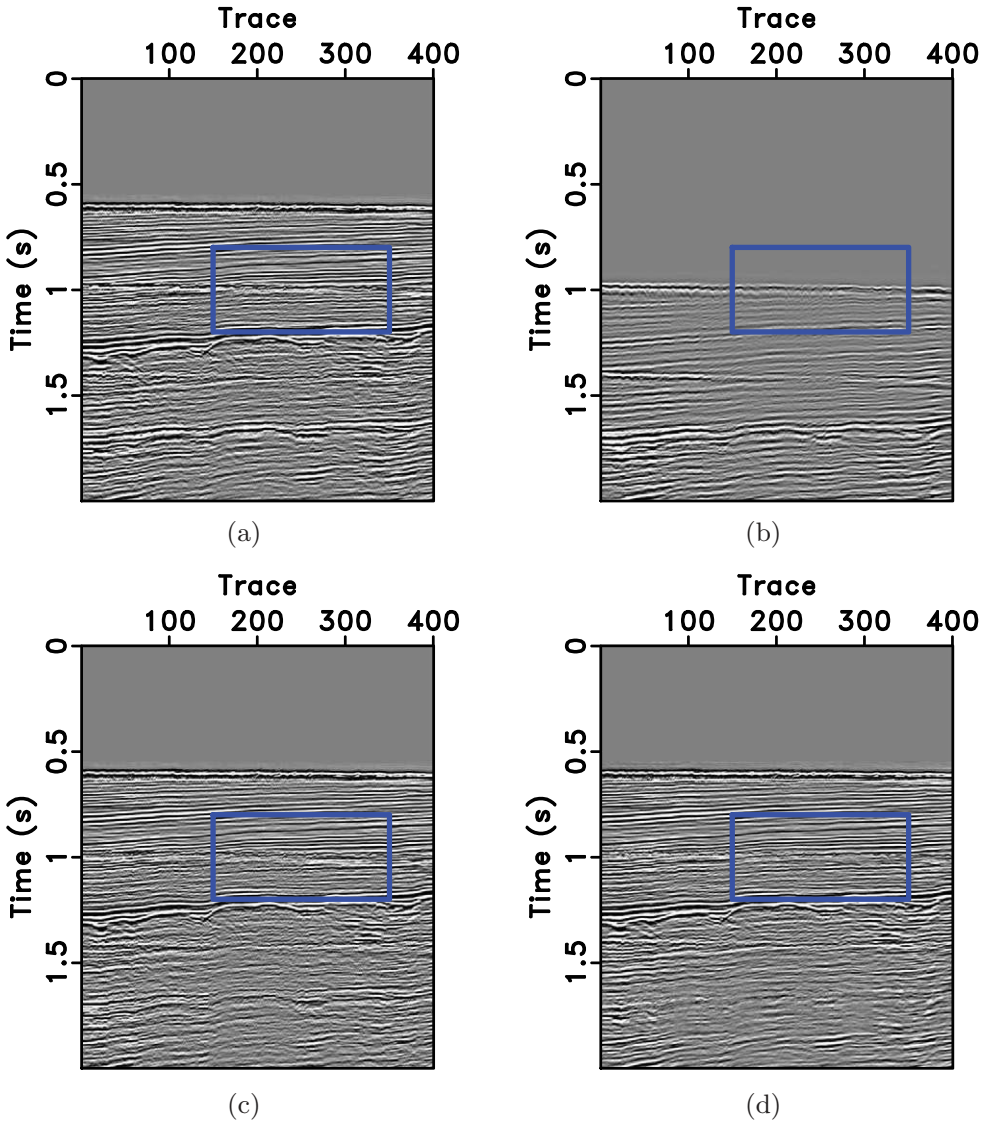

Figure 11. Common offset gather comparison for the marine data example. (a) An offset gather. (b) The predicted multiple. (c) Adaptive subtraction result in the time–space domain. (d) Adaptive subtraction result in the HOSRT domain.

performance using the sparse Radon transform and the HOSRT, respectively. The reconstruction residuals are shown in figures 6(d) and (e). It is obvious that the reconstruction residual of the HOSRT is negligible while the reconstruction residual of the traditional sparse Radon transform is significantly large. Note that the amplitudes of the reconstruction residual sections have been magnified by a factor of two for better comparison. The zoomed zero-offset (first) traces between 0.5 s and 1 s are shown in figure 7(a). The black solid line denotes the original trace (the true trace). The green dashed line corresponds to the Radon transform. The red dot–dashed line corresponds to the HOSRT. It is clear that the black solid and the red dot–dashed lines match very well, but the green dashed line differs too much from the black solid line, indicating a large reconstruction error. The zoomed 50th traces between 1 s and 1.5 s are shown in figure 7(b). The zoomed 100th traces between 1.25 s and 1.75 s are shown in figure 7(c). Both figures 7(b) and (c) show a similar performance using the proposed HOSRT method. Thus, we conclude that for different offsets, the HOSRT can obtain a much better reconstruction result than the Radon transform method. As can be seen from figures 6 and 7, the HOSRT reconstruction is much closer to the original data and damages less amplitude than the Radon transform, thus the HOSRT obtains a better reconstruction result.

The second experiment tries to compare the demultiple and AVO-preserving capabilities between the adaptive subtraction in the time–space domain and that in the HOSRT domain for another selected common shot gather (figure 8). Figures 8(a) and (b) are the original data and the predicted multiples, respectively. Figure 8(c) shows the result of subtracting in the time–space domain and figure 8(d) shows the result of subtracting in the HOSRT domain. For a better comparison, we zoom-in on part of the data and show the detailed comparison in figure 9. Figures 9(a)–(d) correspond to the zoomed sections of figures 8(a)–(d), respectively. The zoomed area is emphasized by the frame boxes as shown in figure 8. From the comparison between subtracting results using the time–space domain method and the HOSRT domain method, in particular from the amplitude preservation comparison as indicated by

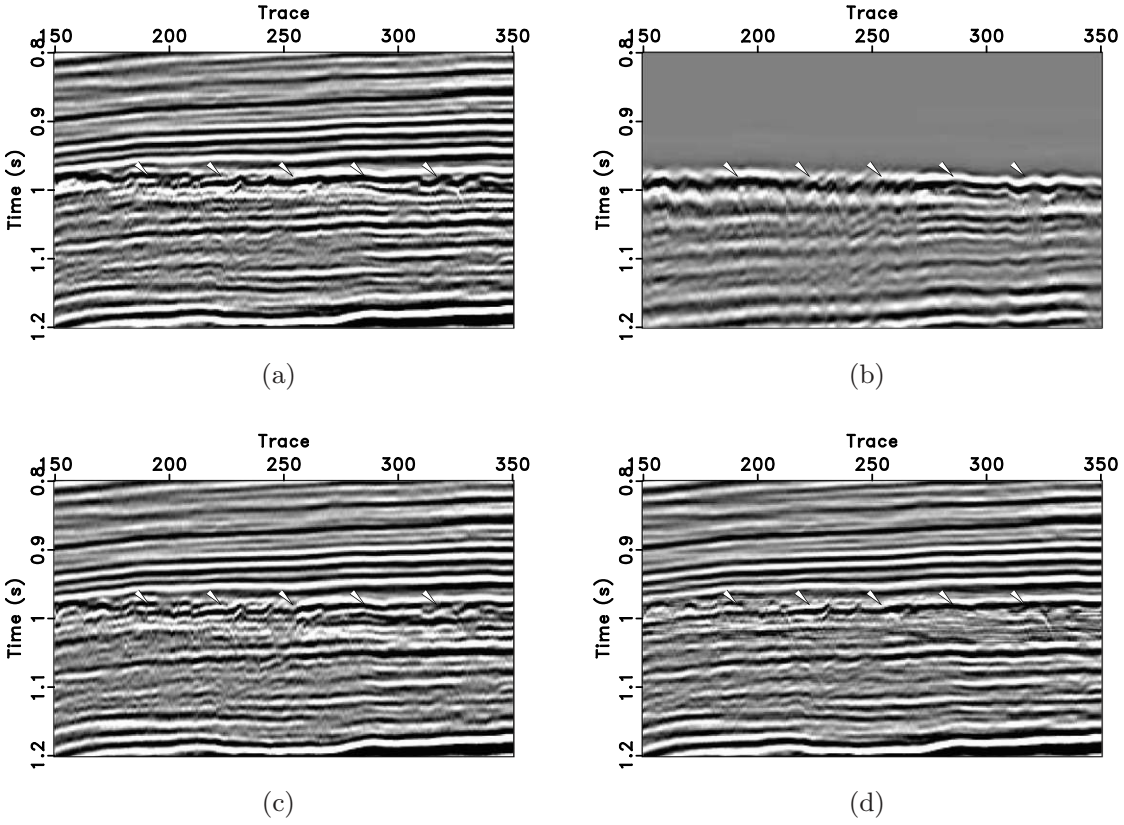

Figure 12. Zoomed common offset gather comparison for the marine data example. (a) Zoomed offset gather. (b) Zoomed predicted multiple. (c) Zoomed result using adaptive subtraction in the time–space domain. (d) Zoomed result using adaptive subtraction in the HOSRT domain.

the arrows in figure 9, we can conclude that the adaptive subtraction in the time–space domain damages the AVO information of primary reflections where multiples intersect with primaries. However, the HOSRT domain approach obtains a successful performance. The continuity of the primaries is improved, and the AVO information is preserved very well using the proposed approach, which is significant for pre-stack seismic inversion. The HOSRT domain of the raw common shot gather and the corresponding nonlinear filter (muting function) to obtain the demultiple result are shown in figure 10 for reference.

A common offset gather is also provided to demonstrate the demultiple performance (figure 11). Figure 11(b) shows the predicted multiples of the marine data as shown in figure 11(a). Figures 11(c) and (d) show the subtracting results using the time–space domain approach and the HOSRT domain approach. The zoomed sections of figures 11(a)–(d) are shown in figures 12(a)–(d), respectively. The zoomed areas are also emphasized by the frame boxes as shown in figure 11. As shown by the zoomed details, the HOSRT approach can suppress most of the multiples and preserve the primaries well while the time–space domain approach causes much more residual multiple energy in the subtracting result, as indicated by the arrows in figure 12. The common offset gather after multiple attenuation using the proposed HOSRT method shows a clear subsurface structure.

At the end of this section, it is worth mentioning that the intention of this paper is to solve the amplitude preservation problem during multiple attenuation when utilizing the high-resolution Radon transform by combining the high-order polynomial transform and the sparse Radon transform. During the adaptive subtraction step, we utilized the nonlinear Butterworth filtering method proposed in Wang (2003a) to improve the separation performance further. We do not intend to provide a better result than the method proposed in Wang (2003a). A comprehensive and fair comparison between our method and the method used in Wang (2003a) is a topic of future investigation.

5. Conclusions

Traditional multiple elimination approaches based on the Radon transform suffer from the low-resolution problem and cannot preserve AVO information. It has been shown that seismic data can be sparsely represented with an overcomplete high-resolution high-order sparse Radon transform (HOSRT). The HOSRT integrates the advantages of both the Radon transform and the orthogonal polynomial transform. It constructs an overcomplete transform domain which not only has the high-resolution Radon transform atoms but also has the orthogonal polynomial atoms. We applied the nonlinear adaptive subtraction method to multiple attenuation by designing a nonlinear Butterworth filter. Compared with the subtraction performance in the time–space domain, the proposed approach can preserve more amplitude details of the primary reflections

because of a better separated structure in the HOSRT domain, in particular when the primaries intersect with the multiples; compared with the subtraction performance using the Radon transform, the proposed approach can preserve AVO information better because of the sparser representation of AVO details in the HOSRT domain.

Acknowledgments

We thank the three anonymous reviewers for their valuable suggestions and comments that greatly improved the manuscript. This work is partially sponsored by the National Natural Science Foundation of China (Nos.41204095), the Science Research Foundation for Returned Overseas Chinese Scholars, State Education Ministry. Yangkang Chen is financially supported by the Texas Consortium for Computational Seismology (TCCS).

References

Chen S, Donoho D L and Saunders M A 2001 Atomic decomposition by basis pursuit *SIAM Rev.* **43** 129–59

Donoho D L, Elad M and Temlyakov V N 2006 Stable recovery of sparse overcomplete representations in the presence of noise *IEEE Trans. Inf. Theory* **52** 6–18

Fomel S 2009 Adaptive multiple subtraction using regularized nonstationary regression *Geophysics* **74** V25–33

Foster D J and Mosher C C 1992 Suppression of multiple reflections using the radon transform *Geophysics* **57** 386–95

Hampson D 1986 Inverse velocity stacking for multiple elimination *J. Can. SEG* **26** 44–55

Hermann F J, Boniger U and Verschuur D J 2007 Non-linear primary-multiple separation with directional curvelet frames *Geophys. J. Int.* **170** 781–99

Johansen T A, Bruland L and Lutro J 1995 Tracking the amplitude versus offset (AVO) by using orthogonal polynomials *Geophys. Prospect.* **43** 245–61

Kabir N and Abma R 2003 Weighted subtraction for diffracted multiple attenuation *73rd Annual Int. Meeting, SEG, Expanded Abstracts* pp 1941–4

Lu W 2003 Adaptive multiple subtraction using independent component analysis *Geophysics* **68** 346–54

Mallat S G and Zhang Z F 1993 Matching pursuit with time-frequency dictionaries *IEEE Trans. Signal Process.* **41** 3397–415

Neelamani R, Baumstein A and Ross W S 2010 Adaptive subtraction using complex-valued curvelet transforms *Geophysics* **75** V51–60

Sacchi M and Ulrych T 1995 High-resolution velocity gathers and offset space reconstruction *Geophysics* **60** 1169–77

Thorson J R and Claerbout J F 1985 Velocity-stack and slant-stack stochastic inversion *Geophysics* **50** 2727–41

Verschuur D J, Berkhout A J and Wapenaar C P A 1992 Adaptive surface-related multiple elimination *Geophysics* **57** 1166–77

Wang Y 2003a Multiple attenuation: coping with the spatial truncation effect in the radon transform domain *Geophys. Prospect.* **51** 75–87

Wang Y 2003b Multiple subtraction using an expanded multichannel matching filter *Geophysics* **68** 346–54

Weglein A B 1999 Multiple attenuation: an overview of recent advances and the road ahead *Leading Edge* **18** 40

Xue Y, Ma J and Chen X 2014 High-order sparse radon transform for AVO-preserving data reconstruction *Geophysics* **79** V13–22