THREE-DIMENSIONAL SPARSE HYPERBOLIC RADON TRANSFORM AND ITS APPLICATION TO DEMULTIPLE

XIANGBO GONG[1,2], SHENCHAO WANG[1] and PAN ZHANG[1]

[1] *College of Geo-Exploration Science and Technology, Jilin University, Changchun, 130026 Jilin, P.R. China. gongxb@jlu.edu.cn*
[2] *Key Laboratory of Applied Geophysics, Ministry of Land and Resources of the People's Republic of China, Changchun, Jilin 130026, P.R. China.*

(Received April 28, 2017; revised version accepted December 16, 2017)

ABSTRACT

Gong, X.B., Wang, S.C. and Zhang, P., 2018. Three-dimensional sparse hyperbolic Radon transform and its application to demultiple. *Journal of Seismic Exploration*, 27: 137-150.

Radon transform (RT) has been widely used in seismic data processing. In this paper, we propose a fast version of three-dimensional (3D) sparse hyperbolic RT (HRT) to eliminate multiples. The adjoint operator transforms the hyperboloids of 3D gather into the points of 3D Radon space, which is used for separation of signal and noise. The forward operator transforms the points of 3D Radon space into the hyperboloids of 3D gather. By using a stretching time axis and the mixed frequency-time domain inversion, we perform the forward and adjoint 3D HRT operators as matrix-matrix multiplications in the frequency domain, which can result in high-resolution Radon panel with high computational efficiency. In addition, a monotone version of fast iterative shrinkage thresholding algorithm (MFISTA) is implemented to accelerate the convergence of the sparse RT-based inversion. Synthetic and field examples of off-shore data demonstrate that our new method can successfully applied in this context.

KEY WORDS: Radon transform, three-dimensional, sparse, demultiple.

INTRODUCTION

Radon transform (RT) for seismic signal processing commonly transforms the time-offset-domain gather into the time-slowness-domain Radon panel. By the moveout difference between signal and noise, we can eliminate the noise or extract the signal after muting in the Radon panel. This technique has been successfully applied to the denoise, demultiple, seismic data interpolation, velocity spectrum, separation of seismic wavefield and simultaneous source separation, etc. (Trad et al., 2002; Wang, 2002, 2003a; Lu, 2013; Ibrahim and Sacchi, 2014; Gong et al., 2014, 2016a, 2016b, 2017; Xue et al., 2014; Ibrahim and Sacchi, 2015; Karimpouli et al.,

0963-0651/18/$5.00

2015; Wang et al., 2015; Zhang et al., 2015; Xue et al., 2016). Another type of RT is named local RT which transforms the gather into different local wavefields. A limited the number of directional wavefields are selected to reconstruct the signal (Sacchi et al., 2004; Yu et al., 2007; Wang et al., 2010).

Generally, two aspects should be considered for RT-based applications. Firstly, the sparsity promotion in Radon panel is an essential factor for practical implementation (Trad et al., 2003; Wang, 2003b). To solve this problem, Thorson and Claerbout (1985) introduced the stochastic inversion for sparse velocity gather. However, their method is computational costly and has difficulty in dealing with the field data. Sacchi and Ulrych (1995) presented an iterative reweighted inversion to construct the high-resolution velocity gather using a Bayesian approach. This approach has been widely used in industrial seismic signal processing. Additionally, the mixed frequency-time domain inversion method is a popular solution for sparse RT (Cary, 1998; Trad et al., 2003; Lu, 2013, Zhang and Lu, 2014a; Gong et al., 2016a). The strategy involves a sparse inversion in the time domain by choosing a l_1 norm to constrain the model and a l_2 norm to constrain the data misfit. The kernel of the RT operator is implemented in the frequency domain by fast Fourier transform (FFT), which both improves the resolution in the Radon panel and preserves high computational efficiency.

Another important aspect for practical applications is to improve the computational efficiency of the RT-based processing. Normally, RT operators performed in the frequency domain can save computation time because the large matrix inversion in the time domain can be converted into several small matrix inversions in the frequency domain (Beylkin, 1987). Abbad et al. (2011) described a fast, modified-parabolic RT, which was based on singular value decomposition. Hu et al. (2013) further improved the computation of hyperbolic RT (HRT) by using a low-rank approximation. In addition, inversion algorithm should also be considered in RT-based inversions. For example, the iterative reweighted least squares (IRLS) algorithm is often used in solving sparse inversion problems (Scales *et al* 1988), but has proved to be inefficient. Lu (2013) developed an accelerated sparse RT by an iterative shrinkage algorithm. Then, Zhang and Lu (2014a) extended this algorithm to three-dimensions and introduced a linear RT for regularization of prestack seismic data. The fast version of the iterative shrinkage algorithm is implemented in RT for faster inversion of sparse matrices (Ibrahim et al., 2015; Gong et al., 2016a).

RT as a mature technique has been applied during the processing of two-dimensional (2D) seismic data. However, for 3D seismic data or data that was gathered over wide azimuths, the operator matrix of 3D RT is much larger than that of 2D RT and it is very costly to inverse the resulting large matrices. Therefore, 3D RT should be designed in a more efficient algorithm to better handle this type of field data. Previous work on 3D RT for seismic signal processing developed the *f-x-y* domain least-squares (LS) RT and applied it to ground roll suppression in orthogonal seismic surveys (Liu and Marfurt, 2004). Hugonnet et al. (2008) introduced a high-resolution 3D

parabolic Radon filter to process dense wide-azimuth gathers in 3D. Verschuur et al. (2012) extended RT to the 4D reconstruction of wide azimuth data. Zhang and Lu (2014a) proposed a high-resolution time-invariant 3D RT by splitting the large matrix-vector-multiplication into two small matrix-matrix multiplications, thereby simplifying the computational complexity of the 3D RT. They applied this method for 2D and 3D prestack, regularized seismic data.

In this paper, we develop a fast version of sparse 3D HRT by combining the stretching time axis, the mixed frequency-time domain inversion strategy and the monotone version of fast iterative shrinkage thresholding algorithm (MFISTA). The adjoint operator of 3D HRT transforms the hyperboloids of 3D gather into points in 3D Radon space, and the forward operator of the 3D HRT transforms the points from 3D Radon space into the hyperboloids in a 3D gather. Synthetic and field data examples demonstrate that our proposition provides robust and efficient results.

THEORY

Due to the hyperbolic characteristic of seismic events in horizontally layered structures, HRT can better focus seismic reflections or diffractions for signal processing. However, the time-variant characteristic of HRT limits its application because of the substantial computational complexity of RT-based operator in the time domain. Pseudo hyperbolic Radon transform (PHRT) is an alternative approach which is based on the integral path defined as $t = \tau + p\sqrt{z^2 + h^2} - z$ (2D case) (Foster and Mosher, 1992; Bickel, 2000; Zhang and Lu, 2014b), where z is the depth parameter, τ is the zero-offset travel time, h is the offset, and p is the slowness. The operators can be implemented in the frequency domain, because it also has a linear relationship between t and τ (time-invariant characteristic). However, the resolution for the Radon panel is unsatisfactory for an arbitrary depth of reflection hyperbolas with the assumption of a constant depth parameter z. The shallower hyperbolas are translated into short hyperbolas and the deeper hyperbolas are translated into short ellipses (Bickel 2000, Zhang and Lu 2014b). The integral path of HRT is defined as $t = \sqrt{\tau^2 + p^2h^2}$ (2D case). It cannot perform RT operator in the frequency domain directly because of its time-variant characteristics. To resolve this problem, a method which stretches the time axis is introduced (Yilmaz, 1989; Sacchi and Ulrych, 1995). Defining $t' = t^2$ and $\tau'=\tau^2$, the path curve of HRT can be rewritten as $t' = \tau' + p^2h^2$, and the time-variant HRT is changed to time-invariant parabolic Radon transform (PRT). Therefore, RT-based operators can be performed in the frequency domain to accelerate the calculation. Note that in the case of a stretching time axis, shallow events with less than one second would be squeezed and deep events with more than one second would be stretched. Thus, an accurate interpolation should be performed to regularize the time samples for Fourier-kernel RT-based operators. Another alternative approach is to replace the fast Fourier transform (FFT) by the nonuniform fast Fourier transform (NFFT) for the seismic data in stretching time axis. (Duijndam et al., 1999).

By using a stretching time axis of the form $t' = t^2$ and $\tau'=\tau^2$, the adjoint operator of 3D HRT is written in the following form:

$$m(\tau', p_x, p_y) = \sum_{h_x} \sum_{h_y} d(t' = \tau' + p_x^2 h_x^2 + p_y^2 h_y^2, h_x, h_y) \ , \tag{1}$$

where $d(t, h_x, h_y)$ presents a 3D gather, h_x and h_y are the offset along the inline and crossline direction, respectively; and p_x and p_y are the slowness along the h_x and h_y direction, respectively, $m(\tau, p_x, p_y)$ presents a 3D Radon space. By eq. (1), the hyperboloids (in normal time axis) or paraboloids (in stretching time axis) of the 3D gather $d(t, h_x, h_y)$ can be projected into the points of the 3D Radon space. The definition is similar to 3D linear RT (LRT) (Liu and Marfurt, 2004; Zhang and Lu, 2014b) and 3D PRT (Hugonnet et al., 2008). Note that the Radon space would be a 4D space $m(\tau, p_x, p_y, p_{xy})$ if an extra parameter, p_{xy}, is employed to represent the squeezed and tilted hyperboloids, which is a more accurate 3D HRT. However, 4D space can result in dramatically increased runtime and strongly ill-conditioned problems (Hugonnet et al., 2008). Thus, we keep the Radon space as 3D one which is enough to represent all the hyperboloids in a 3D gather.

In the case of a stretching time axis, the forward operator of 3D-HRT is written as follows:

$$d(t', h_x, h_y) = \sum_{p_x} \sum_{p_y} m(\tau' = t' - p_x^2 h_x^2 - p_y^2 h_y^2, p) \ . \tag{2}$$

The equation transforms the hyperboloids (in normal time axis) or paraboloids (in stretching time axis) of the 3D Radon space into the points of the 3D gather. Instead of solving the huge computational complexity in the time domain, FFT is taken along the stretching time axis. The adjoint operator of 3D HRT solving in the frequency domain can be written as:

$$M(\omega', p_x, p_y) = \sum_{h_x} \sum_{h_y} D(\omega', h_x, h_y) \mathrm{e}^{\mathrm{i}\omega'(p_x^2 h_x^2 + p_y^2 h_y^2)}. \tag{3}$$

We change the integral order by applying the multiplication rule:

$$M(\omega', p_x, p_y) = \sum_{h_y} \left(\sum_{h_x} \mathrm{e}^{\mathrm{i}\omega' p_x^2 h_x^2} D(\omega', h_x, h_y) \right) e^{i\omega' p_y^2 h_y^2}. \tag{4}$$

The adjoint operator matrix has two different forms for each frequency slice (Zhang and Lu 2014a):

$$\mathbf{M_{vec}} = \mathbf{L}_{xy}^T * \mathbf{D_{vec}} \quad , \tag{5a}$$

$$\mathbf{M} = \mathbf{L}_x^T * \mathbf{D} * \mathbf{L}_y^T \quad , \tag{5b}$$

where $\mathbf{L}_x$ and $\mathbf{L}_y$ are the operators associated the inline and crossline direction, respectively, superscript T denotes the conjugate transpose

operator, symbol * denotes matrix multiplication, $\mathbf{L}_{xy}=\mathbf{L}_y^T \otimes \mathbf{L}_x$, symbol $\otimes$ denotes the Kronecker product of two matrixes, subscript **vec** denotes the matrix vectorization operator, and $\mathbf{M}_{\mathbf{vec}}= \mathbf{vec}(\mathbf{M})$, $\mathbf{D}_{\mathbf{vec}}= \mathbf{vec}(\mathbf{D})$. As noted by Zhang and Lu (2014a), the size of the matrix $\mathbf{L}_{xy}$ is much larger than that of $\mathbf{L}_x$ and $\mathbf{L}_y$, thus eq. (5b) instead of (5a) is involved to solve 3D RT-based operator. The same implementation is taken to Radon space for the forward operator of 3D HRT, thus, the forward and adjoint operator in the matrix format of 3D HRT can be written as:

$$\mathbf{M}_{\text{adj}} = \mathbf{L}_x^T * \mathbf{D} * \mathbf{L}_y^T \quad . \tag{6}$$

$$\mathbf{M}_{\text{adj}} = \mathbf{L}_x^T * \mathbf{D} * \mathbf{L}_y^T \quad . \tag{7}$$

Eq. (7) can only produce relatively low-resolution results because of the limited truncation in discrete computation. We can circumvent this problem by using a linear inversion approach. The cost function of the least squares (LS) solution without constraint is written as:

$$J(\mathbf{M}) = \left\|\mathbf{D} - \mathbf{L}_x * \mathbf{M} * \mathbf{L}_y\right\|_2^2 \quad . \tag{8}$$

Then, LS solution of 3D HRT can be written as:

$$\mathbf{M}_{\text{ls}} = (\mathbf{L}_x^T\mathbf{L}_x)^{-1}\mathbf{L}_x^T * \mathbf{D} * \mathbf{L}_y^T\left(\mathbf{L}_y\mathbf{L}_y^T\right)^{-1}. \tag{9}$$

To improve the resolution of the 3D Radon space, a sparsity-promoting inversion method is introduced and the cost function is written as a mixed l_1-l_2 norm (Trad et al., 2003; Yuan et al., 2015):

$$J(\mathbf{M}) = \left\|\mathbf{D} - \mathbf{L}_x * \mathbf{M} * \mathbf{L}_y\right\|_2^2 + \lambda\|\mathbf{M}\|_1 \quad . \tag{10}$$

The cost function utilizes the l_2-norm to minimize the data error and the l_1-norm to sparse constrained the model. λ is the tradeoff parameter which controls the sparsity of the Radon space and data misfit (Sacchi and Ulrych 1995; Chen et al., 2014).

The resolution for the Radon space solved by eq. (10) is not ideal because all the curvatures of the events are coupling in the case of the sparsity-promoting inversion in the frequency domain (Trad et al., 2003). Consequently, a mixed frequency-time domain strategy is proposed for a high-resolution RT-based sparsity-promoting inversion. This well-known high-resolution strategy is applied to optimize the linear inversion problem in the time domain and apply the RT-based operator in the frequency domain by matrix-matrix multiplications as Radon space is much sparser in the time domain than that in the frequency domain. Only by taking several additional FFTs, the resolution of Radon space and the computational efficiency can be improved. Furthermore, the seismic data can also preserve its waveform by the frequency domain RT operators (Cary, 1998; Trad et al.,

2003; Lu, 2013). Thus, by the mixed frequency-time domain strategy and the sparsity-promoting inversion the cost function is written as:

$$J(\mathbf{m}) = \left\| \mathbf{d} - F^{-1}\mathbf{L}_x * F(\mathbf{m}) * \mathbf{L}_y \right\|_2^2 + \lambda \|\mathbf{m}\|_1 \ , \tag{11}$$

where F and F^{-1} denote the forward and inverse Fourier transform operators, respectively. Many algorithms can be implemented to solve this kind of inversion problem. The popular one is IRLS (Scales et al., 1988) for its simplicity. Another widely used algorithm is preconditioned conjugate gradient algorithm (Trad et al., 2002). Some other solutions such as the greedy algorithm are also well documented (Ng and Perz, 2004; Wang et al., 2010). Lu (2013) introduced an iterative shrinkage thresholding algorithm (ISTA) to further speed up the calculation and it is proved to be effective and robust for this inversion problem. Zhang and Lu (2014a) extended ISTA to 3D LRT and it was successfully applied to regularize the 3D prestack seismic data. To accelerate the inversion calculation, a fast iterative shrinkage thresholding algorithm (FISTA) is adopted to solve the RT-based inversion (Gong et al., 2016a). Additional applications of FISTA to seismic data inversions are during microseismic data denoising by a sparsity constrained time-frequency transform (Rodriguez et al., 2012) and high-resolution amplitude-versus-angle attributes inversion for prestack seismic data (Perez et al., 2013). As Beck and Teboulle noted, FISTA is not a monotone algorithm. To guarantee the monotone of FISTA, an additional judging function is used to monitor the misfit of each iterative computation to obtain a more stable convergence, which only increases a small amount of computational time. In this paper, we employ MFISTA to achieve a robust and stable convergence of RT-based inversion. The process of MFISTA solving 3D HRT is written as following iterative equations, where the LS solution is the initial solution:

$$\mathbf{m}_0 = F^{-1}\left[(\mathbf{L}_x^T\mathbf{L}_x)^{-1}\mathbf{L}_x^T * F(\mathbf{d}) * \mathbf{L}_y^T\left(\mathbf{L}_y\mathbf{L}_y^T\right)^{-1} \right] \ \mathrm{m} \ , \tag{12a}$$

$$\mathbf{x}_1 = \mathbf{m}_0 \ , \tag{12b}$$

$$t_1 = 1 \ . \tag{12c}$$

If $k \geqslant 1$:

$$\mathbf{z}_k = soft\left(\mathbf{x}_k + \frac{F^{-1}\left\{ (\mathbf{L}_x^T\mathbf{L}_x)^{-1}\mathbf{L}_x^T * F\left[\mathbf{d} - F^{-1}\mathbf{L}_x * F(\mathbf{x}_k) * \mathbf{L}_y \right] * \mathbf{L}_y^T\left(\mathbf{L}_y\mathbf{L}_y^T\right)^{-1} \right\}}{\alpha}, \frac{\lambda}{2\alpha} \right) \ , \tag{12d}$$

$$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2} \ , \tag{12e}$$

$$\mathbf{m}_k = \mathrm{argmin}\{J(\mathbf{m}): \mathbf{m} = \mathbf{z}_k, \mathbf{m}_{k-1}\} \ , \tag{12f}$$

$$\mathbf{x}_{k+1} = \mathbf{m}_k + \left(\frac{t_k}{t_{k+1}}\right)(\mathbf{z}_k - \mathbf{m}_k) + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{m}_k - \mathbf{m}_{k-1}) \ , \tag{12g}$$

where k is the current iteration, *soft* is the soft thresholding operator, α is the Lipschtiz constant, $\alpha \geq \max\left(\mathrm{eig}\left(\left(\mathbf{L}_x^T\mathbf{L}_x\right) * \left(\mathbf{L}_y\mathbf{L}_y^T\right)\right)\right)$, which can be evaluated by Rayleigh's power method (Perez et al., 2013), operator max(**a**) is defined to obtain the maximum value of vector **a**, and operator eig(**A**) is defined to obtain the eigenvalue of matrix **A**. Note that, in the step of (11f), the first function value is needed to be computed both FISTA and MFISTA, and second function value $J(\mathbf{m}_{k\text{-}1})$ is already computed in the previous iterate (Beck and Teboulle, 2009), therefore, MFISTA only cost slightly higher than FISTA. After several iterative computations, the high-resolution 3D Radon space $\mathbf{m}_k$ is obtained. If eq. (11) is solved by IRLS, RT-based operator with the weighted matrix $(\mathbf{L}_x^T\mathbf{L}_x + \lambda\mathbf{W}_m^T\mathbf{W}_m)$ or $(\mathbf{L}_y\mathbf{L}_y^T + \lambda\mathbf{W}_m^T\mathbf{W}_m)$ ($\mathbf{W}_m$ denotes weighted matrix for model which varies in each iteration) is needed to calculate inverse matrix for every each iteration, which is very costly; if eq. (11) is solved by MFISTA, the pseudoinverse matrixes $(\mathbf{L}_x^T\mathbf{L}_x)^{-1}$ and $(\mathbf{L}_y\mathbf{L}_y^T)^{-1}$ in each iterative computation are same, thus, we compute those matrixes only once at the first iteration and store them in the computer memory for the other iterations. To be more generalized, these pseudoinverse matrixes at each frequency can be stored for the other gathers with the same geometry acquisition.

EXAMPLES

Synthetic 3D gather example

A synthetic 3D gather is tested to illustrate our proposition. Fig. 1 shows this synthetic 3D gather and its 20th slice in the crossline axis. It contains 1000 time samples (nz) in each trace with a 4 ms sampling interval, its inline (nx) and crossline (ny) lines are 61with a 0.05 km spatial interval. Along inline and crossline directions, its offset range from -1.5 km to 1.5 km. This 3D gather has several low-velocity events (1.8 km/s) indicating multiples and several relatively high-velocity events (2 to 3 km/s) indicating primaries. When 3D HRT is implemented in the time domain, the computational cost is $O\,(nx \times ny \times nz \times np)$ for an adjoint RT operator, where np is the number of discrete slowness. That is a quite heavy computational burden for most of computers. If 3D HRT is solved by using eq. (7) in each frequency slice, the computational cost of one adjoint operator is only $O\,((nx \times np + np \times ny) \times n\omega)$ as it only takes two matrix-matrix multiplications whose dimensions are $(nx \times np)$ and $(np \times ny)$, respectively, where $n\omega$ is the number of discrete frequencies. The adjoint solution, the LS solution, and sparse solution of 3D Radon space are shown in Figs. 2a, 2b and 2c, respectively. As it is noted in the theory section, the mixed frequency-time domain strategy and the sparsity-promoting inversion can greatly improve the resolution of Radon space. Thus, the resolution of Fig. 2c is obviously better than the ones of Figs. 2a and 2b, and it is easy to separate the primaries from the multiplies. In the same computational condition, the runtime in seconds of Figs. 2a, 2b and 2c are 20.7 s, 30.4 s and 263.0 s, where the iteration number of FISTA is 10. Fig. 3 is the result of demultiple by our proposition. The 3D multiple gather and 3D primary gather are shown in Figs. 3a and 3b, respectively. The right parts of Fig. 3 are their corresponding 20th slices in the crossline axis.

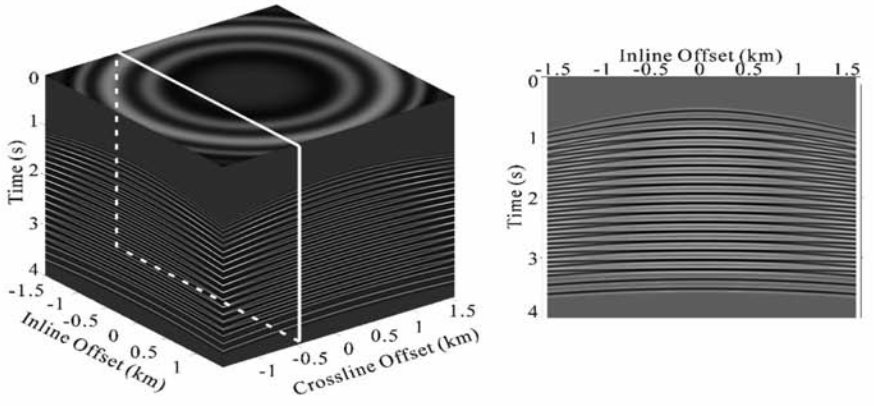

Fig. 1. A synthetic 3D gather with time axis ranges from 0-4 s, and its inline and crossline offset range from -1.5 to 1.5 km. It contains several relatively low-velocity events indicating multiples and relatively high-velocity events indicating primaries. The right part of figure is the 20th slice in the crossline axis.

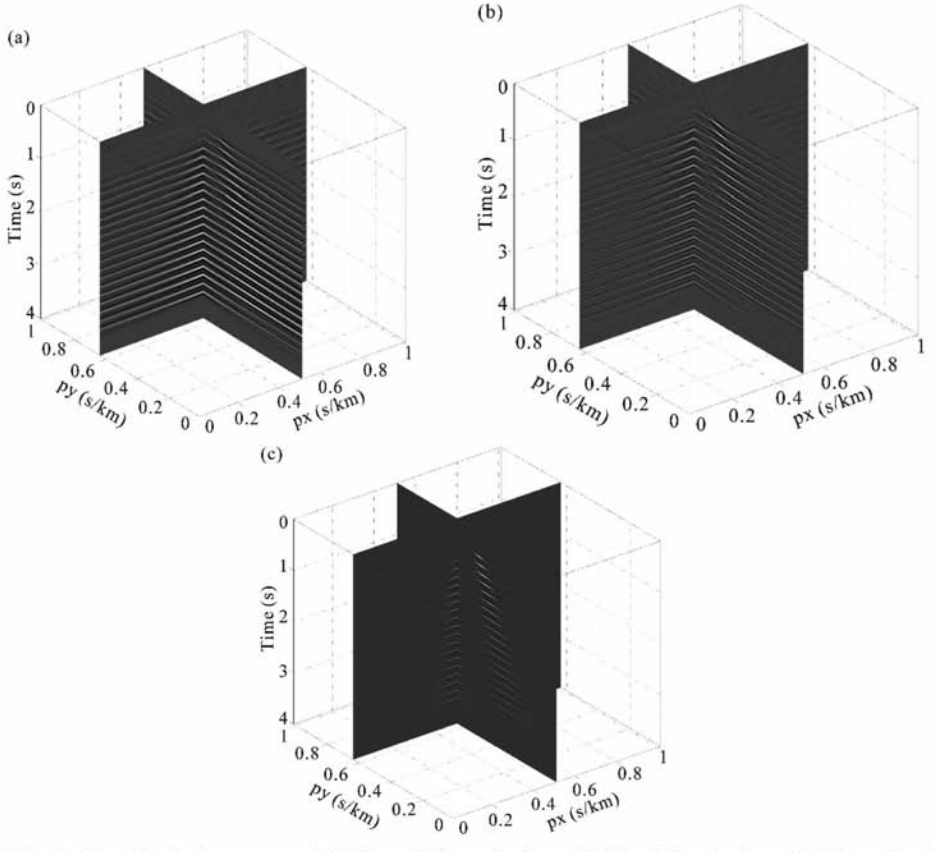

Fig. 2. The 3D Radon space. (a) The adjoint solution; (b) The LS solution; (c) The mixed frequency-time domain sparsity-promoting inversion solution.

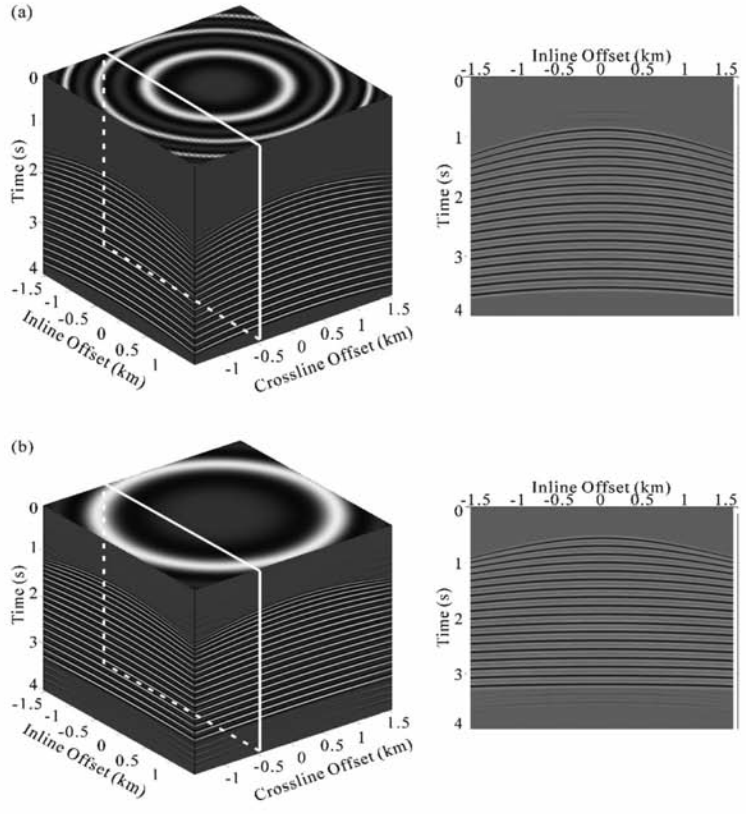

Fig. 3. The result of demultiple. (a) Low-velocity events of the 3D gather indicating multiples; (b) High-velocity events of the 3D gather indicating primaries.

Field data example

Fig. 4 shows a field marine gather which contains 10 cables (the number of inline is 10). The maximum inline offset is 3 km and the trace interval is 0.025 km. The crossline offset (vertical direction of inline) varies from -0.5 km to 0.4 km. The samples in one trace is 2000 with a 4 ms sampling interval. This gather contains primaries, surface-related multiples, internal multiples and other random noise. The 3D Radon space of our proposition is shown in Fig. 5. By muting the energy of multiple in 3D Radon space, 3D primary gather and multiple gather are separated and shown in Figs. 6(a) and 6(b), respectively. Note that some random noise is also attenuated because of the sparsity of Radon space. The inversion error should be considered because 3D HRT is performed based on inversion. Increasing the number of iterations and reducing the tradeoff parameter λ can enhance the fidelity of the recovered data, but also increase the computational cost. Also note that the muting region of demultiple is not easy to be set for 3D data, thus, we choose a same muting region for a group of adjacent slices along *px* or *py* axis. For a better comparison, we zoom in the black rectangle region in Figs. 6a and 6b, which are enlarged to show in Figs. 7a and 7b, respectively.

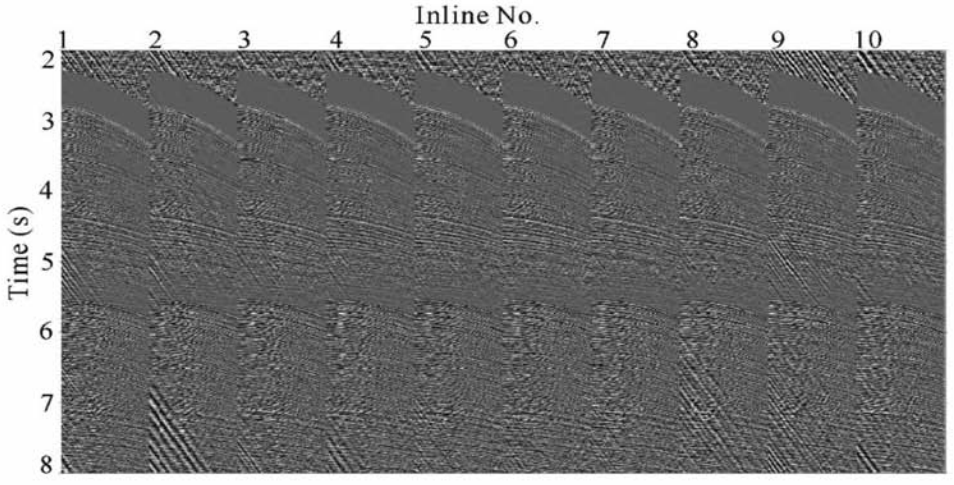

Fig. 4. 3D field data with 10 inlines. Its inline offset range from 0-3km and crossline offset range from -0.5 to 0.4km.

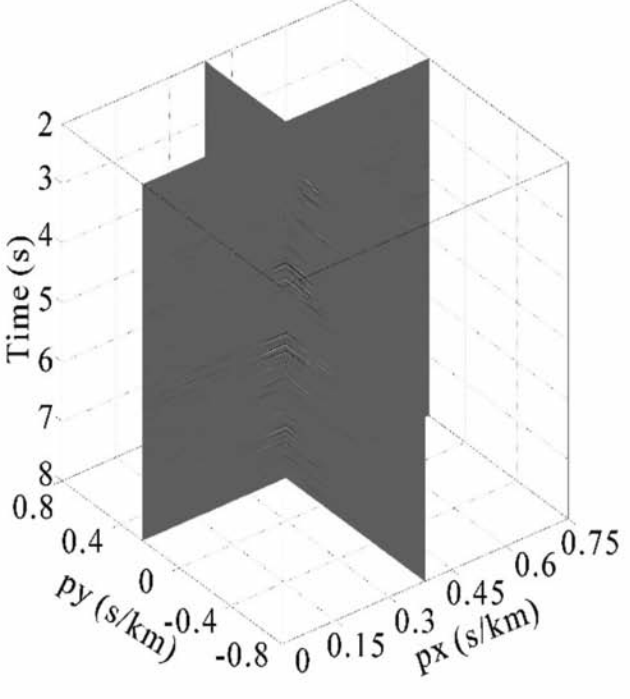

Fig. 5. 3D Radon space by the sparse 3D HRT.

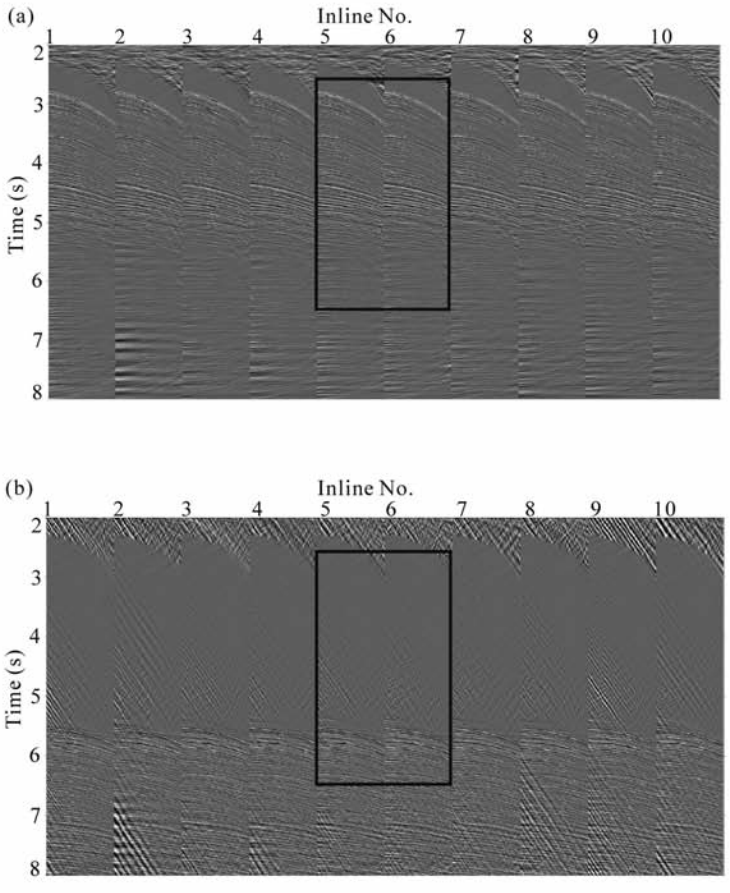

Fig. 6. The result of demultiple. (a) Primaries; (b) Multiples.

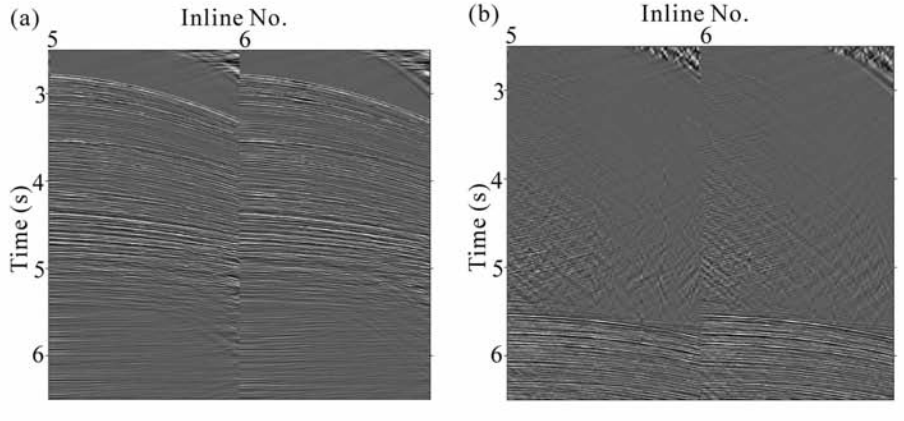

Fig. 7. The enlargement of black rectangle region in Fig. 6. (a) Primaries; (b) Multiples.

CONCLUSIONS

3D HRT considers that the curvatures of events vary along two horizontal directions, which is more suitable for 3D seismic data processing. In this paper, we have proposed a 3D sparse HRT which transforms the 3D gather into a 3D Radon space combining the stretching time axis, the sparsity-promotion strategy and MFISTA. The adjoint and forward operators of 3D HRT in the frequency domain are presented with the form of matrix-matrix multiplications. The mixed frequency-time domain sparsity-promoting inversion and MFISTA are applied to obtain a sparse Radon space. The synthetic and field 3D data examples demonstrate that our proposition is effective and robust. Due to non-orthogonal characteristic of 3D HRT, the computational inversion error is not evitable which needs be further studied in the future.

ACKNOWLEDGEMENTS

This work is supported by the National Nature Science Foundation of China (No. 41674119) and the National "863" Project of China (Nos. 2014AA06A604; 2014AA06A605).

REFERENCES

Abbad, B., Ursin, B. and Porsani, M.J., 2011. A fast, modified parabolic Radon transform. Geophysics, 76: V11-V24.

Beck, A. and Teboulle, M., 2009. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Transact. Image Process., 18: 2419-2434.

Beylkin, G., 1987. Discrete Radon transform. IEEE Trans. Acoust., Speech Signal. Process., 35: 162–172.

Bickel, S.H., 2000. Focusing aspects of the hyperbolic Radon transform. Geophysics, 65: 652–655.

Cary, P.W., 1998. The simplest discrete Radon transform. Expanded Abstr., 68th Ann. Internat. SEG Mtg., New Orleans: 1999-2002.

Chen, Y., Chen, K., Shi, P. and Wang. Y., 2014. Irregular seismic data reconstruction using a percentile-half-thresholding algorithm. J. Geophys. Engineer., 11: 065001.

Duijndam, A.J.W., Schonewille, M.A. and Hindriks, C.O.H., 1999. Reconstruction of seismic signals, irregularly sampled along one spatial direction. Geophysics, 64: 524-538.

Foster, D.J. and Mosher, C.C., 1992. Suppression of multiple reflections using the Radon transform. Geophysics, 57: 386–395.

Gong, X., Han, L. and Li, H., 2014. Anisotropic Radon transform and its application to demultiple. Chin. J. Geophys. 57: 2928-2936.

Gong, X., Yu, S. and Wang, S., 2016a. Prestack seismic data regularization using a time-variant anisotropic Radon transform. J. Geophys. Engineer., 13: 462-469.

Gong, X., Wang, S. and Zhang, T., 2016b. Velocity analysis using high-resolution semblance based on sparse hyperbolic Radon transform. J. App. Geophys., 134: 146-152.

Gong, X., Yu, C. and Wang, Z., 2017. Separation of prestack seismic diffractions using an improved sparse apex-shifted hyperbolic Radon transform. Explor. Geophys., 48: 476-484.

Hu, J., Fomel, S., Demanet, L. and Ying, L., 2013. A fast butterfly algorithm for generalized Radon transforms. Geophysics, 78: U41-U51.

Hugonnet, P., Boelle, J. and Mihoub, M., 2008. High resolution 3D parabolic Radon filtering. Expanded Abstr., 78th Ann. Internat. SEG Mtg., New Orleans: 2492-2496.

Ibrahim, A. and Sacchi, M.D., 2014. Simultaneous source separation using a robust Radon transform. Geophysics, 79: V1-V11.

Ibrahim, A. and Sacchi, M.D., 2015. Fast simultaneous seismic source separation using Stolt migration and demigration operators. Geophysics, 80: WD27-WD36.

Karimpouli, S., Malehmir, A., Hassani, H., Khoshdel, H. and Nabi-Bidhendi, M., 2015. Automated diffraction delineation using an apex-shifted Radon transform. J. Geophys. Engineer., 12: 199-209.

Liu, J. and Marfurt, K.J., 2004. 3-D high resolution Radon transforms applied to ground roll suppression in orthogonal seismic surveys. Expanded Abstr., 74th Ann. Internat. SEG Mtg., Denver: 2144-2147.

Lu, W., 2013. An accelerated sparse time-invariant Radon transform in the mixed frequency-time domain based on iterative 2D model shrinkage. Geophysics, 78: V147–V155.

Ng, M. and Perz, M., 2004. High resolution Radon transform in the t-x domain using "intelligent" prioritization of the Gauss-Seidel estimation sequence. Expanded Abstr., 74th Ann. Internat. SEG Mtg., Denver: 2160–2163.

Perez, D.O., Velis, D.R. and Sacchi, M.D., 2013. High-resolution prestack seismic inversion using a hybrid FISTA least-squares strategy. Geophysics, 78: R185-R195.

Rodriguez, I. V., Bonar, D. and Sacchi, M.D., 2012. Microseismic data denoising using a 3C group sparsity constrained time-frequency transform. Geophysics, 77: V21-V29.

Sacchi, M.D. and Ulrych, T., 1995. High-resolution velocity gathers and offset space reconstruction. Geophysics, 60: 1169–1177.

Sacchi, M.D., Verschuur, D.J. and Zwartjes, P.M., 2004. Data reconstruction by generalized deconvolution. Expanded Abstr., 74th Ann. Internat. SEG Mtg., Denver: 1989–1992.

Scales, J., Gersztenkorn, A. and Treitel, S., 1988. Fast *lp* solution of large sparse, linear systems: Application to seismic travel time tomography. J. Comput. Phys., 75: 314-333.

Thorson, J.R. and Claerbout, J.F., 1985. Velocity-stack and slant-stack stochastic inversion. Geophysics, 50: 2727-2741.

Trad, D., Ulrych, T. and Sacchi, M.D., 2002. Accurate interpolation with high-resolution time-variant Radon transforms. Geophysics, 67: 644–656.

Trad, D., Ulrych, T. and Sacchi, M.D., 2003. Latest views of the sparse Radon transform. Geophysics, 68: 386–399.

Verschuur, E., Vrolijk, J.W. and Tsingas, C., 2012. 4D reconstruction of wide azimuth data (WAZ) using sparse inversion of hybrid Radon transforms. Expanded Abstr., 82nd Ann. Internat. SEG Mtg., Las Vegas: 1-5.

Wang, H., Chen, S., Ren, H., Liang, D., Zhou, H. and She, D., 2015. Seismic data two-step recovery approach combining sparsity-promoting and hyperbolic Radon transform methods. J. Geophys. Engineer., 12: 465-476.

Wang, J., Ng, M. and Perz, M., 2010. Seismic data interpolation by greedy local Radon transform. Geophysics, 75: WB225-WB234.

Wang, Y., 2002. Antialiasing conditions in the delay-time Radon transform. Geophys. Prosp., 50: 665-672.

Wang, Y., 2003a. Multiple attenuation: coping with the spatial truncation effect in the Radon transform domain. Geophys. Prosp., 51: 75-87.

Wang, Y., 2003b. Sparseness-constrained least-squares inversion: application to seismic wave propagation. Geophysics, 68: 1633-1638.

Xue, Y., Ma, J. and Chen, X., 2014. High-order sparse Radon transform for AVO-preserving data reconstruction. Geophysics, 79: V13-V22.

Xue, Y., Yang, J., Ma, J. and Chen, Y., 2016. Amplitude-preserving nonlinear adaptive multiple attenuation using the high-order sparse Radon transform. J. Geophys. Engineer., 13: 207-219.

Yilmaz, O., 1989. Velocity-stack processing. Geophys. Prosp., 37: 357-382.

Yu, Z., Ferguson, J., McMechan, G. and Anno, P., 2007. Wavelet-Radon domain dealiasing and interpolation of seismic data. Geophysics, 72: V41-V49.

Yuan, S., Wang, S., Luo, C. and He, Y., 2015. Simultaneous multitrace impedance inversion with transform-domain sparsity promotion. Geophysics, 80: R71–R80.

Zhang, L., Wang, Y., Zheng, Y. and Chang, X., 2015. Deblending using a high-resolution Radon transform in a common midpoint domain. J. Geophys. Engineer., 12: 167-174.

Zhang, Y. and Lu, W., 2014a. 2D and 3D prestack seismic data regularization using an accelerated sparse time-invariant Radon transform. Geophysics, 79: V165-V177.

Zhang, Y. and Lu, W., 2014b. A 3D constrained hybrid linear-pseudo hyperbolic Radon transform. Expanded Abstr., 84th Ann. Internat. SEG Mtg., Denver: 4350-4354.