SEISMIC DATA DENOISING USING DOUBLE SPARSITY DICTIONARY AND ALTERNATING DIRECTION METHOD OF MULTIPLIERS

LIANG ZHANG[1,2], LIGUO HAN[1], AO CHANG[1], JINWEI FANG[3], PAN ZHANG[1], YONG HU[1] and ZHENGGUANG LIU[2]

[1] *College of Geo-exploration Science and Technology, Jilin University, Changchun 130026, P.R. China. 175050483@qq.com*
[2] *Institute of Geosciences and Info-Physics, Central South University, Changsha, Hunan 410083, P.R. China.*
[3] *CNPC Key Lab of Geophysical Exploration, China University of Petroleum-Beijing, Beijing 102249, P.R. China.*

(Received August 3, 2018; revised version accepted October 11, 2019)

ABSTRACT

Zhang, L., Han, L.G., Chang, A., Fang, J.W., Zhang, P., Hu, Y. and Liu, Z.G., 2020. Seismic data denoising using double sparsity dictionary and alternating direction method of multipliers. *Journal of Seismic Exploration*, 29: 49-71.

Recently, the dictionary learning plays a more and more important role in seismic data denoising. Compared with the fixed-basis transform (e.g., Fourier transform, wavelet transform, curvelet transform, contourlet transform and shearlet transform), the denoising of dictionary learning is better because of adaptive sparse representation of seismic data. However, dictionary learning often produces artifacts due to no prior-constraint structural information. In this paper, we propose a new denoising approach, which has double sparsity and combines the advantage of fixed-basis transform and dictionary learning. The whole work-flow of the new denoising approach is as follows. Firstly, we can obtain sparse coefficients of seismic data via shearlet transform. Secondly, sparse coefficients are divided into some suitable size blocks which are regarded as training sets. Thirdly, the alternating direction method of multipliers (ADMM) is used in sparse coding to update dictionary coefficients. Then, the data-driven tight frame (DDTF) is used in dictionary updating to update dictionary atoms. Again, the ADMM is used to resolve the convex optimization problem, and we reshape output blocks to obtain new sparse coefficients. Finally, the hard-thresholding and inverse shearlet transform are applied to new sparse coefficients to achieve denoising. The synthetic data and field data experiments show that the new denoising approach obtain better result than fixed-basis transform and dictionary learning. In conclusion, the new denoising approach can attenuate artifacts and improve the quality of seismic data denoising.

KEYWORDS: seismic data denoising, shearlet transform, DDTF, ADMM.

0963-0651/20/$5.00

INTRODUCTION

In seismic acquisition, the seismic data often suffers random noise due to environmental factors (e.g., the warble, defoliation, gentle breeze, geophones internal interfere and human speaking). However, the quality of seismic data is very important to the subsequent seismic process, such as seismic migration, inversion, amplitude-versus-angle analysis, interpolation and interpretation (Bunks et al., 1995; Sacchi and Liu, 2005; Hunt et al., 2010; Chen et al., 2015). In other words, the seismic data denoising is a necessary step in seismic data process. Over the past few decades, there are kinds of seismic data denoising approaches applied to random noise attenuation, such as principal components analysis (Hagen, 1982), median filter (Liu et al., 2008), singular value decomposition (Ursin and Zheng, 1985) and sparse representation (Liu et al., 2016). Sparse representation is a highly effective denoising approach, which extracts the characteristic and grasps the information of seismic data via a few sparse coefficients. By analyzing the characteristic of sparse coefficients, we can separate the random noise and complete the denoising (Wright et al., 2009). Owing to the good performance in denoising, Sparse representation becomes more and more popular in many fields besides seismic exploration.

Generally, sparse representation can be divided into two categories: the fixed-basis transform and the dictionary learning. The fixed-basis transform, as the name suggests, is used to obtain sparse coefficients by a fixed-basis function, whereas the dictionary learning is used to obtain sparse coefficients by training. Varieties of fixed-basis transforms have been proposed in the literature for seismic data denoising and interpolation (Vassiliou and Garossino, 1998; Liu and Fomel, 2013; Wu and Castagna, 2017). Gülünay (2003) solves the seismic data missing problem by Fourier transform. Xu et al. (2010) achieve the high dimension interpolation for seismic data via Fourier transform. As the whole time transform, Fourier transform has the defect in time-frequency analysis. Although the short-time Fourier transform solves this deficiency, the time window can not be adjusted adaptively. In order to enhance the accuracy of time-frequency analysis, an adaptive time window transform called wavelet is proposed (Gaci, 2014; Liu et al., 2016; Anvari et al., 2017). Mousavi and Langston (2016) combine continuous wavelet transform with hybrid block thresholding for seismic data denoising. Wang et al. (2017) introduce a wavelet family which refers to generalized beta wavelets for seismic time-frequency analysis. Wavelet has a perfect performance for 1D seismic data but degrades in high dimension. Meanwhile, the more complex events are, the more sparse coefficients wavelet produce. Therefore, the hyper-wavelet transform is developed on the basis of wavelet for the sparse representation of high dimension seismic data (Li and Gao, 2013; Gan et al., 2015; Cao et al., 2015; Karbalaali et al., 2017; Xue et al., 2017). Chen et al. (2014) suppress the blending interference noise in the seislet domain. Zhuang et al. (2014) use time-frequency peak filtering algorithm in the radon domain to attenuate random noise. Zhao et al. (2016) realize seismic random noise attenuation by contourlet and time-frequency peak filtering algorithm. Yang et al. (2017) propose a seismic data denoising

and interpolation approach which relates to curvelet. Liu et al. (2018) apply the shearlet to 2D and 3D seismic data interpolation.

What calls for special attention is that the nature of the signal is varied in form. Using only single fixed-basis is hard to describe different signals well. Thus, there is still room for improvement in the sparse representation of seismic data. Compared with the fixed-basis, dictionary learning provides the adaptive basis to match different signals that it trains the dictionary from seismic data itself with an optimization algorithm. A case in point is the K-singular value decomposition (KSVD). When KSVD trains the dictionary column by column, correlations between data and dictionary are disclosed (Aharon et al., 2006). Hou et al. (2018) propose a seismic data denoising and interpolation approach via KSVD. Nevertheless, KSVD suffers high computational complexity. To accelerate the running speed of dictionary learning, some improved algorithms are proposed (Cai et al., 2014; Nalla and Chalavadi, 2015; Tong et al., 2016; Chen, 2017). DDTF is one of the improved algorithm for KSVD (Zhao and Du, 2017). Owing to striking a balance between speed and accuracy, DDTF is relevant in seismic signal processing (Yu et al., 2015; Siahsar et al., 2017; Zhao et al., 2017). Liang et al. (2014) introduce a seismic data restoration approach via DDTF. Yu et al. (2016) combine the DDTF with Monte Carlo to achieve seismic data denoising and interpolation.

However, dictionary learning often produces artifacts due to no prior-constraint structural information. The double sparsity dictionary (DSD) combines the advantage of fixed-basis transform and dictionary learning, which provides the two-level sparsity for model and attenuates artifacts. (Rubinstein et al., 2010; Ophir et al., 2011). Zhu et al. (2015) combine the KSVD and wavelet to attenuate artifacts of dictionary learning.

In this paper, we propose a new denoising approach, which combines the DDTF with shearlet to form a DSD. The shearlet provides the first sparsity layer while DDTF provides the second sparsity layer. So DSD gets the benefits of both sparse basis and becomes more robust. In addition, the ADMM algorithm is applied to sparse coding and sparsity constraint to be more accurate (Ramaswami et al., 2017). Moreover, both the synthetic data and field data experiments verify the validity of the proposed approach.

THEORY

DSD is based on DDTF and ADMM

In this section, we will give a brief introduction to DSD, which is based on DDTF and ADMM for seismic data denoising. The seismic data denoising via DSD can be formulated as the following optimization problem (Chen et al., 2016):

$$\arg\min_{D,A}\left\|Y-\Phi^{-1}D^{T}A\right\|_{F}^{2},s.t.\begin{cases}\left\|a_{k}\right\|_{0}\le T_{0}\\ D^{T}D=I\end{cases}, \tag{1}$$

where D, A and Φ are matrices standing for dictionary, dictionary coefficients and fixed-basis, respectively. Y is the noise data, a_k is the k-th column of A, T_0 is the bound of nonzero entry in a_k, I denotes the identity matrix, the symbol $\|\bullet\|_F$ denotes the F norm (Frobenius norm), and $\|\bullet\|_0$ indicates l_0 norm (the number of nonzero in the matrix). Furthermore, the superscript –1 and T denote the inverse and transpose, respectively.

The DSD can be divided into two sections. One is sparse coding and the other is dictionary updating. In sparse coding, we input discrete cosine transform (DCT) dictionary as D, after that, we overlook the constant and rewrite eq. (1) as

$$\tilde{a}_k = \arg\min_{a_k} \left\| y_k - \Phi^{-1} D^T a_k \right\|_2^2, s.t. \|a_k\|_0 \le T_0, \tag{2}$$

where $\tilde{a}_k$ indicates the calculation of eq. (2) and y_k denotes the k-th column of Y. Different form Chen et al. (2016), we relax the l_0 constraint problem into l_1 constraint problem in eq. (2) (Donoho, 2006) and rewrite it in a non-constraint form via Lagrange multiplier method (Bertsekas, 1999).

$$\tilde{a}_k = \arg\min_{a_k} \frac{1}{2} \left\| y_k - \Phi^{-1} D^T a_k \right\|_2^2 + \lambda \|a_k\|_1, \tag{3}$$

where λ indicates the regularization parameter to control sparsity (the bigger λ is, the sparser $\tilde{a}_k$ is) and $\|\bullet\|_1$ denotes l_1 norm (the same function as $\|\bullet\|_0$ to produce sparsity). Eq. (3) is a convex optimization problem, so we use ADMM to complete sparse coding in this paper. Here, we obtain the i-th step iteration referring to Boyd et al. (2011):

$$\begin{cases} (a_k)^i = \left[\left(\Phi^{-1} D^T \right)^T \Phi^{-1} D^T + \mu I \right]^{-1} \left[\left(\Phi^{-1} D^T \right)^T y_k + \mu \left(\beta^{i-1} - \Lambda^{i-1} \right) \right] \\ \beta^i = soft\left[(a_k)^i + \Lambda^{i-1}, \frac{\lambda}{\mu} \right] \\ \Lambda^i = \Lambda^{i-1} + (a_k)^i - \beta^i \end{cases}, \tag{4}$$

where β is an auxiliary variable, μ denotes penalty parameter and it is positive to balance the augmented quadratic term, Λ is the Lagrange multiplier, which is a zero matrix at the beginning of iteration, and $soft(\bullet)$ is the soft-thresholding function, defined as follows:

$$soft(x_0, T_s) = \begin{cases} x_0 - T_s, x_0 > T_s \\ 0, -T_s \le x_0 \le T_s \\ x_0 + T_s, x_0 < T_s \end{cases} . \quad (5)$$

At the end of sparse coding, we put all of the updated dictionary coefficient vectors column by column to form a matrix $\tilde{A}$. After that, the proposed approach goes into dictionary updating. Corresponding to Cai et al. (2014), the dictionary updating can be given by:

$$\tilde{D} = \arg\min_D \left\| \Phi Y - D^T \tilde{A} \right\|_F^2, s.t. D^T D = I, \quad (6)$$

where $\tilde{D}$ indicates the calculation of eq. (6), we fix $\tilde{A}$ and decompose the cost function $\left\| \Phi Y - D^T \tilde{A} \right\|_F^2$ as follows:

$$\begin{aligned} &\left\| \Phi Y - D^T \tilde{A} \right\|_F^2 = \sum \left\| \Phi y_k - D^T \tilde{a}_k \right\|_2^2 \\ &= \sum \left((\tilde{a}_k)^T \tilde{a}_k + (\Phi y_k)^T D^T D \Phi y_k - 2(\tilde{a}_k)^T D^T \Phi y_k \right) \\ &= \sum \left((\tilde{a}_k)^T \tilde{a}_k + (\Phi y_k)^T \Phi y_k - 2(D\tilde{a}_k)^T \Phi y_k \right) \\ &= Tr\left(\tilde{A}^T \tilde{A} \right) + Tr\left((\Phi Y)^T \Phi Y \right) - 2Tr\left(D\tilde{A} (\Phi Y)^T \right), \end{aligned} \quad (7)$$

where the symbol $Tr(\bullet)$ represents the trace of matrix. That is, eq. (6) is equivalent to:

$$\tilde{D} = \arg\max_D Tr\left[D\tilde{A} (\Phi Y)^T \right] s.t. D^T D = I, \quad (8)$$

we can solve the eq. (8) referring to Zou et al. (2006). For the optimization problem (M, N are matrices and have the same size):

$$\tilde{M} = \arg\max_M Tr\left(M^T N \right), s.t. M^T M = I, \quad (9)$$

if the singular value decomposition of N can be expressed as:

$$N = U \Sigma V^T, \quad (10)$$

we will obtain the result of eq. (9):

$$\tilde{M} = U V^T, \quad (11)$$

according to eqs. (9)-(11), the result of eq. (8) becomes:

$$\begin{cases} \Phi Y\left(\widetilde{A}\right)^T = U_1 \Sigma_1 V_1^T \\ \widetilde{D} = U_1 V_1^T \end{cases}, \tag{12}$$

where $\Phi Y\left(\widetilde{A}\right)^T = U_1 \Sigma_1 V_1^T$ denotes the singular value decomposition of $\Phi Y\left(\widetilde{A}\right)^T$. When eq. (12) is over, we obtain the updated dictionary $\widetilde{D}$. In view of the correspondence, we implement eq. (3) once again by using ADMM and replace D with $\widetilde{D}$. Similarly, we obtain new dictionary coefficients $\hat{A}$. Finally, we apply the hard-thresholding and inverse shearlet transform to achieve denoising:

$$\begin{cases} \widetilde{S} = \left(\widetilde{D}\right)^T \hat{A} \\ \widetilde{X} = \Phi^{-1}\left[hard\left(\widetilde{S}, T_h\right)\right] \end{cases}, \tag{13}$$

where $\widetilde{S}$, $\widetilde{X}$ and T_h stand for new shearlet coefficients, denoising and hard-thresholding, respectively. The symbol $hard(\bullet)$ is hard-thresholding function, defined as follow:

$$hard\left(x_0, T_h\right) = \begin{cases} x_0, \left|x_0\right| \geq T_h \\ 0, \left|x_0\right| < T_h \end{cases}. \tag{14}$$

Shearlet transform

In this section, we will roughly show the mathematical background of the shearlet transform. Due to its sensitive directivity, we choose shearlet transform as Φ to achieve the excellent denoising. More detail about shearlet transform can be found on Easley et al. (2008).

In the 2-dimension square integrable space, we define a shear matrix N and a scale matrix M as follows (Häuser and Steidl, 2013):

$$N = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}, n \in R \qquad M = \begin{bmatrix} m & 0 \\ 0 & \sqrt{m} \end{bmatrix}, m \in R^+, \tag{15}$$

where n is the shear parameter and m is the scale parameter. Here, We suppose a basis function ψ and its Fourier form Ψ. Besides, Ψ conforms to:

$$\Psi(\omega)=\Psi(\omega_1,\omega_2)=\Psi_1(\omega_1)\Psi_2\left(\frac{\omega_2}{\omega_1}\right),\omega,\omega_1,\omega_2\in R,\omega_1\neq 0, \tag{16}$$

similarly, Ψ_1 is the Fourier form of basis function ψ_1 and Ψ_2 is the Fourier form of basis function ψ_2. In particular, Ψ_1 is the Meyer wavelet and the interval of Ψ_1 is $\Psi_1\subset[-2,-\frac{1}{2}]\cup[\frac{1}{2},2]$. While Ψ_2 is the bump function and the interval of Ψ_2 is $\Psi_2\subset[-1,1]$ (Yi et al., 2009).

At this time, ψ becomes the affine system with composite dilation $\psi_{m,n,j}(x)$ by changing scale parameter m, shear parameter n and translation parameter j:

$$\psi_{m,n,j}(x)=m^{-\frac{3}{4}}\psi\left(M_m^{-1}N_n^{-1}(x-j)\right),m\in R^+,n\in R,j\in R^2, \tag{17}$$

therefore, the sparse representation of seismic data via shearlet transform can be expressed as dot product (Kong and Peng, 2015)

$$S=\left\langle Y,\psi_{m,n,j}\right\rangle, \tag{18}$$

we use $\Phi=\psi_{m,n,j}$ to follow the correspondence of letter in this paper:

$$S=\Phi Y, \tag{19}$$

where S indicates shearlet coefficients and $\langle\bullet\rangle$ denotes dot product. Fig. 1 shows the frequency support of the shearlet transform.

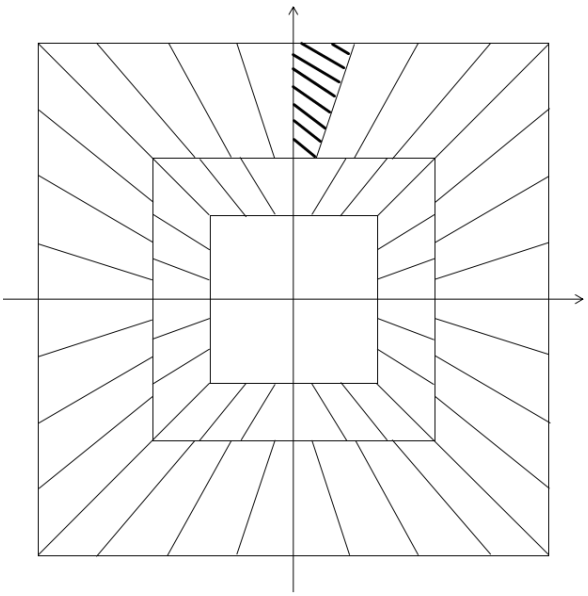


Fig. 1. The frequency support of the shearlet transform.

Now, we can summarize the whole work-flow of the proposed approach as follows:

Algorithm: DSD is based on DDTF and ADMM

Input:
noise data Y
ADMM parameters λ, μ, Λ, β (λ and μ should be tuned, Λ and β are zero matrices at the beginning)
original dictionary D
hard-thresholding T_h (T_h should be tuned)
Output:
updated dictionary $\widetilde{D}$
new shearlet coefficients $\widetilde{S}$
denoising $\widetilde{X}$
Step:
1: Obtain shearlet coefficients S via eq. (23).
2: Choose the appropriate window in shearlet coefficients of each layer to get training sets S_w.
3: Complete sparse coding via eq. (8) and get updated dictionary coefficients $\widetilde{A}$.
4: Complete dictionary updating via eq. (16) and get updated dictionary $\widetilde{D}$.
5: Obtain new dictionary coefficients $\hat{A}$ by solving eq. (8) again.
6: Apply eq. (17) and reshape new training sets $\widetilde{S}_w$ to get denoising $\widetilde{X}$.

The choice of parameters depends on experience. Here, we will introduce how we choose. There are two inner loops in the proposed approach: step 3 and step 5. For step 3, we prefer to set up a fast iteration. It means i should be small. So λ and μ are both 1 in step 3 for all tests. For step 5, we want to obtain a high-accuracy output. Thus we use the variable λ and μ in step 5 for all tests: $\lambda = \frac{\inf\left(\Phi^{-1} D^T \left(y_k\right)^T\right)}{10}$, $\mu = \frac{\inf\left(\left(\beta^i\right)^T\right)}{10}$. Where $\inf(\bullet)$ denotes the sum of the absolute value of the maximum row. Then T_h drops 90% small shearlet coefficients in all tests. And we choose DCT as the original dictionary.

RESULTS AND EXAMPLES

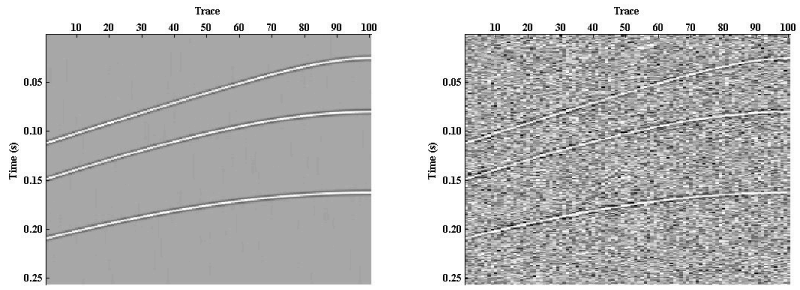
In this section, we apply the proposed approach on synthetic data and field data to test the performance. As a comparison, we also test the performance of shearlet transform, DDTF, DSD-H (the only difference with the proposed approach is that DSD-H uses hard-thresholding to solve eq. (2) while the proposed approach uses ADMM to solve eq. (3)) and illustrate every result in details. To quantify the performance, we define the signal-to-noise ratio (SNR) as follow:

$$SNR = 10\log_{10}\left(\frac{\|X\|_F^2}{\left\|X - \tilde{X}\right\|_F^2}\right), \tag{20}$$

where X indicates the clean data. Inspired by Tang et al. (2012), we use the appropriate window (8×8 in all tests) in data to acquire training sets applied in all examples to improve the robustness of dictionary.

Synthetic data

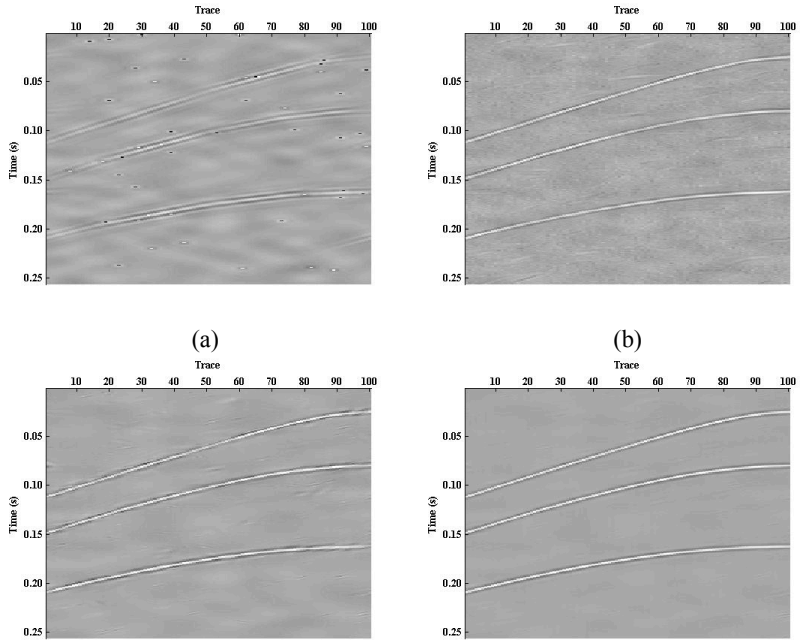
The synthetic data is shown in Fig. 2(a), which contains 100 traces, 256 time samples per trace and the temporal interval is 0.001 s. We add the Gaussian white noise to synthetic data, and the result is shown in Fig. 2(b).



(a) (b)

Fig. 2. Synthetic data: (a) Clean synthetic data; (b) Noise synthetic data.

According to Fig. 2 above, we can see the synthetic data is composed of three high energy events. The Gaussian white noise disturbs events but we still can identify them. We apply shearlet transform, DDTF, DSD-H, the proposed approach on Fig. 2(b) and obtain the denoising shown in Fig. 3.



(a) (b)

(c) (d)

Fig. 3. Synthetic data denoising: (a) Shearlet transform denoising; (b) DDTF denoising; (c) DSD-H denoising; (d) The proposed approach denoising.

In Fig. 3, shearlet transform remains some spot noise. It is noteworthy that the shearlet transform strongly weakens three high energy events. The DDTF still remains much random noise. Furthermore, the DDTF produces some artifacts although it slightly weakens three high energy events. The DSD-H suffers a little artifacts and attenuates much random noise. What is more, DSD-H protects three high energy events well. The proposed approach has little noise and artifacts. Similarly, the proposed approach hardly damages three high energy events. Overall, the proposed approach has the best performance in four methods. Table 1 shows the SNR of synthetic data denoising in Fig. 3.

Table 1. The SNR of synthetic data denoising in Fig. 3.

Method	Noise data (dB)	Denoising (dB)
Shearlet transform	3.0000	12.3398
DDTF	3.0000	15.2621
DSD-H	3.0000	17.1172
The proposed approach	3.0000	20.3204

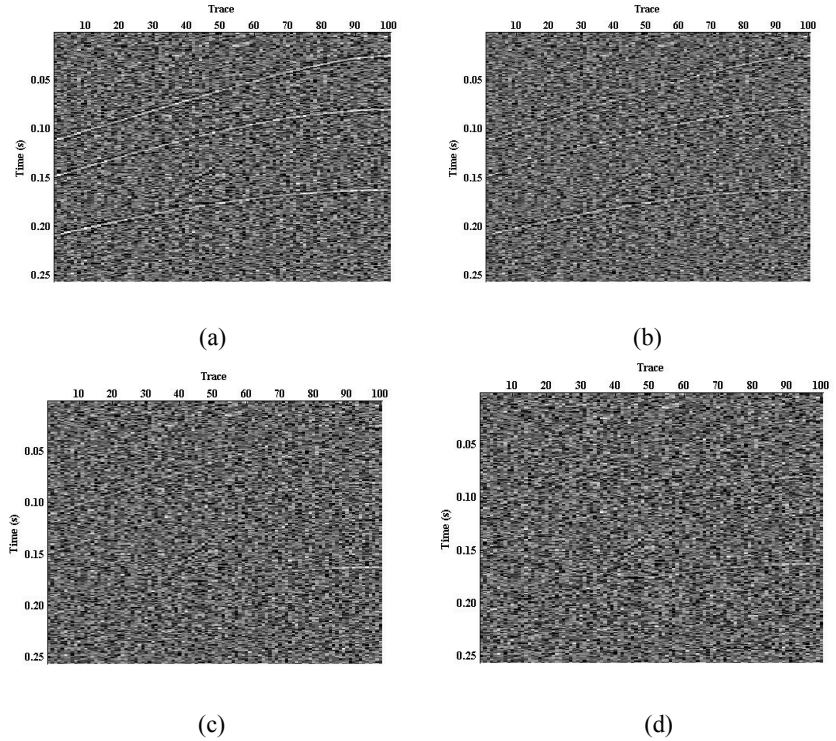
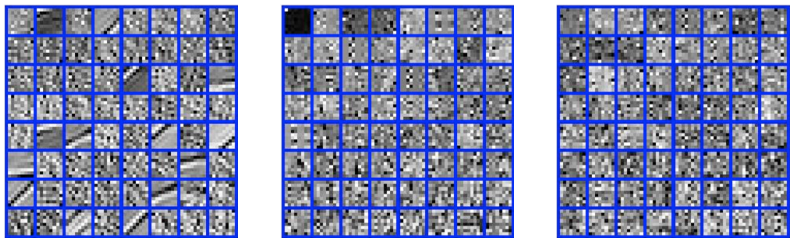


Fig. 4. The discard of synthetic data: (a) The discard of shearlet transform; (b) The discard of DDTF; (c) The discard of DSD-H; (d) The discard of the proposed approach.

As is evident from Table 1, the proposed approach has the best performance compared with shearlet transform, DDTF, DSD-H. Fig. 4 shows the discard of synthetic data by four methods.

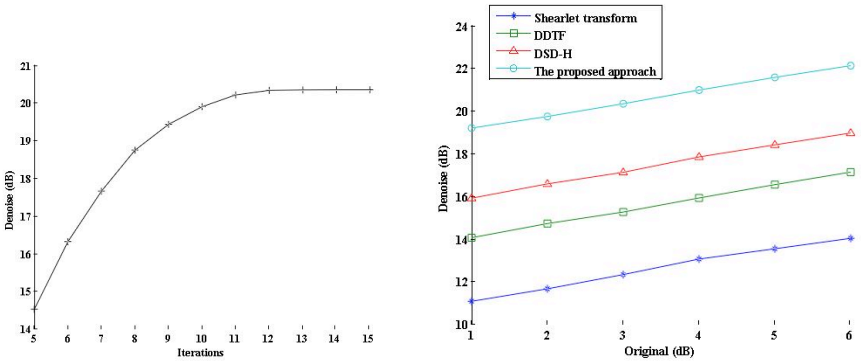
As is clearly seen from Fig. 4, four methods complete denoising with varying degrees of damaging three high energy events while the proposed approach has the least damage. Fig. 5 is the dictionary of synthetic data. Subjected to the space of paper, we only present the finest shearlet coefficients dictionary in all examples.



(a) (b) (c)

Fig. 5. Synthetic data dictionary: (a) DDTF dictionary; (b) DSD-H dictionary (the finest shearlet coefficients); (c) The proposed approach dictionary (the finest shearlet coefficients).

We note down the SNR for every iteration and show it in Fig. 6(a). With the increases of iteration, the growth rate of SNR gradually decreases and it increases hardly in 12 times. Moreover, to make the test results more convincing, four methods are applied in noise synthetic data which range in SNR from 1 to 6, and we show the SNR of denoising in Fig. 6(b). We can see the proposed approach has the best performance throughout.

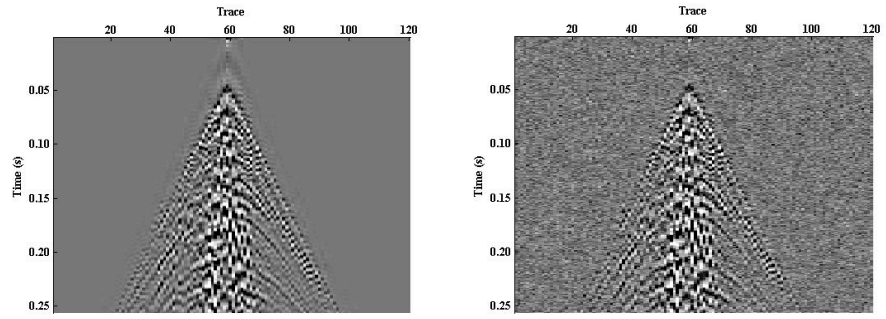


(a) (b)

Fig. 6. Synthetic data SNR curve: (a) The SNR of different iterations; (b) The SNR of four methods.

Complex data

The proposed approach has a good performance in synthetic data, but it is too simple and unknown in field data. For consequent tests, we choose more complex data and field data. The Complex data is shown in Fig. 7(a), which contains 120 traces, 256 time samples per trace and the temporal interval is 0.001 s. We add the Gaussian white noise to complex data, and the result is shown in Fig. 7(b).



(a) (b)

Fig. 7. Complex data: (a) Clean complex data; (b) Noise complex data.

As can be seen from Fig. 7, the complex data is composed of very complex events with different energies. Influenced by Gaussian white noise, the low energy events are disrupted. Similarly, we apply shearlet transform, DDTF, DSD-H, the proposed approach on Fig. 7(b). Afterwards, the denoising of complex data by four methods is shown in Fig. 8.

In Fig. 8, shearlet transform causes the attenuation of low energy events. Furthermore, there is much spot noise in the shearlet transform. DDTF saves the low energy events well. Nevertheless, its denoising remains the most random noise among four methods, especially for artifacts. Compared with shearlet transform and DDTF, DSD-H has a better performance. The random noise and artifacts are less, and the low energy events are protected better. It is obvious that the most random noise is removed by applying the proposed approach. Moreover, low energy events are kept well and artifacts hardly exist in the denoising. In conclusion, the proposed approach has the best performance in four methods. Table 3 shows the SNR of complex data denoising in Fig. 8.

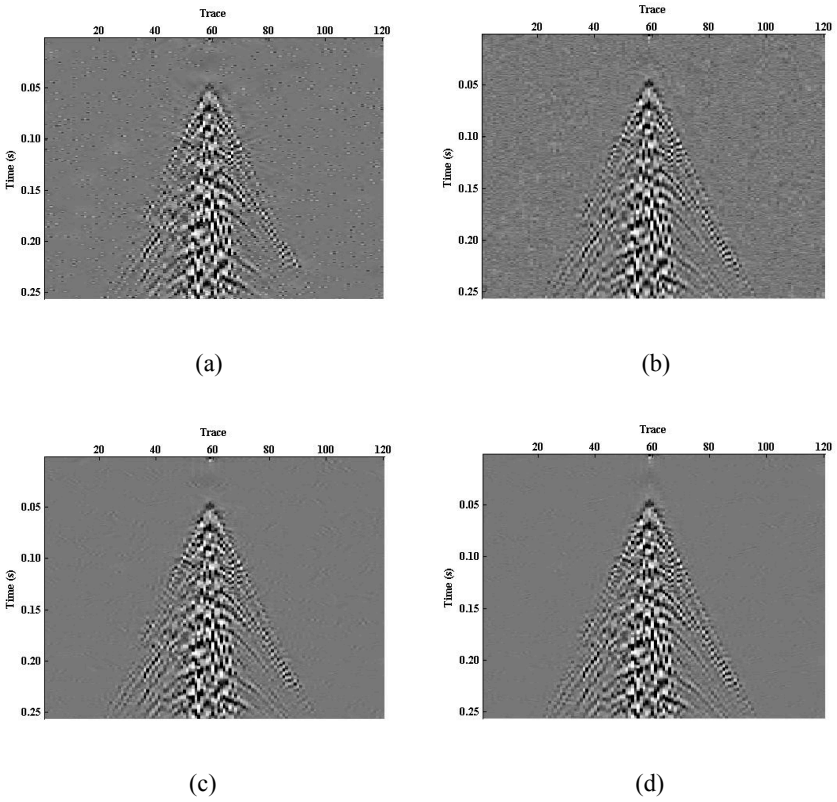


Fig. 8. Complex data denoising: (a) Shearlet transform denoising; (b) DDTF denoising; (c) DSD-H denoising; (d) The proposed approach denoising.

Table 3. The SNR of complex data denoising in Fig. 8.

Method	Noise data (dB)	Denoising (dB)
Shearlet transform	3.0000	6.5939
DDTF	3.0000	8.1835
DSD-H	3.0000	10.9353
The proposed approach	3.0000	13.7304

As is evident from Table 3, we can come to the conclusion which is the same as the synthetic data test. Fig. 9 demonstrates the discard of complex data by shearlet transform, DDTF, DSD-H and the proposed approach.

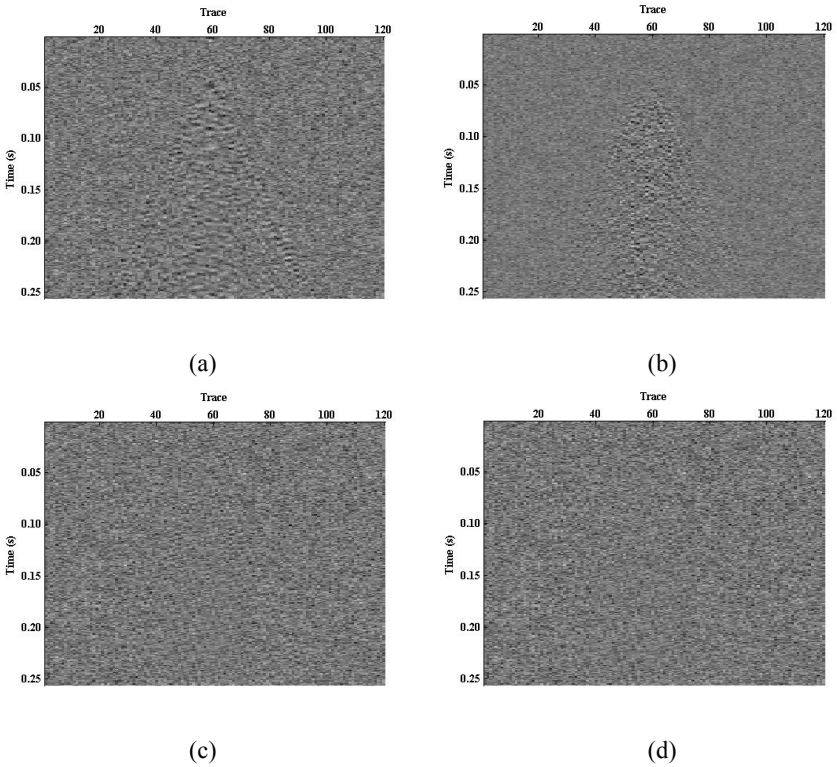


Fig. 9. The discard of complex data: (a) The discard of shearlet transform; (b) The discard of DDTF; (c) The discard of DSD-H; (d) The discard of the proposed approach.

Fig. 9 shows the damage of complex data denoising for four methods. Like the synthetic data test, the proposed approach has the least damage. Fig.10 is the dictionary of complex data.

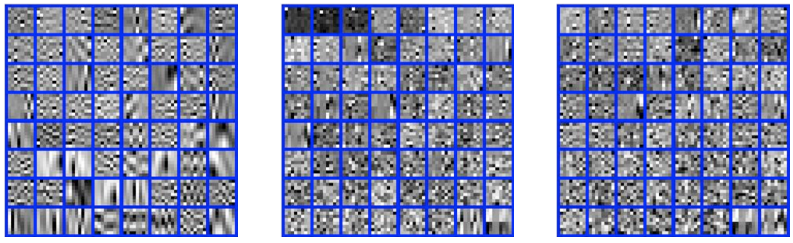
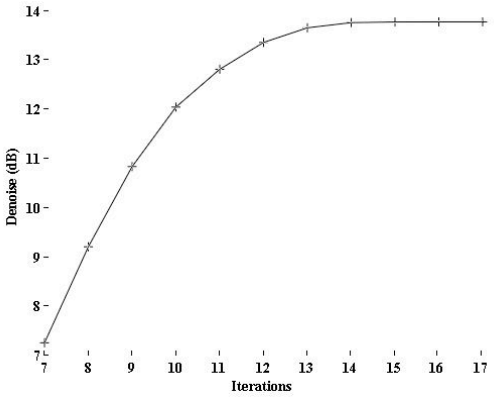
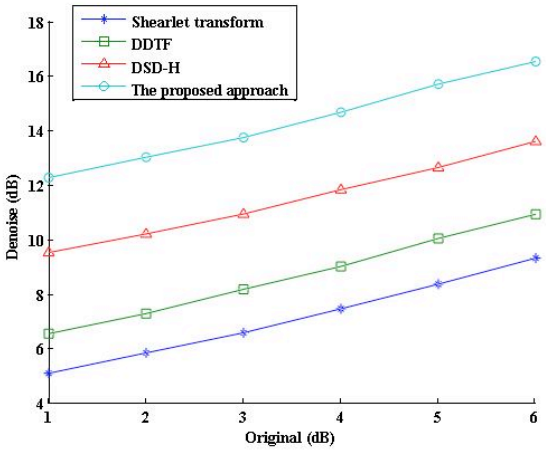


Fig. 10. Complex data dictionary: (a) DDTF dictionary; (b) DSD-H dictionary (the finest shearlet coefficients); (c) The proposed approach dictionary (the finest shearlet coefficients).

Similarly, we show the denoising SNR of different iterations and different initial SNR in Fig. 11 to make the test results more persuasive. Also, we can come to the conclusion which is same as the synthetic data test. The difference is the growth rate of SNR increases hardly in 15 times for complex data.



(a)



(b)

Fig. 11. Complex data SNR curve: (a) The SNR of different iterations; (b) The SNR of four methods.

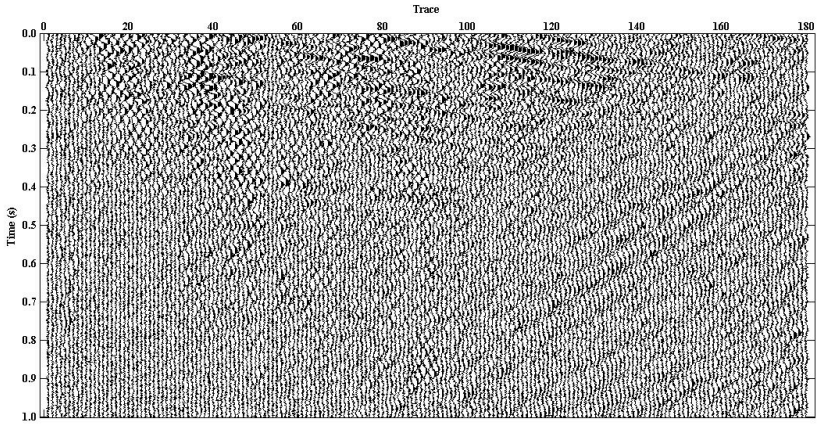


Fig. 12. Field data.

Field data

The field data is a plot of a huge field data, and we show it in Fig. 12, which contains 180 traces, 2000 time samples per trace and the temporal interval is 0.001 s.

The field data contains random noise but we do not know the intensity. Due to the unknown intensity of random noise, we make a rule to the stop the iteration. When the relative error of adjacent iteration less than 10%, the iteration is stopped. Similarly, we apply shearlet transform, DDTF, DSD-H, the proposed approach on Fig. 12 and the denoising is shown in Fig. 13.

As we can see in Fig. 13, shearlet transform weaken the energy of events. In addition, much random noise is remained. Compared with shearlet transform, DDTF remains less random noise and more energy of events. The DSD-H has almost the same performance with the proposed approach in that they both attenuate much random noise. However, the proposed approach saves low energy events better than DSD-H. All in all, the proposed approach has the best performance in four methods. The same as synthetic data and complex data tests, we show the discard of field data in Fig. 14, although it is hard to see the damage of events. We also present the dictionary of field data in Fig. 15.

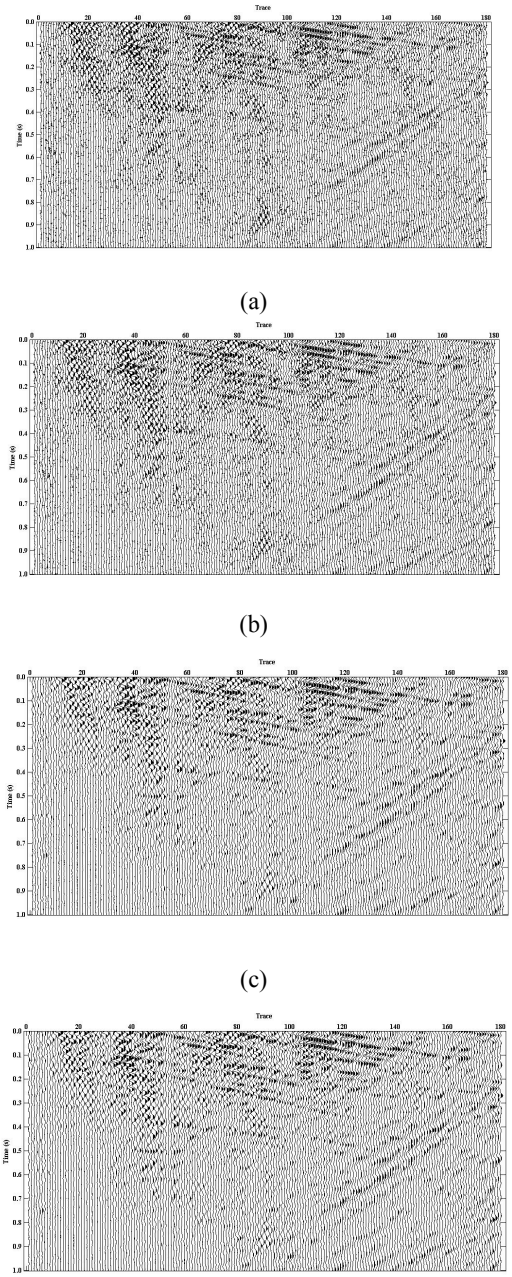
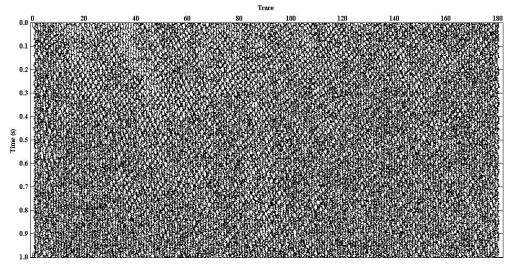
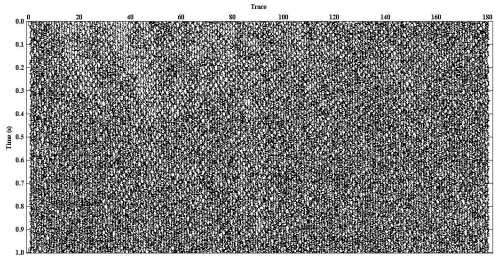


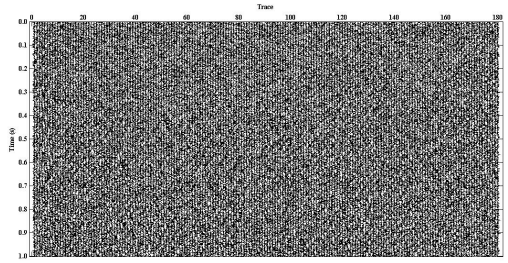
Fig. 13. Field data denoising: (a) Shearlet transform denoising; (b) DDTF denoising; (c) DSD-H denoising; (d) The proposed approach denoising.



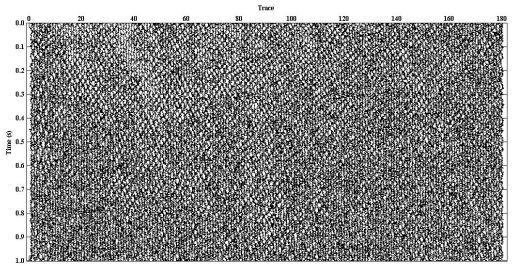
(a)



(b)

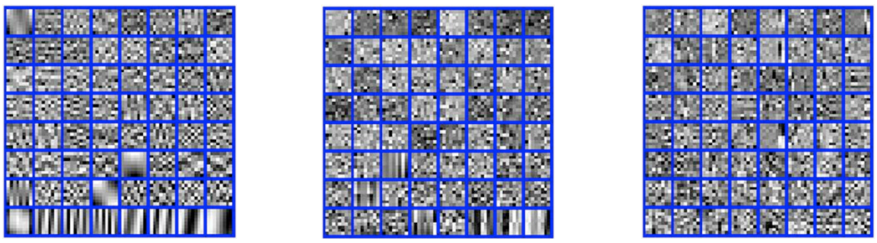


(c)



(d)

Fig. 14. The discard of field data: (a) The discard of shearlet transform; (b) The discard of DDTF; (c) The discard of DSD-H; (d) The discard of the proposed approach.



(a) (b) (c)

Fig. 15. Field data dictionary: (a) DDTF dictionary; (b) DSD-H dictionary (the finest shearlet coefficients); (c) The proposed approach dictionary (the finest shearlet coefficients).

CONCLUSIONS

In this paper, a novel seismic data denoising approach, based on DDTF and ADMM with double sparsity, is presented. Starting from shearlet transform, we can obtain shearlet coefficients of noise data. To improve the robustness of dictionary, we choose the appropriate window in the shearlet coefficients of each layer to get training sets. The ADMM is used in sparse coding to update dictionary coefficients. And then, we update the dictionary atoms in the framework of DDTF. In view of the correspondence, the ADMM is used again to get new training sets. Next, we reshape new training sets to obtain new shearlet coefficients. Finally, we apply the hard-thresholding and inverse shearlet transform to achieve denoising. In the section of examples and results, the proposed approach achieves state-of-the-art results compared to shearlet transform, DDTF, DSD-H. Regardless of synthetic or field data, simple or complex model, the proposed approach has an outstanding performance.

However, there are some drawbacks in the proposed approach, such as computational time and parameter tuning. As many dictionary learning methods, the proposed approach suffers from time-consuming. Although it is faster than KSVD, the proposed approach is slower than DDTF. A probably solution is to find a more effective algorithm to train the dictionary. Moreover, the multiscale and multidirection of shearlet transform cause fussy parameter tuning. The more decomposition level of shearlet transform is, the more fussy parameter tuning is. To balance the SNR of denoising and the decomposition level of shearlet transform is a worth research. Therefore, our future work will focus on the computational time and parameter tuning to improve the proposed approach.

REFERENCES

Aharon, M., Elad, M. and Bruckstein, A., 2006. $ rm k $-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transact. Sign. Process., 54: 4311-4322.

Anvari, R., Siahsar, M.A.N., Gholtashi, S., Kahoo, A.R. and Mohammadi, ?? 2017. Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation. IEEE Transact. Geosci. Remote Sens., 55: 6574-6581.

Bertsekas, D.P., 1999. Nonlinear Programming. Belmont: Athena Scientific, Boston, MA.

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundat. Trends Mach. Learn., 3: 1-122.

Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G., 1995. Multiscale seismic waveform inversion. Geophysics, 60: 1457-1473.

Cai, J.F., Ji, H., Shen, Z. and Ye, G.B., 2014. Data-driven tight frame construction and image denoising. Appl. Computat. Harm. Analys., 37: 89-105.

Cao, J., Zhao, J. and Hu, Z., 2015. 3D seismic denoising based on a low-redundancy curvelet transform. J. Geophys. Engineer., 12: 566.

Chen, Y., Fomel, S. and Hu, J., 2014. Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization. Geophysics, 79(5): V179-V189.

Chen, Y., Zhang, L. and Mo, L., 2015. Seismic data interpolation using nonlinear shaping regularization. J. Seismic Explor., 24: 327-342.

Chen, Y., Ma, J. and Fomel, S., 2016. Double-sparsity dictionary for seismic noise attenuation. Geophysics, 81(2): V103-V116.

Chen, Y., 2017. Fast dictionary learning for noise attenuation of multidimensional seismic data. Geophys. J. Internat., 209: 21-31.

Donoho, D.L., 2006. For most large underdetermined systems of linear equations the minimal l1-norm near-solution is also the sparsest solution. Communic. Pure Appl. Mathemat., 59: 797-829.

Easley, G.R., Labate, D. and Lim, W.Q., 2008. Sparse directional image representations using the discrete shearlet transform. Appl. Computat. Harm. Analys., 25: 25-46.

Gaci, S., 2014. The use of wavelet-based denoising techniques to enhance the first-arrival picking on seismic traces. IEEE Transact. Geosci. Remote Sens., 58: 4558-4563.

Gan, S., Wang, S., Chen, Y., Zhang, Y., and Jin, Z., 2015. Dealiased seismic data interpolation using seislet transform with low-frequency constraint. IEEE Geosci. Remote Sens. Lett., 12: 2150-2154.

Gülünay, N., 2003. Seismic trace interpolation in the Fourier transform domain. Geophysics, 68: 355-369.

Hagen, D.C., 1982. The application of principal components analysis to seismic data sets. Geoexplor., 20: 93-111.

Häuser, S. and Steidl, G., 2013. Convex multiclass segmentation with shearlet regularization. Internat. J. Comput. Mathemat., 90: 62-81.

Hou, S., Zhang, F., Li, X., Zhao, Q. and Dai, H., 2018. Simultaneous multi-component seismic denoising and reconstruction via K-SVD. J. Geophys. Engineer., 15: 681.

Hunt, L., Downton, J., Reynolds, S., Hadley, S., Trad, D. and Hadley, M., 2010. The effect of interpolation on imaging and AVO: A Viking case study. Geophysics, 75(6): WB265-WB274.

Karbalaali, H., Javaherian, A., Dahlke, S. and Torabi, S., 2017. Channel boundary detection based on 2D shearlet transformation: An application to the seismic data in the South Caspian Sea. J. Appl. Geophys., 146: 67-79.

Kong, D. and Peng, Z., 2015. Seismic random noise attenuation using shearlet and total generalized variation. J. Geophys. Engineer., 12: 1024.
Li, Q. and Gao, J., 2013. Contourlet based seismic reflection data non-local noise suppression. J. Appl. Geophys., 95: 16-22.
Liang, J., Ma, J. and Zhang, X., 2014. Seismic data restoration via data-driven tight frame. Geophysics, 79(3): V65-V74.
Liu, Y., Liu, C. and Wang, D., 2008. A 1D time-varying median filter for seismic random, spike-like noise elimination. Geophysics, 74(1): V17-V24.
Liu, Y. and Fomel, S., 2013. Seismic data analysis using local time-frequency decomposition. Geophys. Prosp., 61: 516-525.
Liu, W., Cao, S. and Chen, Y., 2016a. Seismic time-frequency analysis via empirical wavelet transform. IEEE Geosci. Remote Sens. Lett., 13: 28-32.
Liu, W., Cao, S., Chen, Y. and Zu, S., 2016b. An effective approach to attenuate random noise based on compressive sensing and curvelet transform. J. Geophys. Engineer., 13: 135.
Liu, J., Chou, Y. and Zhu, J., 2018. Interpolating seismic data via the POCS method based on shearlet transform. J. Geophys. Engineer., 15: 852-876.
Mousavi, S.M. and Langston, C.A., 2016. Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding. Bull. Seismol. Soc. Am., 106: 1380-1393.
Nalla, P.R. and Chalavadi, K.M., 2015. Iris classification based on sparse representations using on-line dictionary learning for large-scale de-duplication applications. Springer Plus, 4: 238.
Ophir, B., Lustig, M. and Elad, M., 2011. Multi-scale dictionary learning using wavelets. IEEE J. Select. Topics Sign. Process., 5: 1014-1024.
Ramaswami, S., Kawaguchi, Y., Takashima, R., Endo, T. and Togami, M., 2017. ADMM-based audio reconstruction for low-cost-sound-monitoring. 25th Europ.Sign. Process. Conf. (EUSIPCO). doi:10.23919/EUSIPCO.2017.8081410.
Rubinstein, R., Zibulevsky, M. and Elad, M., 2010. Double sparsity: Learning sparse dictionaries for sparse signal approximation. IEEE Transact. Sign. Process., 58: 1553-1564.
Sacchi, M.D. and Liu, B., 2005. Minimum weighted norm wavefield reconstruction for AVA imaging. Geophys. Prosp., 53: 787-801.
Siahsar, M.A.N., Gholtashi, S., Torshizi, E.O., Chen, W. and Chen, Y., 2017. Simultaneous denoising and interpolation of 3D seismic data via damped data-driven optimal singular value shrinkage. IEEE Geosci. Remote Sens. Lett., 14: 1086-1090.
Tang, G., Ma, J.W. and Yang, H.Z., 2012. Seismic data denoising based on learning-type overcomplete dictionaries. Appl. Geophys., 9: 27-32.
Tong, Q., Sun, Z., Nie, Z., Lin, Y. and Cao, J., 2016. Sparse decomposition based on ADMM dictionary learning for fault feature extraction of rolling element bearing. J. Vibroengineer., 18: 5204-5216.
Ursin, B. and Zheng, Y., 1985. Identification of seismic reflections using singular value decomposition. Geophys. Prosp., 33: 773-799.
Vassiliou, A.A. and Garossino, P., 1998. Time-frequency processing and analysis of seismic data using very short-time Fourier transforms. U.S. Patent 5,850,622.
Wang, Z., Zhang, B., Gao, J., Wang, Q. and Liu H., 2017. Wavelet transform with generalized beta wavelets for seismic time-frequency analysis. Geophysics, 82(4): O47-O56.
Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S. and Ma, Y., 2009. Robust face recognition via sparse representation. IEEE Transact. Pattern Analys. Mach. Intellig., 31: 210-227.

灰小偷 12/18/19 2:00 PM

Comment [1]: I am so sorry that I can not find the IEEE title and the volume. I find this reference in this website: https://ieeexplore.ieee.org/document/8081410
And I use this website's citation function. Is this format right?

Wu, L. and Castagna, J.P., 2017. S-transform and Fourier transform frequency spectra of broadband seismic signals. Geophysics, 82(5): O71-O81.

Xu, S., Zhang, Y. and Lambaré G., 2010. Antileakage Fourier transform for seismic data regularization in higher dimensions. Geophysics, 75(6): WB113-WB120.

Xue, Y., Man, M., Zu, S., Chang, F. and Chen, Y., 2017. Amplitude-preserving iterative deblending of simultaneous source seismic data using high-order Radon transform. J. Appl. Geophys., 139: 79-90.

Yang, H., Long, Y., Lin, J., Zhang, F. and Chen, Z., 2017. A seismic interpolation and denoising method with curvelet transform matching filter. Acta Geophys., 65: 1029-1042.

Yi, S., Labate, D., Easley, G.R. and Krim, H., 2009. A shearlet approach to edge analysis and detection. IEEE Transact. Image Process., 18: 929-941.

Yu, S., Ma, J., Zhang, X. and Sacchi, M.D., 2015. Interpolation and denoising of high-dimensional seismic data by learning a tight frame. Geophysics, 80(5): V119-V132.

Yu, S., Ma, J. and Osher, S., 2016. Monte Carlo data-driven tight frame for seismic data recovery. Geophysics, 81(4): V327-V340.

Zhao, Q. and Du, Q., 2017. Constrained data-driven tight frame for robust seismic data reconstruction. Expanded Abstr., 87th Ann. Internat. SEG Mtg., Houston: 4246-4250.

Zhao, X., Li, Y., Zhuang, G., Zhang, C. and Han, X., 2016. 2-D TFPF based on Contourlet transform for seismic random noise attenuation. J. Appl. Geophys., 129: 158-166.

Zhu, L., Liu, E. and McClellan, J.H., 2015. Seismic data denoising through multiscale and sparsity-promoting dictionary learning. Geophysics, 80(6): WD45-WD57.

Zhuang, G., Li, Y., Liu, Y., Lin, H., Ma, H., and Wu, N., 2014. Varying-window-length TFPF in high-resolution Radon domain for seismic random noise attenuation. IEEE Geosci. Remote Sens. Lett., 12: 404-408.

Zou, H., Hastie, T. and Tibshirani, R., 2006. Sparse principal component analysis. J. Computat. Graphic. Statist., 15: 265-286.