IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON DATA MINING FOR INTERNET OF THINGS

Received November 4, 2019, accepted November 29, 2019, date of publication December 23, 2019,
date of current version February 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2959024

Seismic Random Noise Attenuation Based on PCC Classification in Transform Domain

YU SANG[1], JINGUANG SUN[1], XIANGFU MENG[1], HAIBO JIN[2], YANFEI PENG[1], AND XINJUN ZHANG[1]

[1]School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China
[2]School of Software, Liaoning Technical University, Huludao 125105, China

Corresponding author: Yu Sang (sangyu2008bj@sina.com)

This work was supported in part by the National Science Foundation of China under Grant 61602226 and Grant 61772249, in part by the Ph.D. Startup Foundation of Liaoning Technical University of China under Grant 18-1021, and in part by the Basic Research Project of Colleges and Universities of Liaoning Provincial Department of Education under Grant LJ2017FBL004.

ABSTRACT Random noise attenuation of seismic data is an essential step in the processing of seismic signals. However, as the exploration environment is becoming more and more complicated, the energy of valid signals is weaker and the signal to noise (SNR) is much lower, which brings great difficulty to seismic data processing and interpretation. To this end, we propose an unconventional and effective seismic random noise attenuation method based on proximal classifier with consistency (PCC) in transform domain. Firstly, we analyze various transforms for seismic data from traditional wavelet transform and curvelet transform to emerging non-subsampled shearlet transform (NSST) and non-subsampled contourlet transform (NSCT). And, we select the excellent NSST to decompose the noisy seismic data into different sub-bands of frequency and orientation responses. Secondly, unlike traditional sparse transform based seismic denoising methods that often directly use a thresholding operator and corresponding inverse transform to denoise seismic data, our proposed method employs a superior performance PCC to classify the NSST coefficients of seismic data before thresholding operator. The added step can effectively divide the NSST coefficients into reflected useful signal coefficients and noise-related coefficients, which can preserve the edge of reflected signals and keep the information of events intact as much as possible. In addition, we also introduce an adaptive threshold computing method and a soft-thresholding method to achieve seismic data denoising better. Finally, the experimental results on the typical synthetic example and real seismic data show the superior performance of the proposed method.

INDEX TERMS Seismic data, random noise, attenuation, proximal classifier with consistency (PCC), non-subsampled shearlet transform (NSST).

I. INTRODUCTION

Improving the signal-to-noise ratio of seismic exploration data is a key step in processing seismic data, and suppressing random noise in seismic signals is an important link in improving the signal-to-noise ratio of seismic data. In seismic data acquisition from different sensing equipment or networks [1]–[6], many reflection events are polluted by various types of noise, which can be classified into two categories including coherent and incoherent interference. In this paper, we mainly focus on incoherent interference, aka random noise. Random noise attenuation can play an important role in seismic data processing. It can not only improve the fidelity and stability of migration or inversion, but also can make diffraction reduction in the migrated image or profile and facilitate the interpretation for making accurate oil and gas decisions.

Over the past decades, excellent experts in the field of seismic exploration and signal processing at home and abroad have made unremitting efforts to suppress random noise in seismic data. Most of seismic denoising approaches have been presented such as the initial random noise reduction [7], wavelet transform based seismic denosing [8], random noise attenuation by f-x empirical mode decomposition [9], simultaneous dictionary learning and denoising [10], EMD-seislet transform based denoising [11], random noise suppression via variational mode decomposition [12] and fast dictionary learning for noise attenuation of multidimensional seismic

The associate editor coordinating the review of this manuscript and approving it for publication was Tie Qiu.

data [13], and so on. Therein, the sparse transform based seismic random noise attenuation is one of the most effective and widely used methods. Zhang and Lu [8] applied the wavelet transform to remove noise of seismic data. Neelamani *et al.* [14] applied the curvelet transform to attenuate both incoherent and coherent noise in seismic data. Recently, Chen and Fomel [11] proposed a new EMD-seislet transform to denoise seismic data. Liu *et al.* [15] presented a random noise attenuation approach using curvelet transform and compressive sensing theory. Sang *et al.* [16] proposed a transform domain based seismic data denoising algorithm by using a machine learning classification method. All of these approaches have obtained good denoising effects to a certain extent.

In this paper, we present an unconventional and effective technique to attenuate seismic random noise. Firstly, we analyze various transforms for seismic data from traditional wavelet transform and curvelet transform to emerging NSST and NSCT. And, we select the excellent NSST to decompose the noisy seismic data into different sub-bands of frequency and orientation responses. Secondly, unlike traditional sparse transform based methods that often directly use a thresholding operator and corresponding inverse transform to denoise seismic data, our proposed method employs a superior performance proximal classifier with consistency (PCC) to classify the NSST coefficients of seismic data into reflected signal information-related coefficients and noise-related coefficients before thresholding operator, which can preserve the edge of reflected signals and keep the information of events intact as much as possible. Finally, we introduce an adaptive threshold computing method and a soft-thresholding method to achieve seismic data denoising better. We use the typical synthetic example and field acquisition seismic data to verify the effectiveness of the proposed approach. The experimental results show that our approach obtains higher quality assessment of seismic data.

II. PRELIMINARIES AND PROPOSED METHOD

A. ANALYSIS OF VARIOUS TRANSFORMS

Wavelets' success [17] on representing digital signals can be attributed to good sparse presentation for one-dimensional signals which are smooth away from discontinuous points. However, wavelet transform is difficult to deal with multi-dimensional signals with singularities effectively. In contrast, some other image/signal representations, including curvelets [18], contourlet [19], and bandelets [20], can explore the anisotropic regularity of a surface along edges. However, curvelets do not supply with a multi-resolution geometric representation; contourlets consist of very few clear directional features; bandelets transform requires high computational time to search for the optimal geometry.

Subsequently, shearlet transform is presented by Labate *et al.* [21], which is a multi-scale and multi-resolution representation with geometry. The major superiority of the transform is that it can be conducted via a general multi-resolution analysis. Following the success of shearlet transform, non-subsampled shearlet transform (NSST) [22], [23] was presented to maintain the superiority of shearlet transform and make up the deficiencies. NSST uses the non-subsampled Laplacian pyramid filters to take the place of the Laplacian pyramid filters to keep away from the negative effect brought by up-sampling and down-sampling in shearlet transform, which allows NSST to possess multi-scale, multi-directional and shift-invariant characteristics. A typical NSST consists of two parts: multi-scale decomposition and multi-directional decomposition. The multi-scale decomposition is implemented by the non-subsampled Laplacian pyramid filters; and the multi-directional decomposition is carried out by the modified shearing filters. To catch the singularities of images or signals, the non-subsampled Laplacian pyramid filters decompose the sub-band of low-frequency of k times. Thus, it can obtain $k+1$ sub-bands of high frequency and one sub-band of low-frequency. In order to not use sub-sampling, the entire procedure is implemented through a 2-dimensional convolution calculation by mapping shearing filters from the pseudo-polarization grid system to the Cartesian coordinate system, realizing direct processing in transform domain.

At the same time, non-subsampled contourlet transform (NSCT) was proposed by DaCunha et al. [24]. It is another effective multi-scale analysis algorithm and improves the contourlet transform [19]. NSCT applies non-subsampled pyramid decomposition as well as a directional filter bank based on 2-channel fan filter banks to establish directionally sensitive digital filters on several scales. NSCT primarily contains non-subsampled pyramid filter bank (NSPFB) and non-subsampled direction filter bank (NSDFB) in a cascade way. First, NSPFB decomposes the images or signals; then NSDFB takes the generated sub-bands of NSPFB as input to obtain the decomposed results of the initial images or signals in a multi-dimensional way. NSCT conducts K-level directional decomposition on any images or signals to obtain one sub-band of low-frequency and some sub-bands of high-frequency. The sizes of these sub-bands are the same as the initial image. NSCT can be completely reconstructed because NSPFB and NSDFB satisfy the conditions of complete reconstruction. The main drawback of NSCT is that the massive digital sampling is demanded by a directional filter bank, and only the band-limited filters can guarantee the directional selectivity.

For the above mentioned transforms, it can be found in [25] that NSST shows good performance in various experiments. So, we adopt the superior performance NSST to exploit seismic data. See Section II.B.

B. NON-SUBSAMPLED SHEARLET TRANSFORM (NSST) FOR SEISMIC DATA

Shearlet transform [21], [26] can generate almost optimal approximation properties. In the 2-dimensional space, shearlets can be presented as follows:

$$\psi_{a,s,t} = \left|\det M_{a,s}\right|^{-\frac{1}{2}} \psi\left(M_{a,s}^{-1}x - t\right), \tag{1}$$

where a, s and t are respectively scale, orientation and location parameters. $M_{a,s} = \begin{pmatrix} a & \sqrt{a}s \\ 0 & \sqrt{a} \end{pmatrix} = B_s A_s$, $A_s = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$, and $B_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$. The matrix A_s denotes an anisotropic dilation and B_s is a shearing matrix. ψ is the generating function.

The shearlet transform of a function f is defined as

$$SH_\psi f\,(a,s,t) = \langle f, \psi_{a,s,t} \rangle . \tag{2}$$

If the resulting function $\psi_{a,s,t}$ satisfiers the reconstructed conditions, each function $f \in L^2\left(R^2\right)$ can be reconstructed as

$$f = \int_{R^2} \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{a^3} \langle f, \psi_{a,s,t} \rangle \, \psi_{a,s,t} \mathrm{d}a\mathrm{d}s\mathrm{d}t, \tag{3}$$

where $\psi_{a,s,t}$ is defined in the frequency domain $\hat{\psi}\,(\xi_1, \xi_2) = \hat{\psi}_1\,(\xi_1)\,\hat{\psi}_2\left(\xi_2/\xi_1\right)$. $\hat{\psi}_1$ and $\hat{\psi}_2$ are smooth functions. They are inside a support of domain $\left[-2,-1/2\right] \cup \left[1/2,2\right]$ for $\hat{\psi}_1$ and the support of $[-1,1]$ for $\hat{\psi}_2$. So for $\xi = (\xi_1, \xi_2)$,

$$\hat{\psi}_{a,s,t}\,(\xi_1,\xi_2) = a^{3/4} e^{-2\pi i \varsigma t}\, \hat{\psi}_1\,(a\xi_1)\, \hat{\psi}_2\left(a^{-1/2}\left(\xi_2/\xi_1 - s\right)\right), \tag{4}$$

which has support in

$$\left\{(\xi_1,\xi_2):\ \xi_1 \in \left[-2/a, -1/(2a)\right] \cup \left[1/(2a), 2/a\right], \left|\left(\xi_2/\xi_1\right) - s\right| \le \sqrt{a}\right\}. \tag{5}$$

Each $\psi_{a,s,t}$ possesses frequency support on a pair of trapezoids at a few of scales, origin of symmetry, and oriented along a line of slope s. So, the shearlets $\psi_{a,s,t}$ form a series of well-localized waveforms at various scales a, orientations s, and locations t.

Non-subsampled shearlet transform (NSST) is the shift-invariant version of shearlet transform and combines the non-subsampled Laplacian pyramid transform (NSLP) with a few of different combinations of the shearing filters. NSLP can be expressed via iterative processing as follows:

$$NSLP_{j+1} = A_j f = \left(Ah_j^1 \prod_{k=1}^{j-1} Ah_k^0 \right) f, \tag{6}$$

where $NSLP_{j+1}$ is the detail coefficients of the $(j+1)-$ th scale. Ah_k^0 and Ah_j^1 are low-pass and high-pass filters of $NSLP$ at the k-th scale and the j-th scale, respectively. Given data f_a^0 and direction number D_j, the process of the NSST for the fixed j-th scale can be concluded as follows:

1) Use the NSLP to decompose matrix f_a^{j-1} into a data matrix f_a^j of low-pass and a data matrix f_d^j of high pass;
2) Calculate the Fourier transform $\hat{f}_d^j$ in pseudo polar grid, then obtain Pf_d^j;
3) Use a band-pass filtering on Pf_d^j to get $\left\{\hat{f}_{d,k}^j\right\}_{k=1}^{D_j}$;
4) Use inverse fast Fourier transform (FFT) to get NSST coefficients $\left\{f_{d,k}^j\right\}_{k=1}^{D_j}$ in pseudo polar grid.

Fig. 1 shows a framework of the non-subsampled sheartlet transform (NSST). The structure contains a bank of filters

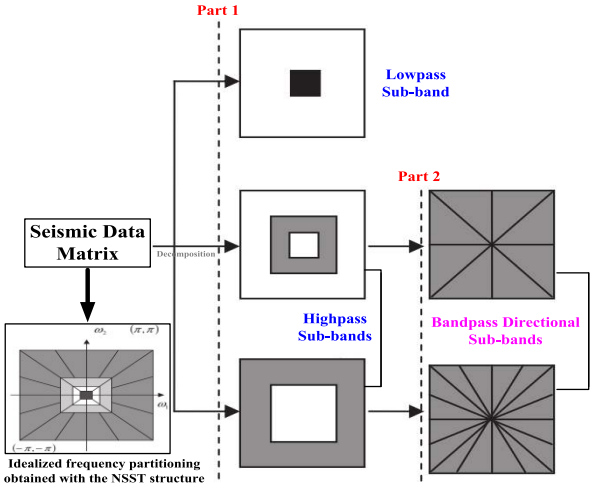


FIGURE 1. Overview of NSST of 2D seismic data.

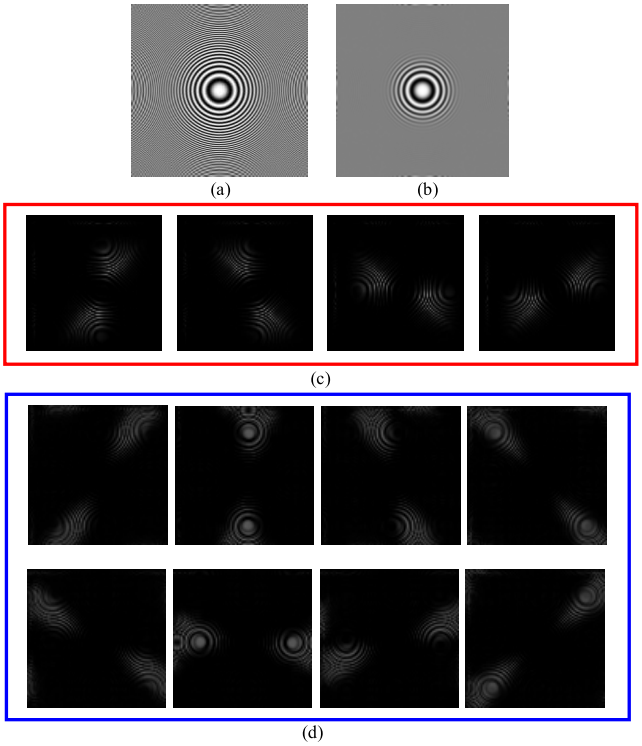


FIGURE 2. The NSST on Zoneplate data: (a) Original Zoneplate. (b) The approximate NSST coefficients (low-pass sub-band). (c) Images of the detail NSST coefficients at scale 2, 4 directions. (d) Images of the detail NSST coefficients at scale 1, 8 directions.

that divides the two-dimentional frequency plane into the sub-bands complained in the lower left quarter of Fig. 1. The 2D seismic data can be divided by NSST into two shift-invariant parts: ① a non-subsampled pyramid structure that guarantees the multi-scale property (Part 1 in Fig. 1) and ② a non-subsampled Directional Filter Bank (DFB) structure that presents directionality (Part 2 in Fig. 1).

Fig. 2 shows analyzed results of applying the two levels NSST on 'Zoneplate' data like the diffracted waves, yielding one low-pass sub-band and multiple high-pass sub-bands.

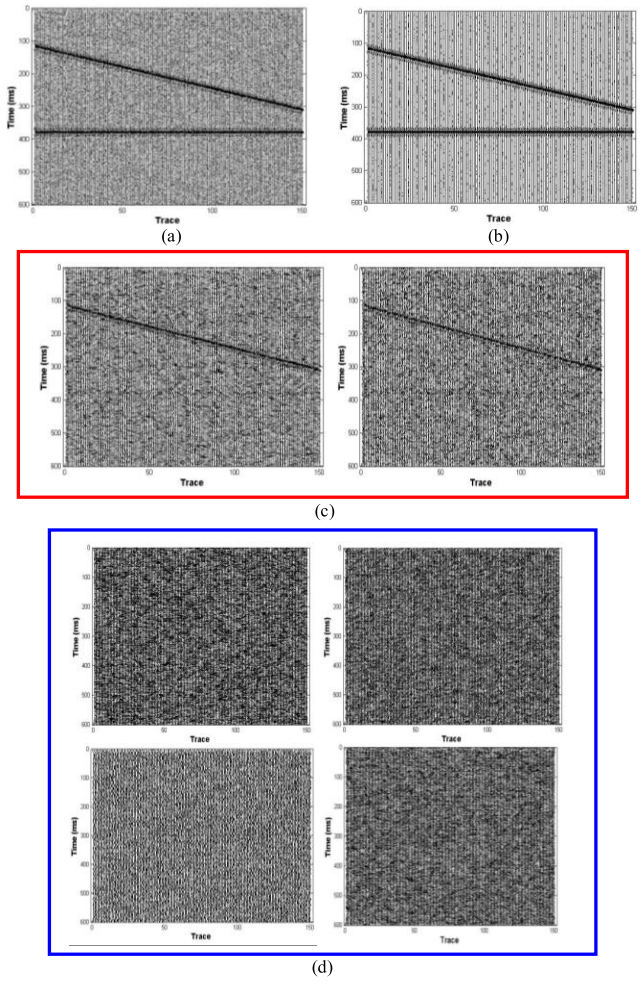


FIGURE 3. The NSST on synthetic seismic data with noise: (a) Noisy seismic data. (b) The approximate NSST coefficients. (c) Images of the detail NSST coefficients at scale 2, 2 directions. (d) Images of the detail NSST coefficients at scale 1, 4 directions.

Here, the numbers of shearing directions are selected to be 8 and 4 from finer to coarser scale.

Fig. 3 shows analyzed results of applying the two levels NSST on a synthetic seismic data with noise (Fig. 3(a)), also yielding a low-pass sub-band (Fig. 3(b)) and multiple high-pass sub-bands (Fig. 3(c-d)). Here, the numbers of shearing directions are also selected to be 4 and 2 from finer to coarser scale. We can see that signals of different frequencies are well separated by different decomposition scales and directions. The effective seismic reflected signals are mainly in the low frequency sub-band (Fig. 3(b)). Most of high frequency random noise is in the 4 shearing directions (Fig. 3(d)). Fig. 3(c) contains partial effective signals and random noise.

C. PROXIMAL CLASSIFIER WITH CONSISTENCY (PCC)

Many machine learning methods have been used different areas [13], [27]. Support vector machines (SVMs) [28], being effective classification and prediction tools for supervised learning, have already made many achievements in practical applications. SVMs establish two parallel hyper-planes; in contrast, the generalized eigenvalue proximal support vector machine (GEPSVM) presented by Mangasarian and Wild [29] achieves the objective of yielding two non-parallel hyper-planes, each of which is near itself class and is far away from the other classes. The potential difference between SVMs and GEPSVM is mainly that SVMs obtain a separated hyper-plane by seeking the solution of one quadratic program problem, whereas, GEPSVM obtains two non-parallel hyper-planes by seeking the solution of two generalized eigenvalue problems. The experimental results on UCI datasets show that GEPSVM is effective in [29].

Based on GEPSVM, Shao *et al.* [30] proposed the proximal classifier with consistency (PCC) algorithm, which also establishes two non-parallel hyper-planes. But, they are different potentially between PCC and GEPSVM. For GEPSVM, the optimization solutions are solved by computing and comparing the two distances from one hyper-plane to two sample points in the training process; while the prediction class of one test point is decided by comparing the two distances from the test point to two hyper-planes in the prediction process. Therefore, the prediction process is inconsistent with the training process. In order to achieve uniformity, PCC attempts to compare the two distances from one testing point or one training point to the two hyper-planes. Besides, the prediction process and the training process use a uniform distance. This not only brings about a low calculation cost of prediction function deservedly, but also logically makes PCC to be more reasonable.

In this section, we introduce the superior performance proximal classifier with consistency (PCC) for transformed seismic data. A two-class classification problem is considered in the n-dimensional space R^n. The training set can be expressed as follows:

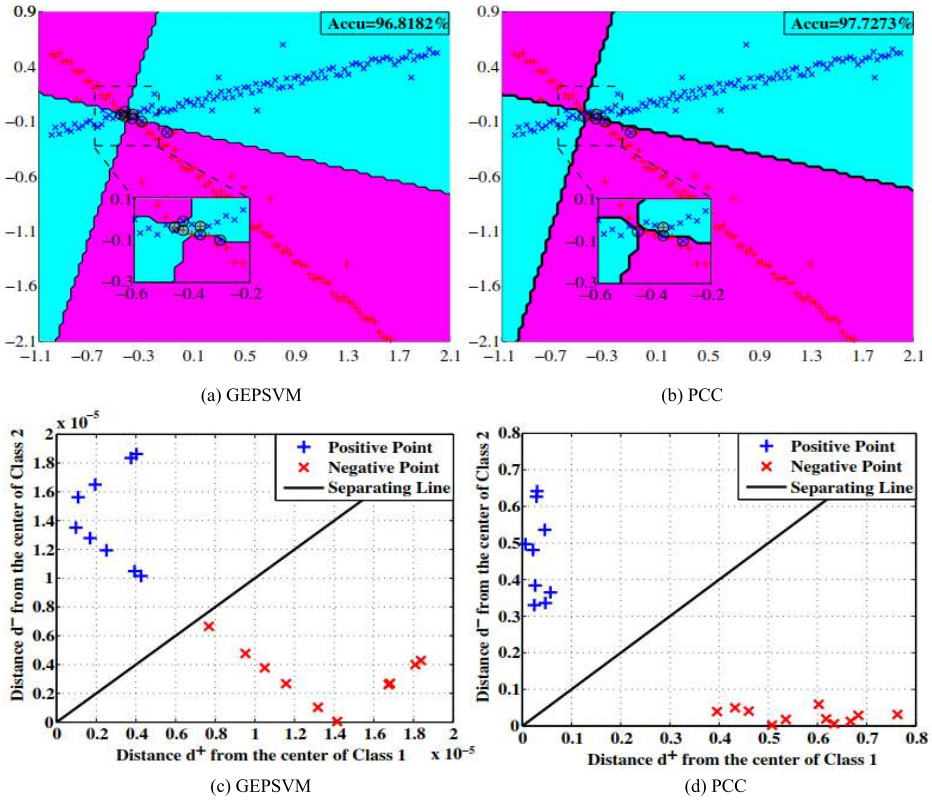
$$\left\{\left(x_1, y_1\right), \cdots, \left(x_p, y_p\right), \left(x_{p+1}, y_{p+1}\right), \cdots, \left(x_{p+q}, y_{p+q}\right)\right\}, \tag{7}$$

where $y_1 = \cdots = y_p = 1$, and $y_{p+1} = \cdots = y_{p+q} = 0$. The positive inputs are $X_+ = \left(x_1, \cdots, x_p\right)^{\mathrm{T}} \in \mathrm{R}^{p\times n}$ and the negative inputs are $X_- = \left(x_{p+1}, \cdots, x_{p+q}\right)^{\mathrm{T}} \in \mathrm{R}^{q\times n}$. The objective of PCC is to search a positive hyper-plane p_+ and a negative hyper-plane p_- simultaneously:

$$\begin{aligned} p_+ &: w_+^{\mathrm{T}} + b_+ = 0, \\ p_- &: w_-^{\mathrm{T}} + b_- = 0. \end{aligned} \tag{8}$$

PCC aims to compute and compare two distances from a point to the two hyper-planes in the process of training and decision. Giving the following expressions:

$$\begin{aligned} u_+ &= \begin{bmatrix} w_+ \\ b_+ \end{bmatrix}, \quad G_+ = [X_+ \; e_+], \; H_+ = G_+^{\mathrm{T}} G_+, \\ u_- &= \begin{bmatrix} w_- \\ b_- \end{bmatrix}, \quad G_- = [X_- \; e_-], \; H_- = G_-^{\mathrm{T}} G_-. \end{aligned} \tag{9}$$



(a) GEPSVM (b) PCC

(c) GEPSVM (d) PCC

FIGURE 4. Results of GEPSVM and PCC on "Cross Planes" data.

The optimal problem of PCC can be presented as follows:

$$\min_{u_+,u_-\neq 0} \frac{u_+^{\mathrm{T}} A u_+ + \delta \|u_+\|^2}{\|u_+\|^2} + \frac{u_-^{\mathrm{T}} B u_- + \delta \|u_-\|^2}{\|u_-\|^2}, \quad (10)$$

where $A = H_+ - vH_-$, $B = H_- - vH_+$, v denotes a positive trade-off parameter found in [31], $\delta > 0$ denotes a weighted factor, and $\|\cdot\|$ is the L2 norm. A new point $x \in \mathrm{R}^n$ can be classified to class i ($i = 1, 0$) when $u_+ = (w_+, b_+)$ and $u_- = (w_-, b_-)$ in (10) are solved:

$$\text{class } i = \arg\min_{i=1,0} \frac{|w_i^{\mathrm{T}} x + b_i|}{\|u_i\|}. \quad (11)$$

where $|\cdot|$ denotes the absolute value.

It has been shown that PCC can well deal with the "Cross Planes" data [29], [32]. Fig. 4 [30] indicates the data and the linear classifiers acquired by PCC and GEPSVM. It lists the 2-D scatter plots of the test data points for this 2-D dataset acquired by GEPSVM and PCC classifiers, respectively. The plots are acquired by plotting points with coordinates $\left(d_i^+, d_i^-\right)$, where d_i^+ and d_i^- are the distances from a testing point x_i to the two hyper-planes. Perceive easily, the classification results of PCC are better than that of GEPSVM from Fig. 4.

D. SEISMIC RANDOM NOISE ATTENUATION USING PCC CLASSIFICATION IN NSST DOMAIN

In this section, we present the new and unconventional seismic random noise attenuation method using PCC classification in NSST domain. NSST is an emerging and excellent multi-scale, multi-direction and optimal sparsity analysis method, which can provide nearly optimal approximation of the decomposed seismic effective signals. First, we decompose the noisy seismic data into different sub-bands of frequency and orientation responses using NSST, i.e., a low-pass sub-band and multiple high-pass sub-bands.

The approximation information of major reflected signals of seismic data is concentrated in the low-pass sub-band, while the high-pass sub-bands contain the noise and other detail reflected signal information-related. Therefore, it is so significant to accurately make a distinction between the noise-related NSST coefficients and reflected useful NSST signal coefficients in high-pass sub-bands. Thus, we propose to use the superior performance PCC to divide the NSST detail coefficients in high-pass sub-bands, as shown in Fig. 5. On this basis, we further extract the useful signal information (as 'red plus' in the black circle of Fig. 5) from the noise-related coefficients (as below separating line in Fig. 5) by introducing an adaptive threshold computing method and a soft-thresholding method after PCC classification.

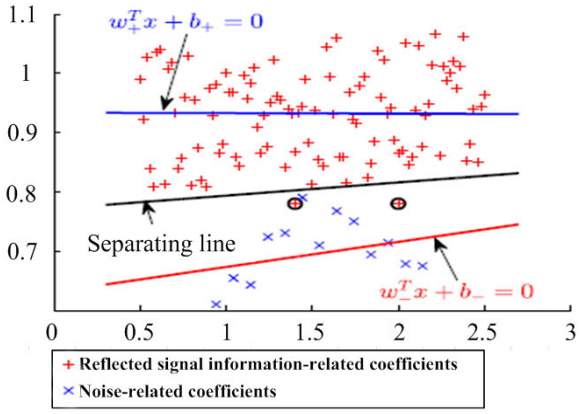


FIGURE 5. Diagram of PCC classification.

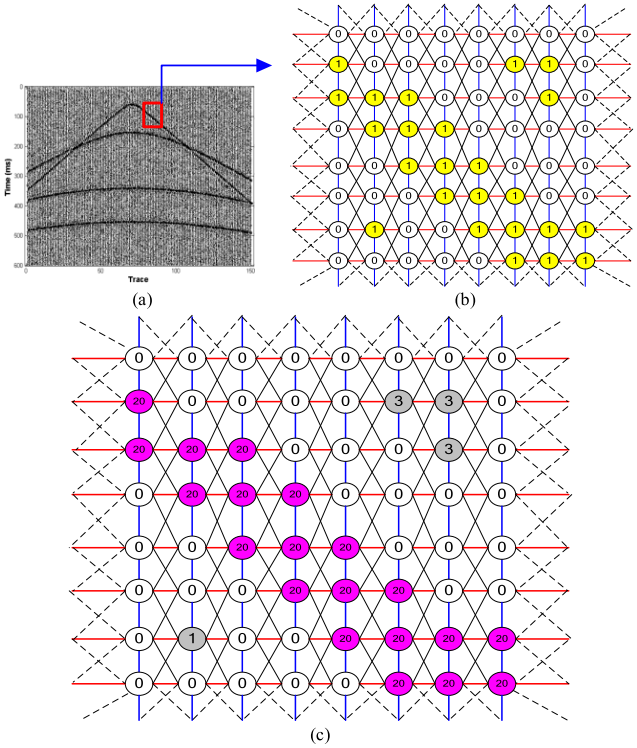


FIGURE 6. Diagram of the spatial geometric regularity: (a) the noisy seismic data; (b) the estimated binary map of the NSST coefficient matrix; (c) support value array.

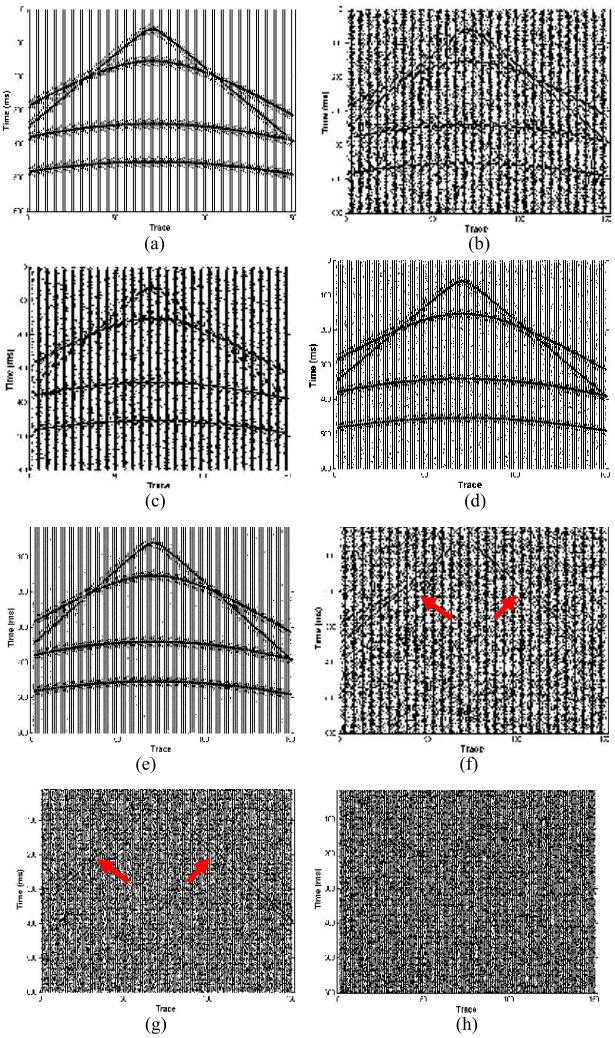


FIGURE 7. Comparison of processing results on noisy synthetic data by different denoising methods. (a) Original synthetic data. (b) Noisy synthetic data. (c, d) Denoised data and removed noise by the wavelet-based th reshold denoising approach. (f, g) Denoised data and removed noise by the curvelet-based threshold denoising approach. (e, h) Denoised data and removed noise by the proposed approach.

The main steps of our presented seismic denoising method can be summarized in detail as follows.

Step1: Conduct a K level NSST decomposition on the noisy seismic data, and acquire one low-pass sub-band and multiple high-pass sub-bands $D_{k,j}\,(k = 1, 2, \cdots, K;\ j = 1, 2, \cdots, J)$. Here, j denotes the orientation of decomposition, J denotes the direction number of decomposition, and k is the level of decomposition. See Section II.B.

Step2: Construct the feature vectors for PCC training by generating preliminary binary map and support value array [33], [34] in high-pass sub-bands $D_{k,j}$. Fig. 6 presents the spatial geometric regularity. Fig. 6(a) is the noisy seismic data, and Fig. 6(b) is the preliminary binary map of red box in Fig. 6(a). For the high-pass sub-bands $D_{k,j}$, we can compute the preliminary binary map matrix Class (x, y),

$$\text{Class}\,(x, y) = \begin{cases} 1, & |\text{C}\,(x, y)| > \tau, \\ 0, & else, \end{cases} \tag{12}$$

where C (x, y) is the NSST coefficients generated by NSST of noisy seismic signals, and (x, y) denotes the location as shown in the preliminary binary map. τ denotes a threshold parameter used to select available large NSST coefficients in the process of constructing the binary map of NSST coefficient matrix. Generally, we choose it by experiment. Fig. 6(c) is the support value array. The support value denotes the summation of all Class $(\cdot)$ which sustain the present binary value Class (x, y) that spatially connects the total number of available NSST coefficients. Thus, we select M NSST

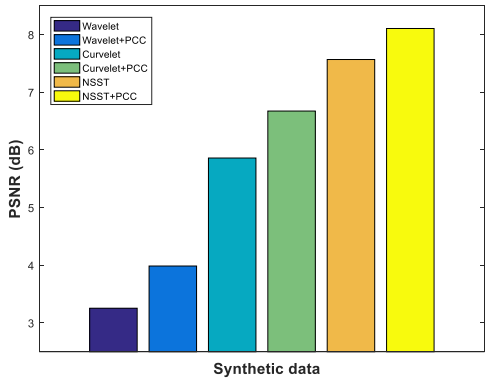


FIGURE 8. Effectiveness of PCC integrated with different transforms.

coefficients corresponding to high-pass sub-bands $D_{k,j}$ with the maximal support value (e.g. the maximal number is 20 of support value '1' as shown in Fig. 6(c)) as the feature vectors $FV^1_{k,j}$, and M NSST coefficients with the support value '0' are randomly selected as the feature vectors $FV^2_{k,j}$. Finally, the preliminary binary map Class (x, y) corresponding to the selected NSST coefficients is viewed as the training objective $OB^1_{k,j}$ and $OB^2_{k,j}$ respectively. The selected feature vectors contain most valid seismic signals.

Step3: Train the PCC model of objective with NSST coefficients. Let $FV^1_{k,j}$ and $FV^2_{k,j}$ be the trained feature vectors, $OB^1_{k,j}$ and $OB^2_{k,j}$ are the trained objective class; then the training samples can be obtained as follows:

$$S_{k,j} = \left\{\left(FV^1_{k,j}, FV^2_{k,j}, OB^1_{k,j}, OB^2_{k,j}\right)\right\}. \tag{13}$$

Therefore, the PCC classification model can be acquired by training $S_{k,j}$.

Step4: Classify the NSST coefficients in high-pass sub-bands $D_{k,j}$ into two categories by using PCC model: reflected useful signal coefficients and noise-related coefficients. So, the NSST coefficient is viewed as the reflected useful signal coefficient while the actual output is 1 according to Formula (11), and the NSST coefficient is viewed as the noise-related coefficient while the actual output is 0.

Step5: Calculate the denoising threshold for each detail sub-band. In this paper, we introduce an adaptive threshold computing method [35] to compute the denoising threshold $T_{k,j}$:

$$T_{k,j} = S_{k,j} \times T^*_{k,j}, \tag{14}$$

where $S_{k,j} = 1 - (J/4) \times \left(E_{k,j} \Big/ \sum_{j=1}^{K} E_{k,j}\right)$ is the compensation factor for modifying the primary denoising threshold $T^*_{k,j} = \delta_{k,j}\sqrt{2\ln N} \times 2^{(k-K)/2}$. $\delta_{k,j}$ is noise standard deviation of detailed sub-bands $D_{k,j}$, and N denotes the number of elements in the noise-related sub-bands. $E_{k,j} = \sum_{d_{k,j}(x,y)\in D_{k,j}} d^2_{k,j}(x, y)$, and $d_{k,j}(x, y)$ is the original noise-related NSST coefficients in high-pass sub-bands $D_{k,j}$.

Step6: Handle the noise-related NSST coefficients in high-pass sub-bands $D_{k,j}$ using the outstanding soft-thresholding approach.

$$\overline{d}_{k,j}(x, y) = \begin{cases} \operatorname{sgn}\left(d_{k,j}(x, y)\right)\left(\left|d_{k,j}(x, y)\right| - T_{k,j}\right), \\ \qquad \left|d_{k,j}(x, y)\right| \geq T_{k,j}, \\ 0, \quad otherwise, \end{cases} \tag{15}$$

where $\overline{d}_{k,j}(x, y)$ denotes the handled NSST coefficients using the soft thresholding approach in high-pass sub-bands.

Step7: Perform inverse NSST transform on the denoised NSST high-pass components, and the original low-pass component to reconstruct the denoised seismic data.

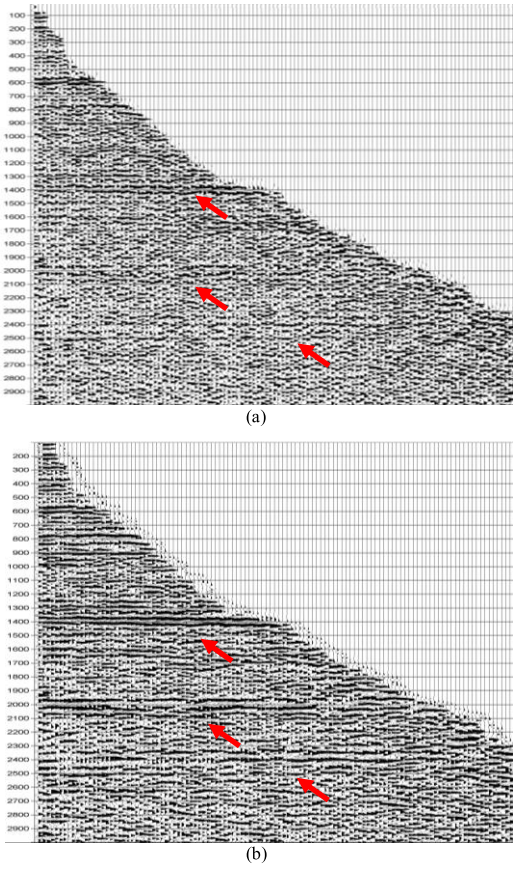


FIGURE 9. CDP gather (a) and the result (b) processed by proposed approach.

III. EXPERIMENTAL RESULTS

To evaluate the performance of our presented approach, we first compare it with the two well-known wavelet transform based seismic denoising method and curvelet transform based seismic denoising method on a typical synthetic example that can be shown in Fig. 7(a), and the time sampling interval is 1 ms and the number of traces is 150. We add Gaussian random noise to the synthetic data, as shown in Fig. 7(b). In our experiments, the threshold parameter τ in Section II.D is set to 15. The numbers of shearing directions are selected to be 4 and 2 from finer to coarser scale. We denoise the noisy synthetic data using the three approaches. Fig. 7(c), 7(d) and 7(e) are the denoised results,

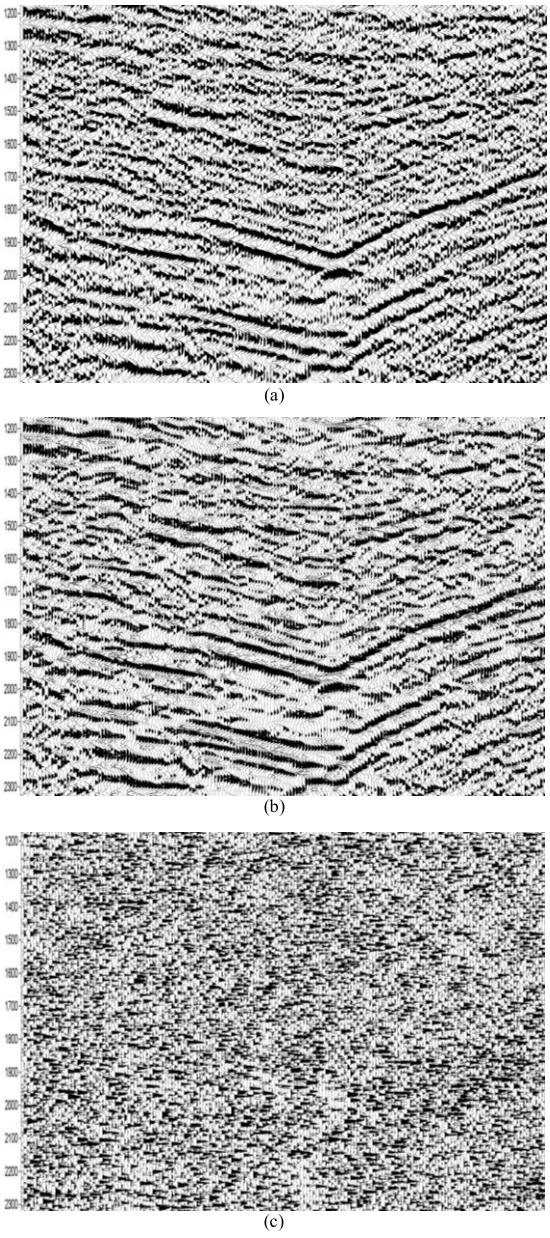


FIGURE 10. Stack profile (a), the result (b) processed by our proposed method and removed noise (c) by the proposed approach.

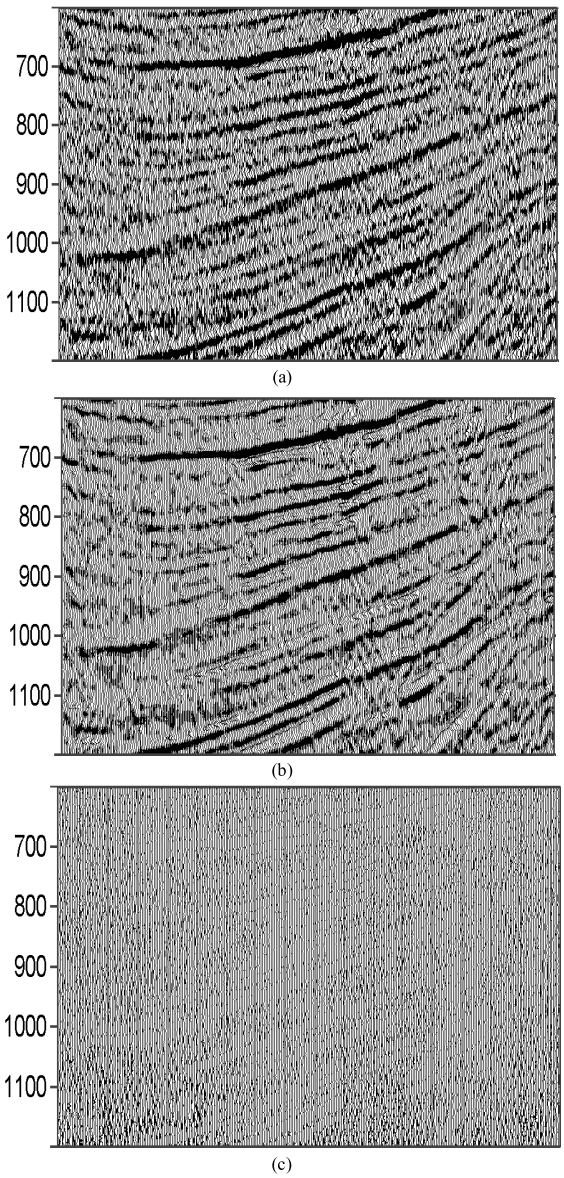


FIGURE 11. Stack profile (a), the result (b) processed by our proposed method and removed noise (c) by the proposed approach.

respectively. From the visual perspective, the result (Fig. 7(e)) by our proposed approach is obviously better than the ones (Fig. 7(c) and 7(d)) by other two approaches. In addition, we also quantitatively evaluate the results using the classic signal to noise ratio (SNR) [36] that can be defined as

$$\mathrm{SNR} = 20 \cdot \log_{10}\left(\|\mathbf{x}_0\|_2 / \|\mathbf{x}_1 - \mathbf{x}_0\|_2\right), \tag{16}$$

where $\|\cdot\|_2$ denotes L2 norm. $\mathbf{x}_0$ is noiseless data, and $\mathbf{x}_1$ is data with noise or denoised data. The SNR values of the results (Fig. 7(b), 7(c), 8(d) and 7(e)) are -2.3405 dB, 3.2548 dB, 5.8596 dB and 7.9543 dB. This indicates that our presented approach can achieve a better denoising result. Fig. 7(f), 7(g) and 7(h) are removed noise by three methods, respectively. We can see that the wavelet transform based and the curvelet transform based methods lose some effective signals (red arrows); whereas, our proposed method does not lose effective signals basically. The reason of attaining such a good result is closely related to conduct an effective PCC classification on NSST detail coefficients using a superior performance classifier. Besides, we summarize the data with different SNR before and after denoising, as shown in Table 1. We can see that the proposed method have best denoising ability, especially for the noisy data of low SNR. Besides, we also compare two outstanding thresholding methods (hard-thresholding and soft-thresholding) for proposed approach, as shown in Table 2. We can see that the soft-thresholding method is a little better than the hard-thresholding method.

TABLE 1. Comparison of data with different signal to noise ratios (SNR) before and after denosing.

Noisy data (dB)	Wavelet-based denosing (dB)	Curvelet-based denosing (dB)	Proposed approach (dB)
-7.3382	-2.5677	0.6259	3.9652
-2.3405	3.2548	5.8596	8.1059
1.6574	7.6549	9.7963	11.7422
5.3364	13.9654	15.1025	16.2377

TABLE 2. Comparison of different thresholding methods for proposed approach.

Noisy data (dB)	Proposed approach (hard-thresholding) (dB)	Proposed approach (soft-thresholding) (dB)
-7.3382	3.6643	3.9652
-2.3405	7.9125	8.1059
1.6574	11.1024	11.7422
5.3364	15.4355	16.2377

Next, we evaluate the effectiveness of PCC on the synthetic data (Fig. 7(b)). We use three approaches (Wavelet transform, Curvelet transform and NSST) and integrate them with PCC. Fig. 8 indicates the comparison results of processing the synthetic data by different approaches. From Fig. 8, we can see that three approaches can be improved significantly when integrated with PCC. Experimental results demonstrate that the improvements are consistent across various transforms.

Furthermore, we apply our proposed method to real seismic data in Liaohe depression, China: a pre-stack common depth point (CDP) gather (Fig. 9(a)) and two stack profile (Figure 10(a) and Figure 11(a)). The processed result (Fig. 9(b)) for the CDP gather show that the reflected signals highlights and the SNR is improved significantly. Fig. 10(b) and Fig. 11(b) show the processed result for the stack profile. We can see clearly that the reflected signals highlights, and there is clearer interlayer information. Fig. 10(c) and Fig. 11(c) indicate the removed noise sections. It is obvious that they do not lose useful signals essentially.

IV. CONCLUSION

We present a new seismic random noise attenuation method based on the non-subsampled shearlet transform (NSST) and proximal classifier with consistency (PCC) classification. The new technique can keep the information of events intact to the maximum degree. The experimental results on the synthetic data and real seismic data show the effectiveness of our proposed method. The fine processing will be more suitable for high precision seismic exploration.

REFERENCES

[1] T. Qiu, "Robustness optimization scheme with multi-population co-evolution for scale-free wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1028–1042, Jun. 2019, doi: 10.1109/TNET.2019.2907243.

[2] T. Qiu, X. Wang, C. Chen, M. Atiquzzaman, and L. Liu, "TMED: A spider Web-like transmission mechanism for emergency data in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8682–8694, May 2018.

[3] Y. Cao, H. Qi, W. R. Zhou, J. Kato, K. Q. Li, X. L. Liu, and J. Gui, "Binary hashing for approximate nearest neighbor search on big data: A survey," *IEEE Access*, vol. 6, pp. 2039–2054, 2018.

[4] X. L. Liu, X. Xie, and X. B. Zhao, "Fast identification of blocked RFID tags," *IEEE Trans. Mobile Comput.*, vol. 17, no. 9, pp. 2041–2054, Jan. 2018.

[5] X. L. Liu, J. N. Cao, and Y. N. Yang, "Fast RFID sensory data collection: Trade-off between computation and communication costs," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1179–1191, Jun. 2019.

[6] X. Liu, X. Xie, J. Liu, D. Yao, J. Cao, K. Li, and S. Wang, "Efficient range queries for large-scale sensor-augmented RFID systems," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 1873–1886, Oct. 2019.

[7] L. L. Canales, "Random Noise Reduction," in *Proc. 54th Annu. Int. Meeting SEG Tech. Program Expanded Abstracts*, Tulsa, OK, USA, 1984, pp. 525–527.

[8] J. H. Zhang and J. M. Lu, "Application of wavelet transform in removing noise and improving resolution of seismic data," (in Chinese), *J. Univ. Petroleum, Ed. Natural Sci.*, vol. 31, no. 12, pp. 1975–1981, 1997.

[9] Y. Chen and J. Ma, "Random noise attenuation by $f - x$ empirical mode decomposition predictive filtering," *Geophysics*, vol. 79, no. 3, pp. V81–V91, 2014.

[10] S. Beckouche and J. W. Ma, "Simultaneous dictionary learning and denoising for seismic data," *Geophysics*, vol. 79, no. 3, pp. A27–A31, 2014.

[11] Y. Chen and S. Fomel, "EMD-seislet Transform," in *Proc. 85th SEG Annu. Int. Meeting, Expanded Abstracts*, New Orleans, LA, USA, 2015, pp. 4775–4778.

[12] W. Liu, S. Cao, and Y. Chen, "Application of variational mode decomposition in random noise attenuation and time-frequency analysis seismic data," in *Proc. EAGE 78th Annu. Int. Conf. Exhib.*, Vienna, Austria, 2016.

[13] Y. Chen, "Fast dictionary learning for noise attenuation of multidimensional seismic data," *Geophys. J. Int.*, vol. 209, no. 1, pp. 21–31, 2017.

[14] R. Neelamani, A. I. Baumstein, D. G. Gillard, M. T. Hadidi, and W. L. Soroka, "Coherent and random noise attenuation using the curvelet transform," *Lead. Edge*, vol. 27, no. 2, pp. 240–248, 2008.

[15] W. Liu, S. Cao, Y. Chen, and S. Zu, "An effective approach to attenuate random noise based on compressive sensing and curvelet transform," *J. Geophys. Eng.*, vol. 13, no. 2, pp. 135–145, 2016.

[16] Y. Sang, J. G. Sun, X. F. Meng, H. B. Jin, Y. F. Peng, and X. J. Zhang, "Non-subsampled shearlet transform based seismic data denoising via proximal classifier with consistency," in *Proc. Int. Conf. Smart Internet Things (SmartIOT)*, Tianjin, China, 2019.

[17] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. New York, NY, USA: Academic, 2008.

[18] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities," *Commun. Pure Appl. Math.*, vol. 57, no. 2, pp. 219–266, Feb. 2004.

[19] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[20] E. L. Pennec and S. Mallat, "Sparse geometric image representations with bandelets," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 423–438, Apr. 2005.

[21] D. Labate, W. Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," *Proc. SPIE*, 2005.

[22] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25–46, Jul. 2008.

[23] B. Hou, X. Zhang, X. Bu, and H. Feng, "SAR image despeckling based on nonsubsampled shearlet transform," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 3, pp. 809–823, Jun. 2012.

[24] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[25] G. Kutyniok, W.-Q. Lim, and R. Reisenhofer, "ShearLab 3D: Faithful digital shearlet transforms based on compactly supported shearlets," *ACM Trans. Math. Softw.*, vol. 42, no. 1, 2016, Art. no. 5.

[26] W.-Q. Lim, "The discrete shearlet transform: A new directional transform and compactly supported shearlet frames," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1166–1180, May 2010.

[27] T. Qiu, H. Wang, K. Li, H. Ning, A. K. Sangaiah, and B. Chen, "SIGMM: A novel machine learning algorithm for spammer identification in industrial mobile cloud computing," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2349–2359, Apr. 2018.

[28] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.
[29] O. L. Mangasarian and E. W. Wild, "Multisurface proximal support vector machine classification via generalized eigenvalues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 69–74, Jan. 2006.
[30] Y. H. Shao, N. Y. Deng, and W. J. Chen, "A proximal classifier with consistency," *Knowl. Syst.*, vol. 49, no. 1, pp. 171–178, 2013.
[31] Y.-H. Shao, N.-Y. Deng, W.-J. Chen, and Z. Wang, "Improved generalized eigenvalue proximal support vector machine," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 213–216, Mar. 2013.
[32] Y. H. Shao, C. H. Zhang, X. B. Wang, and N. Y. Deng, "Improvements on twin support vector machines," *IEEE Trans. Neural Netw.*, vol. 22, no. 6, pp. 962–968, May 2011.
[33] E. J. Balster, Y. F. Zheng, and R. L. Ewing, "Combined spatial and temporal domain wavelet shrinkage algorithm for video denoising," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 220–230, Feb. 2006.
[34] E. J. Balster, Y. F. Zheng, and R. L. Ewing, "Feature-based wavelet shrinkage algorithm for image denoising," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2024–2039, Dec. 2005.
[35] W. Dai, S. Yu, and S. Sun, "Image de-noising algorithm using adaptive threshold based on contourlet transform," *Acta Electronica Sinica*, vol. 35, no. 10, pp. 1939–1943, 2007.
[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

YU SANG received the Ph.D. degree in computer science and technology from the Dalian University of Technology, China, in 2012. He is currently an Associate Professor with Liaoning Technical University. He has published about 20+ research articles. His main research interests include seismic signal processing, artificial intelligence, and computer vision.

JINGUANG SUN was born in 1962. She received the Ph.D. degree from Liaoning Technical University, China, in 2006. She is currently a Full Professor and a Ph.D. Supervisor with Liaoning Technical University. Her research interests include image processing, computer graphics, deep learning, and spatio-temporal data management.

XIANGFU MENG was born in 1981. He received the Ph.D. degree from Northeastern University, China, in 2010. He is currently a Full Professor and a Ph.D. Supervisor with Liaoning Technical University, China. His research interests include spatial data management, recommender systems, and Web database query.

HAIBO JIN received the M.Sc. degree in computer science and technology from Liaoning Normal University, in 2009, and the Ph.D. degree in electronic information and electrical engineering from the Dalian University of Technology, Dalian, China, in 2014. He currently holds a full-time position with Liaoning Technical University. His research interests include the reliability of complex systems, optimal maintenance strategy, and stochastic process.

YANFEI PENG was born in 1975. He received the Ph.D. degree from Liaoning Technical University, China. He is currently an Associate Professor with Liaoning Technical University. He has published about several tens research articles. His main research interests include intelligent information processing and computer vision.

XINJUN ZHANG received the master's degree from Liaoning Technical University, China. She is currently pursuing the degree with the Dalian University of Technology, China. She is currently an Associate Professor with Liaoning Technical University. She has published several research articles. Her main research interests include image processing and multimedia technologies.