

# Compressive Sensing Sparse Sampling Method for Composite Material Based on Principal Component Analysis

Sun Yajie<sup>1, 2, 3\*</sup>, Gu Feihong<sup>3</sup>, Ji Sai<sup>1, 2, 3</sup>, Wang Lihua<sup>4</sup>

1. Jiangsu Engineering Centre of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing 210044, P. R. China;
2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, P. R. China;
3. School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, P. R. China;
4. School of Information and Control, Nanjing University of Information Science and Technology, Nanjing 210044, P. R. China

(Received 20 December 2017; revised 4 March 2018; accepted 30 March 2018)

**Abstract:** Signals can be sampled by compressive sensing theory with a much less rate than those by traditional Nyquist sampling theorem, and reconstructed with high probability, only when signals are sparse in the time domain or a transform domain. Most signals are not sparse in real world, but can be expressed in sparse form by some kind of sparse transformation. Commonly used sparse transformations will lose some information, because their transform bases are generally fixed. In this paper, we use principal component analysis for data reduction, and select new variable with low dimension and linearly correlated to the original variable, instead of the original variable with high dimension, thus the useful data of the original signals can be included in the sparse signals after dimensionality reduction with maximize portability. Therefore, the loss of data can be reduced as much as possible, and the efficiency of signal reconstruction can be improved. Finally, the composite material plate is used for the experimental verification. The experimental result shows that the sparse representation of signals based on principal component analysis can reduce signal distortion and improve signal reconstruction efficiency.

**Key words:** principal component analysis; compressive sensing; sparse representation; signal reconstruction

**CLC number:** TP212.9

**Document code:** A

**Article ID:**1005-1120(2018)02-0282-08

## 0 Introduction

Principal component analysis (PCA) is one of the most commonly used multivariate statistical techniques<sup>[1]</sup> It uses an orthogonal mathematical transformation to convert the observed values of a set of possible dependent variables to principal components, the values that are not linearly related. The number of principal components is less than or equal to the number of original variables. Only when the data is combined with normal distribution, the principal component is independent from each other. PCA is sensitive to the

correlation level between the original variables. It is also known as Hotelling transform, discrete KLT transform or proper orthogonal decomposition in different fields.

By projecting the data into the low dimensional space and obtaining the most possible features of the original data, PCA can be used to deal with the data of high dimension, noisy and high correlation. So far it has developed into a kind of exploratory data analysis and prediction model in terms of feature extraction using covariance or correlation matrix decomposition or using a set of matrix signal values. In recent decades,

\*Corresponding author, E-mail address: syj@nuist.edu.cn.

scholars have looked into the characteristics of PCA extraction and dimension reduction in different disciplines<sup>[2-4]</sup>. Wold et al.<sup>[5]</sup> used cross examination to determine the number of PCA principal components, and a PCA—method for model prediction. Ku<sup>[6]</sup> introduced "time lag transfer" to statistical monitoring, and developed the monitoring method of the previous static PCA to a dynamic PCA method, which was applied to the detection of the disturbance of the dynamic multivariable system.

As long as the signals are sparse, through the sampling rate of compressive sensing (CS) is far lower than that of the traditional Nyquist sampling theorem<sup>[7-9]</sup>. The theory must be premised on the sparsity of the signal, and PCA can be used for data dimensionality reduction. Masiero et al.<sup>[10]</sup> used PCA to find transformations to sparsify signals for CS to retrieve. They dynamically adapted non-stationary real-world signals through the online estimation on their correlation properties in space and time, and then utilized PCA to derive the transformations for CS. Li et al.<sup>[11]</sup> proposed an adaptive block compressive sensing based on edge detection at the encoder, and a smoothed projected Landweber (SPL) reconstruction algorithm based on principal component analysis at the decoder. The reconstruction algorithm used PCA to train a dictionary adapting to image structure with hard thresholding, thus the image blocking effects were eliminated effectively and the reconstructed image quality was improved. Dietz et al.<sup>[12]</sup> presented a real-time dynamic image reconstruction technique, which combined CS and PCA to achieve real-time adaptive radiotherapy with the use of a linac-magnetic resonance imaging system. Li et al.<sup>[13]</sup> proposed an efficient image fusion framework for infrared and visible images on the basis of robust principal component analysis (RPCA) and CS. Compared with several popular fusion algorithms, this framework could extract the infrared targets while retaining the background information in the visible images.

Therefore, the compressive sensing method based on PCA is proposed to provide a better so-

lution to sparse data representation problem of huge amount of ultrasonic phased array signal.

## 1 Principal Component Analysis

PCA is to reduce the dimension of the original data space by constructing a new set of latent variables, and then extract statistical features from the mapping space, therefore to understand the spatial characteristics of the original data. The variables of the new mapping space are composed of linear combination of the original data variables, which greatly reduces the dimension of the projection space. The number of new variables is less than that of the original variables, while still carry useful information of the original data as much as possible. Its contents include the definition and acquisition of main elements, as well as the principal component of the data reconstruction. Since the statistical characteristic vectors of the projection space are orthogonal to each other, the correlations between variables are eliminated, and the complexity of the original process characteristic analysis is simplified. Therefore, this method can effectively identify the most important elements and structures in the data, remove the noise and redundancy, reduce the original complex data, and reveal the simple structure behind the complex data.

Given the original data  $\mathbf{x} = (x_{ij})_{N \times M}$ ,  $x$  is standardized to eliminate the dimensional effects, and the expression is shown as

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad (1)$$

where  $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n hx_{ij}$ , and  $S_j^2 = \frac{1}{n-1} \sum_{i=1}^n h(x_{ij} - \bar{x}_j)^2$ ,  $j=1, 2, \dots, m$ .

The correlation coefficient matrix is calculated between the data variables after standardized operation, and the covariance matrix  $\mathbf{R}$  is

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix} \quad (2)$$

where the element  $r_{jk}$  represents the correlation coefficient of the original variable  $x'_{j_i}$  and  $x'_{k_i}$ , and  $r_{jk} = r_{kj}$ , shown as

$$r_{jk} = \frac{\sum_{k=1}^n (x'_{ki} - \bar{x}'_i) (x'_{kj} - \bar{x}'_j)}{\sqrt{\sum_{k=1}^n (x'_{ki} - \bar{x}'_i)^2 \sum_{k=1}^n (x'_{kj} - \bar{x}'_j)^2}} \quad (3)$$

$i, j = 1, 2, \dots, m$

Jacobi method is used to solve the characteristic equation  $|\lambda I - R| = 0$ , and the eigenvalues of the covariance matrix and the corresponding eigenvectors are obtained. Then it is sorted according to the size of the order, and the characteristic value is recorded as  $[\lambda_1, \lambda_2, \dots, \lambda_m]$ , and the corresponding feature vector is recorded as  $[p_1, p_2, \dots, p_m]$

$$p_i = [p_{i1}, p_{i2}, \dots, p_{im}] \quad (4)$$

Then the main elements are calculated  $t_i = Xp_i$ , where the principal component  $t_i$  on behalf of the projection of the data matrix  $x$  on the direction of the load vector corresponding to the main element.

The contribution rate of each principal components is calculated as  $\frac{\lambda_i}{\sum_{k=1}^m \lambda_k}$ ,  $i = 1, 2, \dots, m$ , as

well as the cumulative contribution rate  $\frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^m \lambda_k}$ ,  $i = 1, 2, \dots, m$ .

In general, the 1th, 2th,  $k$ th principal component corresponding to the eigenvalues of  $\lambda_1, \lambda_2, \dots, \lambda_k$  will be selected, where the cumulative contribution rate of eigenvalues is between 85% and 95%.

In addition, according to the needs, the corresponding dimension (that is, the number of principal components) is selected to composition of the transformation matrix

$$A = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mk} \end{bmatrix} \quad (5)$$

Finally, the new data after dimension reduction is calculated as

$$s = A^T \cdot x \quad (6)$$

## 2 Compressive Sensing Method Based on PCA

CS is a novel theory of sampling and restoration for sparse signal<sup>[7-9]</sup>. As long as the original signals are sparse in the time domain or under some kind of orthogonal transform, the signals can be sampled in a low sampling rate, and the original signals can be reconstructed with high probability.

### 2.1 Sparse representation of signals

CS is based on the premise that the signal must be sparse. When sparsifying the signals, the appropriate sparse transform base according to the signal characteristics is necessary to be selected.

Suppose an original  $x$  signal with the length of  $N$ , the number of signal is  $M$ , a  $N \times M$  dimension matrix can be constructed with the original signal, because there are mutual relationship between the amplitude of each signal in each time point. According to the principal component analysis, the covariance matrix obtained of  $N \times N$  dimension can be used as the sparse transform base for sparse representation of signal. Then, the original signal  $x$  can be expressed as

$$x = \sum_{i=1}^N \theta_i \psi_i \text{ or } x = \Psi \Theta \quad (7)$$

where  $\Psi = [\psi_1, \psi_2, \dots, \psi_N]$  is the transformation matrix of  $N \times N$  dimension,  $\Theta$  the sparse coefficient vector obtained by  $x$  according to the principal component analysis, and must meet the following formula

$$\Theta = \Psi^T x \text{ or } \theta_i = \langle x, \psi_i \rangle = \psi_i^T x \quad (8)$$

PCA reduces a kind of high dimensional data to low dimensional data. Then, a set of new variables in low dimensional replace the original variables in high dimension satisfies the conditions associated with the original ones. Therefore, the new variables can carry the maximum information

of the original ones. PCA can be used to make sparse representation of the signals. Compared with the commonly used sparse representation method, the sparse signals obtained by the proposed method is more closely related to the original signals.

According to Section 1, the sparse coefficient vector of original signal is calculated as

$$\Theta = \Psi^T \mathbf{x} \quad (9)$$

where  $\Psi$  is the transformation of matrix  $\mathbf{A}$  in Section 1.

## 2.2 Projection observation of signals

The core of the compressive sensing theory is to design the measurement matrix, and directly determine whether the compressive sensing can be implemented successfully. If the signal  $\mathbf{x}$  has a sparse representation under an orthogonal transform  $\Psi$ , a measurement matrix  $\Phi$ ,  $\Psi \in \mathbf{R}^{M \times N}$ , which is not related to the transform base  $\Psi$ , and a linear measurement of  $M$  dimension can be obtained

$$y_i = \langle \theta, \varphi_i \rangle \quad (10)$$

Suppose the production measurement vector is  $\mathbf{y} = [y_1, y_2, \dots, y_M]$ , then

$$\mathbf{y} = \Phi \Theta = \Phi \Psi^T \mathbf{x} \quad (11)$$

In order to restore the original signal with high probability, the production measurement matrix  $\Phi$ , which is not related to the sparse transform base  $\Psi$  and satisfied with the restricted isometry property, is needed to be constructed to make production transformation of the signal. Gauss random measurement matrix is not related to the majority of the fixed orthogonal base and satisfies the restricted isometry property, so the Gauss matrix can be used as the projection observation matrix<sup>[14-16]</sup>. For the ultrasonic phased array signal, the Gauss random measurement matrix is multiplied with the sparse coefficient of the phased array signal, and the observation vector of the signal can be obtained.

Suppose the measurement matrix  $\Phi$  is  $M \times N$  dimension, and  $\Theta \in \mathbf{R}^{M \times N}$ , then the general term

$$\Phi(i, j) = \frac{1}{\sqrt{M}} h_{ij} \quad (12)$$

Each element in the matrix is independent to the Gauss distribution with the mean value of 0, and the variance of  $\frac{1}{\sqrt{M}}$ . This matrix is not related to the vast majority of sparse signals, and requires less measurement values in the reconstruction. Gauss random measurement matrix is a matrix with very strong randomness but high uncertainty. For a signal with a length of  $N$  and a sparse degree of  $K$ , only  $M \geq cK \log(\frac{N}{K})$  measured values are needed to recover the original signal with high probability, where  $c$  is a very small constant.

## 2.3 Sparse reconstruction of signals

During the process of compressive sensing, reconstructing the signal  $\mathbf{x}$  from the observations  $\mathbf{y}$  is the inverse problem related to compression sampling, and is called signal reconstruction. By solving Eq. (11), the reconstructed signal can be obtained. This problem is underdetermined with infinite solutions. Candes et al. proved that the underdetermined problem can be solved by solving the minimum  $l_0$ -norm<sup>[13]</sup>, that is,

$$\min_{\Theta_0} \quad \text{s. t.} \quad \mathbf{y} = \Phi \Theta = \Phi \Psi^T \mathbf{x} \quad (13)$$

Eq. (13) is a linear programming problem, and is also a convex optimization problem. Taking the reconstruction error into account, it is converted into a minimum  $l_1$ -norm problem as

$$\min_{\Theta_1} \quad \text{s. t.} \quad \Phi \Theta - \mathbf{y}_2 \leq \epsilon \quad (14)$$

During the process of signal reconstruction, convex optimization algorithm and greedy iterative algorithm are commonly used<sup>[17-18]</sup>. One kind of algorithm is based on convex optimization, mainly by increasing the constraint to obtain the sparsest. And commonly used algorithms are basis pursuit algorithm and gradient projection sparse reconstruction algorithm. The other kind of algorithm is based on greedy iterative algorithm, mainly by the combination of local optimization method to find the non-zero coefficients, in

order to approach the original signal. Commonly used algorithms are matching pursuit algorithm and orthogonal matching pursuit algorithm.

### 3 Experiment and Results

A composite plate is the experimental object. There are nine piezoelectric elements in the linear array arranged on the plate with an equal interval of 12 mm. In signal acquisition, data collection points are 1 024, and sampling frequency is  $f_s = 1\ 000\ 000$  Hz.

One array element is set as a drive to transmit signal, and the other eight elements as the sensor to receive the reflection signal. Each array element stimulates the signal in turns, then each degree corresponds to  $9 \times 8$  signals, and  $9 \times 8 \times 181$  sets of data can be obtained. The  $90^\circ$  direction of the data emitted by No. 0 array element and received by No. 1 array element is selected as the experimental data, and the processing method of other angles is consistent with this. The time domain waveform of the data set is shown in Fig. 1.

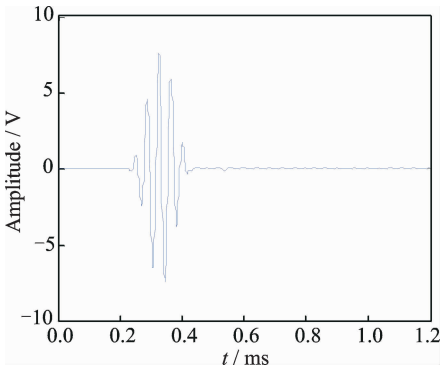


Fig. 1 Waveform of original signal in time domain

At first, PCA is used to deal with the waveform obtained by the  $90^\circ$  direction of the phased array signal emitted by No. 0 array element and received by No. 1 array element. The sparse representation of the original signal is obtained, as shown in Fig. 2. It can be seen that the sparse coefficient of the phased array signal after PCA transform is mostly zero or close to zero, which is consistent with the characteristic of sparse signal.

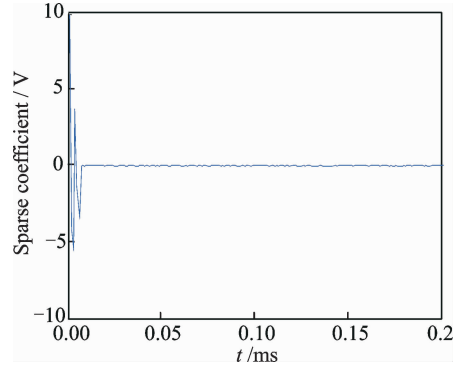


Fig. 2 Sparse coefficient after principal component analysis

Then, the length of ultrasonic phased array data is  $N=1\ 200$ , and the number of observations  $M=400$  is selected to complete the operation of signal projection observation, and the waveform is shown in Fig. 3.

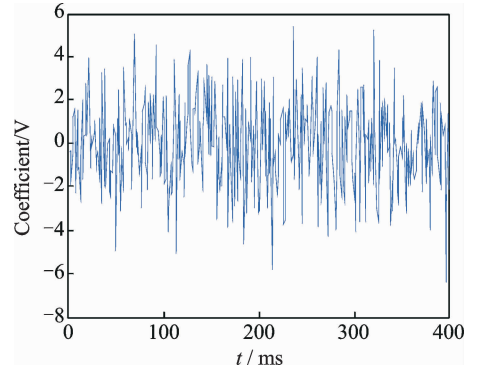


Fig. 3 Signal obtained by projection observation

Finally, the basis pursuit algorithm is used to deal with the ultrasonic phased array signal, and the reconstructed signals obtained are shown in Fig. 4.

### 4 Experimental Error Analysis

The reconstructed signal based on orthogonal matching pursuit algorithm has some differences in the signal waveform, compared with the original phased array signal. In order to analyze the effect of the reconstruction algorithm more accurately, the reconstructed error with different reconstruction algorithm is displayed numerically, as shown in Table 1. The absolute error  $\Delta V$  and the relative error  $\delta$  are calculated as below

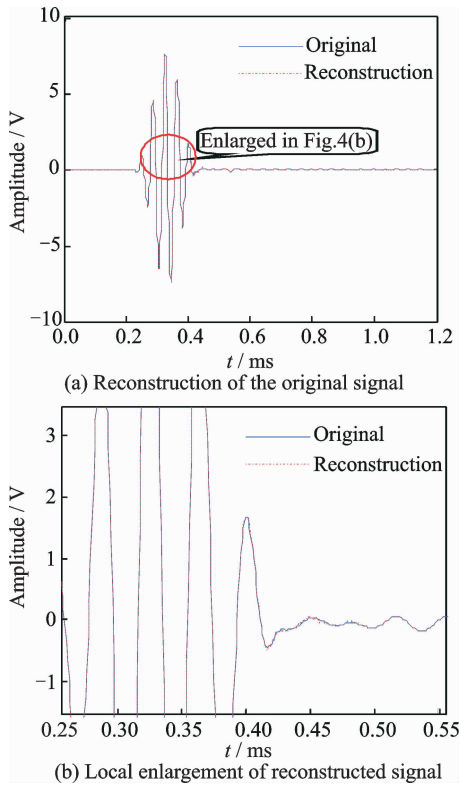


Fig. 4 The signal reconstruction based on basis pursuit algorithm

$$\Delta V = |V_0 - V_1| \quad (15)$$

$$\delta = \left| \frac{V_0 - V_1}{V_1} \right| \times 100\% \quad (16)$$

where  $V_0$  is the amplitude of reconstructed signal at the point of maximum amplitude deviation, and  $V_1$  the amplitude of original phased array signal at the same point.

Table 1 Reconstruction error

Algorithm	Absolute error	Relative error
BP	0.187 1	0.39
GPSR	2.932 5	0.061 6
OMP	0.061 6	0.065 9

In Table 1, GPSR is gradient projection for sparse reconstruction algorithm, and OMP is orthogonal matching pursuit algorithm. Fig. 5 shows the reconstructed signal obtained by BP, GPSR and OMP, respectively.

Table 2 shows the error comparison of some common transform base and the principal component analysis method.

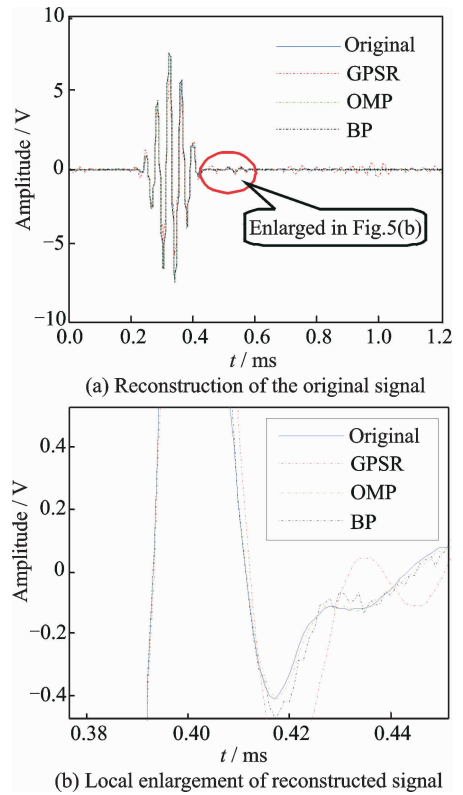


Fig. 5 The signal reconstruction based on BP, GPSR and OMP

Table 2 The error comparison

Orthogonal transformation	Absolute error	Relative error
PCA	0.187 1	0.003 9
PCT	0.894 3	0.018 8
DFT	21.022 9	0.012 7

In Table 2, DCT is discrete cosine transform, and DFT is discrete fourier transform.

The analysis of experimental error indicates that the relative error is relatively lower than that of commonly used method. That is, the proposed method can be applied to signal sparse representation of compressive sensing.

## 5 Conclusions

This paper studies the compressive sensing sparse sampling method based on PCA. This method not only solves the difficulty in storage and processing due to the large amount of data obtained by ultrasonic phased array structural

health monitoring, but also effectively improves the relationship between the original signal and the signal after sparse representation. And the experimental result shows that PCA can be used to reconstruct the signal obtained from the phased array structure health monitoring after sparse representation of the signal with small reconstruction error. In future research, we can choose more optimized projection observation matrix, and more efficient reconstruction algorithm to reconstruct the ultrasonic phased array signal.

### Acknowledgements

This work is supported by the National Natural Science Foundation of China (Nos. 51405241, 61672290), and the Jiangsu Government Scholarship for Overseas Studies and the PAPD Fund.

### References:

- [1] LIU S, GU G, ZHANG Q, et al. Principal component analysis algorithm in video compressed sensing [J]. *International Journal for Light and Electron Optics*, 2014, 125(3): 1149-1153.
- [2] MASIERO R, QUER G, MUNARETTO D, et al. Data acquisition through joint compressive sensing and principal component analysis[C]// *IEEE Conference on Global Telecommunications, USA; IEEE*, 2009 :1271-1276
- [3] ZHANG Y, XU C, LI C, et al. Wood defect detection method with PCA feature fusion and compressed sensing[J]. *Journal of Forestry Research*, 2015, 26(3): 745-751.
- [4] SINGH A, SHARMA L N, DANDAPAT S. Multi-channel ECG data compression using compressed sensing in eigenspace[J]. *Computers in Biology & Medicine*, 2016, 73: 24-37.
- [5] WOLD S. Cross-validatory estimation of the number of components in factor and principal components models[J]. *Technometrics*, 1978, 20(4): 397-405.
- [6] KU W, STORER R H, GEORGAKIS C. Disturbance detection and isolation by dynamic principal component analysis[J]. *Chemometrics and Intelligent Laboratory Systems*, 1995, 30(1): 179-196.
- [7] DONOHO D L. Compressed sensing [J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [8] YIN H, LIU Z. Survey of compressed sensing[J]. *Control and Decision*, 2013, 28(10): 1441-1445.
- [9] SUN Y J, GU F H. Compressive sensing of piezoelectric sensor response signal for phased array structural health monitoring[J]. *International Journal of Sensor Networks*, 2017, 23(4): 258-264.
- [10] LI Ran, GAN Zongliang, ZHU Liangchang. Smoothed projected Landweber image compressed sensing reconstruction using hard thresholding based on principal components analysis[J]. *Journal of Image & Graphics*, 2013(5):504-514. (in Chinese).
- [11] DIETZ B, YIP E, YUN J, et al. Real - time dynamic MR image reconstruction using compressed sensing and principal component analysis (CS-PCA): Demonstration in lung tumor tracking [J]. *Medical Physics*, 2017, 44 (8) :3978
- [12] LI J, SONG M, PENG Y. Infrared and visible image fusion based on robust principal component analysis and compressed sensing [J]. *Infrared Physics & Technology*, 2018, 89: 129-139.
- [13] CANDES E. The restricted isometry property and its implications for compressed sensing[J]. *C R Math Acad Sci Paris* 2008, 346(9-10): 589-592.
- [14] ZHANG B Y, CHEN H H, WANG R C, et al. New random signal generating method of multiple excitation vibration system based on white noises(in chinese)[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2017, 49(6): 839-844.
- [15] CHEN T, ZHU G, LIU Y. Optimization research of Gauss matrices in compressive sensing[J]. *Application Research of Computers*, 2014, 31(12): 3599-3602.
- [16] QI M X, CHEN S J, ZHOU S P, et al. Crack Detection in Pipes with Different Bend Angles Based on Ultrasonic Guided Wave[J]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2017, 34(3): 318-325.
- [17] LI Y, HE S. Blind signal separation algorithm for non-negative matrix factorization based on projected gradient[J]. *Computer Engineering*, 2016, 42(2): 104-107.
- [18] ZHU Y, ZHU Y, PENG Y. Adaptive compressive sensing and tracking of dynamic sparse spectrum[J]. *Journal of Signal Processing*, 2016, 32(3): 341-348.

Dr. **Sun Yajie** received M. Sc. degree in 2006, and her Ph. D. degree in 2010 from Nanjing University of Aeronautics

and Astronautics. Now she is an associate professor in Nanjing University of Information Science and Technology. Her current research interests include structural health monitoring and signal processing.

Ms. **Gu Feihong** received B. Sc. degree in 2015 from Binjiang College of Nanjing University of Information Science and Technology. Now she is a postgraduate in Nanjing University of Information Science and Technology. Her current research interests include structural health monitoring and signal processing.

Dr. **Ji Sai** received B. Sc. degree from Nanjing Institute of

Meteorology in 1999, his M. Sc. degree from Nanjing University of Aeronautics and Astronautics in 2006, and his Ph. D. degree from Nanjing University of Aeronautics and Astronautics in 2014. Now he is an associate professor in Nanjing University of Information Science and Technology. His current research interests include wireless sensor networks and computer measurement and control.

Ms. **Wang Lihua** is now a senior experimentalist in Nanjing University of Information Science and Technology. Her current research interests include engineering test technology, mechanical design and manufacture.

(Production Editor: Zhang Bei)



