Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **5**, September **2016**

A New Undetermined Blind Source Separation Method based on Orthogonal Projection

YanFei Jia[1,2],Shanshan Liu[3,1],Xiaodong Yang[1]

[1]College of Information and Communication Engineering
Harbin Engineering University
No.145 Nantong Street,Nangang District, Harbin City, China
jia_yanfei@163.com

[2]College of Electric and Information Engineering
Beihua University
No.1 Xinshan Street, Longtan District, Jilin City, China

[3]College of Information Engineering
Northeast Dianli University
No.169 Chuangchun Road,Chuanying District, Jilin City, China

Received April, 2016; revised June, 2016

Abstract. *The undetermined blind source separation algorithm based on compressed sensing applies the estimated matrix as the basic matrix, and constructs the sensing matrix through identity matrix with n order to extend the element of estimated matrix. The n here signifies the number of sample points. Therefore, with the increase of the number of sample points, the dimension of sensing matrix will be larger, so more computing memory will be needed, leading to the decrease in the computation speed of the algorithm based on compressed sensing. To overcome the problem, an optimized algorithm based on orthogonal projection is proposed. The new algorithm firstly transforms the observed signals to frequency domain by Fourier transform, and then separately computes the projection of the real and imaginary parts of frequency domain signals in every column direction of the estimated matrix. Subsequently, the column whose frequency signal has the maximal projection as separated vector is determined and utilized to separate the source signal of frequency domain. In the end, inverse Fourier transform is used to transform the frequency domain signal into time domain signal so as to realize the estimation of source signal. The simulation results prove that the proposed algorithm is featured with faster computation speed and lower error compared with the algorithm based on the compressed sensing.*

Keywords: Undetermined blind source separation, Orthogonal projection, Compressed sensing, Sensing matrix.

1. **Introduction.** The undetermined blind source separation is a method that can separate the source signal from the observed signal without any prior knowledge of source signal or mixing matrix in the condition where the number of source signal is greater than that of the observed signal[1-2]. It has been widely applied in processing speech signal , fault diagnosis, electromagnetic interference, biomedical signal and processing communication signal , etc.[3-8]. The procedure of undetermined blind source separation can be divided into two stages: the first is to estimate the mixing matrix with approaches like clustering [9-10], support vector machine [11], potential function [12], tensor decomposition [13] etc. The second is to estimate the source signal by means of linear programming [14], the shortest path decomposition [1], compressed sensing [15], etc [16-18], based on the

estimated matrix obtained in the first step. The current research on undetermined blind source separation mainly focuses on estimating the mixing matrix and reconstructing the source signal. In this paper, how to reconstruct the source signal is mainly studied.

Based on compressed sensing, the undetermined blind source separation utilizes identity matrix to extend the estimated matrix to construct sensing matrix. The dimension of sensing matrix is (where m and n represent respectively the number of row and the column of estimated matrix, and refers to the number of sample points). Hence, if the sample point is large, the dimension of sensing matrix will become large accordingly, and a lot of computing memory will be needed. Consequently, the ordinary computer cannot compute the sensing matrix when the number of sample points exceed five thousand. To reduce the memory requirement, the estimated matrix is adopted directly as the sensing matrix without changing the observed signal into one column signal. In addition, the orthogonal projection method is utilized to find out the optimal separated vector to reconstruct the corresponding source signal.

The rest of the paper is organized as follows: Section 2 undetermined blind source separation model is introduced. Section 3 describes the undetermined blind source separation based on compressed sensing. The proposed algorithm is introduced in details in Section 4. Simulation results analysis and discussion are presented in Section 5. In Section 6, the work of this paper is concluded and the drawback and the future work are proposed.

2. **Model.** In this paper, the following undetermined blind source separation model without noisy is considered

$$x(t) = As(t) \qquad t = 1, 2, \cdots, N \tag{1}$$

where $x(t)= [x_1(t),x_2(t),\cdots,x_m(t)]^{\mathrm{T}}$ signifies the vector of observed signal with index t, $A \in R^{m\times n}$ refers to the unknown mixing matrix where the value of m is smaller than n, $s(t)= [s_1(t),s_2(t),\cdots,s_n(t)]^{\mathrm{T}}$ referring to the vector of source signal. The undetermined blind source separation aims to estimate the source signal from the observed signal without any knowledge about the mixing matrix and source signal. If the inverse of mixing matrix A exists, will be estimated and the source signal will be obtained through $s(t) = A^{-1}x(t)$. However, Due to $m < n$, the inverse of will not exist, and the problem will become a non-deterministic polynomial and hard ill-conditioned problem. Therefore, different from the traditional blind source separation, the solution procedure of the undetermined blind source separation can not estimate the source signal directly with the separated matrix, which is the inverse of mixing matrix. In undetermined blind source separation algorithm, the mixing matrix and the source signal must be estimated separately [1-2]. In the undetermined blind source separation, the source signal is assumed to be sparse signal, where only a few sample points are non zero or far larger than zero and the others approach to zero or equal to zero at the same time.

3. **Undetermined Blind Source Separation Based on Compressed Sensing.** As a new technique to sample signal with below Nyquist rate, compressed sensing can reconstruct the signal from the sample signal with lower sample rate[19-22]. The model of compressed sensing is as follows [22].

$$y = \Phi x = \Phi\psi s = \Theta s \tag{2}$$

where $y = \Phi x = \Phi\psi s = \Theta s$ refers to the compressed signal with m length.x signifies one dimension original signal with n length, which is sparse or can be expressed as sparse in some sparse basis. $y = \Phi x = \Phi\psi s = \Theta s$ means the measured matrix with $m < n.\psi$ is sparse basis matrix that ensures the original signal to be sparse while s signifies sparse vector; and $\Theta \in R^{m*n}$ indicates the sensing matrix.

The recovery of source signal is realized by applying the method of compression and reconstruction, which is expressed by the inverse operation of formula (2). By contrasting the compression sensing mathematical model (2) to the undetermined blind source separation mathematical model (1), it can be confirmed that the principles of the two models are basically the same. It can be observed that the signal can be reconstructed with good sparse in the undetermined blind source separation of the source signal while the compressed sensing algorithm requires the sparse source signal. Hence, the compressed sensing reconstruction algorithm is also suitable for the signal recovery in the blind source separation [15]. The reconstruction process of compressed sensing is to recover the signal , which is similar to the estimation of the source signal in undetermined blind source separation. Therefore, the compressed sensing method can be adopted to estimate the source signal in the condition that the mixing matrix has been estimated in undetermined blind source separation [15].

The undetermined blind source separation based on compressed sensing [15] includes two stages: the first is to construct one-dimensional mixing signal. For example, the number of observed signals is assumed to be $m = 2$ with t length. The one-dimensional mixing signal with $2t$ length will thus become the reconstruction process of compressed sensing to recover the signal x ,which is similar to the estimation of the source signal in undetermined blind source separation. The one dimensional mixing signal is $mt \times 1$ vector according to the compressed sensing principle, and the corresponding source signal will be $nt \times 1$ vector. Thus, the sensing matrix Θ will be $mt \times nt$ dimension.To make the sensing matrix meet the demand, the identity matrix with t order is utilized to extend the element of estimated matrix $\tilde{A}$ so as to construct the sensing matrix. The source signal reconstruction can be expressed as[15]:

$$y = \Theta s = \begin{pmatrix} \Lambda_{11}, \Lambda_{12}, \cdots, \Lambda_{1n} \\ \vdots \\ \Lambda_{m1}, \Lambda_{m2}, \cdots, \Lambda_{mn} \end{pmatrix} s \tag{3}$$

where $y = (y_{11}, y_{12}, \cdots, y_{1t}, \cdots y_{m1}, y_{m2}, \cdots, y_{mt})^T$,$\Lambda_{ij} = E_t a_{ij}$, E_t refers to identity matrix with t order. t signifies the number of sample points.a_{ij} means the element of $\tilde{A}$. $\mathbf{s}$= $(s_{11}, s_{12}, \cdots s_{1t}, \cdots, s_{n1}, \cdots s_{nt})^{\mathrm{T}}$ indicates the reconstruction signal.

This algorithm focally considers the sensing matrix as a redundant dictionary, where orthogonal matching pursuit algorithm is applied to each atom corresponding to each column of the redundant dictionary so as to find the optimal matching source signal points.

4. **Proposed method.** From the above analysis on the undetermined blind source separation based on the compressed sensing in section 3, it can be found that in (3) the sensing matrix in the signal reconstruction is extended through the identity matrix with t order. t signifies the number of sample points, so the dimension of the sensing matrix increases with the rise of sampling points. When it reaches a certain number of dimensions, it may be difficult to successfully construct the sensing matrix due to the limitated equipment memory. To solve this problem, the following optimized algorithm is applied to reconstruct signal in undetermined blind source separation.

In the signal reconstruction stage, the undetermined blind source separation based on compressed sensing [15] changes the multidimensional observed signal into one dimensional signal by filling all sample points into one column, and constructing sensing matrix with the size which is t times larger than that of the mixing matrix, where t is the number of sample points. In this way, the algorithm needs more computing memory to store the sensing matrix accordingly. It cannot even reconstruct the source signal with the ordinary computation for large sample.

To overcome the above problems, the dimensional of the observed signal and sensing matrix is not changed in the signal reconstruction stage. The column of estimated matrix and the frequency signal of observed signal obtained by Fourier transform are applied as basis vector and compressed signal respectively. It can be expressed as

$$y = As \tag{4}$$

where $y \in R^{m*t}$ refers to the frequency signal of observed signal with the same size; $A \in R^{m*n}$ signifies the sensing matrix. It can be found that formula (4) is almost the same as (3), but the sizes of sensing matrix $\Theta \in R^{mt*nt}$ in (3) are much larger than those of A in (4), so in the same condition, the memory to store the sensing matrix Θ in (3) is t times than in (4), in which t is the number of sample point. To conclude, formula (3) increases the load of computation and reduces the computation speed.

The orthogonal projection method is utilized to compute the projection of observed signal y in the column direction of estimated matrix and the estimated matrix column that corresponds to the maximal projection is adopted to construct a new estimated matrix. Subsequently, the source signal in frequency domain can be reconstructed by multiplying the inverse of the new estimated matrix and the frequency signal of observed signal. The whole procedure is as follows.

A. Set the maximal iterations;
B. Transform the observed signals to frequency domain, and normalize the transformed signal;
C. Apply potential function method to estimate the mixing matrix and obtain the estimated matrix;
D. Set the residual value of to be the observed signal $y(:,k)$ in (4). Here k represents the current iteration number;
E. Calculate the projection value of real and imaginary parts respectively on each column vector of mixing matrix;
F. Sort the projection value of real and imaginary parts respectively in a descending order, and utilize the column vectors that correspond to the first m elements of sorted projection to construct new estimated matrix β_r ,and β_r;
G. Reconstruct the real and imaginary parts of source signal respectively by the following formulas

$$\begin{aligned} &s_r(j,k) = \beta_r^{-1} r_r \\ &s_i(j,k) = \beta_i^{-1} r_i \\ &s_r(others,k) = 0 \\ &s_i(others,k) = 0 \end{aligned} \tag{5}$$

where s_r and s_i are the real and imaginary parts of recovered signal respectively,r_r and r_i refer to the real and imaginary parts of r ; $j \in R^{m*1}$ signifies the column vector index of the first m elements of sorted projection.
H. When $k = k + 1$, return to step D. Stop until the current iteration equals to the maximal iteration.

5. **Simulations and Results.** In this paper, Fourier transform is utilized to transform the time domain of observed signal that is not sparse to frequency domain to construct the sparse observed signal in frequency, and in the end the inverse Fourier transform is utilized to transform the estimated signals in frequency domain to time domain.The frequency domain signal is normalized and the mixing matrix is estimated with the method based on potential function.

Since the undetermined blind source separation algorithm based on compressed sensing cannot be applied to a large number of sampling points, in this paper firstly, a small sample is firstly adopted to verify the validity of the algorithm. A sinusoidal signal and two demodulated signals with different frequencies are used as the original signal. The figure 1 describes the original signal, while the figure 2 for the mixed signal (observed signals). The undetermined blind source separation method based on compressed sensing and the proposed method is used to recover the source signal respectively.The figure 3 refers to the separated signal from the proposed algorithm. By comparing the original signal in figure 1 with the separated signal in figure 3, it can be seen that their waveforms are quite similar except amplitude, which is one of inherent indetermination of blind source separation. this proves that the proposed method is effective.

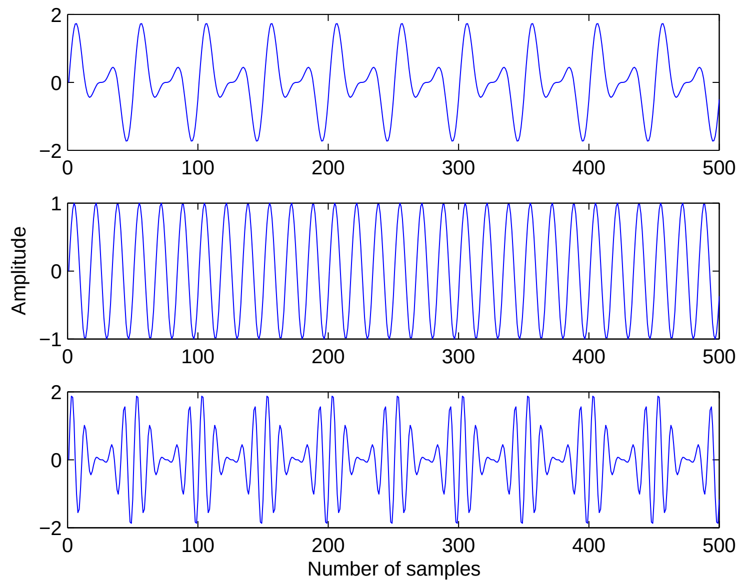

FIGURE 1. Original signals

To further illustrate the effectiveness of this algorithm, the signal to interference ratio (SIR) of quantized form is introduced [15] to offer a more accurate description of the reconstructed signal. The first SIR of ith channel signal can be computed by:

$$SIR_i = 10\lg\frac{\|S_i\|_2}{||S_i - \tilde{S}_i\ \ ||_2} \tag{6}$$

where S_i is the ith source signal and $\tilde{S}_i$ is the estimated signal (recovered signal) of ith source signal $\tilde{S}$. The bigger the value of SIR is, the smaller the error will be. Therefore, when the SIR value is relatively high, the signal recovery error is relatively small. Generally, for $SIR > 10$, the recovered signal is considered to be better.

Table 1 displays a comparison of SIR between the algorithm base on compressed sensing [15] and the proposed algorithm. From table I, it can be observed that the SIR value of the proposed algorithm is bigger than that of the algorithm[15] for all source signals, indicating that the proposed algorithm has smaller error. Table 2 displays the time consumption of different numbers of samples for recovered signal after estimating mixing matrix. By analyzing the table 2, it can be observed that for any number of samples, the time consumption of the proposed algorithm is smaller than the algorithm based on compressed sensing, which is just about one tenth of that for the algorithm based on

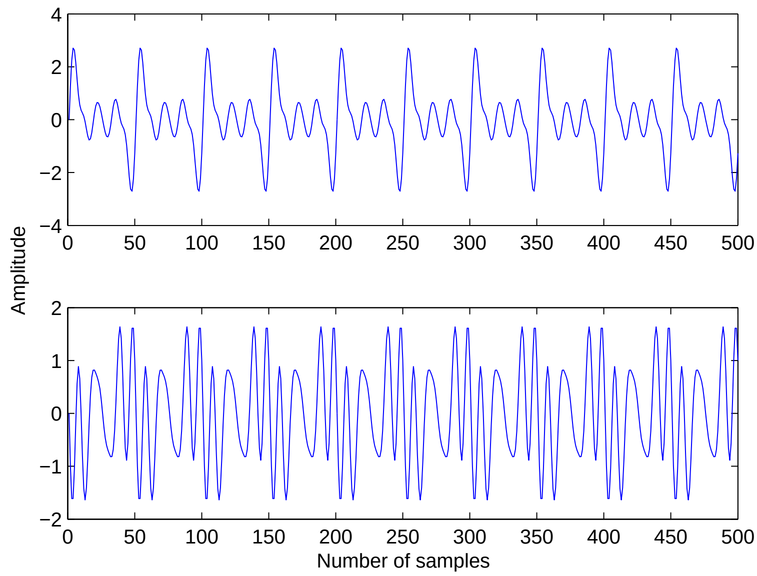

FIGURE 2. Mixing signals

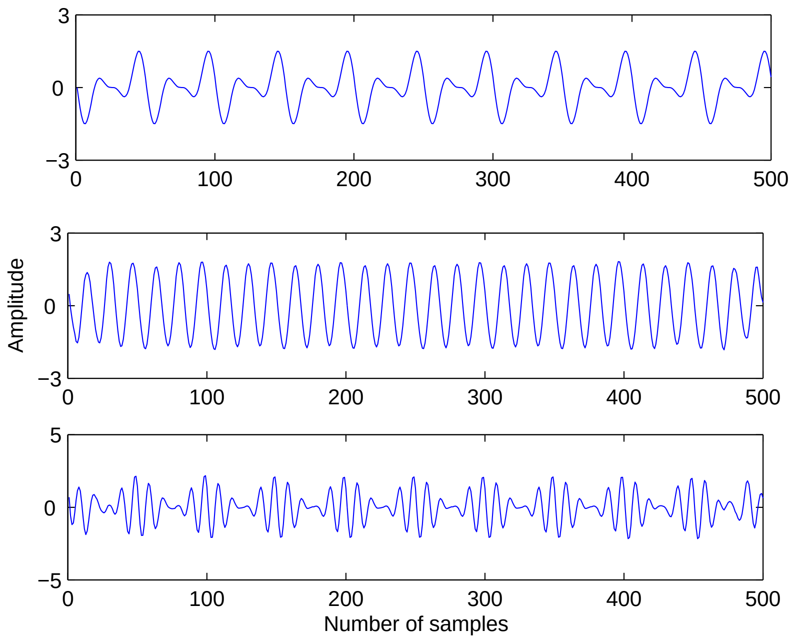

FIGURE 3. Separated signals

compressed sensing. This verifies that the proposed algorithm owns faster computing speed and reduces the complexity of the algorithm based on compressed sensing.

In order to verify the effectiveness of the algorithm for practical signal, in the paper, the voice signal paths of three different people is randomly selected from the speech database TIMIT as input. Similarly, the potential function method is utilized to estimate the mixing matrix. Figure 4 demonstrates the speech signal waveform, while Figure 5 shows the mixed signal. Figure 6 illustrates recovered signal by proposed algorithm. Table 3 and Table 4 describe the SIR and time consumption of the proposed algorithm separately.

By comparing the figure 4 with figure 6, table 3 and table 4, it can be concluded that the proposed algorithm is effective for practical speech signal.

Table 1. SIR of methods

Original signals	S1	S2	S3
Algorithm [15]	16.7074dB	19.6794dB	18.2420dB
Proposed algorithm	25.8273dB	27.3002dB	25.7394dB

Table 2. Time consumption with different samples

Number of samples	300	500	1000	1500
Algorithm [15]	1.284s	3.796s	9.834s	12.573s
Proposed algorithm	0.3411s	0.627s	0.912s	1.629s

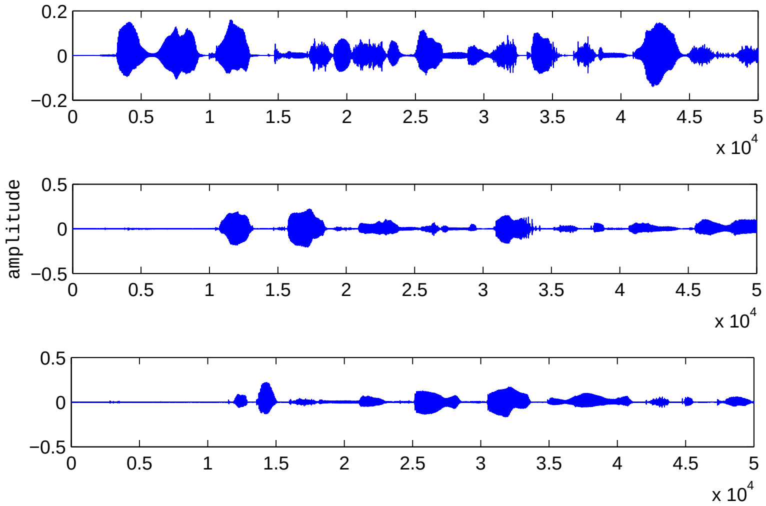

Figure 4. Original speech signals

Table 3. SIR of Proposed signal with speech signal

Original signals	S1	S2	S3
Proposed algorithm	14.463dB	13.112dB	17.128dB

The same speech signal is utilized to test the algorithm [15] through computer with 2G memory. However, the results cannot be obtained in an hour. Consequently, the algorithm [15] for the speech signal with a large number of samples is not compared. According to previous experience, if the number of samples excesses five thousand, the results cannot be obtained in an hour if utilizes the algorithm based on compressed sensing.

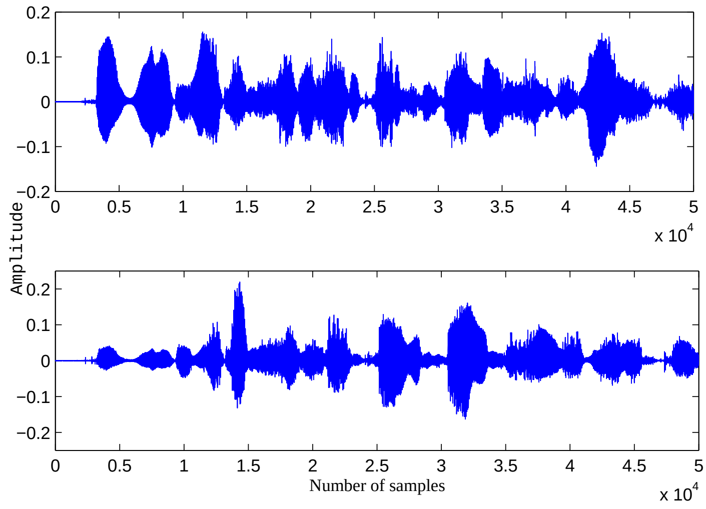

FIGURE 5. Mixing speech signals

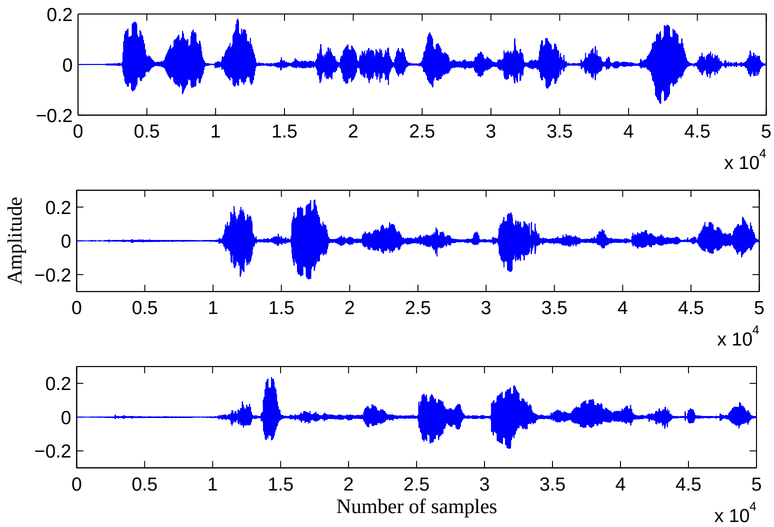

FIGURE 6. Separated speech signals

TABLE 4. Time consumption of different samples with speech signal

Number of samples	3000	10000	25000	40000
Proposed algorithm	0.829s	2.583s	5.015s	7.792s

6. **Conclusions.** In this paper, an improved undetermined blind source separation based on orthogonal projection is proposed. In particular, it directly takes estimated matrix as sensing matrix to recover the source signal with sample point one by one. In this way, the requirement of memory can be lowered. Compared with the undetermined blind source separation based on compressed sensing, the proposed algorithm shows faster computing speed and fewer errors in the same condition. However, the drawback is that the proposed algorithm can not be suitable for the source signals whose number of non zero signals is

more than the number of observed signals in the same time, which will be our further study.

REFERENCES

[1] H. Bofill, M. Zibulevsky, Blind separation of more sources than mixtures using sparsity of their short-time Fourier transform, *in Proc. Of int .Workshop on ICA and BSS*, Helsinki, pp.87–92, 2000.

[2] P. Bofill, M. Zibulevsky, Undetermined blind source separation using sparse representation, *Signal Processing*, vol.81, no.11, pp.2353–2362, 2001.

[3] G. Hui, F. Y. Qing, S. D. Lin, and L. Yan, Method to separate electromagnetic interference sources based on underdetermined blind source separation, *Journal of Jilin University(Engineering and Technology Edition)*, vol.45, no.4, pp.1329–1335, 2005.

[4] Z. Chao, G. Qi, G. L. Jie, and L. H. Lin, Underdetermined blind source separation method of rotating machinery faults based on local characteristic-scale decomposition, *Journal of Yanshan University*, vol.38, no.2, pp.168–174, 2014.

[5] L. L. Cui, C. G. Wu, G. Q. Ma, and H. Q. Wang, Diagnosis of Roller Bearings Compound Fault Using Underdetermined Blind Source Separation Algorithm Based on Null-Space Pursuit, *Shock and Vibration*, vol.2015, pp.1–8, 2015.

[6] J. Y. Hao,and X. P. Fei, Underdetermined blind source separation applied to the ultra-wideband Communication signals, *Journal of Nanjing University of Post and Telecommunications(Natural science)*, vol.31, no.1, pp.23–28, 2011.

[7] S. P. Lv, C. C. Liu, C. Zhang, and J. A. Wen, Underdetermined Blind Speech Signal Separation Method Based on the Improved Shortest Path Method, *Advance in swarm and computational intelligence*, vol.9141, pp.397–404, 2015.

[8] Z. Koldovsky, A. H. Phan, P. Tichavsky, and A. Chichocki, A treatment of EEG data by underdetermined blind source separation for motor imagery classification, *20th European Signal processing conference*, Bucharest, pp.27–31, 2012.

[9] J. Thiagrajan, K. N. Ranmamurthy, and A. Spanias, Mixing matrix estimation using discriminative is clustering for blind source separation, *Digital Signal Processing*, vol.23, no.1, pp.9–18, 2012.

[10] X. J. Bi, R. J. Gong, Undetermined blind mixing matrix estimation algorithm based on mixing clustering and mesh density, *System Engineering and Electronics*, vol.34, no.3, pp.614–618, 2012.

[11] R. H. Li, Z. Y. Yang, M. Zhao, and S. L. Xie, SVM based Undetermined blind source separation, *Journal of electronics and Information Technology* , vol.31, no.2, pp.320–322, 2009.

[12] W. H. Fu, L. Wang, and L. F. Ma, Improved laplace mixed model potential function algorithm for UBSS, *Journal of Xidian University* , vol.41, no.6, pp.1–5, 2014.

[13] Y. L. Zhang, S. T. Lou, W. T. Zhang, Estimation of undetermined mixture matrix in blind source separation based on tensor decomposition, *System Engineering and Electronics*, vol.33, no.8, pp.1073–1076, 2011.

[14] Y. Q. Li, S. Amari, A. Cichocki, and D. W. C. Ho, Underdetermined Blind Source Separation Based on Sparse Representation, *IEEE Trans. on Signal Processing*, vol.54, no.2, pp.423–437, 2006.

[15] L. N. Li, Q. X. Cheng, and X. Y. Gan, Under-determined blind source separation based on potential function and compressive sensing, *Journal of computer applications*, vol.34, no.3, pp.658–662, 2014.

[16] S. N. Ge, J. Han, M. Han, Nonnegative mixture for underdetermined blind source separation based on a tensor algorithm, *Circuits System signal process*, vol.34, no.9, pp.2935–2950, 2015.

[17] J. D. Sun, Y. X. Li, J. T. Wen, and S. G. Yang, Novel mixing matrix estimation approach in underdetermined blind source separation, *Neurocomputing*,[Online].pp. 1-10. Available: http://dx.doi.org/10.1016/j.neucom.2015.08.008i, 2015.

[18] W. Li, H. Yang, Blind Source Separation of Underdetermined Mixtures Based on Local Mean Decomposition and Conjugate Gradient Algorithm, *J. of communication*, vol.9, no.5, pp.425–432, 2014.

[19] H. C. Huang, F. C. Chang, Error Resilience for Compressed Sensing with Multiple-Channel Transmission, *J. of Information Hiding and Multimedia Signal Processing*, vol.6, no.5, pp.847–856, 2015.

[20] X.Zhang, A. Wang, B.Zeng, L.Lei, Adaptive Distributed Compressed Video Sensing, *J. of Information Hiding and Multimedia Signal Processing*, vol.5, no.1, pp.98–106, 2014.

[21] J. S. Pan, W. Li, C. S. Yang, L. J. Yan, Image steganography based on subsampling and compressive sensing, *Multimedia Tools and Applications*, vol.74, no.21, pp.9191–9205, 2015.

[22] K. Gitta(eds.), *Compressed sensing: theory and application*, New York: Cambridge Univ, 2012.