Front. Electr. Electron. Eng. China 2008, 3(3): 338–342
DOI 10.1007/s11460-008-0015-5

RESEARCH ARTICLE

Weihong FU, Xiaoniu YANG, Nai'an LIU, Xingwen ZENG

# Blind source separation algorithm for communication complex signals in communication reconnaissance



**Abstract** Most blind source separation algorithms are only applicable to real signals, while in communication reconnaissance processed signals are complex. To solve this problem, a blind source separation algorithm for communication complex signals is deduced, which is obtained by adopting the Kullback-Leibler divergence to measure the signals' independence. On the other hand, the performance of natural gradient is better than that of stochastic gradient, thus the natural gradient of the cost function is used to optimize the algorithm. According to the conclusion that the signal's mixing matrix after whitening is orthogonal, we deduce the iterative algorithm by constraining the separating matrix to an orthogonal matrix. Simulation results show that this algorithm can efficiently separate the source signals even in noise circumstances.

**Keywords** communication reconnaissance, blind source separation, complex signal

## 1 Introduction

Communication reconnaissance is one of the key technologies in electronic warfare. The process of analyzing and recognizing enemy communication signals without any prior information is called blind reconnaissance. To that end, we must separate source signals from mixed signals, that is, we need to estimate every source signal in time domain, which is called blind source separation (BSS). Recently, BSS is applied in wireless communication, image processing, seismic signals processing, array signal processing and biomedicine signal processing. While only few researches [1–3] applied the BSS technology into communication reconnaissance. Furthermore, many known algorithms are applicable to real valued signals [4] only instead of complex signals. However, in communication reconnaissance, processed signals are often complex signals. Thus, a blind source separation algorithm for communication with complex signals is presented. The simulation results prove the validity of the algorithm.

Translated from *Journal of Huazhong University of Science and Technology* (*Natural Science Edition*), 2007, 35(4): 33–36 [译自: 华中科技大学学报 (自然科学版)]

Weihong FU (✉), Nai'an LIU, Xingwen ZENG
State Key Laboratory of Integrated Service Networks, Xidian Univeristy, Xi'an 710071, China
E-mail: whfu@mail.xidian.edu.cn

Xiaoniu YANG
National Laboratory of Information Control Technology for Communication System, Jiangnan Electronic Communications Research Institute, Jiaxing 314033, China

## 2 Problem presentation

Assume the vector $\boldsymbol{x}(t)=[x_1(t)\ \ x_2(t)\ \ \cdots\ \ x_n(t)]^{\mathrm{T}}$ is the observed signals from $n$ antennas and the vector $\boldsymbol{s}(t)=[s_1(t)\ \ s_2(t)\ \ \cdots\ \ s_m(t)]^{\mathrm{T}}$ is the source signals from $m$ sources, where $x_i(t)$ denotes the observed signal of $i$th $(i=1, 2,\ldots,n)$ receiving antenna at $t$ time and $s_i(t)$ is the transmitted signals of the $i$th source at $t$ time. Here $\boldsymbol{x}(t)$ and $\boldsymbol{s}(t)$ are complex vectors. Assume that every source $s_i(t)$ $(i=1, 2,\ldots,m)$ is independent. If the mixing matrix is $\boldsymbol{H}\in C^{n\times m}$, then $\boldsymbol{x}(t)=\boldsymbol{H}\boldsymbol{s}(t)+\boldsymbol{v}$, where $\boldsymbol{v}$ is white Gaussian noise. We assume $n\geqslant m$.

The blind source separation technology is, according to the observed signal vector $\boldsymbol{x}(t)$, finding a separating matrix $\boldsymbol{W}$ and making every component of the separating signals $\boldsymbol{y}(t)=\boldsymbol{W}\boldsymbol{x}(t)$ as independent as possible. The paper studies the linear instantaneous mixing model of communication signals. Therefore, for the convenience of description, we will omit the parameter $t$. And the circumstance noise is considered so faint that it can be omitted.

## 3 Algorithm deduction

### 3.1 Selection of cost function

To get the good estimation $\boldsymbol{y}=\boldsymbol{W}\boldsymbol{x}$ of the source signal $\boldsymbol{s}$, we introduce an object function or loss function $\rho(\boldsymbol{y};\boldsymbol{W})$ about the separating signal vector $\boldsymbol{y}$ and the separating matrix $\boldsymbol{W}$. The expectation of this function is called risk function (or cost function): $R(\boldsymbol{W}) = E\{\rho(\boldsymbol{y};\boldsymbol{W})\}$, which is the measurement of independence of the components of the separating signal vector $\boldsymbol{y}$. When the components of the vector $\boldsymbol{y}$ is independent of each other, the risk function $R(\boldsymbol{W})$ will be minimized. The Kullback-Leibler (K-L) divergence is used to measure the independence, where $p_y(\boldsymbol{y};\boldsymbol{W})$ denotes the probability density function of the random variable $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$, and $q(\boldsymbol{y})$ is the other kind of probability density function of $\boldsymbol{y}$, then the risk function is [5]

$$\begin{aligned} R(\boldsymbol{W}) &= E\{\rho(\boldsymbol{y};\boldsymbol{W})\} = K_{pq}(\boldsymbol{W}) \\ &= K\left[p_y(\boldsymbol{y};\boldsymbol{W})\|q(\boldsymbol{y})\right] \\ &= \int p_y(\boldsymbol{y};\boldsymbol{W})\log\frac{p_y(\boldsymbol{y};\boldsymbol{W})}{q(\boldsymbol{y})}\,\mathrm{d}\boldsymbol{y}. \end{aligned} \tag{1}$$

The K-L divergence gives the difference measurement of two probability density function. When $q(\boldsymbol{y})$ is the true probability density function $p_s$ of the source signals, if $\boldsymbol{W} = \boldsymbol{H}^{-1}$, then $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x} = \boldsymbol{W}\boldsymbol{H}\boldsymbol{s} = \boldsymbol{s}$, and $p_y(\boldsymbol{y};\boldsymbol{W}) = p_s(\boldsymbol{y})$. Therefore, $K_{pq}(\boldsymbol{W})$ is an appropriate cost function. From Eq. (1), for the generic $q(\boldsymbol{y})$, we can have

$$\begin{aligned} R(\boldsymbol{W}) = E\{\rho(\boldsymbol{y};\boldsymbol{W})\} &= \int p_y(\boldsymbol{y};\boldsymbol{W})\log\frac{p_y(\boldsymbol{y};\boldsymbol{W})}{q(\boldsymbol{y})}\mathrm{d}\boldsymbol{y} \\ &= -\boldsymbol{H}(\boldsymbol{y}) - \sum_{i=1}^{n} E\{\log(q_i(y_i))\}, \end{aligned} \tag{2}$$

where

$$\boldsymbol{H}(\boldsymbol{y}) = -\int p_y(\boldsymbol{y};\boldsymbol{W})\log p_y(\boldsymbol{y};\boldsymbol{W})\mathrm{d}\boldsymbol{y}$$

denotes the differential entropy of $\boldsymbol{y}$. Assume that the probability density function of $\boldsymbol{z}$ is $p_z(\boldsymbol{z})$. Because

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{z},\ p_y(\boldsymbol{y}) = p_z(\boldsymbol{W}^{-1}\boldsymbol{y})|\det\boldsymbol{W}^{-1}|,$$

then

$$\begin{aligned} H(\boldsymbol{y}) &= -\int p_z(\boldsymbol{W}^{-1}\boldsymbol{y})|\det\boldsymbol{W}^{-1}|\left[\log(p_z(\boldsymbol{W}^{-1}\boldsymbol{y}))\right. \\ &\quad \left. + \log|\det\boldsymbol{W}^{-1}|\right]\mathrm{d}(\boldsymbol{W}\boldsymbol{z}) \\ &= -\int p_z(\boldsymbol{z})\left[\log(p_z(\boldsymbol{z})) + \log|\det\boldsymbol{W}^{-1}|\right]\mathrm{d}\boldsymbol{z} \\ &= H(\boldsymbol{z}) + \log|\det(\boldsymbol{W})|. \end{aligned}$$

Because $H(\boldsymbol{z})$ is independent of $\boldsymbol{W}$, it can be omitted. Then the final cost function is

$$\begin{aligned} R(\boldsymbol{W}) &= K_{pq}(\boldsymbol{W}) = E\{\rho(\boldsymbol{y};\boldsymbol{W})\} \\ &= -\log|\det(\boldsymbol{W})| - \sum_{i=1}^{m} E\{\log(q_i(y_i))\}. \end{aligned} \tag{3}$$

### 3.2 Whitening reprocessing of observed signals

The spatial decorrelation plays an important role in signal processing. It is considered a necessary (but not sufficient) condition of random independence rule. After pre-whitening, blinding source separation often becomes easier and the characteristics are more excellent. If the covariance matrix of rand vector $\boldsymbol{z}$ with zero mean is an identity matrix, i.e., $\boldsymbol{R}_{zz} = \boldsymbol{I}_n$, then we say the vector $\boldsymbol{z}$ is white. Whitening the observed signal vector $\boldsymbol{x}$ aims to find a whitening matrix $\boldsymbol{Q} \in C^{m\times n}$ and make the covariance matrix of the vector $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{x}$ an identity matrix. The whitening matrix $\boldsymbol{Q}$ can be found by singular value decomposition [5]. Assume the number of source signals $m$ is known or has been estimated, and then the whitening matrix $\boldsymbol{Q}$ is as follows:

$$\boldsymbol{Q} = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}_x^{-1/2}\tilde{\boldsymbol{V}}^{\mathrm{H}}, \tag{4}$$

where $\boldsymbol{U}$ is an arbitrary unitary matrix, $\tilde{\boldsymbol{\Lambda}}_x$ is the four-square matrix consisting of the first $m$ rows and first $m$ columns of the matrix $\boldsymbol{\Lambda}_x$, and $\tilde{\boldsymbol{V}}$ is the matrix consisting of the first $m$ columns of $\boldsymbol{V}$. It can be proven that the vector $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{x}$ is white. Because the matrix $\boldsymbol{U}$ is an arbitrary unitary matrix, the whitening matrix $\boldsymbol{Q}$ is not an exclusive matrix. Besides, it can be seen from Eq. (4) that after whitening, the dimension of the observed vector is reduced from $n$ into $m$.

### 3.3 Iteration algorithm of blind separation

After pre-whitening the observed signal vector, we can find the separating matrix $\boldsymbol{W}$ by whitened signal vector $\boldsymbol{z}$ and obtain the separated signal vector $\boldsymbol{y}$. Because $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{H}\boldsymbol{s}$, the global mixing matrix $\boldsymbol{B}$ is equal to $\boldsymbol{Q}\boldsymbol{H}$ and $\boldsymbol{B}$ is a non-singular square matrix. We assume that the covariance matrix of the source signal vector is identity matrix, i.e., $\boldsymbol{R}_{ss} = \boldsymbol{I}_m$. For $\boldsymbol{R}_{zz} = E\{\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}\} = E\{\boldsymbol{B}\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\} = \boldsymbol{I}_m$, $\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}} = \boldsymbol{I}_m$. Therefore, the total mixing matrix $\boldsymbol{B}$ is orthogonal. Under the condition, it only needs to find an orthogonal matrix $\boldsymbol{W}$ ($\boldsymbol{W}^{-1} = \boldsymbol{W}^{\mathrm{H}}$) to estimate the source signals from the signals $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{z}$. Certainly, the covariance matrix $\boldsymbol{R}_{ss}$ of the source signal vector is not equal to $\boldsymbol{I}_m$ but $\boldsymbol{R}_{ss} = \boldsymbol{D} = \mathrm{diag}(d_1,d_2,\ldots,d_m)$. Here the source signal vector $\boldsymbol{s}$ can be decomposed into $\boldsymbol{s} = \boldsymbol{D}^{1/2}\tilde{\boldsymbol{s}}$ and $\boldsymbol{R}_{\tilde{s}\tilde{s}} = \boldsymbol{I}_m$. If we can regard $\tilde{\boldsymbol{s}}$ as a source signal vector and $\tilde{\boldsymbol{B}} = \boldsymbol{Q}\boldsymbol{H}\boldsymbol{D}^{1/2}$ as $\boldsymbol{B}$, then we can obtain the same conclusion. Thus in the following sections, we always consider $\boldsymbol{R}_{ss} = \boldsymbol{I}_m$.

To accurately separate source signals, the object function $\rho(\boldsymbol{y};\boldsymbol{W})$ should be minimized. Therefore, the relative gradient learning algorithm is adopted. Then

$$\Delta\boldsymbol{W}(k) = \boldsymbol{W}(k+1) - \boldsymbol{W}(k) = -\eta(k)\frac{\partial\rho(\boldsymbol{y};\boldsymbol{W})}{\partial\boldsymbol{W}}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W}. \tag{5}$$

From Eqs. (3) and (5), we have

$$\Delta \boldsymbol{W}(k)=\eta(k)\left[\frac{\partial \log|\det \boldsymbol{W}(k)|}{\partial \boldsymbol{W}(k)} + \sum_{i=1}^{m}\left(\frac{\partial \log(q_i(\boldsymbol{y}_i(k)))}{\partial \boldsymbol{y}_i(k)}\frac{\partial \boldsymbol{y}_i(k)}{\partial \boldsymbol{W}(k)}\right)\right]\boldsymbol{W}(k)^{\mathrm{H}}\boldsymbol{W}(k). \quad (6)$$

It can be deduced from $\boldsymbol{y}(k) = \boldsymbol{W}(k)\boldsymbol{z}(k)$ that

$$\Delta \boldsymbol{W}(k)=\eta(k)\left(\boldsymbol{I}-\boldsymbol{f}[\boldsymbol{y}(k)]\boldsymbol{y}^{\mathrm{H}}(k)\right)\boldsymbol{W}(k), \quad (7)$$

where $\boldsymbol{f}[\boldsymbol{y}(k)] = [f_1(y_1), f_2(y_2), \ldots, f_m(y_m)]^{\mathrm{T}}$. Equation (7) denotes the iterative algorithm without any restriction on $\boldsymbol{W}$. As mentioned before, after pre-whitening, the total mixing matrix is an orthogonal matrix. Therefore, we suppose the separating matrix $\boldsymbol{W}$ is always orthogonal in the iterative process. Assume the matrix $\boldsymbol{W}(k)$ is orthogonal (the matrix $\boldsymbol{W}(k)$ can be initialized as the identity to satisfy the orthogonalization), for $\Delta \boldsymbol{W} = \boldsymbol{\varepsilon} \boldsymbol{W}(k)$, where $\boldsymbol{\varepsilon}$ is a small variational matrix, then

$$\boldsymbol{W}(k+1) = \boldsymbol{W}(k)+\boldsymbol{\varepsilon}\boldsymbol{W}(k)$$

and

$$\boldsymbol{W}(k+1)\boldsymbol{W}(k+1)^{\mathrm{H}} = \boldsymbol{I}+\boldsymbol{\varepsilon}+\boldsymbol{\varepsilon}^{\mathrm{H}}+O(\boldsymbol{\varepsilon}) = \boldsymbol{I} \Rightarrow \boldsymbol{\varepsilon} = -\boldsymbol{\varepsilon}^{\mathrm{H}}.$$

Therefore, the matrix $\boldsymbol{\varepsilon}$ must have skew symmetry to maintain the matrix $\boldsymbol{W}(k)$ orthogonal. Combining Eq. (7), the matrix $\boldsymbol{\varepsilon}$ can be selected as

$$\boldsymbol{\varepsilon} = \eta(k)[\boldsymbol{y}(k)\boldsymbol{f}^{\mathrm{H}}[\boldsymbol{y}(k)] - \boldsymbol{f}[\boldsymbol{y}(k)]\boldsymbol{y}^{\mathrm{H}}(k)].$$

The iterative equation of the orthogonal separating matrix $\boldsymbol{W}$ is as follows:

$$\begin{cases} \boldsymbol{W}(k+1)=\boldsymbol{W}(k)+\eta(k)\left[\boldsymbol{y}(k)\boldsymbol{f}^{\mathrm{H}}[\boldsymbol{y}(k)] - \boldsymbol{f}[\boldsymbol{y}(k)]\boldsymbol{y}^{\mathrm{H}}(k)\right]\boldsymbol{W}(k), \\ \boldsymbol{y}(k)=\boldsymbol{W}(k)\boldsymbol{z}(k). \end{cases} \quad (8)$$

We can also substitute the time average value for instantaneous value, and obtain the batch algorithm as follows:

$$\begin{cases} \boldsymbol{W}(k+1)=\boldsymbol{W}(k)+\eta(k)\left[\langle\boldsymbol{y}(k)\boldsymbol{f}^{\mathrm{H}}[\boldsymbol{y}(k)]\rangle - \langle\boldsymbol{f}[\boldsymbol{y}(k)]\boldsymbol{y}^{\mathrm{H}}(k)\rangle\right]\boldsymbol{W}(k), \\ \boldsymbol{y}(k)=\boldsymbol{W}(k)\boldsymbol{z}(k), \end{cases} \quad (9)$$

where the symbol $\langle\cdot\rangle$ denotes the time average value of the variable.

From the deduction, we know that the function $f_i(y_i)$ is only related to the distributed function $q(\boldsymbol{y})$. The optimal selection of the function $q(\boldsymbol{y})$ is the real probability density function of the source signal, which cannot be obtained. Thus we must select a function approximating to the real probability density function. From the whole deduction process, we find that we just use the function $f_i(y_i)$ ($i = 1,2,\ldots,m$). Therefore, the selection of $f_i(y_i)$ is very important.

### 3.4 Selection of function $f_i(y_i)$

The selection of the function $f_i(y_i)$ relates to the stability of the separated algorithm. The stability of the iterative algorithm in Eqs. (8) and (9) has been analyzed in many papers [5–7]. Here we just give the conclusion. The stable condition of the algorithm is as follows:

$$\chi_i+\chi_j>0, \quad (10)$$

where $\chi_i=E\{f_i'(y_i)\}-E\{f_i(y_i)y_i\}$, and $f_i'(y_i)$ is the derivative function of $f_i(y_i)$. If the equation $\chi_i>0$ is true for any $y_i$, the algorithm is stable. Because the communication signal is sub-Gaussian signal where the kurtosis is negative, i.e., $E(|s_i|^4) - 3E(|s_i|^2)^2 < 0$, we can select the function as $f_i(y_i) = f(y_i) = |y_i|^2 y_i + \alpha y_i$ ($\alpha \geqslant 0$). It can be proven that this function satisfies the stable condition of Eq. (10).

The blind separation for complex communication can be concluded as follows:

**Step 1** Pre-whitening the observed signal $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{x}$, where whitening matrix $\boldsymbol{Q}$ can be obtained according to Eq. (4).

**Step 2** Using Eq. (9) to separate the whitening data $\boldsymbol{z}$.

## 4 Simulation results and conclusions

We define the following performance index $P$ to evaluate the separating effect of the separate algorithm [5].

$$P=\frac{1}{m}\sum_i\left\{\left(\sum_j\frac{|\boldsymbol{G}(i,j)|}{\max_j(|\boldsymbol{G}(i,j)|)}-1\right)+\left(\sum_j\frac{|\boldsymbol{G}(j,i)|}{\max_j(|\boldsymbol{G}(j,i)|)}-1\right)\right\},$$

where $\boldsymbol{G}(i, j)$ is the $(i, j)$ element of the global system matrix $\boldsymbol{G} = \boldsymbol{WQH}$. Ideally, there is only a nonzero element in each column and each row of the matrix $\boldsymbol{G}$, which is called quasi-identity. If the matrix $\boldsymbol{G}$ is quasi-identity, the separating system can effectively separate every source signal. The characteristic of the separated signals is just the same as the source signals up to the order and the scale of ambiguity. The performance index $P$ evaluates the comparability of the matrix $\boldsymbol{G}$ and quasi-identity. Ideally, $P = 0$.

The algorithm researched in the paper is designed for communication signals, thus we use the mixing model of uniform circular array which has six elements [8] when we simulate the algorithm. At the same time we process the analytic signals. The radius $R$ of the circular array is half of the minimum wave length.

Assume there are three source signals, 2ASK signal (carrier frequency is 610 kHz), 4ASK signal (carrier frequency is 690 kHz), and 8PSK signal (carrier frequency is 770 kHz) respectively. It needs to explain that the carrier frequency of the signals is decreased because of the limitation of computer computation ability. But it does not affect the evaluation of the algorithm. On the other hand, we ignore the noise during the deduction of the algorithm. To validate the efficiency of the algorithm in noisy circumstance, we add the white Gaussian noise to the observed source during simulation. The power of the noise is 0.01, and that of the source signal is 1 (i.e., signal to noise ratio (SNR) is 20 dB). The rate of data is 70 kBd/s, and the sampling rate is 560 kHz. The angles of incidence of three sources are 40°, 60°, and 80° respectively.

Figure 1 shows the $P$ variation of the algorithm in the paper. The learning rate $\eta(k) = 0.08$. It can be seen from Fig. 1 that the $P$ is about 0.025 when the algorithm starts converging, which is very close to zero. The global matrix $\boldsymbol{G}$ is:

$$\boldsymbol{G} = \begin{bmatrix} 0.9830 & 0.0025 & 0.0050 \\ 0.0164 & 0.0077 & 0.9933 \\ 0.0194 & 0.9847 & 0.0101 \end{bmatrix}.$$

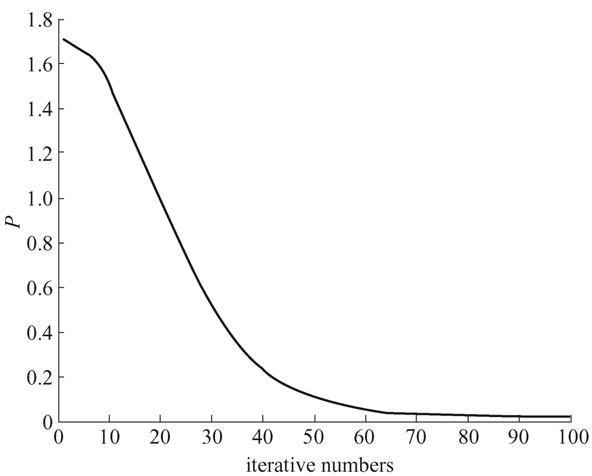


**Fig. 1** System performance index

From the above equation, we can see that $\boldsymbol{G}$ is very close to quasi-identity. Therefore, the blind separation algorithm of complex signals deduced in the paper can separate the source signals from the observed signals. The waveforms of source signals, separated signals and observed signals are shown in Figs. 2–4 respectively. The algorithm can preferably resume the waveform of the source signal and the algorithm is stable and credible.

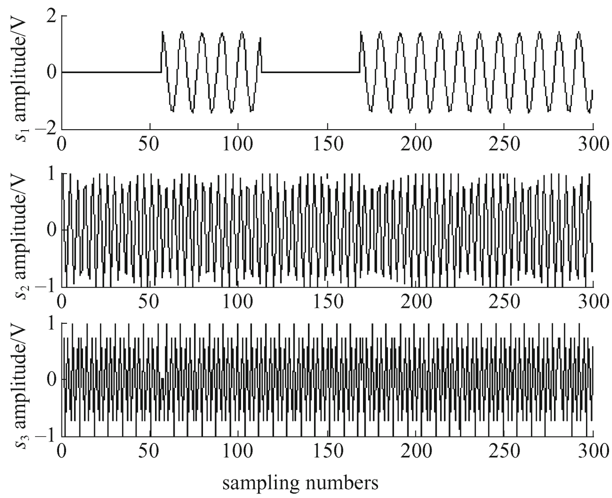


**Fig. 2** Waveform of source signals

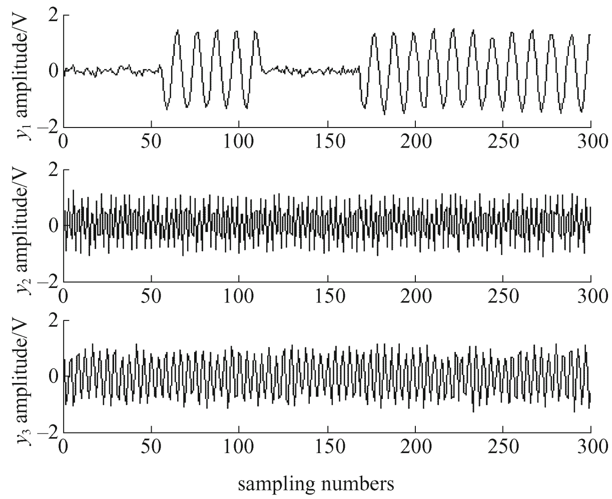


**Fig. 3** Waveform of separated signals

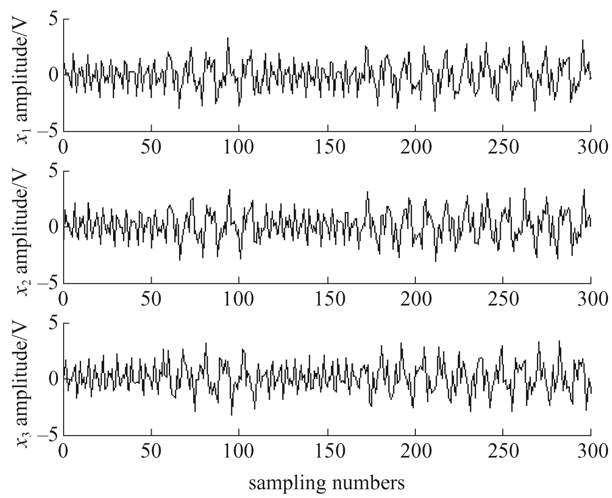


**Fig. 4** Waveform of first three mixed signals

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (Grant No. 60672038).

## References

1. Xu Shimin, Chen Pengju. Separation method of communication signals with overlap frequency. Aerospace Electronic Warfare, 2004, (5): 53–55 (in Chinese)
2. Xu Shimin. Research on the application of blind source separation in the communication reconnaissance system of wide IF band. Electronic Warfare Technology, 2005, 20(4): 12–14,18 (in Chinese)
3. Fu Weihong, Yang Xiaoniu, Liu Nai'an, et al. Communication signal blind reconnaissance technology based on probability density estimation blind sources separation. Journal of Huazhong University of Science and Technology (Nature Science), 2006, 34(10): 24–27 (in Chinese)
4. Hyvarinen A, Karhunen J, Oja E. Independent Component Analysis. New York: John Wiley & Sons, 2001, 209–212
5. Cichocki A, Amari S. Adaptive blind signal and image processing. New York: John Wiley& Sons, 2002
6. Cardoso J F, Laheld B H. Equivariant adaptive source separation. IEEE Transactions on Signal Processing, 1996, 44(12): 3017–3030
7. Amari S, Chen T P, Cichocki A. Stability analysis of learning algorithms for blind source separation. Neural Networks, 1997, 10(9): 1345–1351
8. Yang Xiaoniu, Lou Caiyi, Xu Jianliang. Principle and Application of Software Radio. Beijing: Publishing House of Electronics Industry, 2001