Available online at www.sciencedirect.com
ScienceDirect
Progress in Natural Science 18 (2008) 751–755
Progress in Natural Science
www.elsevier.com/locate/pnsc

New method for signal encryption using blind source separation based on subband decomposition

Zuyuan Yang *, Guoxu Zhou, Zongze Wu, Jinlong Zhang

School of Electrics and Information Engineering, South China University of Technology, Guangzhou 510641, China

Received 20 November 2007; received in revised form 22 January 2008; accepted 24 January 2008

Abstract

A novel cryptosystem based on subband decomposition independent component analysis (SDICA) is proposed in this work, where no assumption of independence for the ciphers and the plaintexts is required. In the proposed cryptosystem, the encryption is asynchronous, i.e. the plaintexts are mixed mutually firstly and then mixed with the ciphers. In addition, the decryption is asynchronous, such that the decryption accuracy of the plaintexts can be enhanced. Some special information about the original mixing matrix is used for solving the indeterminacy of the permutation and scale of columns of the recovered mixing matrix in SDICA, instead of the characteristics of the plaintexts. Simulations are given to illustrate security and availability of our cryptosystem.

Keywords: SDICA; Signal encryption; BSS

1. Introduction

With the fast development of network technology, the security of information has attracted more and more attention, especially for speech communication and image transmission. A variety of signal encryption techniques have been developed, such as Triple-DES and RSA [1]. The analogue encryption is another encryption technique used in speech communication, both in time domain and the other in transform domain [2]. In addition, some new methods have also been developed [3,4]. Recently, based on the fact that the decrypted signal can be an efficient estimation of the original speech, a technique using blind source separation (BSS) has been applied to speech encryption [5]. In that cryptosystem, the difficulty of solving the underdetermined BSS problem is utilized sufficiently to ensure the security, and the decryption method is also very convenient via the independent component analysis (ICA) [6]. Therefore it is a significant technique and may be applied to practical speech cryptosystem in future. However, there still exist several shortcomings when it is applied in the real world. Firstly, speech signals are often sparse, and the recoverability for the underdetermined mixing model of sparse signal is exploited widely in Ref. [7–10]. In Ref. [11], the authors proposed a mathematical theory to analyze robustness of the overcomplete representations in noise environment. They found that, in the overcomplete case, the separation quality would be improved if the sparsity of the sources became more strict. That is to say, the security of the cryptosystem may be destroyed by the sparsity characteristics. Secondly, when the ICA method is applied for decryption in Ref. [5], the plaintext signals and the ciphers must satisfy the assumptions that they are mutually independent and there exists one Gaussian signal at most. However, the statistical independence is a very strong assumption, and there are a lot of real problems in which the assumption is violated. For example, the brain source signals are generally not completely inde-

* Corresponding author. Tel./fax: +86 20 87114709.
E-mail address: yangzuyuan@yahoo.com.cn (Z. Yang).

1002-0071/$ - see front matter © 2008 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.
doi:10.1016/j.pnsc.2008.01.020

pendent except for some high-frequency subbands, according to the observed electroencephalograph (EEG) data [12]. For these signals, they cannot be decrypted efficiently using Lin's cryptosystem.

Therefore, how to ensure the security and availability of the cryptosystem based on BSS for sparse or dependent signal is an open question. As mentioned above, the security of the encryption scheme based on BSS/ICA relies mainly on the underdetermined mixing model. Instead, the determined mixing model with strong noises generated by Lorenz chaotic system will be used in the proposed cryptosystem. The plaintexts are mutually mixed firstly and then mixed with the ciphers, and the corresponding security is ensured by this asynchronous mixing method. Also, the decryption accuracy is improved. For the dependent sources, SDICA algorithm is used for decryption, where the mixing matrix is extracted firstly and then the whole source (not only the subband signals) is reconstructed. Note that there are permutation and scale ambiguities for columns of the extracted mixing matrix, and it will be reconstructed according to the information of the original mixing matrix.

2. Mixing model and SDICA

2.1. Linear mixing model with additive noise

The typical linear mixing model for BSS/ICA with m sources and n sensors is

$$H(t) = AS(t), \tag{1}$$

where $H(t) = [h_1(t), \cdots, h_n(t)]^T$ denotes the observation, $A \in R^{n\times m}$ denotes the mixing matrix, and $S(t) = [s_1(t), \cdots, s_m(t)]^T$ denotes the source. Considering the additive noises, Eq. (1) can be rewritten as

$$H(t) = AS(t) + V(t), \tag{2}$$

where $V(t) = [v_1(t), \cdots, v_n(t)]^T$ denotes the additive noise. This study is on the assumption that $m = n$ and the additive noises are strong. Note that if $m > n$ and neglects the noises, Eq. (2) is just the underdetermined mixing model used in Ref. [5].

2.2. Subband decomposition ICA

Since many real signals are correlated or dependent but have independent sub-components, the assumption that the sources are mutually independent will be relaxed as follows:

(1) All of the sources may be dependent, but they can be represented as the sum of several sub-components as

$$s_i(t) = s_{i1}(t) + s_{i2}(t) + \cdots + s_{iL}(t), \tag{3}$$

where $s_{ik}(t)$, $k = 1, \cdots, L$ are the narrow-band sub-components.

(2) There exist such sub-components which are statistically independent. Based on the assumptions mentioned above, we can design the linear time-invariant filter T_k which can extract the independent sub-components according to Eq. (3), i.e.

$$S_k(t) = T_k[S(t)] = [s_{1k}(t), \cdots, s_{mk}(t)]^T, \tag{4}$$

Note that by applying T_k to Eq. (1), the mixing sub-components can be extracted as

$$H_k(t) = T_k[H(t)] = A \cdot T_k[S(t)] = AS_k(t), \tag{5}$$

where $H_k(t) = [h_{1k}(t), \cdots, h_{mk}(t)]^T$.

We can see that the mixing matrix corresponding to the sources and their sub-components keeps the same. It can be calculated by applying ICA to the mixing signals $H_k(t)$ of the sub-components, and then the sources can be reconstructed by the same recovered mixing matrix.

3. Proposed cryptosystem

Lorenz circuit will be introduced firstly before the proposed cryptosystem is exploited, for that the corresponding chaotic system is bounded, determinate and sensitive to the initial value. The state equations are [13]

$$\begin{cases} \dot{x} = \sigma_1(y - x) \\ \dot{y} = \sigma_2 x - xz - y \\ \dot{z} = xy - \sigma_3 z \end{cases} \tag{6}$$

where $\sigma_1 = 10$, $\sigma_2 = 28$, $\sigma_3 = 8/3$. According to (6), two uncontrolled trajectories generated using fourth-order Runge-Kutta method with the same options except for the initial values are shown in Fig. 1.

3.1. Encryption

Considering m dependent sources $S(t)$ with the same number T of the samples, a full-rank matrix $A \in R^{m\times m}$ and the ciphers $V(t) \in R^{m\times T}$, then the plaintexts $S(t)$ can be encrypted into the ciphertexts $H(t)$ as follows:

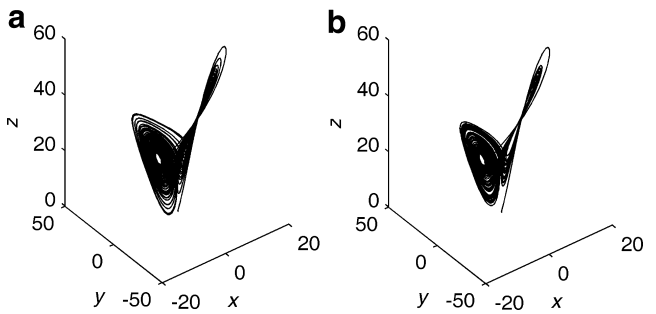

Fig. 1. Two uncontrolled trajectories generated from Lorenz circuit. (a) The initial value is [0.5863 0.4705 0.9392]; (b) the initial value is [0.5864 0.4706 0.9393].

$$\begin{cases} \tilde{Z}(t) = A(S(t) + \eta C(t)) \\ H(t) = \tilde{Z}(t) + \mu(\Phi(V(t)) + V(t)) \end{cases} \quad (7)$$

Here, $V(t)$ is usually generated pseudo-randomly by a secret seed I_0; $C(t) = [c_1(t), \cdots, c_m(t)]^T$ is the auxiliary signal such that all of the entries of $S(t) + \eta C(t)$ are nonzero (η is the scaling coefficient), and it is often constructed using the same function for simplicity (e.g. $c_i(t) = \sin(t)$); $\mu \neq 0$ is the scaled coefficient such that the plaintexts can be well masked; $\Phi(V(t)) = [\phi_1(t), \cdots, \phi_m(t)]^T$ is constructed using chaotic series generated by (6), and the corresponding operation flow is given below: given $V \in R^{m\times T}$, let $\bar{v} = vec(V) \in R^{1\times(m\times T)}$, and $\hat{V} = mat(\bar{v}) \in R^{3\times q}$, where *vec* means converting a matrix into a vector, i.e. $\bar{v}(1, i + (t-1)\times m) = v_i(t)$, $i = 1, \cdots, m$, $t = 1, \cdots, T$, and *mat* is just the inverse operation, $q = (m \times T + p)/3$, p is the smallest nonnegative integer such that $\mathrm{mod}(m \times T + p, 3) = 0$. Note that if $p \neq 0$, then there will be p zeros inserted into the end of $\bar{v}$ before calculating $\hat{V}$. Then, different columns of $\hat{V}$ are used as different initial values of Lorenz system in (6). There will be q different chaotic series with three dimensions generated, where the solving options are assumed to be the same except for the initial value. Extracting one component (three dimensions) from each series at the same time instant t_k, such that a new matrix $\hat{U} \in R^{3\times q}$ is constructed. Let $\bar{u} = vec(\hat{U}) \in R^{1\times(3\times q)}$, and $U = mat(\bar{u}) \in R^{m\times T}$, then $\Phi(V) = U$ is generated. Note that if $p \neq 0$, then the last p entries of $\bar{u}$ will be discarded before calculating U. In addition, the sizes of $\Phi(V)$ and V are the same, such that the element $v_i(t)$ is mapped into $\phi_i(t)$ one by one.

The model (7) represents the linear mixing method with strong noise which is intractable to solve. Because $\Phi(V(t))$ is very sensitive to $V(t)$, the encryption of the plaintexts relies mainly on the cipher, such that the security of the sources can be ensured if the mixing matrix and the ciphers are selected properly. For the following decryption, the secret seed I_0 can be transmitted to the receivers through an absolutely secure channel. The characteristics of the mixing matrix, the parameters t_k, μ, η and the constructed function of $C(t)$ can be inserted into the head data of the encrypted signal in a definite format for transmission.

3.2. Decryption

Combining with the m received encrypted signals $H(t)$, the secret seed I_0 which can regenerate the ciphers $V(t)$, the parameters t_k, μ, η, and the constructed function of $C(t)$, the plaintexts (i.e. the source signals) can be decrypted as

$$\begin{cases} \tilde{Z}(t) = H(t) - \mu(\Phi(V(t)) + V(t)) \\ \tilde{S}(t) = W\tilde{Z}(t) - \eta C(t) \end{cases} \quad (8)$$

where $\tilde{S}(t) = [\tilde{s}_1(t), \cdots, \tilde{s}_m(t)]^T$ denote the estimations of the sources, i.e. the decrypted signals; $W \in R^{m\times m}$ is the inverse of the mixing matrix A, and it is calculated by SDICA to the mixing signals $\tilde{Z}(t)$ (supposed that the sources have independent subbands); $C(t)$ and $\Phi(V(t))$ are recalculated with the same method in encryption. The indeterminacy of the permutation and the scale of the decrypted sources or the columns of the recovered mixing matrix are decided by the prior information about the mixing matrix which is transmitted together with the encrypted signals.

3.3. Selection of ciphers and mixing matrix

There are several rules for the selection of the key signals and the mixing matrix:

(1) The key signals should be selected such that the plaintexts can be well masked both in the time domain and the frequency domain; (2) the mixing matrix is invertible and nontrivial (i.e., there exist at least two nonzero entries in each row); (3) the mixing matrix is diagonally dominant and the columns are with unit length.

Consider the first rule, if the sources cannot be well masked in the time domain, then the insecurity is obvious. When it happens in the frequency domain, then one may extract some sub-components of the sources and the mixing matrix by applying SDICA to the observations directly, thus decreasing the immunity of the cryptosystem. For the second rule, the invertibility of the mixing matrix is the basic request for the ICA algorithm [14], and we assume that it is nontrivial to enhance the security of the source signals. For the third rule, it is mainly utilized for recovering the exact mixing matrix, such that solving the indeterminacy of the permutation and scale of the decrypted sources. Note that other schemes can also be used, such as the anti-diagonal dominant property.

3.4. Analysis of security

Three ordinary attacks to the cryptosystem are considered in this study: (1) ciphertext-only attack; (2) known-plaintext attack; (3) chosen-plaintext attack.

Consider the rules for the selection of the key signals and the mixing matrix as mentioned above, the proposed cryptosystem is secure under the first attack obviously. Under the rest attacks, it is also secure for that the mixing matrix is dynamic for different encryption processes, and the auxiliary signals $C(t)$ can make sure that all of the entries of $S(t) + \eta C(t)$ are nonzero for any $S(t)$. As a result, the plaintexts and the ciphers are mixed mutually sufficiently. Furthermore, because the values of the chaotic series at time t_k are quite sensitive to the initial values (Table 1), and $\Phi(V(t))$ is sensitive to $V(t)$, it is hard to attack the proposed cryptosystem using numerical methods (e.g., similar cipher attack). Note that the subband independence of the sources may not be affected by the auxiliary signals $C(t)$ if the constructed function of $C(t)$ is selected properly. The less sparse the mixing matrix is, the more secure the cryptosystem is, because the sources can be mixed more sufficiently.

Table 1
Ten groups of values of Lorenz chaotic series at time $t_0 = 0$ and $t_k = 1500$, respectively

Group	t_0			t_k		
	$x(0)$	$y(0)$	$z(0)$	$x(k)$	$y(k)$	$z(k)$
No. 1	0.5855	0.4697	0.9384	0.4862	−1.3728	22.4608
No. 2	0.5856	0.4698	0.9385	−9.2051	2.5195	38.2909
No. 3	0.5857	0.4699	0.9386	13.2653	11.0948	35.4642
No. 4	0.5858	0.4700	0.9387	7.6720	2.6090	31.7520
No. 5	0.5859	0.4701	0.9388	4.3021	−4.8085	32.7945
No. 6	0.5860	0.4702	0.9389	−2.9506	−4.7708	19.5858
No. 7	0.5861	0.4703	0.9390	10.5195	16.9948	19.5712
No. 8	0.5862	0.4704	0.9391	4.7766	2.8731	25.6602
No. 9	0.5863	0.4705	0.9392	−9.2637	−1.2615	35.7402
No. 10	0.5864	0.4706	0.9393	−8.7973	−2.4722	33.9384

4. Simulation

Four dependent speech signals with 6000 samples were utilized to validate the cryptosystem in this study (Fig. 2), and the corresponding correlation matrix is

$$CORR = \begin{bmatrix} 1.0000 & 0.2789 & 0.2739 & 0.3602 \\ 0.2789 & 1.0000 & 0.2776 & 0.2426 \\ 0.2739 & 0.2776 & 1.0000 & 0.3020 \\ 0.3602 & 0.2426 & 0.3020 & 1.0000 \end{bmatrix} \tag{9}$$

At the stage of encryption, the key signals (ciphers) are generated between −1 and 1 (Fig. 3), the parameters are $c = 1$, $\mu = 0.1$, $t_k = 1500$, $\eta = 0.1$, $c_i(t) = \sin(t/100)$, and the mixing matrix is

$$A = \begin{bmatrix} 0.9473 & 0.1185 & 0.2589 & -0.1612 \\ 0.2895 & 0.9011 & -0.1936 & 0.1366 \\ -0.0813 & -0.1316 & 0.9340 & 0.3764 \\ -0.1107 & 0.3942 & -0.1522 & 0.9020 \end{bmatrix} \tag{10}$$

The following index signal-to-noise ratio (SNR) is used to quantify the mask of the sources

$$SNR_i = 10 \log \frac{E[s_i(t)]^2}{E[(h_i(t) - s_i(t))^2]} \tag{11}$$

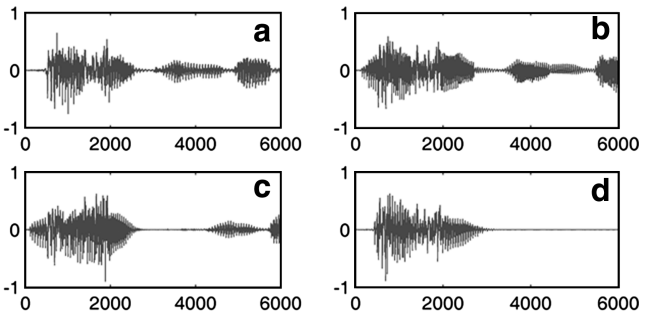

Fig. 2. Four source signals. (a) $s_1(t)$; (b) $s_2(t)$; (c) $s_3(t)$; (d) $s_4(t)$.

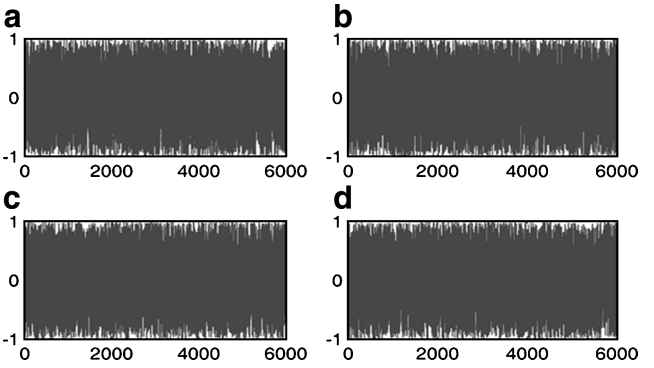

Fig. 3. Four key signals. (a) $v_1(t)$; (b) $v_2(t)$; (c) $v_3(t)$; (d) $v_4(t)$.

It can also be used to quantify the decryption after changing $h_i(t)$ into the last output $\tilde{s}_i(t)$. Obviously, a better cryptosystem asks for a lower SNR for encryption and a higher SNR for decryption.

Fig. 4 shows the encrypted signals $H(t)$. At the stage of decryption, the direct ICA algorithm and the SDICA algorithm are used separately. The encryption and decryption indices using these two methods are shown in Table 2. The corresponding recovered mixing matrices $\hat{A}$, $\tilde{A}$ and the decrypted source signals are as follows:

$$\hat{A} = \begin{bmatrix} 0.4372 & 0.7437 & 0.3188 & 0.5759 \\ 0.4463 & 0.3171 & 0.6683 & -0.6931 \\ 0.5399 & 0.4233 & -0.6456 & -0.0263 \\ 0.5641 & -0.3942 & 0.1881 & -0.4327 \end{bmatrix} \tag{12}$$

$$\tilde{A} = \begin{bmatrix} 0.9469 & 0.1387 & 0.2042 & -0.2009 \\ 0.2934 & 0.8954 & -0.2537 & 0.0039 \\ -0.0747 & -0.1143 & 0.9305 & 0.4290 \\ -0.1078 & 0.4075 & -0.1677 & 0.8807 \end{bmatrix} \tag{13}$$

Eq. (9) shows that the sources (i.e., the plaintexts) are not mutually independent, and it may fail using the cryptosystem in Ref. [5]. Table 2 and Fig. 4 show that the sources are well masked in the proposed cryptosystem. They are decrypted using SDICA with a high accuracy, but the decryption using direct ICA seems to fail according to the very low SNR. It can also be verified by the recovered mixing matrices (see Eqs. (12) and (13)). Figs. 5 and 6 show the decrypted signals using SDICA and ICA, respectively.

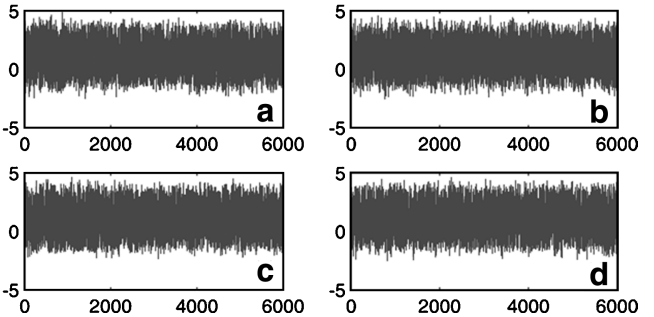

Fig. 4. Four encrypted signals. (a) $h_1(t)$; (b) $h_2(t)$; (c) $h_3(t)$; (d) $h_4(t)$.

Table 2
SNRs(db) of encryption and decryption by ICA, SDICA

	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$
$h_1(t)$	−54.1178	−51.6730	−52.2098	−54.9338
$h_2(t)$	−53.9602	−51.6402	−52.0532	−54.8175
$h_3(t)$	−54.0392	−51.6566	−52.2761	−54.9479
$h_4(t)$	−54.0432	−51.6778	−52.1862	−54.9692
ICA	−2.4300	0.9889	−3.2773	−0.6219
SDICA	53.2003	39.7474	71.2014	52.1861

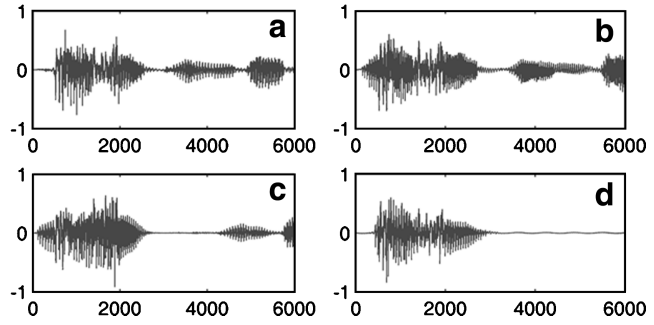

Fig. 5. Four decrypted signals using SDICA. (a) $y_1(t)$; (b) $y_2(t)$; (c) $y_3(t)$; (d) $y_4(t)$.

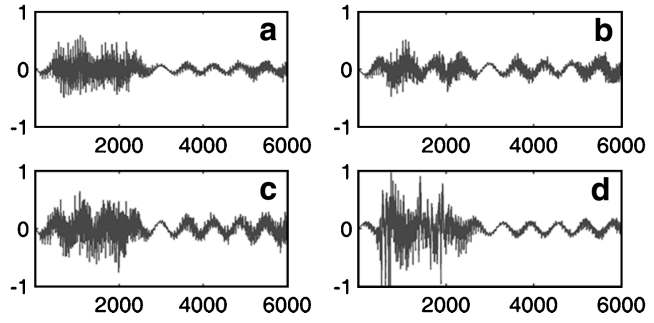

Fig. 6. Four decrypted signals using ICA directly. (a) $y_1(t)$; (b) $y_2(t)$; (c) $y_3(t)$; (d) $y_4(t)$.

5. Conclusions

In this work, the scheme based on SDICA is used for the decryption of the dependent plaintexts, and the mixing model with strong noises is used for encryption where the mixing method of the plaintexts and the key signals is asynchronous. Just like the normal ICA algorithm, in the SDICA algorithm, there is still the indeterminacy of the permutation and scale for columns of the recovered mixing matrix. In order to obtain the exact mixing matrix, some information about the original mixing matrix is used, such as the property of the diagonal dominance.

There exist several methods for SDICA, such as the scheme based on the global mixing–unmixing matrices analysis [12] and the wavelet packet approach [15], and so on. If there is no independent subband in the original sources, the auxiliary signals $C(t)$ can be constructed properly into the independent subband (i.e. for $\forall i,j$, $c_i(t)$, $c_j(t)$ are high-frequency subbands which are mutually independent), as long as the frequency of the sources are limited. Then, the ciphers are required to mask auxiliary signals additionally in the frequency domain. Furthermore, in Ref. [16], a novel method is proposed and proved based on a modified Stone's conjecture. This is a new BSS method using second order statistics. Original Stone's conjecture is evidenced to be false and is modified mathematically. It establishes a new reliable basis of BSS. Thanks to this new method, the auxiliary signals $C(t)$ can be constructed into that with different temporary structure, such that the plaintexts can even be decrypted exactly.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant Nos. U0635001, 60674033, 60774094), and the Natural Science Fund of Guangdong Province, China (Grant No. 05006508).

References

[1] Menezes A, Van Oorschot P, Vanstone S. Handbook of applied cryptography. Boca Raton, FL: CRC; 1996.
[2] Ma FL, Cheng J, Wang YM. Wavelet transform-based analogue speech scrambling scheme. Electron Letts 1996;32:719–21.
[3] Manjunath G, Anand GV. Speech encryption using circulant transformations. IEEE Int Conf Multimedia Expo 2002;1:553–6.
[4] Borujeni SE. Speech encryption based on fast Fourier transform permutation. IEEE Int Conf Ele Cir Syst 2000;1:290–3.
[5] Lin QH, Yin FL, Mei TM, et al. A blind source separation based method for speech encryption. IEEE Trans Cir Syst I 2006;53:1320–8.
[6] Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 1999;10:626–34.
[7] He ZS, Xie SL, Fu YL. Sparse representation and blind source separation of ill-posed mixtures. Sci China (Series F-Inform Sci) 2006;49:639–52.
[8] He ZS, Xie SL, Ding SX, et al. Convolutive blind source separation in the frequency domain based on sparse representations. IEEE Trans Audio Speech Lang Proc 2007;15(5):1551–63.
[9] Li YQ, Amari S, Cichocki A, et al. Probability estimation for recoverability analysis of blind source separation based on sparse representation. IEEE Trans Inform Theory 2006;52(7):3139–52.
[10] He ZS, Xie SL, Fu YL. Sparsity analysis of signals. Progr Nat Sci 2006;16(8):879–84.
[11] He ZS, Xie SL, Zhang LQ, et al. A note on Lewicki-Sejnowski gradient for learning overcomplete representations. Neural Comput 2008;20:636–43.
[12] Tanaka A, Cichocki A. Subband decomposition independent component analysis and new performance criteria. ICASSP 2004:541–4.
[13] Lü JH, Chen GR, Cheng DZ, et al. Bridge the gap between the Lorenz system and the Chen system. Int J Bifurcation Chaos 2002;12(12):2917–26.
[14] Comon P. Independent component analysis, a new concept? Signal Process 1994;36:287–314.
[15] Kopriva I, Seršić D. Wavelet packet approach to blind separation of statistically dependent sources. Neurocomputing 2007;71:1642–55.
[16] Xie SL, He ZS, Fu YL. A note on Stone's conjecture of blind signal separation. Neural Comput 2005;17:321–30.