Independent Component Analysis by Evolutionary Neural Networks

Independent Component Analysis by Evolutionary Neural Networks

Xiang-Yan Zeng*, Yen-Wei Chen**, Zensho Nakao**, Katsumi Yamashita**

Abstract

In this paper, we propose an evolutionary neural network for blind source separation (BSS). The BSS is the problem to obtain the independent components of original source signals from mixed signals. The original sources that are mutually independent and are mixed linearly by an unknown matrix are retrieved by a separating procedure based on Independent Component Analysis (ICA). The goal of ICA is to find a separating matrix so that the separated signals are as independent as possible. In neural realizations, separating matrix is represented as connection weights of networks and usually updated by learning formulae. The effectiveness of the algorithms, however, is affected by the neuron activation functions that depend on the probability distribution of the signals. In our method, the network is evolved by Genetic Algorithm (GA) that does not need activation functions and works on evolutionary mechanism. The kurtosis that is a simple and original criterion for independence is used as the fitness function of GA. After learning, the network can be used to separate additional signals different from training set but mixed by the same matrix. The applicability of the proposed method for blind source separation is demonstrated by the simulation results.

Keywords: Blind source separation, Independent Component Analysis, Genetic Algorithm, Kurtosis

1. INTRODUCTION

The problem of blind source separation (BSS) has arisen in many areas such as image restoration, speech recognition, but we can take all of these as generalized signal processing. In BSS, the goal is to extract independent components of source signals from their linear mixtures without knowing the mixing procedure, so the problem is tightly related to Independent Component Analysis (ICA). Various neural network algorithms have been proposed for solving this problem [2,5,6]. The performance of these algorithms, however, is much affected by the update formula for the separating matrix or connection weights of neural networks. The update formulae are derived from the minimization of the dependence among the output components and usually include neuron activation functions that depend on the distribution of input signals. In this paper we propose an algorithm to get the separating matrix by using genetic algorithm (GA). GA is well known for the ability to search for the optimum solution of a problem by using genetic operators and a fitness function. ICA can be taken as a problem of minimization of the dependence among signals. Our method is free from the selection of activation function and works generally on the evolutionary mechanism. Some research has been done in applying GA to the BSS [4], in which the Kullback-Leibler entropy is computed by analyzing the signals one by one and therefore needs exhausted computation. We use the absolute value of the kurtosis as the fitness function that is a simple and original criterion for independence. We shall show in the simulation experiment that this method is effective for ICA in the aspects of computational cost and accuracy.

Received on: Dec 6, 1999

*Master student, Department of EEE, Univ. of the Ryukyus
** Department of EEE, Univ. of the Ryukyus

+ Part of the paper will be presented at SPIE(San-Jose, U.S.A), Jan. 24-28,1999.

2. BLIND SEPARATION BY NEURAL NETWORKS

2.1 Data Model For the Problem

The data model used in defining BSS problem or ICA is considered as following.

Assume that there exist unknown source signals s_i, $i = 1,\ldots,N$ which are mutually independent and each one with a zero mean. It is assumed that the sources are mixed with an unknown $N \times N$ matrix $\boldsymbol{B}$:

$$\boldsymbol{x} = \boldsymbol{Bs} \quad (1)$$

where $\boldsymbol{s} = [s_1,\ldots,s_N]^T$ and $\boldsymbol{x} = [x_1,\ldots,x_N]^T$ are N-dimensional source and mixed signals.

In BSS, the task is to retrieve the source signals from $\boldsymbol{x}$ without knowing the mixing matrix $\boldsymbol{B}$. It is impossible to get the original sources $\boldsymbol{s}$; however, we can find a separating matrix $\boldsymbol{W}$:

$$\boldsymbol{z} = \boldsymbol{Wx} \quad (2)$$

so that $\boldsymbol{z} = [z_1,\ldots,z_N]^T$ is an estimate of $\boldsymbol{s}$ in the sense that each component of $\boldsymbol{s}$ may appear in any component of $\boldsymbol{z}$ with a scalar factor. The criterion of the optimal $\boldsymbol{W}$ is that the components of $\boldsymbol{z}$ are mutually independent. Therefore, the problem becomes to find the independent components of $\boldsymbol{z}$

2.2. The Kurtosis as Criterion of Independence

The basic idea of ICA is to minimize the dependence among the output components. In many cases, average mutual information (MI) evaluated by Gram-Charlier expansion or Edgeworth expansion is usually used to measure the dependence. Both of the two expansions are based on higher-order statistics and usually lead to rather complicated adaptive separation algorithm[1,2]. In this paper, a much simpler function, the sum of the absolute values of the kurtosis as shown in Eq. (3), is used as the fitness function of GA.

$$\mathrm{J}(\boldsymbol{z}) = \sum_{i=1}^{N} \left| \mathrm{E}\{z_i^4\} - 3\mathrm{E}^2\{z_i^2\} \right| \quad (3)$$

The dependence among the components $z_1,\ldots,z_N$ is minimized when $\mathrm{J}(\boldsymbol{z})$ is maximized for a separating matrix $\boldsymbol{W}$ under some constraints. This conclusion is from the following theorem[3] proven by Moreau and Macchi.

Theorem 1: Assume that the kurtosis of all the independent sources s_i have the same sign, and consider vectors $\boldsymbol{z} = \boldsymbol{Hs}$ with $s = (s_1,\ldots,s_N)^T$ the sources and $\boldsymbol{H}$ a invertible matrix. If there exists $\mathrm{E}\{\boldsymbol{zz}^{\mathrm{T}}\} = \mathbf{I}$, which means the components of $\boldsymbol{z}$ are whitened vectors, then function J satisfies:

(1) $\mathrm{J}(\boldsymbol{z})$ is invariant by scale changes in any of the elements z_i;

(2) $\mathrm{J}(\boldsymbol{z}) \le \mathrm{J}(\boldsymbol{s})$;

(3) $\mathrm{J}(\boldsymbol{z}) = \mathrm{J}(\boldsymbol{s})$ if and only if any component of $\boldsymbol{z}$ appears as a component of $\boldsymbol{s}$ with a scalar factor.

In many practical applications, the sources belong to the same kind of signals and their kurtosis have the same sign. The conclusion (3) means that of all the signals $\boldsymbol{z} = \boldsymbol{Hs}$ with independent sources $\boldsymbol{s}$, the one that maximizes J is $\boldsymbol{s}$ itself or the permutation of its components with scalar factors. So we can recover the independent components from the mixture signals $\boldsymbol{x}$ by maximizing function $\mathrm{J}(\boldsymbol{z})$ on the premise of $\mathrm{E}\{\boldsymbol{zz}^{\mathrm{T}}\} = \mathbf{I}$.

2.3. Whitening

Whitening $\boldsymbol{z}$ itself in the parallel and iterative GA is costly in computation , and we can realize it in the next way.

Before the separation, the signal vector $\boldsymbol{x}$ is prewhitened by transformation

$$\boldsymbol{v} = \boldsymbol{Vx} \quad (4)$$

such that

$$\mathrm{E}\{\boldsymbol{vv}^{\mathrm{T}}\} = \mathbf{I} \quad (5)$$

The whitening matrix $\boldsymbol{V}$ can be obtained from a neural algorithm

$$\boldsymbol{V}_{k+1} = \boldsymbol{V}_k - \mu(\boldsymbol{v}_k\boldsymbol{v}_k^{\mathrm{T}} - \boldsymbol{I})\boldsymbol{V}_k \quad (6)$$

where μ is the learning rate, and k denotes the learning iteration.

The whitened vector $\mathbf{v}$ is then transformed into output $\boldsymbol{z}$ by matrix $\boldsymbol{W}$

$$\boldsymbol{z} = \boldsymbol{Wv} \quad (7)$$

If we constrain that $\boldsymbol{W}$ is orthogonal, then

$$\mathrm{E}\{\mathrm{zz}^{\mathrm{T}}\} = \mathrm{W}\,\mathrm{E}\{\mathrm{vv}^{\mathrm{T}}\}\mathrm{W}^{\mathrm{T}} = \mathrm{WIW}^{\mathrm{T}} = \mathrm{I} \quad (8)$$

2.4. Evolutionary Neural Networks for ICA

In the neural realization of ICA, the whitened signal vector $\boldsymbol{v}$ and the separated signal $\boldsymbol{z}$ are inputs and outputs of the network with the separating matrix $\boldsymbol{W}$ as the connection weights. The GA algorithm used to find the separating matrix $\boldsymbol{W}$ is as follows:

(1) Initialize the population of the separating matrix to

random values in $[-1,1]$;

(2) In each generation, repeat (3)～(6);

(3) Calculate the fitness of each individual by kurtosis defined by Eq. (3).

(4) Get the best individuals and put them directly to the next generation;

(5) Perform numerical crossover as shown in Fig. 1 to the individuals selected by Roulette algorithm.

(6) Perform mutation to the random selected elements of matrices with random values in $[-1,1]$.

Parent matrices: $W^{p1} = \left[w_{ij}^{p1}\right]_{3\times3}$;

$$W^{p2} = \left[w_{ij}^{p2}\right]_{3\times3}$$

Child matrices: $W^{c1} = \left[w_{ij}^{c1}\right]_{3\times3}$;

$$W^{c2} = \left[w_{ij}^{c2}\right]_{3\times3}$$

where

$$w_{ij}^{c1} = (\alpha1 \times w_{ij}^{p1} + \alpha2 \times w_{ij}^{p2})/2$$

$$w_{ij}^{c2} = (\alpha1 \times w_{ij}^{p1} - \alpha2 \times w_{ij}^{p2})/2$$

$$0 < \alpha1 = random(), \alpha2 = random() < 1$$

Figure 1　Numerical crossover

3. SIMULATION

In order to check the effectiveness of the proposed method, we simulate the separation by using synthetic original signals and a mixing matrix whose elements are random values in the range $[-1,1]$. One example of the sources is

$$\begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix} = \begin{bmatrix} random() \\ 0.1\sin(400t)\cos(30t) \\ 0.01 sign[\sin(500t) + 9\cos(40t)] \end{bmatrix}$$

where $random()$ is a noise source with a uniform distribution in the range $[-1,1]$.

The simulation consists of two stages.

(1) Train the network by GA using 1000 signals sampled at $t = 0,1,\ldots,999$ to get the separating weight matrix W for which the outputs are as independent as possible.

(2) Separate 1000 test signals sampled at $t = 1000,\ldots,1999$ by the trained neural network.

The parameters used for separation is shown in Table 1.

The simulation results are shown in Fig. 2 ～ Fig. 4. The three signals s_1, s_2, s_3 in Fig. 2 represent the sources. The mixed signals x_1, x_2, x_3 are shown in Fig. 3 and the separated signals z_1, z_2, z_3 are depicted in Fig. 4. In the simulation experiment, the separated signals z_1, z_2, z_3 correspond to the source signals s_3, s_1, s_2 respectively. Though the scales of z_1, z_2, z_3 are different from those of s_3, s_1, s_2, we can see that the signals are almost clearly recovered.

Table 1　GA parameters

Number of generations	300
Population size	50
Elite number	5
Crossover rate of each individual	0.9
Mutation rate of each element except elite	0.08

4. CONCLUSION

In this paper, we propose an evolutionary neural network for ICA problem. In neural realizations of ICA, the connection weights are usually updated by formulae which are derived from the maximization or minimization of some criterions. In many cases, the learning formulae are obtained through many approximations and depend on the activation function. In our method, however, GA instead of learning formulae is used to evolve the connection weights. The goal of GA is to search for the optimal connection weights which can separate the signals as independent as possible. The separating ability of the algorithm is related to the fitness function which is used to measure the independence among signals. We use the absolute value of kurtosis as the fitness function that is a rather simple and original statistics criterion. The simulation is performed for synthetic original sources and a random mixing matrix. The applicability of the algorithm to ICA is demonstrated by experimental results.

References

[1] G. Deco, and D. Obradovic. "Independent Component Analysis: General Formulation and Linear Case", *An*

Information-theoretic Approach to Neural Computing, Springer. 1996.

[2] S.Amari, A. Cichocki, and H. H. Yang. "A New Learning Algorithm for Blind Signal Separation", *Advances in Neural Information Processing Systems , 8:757-763(1995).*

[3] E. Moreau, and O.Macchi. "New Self-adaptive Algorithms for Source Separation Based on Contrast Functions", In *Proc. IEEE Signal Proc. Workshop on Higher Order Statistics*, pp. 215-219, June 1993,

[4] M. Yoshioka, and S. Omatsu. "Siganl Separation Method Using Independent Component Analysis", In *Proc. The Fifth International Conf. On Neural Information*, pp. 891-894, Oct. 1998.

[5] C.Jutten, and J.Herault. "Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture", *Signal Processing*, vol. 24, no. 1, pp.1-10, July 1991.

[6] A.Bell, and T. Sejnowski. " An information-maximization approach to blind separation and blind deconvoution", *Neural Computation*, ol.7, pp. 1129-1159,1995.

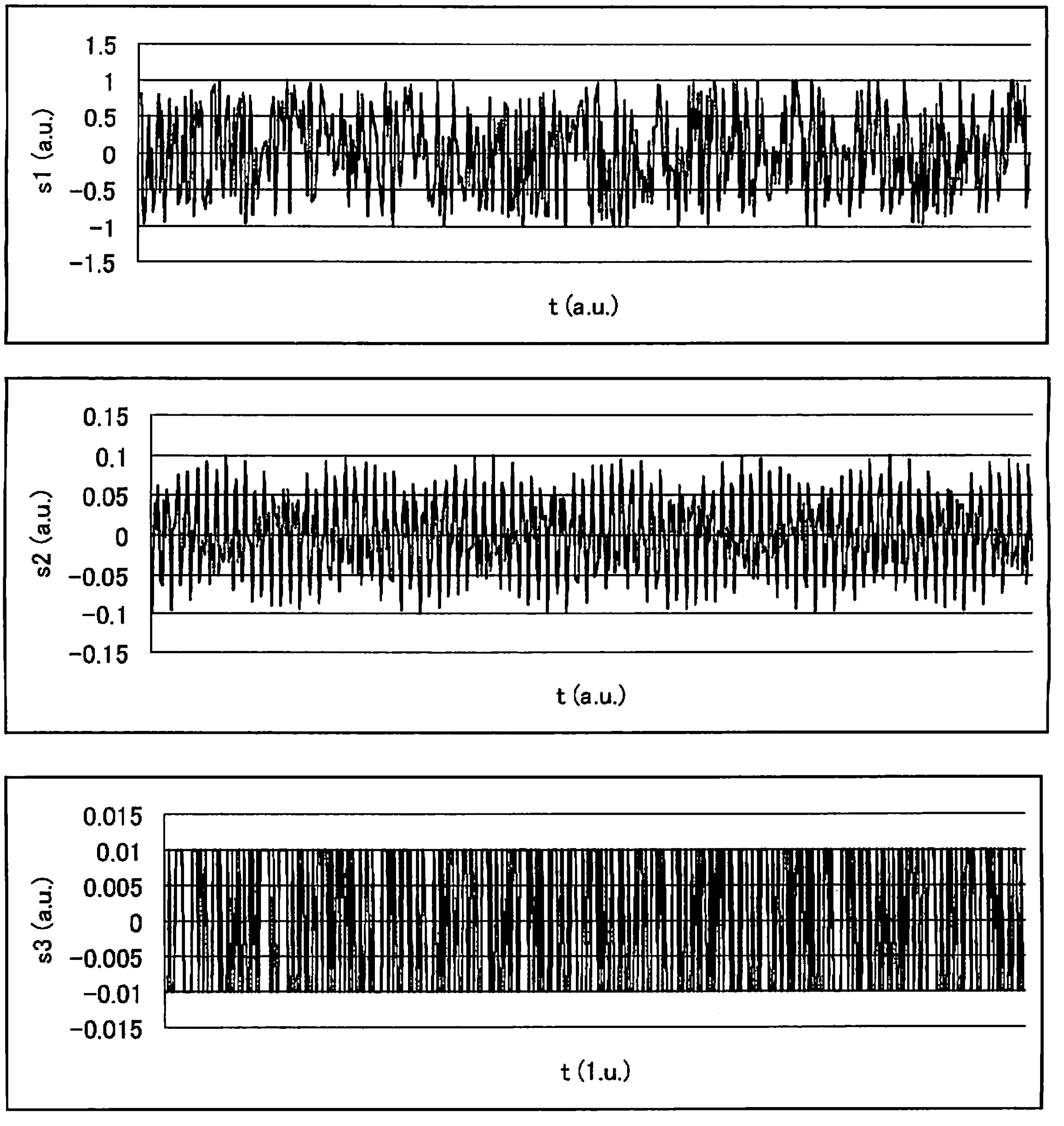

Figure 2　The original source signals

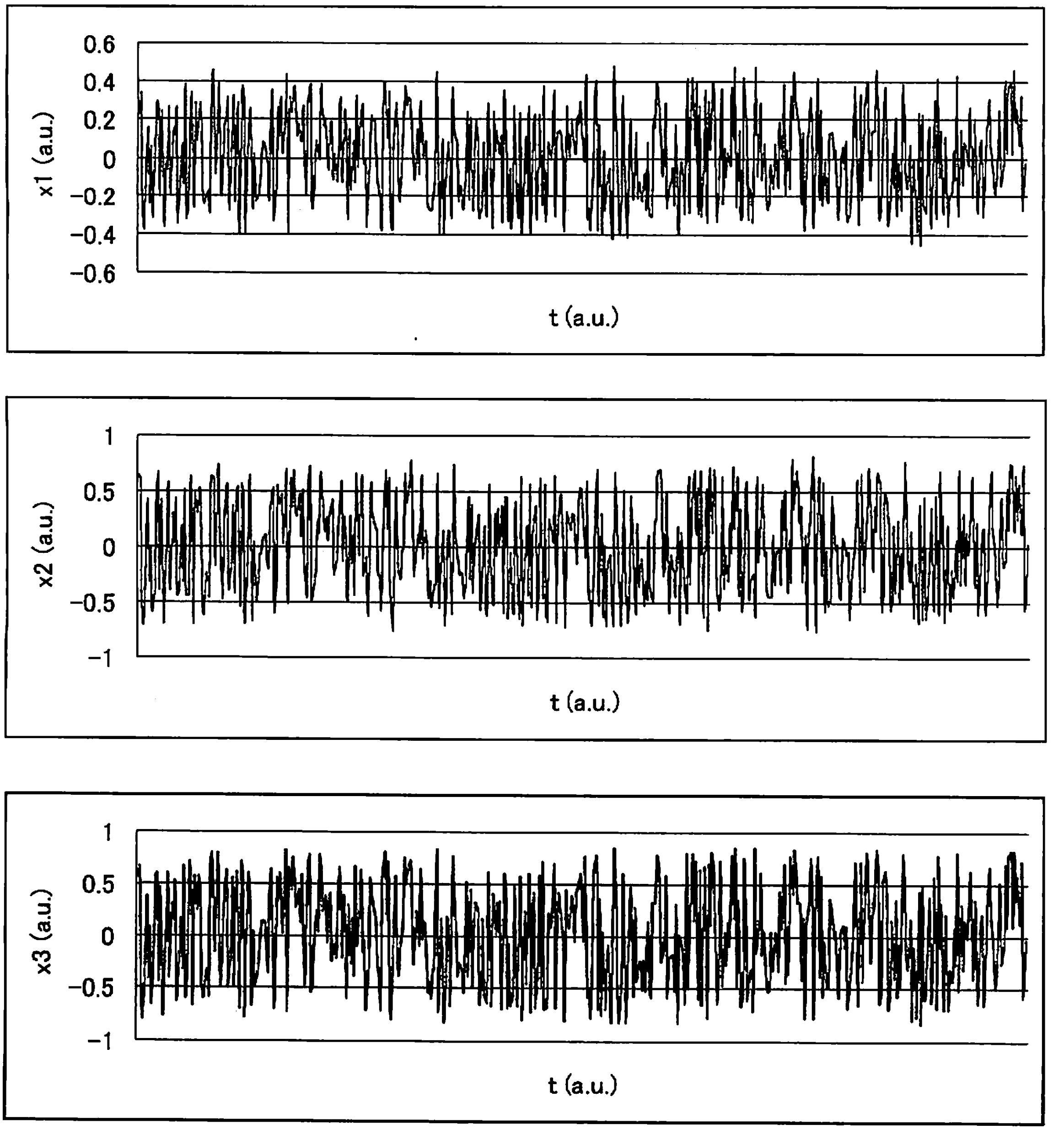

Figure 3 The signals mixed by a random matrix

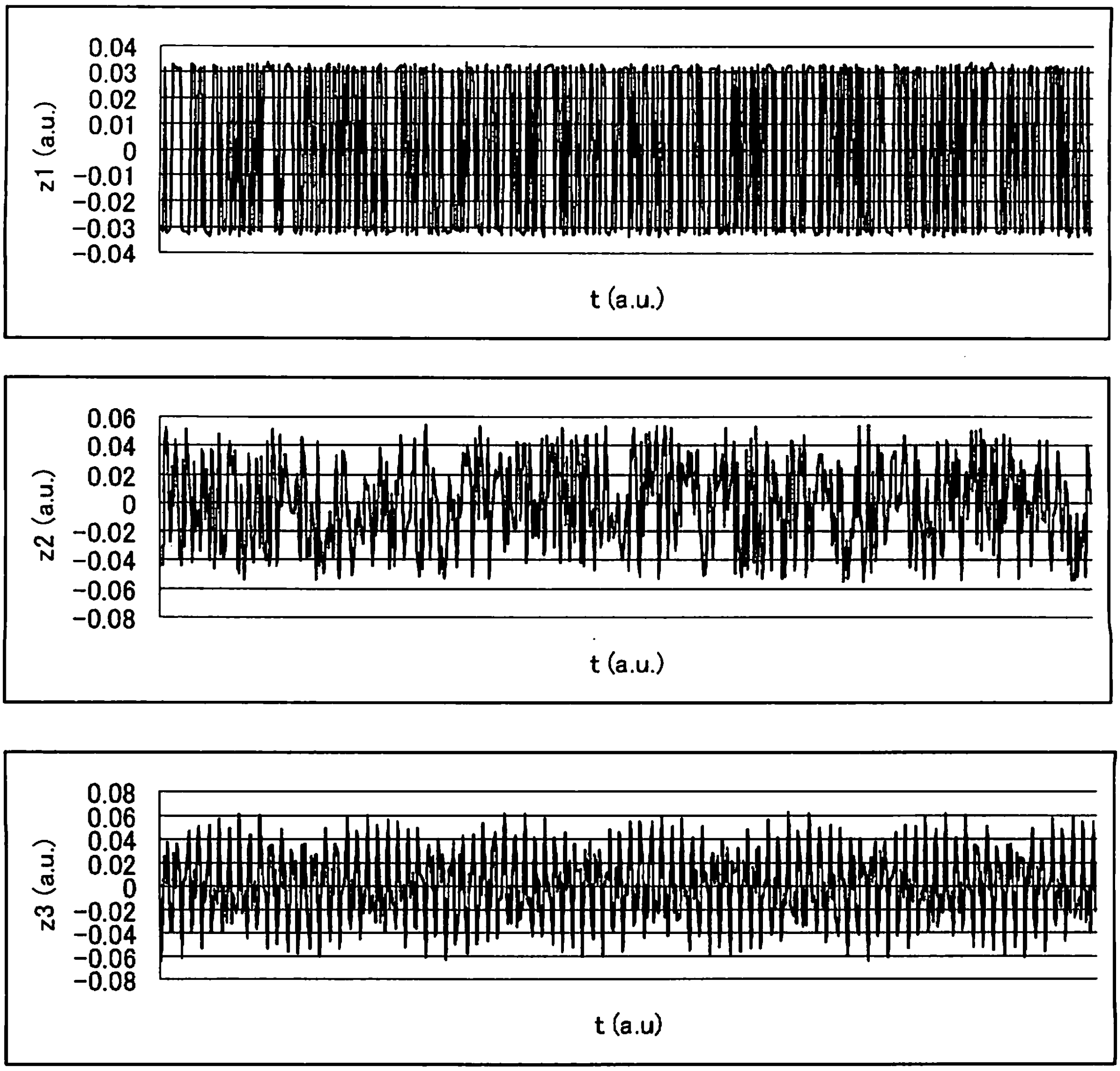

Figure 4 The separated signals