A FIXED-POINT ALGORITHM FOR BLIND SEPARATION OF TEMPORALLY CORRELATED SOURCES

Zhenwei Shi[1,2], *Dan Zhang*[2], *Changshui Zhang*[2]

1. Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics Beijing 100083, P.R. China
2. Department of Automation, Tsinghua University, Beijing 100084, P.R. China
{shizhenwei, zcs}@mail.tsinghua.edu.cn, dan-zhang05@mails.tsinghua.edu.cn

ABSTRACT

In this paper we develop a new method for blind separation of temporally correlated sources, possibly dependent signals from linear mixtures of them. The proposed algorithm is based on the mutual independency of the innovations of source signals instead of original signals. This algorithm takes into account both the temporal structure and the high-order statistics of source signals and in contrast to the most known blind separation algorithms only exploiting the second order statistics or the non-Gaussianity. In this framework, a fixed-point algorithm is introduced. The fixed-point algorithm is computationally very simple, converge fast, and does not need choose any learning step sizes. Extensive computer simulations with speech signals and images confirm the validity and high performance of the proposed algorithm.

1. INTRODUCTION

Blind source separation (BSS) or independent component analysis (ICA) has been widely applied to many areas, such as speech and image enhancement, wireless communication, data mining and biomedical signal processing. There have existed a number of learning algorithms for blind source separation [2, 5]. Generally speaking, BSS algorithms use only the high-order statistics criteria, such as FastICA [5] or use the temporal structure alone, such as SOBI [1]. The most known and efficient techniques assume that primary sources are statistically independent, however in many applications the sources are not completely independent and developed BSS or ICA algorithms may fail to separate signals of interest. But results obtained by these methods are likely to improve if one exploits all the available information on the sources, such as the high-order statistics and the temporal structure.

In this paper, we propose a fixed-point algorithm based on the mutual independency of the innovations of source signals instead of original signals. The purpose is to make use of both the temporal structure and the high-order statistics of source signals for obtaining more efficient learning algorithm. The proposed fixed-point algorithm has advantages in blind separation of temporally correlated sources, possibly dependent signals.

The work was supported by the National Science Foundation of China (60605002, 10571018) and by the Chinese Postdoctoral Science Foundation (2005038075).

2. PROBLEM FORMULATION

For our purpose, the problem of blind source separation can be formulated as follows: we observe sensor signals $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^T$ described by matrix equation:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{A}$ is an $n \times n$ full-rank unknown mixing matrix, $\mathbf{s}(t) = (s_1(t), \ldots, s_n(t))^T$ is a vector of unknown temporally correlated sources with zero mean. We assume that source signals have temporal structures and can be modelled by linear autoregressive models as:

$$s_i(t) = \sum_{\tau=1}^{M} \alpha_\tau^i s_i(t-\tau) + \delta s_i(t), \tag{2}$$

where $\delta s_i(t)$ is zero-mean, i.i.d (white) time series called innovations.

The task of BSS is to construct an $n \times n$ separating matrix $\mathbf{B}$ given just the sensor signals, such that the output vector $\mathbf{y}(t) = (y_1(t), \ldots, y_n(t))^T = \mathbf{B}\mathbf{x}(t)$ recovers the n source signals up to scaling and permutation [5].

3. A FIXED-POINT ALGORITHM FOR INDEPENDENT INNOVATION ANALYSIS

3.1. Derivation of the learning algorithm

Provided that the measured sensor signals $\mathbf{x}$ have already been followed by an $n \times n$ whitening filter $\mathbf{V}$ such that the components of $\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t)$ are unit variance and uncorrelated, the BSS problem is reduced to the search for an $n \times n$ orthonormal matrix $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_n)^T$ and the total separating matrix is given as $\mathbf{B} = \mathbf{W}\mathbf{V}$.

We suggest to use the mutual independency of the innovation signals as a criterion for finding the separating matrix. The analysis of innovation signals provides us a way to explore both the temporal structures and high-order statistics of source signals. Thus, we formulate the blind separation of temporally correlated sources in the framework of independent innovation analysis (IIA).

If the innovations are mutual independent, the covariances $E\{f(\delta y_i(t))g(\delta y_j(t))\}$ are all zero ($i \neq j$), where f and g are different odd functions [5]. In other words, the covariance matrix of $f(\delta y_i(t))$ and $g(\delta y_j(t))$:

$$\begin{aligned}\mathbf{C_{fg}} &= E\{\mathbf{f}(\delta\mathbf{y}(t))\mathbf{g}(\delta\mathbf{y}(t)^T)\} \\ &= diag(E\{f(\delta y_1(t))g(\delta y_1(t))\}, \ldots, E\{f(\delta y_n(t))g(\delta y_n(t))\})\end{aligned}$$

is a non-singular diagonal matrix.

The above diagonalization principle can be expressed in the following way:

$$\mathbf{\Lambda C_{fg}} = \mathbf{I}, \tag{3}$$

where $\mathbf{\Lambda}$ is a non-singular diagonal matrix and $\mathbf{I}$ is the identity matrix. By multiplying the above equation by separating matrix $\mathbf{W}$, we obtain a fixed-point equation about $\mathbf{W}$:

$$\mathbf{W} = \mathbf{\Lambda C_{fg} W}. \tag{4}$$

This expression suggests the following fixed-point iterative learning algorithm:

$$\mathbf{W} \leftarrow \mathbf{\Lambda C_{fg} W}. \tag{5}$$

Then the symmetric orthogonalization can be used to keep separating matrix $\mathbf{W}$ orthogonally [5]:

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{-\frac{1}{2}}\mathbf{W}. \tag{6}$$

The forms of the activation functions f and g depend on the distributions of the innovations. Typically, we can choose $f(u) = u$ and $g(u) = \tanh(au)$ $(a \geq 1)$ for super-Gaussian innovations. Thus, after the data $\mathbf{x}$ is whitened, the IIA algorithm is then as follows. At every step, first estimate the autoregressive constants α_τ^i in equation (2) for the time series given by $\mathbf{W}\tilde{\mathbf{x}}$. Then do the fixed-point iteration in (5) and the symmetric orthogonalization in (6).

A simple special case of the method is obtained when the autoregressive model has just one predicting term [4]:

$$\hat{y}_i(t) = \alpha_1^i y_i(t-1). \tag{7}$$

The lag need not be equal to 1, but this is the basic case. The parameter α_1^i in the algorithm can then be estimated simply by a least-squares method as [4]:

$$\hat{\alpha}_1^i = \mathbf{w}_i^T E\{\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}(t-1)^T\}\mathbf{w}_i. \tag{8}$$

3.2. Outline of the algorithm

The proposed algorithm involves the following steps.

1. Remove the mean from the data and whiten it to give $\tilde{\mathbf{x}}$. Choose an initial separating matrix $\mathbf{W}$. Typically, let $\mathbf{\Lambda} = \mathbf{I}$.
2. Compute estimates of the source signals as $\mathbf{y}(t) = \mathbf{W}\tilde{\mathbf{x}}(t)$.
3. Compute estimates of the autoregressive coefficients $\hat{\alpha}_\tau^i$, for example, by a classical least-squares method.
4. Compute estimates of the innovations as $\delta y_i(t) = y_i(t) - \sum_{\tau=1}^M \hat{\alpha}_\tau^i y_i(t-\tau)$.
5. Choose the activation functions f and g. For example, take $f(u) = u$ and take $g(u) = \tanh(u)$ when innovations are super-Gaussian.
6. Do a fixed-point iterative step:
$$\mathbf{W} \leftarrow \mathbf{C_{fg} W}. \tag{9}$$
7. Orthogonalize $\mathbf{W}$ by
$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{-\frac{1}{2}}\mathbf{W}. \tag{10}$$
8. If not converged, go back to step 2.

Note that convergence means that the old and new values of $\mathbf{w}_i$ point in the same direction, i.e. their dot-product is (almost) equal to 1 $(i = 1, 2, \ldots, n)$.

4. SIMULATIONS

In order to verify the effectiveness of our algorithm, we consider the separation of the following three sets of source signals:

Experiment 1: one voice signal (dyrcj_voice.wav) and one music signal (dyrcj_piccolo.wav) (taken from [7]) (available at http://www.au.tsinghua.edu.cn/szll/bodao /zhangchangshui /bigeye/member/zyghtm/Experiments.htm).

Experiment 2: ten natural speech signals of ten different persons pronouncing the same sentence (10halo.mat: available at http://www.bsp.brain.riken.jp/index.php). The source signals are highly correlated.

Experiment 3: four nature images (available at http://www.bsp.brain.riken.jp/index.php). The correlations between the four source signals are not close to zero.

For all of the experiments the standard ICA algorithms based on high-order statitics such as FastICA [5] does not work well. Methods based on the time-dependency information alone like SOBI [1] performs poorly in Experiment 2 and fails with the data in Experiment 3. Thus, for the goal of comparison, we test three batch algorithms including complexity pursuit using the symmetric orthogonalization (CPSYM)

Table 1. The average performance indexes **PI** and CPU-time in seconds for different algorithms in the Experiment 1.

Algorithms	**CPDEF**	**FPIIA**
PI	0.0375	0.0306
CPU − time	29.41	3.47

[4], complexity pursuit using the deflation scheme (CPDEF) [4] and the fixed-point algorithm for independent innovation analysis in this paper (FPIIA).

Remark 1: *Complexity pursuit [4] combines both of non-Gaussianity and time-correlations, which is an extension of projection pursuit [3] to signals with time structure. The method is closely related to blind separation of time-dependent source signals and ICA.*

Remark 2: *In the paper [6], we have presented a fixed-pointed algorithm for complexity pursuit, and if the signals have no time dependencies, the method reduces to the well-known FastICA algorithm [6]. Unfortunately, the algorithm does not performs more efficiently than original complexity pursuit [4] in the experiments (such as the performance index). Thus, the algorithm does not appear in the comparison.*

The nonlinearity is chosen as $\tanh(\cdot)$ and the step size is taken equal to 1 in the CPSYM and the CPDEF. We choose a first order AR model (the lag $\tau = 1$) in the three algorithms. The performance index defined as [2]:

$$\mathbf{PI} = \sum_{i=1}^{n}(\sum_{j=1}^{n}\frac{|p_{ij}|}{\max_k |p_{ik}|} - 1) + \sum_{j=1}^{n}(\sum_{i=1}^{n}\frac{|p_{ij}|}{\max_k |p_{kj}|} - 1), \tag{11}$$

where p_{ij} is the ijth element of $n \times n$ matrix $\mathbf{P} = \mathbf{WVA}$. The larger the value **PI**, the poorer the performance. We give here the mean values of 100 independent trials for unmixing error and time consumption (measured on a 1.8 GHz Pentium IV PC).

4.1. Experiment 1

In the first experiment the mixing matrix $\mathbf{A}$ is chosen to be nearly singular:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 2.00001 \end{pmatrix}.$$

This means that all sensor signals look almost identical. Such a situation may occur when sensors are located very close to each other. The CPSYM does not work in this ill-conditioned problem. However, the CPDEF and the FPIIA can achieve the BSS. Table 1 gives the results. Obviously, the proposed fixed-point algorithm performs more efficiently in both convergence speech and steady-state accuracy. The convergence of the CPDEF is quite slow. The plots of original sources, their mixtures and separated signals by the FPIIA are shown in Figure 1.

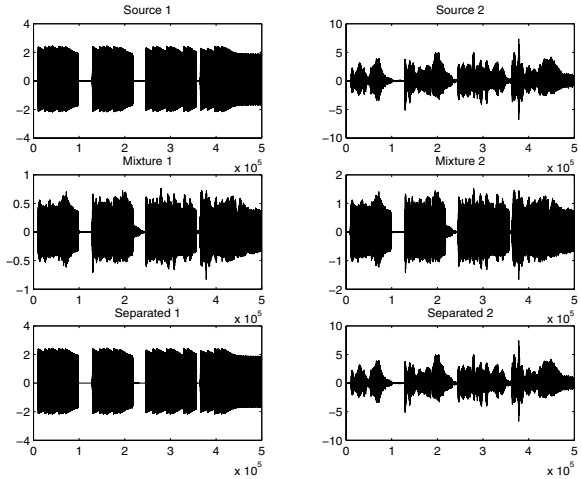

Fig. 1. Plots illustrating Experiment 1: two original signals, two mixtures and two separated signals by our algorithm.

Table 2. The average performance indexes **PI** and CPU-time in seconds for different algorithms in the Experiment 2.

Algorithms	**CPSYM**	**FPIIA**
PI	4.5662	4.5094
CPU − time	10.62	1.03

4.2. Experiment 2

In the second experiment ten natural speech signals have been mixed using randomly chosen 10×10 nonsingular mixing matrix. In this case, the CPDEF with the deflation scheme provides very poor performance perhaps because the estimation errors in the sources that are estimated first accumulate and increase the errors in the later estimated sources. Thus, it is excluded from this comparison. The results are shown in Table 2.

In the case, we can see that the FPIIA and the CPSYM degrade in performance but the FPIIA still works better. The CPSYM can achieve the BSS but computational load is about 10 times larger than for the FPIIA. From a subjective listening point of view, the separation of the ten nature speech example by our algorithm is remarkable for the high intelligibility of the recovered sentences, in spite of some crosstalk. Figure 2

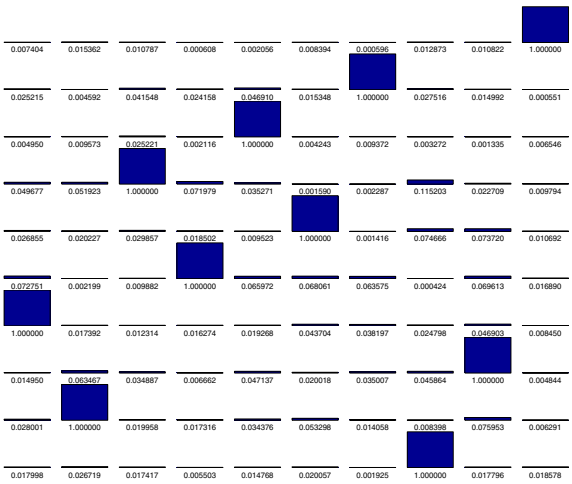

Fig. 2. Plots illustrating Experiment 2: performance matrix $\mathbf{P}$ for the separation of ten sources by our algorithm after normalizing.

Table 3. The average performance indexes $\mathbf{PI}$ and CPU-time in seconds for different algorithms in the Experiment 3.

$\mathbf{Algorithms}$	$\mathbf{CPSYM}$	$\mathbf{FPIIA}$
$\mathbf{PI}$	0.9996	0.6994
$\mathbf{CPU-time}$	3.50	0.83

shows the performance of the matrix $\mathbf{P}$ by our algorithm after normalizing.

4.3. Experiment 3

In this simulation, the source signals sampled from four nature images have been mixed using randomly chosen 4×4 nonsingular mixing matrix. The CPDEF fails with these data due to the same reason in the second experiment and does not appear in this comparison. The results are depicted in Table 3. The superiority of the FPIIA over CPSYM is straightforward. Figure 3 shows four original images, four mixtures and four separated images by our algorithm.

In short, the proposed fixed-point algorithm outperforms the other two algorithms in both convergence speed and performance indexes.

5. CONCLUSIONS

In this paper we have presented a fixed-point algorithm for blind separation of temporally correlated sources. Contrast to other gradient based algorithms, the algorithm is computationally simple, provides very fast convergence and does not need choose any learning step sizes. This means that the algorithm is a very appealing method and easy to use. We have confirmed by extensive computer simulations that the derived algorithm has shown excellent performance when the source signals are temporally correlated even if they are statistically dependent. Further theoretical endeavors of the fixed-point learning algorithm are subjects for future study.

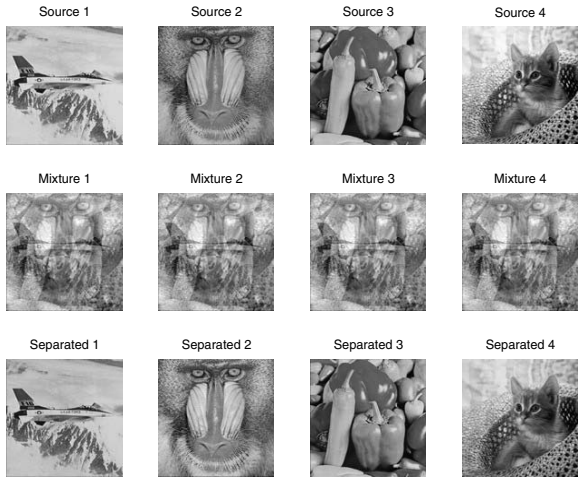

Fig. 3. Plots illustrating Experiment 3: four original images, four mixtures and four separated images by our algorithm.

6. REFERENCES

[1] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique based on second order statistics," *IEEE Transactions on Signal Processing*, vol. 45, pp. 434-444, 1997.

[2] A. Cichocki and S. Amari, "Adaptive Blind Signal and Image Processing," Wiley, 2002.

[3] J.H. Friedman and J.W. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Transactions of Computers*, C-23, vol. 9, pp. 881-890, 1974.

[4] A. Hyv$\ddot{a}$rinen, "Complexity pursuit: separating interesting components from time-series," *Neural Computation*, vol. 13, no.4, pp. 883-898, 2001.

[5] A. Hyv$\ddot{a}$rinen, J. Karhunen, and E. Oja, "Independent component analysis," Wiley, 2001.

[6] Z. Shi, H. Tang, and Y. Tang, "A fast fixed-point algorithm for complexity pursuit," *Neurocomputing*, vol. 64, pp. 529-536, 2005.

[7] Y.G. Zhang and C.S. Zhang, "Separation of music signals by harmonic structure modeling," Advances in Neural Information Processing Systems, pp. 1617-1624, 2006.