ON THE GENERALIZATION OF BLIND SOURCE SEPARATION ALGORITHMS FROM INSTANTANEOUS TO CONVOLUTIVE MIXTURES

Tiemin Mei, Alfred Mertins and Fuliang Yin[*]

Institute for Signal Processing, University of Lübeck, 23538 Lübeck, Germany

ABSTRACT

Many convolutive blind source separation (BSS) approaches are generalized from instantaneous BSS methods in either time or frequency domain. In this paper, we establish in a general way the inner relationship between the time-domain instantaneous BSS and the frequency-domain convolutive BSS. From this point of view, the time-domain approaches for instantaneous mixture separation are generalized to those for convolutive mixture separation in the frequency domain. Two examples are given to illustrate the feasibility of the proposed approach.

1. INTRODUCTION

Blind source separation (BSS) is to recover a set of unknown signal sources from observations that are unknown mixtures of those sources. A challenging situation for BSS is that mixing processes are convolutive, where observations are the combinations of the unknown filtered versions of the signal sources. The problem has attracted extensive research work in the research communities due to its many potential applications, such as audio processing, image processing, communication systems and biomedical signal processing.

As BSS of instantaneous mixtures is concerned, many efficient approaches have been proposed and used in practical applications. The BSS of instantaneous mixtures is just a special case of convolutive mixtures, so it is quite natural to generalize these successful BSS approaches for instantaneous mixtures to the separation of convolutive mixtures.

It has been investigated to generalize BSS approaches for instantaneous mixtures to BSS of convolutive mixtures in two different ways. One is to achieve convolutive mixture BSS directly in the time domain, and the other is to work in the frequency domain. In the time domain, the objective function of instantaneous BSS is usually revised so as to adapt to the convolutive case. Examples are the joint block diagonalization approach [1], Amari's method derived from the natural gradient algorithm [2], and that in [4]. However, these approaches are not very efficient for long mixing channels, such as those in well-known cocktail party problems, where the mixing channels may have $500-2000$ taps or more if modelled by FIR filters.

Frequency-domain approaches have been considered as the most promising technique for convolutive BSS, especially for cases of long mixing channels. In the frequency domain, there are two different ways to exploit instantaneous BSS for convolutive BSS. One way is to apply firstly an instantaneous BSS approach to every frequency bin (sub-signal) of convolutive mixtures. Here, the convolutive mixtures are separated frequency bin by frequency bin. Secondly, some measures are taken to align these sub-signals to overcome the local permutation. Lastly they are reconstructed into source signals [5][6][8]. The other way is to apply the objective function for instantaneous BSS to the frequency model of convolutive mixtures and to integrate this frequency-dependent objective function in order to generate an objective function that is a function of the time-domain parameters of the separation system. The optimization leads directly to the convolutive BSS [9][10].

In the following, we establish a general way how to generalize a time-domain instantaneous BSS algorithm directly to a frequency-domain convolutive BSS algorithm.

2. MODEL OF INSTANTANEOUS BSS

We consider the N-by-N case, that is, there are N signal sources and N observed signals. We assume that the sources are complex valued and are of zero mean, and that the mixing system is instantaneous. The mixing process can be described as follows:

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

where $\mathbf{s}(n) = [\, s_1(n),\, s_2(n), \ldots,\, s_N(n)\,]^{\mathrm{T}}$ are the signal sources, $\mathbf{x}(n) = [x_1(n), x_2(n), \ldots, x_N(n)]^{\mathrm{T}}$ are the observed signals and A is the mixing matrix which is assumed

This work has been supported by the German Research Foundation under Grant No. ME 1170/1. In addition, T. Mei is on leave from the School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110168 China. [*]F. Yin is with the School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116023 China. (mei@isip.uni-luebeck.de; mertins@isip.uni-luebeck.de; flyin@dlut.edu.cn)

978-1-4244-2241-8/08/$25.00 ©2008 IEEE

to be nonsingular and time invariant. The task of BSS is to recover the sources from the observations in the form:

$$\mathbf{y}(n) = W\mathbf{x}(n) \tag{2}$$

where $\mathbf{y}(n) = [y_1(n), y_2(n), ..., y_N(n)]^{\mathrm{T}}$ is the output of the separation system, and *W* is the matrix describing the separation network. Combining (1) and (2) gives:

$$\mathbf{y}(n) = G\mathbf{s}(n) \tag{3}$$

where $G = WA$, which is the transform matrix from $\mathbf{s}(n)$ to $\mathbf{y}(n)$. Separation is considered to be successful if we can find a matrix *W* such that *G* is the product of a diagonal matrix *D* and a permutation matrix *P*.

Consider a real-valued objective function $\Phi(W, W^*)$, where * denotes complex conjugation. Under the assumption that the sources satisfy all the required separability conditions for the given objective function, we obtain the gradient-based learning rule

$$W^{l+1} = W^l - \mu\nabla\Phi(W^l, W^{l^*}) \tag{4}$$

where l is the iteration index and

$$\nabla\Phi(W, W^*) = \frac{\mathrm{d}\Phi(W, W^*)}{\mathrm{d}W^*}. \tag{5}$$

According to the relationship between the ordinary gradient and the natural gradient [3],

$$\nabla_{\mathrm{natural}}\Phi(W, W^*) = \nabla\Phi(W, W^*)W^{\mathrm{H}}W, \tag{6}$$

where the superscript $^{\mathrm{H}}$ denotes Hermitian transposition, we get the natural gradient-based algorithm:

$$W^{l+1} = W^l - \mu\nabla\Phi(W^l, W^{l^*})W^{l^{\mathrm{H}}}W^l. \tag{7}$$

Whatever the objective function is, we will get the above type of gradient-based learning rule. For instance, if the objective function is defined on the basis of Kullback-Leibler divergence, we get Amari's natural gradient based algorithm [2]; if the objective function is defined as the joint diagonalization of correlation matrices, we get the decorrelation-based BSS algorithm [11][12].

3. MODEL OF CONVOLUTIVE BSS

We still consider the N-by-N case, that is, there are N signal sources, N observation signals and N separated signals as well. The mixing channels are assumed to be FIR of length L, and the separation channels are also FIR and their length (M) is chosen so that $M \geq (N-1)(L-1)+1$ in order to achieve satisfying performance. Also we have the following assumptions regarding the sources and the mixing processes [10]:

$(A1)$ Signal sources $\mathbf{s}(n) = [s_1(n), s_2(n), ..., s_N(n)]^{\mathrm{T}}$ are real, zero mean and independent of each other.

$(A2)$ The signal sources $\mathbf{s}(n)$ are non-stationary or non-Gaussian, which means that their auto-power spectral densities are time-varying in nature or they have nonzero high-order cumulants.

$(A3)$ The mixing system $A(n) = [a_{ij}(n)]_{N\times N}$ is linear and time invariant (LTI), where $a_{ij}(n)$ is the impulse response of the channel from source $s_j(n)$ to observation $x_i(n)$.

$(A4)$ The transfer matrix of the mixing system $\mathbf{A}(z) = \sum_{n=0}^{L-1} A(n)z^{-n}$ is nonsingular on the unit circle in the complex plane.

Assumptions $(A1)$ and $(A3)$ are the basic conditions for BSS; Assumptions $(A2)$ and $(A4)$ are necessary for the separation of sub-signals at a given frequency.

The noise-free convolute mixing model is given as follows,

$$\mathbf{x}(n) = A(n) * \mathbf{s}(n) = \sum_{l=0}^{L-1} A(l)\mathbf{s}(n-l) \tag{8}$$

where $*$ denotes the convolution operation, $\mathbf{s}(n)$ is the signal source vector, $\mathbf{x}(n)$ is the mixture vector, and $A(n) = [a_{ij}(n)]_{N\times N}$ is the mixing matrix.

The separation system output $\mathbf{y}(n) = [y_1(n), y_2(n), \ldots \ldots, y_N(n)]^{\mathrm{T}}$ is given as

$$\mathbf{y}(n) = H(n) * \mathbf{x}(n) = \sum_{l=0}^{M-1} H(l)\mathbf{x}(n-l) \tag{9}$$

where $H(n) = [h_{ij}(n)]_{N\times N}$ is the separation matrix and $h_{ij}(n)$ denotes the impulse response of the FIR channel from $x_j(n)$ to output $y_i(n)$. From (8) and (9), we have

$$\mathbf{y}(n) = H(n) * A(n) * \mathbf{s}(n) = G(n) * \mathbf{s}(n) \tag{10}$$

with $G(n) = H(n) * A(n)$. Equivalently, in the z-domain, we have

$$\mathbf{Y}(z) = \mathbf{G}(z)\mathbf{S}(z). \tag{11}$$

BSS is considered to be successful if the output $\mathbf{y}(n)$ is at most a permuted and filtered version of the signal sources $\mathbf{s}(n)$, in which case $\mathbf{G}(z)$ is a product of a permutation matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}(z)$:

$$\mathbf{G}(z) = \mathbf{P}\mathbf{D}(z). \tag{12}$$

We use the short-time Fourier transform (STFT) to describe the mixing and separating processes, which, based on (8) and (9), are given as follows if the boundary effect of the linear convolution is negligible:

$$\mathbf{X}(n, e^{j\omega}) = \mathbf{A}(e^{j\omega})\mathbf{S}(n, e^{j\omega}) \tag{13}$$

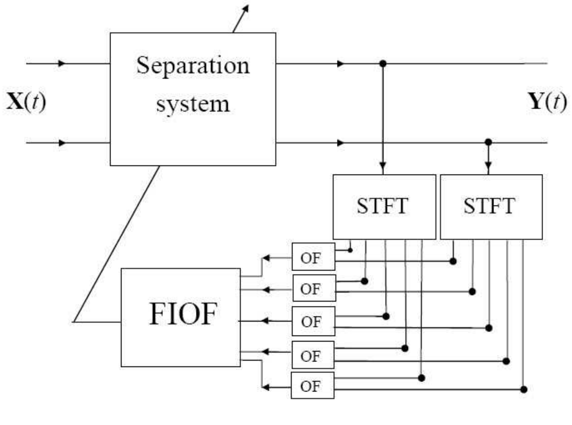

Fig. 1. This diagram illustrates the relationship between instantaneous and convolutive BSS.

and

$$\mathbf{Y}(n, e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(n, e^{j\omega}) \tag{14}$$

where n is the time index which describes the short-time signal spectra in different time windows. If the original signal sources $\mathbf{s}(n)$ are nonstationary, then the sub-signal sources $\mathbf{S}(n, e^{j\omega})$ are also nonstationary.

By applying the instantaneous-BSS objective function $\Phi(W, W^*)$ to every frequency bin of the short-time Fourier transform model (14) and then performing integration, we get the so called frequency-domain integrated objective function (FIOF):

$$\Psi(H(n)|_{n=0,1,\ldots,M-1}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega}))d\omega. \tag{15}$$

This can be seen more clearly in Fig. 1.

For the gradient of the FIOF with respect to the separation filter matrices $H(n)$, we obtain

$$\begin{aligned}&\frac{\partial\Psi(H(n)|_{n=0,1,\ldots,M-1})}{\partial H(n)}\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\nabla_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)e^{j\omega n}\right.\\ &\quad\left.+\nabla^*_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)e^{-j\omega n}\right]d\omega\end{aligned} \tag{16}$$

where

$$\nabla_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right) = \left[\frac{\partial\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)}{\partial H^*_{pq}(e^{j\omega})}\right],$$

$$\nabla^*_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right) = \left[\frac{\partial\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)}{\partial H_{pq}(e^{j\omega})}\right].$$

According to [13], the natural gradient is

$$\begin{aligned}&\left[\frac{\partial\Psi(H(n)|_{n=0,1,\ldots,M-1})}{\partial H(n)}\right]_{\mathrm{natural}}\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\nabla_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)e^{j\omega n}\right.\\ &\quad\left.+\nabla^*_{\mathbf{H}}\Phi\left(\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega}), \mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\right)e^{-j\omega n}\right]\mathbf{H}^{\mathrm{H}}(\mathrm{e}^{\mathrm{j}\omega})\mathbf{H}(\mathrm{e}^{\mathrm{j}\omega})\mathrm{d}\omega.\end{aligned} \tag{17}$$

Based on the natural gradient, we obtain the learning rule as follows:

$$H^{l+1}(n) = H^l(n) - \mu\left[\frac{\partial\Psi(H^l(n)|_{n=0,1,\ldots,M-1})}{\partial H^l(n)}\right]_{\mathrm{natural}}. \tag{18}$$

The answer to the question of why this FIOF-based BSS approach can avoid the permutation problem is that the time-domain impulse responses have been limited by length, which is equivalent to the frequency-domain smoothness constraints on the unmixing filters [6][7].

4. EXAMPLES

We present two concrete examples to show the validity and feasibility of this method.

4.1. Extending Amari's and Cichocki's natural-gradient based algorithm to convolutive mixture separation

Amari's and Cichocki's natural-gradient based algorithm is one of the well known BSS approaches for instantaneous mixtures. It is also generalized to the blind separation of convolutive mixtures in different ways.

For the instantaneous mixing cases, Amari *et al.* proposed an algorithm based on the KL-divergence, in which the objective function is given as

$$\Phi(W) = -\frac{1}{2}\log\left(\det(W^{\mathrm{T}}W)\right) - \sum_{i=1}^{N}\log p_i(y_i). \tag{19}$$

Based on the above objective function, the natural-gradient based approach for instantaneous mixtures was derived as follows [2]:

$$W^{l+1} = W^l + \mu\left(\mathbf{I} - \mathbf{f}(\mathbf{y}(l))\mathbf{y}^{\mathrm{T}}(l)\right)W^l \tag{20}$$

where l is a time index and iteration indicator, and $\mathbf{I}$ is the identity matrix. The term

$$\mathbf{f}(\mathbf{y}(l)) = [f_1(y_1(l)), f_2(y_2(l)), \ldots, f_N(y_N(l))]^{\mathrm{T}}$$

with

$$f_i(y_i(l)) = -\frac{d\log(p_i(y_i(l)))}{dy_i(l)} = -\frac{p_i'(y_i(l))}{p_i(y_i(l))},$$

which depends on the p.d.f of the sources, is referred to as the activation function.

Following the idea presented in Section 3, we define the following frequency-domain integrated objective function [10]:

$$\Psi(l, H(n)\,|_{n=0,1,\ldots,M-1}) = -\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(\det\left(\mathbf{H}^{\mathrm{H}}(e^{j\omega})\mathbf{H}(e^{j\omega})\right)\right)d\omega - \frac{1}{2\pi}\sum_{i=1}^{N}\int_{-\pi}^{\pi}\log p_i\left(\left|y_i(l, e^{j\omega})\right|\right)d\omega. \quad (21)$$

We obtain the natural-gradient based adaptive learning rule as follows:

$$H^{l+1}(n) = H^{l}(n) + \mu \times \frac{1}{2\pi}\int_{-\pi}^{\pi}[\mathbf{I} - \mathbf{F}(\mathbf{Y}(l, e^{j\omega}))\mathbf{Y}^{\mathrm{H}}(l, e^{j\omega})]\mathbf{H}^{l}(e^{j\omega})e^{j\omega n}d\omega \quad (22)$$

where,

$$\mathbf{F}\left(\mathbf{Y}(l, e^{j\omega})\right) = [f_1(y_1(l, e^{j\omega})), ..., f_N(y_N(l, e^{j\omega}))]^{\mathrm{T}}$$

and

$$f_p(y_p(l, e^{j\omega})) = -\frac{\partial\left(\log p_p\left(\left|y_p(l, e^{j\omega})\right|\right)\right)}{\partial\left|y_p(l, e^{j\omega})\right|}e^{j\theta\left(y_p(l, e^{j\omega})\right)}$$

is the activation function with

$$\theta\left(y_p(l, e^{j\omega})\right) = \arg\left(y_p(l, e^{j\omega})\right)$$

being the phase of $y_p(l, e^{j\omega})$.

4.2. Extending correlation-based algorithms to convolutive mixture separation

For instantaneous mixtures, correlation-based BSS approaches have been studied by many researchers, so those works provide a solid background for studying correlation-based approaches for convolutive mixture separation [11][12].

For nonstationary sources, we define the objective function with correlation matrix $R_{\mathbf{yy}}(l) = E[\mathbf{y}^{\mathrm{T}}(l)\mathbf{y}(l)]$ on the basis of Hadamard's inequality as

$$\Phi(W) = \Phi(R_{\mathbf{yy}}(l)) = \frac{1}{2}\log\left[\frac{\det[D_{\mathbf{yy}}(l)]}{\det[R_{\mathbf{yy}}(l)]}\right] \quad (23)$$

where $D_{\mathbf{yy}}(l)$ is a diagonal matrix whose diagonal elements are just those of the correlation matrix $R_{\mathbf{yy}}(l)$.

The natural-gradient based online learning rule for instantaneous mixture separation is given as

$$W^{l+1} = W^{l} - \mu[D_{\mathbf{yy}}^{-1}(l)R_{\mathbf{yy}}(l) - \mathbf{I}]W^{l} \quad (24)$$

where $\mathbf{I}$ is the identity matrix.

Applying the objective function (23) to every frequency bin of the convolutive model in (14) and performing integration, we obtain the FIOF for convolutive mixture separation:

$$\Psi(H(n)|_{n=0,1,\ldots,M-1}) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(\frac{\det\left[\mathbf{D}_{\mathbf{YY}}(l,\omega)\right]}{\det\left[\mathbf{P}_{\mathbf{YY}}(l,\omega)\right]}\right)d\omega \quad (25)$$

where $\mathbf{P}_{\mathbf{YY}}(l,\omega)$ is the instant correlation matrix of subsignals $\mathbf{Y}(l, e^{j\omega})$, which can also be interpreted as the instant power spectral density matrix of the separated sources $\mathbf{y}(n)$; $\mathbf{D}_{\mathbf{YY}}(l,\omega)$ is a diagonal matrix with the diagonal elements of $\mathbf{P}_{\mathbf{YY}}(l,\omega)$ as its diagonal elements.

Deducing in the same way as in Section 3, we obtain the following correlation-based learning rule for convolutive mixture separation:

$$H^{l+1}(n) = H^{l}(n) - \mu \times \quad (26)$$
$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\mathbf{D}_{\mathbf{YY}}^{-1}(l,\omega)\mathbf{P}_{\mathbf{YY}}(l,\omega) - \mathbf{I}\right]\mathbf{H}^{l}(e^{j\omega})e^{j\omega n}d\omega.$$

5. SIMULATIONS

In this section, we investigate the question of how the separation performance of the FIOF-based approaches depends on the signal-to-interference-ratios (SIRs) of the inputs of the separation system. Sawada's data of speech signals [16] (the case of two sources) are used in this experiment. The cross-channel components are multiplied by factors to adjust the mixing depth of two sources to obtain input mixtures of different SIRs. The relationship between the averaged input SIR's and the averaged output SIR's is shown in Fig. 2. It shows that it is difficult to separate the mixtures of very low SIRs. This result was obtained with the algorithm (22) (algorithm (26) will give a similar result). The corresponding parameters are as follows: the length of unmixing filters is 512; the fast Fourier transform size is 2048; the learning rate is $\mu = 0.01 - (0.01 - 0.0001)t/t_{\max}$ (where t is the iteration index and $t_{\max}$ is the maximum number of iterations).

6. CONCLUSIONS

In this paper, we presented a general relationship between the BSS algorithms for instantaneous and convolutive mixtures by transforming the time-domain instantaneous BSS to frequency-domain convolutive BSS and integrating the frequency-domain criterion. So we can fully use the results provided for instantaneous BSS. This kind of transform is different from the frequency-domain implementation of a time-domain convolutive BSS algorithm. It is also different

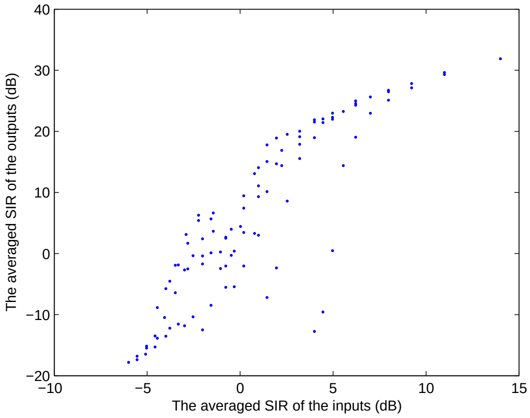

Fig. 2. The relationship between the signal-to-interference-ratios of the inputs and outputs of the separation system.

from applying a time-domain instantaneous BSS algorithm directly to the frequency-domain model of convolutive mixtures. It is a kind of hybrid method of time-domain and frequency-domain approaches. So it has the advantages of both time-domain and frequency-domain approaches: efficient computation and no local permutation.

7. REFERENCES

[1] H. Bousbia-Salah, A. Belouchrani, K. Abed-Meraim, “Blind separation of convolutive mixtures using joint block diagonalization,” *Sixth International Symposium on Signal Processing and its Applications*, vol. 1, pp.13-16, 2001.

[2] S. Amari and A. Cichocki, “Adaptive blind signal processing-neural network approaches,” *Proc. of the IEEE*, vol.86, no.10, pp.2026-2048, 1998.

[3] S. Amari, “Natural gradient works efficiently in learning,” *Neural Computation*, vol. 10, pp. 251–276, 1998.

[4] M. Kawamoto and K. Matsuoka et al, “A method of blind separation for convolved non-stationary signals,” *Neurocomputing*, vol.22, pp.157-171, 1998.

[5] P. Smaragdis, “Blind separation of convolved mixtures in the frequency domain,” *Neurocomputing*, vol. 22, pp.21-34, 1998.

[6] L. Parra, C. Spence, “Convolutive blind separation of non-stationary sources,” *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 3, pp. 320-327, 2000.

[7] T. Mei, A. Mertins, F. Yin, J. Xi, and J.F. Chicharo, “Blind source separation for convolutive mixtures based on the joint diagonalization of power spectral density matrices,” *Signal Processing*, vol. 88, no. 8, pp. 1990-2007, 2008.

[8] K. Rahbar and J. Reilly, “A frequency domain method for blind source separation of convolutive audio mixtures,” IEEE Trans. Speech and Audio Processing, vol. 13, no. 5, pp. 832-844, 2005.

[9] M. Kawamoto, Y. Inouye, “Blind deconvolution of MIMO-FIR systems with colored inputs using second-order statistics,” *IEICE Trans. Fundamentals*, vol. E86-A, no. 3, pp. 597-604, 2003.

[10] T. Mei, J. Xi, F. Yin, A. Mertins and J. F. Chicharo, “Blind source separation based on time-domain optimization of a frequency-domain independence criterion,” *IEEE Trans. on Audio,Speech and Language Processing*, vol. 14, no. 6, pp. 2075-2085, 2006.

[11] S. Choi, A. Cichocki, and S. Amari, “Equivariant nonstationary source separation,” *Neural Networks*, vol. 15, pp. 121–130, 2002.

[12] F. Yin, T. Mei, J. Wang, ”Blind source separation based on decorrelation and nonstationarity,” *IEEE Transactions on Circuits and Systems I*, vol. 54, no. 5, pp. 1150-1158, 2007.

[13] I. Sabala, A. Cichocki, S. Amari, “Relationships between Instantaneous blind source separation and multichannel blind deconvolution,” *IEEE World Congress on Computational Intelligence, Neural Networks Proceedings*, vol.1, pp.39-44, 1998.

[14] T. Lee, A. J. Bell and R. Orglmeister, “Blind source separation of real world signals,” *International Conference on Neural Networks*, vol. 4, pp. 2129-2134, 1997.

[15] T. Mei, A. Mertins, Fuliang Yin, J. Xi, J. F. Chicharo, “Blind source separation for convolutive mixtures based on the joint diagonalization of power spectral density matrices,” *Submmited to Signal Processing*.

[16] http://www.kecl.ntt.co.jp/icl/signal/sawada/demo/bss2to4/index.html