A distortion free learning algorithm for feedforward BSS and its comparative study with feedback BSS

2006 International Joint Conference on Neural Networks
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

A Distortion Free Learning Algorithm for Feedforward BSS and Its Comparative Study with Feedback BSS

Akihide Horita, *Student Member, IEEE*, Kenji Nakayama, Akihiro Hirano *Members, IEEE*, and Yasuhiro Dejima *Non Member, IEEE*

***Abstract*— Source separation and signal distortion are theoretically analyzed for the FF-BSS systems implemented in both the time and frequency domains and the FB-BSS system. The FF-BSS systems have some degree of freedom, and cause some signal distortion. The FB-BSS has a unique solution for complete separation and distortion free. Next, the condition for complete separation and distortion free is derived for the FF-BSS systems. This condition is applied to the learning algorithms. Computer simulations by using speech signals and stationary colored signals are carried out for the conventional methods and the new learning algorithms employing the proposed distortion free constraint. The proposed method can drastically suppress signal distortion, while maintaining high separation performance. The FB-BSS system also demonstrates good performances. The FF-BSS systems and the FB-BSS system are compared based on the transmission time difference in the mixing process. Location of the signal sources and the sensors are rather limited in the FB-BSS system.**

I. Introduction

Signal processing, including noise cancellation, echo cancellation, equalization of transmission lines, estimation and restoration of signals have become a very important research area. In some cases, we do not have enough information about signals and their interference. Furthermore, their mixing and transmission processes are not well known in advance. In these kind of situations, blind source separation (BSS) technology using statistical properties of signal sources have become very important [1]-[5].

Since, in many applications mixing processes are convolutive mixtures, several methods in the time domain and the frequency domain have been proposed. Two kinds of proposed network structures are feedforward (FF) and feedback (FB) structures. Separation performance is highly dependent on the signal sources and the transfer functions in the mixture [7]-[10],[13],[14],[16]-[18].

BSS learning algorithms make the output signals to be statistically independent. This direction cannot always guarantee distortion free separation. A method in which the distance between the observed signals and the separated signals is added to the cost function has been proposed[11]. However, since the observations include many kinds of signal sources, it is difficult to suppress signal distortion. Furthermore, even though signal distortion in BSS systems is an important problem, it has not been addressed well up to now [19].

Akihide Horita, Kenji Nakayama, Akihiro Hirano and Yasuhiro Dejima are with the Graduate School of Natural Science and Technology, Kanazawa University, Kakuma-machi, Kanazawa, 920-1192, JAPAN (email: horita@leo.ec.t.kanazawa-u.ac.jp, nakayama@t,kanazawa-u.ac.jp).

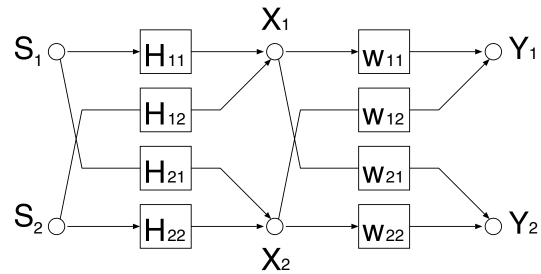

Fig. 1. FF-BSS system with 2 signal sources and 2 sensors.

In this paper, first, an evaluation measure of signal distortion is discussed. Secondly, conditions for source separation and distortion free are derived. Based on these conditions, convergence properties are analyzed. Furthermore, a new learning algorithm for FF-BSS systems is proposed. We analyze the performance of our new learning algorithm in comparison with conventional methods, through computer simulations. The simulation results support our theoretical analysis.

II. FF-BSS Systems for Convolutive Mixture

A. Network Structure and Equations

For simplicity, 2 signal sources and 2 sensors are used. A block diagram is shown in Fig.1. The observations and the output signals are given by:

$$x_j(n) = \sum_{i=1}^{2}\sum_{l=0}^{K_h-1} h_{ji}(l)s_i(n-l), j=1,2 \quad (1)$$

$$y_k(n) = \sum_{j=1}^{2}\sum_{l=0}^{K_w-1} w_{kj}(l)x_j(n-l), k=1,2 \quad (2)$$

B. Learning Algorithm in Time Domain

The learning algorithm is derived following a natural gradient algorithm using a mutual information as a cost function [6].

$$w_{kj}(n+1,l) = w_{kj}(n,l) + \eta\{w_{kj}(n,l) - \sum_{\substack{p=1\\ \neq j}}^{2}\sum_{q=0}^{K_w-1} \varphi(y_k(n))y_p(n-l+q)w_{pj}(n,q)\} \quad (3)$$

$$\varphi(y_k(n)) = \frac{1-e^{-y_k(n)}}{1+e^{-y_k(n)}} \quad (4)$$

The learning rate is represented by η.

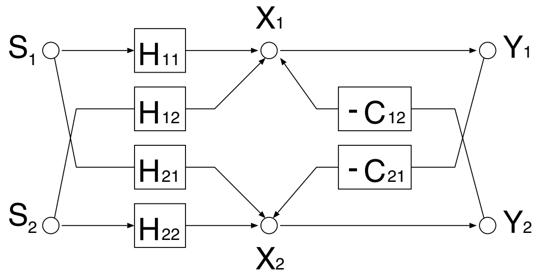

Fig. 2. FB-BSS system with 2 signal sources and 2 sensors.

C. Learning Algorithm in Frequency Domain

Filter coefficients in the separation block are trained according [6],[12],[15],

$$\begin{aligned}\boldsymbol{W}(r+1,m) &= \boldsymbol{W}(r,m) \\ &+\eta[\boldsymbol{I}-\langle\boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) \quad (5)\\ \Phi(\boldsymbol{Y}(r,m)) &= \frac{1}{1+e^{-\boldsymbol{Y}^R(r,m)}}+\frac{j}{1+e^{-\boldsymbol{Y}^I(r,m)}} \quad (6)\end{aligned}$$

r is the block number used in FFT, and m indicates the frequency point in each block. $<>$ is an averaging operation. $\boldsymbol{W}(r,m)$ is the weight matrix of the r-th FFT block and the m-th frequency point. Its (k,j) element, represented by $W_{kj}(r,m)$, is the connection from the j-th observation to the k-th output. $\boldsymbol{Y}(r,m)$ is the output vector of the r-th FFT block and the m-th frequency point. Its k-th element, represented by $Y_k(r,m)$, is the k-th output. $\boldsymbol{Y}^R(r,m)$ and $\boldsymbol{Y}^I(r,m)$ indicate the real part and the imaginary part.

It has been reported that Eq.(5) includes a needless constraint which transforms the outputs into white signals, and as a result a learning algorithm that avoids whitening has been proposed[15].

$$\begin{aligned}\boldsymbol{W}(r+1,m) &= \boldsymbol{W}(r,m)+\eta[diag(\langle\boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle) \\ &-\langle\boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) \quad (7)\end{aligned}$$

In this paper, systems following Eqs.(5) and (7) will be referred to as FF-BSS freq.(1) and (2), respectively.

III. FB-BSS Systems for Convolutive Mixture

A. Network Structure and Equations

Figure 2 shows the FB-BSS system proposed by Jutten et al [1]. The mixing stage has a convolutive structure. The blocks C_{kj} consist of an FIR filter.

The observations and the output signals are expressed as follows:

$$x_j(n) = \sum_{i=1}^{N}\sum_{m=0}^{M_{ji}-1} h_{ji}(m)s_i(n-m) \quad (8)$$

$$y_k(n) = x_k(n)-\sum_{\substack{j=1\\ \neq k}}^{N}\sum_{l=0}^{L_{kj}-1} c_{kj}(l)y_j(n-l) \quad (9)$$

B. Learning Algorithm

In the FB-BSS system, a learning algorithm in the time domain is used [4]. The following learning algorithm has been derived by assuming several conditions [14],[16]. The signal sources $S_1(z)$ and $S_2(z)$ are located close to the sensors $X_1(z)$ and $X_2(z)$. Therefore, the time delays of $H_{ji}(z), i \neq j$ are slightly longer than those of $H_{ii}(z)$. Furthermore, the amplitude responses of $H_{ji}(z), i \neq j$ are smaller than those of $H_{ii}(z)$.

$$\begin{aligned}c_{kj}(n+1,l) &= c_{kj}(n,l) \\ &+ \eta f(y_k(n))g(y_j(n-l)) \quad (10)\end{aligned}$$

$f(y_k(n))$ and $g(y_j(n-l))$ are odd functions.

The above conditions are acceptable, when the distance between the sources or the sensors is large enough. When the distance is not large enough, the separation performance is unsatisfactory, because there are not enough time delays. This point will be analyzed in Sec.VIII-E.

IV. Criterion of Signal Distortion

How to evaluate signal distortion in the BSS systems is one of the problems. Learning algorithms used in BSS systems make the output signals to be statistically independent. Estimation of the mixing process is not taken into account. Especially, in convolutive mixtures, the output signals are not guaranteed to approach to the sources. Therefore, the signal sources observed at the sensors are taken into account as a criterion for the signal distortion [4],[19].

The signal distortion is evaluated by several measures, based on the signals, the transfer functions and their amplitude responses. The transfer functions from the i-th source to the k-th output $A_{ki}(e^{j\omega})$ is compared to that from the i-th source to the j-th sensor $H_{ji}(e^{j\omega})$. The output signals $A_{ki}(e^{j\omega})S_i(e^{j\omega})$ is compared to the observed signal $H_{ji}(e^{j\omega})S_i(e^{j\omega})$. Four kinds of measures are shown below.

$$\begin{aligned}\sigma_{d1a} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega}) \\ &- A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (11)\\ \sigma_{d1b} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})S_i(e^{j\omega})| \\ &- |A_{ki}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \quad (12)\\ \sigma_1 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (13)\\ SD_{1x} &= 10\log_{10}\frac{\sigma_{d1x}}{\sigma_1}, x=a,b \quad (14)\\ \sigma_{d2a} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})-A_{ki}(e^{j\omega})|^2 d\omega \quad (15)\\ \sigma_{d2b} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})|-|A_{ki}(e^{j\omega})|)^2 d\omega \quad (16)\\ \sigma_2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})|^2 d\omega \quad (17)\\ SD_{2x} &= 10\log_{10}\frac{\sigma_{d2x}}{\sigma_2}, x=a,b \quad (18)\end{aligned}$$

Since FF-BSS systems cannot control the output signal level, the output signal level might differ from the criteria. In order to neglect this scaling effect in the calculation of SD_{1x} and SD_{2x}, the following conditions are guaranteed.

$$\begin{aligned} & \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega}))|^2 d\omega \\ = & \frac{1}{2\pi}\int_{-\pi}^{\pi}|A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \end{aligned} \quad (19)$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})|^2 d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}|A_{ki}(e^{j\omega})|^2 d\omega \quad (20)$$

V. Source Separation and Signal Distortion in FF-BSS systems

A. Source Separation and Signal Distortion

For simplicity, an FF-BSS system with 2-sources and 2-sensors, shown in Fig.1, is used. Furthermore, $S_i(z)$ is assumed to be separated at the output $Y_i(z)$. This does not lose generality. Taking the signal distortion criterion into account, the condition for distortion-free source separation can be expressed as follows:

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (21)$$

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (22)$$

$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (23)$$

$$W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (24)$$

The above equations imply two conditions. First, complete source separation, that is the non-diagonal elements are all zero, as shown in Eqs.(22) and (23). Secondly, distortion free, that is the diagonal elements are $H_{ii}(z)$ as shown in Eqs.(21) and (24).

The conventional learning algorithm given by Eqs.(3)-(7) satisfies only Eqs.(22) and (23). Equations (21) and (24) are not satisfied. Therefore, distortion free is not guaranteed in general.

B. Signal Distortion in FF-BSS Systems Trained in Frequency Domain

In the frequency domain, there is a weighting effect. From Eqs.(5), (7), it follows that the correction of the weights are proportional to the product of the outputs $Y_i(r,m)Y_j(r,m)$. In the early stage of the learning process, the output $Y_i(r,m)$ includes both $S_i(r,m)$ and $S_j(r,m)$. Therefore, the correction of the weights are proportional to $S_i(r,m) + S_j(r,m)$. On the other hand, as the learning makes progress, the output $Y_i(r,m)$ includes mainly $S_i(r,m)$. Therefore, the correction of the weights are proportional to $S_i(r,m) \times S_j(r,m)$. If the signal sources are all speech, their spectra are similar to each other. In this case, the spectra of $S_i(r,m) + S_j(r,m)$ and $S_i(r,m) \times S_j(r,m)$ are similar to those of the signal sources. If the signal sources locate in different frequency bands, e.g. as in music, their spectra are not similar. In this case, $S_i(r,m)+S_j(r,m)$ and $S_i(r,m)\times S_j(r,m)$ have many peaks. Therefore, the learning process amplifies some part of spectrum of the signal source, which is not dominant. It can be expected that weighting will cause signal distortion, and makes source separation more difficult.

C. Learning Algorithm in Time Domain Suppressing Signal Distortion by Using Observation as Criteria

A learning algorithm for reducing signal distortion has been proposed [11]. The cost function includes the distance between the observed signals and the output signals. This means that the output signals are forced to approach to the observed signals. The update equation is given by

$$\begin{aligned} & \boldsymbol{w}(n+1,l) = \boldsymbol{w}(n,l) \\ + & \alpha \sum_{m=0}^{K_w-1}[\boldsymbol{I}\delta(n-m) - \langle \Phi(\boldsymbol{y}(n))\boldsymbol{y}^T(n-l+m)\rangle \\ - & \beta(\boldsymbol{y}(n) - \boldsymbol{x}(n))\boldsymbol{y}^T(n-l+m)]\boldsymbol{w}(n,m) \end{aligned} \quad (25)$$

$$\varphi(\boldsymbol{y}(n)) = \frac{1-e^{-\boldsymbol{y}(n)}}{1+e^{-\boldsymbol{y}(n)}} \quad (26)$$

In this method, the output signals $Y_i(z) = A_{ii}(z)S_i(z) + A_{ij}(z)S_j(z)$ tend to approach to the observed signals $X_i(z) = H_{ii}(z)S_i(z)+H_{ij}(z)S_j(z)$. When $S_i(z)$ and $S_j(z)$ are statistically independent, $A_{ii}(z)$ and $A_{ij}(z)$ are able to approach to $H_{ii}(z)$ and $H_{ij}(z)$, respectively. The former guarantees distortion free, however, the latter avoids source separation.

VI. Source Separation and Signal Distortion in FB-BSS Systems

There are two cases, in which possible solutions for perfect separation exist, as shown below:

$$(1)\ C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (27)$$

$$(2)\ C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \quad C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \quad (28)$$

It is assumed that the delay times of $H_{ii}(z)$ are shorter than those of $H_{ji}(z)$. This means that in Fig.2, the sensor of X_i is located close to S_i. From this assumption, the solutions in case (1) become causal systems. On the other hand, the solutions in case (2) are noncausal.

When $C_{kj}(z)$ satisfy the separation conditions Eqs.(27), the output signals are given by

$$Y_1(z) = H_{11}(z)S_1(z) \quad (29)$$

$$Y_2(z) = H_{22}(z)S_2(z) \quad (30)$$

They are exactly the same as the criteria of the signal distortion discussed in Sec.IV Therefore, the FB-BSS systems have a unique solution, which satisfies both source separation as well as the distortion free simultaneously. Thus, in the FB-BSS systems, if complete signal separation is achieved, distortion free is also automatically satisfied.

VII. Distortion Free Condition and Its Application to Learning Algorithm for FF-BSS Systems

A. FF-BSS Systems Trained in Frequency Domain

From the relations of Eqs.(22) and (23), $H_{ji}(z)$ are expressed.

$$H_{12}(z) = -\frac{W_{12}(z)}{W_{11}(z)}H_{22}(z) \quad (31)$$

$$H_{21}(z) = -\frac{W_{21}(z)}{W_{22}(z)}H_{11}(z) \quad (32)$$

By substituting the above equations into Eqs.(21) and (24), $H_{ji}(z)$ can be removed, and the following equations consisting only of $W_{kj}(z)$ can be obtained.

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z) \quad (33)$$

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{11}(z) \quad (34)$$

From these equations, $W_{11}(z) = W_{22}(z)$ is derived. Therefore, the above equations result in

$$W_{jj}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \quad (35)$$
$$j = 1, 2, k = 1, 2, j \neq k$$

This 2nd-order equation expresses the condition for both complete source separation and distortion free. This equation is solved for $W_{11}(z)$ and $W_{22}(z)$ as follows:

$$W_{jj}(z) = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2}, j = 1, 2 \quad (36)$$

This constraint can be included in the learning processes for the FF-BSS systems in the frequency domain. Eqs.(5) and (7). Furthermore, $\tilde{W}_{jj}(z)$, which are determined from Eq.(36) are taken into account to some extent, as shown in the following.

$$\tilde{W}_{jj}(r+1, m) = (1-\alpha)W(r+1, m) + \alpha\frac{1 + \sqrt{1 + 4W_{12}(r, m)W_{21}(rm)}}{2} \quad (37)$$

B. Learning Algorithm with Constraint in Time Domain

In this section, a new learning algorithm for FF-BSS systems, trained in the time domain, is proposed. The constraint given by Eq.(36) is taken into account in the learning process. Equation (36) is rewritten as follows:

$$(2W_{jj}(z) - 1)^2 = 1 + 4W_{12}(z)W_{21}(z) \quad (38)$$

This constraint is used in the learning process as follows: Given $W_{12}(z)$ and $W_{21}(z)$, $W_{jj}(z)$ are obtained so as to satisfy the relation of Eq.(38).

The condition for the distortion free source separation holds complete separation and distortion free. However, the learning of the separation block starts from an initial guess. Therefore, in the early stage of the learning process, the signal sources are not well separated. Taking this situation into account, the constraint of Eq.(38) is gradually imposed as the learning process progresses. The following learning algorithm is proposed.

$$w_{kj}(n+1, l) = w_{kj}(n) + \eta\{w_{kj}(n) - \sum_{o=0}^{K_w-1}\sum_{p=1}^{2}\phi(y_k(n))y_p(n-o+p)w_{kp}(n, o)\} \quad (39)$$

$$w_{jj}(N+1, l) = (1-\alpha)w_{jj}(n+1, l) + \alpha\tilde{w}_{jj}(n+1) \quad (40)$$

$\tilde{w}_{jj}(n+1)$ is determined so as to satisfy the relation of Eq.(38). α is set to a small positive number.

C. Generalization of the Constraint

The distortion free constraint can be extended to more than two channels. The details will be described in Appendix.

TABLE I
Abbreviation for FF-BSS with Different Learning and Implementation.

FF-BSS time	Eqs.(3)-(4) [6]
FF-BSS time(DF)	Eqs.(3)-(4) with new distortion free constraint Eq.(40)
FF-BSS time(MDP)	Eqs.(25)-(26) [11]
FF-BSS freq.(1)	Eqs.(5)-(6) [6],[12],[15]
FF-BSS freq.(1+DF)	Eqs.(5)-(6) with new distortion free constraint Eq.(37)
FF-BSS freq.(2)	Eqs.(7),(6) [15]
FF-BSS freq.(2+DF)	Eqs.(7),(6) with new distortion free constraint Eq.(37)

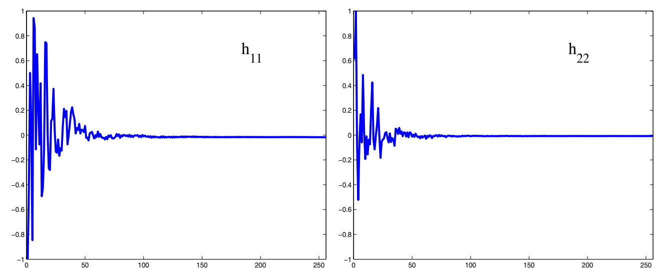

Fig. 3. Impulse responses of $H_{11}(z)$ and $H_{22}(z)$

VIII. Simulation and Discussion

A. Learning Methods and Their Abbreviation

In this paper, many kinds of learning methods will be compared. They are summarized in Table I.

B. Simulation Conditions

Two sources and two sensors are used. Impulse responses of direct paths $H_{11}(z)$ and $H_{22}(z)$ in the mixture are shown in Fig.3. The transfer function of the cross paths are related to the direct paths as $H_{ji}(z) = 0.9z^{-1}H_{ii}(z)$. Speeches and colored signals, generated through 2nd-order AR models, are used as sources. The FFT size is 256 points in the frequency domain training. FIR filters with 256 taps are used in the FF-BSS system, trained in the time domain and in the FB-BSS system. The initial guess of the separation block are $W_{11}(z) = W_{22}(z) = 1$ and $W_{kj}(z) = 0, k \neq j$, in the

TABLE II

COMPARISON AMONG EIGHT KINDS OF BSS SYSTEMS FOR SPEECH SIGNALS.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS time	12.2	5.56	0.25	-2.94	0.57	-3.82
FF-BSS time (DF)	8.33	4.33	-12.1	-16.2	-15.4	-19.9
FF-BSS time (MDP)	3.98	2.90	-10.3	-13.6	-8.24	-12.3
FF-BSS freq.(1)	7.37	2.61	-6.23	-8.67	-2.82	-2.95
FF-BSS freq. (1+DF)	9.68	6.38	-13.5	-18.1	-15.1	-18.3
FF-BSS freq.(2)	18.8	10.5	-10.9	-16.5	-11.9	-15.0
FF-BSS freq. (2+DF)	18.7	10.1	-25.8	-30.0	-18.4	-21.1
FB-BSS	14.1	9.24	-14.5	-17.9	-14.7	-17.3

TABLE III

DIFFERENCE BETWEEN $H_{ji}(z)$ AND $A_{ji}(z)$ FOR EIGHT KINDS OF BSS SYSTEMS FOR SPEECH SIGNALS.

Methods	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS time	2.44	-0.25	2.03	-0.63
FF-BSS time (DF)	-1.14	-5.00	-5.83	-8.12
FF-BSS time (MDP)	-9.36	-10.3	-8.53	-9.39
FF-BSS freq. (1)	3.80	-4.87	-0.42	-0.55
FF-BSS freq. (1+DF)	-0.21	6.51	-4.08	-5.11
FF-BSS freq. (2)	3.89	-5.57	-0.31	-1.00
FF-BSS freq. (2+DF)	3.99	-5.39	-0.03	-0.76
FB-BSS	0.38	-2.70	-0.59	-4.67

FF-BSS system, and $C_{12}(z) = C_{21}(z) = 0$ in the FB-BSS system.

Source separation is evaluated by the following two signal-to-interference ratios SIR_1 and SIR_2. Here, $S_1(z)$ and $S_2(z)$ are assumed to be separated in $Y_1(z)$ and $Y_2(z)$, respectively. However, it does not lose generality.

$$\sigma_{s1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})S_1(e^{j\omega})|^2 + |A_{22}(e^{j\omega})S_2(e^{j\omega})|^2)d\omega \quad (41)$$

$$\sigma_{i1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})S_2(e^{j\omega})|^2 + |A_{21}(e^{j\omega})S_1(e^{j\omega})|^2)d\omega \quad (42)$$

$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}}{\sigma_{i1}} \quad (43)$$

$$\sigma_{s2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})|^2 + |A_{22}(e^{j\omega})|^2)d\omega \quad (44)$$

$$\sigma_{i2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})|^2 + |A_{21}(e^{j\omega})|^2)d\omega \quad (45)$$

$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}}{\sigma_{i2}} \quad (46)$$

C. BSS Performance for Speech Signals

Evaluation measures are summarized in Table II. We set $i = j = k$ in Eqs.(11)-Eq.(18) with respect to signal distortion evaluations, because $S_1(z)$ and $S_2(z)$ are assumed to be separated in $Y_1(z)$ and $Y_2(z)$, respectively.

FF-BSS time has the worst performance in signal distortion SD_{ix}. FF-BSS time(MDP) can improve signal distortion. However, as discussed in Sec.V-C, due to the residual cross terms, the signal to interference ratio SIR_i is not good. FF-BSS time(DF) can improve SD_{ix} and also SIR_i. Compared to FF-BSS time, SIR_i is slightly reduced. However, FF-BSS time gains SIR_i by signal distortion. This means the frequency component in mainly the high frequency band is amplified, and signal distorted, at the same time, the signal power is increased resulting in relatively high SIR_i.

Regarding the frequency domain implementation, FF-BSS freq.(1) is not good in signal distortion due to the identity matrix in Eq.(5), which causes whitening as discussed in Sec.II-C. On the other hand, FF-BSS freq.(2) can improve signal distortion due to the weighting effects as discussed in Sec.V-B. However, by using the proposed distortion free constraint, FF-BSS freq.(1+DF) and FF-BSS freq.(2+DF) can improve more from each counterpart. SIR_i of FF-BSS freq.(2) and FF-BSS freq.(2+DF) are almost the same. However, SIR_i with the lower SD_{ix} has the higher accuracy.

On the contrary, FB-BSS demonstrates good performances in SIR_i and SD_{ix} due to a unique solution for complete separation and distortion free, as discussed in Sec.VI. However, its performances are slightly lower than those of FF-BSS freq.(2+DF).

The residual cross terms are more investigated here. As discussed in Sec.V-C, in the learning process of FF-BSS time(MDP), $A_{jj}(z)$ and $A_{ji}(z)$ tend to approach to $H_{jj}(z)$ and $H_{ji}(z)$, respectively. The former is evaluated in SD_{ix} and the latter is evaluated in SIR_i. Here, difference between $H_{ji}(z)$ and $A_{ji}(z)$ is evaluated in detail. Table III shows the numerical data, where SD_{ix} are calculated from Eqs.(11) through (18), setting the indices to be $i \neq j = k$. Small values mean they are similar to each other. Since FF-BSS time(MDF) has the smallest values, especially in SD_{1x}, that is in the frequency band, where the signal components are mainly located, then the theoretical discussion in Sec.V-C can be supported.

D. Stationary Colored Signals with Different Frequency Bands

2nd-order AR models are used for generating stationary colored signals. Spectra of $H_{11}(z)S_1(z)$ and $H_{22}(z)S_2(z)$ are shown in Fig.4. The frequency components are mainly located around 1/4($\times f_s$) and 3/4($\times f_s$), f_s is a sampling frequency. SIR_i and SD_{ix} for eight kinds of BSS systems are summarized in Table IV.

In the time domain implementation, FF-BSS time still is not good in signal distortion. The spectra of $A_{11}(z)S_1(z)$ and $A_{22}(z)S_2(z)$ are shown in Fig.5. They are whitened compared to those of Fig.4. Although it has relatively high SIR_i, these values have no meaning due to large signal distortion. FF-BSS time(MDF) shows good performance in SD_{ix}, however, it is not good in SIR_i. On the other hand, FF-BSS time(DF) can drastically improve SD_{ix} and also guarantee high SIR_i. From Fig.6, the spectra of $A_{ii}(z)S_i(z)$ are similar to that of $H_{ii}(z)S_i(z)$. This means signal distortion is small.

TABLE IV
COMPARISON AMONG EIGHT KINDS OF BSS SYSTEMS FOR STATIONARY COLORED SIGNALS WITH DIFFERENT FREQUENCY BANDS.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS time	9.49	7.07	-0.28	-3.11	-0.69	-4.99
FF-BSS time (DF)	8.05	4.07	-14.4	-16.6	-10.4	-13.2
FF-BSS time (MDP)	4.49	2.20	-15.7	-18.8	-13.7	-16.5
FF-BSS freq.(1)	4.93	4.32	-3.62	-5.03	-3.67	-4.28
FF-BSS freq. (1+DF)	7.63	4.06	-18.4	-20.8	-10.8	-14.1
FF-BSS freq.(2)	5.22	2.82	-12.2	-14.3	-7.42	-9.44
FF-BSS freq. (2+DF)	5.67	2.72	-13.7	-15.7	-8.27	-11.1
FB-BSS	16.5	7.19	-16.7	-18.6	-10.4	-13.6

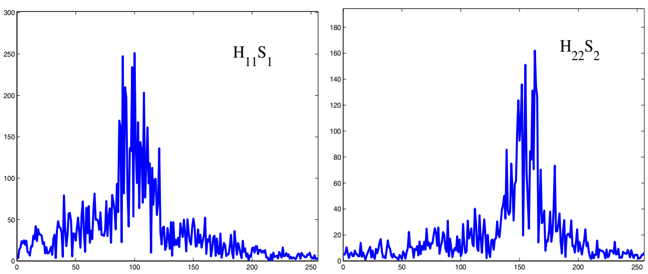

Fig. 4. Spectra of $H_{ii}(z)S_i(z)$ for colored signals.

In the frequency domain implementation, FF-BSS freq.(1) is also not good in signal distortion. FF-BSS freq.(1+DF) can drastically improve SD_{ix}, while keeping high SIR_i. The spectra of $A_{ii}(z)S_i(z)$ are shown in Fig.7. They are also similar to those of Fig.6. FF-BSS freq.(2) can provide relatively good performances. From Fig.8, the spectrum of $A_{22}(z)S_2(z)$ includes relatively large component around 1/4($\times f_s$), due to the weighting effects as discussed in Sec.V-B. In this case, since the frequency components are located in the different frequency bands, then the spectrum around the frequency band, where the other signal source is dominant, is amplified. FF-BSS freq.(2+DF) can slightly improve SD_{ix}, while keeping almost the same SIR_i as FF-BSS freq.(2).

Regarding the FF-BSS systems, FF-BSS time(DF) and FF-BSS freq.(1+DF) are better than the other methods. On the contrary, FB-BSS can achieve very good performances. Its SIR_i and SD_{ix} are higher than those of the FF-BSS systems.

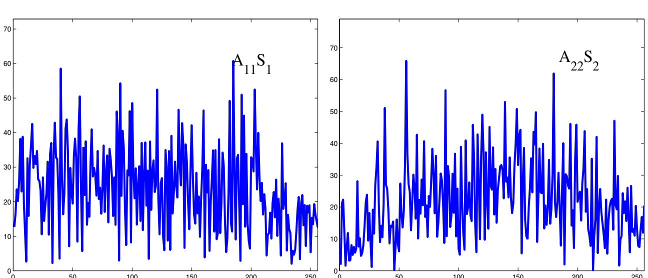

Fig. 5. Spectra of $A_{ii}(z)S_i(z)$ in FF-BSS time for colored signals.

E. Comparison between FF-BSS and FB-BSS

From the above comparison in SIR_i and SD_{ix}, the FB-BSS system can always provide good performance compared to the FF-BSS systems. Because the FB-BSS has a unique solution for complete source separation and distortion free. However, it requires some conditions for learning convergence. Especially, difference between the transmission time through the direct path and the cross path is very important.

In this section, effect of the transmission time difference on the performance of the FF-BSS systems and the FB-BSS system is analyzed in detail. We set the transmission time for the direct path to be zero, and for the cross path to be τ.

Figure 9 shows SIR_1 for FF-BSS time(DF) and FB-BSS by using the speech signals. The transmission time τ is changed from 0 to 125μsec. The sampling frequency is 8kHz, then a sampling period is 125μsec. Since the feedback $-C_{jk}(z)$ includes one sample delay, therefore, the time difference between $H_{11}(z)$ and $H_{21}(z)(-C_{12}(z))$ becomes $125 + \tau\mu sec$. This time difference is very important in the FB-BSS learning process. So, the horizontal axis indicates $125 + \tau\mu sec$.

From this figure, FF-BSS time(DF) is not affected by the time difference, because this condition is not required for learning convergence. On the other hand, FB-BSS is affected, and SIR_1 decreases as the time difference becomes small. The cross point is about 215μsec, which is $125 + 90\mu sec$. Thus, if the transmission time difference between $H_{ji}(z)$ and $H_{ii}(z)$ is less than 90μsec, then the FF-BSS systems are better than the FB-BSS system, and vice versa.

IX. CONCLUSIONS

In this paper, source separation and signal distortion have been theoretically analyzed for the FF-BSS systems implemented in both the time and frequency domains and the

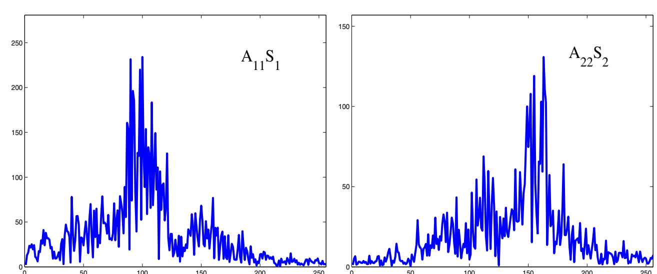

Fig. 6. Spectra of $A_{ii}(z)S_i(z)$ in FF-BSS time(DF) for colored signals.

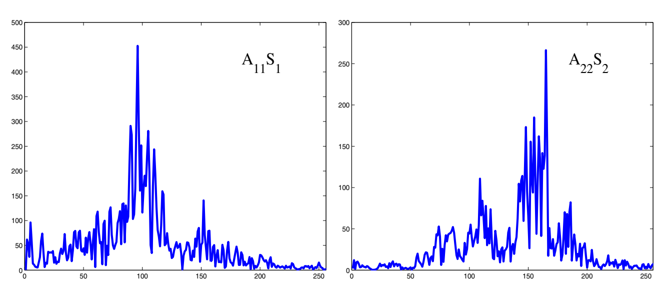

Fig. 7. Spectra of $A_{ii}(z)S_i(z)$ in FF-BSS freq.(1+DF) for colored signals.

FB-BSS system. The FF-BSS systems have some degree of freedom, and cause some signal distortion. The FB-BSS has a unique solution for complete separation and distortion free. Next, the condition for complete separation and distortion free has been derived for the FF-BSS systems. This condition has been applied to the learning algorithms. Computer simulations by using speech signals and stationary colored signals have been carried out for the conventional methods and the new learning algorithms employing the proposed distortion free constraint. The proposed method can drastically suppress signal distortion, while maintaining high separation performance. The FB-BSS system also has demonstrated good performances. The FF-BSS systems and the FB-BSS system have been compared based on the transmission time difference in the mixing process. As a result, location of the signal sources and the sensors are rather limited in the FB-BSS system.

APPENDIX

The constraint is extended to more than two channels. The condition for distortion-free source separation can be expressed as follows:

$$\begin{aligned} \boldsymbol{W}(z)\boldsymbol{H}(z) &= \boldsymbol{\Lambda(z)} & (47)\\ \boldsymbol{W}(z)(\boldsymbol{\Lambda}(z)+\boldsymbol{\Gamma}(z)) &= \boldsymbol{\Lambda}(z) & (48)\\ \boldsymbol{\Gamma}(z) &= (\boldsymbol{W}^{-1}(z)-\boldsymbol{I})\boldsymbol{\Lambda(z)} & (49)\\ \boldsymbol{\Lambda}(z) &= \mathrm{diag}[\boldsymbol{H}(z)] & (50)\\ \boldsymbol{\Gamma}(z) &= \boldsymbol{H}(z)-\boldsymbol{\Lambda}(z) & (51) \end{aligned}$$

The diagonal elements of Eq.(49) are given by:

$$\begin{aligned} \mathrm{diag}[\boldsymbol{\Gamma}(z)] &= \mathrm{diag}[(\boldsymbol{W}^{-1}(z)-\boldsymbol{I})\boldsymbol{\Lambda(z)}] & (52)\\ &= \mathrm{diag}[(\hat{\boldsymbol{W}}(z)-|\boldsymbol{W}(z)|\boldsymbol{I})\boldsymbol{\Lambda(z)}] & (53)\\ &= \boldsymbol{0} & (54) \end{aligned}$$

where, $\hat{\boldsymbol{W}}(z)$ is the adjugate matrix of $\boldsymbol{W}(z)$. The following relation is derived from the above equations.

$$\begin{aligned} diag[\hat{\boldsymbol{W}}(z)]-|\boldsymbol{W}(z)|\boldsymbol{I} &= \boldsymbol{0} & (55)\\ |\boldsymbol{W}_{ii}(z)|-|\boldsymbol{W}(z)| &= 0 & (56) \end{aligned}$$

where, $\boldsymbol{W}_{ii}(z)$ is the minor, removing the i-th row and i-th column. Similar to the case of two channels, $W_{ii}(z) = W_{jj}(z)$ is held, and the constraint can be obtained from the above equations. However, solving the above equation is computational expensive. Therefore, we consider an approximate calculation method. By deploying $W(z)$ to cofactors, the following relation can be obtained.

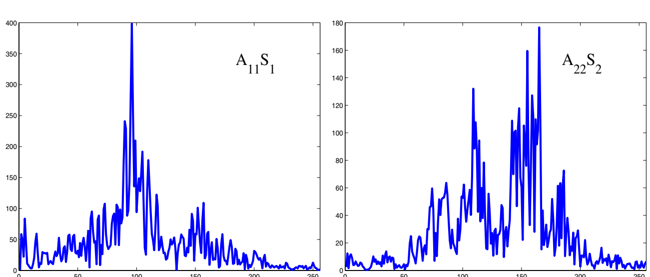

Fig. 8. Spectra of $A_{ii}(z)S_i(z)$ in FF-BSS freq.(2) for colored signals.

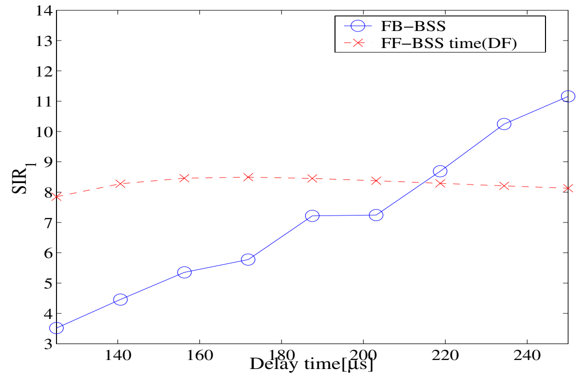

Fig. 9. Effects of transmission time difference on SIR_1 in FF-BSS time(DF) and FB-BSS.

$$w_{ii}(z) = 1+\boldsymbol{w}_{ix}^T(z)\boldsymbol{W}_{ii}^{-1}(z)\boldsymbol{w}_{xi}(z) \quad (57)$$

where, $\boldsymbol{w}_{xi}(z)$ and $\boldsymbol{w}_{ix}(z)$ are:

$$\begin{aligned} \boldsymbol{w}_{xi}(z) &= [w_{1i}, w_{2i}, \cdots, w_{yi}, \cdots, w_{Ni}]^T & (58)\\ \boldsymbol{w}_{ix}(z) &= [w_{i1}, w_{i2}, \cdots, w_{iy}, \cdots, w_{iN}]^T & (59) \end{aligned}$$

where $y \neq i$. $\boldsymbol{W}_{ii}^{-1}(z)$ includes , however, they are updated very slowly, because a small learning rate is usually used. This means w_{ii} can be treated as constant in $W_{ii}^{-1}(z)$. Thus, by using w_{ii} of Eq.(51) can approximately guarantee distortion free.

REFERENCES

[1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," Signal Processing, 24, pp.1-10, 1991.
[2] P.Comon, C.Jutten and J.Herault, "Blind separation of sources, Part II: Problems statement," Signal Processing, 24, pp.11-20, 1991.
[3] E.Sorouchyari, "Blind separation of sources, Part III: Stability analysis," Signal Processing, 24, pp.21-29, 1991.
[4] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," Signal Processing, vol.45, no.2, pp.209–229, March 1995.
[5] A.Cichocki, S.Amari, M.Adachi, W.Kasprzak, "Self-adaptive neural networks for blind separation of sources," Proc. ISCAS'96, Atlanta, pp.157-161, 1996.
[6] S.Amari, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation," Neural Networks, vol.10, no.8, pp.1345-1351, 1997.
[7] C.Simon, G.d' Urso, C.Vignat, Ph.Loubaton and C.Jutten, "On the convolutive mixture source separation by the decorrelation approach," IEEE Proc. ICASSP'98, Seattle, pp.IV-2109–2112, May 1998.
[8] S.Cruces and L.Castedo, "A Gauss-Newton methods for blind source separation of convolutive mixtures," IEEE Proc. ICASSP'98, Seattle, pp.IV2093–2096, May 1998.
[9] L.Parra and C.Spence, "Convolutive blind separation of nonstationary source," IEEE Trans. Speech Audio Processing, vol.8, pp.320-327, May 2000.
[10] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation," IEEE Proc. ICASSP'01, MULT-P3.3, May 2001.
[11] K.Matsuoka and S.Nakashima, "Minimal distortion principle for blind source separation," Proc. ICA2001, pp.722-727, 2001.
[12] I.Kopriva, Z.Devcic and H.Szu, "An adaptive short-time frequency domain algorithm for blind separation of nonstationary convolved mixtures," IEEE INNS Proc. IJCNN'01, pp.424-429, July 2001.

[13] K.Nakayama, A.Hirano and T.Sakai, ”An adaptive nonlinear function controlled by kurtosis for blind source separation,” IEEE INNS, Proc. IJCNN’2002, Honolulu, Hawaii, pp.1234-1239, May 2002.
[14] K.Nakayama, A.Hirano and A.Horita, ”A learning algorithm for convolutive blind source separation with transmission delay constraint,” IEEE INNS, Proc. IJCNN’2002, Honolulu, Hawaii, pp.1287-1292, May 2002.
[15] S.Araki, R.Mukai, S.Makino, T.Nishikawa, H.Saruwatari, ”The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech,” IEEE Trans. Speech and Audio Processing, vol.11, no.2, pp.109-116, March 2003.
[16] K.Nakayama, A.Hirano and A.Horita, ”A learning algorithm with adaptive exponential stepsize for blind source separation of convolutive mixtures with reverberations,” IEEE INNS, Proc. IJCNN’2003 July 2003.
[17] K. Nakayama, A.Hirano and T.Sakai, ”An adaptive nonlinear function controlled by estimated output pdf for blind source separation,” Proc. ICA2003, Nara, Japan, pp.427-432, April 2003.
[18] A. Cichocki and S. Amari, ”Adaptive blind signal and image processing” John Wiley & Sons, Ltd, April 2003.
[19] K.Nakayama, A.Hirano and Y.Dejima, ”Analysis of signal separation and distortion in feedforward blind source separation for convolutive mixture,” Proc. MWSCAS2004, Hiroshima, Japan, pp. III-207-III-210, July 2004.
[20] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, ”Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra” IJCNN2005, Montreal, Canada, August 2005.