A LEARNING ALGORITHM WITH DISTORTION FREE CONSTRAINT AND COMPARATIVE STUDY FOR FEEDFORWARD AND FEEDBACK BSS

Akihide HORITA, Kenji NAKAYAMA, Akihiro HIRANO, and Yasuhiro DEJIMA

Graduate School of Natural Science and Technology, Kanazawa Univ.
Kakuma-machi, Kanazawa, 920-1192, JAPAN
phone: + (81) (0)76-234-4896, fax: + (81) (0)76-234-4900,
email:horita@leo.ec.t.kanazawa-u.ac.jp, nakayama@t.kanazawa-u.ac.jp

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 signal distortion free. Next, the condition for complete separation and signal distortion free is derived for the FF-BSS systems. This condition is applied to the learning algorithms. Computer simulations by using speech 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.

1. 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],[2].

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 [4],[7],[9],[10].

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[5]. 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. Therefore, we have discussed an evaluation measure of signal distortion. Conditions for source separation and signal distortion free have been derived. Based on these conditions, convergence properties have been analyzed[11].

In this paper, a new learning algorithm with distortion free constraint for the FF-BSS systems is proposed. The proposed method is compared to the conventional FF-BSS and FB-BSS systems through computer simulations. Although the FB-BSS system has good performance in both source separation and distortion free, it has some limitation depending on transmission time difference. This limitation is also analyzed.

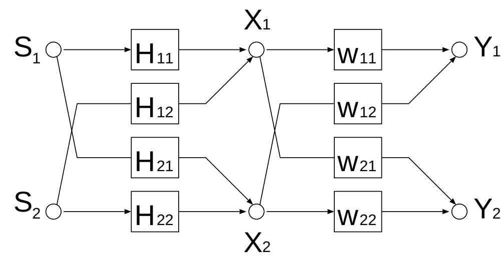

Figure 1: FF-BSS system with 2 signal sources and 2 sensors.

2. FF-BSS SYSTEMS FOR CONVOLUTIVE MIXTURE

2.1 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:

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

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

2.2 Learning Algorithm in Time Domain

The learning algorithm is derived following a natural gradient algorithm using a mutual information as a cost function [3].

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

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

The learning rate is represented by η.

2.3 Learning Algorithm in Frequency Domain

Filter coefficients in the separation block are trained according [3],[6],[8],

$$\boldsymbol{W}(r+1,m) = \boldsymbol{W}(r,m) + \eta[\boldsymbol{I} - \langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m)) \boldsymbol{Y}^H(r,m) \rangle] \boldsymbol{W}(r,m) \quad (5)$$

$$\Phi(\boldsymbol{Y}(r,m)) = \frac{1}{1+e^{-\boldsymbol{Y}^R(r,m)}} + \frac{j}{1+e^{-\boldsymbol{Y}^I(r,m)}} \quad (6)$$

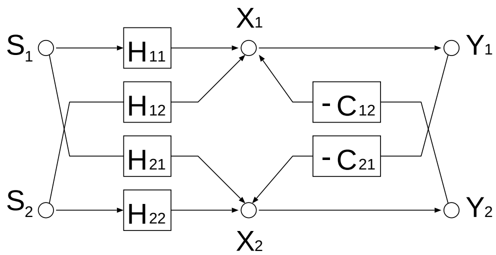

Figure 2: FB-BSS system with 2 signal sources and 2 sensors.

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 whiting has been proposed[8].

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

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

3. FB-BSS SYSTEMS FOR CONVOLUTIVE MIXTURE

3.1 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:

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

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

3.2 Learning Algorithm

In the FB-BSS system, a learning algorithm in the time domain is used [2]. The following learning algorithm has been derived by assuming several conditions [7],[9]. 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 $H_{ii}(z)$.

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

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

4. 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 [2].

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.

$$\begin{aligned}\sigma_{d1a} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega}) \\ &- A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (11)\end{aligned}$$

$$\begin{aligned}\sigma_{d1b} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})S_i(e^{j\omega})| \\ &- |A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (12)\end{aligned}$$

$$\sigma_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (13)$$

$$SD_{1x} = 10\log_{10}\frac{\sigma_{d1x}}{\sigma_1}, x=a,b \qquad (14)$$

$$\sigma_{d2a} = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega}) - A_{ki}(e^{j\omega})|^2 d\omega \qquad (15)$$

$$\sigma_{d2b} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})| - |A_{ki}(e^{j\omega})|)^2 d\omega \qquad (16)$$

$$\sigma_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})|^2 d\omega \qquad (17)$$

$$SD_{2x} = 10\log_{10}\frac{\sigma_{d2x}}{\sigma_2}, x=a,b \qquad (18)$$

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}, average power of $H_{ji}(z)S_i(z)$, $A_{ki}(z)S_i(z)$, $H_{ji}(z)$, and $A_{ki}(z)$ are normalized.

5. SOURCE SEPARATION AND SIGNAL DISTORTION IN FF-BSS SYSTEMS

5.1 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:

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \qquad (19)$$

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \qquad (20)$$

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

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

The above equations imply two conditions. First, complete source separation, that is the non-diagonal elements are all zero, as shown in Eqs.(20) and (21). Secondly, signal distortion free, that is the diagonal elements are $H_{ii}(z)$ as shown in Eqs.(19) and (22).

The conventional learning algorithm given by Eqs.(3)-(7) satisfies only Eqs.(20) and (21). Equations (19) and (22) are not satisfied. Therefore, signal distortion free is not guaranteed in general.

5.2 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.

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

A learning algorithm for reducing signal distortion has been proposed [5]. 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

$$\begin{aligned} & \boldsymbol{w}(n+1,l) = \boldsymbol{w}(n,l) \\ + \quad & \alpha \sum_{m=0}^{K_w-1} [\boldsymbol{I}\delta(n-m) - \langle \Phi(\boldsymbol{y}(n))\boldsymbol{y}^T(n-l+m)\rangle \\ - \quad & \beta(\boldsymbol{y}(n) - \boldsymbol{x}(n))\boldsymbol{y}^T(n-l+m)]\boldsymbol{w}(n,m) \end{aligned} \tag{23}$$

$$\varphi(\boldsymbol{y}(n)) = \frac{1-e^{-\boldsymbol{y}(n)}}{1+e^{-\boldsymbol{y}(n)}} \tag{24}$$

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.

6. SOURCE SEPARATION AND SIGNAL DISTORTION IN FB-BSS SYSTEMS

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

$$(1)\ C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \qquad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \tag{25}$$

$$(2)\ C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \qquad C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \tag{26}$$

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.(25), the output signals are given by

$$Y_1(z) = H_{11}(z)S_1(z) \tag{27}$$

$$Y_2(z) = H_{22}(z)S_2(z) \tag{28}$$

They are exactly the same as the criteria of the signal distortion discussed in Sec.4 Therefore, the FB-BSS systems have a unique solution, which satisfies both source separation as well as the signal distortion free simultaneously. Thus, in the FB-BSS systems, if complete signal separation is achieved, signal distortion free is also automatically satisfied.

7. DISTORTION FREE CONDITION AND ITS APPLICATION TO LEARNING ALGORITHM FOR FF-BSS SYSTEMS

7.1 FF-BSS Systems Trained in Frequency Domain

From the relations of Eqs.(20) and (21), $H_{ji}(z)$ are expressed.

$$H_{12}(z) = -\frac{W_{12}(z)}{W_{11}(z)}H_{22}(z) \tag{29}$$

$$H_{21}(z) = -\frac{W_{21}(z)}{W_{22}(z)}H_{11}(z) \tag{30}$$

By substituting the above equations into Eqs.(19) and (22), $H_{ji}(z)$ can be removed, and the following equations consisting only of $W_{kj}(z)$ can be obtained.

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z) \tag{31}$$

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{11}(z) \tag{32}$$

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

$$\begin{aligned} & W_{jj}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \\ & j = 1,2, k = 1,2, j \neq k \end{aligned} \tag{33}$$

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

$$W_{jj}(z) = \frac{1 \pm \sqrt{1+4W_{12}(z)W_{21}(z)}}{2}, j = 1,2 \tag{34}$$

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.(34) are taken into account to some extent, as shown in the following.

$$\begin{aligned} \tilde{W}_{jj}(r+1,m) &= (1-\alpha)W(r+1,m) \\ &+ \alpha\frac{1+\sqrt{1+4W_{12}(r,m)W_{21}(rm)}}{2} \end{aligned} \tag{35}$$

7.2 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.(34) is taken into account in the learning process. Equation (34) is rewritten as follows:

$$(2W_{jj}(z)-1)^2 = 1 + 4W_{12}(z)W_{21}(z) \tag{36}$$

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 approximate the relation of Eq.(36).

The condition for the distortion free source separation holds complete separation and signal 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.(36) is gradually

Table 1: Abbreviation for FF-BSS with Different Learning and Implementation.

FF-BSS time	Eqs.(3)-(4) [3]
FF-BSS time(DF)	Eqs.(3)-(4) with new distortion free constraint Eq.(38)
FF-BSS time(MDP)	Eqs.(23)-(24) [5]
FF-BSS freq.(1)	Eqs.(5)-(6) [3],[6],[8]
FF-BSS freq.(1+DF)	Eqs.(5)-(6) with new distortion free constraint Eq.(35)
FF-BSS freq.(2)	Eqs.(7),(6) [8]
FF-BSS freq.(2+DF)	Eqs.(7),(6) with new distortion free constraint Eq.(35)

imposed as the learning process progresses. The following learning algorithm is proposed.

$$w_{kj}(n+1,l) = w_{kj}(n) + \eta\{w_{kj}(n) - \sum_{o=0}^{K_w-1}\sum_{p=1}^{2}\phi(y_k(n))y_p(n-o+p)w_{kp}(n,o)\} \quad (37)$$

$$w_{jj}(N+1,l) = (1-\alpha)w_{jj}(n+1,l) + \alpha\tilde{w}_{jj}(n+1) \quad (38)$$

$\tilde{w}_{jj}(n+1)$ is determined so as to approximate the relation of Eq.(36). α is set to a small positive number.

8. SIMULATION AND DISCUSSION

8.1 Learning Methods and Their Abbreviation

In this paper, many kinds of learning methods will be compared. They are summarized in Table 1.

8.2 Simulation Conditions

Two sources and two sensors are used. The transfer function of the cross paths are related to the direct paths as $H_{ji}(z) = 0.9z^{-1}H_{ii}(z)$. Speeches 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 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.

$$\sigma_{s1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})S_1(e^{j\omega})|^2 + |A_{22}(e^{j\omega})S_2(e^{j\omega})|^2)d\omega \quad (39)$$

$$\sigma_{i1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})S_2(e^{j\omega})|^2 + |A_{21}(e^{j\omega})S_1(e^{j\omega})|^2)d\omega \quad (40)$$

$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}}{\sigma_{i1}} \quad (41)$$

$$\sigma_{s2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})|^2 + |A_{22}(e^{j\omega})|^2)d\omega \quad (42)$$

$$\sigma_{i2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})|^2 + |A_{21}(e^{j\omega})|^2)d\omega \quad (43)$$

$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}}{\sigma_{i2}} \quad (44)$$

Table 2: 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

8.3 BSS Performance for Speech Signals

Evaluation measures are summarized in Table 2. 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.5.3, 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.2.3. On the other hand, FF-BSS freq.(2) can improve signal distortion due to the weighting effects on the filter coefficient updation as discussed in Sec.5.2. 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 signal distortion free, as discussed in Sec.6. 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.5.3, 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 3 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.5.3 can be supported.

8.4 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

Table 3: 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

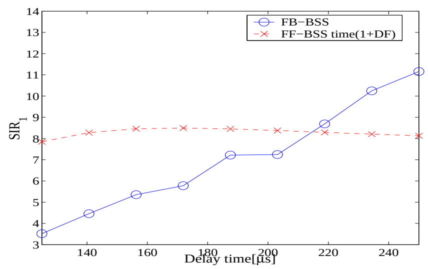

Figure 3: Effects of transmission time difference on SIR_1 in FF-BSS time(DF) and FB-BSS.

the FF-BSS systems. Because the FB-BSS has a unique solution for complete source separation and signal 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 and the FB-BSS systems 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 3 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 $125\mu sec$. The sampling frequency is 8kHz, then a sampling period is $125\mu sec$. 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 $215\mu sec$, which is $125 + 90\mu sec$. Thus, if the transmission time difference between $H_{ji}(z)$ and $H_{ii}(z)$ is less than $90\mu sec$, then the FF-BSS systems are better than the FB-BSS system, and vice versa.

9. APPLICATION TO MULTI-CHANEL BSS

The proposed method can be applied to more than two signal sources. The same constraint as Eq.(34) can be derived, which are expressed by using vectors and matrices. Although some approximations are required, performances of signal source separation and signal distortion free are not affected. Simulations of three signal sources have been already carried out and the performances are better than those of the conventional methods. Due to the page limitation, the details are omitted here.

10. 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 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 signal distortion free. Next, the condition for complete separation and signal 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 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.

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] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," Signal Processing, vol.45, no.2, pp.209–229, March 1995.

[3] 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.

[4] L.Parra and C.Spence, "Convolutive blind separation of nonstationary source," IEEE Trans. Speech Audio Processing, vol.8, pp.320-327, May 2000.

[5] K.Matsuoka and S.Nakashima, "Minimal distortion principle for blind source separation," Proc. ICA2001, pp.722-727, 2001.

[6] 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.

[7] 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, pp.1287-1292, May 2002.

[8] 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.

[9] 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.

[10] 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.

[11] 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," IJCNN'05, Montreal, Aug. 2005.