A Learning Algorithm with Distortion Free Constraint for Multi-Channel BSS Systems and Performance Analysis Based on Their Structures
畳み込み形多チャネルBSSにおける信号歪み抑制学習法及び回路形式に基づく考察

Akihide Horita† Kenji Nakayama† Akihiro Hirano†
堀田 明秀† 中山 謙二† 平野 晃宏†

†Division of Electrical Engineering and Computer Science
Graduate School of Natural Science and Technology, Kanazawa Univ.
金沢大学大学院 自然科学研究科 電子情報科学専攻
E-mail: horita@leo.ec.t.kanazawa-u.ac.jp, nakayama@t.kanazawa-u.ac.jp

ABSTRACT

Feed-Forward (FF-) and FeedBack (FB-) structures have been proposed for Blind Source Separation (BSS) systems. FF-BSS systems have some degree of freedom in the solution space, and signal distortion is likely to occur in convolutive mixtures. Previously, a condition for complete separation and distortion free has been derived for 2-channel FF-BSS system. This condition has been applied to the learning algorithms as a distortion free constraint in both the time and frequency domains.

In this paper, the condition is further extended to multiple channel FF-BSS systems. This condition requires a high computational complexity to be applied to the learning process as a constraint. An approximate constraint is proposed in order to relax the high computational load. In comparison with the original constraint, computer simulations have demonstrated that the approximation can obtain similar performances with respect to source separation as well as signal distortion using speech signals. Furthermore, the performances can be improved compared to the conventional methods. On the other hand, FB-BSS structure is hard to cause signal distortion. However, FB-BSS system requires a condition for the transmission time difference in the mixing process. FF-BSS systems and 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.

あらまし

ブラインドソースセパレーション(BSS)においてフィードフォワード(FF-)形とフィードバック(FB-)形の２つの回路形式が提案されている．FF-BSSでは自由度が存在し，信号歪みが生じる可能性が高い．そこでこれまで2チャンネルの畳み込み形FF-BSSに対して信号歪み抑制形学習法を提案してきた．

本稿では，畳み込み形多チャネルBSSに対して信号歪み抑制形学習法を提案する．これは完全な信号源分離と信号無歪みの条件から制約条件を導き，従来の学習法に組み合わせている．制約条件では近似式を導入している．シミュレーションにより，従来法に比べて信号歪みを大幅に抑制することが出来ることを確認した．一方，FB-BSSは本質的に信号歪みを発生しにくい構造であるが，信号源からセンサーまでの遅延時間に条件が課せられる．この観点から，FF-BSSとFB-BSSの分離性能について比較を行い，各々が有効に適用できる範囲を明らかにした．

1 Introduction

Signal processing, including noise cancellation, echo cancellation, equalization of transmission lines, estimation and restoration of signals has become a very important research area. However, often information with respect to the signals itself and their interference is insufficient. Furthermore, their mixing and transmission processes are not well known in advance. In these kinds of situations, blind source separation (BSS) technology using statistical properties of signal sources have become very important [1]-[3].

In many applications, mixing processes are convolutive mixtures. Therefore, separation processes require convolutive models. Various methods for separating sources in the time domain and the frequency domain have been proposed. Their separation performance is highly dependent on the signal sources and the transfer functions in the mixture [5],[6],[12].

BSS learning algorithms make the output signals statistically independent. However, this approach cannot always guarantee distortion free separation. Therefore, a new learning algorithm with a constraint based on the condition for complete source separation and distortion free for two channels has been proposed [14].

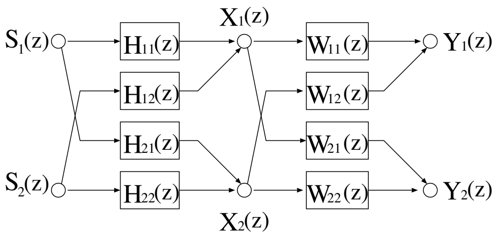

Figure 1: BSS system with 2 signal sources and 2 sensors.

In this paper, the learning algorithm is extended to more than two channels. Moreover, since the calculation is computationally expensive, an approximation is proposed. Furthermore, comparing to FB-BSS system, their characteristics are confirmed.

2 FF-BSS System for Convolutive Mixture

2.1 Network Structure and Equations

A block diagram of an FF-BSS system (2 signal sources and 2 sensors) is shown in Fig. 1. The mixing stage has a convolutive structure. The blocks $W_{kj}(z)$ consist of an FIR filter. The observations $\boldsymbol{x}(n)$ and the output signals $\boldsymbol{y}(n)$ are given by:

$$\boldsymbol{x}(n) = \sum_{l=0}^{K_h-1} \boldsymbol{h}(l)\boldsymbol{s}(n-l) \tag{1}$$

$$\boldsymbol{y}(n) = \sum_{l=0}^{K_w-1} \boldsymbol{w}(n,l)\boldsymbol{x}(n-l) \tag{2}$$

$$\boldsymbol{s}(n) = [s_1(n), \cdots, s_N(n)]^T \tag{3}$$

$$\boldsymbol{x}(n) = [x_1(n), \cdots, x_N(n)]^T \tag{4}$$

$$\boldsymbol{y}(n) = [y_1(n), \cdots, y_N(n)]^T \tag{5}$$

$$\boldsymbol{h}(l) = \begin{bmatrix} h_{11}(l) & \cdots & h_{1N}(l) \\ \vdots & \ddots & \vdots \\ h_{N1}(l) & \cdots & h_{NN}(l) \end{bmatrix} \tag{6}$$

$$\boldsymbol{w}(n,l) = \begin{bmatrix} w_{11}(n,l) & \cdots & w_{1N}(n,l) \\ \vdots & \ddots & \vdots \\ w_{N1}(n,l) & \cdots & w_{NN}(n,l) \end{bmatrix} \tag{7}$$

where $\boldsymbol{s}(n)$ is a signal source, $\boldsymbol{h}(l)$ is a mixing system and $\boldsymbol{w}(n,l)$ is a separation system. In the z-domain, the above equations can be expressed by:

$$\boldsymbol{X}(z) = \boldsymbol{H}(z)\boldsymbol{S}(z) \tag{8}$$

$$\boldsymbol{Y}(z) = \boldsymbol{W}(z)\boldsymbol{X}(z) \tag{9}$$

The relation between the signal sources and the outputs is defined by:

$$\boldsymbol{Y}(z) = \boldsymbol{W}(z)\boldsymbol{H}(z)\boldsymbol{S}(z) = \boldsymbol{A}(z)\boldsymbol{S}(z) \tag{10}$$

2.2 Learning Algorithm in Time Domain

Previously, a learning algorithm for separating sources based on a natural gradient method using mutual information as a cost function has been proposed [4]. This learning algorithm in the time domain is can be given by:

$$\begin{aligned} \boldsymbol{w}(n+1,l) &= \boldsymbol{w}(n,l) + \eta \sum_{q=0}^{K_w-1} [\boldsymbol{I}\delta(n-q) \\ &\quad -\langle \boldsymbol{\Phi}(\boldsymbol{y}(n))\boldsymbol{y}^T(n-l+q)\rangle]\boldsymbol{w}(n,q) \end{aligned} \tag{11}$$

$$\boldsymbol{\Phi}(\boldsymbol{y}(n)) = [\Phi(y_1(n)), \cdots, \Phi(y_N(n))]^T \tag{12}$$

$$\Phi(y_k(n)) = \frac{1-e^{-y_k(n)}}{1+e^{-y_k(n)}} \tag{13}$$

The learning rate is represented by η. $<>$ is an averaging operation. $\delta(n)$ is Dirac's delta function, where $\delta(0)=1$ and $\delta(n)=0 \ \ (n \neq 0)$.

2.3 Learning Algorithm in Frequency Domain

The same learning algorithm in the frequency domain using FFT, is defined as [4],[8],[10]:

$$\begin{aligned} \boldsymbol{W}(r{+}1,m) &= \boldsymbol{W}(r,m) + \eta[diag(\langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle) \\ &\quad -\langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) \end{aligned} \tag{14}$$

$$\boldsymbol{\Phi}(\boldsymbol{Y}(r,m)) = [\Phi(Y_1(r,m)), \cdots, \Phi(Y_N(r,m))]^T \tag{15}$$

$$\Phi(Y_k(r,m)) = \frac{1}{1+e^{-Y_k^R(r,m)}} + \frac{j}{1+e^{-Y_k^I(r,m)}} \tag{16}$$

The parameter r is the block number used in the FFT, and m indicates the frequency point within each block. $\boldsymbol{W}(r,m)$ is the weight matrix of the r-th FFT block and the m-th frequency point. $\boldsymbol{Y}(r,m)$ is the output of the r-th FFT block and the m-th frequency point. The function diag($\cdot$) is the diagonal matrix of $\cdot$. $Y_k^R(r,m)$ and $Y_k^I(r,m)$ represent the real part and the imaginary part, respectively.

3 FB-BSS System for Convolutive Mixture

3.1 Network Structure and Equations

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

The observations and the output signals are expressed as follows:

$$\boldsymbol{x}(n) = \sum_{l=0}^{K_h-1} \boldsymbol{h}(l)\boldsymbol{s}(n-l) \tag{17}$$

$$\boldsymbol{y}(n) = \boldsymbol{x} - \sum_{l=1}^{K_c-1} \boldsymbol{c}(n,l)\boldsymbol{y}(n-l) \tag{18}$$

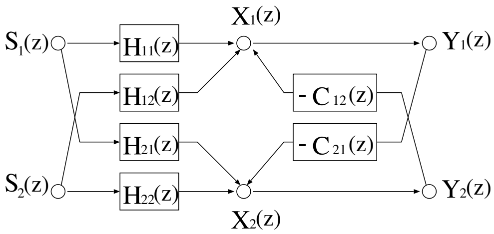

Figure 2: FB-BSS system with 2 signal sources and 2 sensors.

$$\boldsymbol{c}(n,l) = \begin{bmatrix} 0 & \cdots & c_{1N}(n,l) \\ \vdots & \ddots & \vdots \\ c_{N1}(n,l) & \cdots & 0 \end{bmatrix} \quad (19)$$

In the z-domain, the above equations can be expressed by:

$$\boldsymbol{X}(z) = \boldsymbol{H}(z)\boldsymbol{S}(z) \quad (20)$$

$$\boldsymbol{Y}(z) = \boldsymbol{X}(z) - \boldsymbol{C}(z)\boldsymbol{Y}(z) \quad (21)$$

The relation between the signal sources and the outputs is defined by:

$$\boldsymbol{Y}(z) = (\boldsymbol{I} + \boldsymbol{C}(z))^{-1}\boldsymbol{H}(z)\boldsymbol{S}(z) = \boldsymbol{A}(z)\boldsymbol{S}(z) \quad (22)$$

3.2 Learning Algorithm

The following learning algorithm has been derived by assuming several conditions [9],[11]. The signal sources $S_1(z)$ and $S_2(z)$ are located close to the sensors of $X_1(z)$ and $X_2(z)$, respectively. 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 $H_{ii}(z)$.

$$c_{jk}(n+1,l) = c_{jk}(n,l) + \eta f(y_j(n))g(y_k(n-l)) \quad (23)$$

$f(y_j(n))$ and $g(y_k(n-l))$ 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.7.4.

4 Criterion for Signal Distortion

In this paper, only signal distortion caused by the BSS systems itself is considered [2], [13]. Therefore, four kinds of measures (SD_{ix} $i = 1,2$ $x = a,b$), as shown below, are applied:

$$SD_{1x} = 10\log_{10}\frac{\sigma_{d1x}}{\sigma_1}, x = a,b \quad (24)$$

$$SD_{2x} = 10\log_{10}\frac{\sigma_{d2x}}{\sigma_2}, x = a,b \quad (25)$$

$$\sigma_{d1a} = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega}) - A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (26)$$

$$\sigma_{d1b} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})S_i(e^{j\omega})| - |A_{ki}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \quad (27)$$

$$\sigma_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (28)$$

$$\sigma_{d2a} = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega}) - A_{ki}(e^{j\omega})|^2 d\omega \quad (29)$$

$$\sigma_{d2b} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ji}(e^{j\omega})| - |A_{ki}(e^{j\omega})|)^2 d\omega \quad (30)$$

$$\sigma_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ji}(e^{j\omega})|^2 d\omega \quad (31)$$

Since FF-BSS systems are unable to 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 average powers of $H_{ji}(z)S_i(z)$, $A_{ki}(z)S_i(z)$, $H_{ji}(z)$, and $A_{ki}(z)$ are normalized.

5 Learning Algorithm with Distortion Free Constraint for 2-Channels

FF-BSS systems have some degree of freedom in the solution space, and signal distortion is likely to occur in convolutive mixtures. A constraint for complete separation and distortion free has been derived for 2-channel FF-BSS systems [14]. The condition is given by:

$$W_{jj}{}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \quad (32)$$
$$j = 1,2, k = 1,2, j \neq k$$

This condition has been applied to the learning algorithm as a distortion free constraint. The learning algorithm with the distortion free constraint trained in time domain is given by:

Step 1 : Update $w_{jj}(n,l)$ and $w_{jk}(n,l)$ following Eqs. (11) through (13) resulting in $w_{jj}(n+1,l)$ and $w_{jk}(n+1,l)$.

Step 2 : Modify $w_{jj}(n,l)$ following:

$$w_{jj}(n+1,l) = (1-\alpha)w_{jj}(n+1,l) + \alpha\tilde{w}_{jj}(n,l) \quad (0 < \alpha \leq 1) \quad (33)$$

$w_{jj}(n+1,l)$ in the left-hand side is the modified version and $\tilde{w}_{jj}(n,l)$ is determined by Eq. (32).

The learning algorithm with the distortion free constraint trained in frequency domain is given by:

Step 1 : Update $W_{jj}(r,m)$ and $W_{jk}(r,m)$ following Eq. (14) through (16), resulting in $W_{jj}(r+1,m)$ and $W_{jk}(r+1,m)$.

Step 2 : Modify $W_{jj}(r+1,m)$ following:

$$W_{jj}(r+1,m) = (1-\alpha)W_{jj}(r+1,m) + \alpha\frac{1+\sqrt{1+4W_{12}(r,m)W_{21}(r,m)}}{2} \quad (34)$$

$$(0 < \alpha \leq 1)$$

$W_{jj}(r+1,m)$ in the left-hand side is the modified version.

6 Distortion Free Learning Algorithm for Multi-Channel BSS

6.1 Distortion Free Constraint

The constraint described in Sec. 5 is extended to more than two channels. The condition for distortion free source separation can be expressed as follows:

$$\boldsymbol{W}(z)\boldsymbol{H}(z) = \boldsymbol{\Lambda(z)} \quad (35)$$

$$\boldsymbol{\Lambda}(z) = \text{diag}[\boldsymbol{H}(z)] \quad (36)$$

Let $\boldsymbol{\Gamma}(z)$ be a matrix having the non-diagonal elements of $\boldsymbol{H}(z)$:

$$\boldsymbol{\Gamma}(z) = \boldsymbol{H}(z) - \boldsymbol{\Lambda}(z) \quad (37)$$

Substituting $\boldsymbol{H}(z)$ in Eq. (37) into Eq. (35):

$$\boldsymbol{W}(z)(\boldsymbol{\Lambda}(z) + \boldsymbol{\Gamma}(z)) = \boldsymbol{\Lambda}(z) \quad (38)$$

Solving this equation for $\boldsymbol{\Gamma}(z)$:

$$\boldsymbol{\Gamma}(z) = \boldsymbol{W}^{-1}(z)(\boldsymbol{I} - \boldsymbol{W}(z))\boldsymbol{\Lambda(z)} \quad (39)$$

$$= (\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda(z)} \quad (40)$$

From Eq. (37), it follows:

$$\text{diag}[\boldsymbol{\Gamma}(z)] = \text{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda(z)}] = \boldsymbol{0} \quad (41)$$

Since $\boldsymbol{\Lambda(z)}$ is the diagonal matrix, the above equation can be rewritten as:

$$\text{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})] = \boldsymbol{0} \quad (42)$$

This condition holds, when the diagonal elements of $\boldsymbol{W}^{-1}(z)$ are 1. The inverse matrix is generally expressed by:

$$\boldsymbol{W}^{-1}(z) = \frac{\text{adj}\,\boldsymbol{W}(z)}{\det \boldsymbol{W}(z)} \quad (43)$$

$\text{adj}\,\boldsymbol{W}(z)$ is the adjugate matrix of $\boldsymbol{W}(z)$. Since, the diagonal elements of $\boldsymbol{W}^{-1}(z)$ equal 1, then:

$$\text{diag}[\boldsymbol{W}^{-1}(z)] = \frac{\text{diag}[\text{adj}\,\boldsymbol{W}(z)]}{\det \boldsymbol{W}(z)} = \boldsymbol{I} \quad (44)$$

Let the j-th diagonal element of $\text{adj}\,\boldsymbol{W}(z)$ be $\hat{W}_{jj}(z)$:

$$\frac{\hat{W}_{jj}(z)}{\det \boldsymbol{W}(z)} = 1 \quad (45)$$

$$\hat{W}_{jj}(z) = \det \boldsymbol{W}(z) \quad (46)$$

$\hat{W}_{jj}(z)$ is also a cofactor of $\boldsymbol{W}(z)$.
This is a general expression for the distortion free constraint for multi-channel BSS systems.

6.2 Learning Algorithm

Solving Eq. (46) is computationally expensive. Therefore, we introduce an approximation formula for this calculation.

$\det \boldsymbol{W}(z)$ is further expressed by:

$$\det \boldsymbol{W}(z) = \sum_{k=1}^{N} W_{jk}(z)(-1)^{j+k} \det \boldsymbol{M}_{jk}(z) \quad (47)$$

$\boldsymbol{M}_{jk}(z)$ is an $(N-1) \times (N-1)$ minor matrix. $\hat{W}_{jj}(z)$ is also:

$$\hat{W}_{jj}(z) = (-1)^{2j} \det \boldsymbol{M}_{jj}(z) = \det \boldsymbol{M}_{jj}(z) \quad (48)$$

From Eqs.(46), (47) and (48), we obtain:

$$\det \boldsymbol{M}_{jj}(z) = \sum_{k=1}^{N} W_{jk}(z)(-1)^{j+k} \det \boldsymbol{M}_{jk}(z) \quad (49)$$

In this equation, $W_{jj}(z)$ is extracted:

$$\det \boldsymbol{M}_{jj}(z)(1 - W_{jj}(z)) = \sum_{\substack{k=1 \\ \neq j}}^{N} W_{jk}(z)(-1)^{j+k} \det \boldsymbol{M}_{jk}(z) \quad (50)$$

The right hand side is further rewritten by:

$$-\boldsymbol{w}_{rj}^T(z)\,\text{adj}\,\boldsymbol{M}_{jj}(z)\boldsymbol{w}_{cj}(z) \quad (51)$$

where $\boldsymbol{w}_{cj}(z)$ and $\boldsymbol{w}_{rj}(z)$ are:

$$\boldsymbol{w}_{cj}(z) = [W_{1j}(z), W_{2j}(z), \cdots, W_{Nj}(z)]^T \quad (52)$$

$$\boldsymbol{w}_{rj}(z) = [W_{j1}(z), W_{j2}(z), \cdots, W_{jN}(z)]^T \quad (53)$$

$\boldsymbol{w}_{cj}(z)$ and $\boldsymbol{w}_{rj}(z)$ do not include $W_{jj}(z)$.

Finally,

$$W_{jj}(z) = 1 + \boldsymbol{w}_{rj}^T(z)\frac{\text{adj}\,\boldsymbol{M}_{jj}(z)}{\det \boldsymbol{M}_{jj}(z)}\boldsymbol{w}_{cj}(z) \quad (54)$$

$$= 1 + \boldsymbol{w}_{rj}^T(z)\boldsymbol{M}_{jj}^{-1}(z)\boldsymbol{w}_{cj}(z) \quad (55)$$

Equation (55) is not an explicit solution, because the matrix $\boldsymbol{M}_{jj}^{-1}(z)$ includes $W_{kk}(z)$ $(j \neq k)$. However, it can be expected that the update changes of $W_{kk}(z)$ are very little, because usually a small learning rate is applied. Therefore, in order to ensure high distortion free source separation regarding more than two channels, Eq. (55) can be used by treating the $W_{kk}(z)$ from $\boldsymbol{M}_{jj}^{-1}(z)$ as constants.

The learning algorithm with the distortion free constraint trained in frequency domain is given by:

Step 1 : Update $W_{jj}(r,m)$ and $W_{jk}(r,m)$ following Eq. (14) through (16), resulting in $W_{jj}(r+1,m)$ and $W_{jk}(r+1,m)$.

Step 2 : Modify $W_{jj}(r+1,m)$ following:

$$W_{jj}(r+1,m) = (1-\alpha)W_{jj}(r+1,m) + \alpha\tilde{W}_{jj}(r,m)$$

$$(0 < \alpha \leq 1) \quad (56)$$

Table 1: Abbreviations of the applied learning algorithms.

FF-BSS TIME	Eqs. (11)-(13) [4]
FF-BSS TIME (DF)	Eqs. (11)-(13) with the distortion free constraint Eq. (33)
FF-BSS TIME (ADF)	Eqs. (11)-(13) with the distortion free constraint applying Eq. (57)
TIME (MDP)	A conventional distortion free learning method [7]
FF-BSS FREQ	Eqs. (14)-(16) [10]
FF-BSS FREQ (ADF)	Eqs. (14)-(16) with the distortion free constraint applying Eq. (56)
FB-BSS	Eq. (23) [9], [11]

$W_{jj}(r+1,m)$ in the left-hand side is the modified version and $\tilde{W}_{jj}(r,m)$ is determined by Eq. (55).

Regarding the learning algorithm trained in time domain, since it is difficult to calculate Eq. (55) in the time domain, the constraint is calculated in the frequency domain using the separation system which transforms $w_{jk}(n,l)$ into frequency domain.

Step 1 : Update $w_{jj}(n,l)$ and $w_{jk}(n,l)$ following Eqs. (11) through (13) resulting in $w_{jj}(n+1,l)$ and $w_{jk}(n+1,l)$.

Step 2 : Modify $w_{jj}(n,l)$ following:

$$w_{jj}(n+1,l) = (1-\alpha)w_{jj}(n+1,l) + \alpha\tilde{w}_{jj}(n,l) \quad (0 < \alpha \leq 1) \qquad (57)$$

$w_{jj}(n+1,l)$ in the left-hand side is the modified version and $\tilde{w}_{jj}(n,l)$ is Eq. (55) calculated in frequency domain.

7 Simulations and Discussions

7.1 Learning Methods and Their Abbreviations

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

7.2 Simulation Conditions

Mixture systems simulating actual acoustic spaces are applied. Speeches are used as sources. The FFT size is set to 256 points for training in the frequency domain. FIR filters with 256 taps are used for training in the time domain. The initial guess for the separation blocks are $W_{jj}(z) = 1$ and $W_{kj}(z) = 0, k \neq j$ for FF-BSS systems and $C_{kj}(z) = 0$ for FB-BSS system.

Source separation is evaluated by the following two signal-to-interference ratios SIR_1 and SIR_2. Here, the sources $S_i(z)$ are assumed to be separated at the outputs $Y_i(z)$. However, this does not lose generality.

$$\sigma_{s1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{N}|A_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (58)$$

Table 2: Comparison among BSS systems for speech signals.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
TIME	13.3	7.58	0.50	-2.48	-0.71	-4.65
TIME (ADF)	8.00	5.24	-14.0	-17.9	-17.1	-20.5
TIME (MDP)	6.25	4.36	-8.02	-11.2	-10.2	-14.1
FREQ	17.0	9.58	-14.1	-21.1	-14.6	-17.2
FREQ (ADF)	17.1	9.12	-26.7	-34.7	-19.9	-21.7
FB-BSS	13.6	6.82	-15.4	-18.9	-17.0	-19.7

$$\sigma_{i1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}|A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (59)$$

$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}}{\sigma_{i1}} \qquad (60)$$

$$\sigma_{s2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{N}|A_{ii}(e^{j\omega})|^2 d\omega \qquad (61)$$

$$\sigma_{i2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}|A_{ki}(e^{j\omega})|^2 d\omega \qquad (62)$$

$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}}{\sigma_{i2}} \qquad (63)$$

7.3 Source Separation and Signal Distortion

Simulation is performed for three sources and three sensors. Evaluation measures are summarized in Table 2. We have set $i = j = k$ in Eqs. (24)-(31) with respect to the signal distortion evaluations, because $S_1(z)$,$S_2(z)$ and $S_3(z)$ are assumed to be separated at $Y_1(z)$,$Y_2(z)$ and $Y_3(z)$, respectively.

The conventional learning algorithm FF-BSS TIME performs worst regarding the signal distortion measures SD_{ix}. TIME (MDP) can improve signal distortion. However, the signal-to-interference ratios SIR_i are unsatisfactory. TIME (ADF) can improve SD_{ix} as well as SIR_i. Compared to TIME, the evaluation values for SIR_i are slightly lower. However, in FF-BSS TIME, signal distortion is caused by the amplification of high frequency bands in an attempt to make the signal sources statistically independent. Consequently, this amplification contributes to higher, but somewhat blurred SIR_i values.

FF-BSS FREQ is better than the time domain implementations regarding source separation. Regarding signal distortion, FF-BSS FREQ is slightly worse than the learning algorithm trained in the time domain incorporating a distortion free constraints. On the other hand, by using the proposed distortion free constraint, FF-BSS FREQ (ADF) can drastically improve from FF-BSS FREQ regarding signal distortion, while maintaining high signal separation performances.

On the other hand, FB-BSS system provide good performances regarding source separation SIR_i as well as signal distortion SD_{ix} compared to the FF-BSS sys-

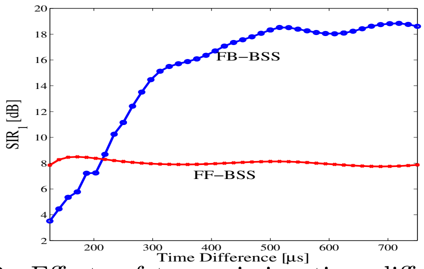

Figure 3: Effects of transmission time difference on SIR_1 in FF-BSS TIME (DF) and FB-BSS.

tems trained in the time domain.

7.4 Effect of Transmission Time Difference

Simulation is performed for two sources and two sensors. 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 trained in time domain. 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 3 shows SIR_1 for FF-BSS TIME (DF) and FB-BSS by using the speech signals. 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.

8 Conclusion

The constraint proposed for 2-channel FF-BSS systems, is extended to multiple channel FF-BSS systems. Furthermore, this constraint is approximated in order to relax the computational load. In comparison with the original constraint, computer simulations have demonstrated that the approximation can obtain a similar performance with respect to source separation as well as signal distortion using speech signals. On the other hand, FB-BSS system also has demonstrated good performances. FF-BSS systems and 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 FB-BSS system.

A part of this research has been supported by Grant-in-Aid for Scientific Research ((C)17560335).

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] 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.

[4] 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.

[5] 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.

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

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] A. Horita, K. Nakayama, A. Hirano and Y. Dejima,"A distortion free learning algorithm for feedforward BSS and its comparative study with feedback BSS", IEEE&INNS, Proc., IJCNN2006, Vancouver, pp.7642-7649, July 2006.

[15] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "A learning algorithm with distortion free constraint and comparative study for feedforward and feedback BSS", Proc., EUSIPCO2006, Florence, Italy, July 2006.