A DISTORTION FREE LEARNING ALGORITHM FOR MULTI-CHANNEL CONVOLUTIVE BLIND SOURCE SEPARATION

Akihide HORITA, Kenji NAKAYAMA, and Akihiro HIRANO

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
web: http://leo.ec.t.kanazawa-u.ac.jp/˜nakayama/

ABSTRACT

In this paper, source separation and signal distortion are theoretically analyzed in multi-channel convolutive blind source separation (BSS) systems implemented in both the time and the frequency domains. Feedforward (FF-) BSS systems have some degrees of freedom in the solution space. Therefore, signal distortion is likely to occur. A condition for complete separation and distortion free is proposed for multi-channel convolutive FF-BSS systems. This condition is incorporated in learning algorithms as a distortion free constraint. The condition for weights in a separation block cannot be expressed in an explicit form. Therefore, an approximation expression is introduced. Computer simulations using speech signals and stationary colored signals have been carried out for conventional methods and the new learning algorithms. The proposed method can drastically suppress signal distortion, while maintaining a high source separation performance.

1. INTRODUCTION

In many applications of Blind Source Separation (BSS) systems, the mixing processes are convolutive mixtures. Therefore, separation processes require convolutive models. Various methods for separating sources in the time domain and in the frequency domain have been proposed [1],[2],[4]. BSS learning algorithms make the output signals to be statistically independent. However, this approach does not always guarantee distortion free separation for Feedforward (FF-) BSS systems. Several methods have been proposed for suppressing signal distortion. One of them includes the distance between the observed signals and the separated signals in the cost function [3]. The other method suppresses signal distortion by multiplying the inverse matrix of the separation block [5]. Furthermore, we have proposed a distortion free learning algorithm for 2-channel FF-BSS systems [7],[8].

In this paper, the proposed method for 2-channel FF-BSS systems is extended to general multi-channel FF-BSS systems. First, evaluation measures for signal distortion are discussed. Secondly, conditions for complete source separation and distortion free are derived for multi-channel convolutive FF-BSS systems. Based on these conditions, an extended learning algorithm with a distortion free constraint is proposed. Computer simulation will be demonstrated by using speech signals and stationary colored signals in order to confirm usefulness of the proposed method.

2. BSS SYSTEMS FOR CONVOLUTIVE MIXTURE

2.1 Network Structure and Equations

A block diagram of a feedforward (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.

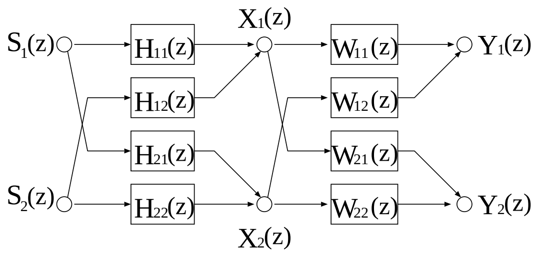

Figure 1: BSS system with 2 signal sources and 2 sensors.

In the z-domain, relations among the source signals, the observations and the output signals are expressed by:

$$\boldsymbol{X}(z) = \boldsymbol{H}(z)\boldsymbol{S}(z) \quad (1)$$
$$\boldsymbol{Y}(z) = \boldsymbol{W}(z)\boldsymbol{X}(z) \quad (2)$$
$$\boldsymbol{Y}(z) = \boldsymbol{W}(z)\boldsymbol{H}(z)\boldsymbol{S}(z) = \boldsymbol{A}(z)\boldsymbol{S}(z) \quad (3)$$

2.2 Learning Algorithm in Time Domain

The conventional learning algorithm [2] is applied to our method.

$$\boldsymbol{w}(n+1,l) = \boldsymbol{w}(n,l) + \eta \sum_{q=0}^{K_w-1} [\boldsymbol{I}\delta(n-q) - < \boldsymbol{\Phi}(\boldsymbol{y}(n))\boldsymbol{y}^T(n-l+q) >]\boldsymbol{w}(n,q) \quad (4)$$
$$\boldsymbol{\Phi}(\boldsymbol{y}(n)) = [\Phi(y_1(n)), \cdots, \Phi(y_N(n))]^T \quad (5)$$
$$\Phi(y_k(n)) = \frac{1-e^{-y_k(n)}}{1+e^{-y_k(n)}} \quad (6)$$

l is the tap number in the FIR filters, η is a learning rate, $<>$ is an averaging operation and $\delta(n)$ is the Dirac's delta function, where $\delta(0)=1$ and $\delta(n)=0, n \neq 0$.

2.3 Learning Algorithm in Frequency Domain

The learning algorithm in the frequency domain has been proposed in [2],[4],[6]:

$$\begin{aligned} \boldsymbol{W}(r+1,m) &= \boldsymbol{W}(r,m) \\ &\quad +\eta[\boldsymbol{I} - \langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) \quad (7) \\ \boldsymbol{\Phi}(\boldsymbol{Y}(r,m)) &= [\Phi(Y_1(r,m)),\cdots,\Phi(Y_N(r,m))]^T (8) \\ \Phi(Y_k(r,m)) &= \frac{1}{1+e^{-Y_k^R(r,m)}} + \frac{j}{1+e^{-Y_k^I(r,m)}} \quad (9) \end{aligned}$$

r is the frame number used in the FFT and m indicates the frequency point in each frame. $Y_k^R(r,m)$ and $Y_k^I(r,m)$ represent the real part and the imaginary part, respectively.

It has been reported that the learning algorithm given in Eq.(7) transforms the outputs into white signals. A learning algorithm that avoids whitening has been proposed in [6]:

$$\begin{aligned} \boldsymbol{W}(r+1,m) &= \boldsymbol{W}(r,m) \\ &+ \eta[\mathrm{diag}(\langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle) \\ &- \langle \boldsymbol{\Phi}(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) (10) \end{aligned}$$

In this paper, the learning algorithms using Eqs.(7) and (10) will be referred to as FREQ(1) and FREQ(2), respectively.

3. ANALYSIS OF SOURCE SEPARATION AND SIGNAL DISTORTION

3.1 Criterion for Signal Distortion

Learning algorithms applied in BSS systems make the output signals to be statistically independent, and an estimation of the mixing process is not taken into account. In this paper, we aim at suppressing signal distortion caused in the separation block. Therefore, the signal sources included in the observed signals $H_{ii}(z)S_i(z)$ or $H_{ji}(z)S_i(z)(j \neq i)$ are taken into account as criteria for signal distortion [7],[8].

3.2 Conditions for Source Separation and Distortion Free

For simplicity, a BSS system with two sources and two sensors, as shown in Fig.1, is used. Furthermore, the sources $S_i(z)$ are assumed to be separated at the outputs $Y_i(z)$. This does not lose generality. Considering the criterion of signal distortion as defined in Sec.3.1, the condition for source separation and distortion free can be expressed as follows:

Source separation : Non diagonal elements of $\boldsymbol{A}(z)$ are zero.

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (11)$$
$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (12)$$

Distortion free : Diagonal elements of $\boldsymbol{A}(z)$ are $H_{ii}(z)$.

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (13)$$
$$W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (14)$$

3.3 Signal Distortion

The conventional learning algorithm given by Eqs.(4)-(10) can satisfy Eqs.(11) and (12). However, Eqs.(13) and (14) are not guaranteed to be satisfied. Therefore, by applying this algorithm, signal distortion may occur.

Since the number of the equations for source separation is two, while the number of variables is four, there exist some degrees of freedom. This means the output signal spectra can be changed so as to be statistically independent to each other. Therefore, it is very likely that signal distortion will occur. This analysis is valid for BSS systems trained in both the time and the frequency domains [7],[8].

4. A CONSTRAINT OF SOURCE SEPARATION AND DISTORTION FREE FOR MULTI-CHANNEL FF-BSS

4.1 Derivation of A Constraint

The conditions for the source separation and the distortion free given by Eqs.(11) through (14), can be extended to a multi-channel case by expressing in a matrix form as follows:

$$\boldsymbol{W}(z)\boldsymbol{H}(z) = \boldsymbol{\Lambda}(z) \quad (15)$$
$$\boldsymbol{\Lambda}(z) = \mathrm{diag}[\boldsymbol{H}(z)] \quad (16)$$

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

$$\boldsymbol{\Gamma}(z) = \boldsymbol{H}(z) - \boldsymbol{\Lambda}(z) \quad (17)$$

Substituting $\boldsymbol{H}(z)$ in Eq.(17) into Eq.(15):

$$\boldsymbol{W}(z)(\boldsymbol{\Lambda}(z) + \boldsymbol{\Gamma}(z)) = \boldsymbol{\Lambda}(z) \quad (18)$$

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

$$\begin{aligned} \boldsymbol{\Gamma}(z) &= \boldsymbol{W}^{-1}(z)(\boldsymbol{I} - \boldsymbol{W}(z))\boldsymbol{\Lambda}(z) \quad (19) \\ &= (\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda}(z) \quad (20) \end{aligned}$$

From Eq.(17), it follows:

$$\mathrm{diag}[\boldsymbol{\Gamma}(z)] = \mathrm{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda}(z)] = \boldsymbol{0} \quad (21)$$

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

$$\mathrm{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})] = \boldsymbol{0} \quad (22)$$

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

$$\boldsymbol{W}^{-1}(z) = \frac{\mathrm{adj}\boldsymbol{W}(z)}{\det \boldsymbol{W}(z)} \quad (23)$$

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

$$\mathrm{diag}[\boldsymbol{W}^{-1}(z)] = \frac{\mathrm{diag}[\mathrm{adj}\boldsymbol{W}(z)]}{\det \boldsymbol{W}(z)} = \boldsymbol{I} \quad (24)$$

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

$$\frac{\hat{W}_{jj}(z)}{\det \boldsymbol{W}(z)} = 1 \quad (25)$$
$$\hat{W}_{jj}(z) = \det \boldsymbol{W}(z) \quad (26)$$

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

4.2 Approximation Formula for Distortion Free Constraint

Solving Eq.(26) is computationally expensive. Therefore, we introduce an approximation formula for this calculation.

Equation (26) can be rewritten as follows:

$$W_{jj}(z) = 1 + \boldsymbol{w}_{row,j}^T(z)\boldsymbol{M}_{jj}^{-1}(z)\boldsymbol{w}_{col,j}(z) \quad (j = 1, \cdots, N) \qquad (27)$$

Derivation of the above equation is given in Appendix A. $\boldsymbol{M}_{jj}(z)$ is an $(N-1)\times(N-1)$ minor matrix, i.e. removing the j-th row and the j-th column from $\boldsymbol{W}(z)$. The vectors $\boldsymbol{w}_{col,j}(z)$ and $\boldsymbol{w}_{row,j}(z)$ are given by

$$\boldsymbol{w}_{col,j}(z) = [W_{1j}(z), W_{2j}(z), \cdots, W_{Nj}(z)]^T \qquad (28)$$

$$\boldsymbol{w}_{row,j}(z) = [W_{j1}(z), W_{j2}(z), \cdots, W_{jN}(z)]^T \qquad (29)$$

$\boldsymbol{w}_{col,j}(z)$ and $\boldsymbol{w}_{row,j}(z)$ do not include $W_{jj}(z)$. Equation (27) is not an explicit solution, because $W_{kk}(z)$ is included in $\boldsymbol{M}_{jj}^{-1}(z)(j \neq k)$, which is used to calculate $W_{jj}(z)$. However, it can be expected that the update changes of $W_{kk}(z)$ are very small, because a small learning rate is usually employed. Therefore, Eq.(27) can be used by treating the $W_{kk}(z)$ in $\boldsymbol{M}_{jj}^{-1}(z)$ as constants.

4.3 Combination of Learning Algorithm and Constraint

In the frequency domain, the constraints given by Eq.(27) is combined with the conventional learning algorithms given by Eqs.(7) through (10). First, the weights $W_{ji}(z)$ are adjusted in these separation learning algorithms. Next, $W_{jj}(z)$ are calculated by using the constraints. The new $W_{jj}(z)$ are used to modify the previous $W_{jj}(z)$ as follows:

Step 1 : Update $W_{jj}(r,m)$ and $W_{jk}(r,m)$ following Eqs.(7) or (10), resulting in $W_{jj}(r+1,m)$ and $W_{jk}(r+1,m)$.

Step 2 : Calculate $W_{jj}(r+1,m)$ by Eq.(27). Here, $W_{jj}(r+1,m)$ is changed to $\tilde{W}_{jj}(r+1,m)$ for convenience.

$$\tilde{W}_{jj}(r+1,m) = 1 + \boldsymbol{w}_{row,j}^T(r+1,m)\boldsymbol{M}_{jj}^{-1}(r+1,m)\boldsymbol{w}_{col,j}(r+1,m) \qquad (30)$$

Step 3 : Modify $W_{jj}(r+1,m)$ following

$$W_{jj}(r+1,m) = (1-\alpha)W_{jj}(r+1,m) + \alpha\tilde{W}_{jj}(r+1,m) \quad (0 < \alpha \leq 1) \qquad (31)$$

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

In the time domain, the equivalent process above can be carried out. However, calculating Eq.(27) in the time domain is rather complicated. Therefore, it is desirable to transform the weights $\boldsymbol{w}(n+1,l)$ updated by Eq.(4) into the frequency domain, and calculate $\tilde{W}_{jj}(z)$ by Eq.(27). $\tilde{W}_{jj}(z)$ are further transformed into $\tilde{w}_{jj}(n+1,l)$.

5. SIMULATIONS AND DISCUSSIONS

5.1 Learning Methods and Their Abbreviations

Several kinds of learning methods will be compared. They are summarized in Table 1.

Table 1: Abbreviations of applied learning algorithms.

TIME	Eqs.(4)-(6) [2]
TIME(DF)	Eqs.(4)-(6) with constraint Eq.(27)
TIME(MDP)	[3]
FREQ(1)	Eqs.(7)-(9) [2],[4],[6]
FREQ(1+PB)	Eqs.(7)-(9) and [5]
FREQ(1+DF)	Eqs.(7)-(9) with constraint Eq.(27)
FREQ(2)	Eqs.(8)-(10) [6]
FREQ(2+PB)	Eqs.(8)-(10) and [5]
FREQ(2+DF)	Eqs.(8)-(10) with constraint Eq.(27)

5.2 Simulation Setup

Simulations are performed for 2 channels and 3 channels. A mixture block, simulating actual acoustic spaces by using 256 tap FIR filters, is applied. Speeches and stationary colored signals, generated by 2nd-order AR models, are used as sources. The FFT size is 256 points for implementing in the frequency domain. FIR filters with 256 taps are used in the time domain. The initial guess for the separation blocks are $W_{jj}(z) = 1$ and $W_{kj}(z) = 0, k \neq j$. η in the learning algorithms Eqs.(4), (7) and (10) is set to $1.0 \times 10^{-6} \sim 7.0 \times 10^{-6}$. α in the proposed learning algorithm Eq.(31) is set to 1. In the time domain implementation, whose equations are omitted here, α is set to 3.0×10^{-4}.

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 $Y_i(z)$. This does not lose generality.

$$\sigma_{s1}^2 = \frac{1}{2\pi}\sum_{i=1}^{N}\int_{-\pi}^{\pi}|A_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (32)$$

$$\sigma_{i1}^2 = \frac{1}{2\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}\int_{-\pi}^{\pi}|A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \qquad (33)$$

$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}^2}{\sigma_{i1}^2} \qquad (34)$$

$$\sigma_{s2}^2 = \frac{1}{2\pi}\sum_{i=1}^{N}\int_{-\pi}^{\pi}|A_{ii}(e^{j\omega})|^2 d\omega \qquad (35)$$

$$\sigma_{i2}^2 = \frac{1}{2\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}\int_{-\pi}^{\pi}|A_{ki}(e^{j\omega})|^2 d\omega \qquad (36)$$

$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}^2}{\sigma_{i2}^2} \qquad (37)$$

σ_{s1}^2 and σ_{s2}^2 are the power of the signal source at the output and the related transfer function, respectively. σ_{i1}^2 and σ_{i2}^2 are the power of the interference signal and the related transfer function, respectively.

In evaluating the signal distortion, we assume that $S_i(z)$ is dominant in $X_i(z)$, and $H_{ii}(z)S_i(z)$ can be used as the criterion. Furthermore, $S_i(z)$ is assumed to be separated at $Y_i(z)$. Based on these assumptions, four different measures, as shown below, are used to evaluate signal distortion.

$$SD_{1x} = 10\log_{10}\frac{\sigma_{d1x}^2}{\sigma_1^2}, x = a, b \quad (38)$$

$$SD_{2x} = 10\log_{10}\frac{\sigma_{d2x}^2}{\sigma_2^2}, x = a, b \quad (39)$$

$$\sigma_{d1a}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ii}(e^{j\omega})S_i(e^{j\omega}) - A_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (40)$$

$$\sigma_{d1b}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ii}(e^{j\omega})S_i(e^{j\omega})| - |A_{ii}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \quad (41)$$

$$\sigma_1^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (42)$$

$$\sigma_{d2a}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ii}(e^{j\omega}) - A_{ii}(e^{j\omega})|^2 d\omega \quad (43)$$

$$\sigma_{d2b}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|H_{ii}(e^{j\omega})| - |A_{ii}(e^{j\omega})|)^2 d\omega \quad (44)$$

$$\sigma_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H_{ii}(e^{j\omega})|^2 d\omega \quad (45)$$

Since BSS systems are unable to control amplitude of the output signals, the output signals may differ from the criteria in amplitude. In order to neglect this scaling effect in calculating SD_{1x} and SD_{2x}, the average power of $H_{ii}(e^{j\omega})S_i(z)$, $A_{ii}(e^{j\omega})S_i(z)$, $H_{ii}(e^{j\omega})$, and $A_{ii}(e^{j\omega})$ are normalized to unity.

Since the signal distortion SD_{ix} are evaluated based on the difference between the criterion and the simulation results, the smaller SD_{ix} indicate the lower signal distortion. The signal source separation SIR_i are evaluated by a signal to interference ratio. Therefore, basically speaking, the larger SIR_i assume the higher separation performance. However, when the signal distortion is large, the separated signal includes the distortion components, and SIR_i are not reliable. SIR_i should be evaluated under lower signal distortion.

5.3 Two Channel Speech Signal Sources

Simulation results for 2 channel speech signals are summarized in Table 2.

In the time domain implementation, TIME is not good in the signal distortion. SIR_i in TIME, which are higher than the others, have no meaning due to the large signal distortions. TIME(MDP) can improve the signal distortion as expected. However, source separation is not good because the observations include both sources.

On the other hand, the proposed method TIME(DF) can improve the signal distortion, while maintaining relatively high SIR_i.

In the frequency domain implementation, the conventional method FREQ(1) is not good in both SIR_i and SD_{ix}. FREQ(2) can improve both SIR_i and SD_{ix} compared with FREQ(1). However, its SD_{ix} are not sufficient. Even though FREQ(1+PB) and FREQ(2+PB) can improve SIR_i and SD_{ix}, the proposed method FREQ(2+DF) can provide higher performances.

Table 2: Comparison of learning algorithms of BSS systems for 2 channel speech signals.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
TIME	12.2	5.56	0.25	-2.94	0.57	-3.82
TIME(MDP)	3.98	2.90	-10.3	-13.6	-8.24	-12.3
TIME(DF)	8.31	4.33	-12.5	-16.7	-15.4	-19.8
FREQ(1)	7.37	2.61	-6.23	-8.67	-2.82	-2.95
FREQ(1+PB)	15.2	6.59	-23.1	-27.1	-16.9	-18.8
FREQ(1+DF)	9.68	6.38	-13.5	-18.1	-15.1	-18.3
FREQ(2)	13.0	12.9	-9.43	-15.1	-10.9	-13.9
FREQ(2+PB)	14.3	11.6	-16.9	-20.6	-16.9	-18.9
FREQ(2+DF)	19.8	11.8	-24.6	-28.1	-17.9	-20.6

5.4 Two Channel Stationary Colored Signal Sources

2nd-order AR models are used to generate stationary colored signals. The frequency components are mainly located around $1/8f_s$ and $3/8f_s$. SIR_i and SD_i are summarized in Table 3.

Table 3: Comparison of learning algorithms of BSS systems for 2 channel stationary colored signals.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
TIME	9.49	7.07	-0.28	-3.11	-0.69	-4.99
TIME(MDP)	4.49	2.20	-15.7	-18.8	-13.7	-16.5
TIME(DF)	8.05	4.07	-14.4	-16.6	-10.4	-13.2
FREQ(1)	4.93	4.32	-3.62	-5.03	-3.67	-4.28
FREQ(1+PB)	3.22	1.51	-12.7	-14.3	-7.83	-9.38
FREQ(1+DF)	7.63	4.06	-18.4	-20.8	-10.8	-14.1
FREQ(2)	5.22	2.82	-12.2	-14.3	-7.42	-9.44
FREQ(2+PB)	5.35	2.62	-14.1	-15.8	-8.39	-10.7
FREQ(2+DF)	5.67	2.72	-13.7	-15.7	-8.27	-11.1

In the time domain implementations, the signal distortions in TIME are not good. The output signal spectra outside the signal bands are extremely amplified. This is due to the degree of freedom in the FF-BSS systems as discussed in Sec.3. The signal distortions can be reduced by TIME(MDP). However, SIR_i are not good as shown in Table 3. The proposed method TIME(DF) can improve both SIR_i and SD_{ix}.

In the frequency domain implementations, FREQ(2) is more efficient in SD_{ix} compared to FREQ(1). Even though FREQ(1+PB) can improve SD_{ix}, FREQ(1+DF) is more efficient. FREQ(2+PB) and FREQ(2+DF) slightly improve SD_{ix}. Their differences are small. In this case, FREQ(1+DF) can provide the best performances in both SIR_i and SD_{ix}.

5.5 Three Channel Speech Signal Sources

Simulation results are listed in Table 4. Similar properties both in the time and frequency domains are recognized as in the 2-channel speech signal. FREQ(1+PB) and FREQ(2+PB) can improve both SIR_i and SD_{ix}. Still FREQ(2+DF) can improve more in SD_{ix}.

Table 4: Comparison of learning algorithms of BSS systems for 3 channel 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(MDP)	6.25	4.36	-8.02	-11.2	-10.2	-14.1
TIME(DF)	8.00	5.24	-14.0	-17.9	-17.1	-20.5
FREQ(1)	11.2	5.32	-10.4	-16.1	-4.39	-4.52
FREQ(1+PB)	15.9	8.61	-19.5	-24.2	-17.7	-19.7
FREQ(1+DF)	8.50	7.36	-14.7	-21.1	-16.3	-18.7
FREQ(2)	17.2	10.3	-13.4	-19.4	-13.9	-16.4
FREQ(2+PB)	16.3	9.68	-17.3	-21.1	-18.0	-19.9
FREQ(2+DF)	16.1	9.56	-24.1	-28.4	-20.1	-21.7

5.6 Comparison of Learning Algorithms

FREQ(MDP) is not good in source separation. FREQ(1+PB) and FREQ(2+PB) are useful in both SIR_i and SD_{ix}. However, they are not good for the stationary signal sources. Furthermore, either FREQ(1+DF) or FREQ(2+DF) can provide higher performances compared to them. FREQ(1+DF) is effective for stationary signals, and FREQ(2+DF) is useful for non-stationary signals. Furthermore, we have confirmed that FREQ(2+DF) can provide good performances for a combination of stationary and non-stationary signals.

6. CONCLUSION

The condition for complete source separation and distortion free is proposed for multi-channel convolutive FF-BSS systems. This condition is combined with the learning algorithms. Computer simulations demonstrate the proposed method can improve in both the source separation and the signal distortion.

REFERENCES

[1] 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.

[2] 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.

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

[4] 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.

[5] N.Murata, S.Ikeda and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals", Neurocomputing, vol.41, pp.1-24, Oct. 2001.

[6] 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.

[7] 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, pp.1257-1262 Montreal, Canada, August 2005.

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

A. DERIVATION OF EQ.(27)

$\det \boldsymbol{W}(z)$ is given by

$$\det \boldsymbol{W}(z) = \sum_{k=1}^{N} W_{jk}(z)(-1)^{j+k} \det \boldsymbol{M}_{jk}(z) \quad (46)$$

In general, $\hat{W}_{jj}(z)$ is expressed by

$$\hat{W}_{jj}(z) = (-1)^{2j} \det \boldsymbol{M}_{jj}(z) = \det \boldsymbol{M}_{jj}(z) \quad (47)$$

From Eqs.(26), (46) and (47), we obtain

$$\det \boldsymbol{M}_{jj}(z) = \sum_{k=1}^{N} W_{jk}(z)(-1)^{j+k} \det \boldsymbol{M}_{jk}(z) \quad (48)$$

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

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

$\det \boldsymbol{M}_{jk}(z)$ is further rewritten as:

$$\det \boldsymbol{M}_{jk}(z) = \sum_{\substack{l=1 \\ \neq j}}^{N} W_{lj}(z)\kappa(j,k,l) \det \boldsymbol{m}_{lj}(z) \quad (50)$$

$$\kappa(j,k,l) = \begin{cases} (-1)^{l+j} & (l < k) \\ -(-1)^{l+j} & (l \geq k) \end{cases} \quad (51)$$

$\boldsymbol{m}_{lj}(z)$ is an $(N-2) \times (N-2)$ minor matrix, i.e. the l-th row and the j-th column are removed from $\boldsymbol{M}_{jk}(z)$. Therefore, the right hand side of Eq.(49) is rewritten as:

$$-\sum_{\substack{k=1 \\ \neq j}}^{N} W_{jk}(z) \sum_{\substack{l=1 \\ \neq j}}^{N} W_{lj}(z)\kappa(k,k,l) \det \boldsymbol{m}_{lj}(z) \quad (52)$$

$$= -\boldsymbol{w}_{row,j}^T(z)\boldsymbol{M}_{jj}(z)\boldsymbol{w}_{col,j}(z) \quad (53)$$

Finally, this results in the equation given in Eq.(27).

$$W_{jj}(z) = 1 + \boldsymbol{w}_{row,j}^T(z)\frac{\boldsymbol{M}_{jj}(z)}{\det \boldsymbol{M}_{jj}(z)}\boldsymbol{w}_{col,j}(z) \quad (54)$$

$$= 1 + \boldsymbol{w}_{row,j}^T(z)\boldsymbol{M}_{jj}^{-1}(z)\boldsymbol{w}_{col,j}(z) \quad (55)$$