社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

多チャネル信号源と畳み込み混合に対するフィードフォワード形BSSにおける信号歪み抑制形学習法

堀田 明秀† 中山 謙二† 平野 晃宏† 出島 康宏†

† 金沢大学大学院 自然科学研究科
920-1192 石川県金沢市角間町
E-mail: †{horita,dejima}@leo.ec.t.kanazawa-u.ac.jp, ††{nakayama,hirano}@t.kanazawa-u.ac.jp

あらまし 畳み込み混合過程におけるフィードフォワード (FF-) 形ブラインドソースセパレーション (BSS) では自由度が存在するため信号歪みが生じる．我々は，信号源ーセンサーが 2 チャンネルの場合において，完全分離と信号無歪みの条件を制約条件として課す信号歪み抑制学習アルゴリズムを時間領域，周波数領域の FF-BSS に対して提案してきた．本稿では，信号歪み抑制の制約条件を多チャンネルに拡張し，かつ，計算の複雑さを軽減するために制約条件を近似する方式を提案する．音声を用いたコンピュータシミュレーションによってその近似制約方式と厳密制約方式がほぼ同等の分離性能と信号歪み抑制が得られることを確認した．また，3 チャンネルにおいても，従来方式より特性が改善されることを確認した．
キーワード ブラインドソースセパレーション 信号歪み 収束性 学習アルゴリズム 畳み込み

A Distortion Free Learning Algorithm for Feed-Forward BSS with Convolutive Mixture and Multi-Channel Signal Sources

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

† Graduate School of Natural Science and Technology, Kanazawa Univ.
Kakuma-machi, Kanazawa, 920-1192
E-mail: †{horita,dejima}@leo.ec.t.kanazawa-u.ac.jp, ††{nakayama,hirano}@t.kanazawa-u.ac.jp

Abstract Feed-forward Blind Source Separation (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. 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-BSSs. This condition requires the 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 conventionals for three channels.
Key words Blind source separation, Signal distortion, Convergence, Learning algorithm, Convolutive

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

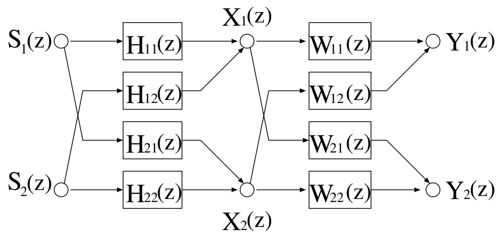

図 1 BSS system with 2 signal sources and 2 sensors.

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], [10].

BSS learning algorithms make the output signals statistically independent. However, this approach cannot always guarantee distortion free separation. A method for suppressing signal distortion in which the distance between the observed signals and the separated signals is added to the cost function has been proposed [7]. However, since the observations may consist of many signal sources, by applying this method, separation itself becomes very difficult. To begin with, even though signal distortion in BSS systems is an important problem, it has not been well addressed until recently. Previously, an evaluation measure of signal distortion has been discussed, and conditions for source separation as well as distortion free have been derived. Taking these conditions in consideration, convergence properties have been analyzed. Furthermore, a new learning algorithm with a constraint based on these conditions for two channels has been proposed [13].

In this paper, the distortion free constraint is extended to more than two channels. Moreover, since the calculation is computationally expensive, an approximation is proposed.

2 BSS Systems for Convolutive Mixture

2.1 Network Structure and Equations

A block diagram of a 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:

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

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

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

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

2.3 Learning Algorithm in Frequency Domain

The same learning algorithm in the frequency domain using FFT, is defined as [4], [8], [9]:

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

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

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

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 Criterion for Signal Distortion

In this paper, the signals $H_{ii}(z)S_i(z)$ or $H_{ji}(z)S_i(z)$ $(j \neq i)$ are taken into account as a criterion for signal distortion [2], [11]. These signals are unmixed, i.e. they follow a single path from the original sources to the sensors and resemble the original situation closest. In other words, only signal distortion caused by the BSS systems itself is considered.

Signal distortion can also be evaluated through the transfer functions, i.e. the transfer function from the i-th source to the k-th output $A_{ki}(z)$ is compared to the one from the i-th source to the j-th sensor $H_{ji}(z)$. Furthermore, signal distortion can be evaluation in two different ways, i.e. the amplitude responses including the phase and amplitude responses excluding the phase with respect to the signals and transfer functions. As a result, four kinds of measures (SD_{ix} $i = 1, 2$ $x = a, b$), as shown below, are applied:

$$SD_{1x} = 10\log_{10}\frac{\sigma_{d1x}}{\sigma_1}, x = a, b \quad (17)$$

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

$$\sigma_{d1a} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega}) - A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (19)$$

$$\sigma_{d1b} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})S_i(e^{j\omega})| - |A_{ki}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \quad (20)$$

$$\sigma_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (21)$$

$$\sigma_{d2a} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega}) - A_{ki}(e^{j\omega})|^2 d\omega \quad (22)$$

$$\sigma_{d2b} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})| - |A_{ki}(e^{j\omega})|)^2 d\omega \quad (23)$$

$$\sigma_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})|^2 d\omega \quad (24)$$

Since the 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.

Smaller values for the evaluations SD_{ix} $i = 1, 2$ $x = a, b$ indicate better performances with respect to signal distortion.

4 Source Separation and Signal Distortion

4.1 Source Separation and Signal Distortion

For simplicity, a BSS system with 2-sources and 2-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, the condition for distortion free source separation can be expressed as follows:

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (25)$$

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (26)$$

$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (27)$$

$$W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (28)$$

The conditions for complete source separation are expressed by Eqs. (26) and (27). These equations imply that all non-diagonal elements of $\boldsymbol{A}(z)$ are zero. The conditions for distortion free are expressed by Eqs. (25) and (28). These equations imply that the diagonal elements of $\boldsymbol{A}(z)$ equal $H_{ii}(z)$.

The conventional learning algorithm given by Eqs. (11)-(14) satisfies only Eqs. (26) and (27). Equations (25) and (28) are not guaranteed to be satisfied. Therefore, by applying this algorithm, signal distortion may occur.

4.2 Distortion Free Condition and Its Application to Learning Algorithm

Equations (26) and (27) are rewritten, expressing $H_{ji}(z)$ as follows:

$$H_{ji}(z) = -\frac{W_{ji}(z)}{W_{jj}(z)}H_{ii}(z) \quad (29)$$

By substituting the above equations into Eqs. (25) and (28), $H_{ji}(z)$ can be removed, and the following equations consisting only of $W_{kj}(z)$ can be obtained:

$$W_{jj}{}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \quad (30)$$
$$j = 1, 2, k = 1, 2, j \neq k$$

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

$$W_{jj}(z) = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2}, j = 1, 2 \quad (31)$$

This condition can be included in the learning processes for BSS systems in the time domain as well as trained in the frequency domain as a distortion free constraint. In this method, $W_{12}(z)$ and $W_{21}(z)$ are still being calculated as in the conventional methods, following Eq. (11) and (14). However, by using $W_{12}(z)$ and $W_{21}(z)$, $W_{jj}(z)$ are obtained so as to satisfy Eq. (30) i.e. from Eq. (31).

It should be noted, that in the early stage of the learning process, the signal sources are not well separated, because the separation block starts from an initial guess. Taking this situation into account, the constraint of Eq. (31) is gradually imposed as the learning process makes progress.

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

$$w_{jk}(n+1,l) = \tilde{w}_{jk}(n+1,l) \qquad (j \neq k) \qquad (32)$$

$$w_{jj}(n+1,l) = (1-\alpha)\tilde{w}_{jj}(n+1,l) + \alpha\bar{w}_{jj}(n,l) \qquad (33)$$

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

where $\tilde{w}_{jk}(n+1,l)$ are the separation system updated by Eqs. (11)-(13) following the conventional method, and $\bar{w}_{jj}(n,l)$ is determined by Eq. (30).

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

$$W_{jk}(r+1,m) = \tilde{W}_{jk}(r+1,m) \qquad (j \neq k) \qquad (34)$$

$$W_{jj}(r+1,m) = (1-\alpha)\tilde{W}_{jj}(r+1,m) + \alpha\frac{1+\sqrt{1+4W_{12}(r,m)W_{21}(r,m)}}{2} \qquad (35)$$

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

where $\tilde{W}_{jk}(r+1,m)$ are the separation systems updated by Eq. (14).

4.3 Learning Algorithm in Time Domain Suppressing Signal Distortion by Using Observations as Criteria

Previously, a learning algorithm for reducing signal distortion has been proposed [7]. The cost function as defined in Sec. 2.2 has been extended by including the distance between the observed signals and the output signals. Therefore, the output signals are forced to approach to the observed signals. The resulting update equation for the filter coefficients is given by:

$$\begin{aligned} &\boldsymbol{w}(n+1,l) = \boldsymbol{w}(n,l) \\ &+ \eta \sum_{q=0}^{K_w-1} [\boldsymbol{I}\delta(n-q) - \langle \boldsymbol{\Phi}(\boldsymbol{y}(n))\boldsymbol{y}^T(n-l+q)\rangle \\ &- \mu(\boldsymbol{y}(n) - \boldsymbol{x}(n))\boldsymbol{y}^T(n-l+q)]\boldsymbol{w}(n,q) \end{aligned} \qquad (36)$$

In this method, the output signals $Y_i(z) = A_{ii}(z)S_i(z) + A_{ij}(z)S_j(z)$ are stimulated to approach to the observed signals $X_i(z) = H_{ii}(z)S_i(z) + H_{ij}(z)S_j(z)$. Since it is assumed that $S_i(z)$ 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. According to our discussion from Sec. 4.1, if $A_{ii}(z)$ approaches $H_{ii}(z)$, then distortion free is guaranteed. Furthermore, in the same section, it was also concluded that the non-diagonal elements of $\boldsymbol{A}(z)$ should be equal to zero in order to satisfy source separation. However, here $A_{ij}(z)$ tends to approach to $H_{ij}(z)$. Therefore, this algorithm might achieve low signal distortion, but perform poor with respect to signal source separation due to these residual cross terms.

5 Generalization of Proposed Constraint

The constraint described in Sec. 4.2 is extended to more than two channels. The condition for distortion free source separation can be expressed as follows:

$$\boldsymbol{W}(z)\boldsymbol{H}(z) = \boldsymbol{\Lambda}(\boldsymbol{z}) \qquad (37)$$

$$\boldsymbol{\Lambda}(z) = \mathrm{diag}[\boldsymbol{H}(z)] \qquad (38)$$

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

$$\boldsymbol{\Gamma}(z) = \boldsymbol{H}(z) - \boldsymbol{\Lambda}(z) \qquad (39)$$

It satisfies:

$$\boldsymbol{W}(z)(\boldsymbol{\Lambda}(z) + \boldsymbol{\Gamma}(z)) = \boldsymbol{\Lambda}(z) \qquad (40)$$

Furthermore,

$$\boldsymbol{\Gamma}(z) = \boldsymbol{W}^{-1}(z)(\boldsymbol{I} - \boldsymbol{W}(z))\boldsymbol{\Lambda}(\boldsymbol{z}) \qquad (41)$$

$$= (\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda}(\boldsymbol{z}) \qquad (42)$$

From Eq. (39):

$$\mathrm{diag}[\boldsymbol{\Gamma}(z)] = \mathrm{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})\boldsymbol{\Lambda}(\boldsymbol{z})] = \boldsymbol{0} \qquad (43)$$

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

$$\mathrm{diag}[(\boldsymbol{W}^{-1}(z) - \boldsymbol{I})] = \boldsymbol{0} \qquad (44)$$

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

$$\boldsymbol{W}^{-1}(z) = \frac{\mathrm{adj}\,\boldsymbol{W}(z)}{\det \boldsymbol{W}(z)} \qquad (45)$$

$\mathrm{adj}\,\boldsymbol{W}(z)$ is the adjugate matrix of $\boldsymbol{W}(z)$, and denoted $\hat{\boldsymbol{W}}(z)$. Since, a diagonal element of $\boldsymbol{W}^{-1}(z) = \hat{\boldsymbol{W}}(z)/\det \boldsymbol{W}(z)$, then:

$$\frac{\hat{\boldsymbol{W}}(z)}{\det \boldsymbol{W}(z)} = 1 \qquad (46)$$

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

$$\det \boldsymbol{W}(z) = \sum_{j=1}^{N} W_{ij}(z)(-1)^{i+j}\det \boldsymbol{M}_{ij}(z) \qquad (47)$$

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

$$\hat{\boldsymbol{W}}(z) = (-1)^{2i}\det \boldsymbol{M}_{ij}(z) = \det \boldsymbol{M}_{ij}(z) \qquad (48)$$

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

$$\det \boldsymbol{M}_{ij}(z) = \sum_{j=1}^{N} W_{ij}(z)(-1)^{i+j}\det \boldsymbol{M}_{ij}(z) \qquad (49)$$

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

$$\det \boldsymbol{M}_{ii}(z)(1 - W_{ii}(z)) = \sum_{\substack{j=1 \\ \neq i}}^{N} W_{ij}(z)(-1)^{i+j}\det \boldsymbol{M}_{ij}(z) \qquad (50)$$

The right hand side is further rewritten by:

$$-\boldsymbol{w}_{ix}^T(z)\,\mathrm{adj}\,\boldsymbol{M}_{ij}(z)\boldsymbol{w}_{xi}(z) \tag{51}$$

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

$$\boldsymbol{w}_{xi}(z) = [W_{1i}(z), \cdots, W_{yi}(z), \cdots, W_{Ni}(z)]^T \tag{52}$$

$$\boldsymbol{w}_{ix}(z) = [W_{i1}(z), \cdots, W_{iy}(z), \cdots, W_{iN}(z)]^T \tag{53}$$

and $y \neq i$. Finally,

$$W_{jj}(z) = 1 + \boldsymbol{w}_{ix}^T(z)\frac{\mathrm{adj}[\boldsymbol{M}_{ij}(z)]}{\det[\boldsymbol{M}_{ij}(z)]}\boldsymbol{w}_{xi}(z) \tag{54}$$

$$= 1 + \boldsymbol{w}_{ix}^T(z)\boldsymbol{M}_{ii}^{-1}(z)\boldsymbol{w}_{xi}(z) \tag{55}$$

Equation (55) is not an explicit solution, because $\boldsymbol{M}_{jj}^{-1}(z)$ 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:

$$W_{jk}(r+1,m) = \tilde{W}_{jk}(r+1,m) \qquad (j \neq k) \tag{56}$$

$$W_{jj}(r+1,m) = (1-\alpha)\tilde{W}_{jj}(r+1,m) + \alpha\bar{W}_{jj}(r,m) \tag{57}$$

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

where $\tilde{W}_{jk}(r+1,m)$ are the separation systems updated by Eq. (14) and $\bar{W}_{jj}(r,m)$ 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.

$$w_{jk}(n+1,l) = \tilde{w}_{jk}(n+1,l) \qquad (j \neq k) \tag{58}$$

$$w_{jj}(n+1,l) = (1-\alpha)\tilde{w}_{jj}(n+1,l) + \alpha\bar{w}_{jj}(n,l) \tag{59}$$

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

where $\tilde{w}_{jk}(n+1,l)$ are the separation system updated by Eqs. (11)-(13) following the conventional method, and $\bar{w}_{jj}(n,l)$ is Eq. (55) calculated in frequency domain.

6 Simulations and Discussion

6.1 Learning Methods and Their Abbreviations

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

6.2 Simulation Conditions

Simulations are performed for the following two cases:

（1） Two sources and two sensors.

（2） Three sources and three sensors.

表 1 Abbreviations of the applied learning algorithms.

TIME	Eqs. (11)-(13) [4]
TIME (DF)	Eqs. (11)-(13) with the distortion free constraint Eqs. (32),(33)
TIME (ADF)	Eqs. (11)-(13) with the distortion free constraint applying Eqs. (58),(59)
TIME (MDP)	Eq. (36) [7]
FREQ	Eqs. (15)-(14) [9]
FREQ (DF)	Eqs. (15)-(14) with the distortion free constraint Eqs. (34),(35)
FREQ (ADF)	Eqs. (15)-(14) with the distortion free constraint applying Eqs. (56),(57)

表 2 Comparison among seven kinds of BSS systems for speech signals in the case 1.

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 (DF)	8.33	4.33	-12.1	-16.2	-15.4	-19.9
TIME (ADF)	8.31	4.33	-12.5	-16.7	-15.4	-19.8
TIME (MDP)	3.98	2.90	-10.3	-13.6	-8.24	-12.3
FREQ	18.8	10.5	-10.9	-16.5	-11.9	-15.0
FREQ (DF)	18.7	10.1	-25.8	-30.0	-18.4	-21.1
FREQ (ADF)	18.7	10.1	-25.8	-30.0	-18.4	-21.2

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 $W_{kj}(z) = 0, k \neq j$.

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.

$$\sigma_{s1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{N}|A_{ii}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \tag{60}$$

$$\sigma_{i1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}|A_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \tag{61}$$

$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}}{\sigma_{i1}} \tag{62}$$

$$\sigma_{s2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{N}|A_{ii}(e^{j\omega})|^2 d\omega \tag{63}$$

$$\sigma_{i2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=1}^{N}\sum_{\substack{i=1\\ \neq k}}^{N}|A_{ki}(e^{j\omega})|^2 d\omega \tag{64}$$

$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}}{\sigma_{i2}} \tag{65}$$

Larger values for the evaluations SIR_i $(i = 1, 2)$ indicate better performances with respect to source separation.

6.3 Performance in Case 1

Evaluation measures in case 1 are summarized in Table 2.

表 3 Comparison among seven kinds of BSS systems for speech signals in the case 2.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
TIME	-1.07	5.98	0.54	-2.46	-0.77	-4.71
TIME (ADF)	5.44	4.54	-14.0	-18.2	-17.4	-20.8
TIME (MDP)	4.17	3.99	-8.56	-11.9	-10.5	-14.3
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

We have set $i = j = k$ in Eqs. (17)-(24) with respect to the signal distortion evaluations, because $S_1(z)$ and $S_2(z)$ are assumed to be separated at $Y_1(z)$ and $Y_2(z)$, respectively.

The conventional learning algorithm TIME performs worst regarding the signal distortion measures SD_{ix}. TIME (MDP) can improve signal distortion. However, as discussed in Sec. 4. 3, due to the residual cross terms, the signal to interference ratios SIR_i are unsatisfactory. TIME (DF) can improve SD_{ix} as well as SIR_i. Compared to TIME, the evaluation values for SIR_i are slightly lower. However, in 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.

FREQ is better than the time domain implementations regarding source separation. Regarding signal distortion, 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, FREQ (DF) can drastically improve from FREQ regarding signal distortion, while maintaining high signal separation performances.

The learning algorithms applying the distortion free constraint following the approximation, i.e. TIME (ADF) and FREQ (ADF), are able to obtain the same performances regarding source separation as well as signal distortion as their counterparts. TIME (ADF) and FREQ (ADF) apply an approximation method for calculating the distoriton free constraint by treating the $W_{kk}(z)$ in $\boldsymbol{M}_{jj}^{-1}(z)$ from Eq. (55) as constants. The simulation results demonstrate that the approximation can be applied succesfully.

6. 4 Performance in Case 2

Evaluation measures in case 2 are summarized in Table 3.

In the time domain implementation, TIME(ADF) is the best regarding both signal separation and signal distortion. TIME(MDP) is inferior to the proposed, its signal distortion is larger than TIME(ADF) by about 6dB. In the case of the frequency domain, the proposed approach FREQ(ADF) still can reduce signal distortion from the conventional method, while maintaining almost the same separation performances. From these results, it can be concluded that the approximated constraint is useful for multi-channel FF-BSS systems.

7 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. Furthermore, good performances are also obtained in situations using three channels.

文 献

[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 nonstationary 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] 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.

[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] 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.

[12] 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.

[13] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "A Distortion Free Learning Algorithm for Feedforward BSS and Its Comparative Study with Feedback BSS" IJCNN2006, Vancouver, BC, Canada, July 2006. to be presented.