COMPLEX ICA FOR DIRECTION FINDING AND SEPARATION OF BROADBAND SOUND SOURCES – HIGH-QUALITY SIGNAL SEPARATION USING AN INVERSE FILTER –

Noboru Nakasako and Hisanao Ogura

School of Biology-Oriented Sc. & Tech., Kinki Univ.
Nishi-Mitani 930, Uchita-cho, Naga-gun, Wakayama 649-6493 Japan.
nakasako,ogura@info.waka.kindai.ac.jp

ABSTRACT

Recently, the problem of independent component analysis and/or blind signal separation becomes a very popular and emerging field of research, because the problem contains many potential applications, e.g., speech/image enhancement and/or recognition, noise reduction and so on. However, there are problems in a case when observations have time differences between source signals and are affected by system (room etc.). The authors developed the blind separation of narrowband complex signals with time delay by means of complex Hermite moments and unitary transformations.

In this paper we first estimate the time delays or the distances between receivers and sound sources as well as their power spectra. This method enables us to find the sound source positions or the directions of arrival sounds. Further, we construct an inverse filter for real mixed signals to reproduce the high-quality separated signals, while the usual complex ICA requires synthesizing all frequency components of separated signals. Finally, through simulation, we demonstrate the high-quality blind separation as well as the effectiveness for the estimation of the source parameters.

1. INTRODUCTION

Recently, the problem of independent component analysis (ICA) and/or blind source separation (BSS)[1, 2, 3, 4] becomes a very popular and emerging field of research, because the problem contains many potential applications, e.g., speech/image enhancement and/or recognition, noise reduction and so on. In such a problem, a priori information we can only utilize is the statistical independency between source signals.

At the early stage of research on the blind separation, most of works have been devoted to instantaneous mixture of stochastic independent signals. The authors also proposed the ICA solution based on the evaluation function composed of Hermite moment and applied it to several examples[5, 6, 7, 8].

However, there are problems in a case when observations have time differences between source signals and are affected by system (room etc.). Recently, many works have been reported devoted to applications of blind separation to more real-world conditions, where mixed signals are time-delayed and/or convolved[9, 10, 11, 12, 13, 14]. The authors also developed the blind separation of narrowband complex signals with time delay by means of complex Hermite moments and unitary transformations[15, 16, 17]. In fact, in real-world applications wideband signals are filtered and delayed before reaching sensors.

This work was supported in part by JSPS Grants-in-Aid for Scientific Research (No.13650472)

On the other hand, in many actual fields of engineering, the source localization problem has been very attractive and a number of approaches have been developed, e.g., MUSIC[18], Spatio-temporal method[19, 20], RSS and CSM[21, 22]. Although there have been few approaches based on the statistical independency, the sound source localization method was proposed using the estimated parameters by ICA[23].

In this paper, along the same line as our previous work, we first estimate the time delays or the distances between receivers and sound sources as well as their power spectra. This method enables us to find the sound source positions or the directions of arrival sounds. In our method, it is easy to understand the physical meaning (i.e., magnitude and phase information) comparing with the other ICA methods, since our method parameterizes the separating matrix. Further, we can easily construct an inverse filter for (over-all) real mixed signals to reproduce the high-quality separated signals, while the usual complex ICA requires synthesizing all frequency components of separated signals. Finally, through simulation, we demonstrate the high-quality blind separation as well as the effectiveness for the estimation of the source parameters.

2. PARAMETER ESTIMATION FOR NARROW-BANDED MIXTURES WITH TIME DIFFERENCES

Mixing process Let $\boldsymbol{W}(m) = (W_1(m), W_2(m))^T$ be the normalized ($\langle W_j(m)\rangle = 0,\ \ \langle W_j^2(m)\rangle = 1, \quad j = 1, 2$) and independent real signals. Suppose that they are mixed as $\boldsymbol{X}'(m) = (X_1'(m), X_2'(m))^T$ with time differences:

$$X_1'(m) = aW_1(m-p) + bW_2(m-q) \tag{1}$$
$$X_2'(m) = cW_1(m-r) + dW_2(m-s) \tag{2}$$

Here, $a, b, c, d \geq 0$ denote the amplitudes of original signals in mixtures, and the time delays p, q, r, s are assumed to be all positive integers. We confine our method to 2D case for simplicity though we can in principle extend our method to more than 3D case.

2.1. Narrow-banded signals of time-shifted mixtures and orthogonalization

The technique of BSS enables us to estimate the unknown parameters a, b, c, d and p, q, r, s, except for indeterminacy in ICA. More concretely, the amplitude ratios $a/c, b/d$ and the phase differences $p-r, q-s$ are obtained through BSS for complex signals.

Complex narrow-band filter First, let us transform the mixture $X'_j(m)$ $(j=1,2)$ into the complex signal $X'_{j\nu}(m)$, by using the complex filter of 1st order with the center angular frequency Λ_ν(rad) (=$2\pi f_\nu/(\text{sampling rate})$, f_ν: center frequency (Hz))

$$X'_{j\nu}(m) = \sqrt{\alpha}\mathrm{e}^{\mathrm{i}\Lambda_0} X'_{j\nu}(m-1) + X'_j(m) \qquad (3)$$
$$(\alpha = \mathrm{e}^{-K}, K > 0).$$

Here, K (or α) is a parameter representing the bandwidth and $\Delta\Lambda \simeq K$. Generally denoting such a narrow-banded complex signal as $X'_{\nu j}(m) = \mathrm{e}^{im\Lambda_\nu} x'_{\nu j}(m)$, $x'_{\nu j}(m)$ represents the complex time series slowly changing with a time constant about $m_0 = 2/K$.

Narrow-banded complex signal The narrow-banded complex signals $\boldsymbol{X}'_\nu(m) = [X'_{\nu 1}(m), X'_{\nu 2}(m)]^T$, $\boldsymbol{W}'_\nu(m) = [W'_{\nu 1}(m), W'_{\nu 2}(m)]^T$ can be approximately expressed as

$$\boldsymbol{X}'_\nu(m) \simeq \mathrm{e}^{\mathrm{i}\Lambda_\nu m}\boldsymbol{x}'_\nu(m), \quad \boldsymbol{W}'_\nu(m) \simeq \mathrm{e}^{\mathrm{i}\Lambda_\nu m}\boldsymbol{w}'_\nu(m) \qquad (4)$$

with $\boldsymbol{x}'_\nu(m) = [x'_{\nu 1}(m), x'_{\nu 2}(m)]^T$, $\boldsymbol{w}'_\nu(m) = [w'_{\nu 1}(m), w'_{\nu 2}(m)]^T$.

Although the narrow-banded $W'_{\nu j}(m), j = 1,2$ are not normalized, let $S_j(\Lambda_\nu)$ be the power spectrum of original signal $W_j(m)$, then the variance is approximately proportional to $S_j(\Lambda_\nu)$ (A: constant):

$$\langle |W'_{\nu j}(m)|^2 \rangle \simeq A S_j(\Lambda_\nu) \qquad j = 1,2 \qquad (5)$$

As these parameters are actually unknown, for convenience, let $P_\nu = \sqrt{AS_1(\Lambda_\nu)}$ and $Q_\nu = \sqrt{AS_2(\Lambda_\nu)}$ in the formulation below. Thus, $W_{\nu j}(m)$, $j = 1,2$ are supposed to be normalized independent signals where $W'_{\nu 1}(m) = P_\nu W_{\nu 1}(m)$ and $W'_{\nu 2}(m) = Q_\nu W_{\nu 2}(m)$ (i.e., $\langle \overline{W_{\nu j}(m)} W_{\nu k}(m) \rangle = \delta_{jk}, \quad j,k = 1,2$).

Narrow-banded complex mixtures Let us denote the narrow-banded signals of the ν-th center frequency Λ_ν as $\boldsymbol{X}'_\nu(m)$ and $\boldsymbol{W}_\nu(m)$, their relationship can be approximated in the following form as:

$$\boldsymbol{X}'_\nu(m) = \begin{bmatrix} X'_{\nu 1}(m) \\ X'_{\nu 2}(m) \end{bmatrix}$$
$$\simeq \begin{bmatrix} a\mathrm{e}^{-\mathrm{i}\Lambda_\nu p} & b\mathrm{e}^{-\mathrm{i}\Lambda_\nu q} \\ c\mathrm{e}^{-\mathrm{i}\Lambda_\nu r} & d\mathrm{e}^{-\mathrm{i}\Lambda_\nu s} \end{bmatrix} \begin{bmatrix} W'_{\nu 1}(m) \\ W'_{\nu 2}(m) \end{bmatrix} \qquad (6)$$
$$= \boldsymbol{K}(\mathrm{e}^{-\mathrm{i}\Lambda_\nu})\boldsymbol{W}_\nu(m) \qquad (7)$$
$$\boldsymbol{K}_\nu = \boldsymbol{K}(\mathrm{e}^{-\mathrm{i}\Lambda_\nu}) \equiv \begin{bmatrix} P_\nu a\mathrm{e}^{-\mathrm{i}\Lambda_\nu p} & Q_\nu b\mathrm{e}^{-\mathrm{i}\Lambda_\nu q} \\ P_\nu c\mathrm{e}^{-\mathrm{i}\Lambda_\nu r} & Q_\nu d\mathrm{e}^{-\mathrm{i}\Lambda_\nu s} \end{bmatrix} \qquad (8)$$

These approximately represent the instantaneous mixture by the complex matrix. The parameters $a, b, c, d; p, q, r, s$ of mixing matrix $\boldsymbol{K}$ don't depend on the parameter ν representing the center frequency, but their amplitudes P_ν and Q_ν depend on ν through power spectra of two independent signals $W_1(m)$ and $W_2(m)$, respectively.

Ortho-normalization of narrow-banded time-shifted mixtures The index ν may be neglected in the formulation below for simplicity though ν is corresponding to the narrow-banded signals of the ν-th center frequency Λ_ν. Let us introduce the orthogonalization of correlation matrix, to orthogonalize the narrow-banded complex mixtures similar to Principal Component Analysis (PCA). First, let the correlation matrix of narrow-banded complex mixtures be

$$\boldsymbol{R} = \langle \boldsymbol{X}'\boldsymbol{X}'^* \rangle \simeq \boldsymbol{K}\boldsymbol{K}^* \qquad (9)$$

with $A^* = \overline{A}^T$:Hermite conjugate.

Suppose that $\boldsymbol{R}$ is orthogonalized by a unitary matrix $\boldsymbol{U}$ as:

$$\boldsymbol{U}^*\boldsymbol{R}\boldsymbol{U} = \boldsymbol{\sigma}^2, \boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}, \boldsymbol{\sigma}^2 = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \qquad (10)$$

The solution of eigenvalue problem yields $\boldsymbol{U}$ and $\boldsymbol{\sigma}^2$, and then by the transformation

$$\boldsymbol{X} = \boldsymbol{\sigma}^{-1}\boldsymbol{U}^*\boldsymbol{X}' \qquad (11)$$

the ortho-normalized (standardized) narrow-banded complex mixtures $\boldsymbol{X}$ can be obtained, where $\langle \boldsymbol{X}\boldsymbol{X}^* \rangle = \boldsymbol{I}$.

2.2. ICA for narrow-banded mixtures (summary)[15]

Next, the ICA technique on complex mixtures is summarized to separate the ortho-normalized and narrow-banded mixtures $\boldsymbol{X}$.

Complex Hermite moment For a normalized complex random variable Z,

$$M_{mn} \equiv \langle H_{mn}(Z, \overline{Z}) \rangle, \quad m, n = 0, 1, 2, \cdots \qquad (12)$$

is called complex Hermite moment with respect to Z. Here, $H_{mn}(Z, \overline{Z})$ denotes a complex Hermite polynomial [24, 25], and later only $H_{nn}(z, \overline{z})$ is used, which of lower orders are:

$$H_{00} = 1, \ \ H_{11}(z, \overline{z}) = |z|^2 - 1 \qquad (13)$$
$$H_{22}(z, \overline{z}) = |z|^4 - 4|z|^2 + 2 \qquad (14)$$
$$H_{33}(z, \overline{z}) = |z|^6 - 9|z|^4 + 18|z|^2 - 6 \qquad (15)$$

When Z is normalized, the Hermite moment with order lower than 2 is always 0: $\langle H_{mn}(Z) \rangle = 0 (m + n = 1, 2)$. All of the complex Hermite moments are 0 excerpt for $m = n = 0$, in a special case where Z is a complex Gaussian variable. This is because of the orthogonality of complex Hermite polynomial.

Unitary transformation of complex signal First, let a 2D unitary matrix be denoted by $\boldsymbol{U}(g) \in SU(2)$. Here, the totality $SU(2)$ of 2D unitary transformations satisfying det$\boldsymbol{U} = 1$ forms a group, which we denote by G for short. Let $g \in G$ be an element of G, and let the matrices corresponding to g, e (e: identity) be denoted by $\boldsymbol{U} = \boldsymbol{U}(g), \boldsymbol{U}(e) = \boldsymbol{I}$. We note the inversion, $g_1 g_2 \to \boldsymbol{U}(g_1)\boldsymbol{U}(g_2)$, $g^{-1} \to \boldsymbol{U}^{-1}(g) = \boldsymbol{U}^*(g)$.

By a unitary matrix $\boldsymbol{U}(g)$, the signal $\boldsymbol{X}$ is transformed as

$$\boldsymbol{X}(g) = \boldsymbol{U}(g)\boldsymbol{X}, \qquad (16)$$
$$X_\nu(g) = \sum_{\mu=1}^{2} u_{\nu\mu}(g) X_\mu, \ (X_\mu : \text{observed mixture}) \qquad (17)$$

where $\boldsymbol{Z}(g)$ designates explicitly that it is a function of g.

We represent the matrix $\boldsymbol{U}(g)$ in terms of the three parameters (α, β, γ)[26]. Hence, we also represent the matrix and $g = (\alpha, \beta, \gamma)$ by $\boldsymbol{U}(g) = \boldsymbol{U}(\alpha, \beta, \gamma)$,

$$\boldsymbol{U}(\alpha, \beta, \gamma) \equiv \boldsymbol{A}(\gamma)\boldsymbol{B}(\beta)\boldsymbol{A}(\alpha), \qquad (18)$$

$$\boldsymbol{A}(\alpha) \equiv \begin{bmatrix} \mathrm{e}^{\mathrm{i}\frac{\alpha}{2}} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\frac{\alpha}{2}} \end{bmatrix}, \boldsymbol{B}(\beta) \equiv \begin{bmatrix} \cos\frac{\beta}{2} & \mathrm{i}\sin\frac{\beta}{2} \\ \mathrm{i}\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{bmatrix} \qquad (19)$$

with $0 \le \alpha < 2\pi,\ \ 0 \le \beta < \pi,\ \ -2\pi \le \gamma < 2\pi$.

Evaluation function for complex mixtures Using the Hermite moment, let us denote an evaluation function with respect to the signals $X_\nu(g)$ transformed from the observations $\boldsymbol{X}$ as:

$$\mathcal{Q}_{22}(g) = \sum_{\nu=1}^{2} \langle H_{22}(X_\nu(g), \overline{X}_\nu(g)) \rangle^2$$
$$= \sum_{\nu=1}^{2} \langle [|X_\nu(g)|^4 - 4|X_\nu(g)|^2 + 2] \rangle^2 \qquad (20)$$

This is a function of unitary transformation $g = (\alpha, \beta, \gamma)$.

By maximizing $\mathcal{Q}_{22}(g)$ with respect to g: $\boldsymbol{U}(g_m)$, the separation matrix in ICA can be obtained, and the maximization can be made by a gradient method.

Estimation of mixing matrix and indeterminacy Let $\boldsymbol{U}(g_0)$, $g_0 = (\alpha_0, \beta_0, \gamma_0)$ be a unitary **mixing matrix** (unknown) defining the orthonormalized mixtures, the **separation matrix** separating them into the independent signals are given in a form as:

$$\boldsymbol{U}(gg_0) = \boldsymbol{U}(g)\boldsymbol{U}(g_0) \qquad (21)$$
$$= \boldsymbol{A}(\gamma)\boldsymbol{B}(\beta)\boldsymbol{A}(\alpha + \gamma_0)\boldsymbol{B}(\beta_0)\boldsymbol{A}(\alpha_0) \qquad (22)$$

Since the first and the last matrices $\boldsymbol{A}(\gamma)$ and $\boldsymbol{A}(\alpha_0)$ are orthogonal, they don't contribute to the absolute value of Eq.(22). In the estimation process of $\boldsymbol{U}(g)$, γ may be set to 0 and the angle α_0 of the final matrix cannot be estimated from g_m. This means that the phase factor is not related with the component separation.

Thus, g_m can be searched for with respect to two variables $g = (\alpha, \beta, 0)$ as:

$$\boldsymbol{U}(gg_0) = \boldsymbol{U}(g)\boldsymbol{U}(g_0) = \pm\boldsymbol{A}(\alpha_0) = \pm \begin{bmatrix} e^{i\frac{\alpha_0}{2}} & 0 \\ 0 & e^{-i\frac{\alpha_0}{2}} \end{bmatrix}$$
$$= \pm i\boldsymbol{I}'\boldsymbol{A}(\alpha_0) = \pm i \begin{bmatrix} 0 & e^{-i\frac{\alpha_0}{2}} \\ e^{i\frac{\alpha_0}{2}} & 0 \end{bmatrix} \qquad (23)$$
$$\boldsymbol{U}(g_0) = \pm\boldsymbol{U}^*(g)\boldsymbol{A}(\alpha_0), \pm\, i\boldsymbol{U}^*(g)\boldsymbol{I}'\boldsymbol{A}(\alpha_0) \qquad (24)$$

Estimation of separation matrix and narrow-banded separated signal The estimation of $\boldsymbol{U}(g_m)$ yields the objective (narrow-banded) separated signals $\boldsymbol{V}$ as:

$$\boldsymbol{V} = \boldsymbol{X}(g_m) \equiv \boldsymbol{U}(g_m)\boldsymbol{X} = \boldsymbol{U}(g_m)\boldsymbol{\sigma}^{-1}\boldsymbol{U}^*\boldsymbol{X}' \qquad (25)$$

with $\boldsymbol{U}^* \equiv \overline{\boldsymbol{U}}^T$. A mentioned above, we see that only (α, β) of $g_m = g = (\alpha, \beta, \gamma)$ in $\boldsymbol{U}(g_m)$ is determined upon maximization and that γ remains arbitrarily or uncertainly. This corresponds to the natural indeterminacy that we can arbitrarily choose a phase of independent complex variables. That is,

$$\boldsymbol{U}(g_m) = \boldsymbol{U}(\alpha, \beta, 0) = \boldsymbol{B}(\beta)\boldsymbol{A}(\alpha) \qquad (26)$$
$$\boldsymbol{U}(\alpha, \beta, \gamma) = \boldsymbol{A}(\gamma)\boldsymbol{U}(g_m) = \boldsymbol{A}(\gamma)\boldsymbol{B}(\beta)\boldsymbol{A}(\alpha) \qquad (27)$$

$\boldsymbol{U}(g_m)$ denotes a part determined by maximization and $\boldsymbol{A}(\gamma)$ the matrix representing the indeterminate phase factor $e^{\pm i\gamma/2}$ of independent components. Figure 1 shows a flow diagram of the blind source separation for complex signals. The real observations are first narrow-banded, orthogonalized and then transformed by a unitary matrix.

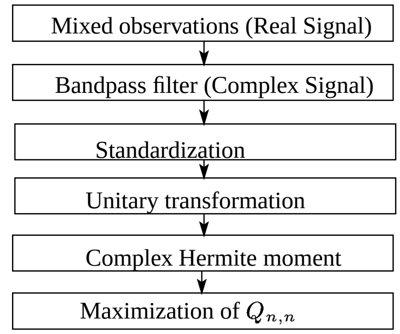

Fig. 1. A flow of blind source separation for convolved signals.

2.3. Estimation of mixing matrix and sound source direction

Estimation of mixing matrix The separated signals $\boldsymbol{V}$ can be calculated from the narrow-banded (mixed) observations $\boldsymbol{X}'$ by Eq.(25). Thus, inversely, we can estimate the mixing matrix $\boldsymbol{K}$ defined in Eq.(7) and transforms the independent signals $\boldsymbol{W}$ into the observations $\boldsymbol{X}'$. That is, from the relation:

$$\boldsymbol{V} = \boldsymbol{A}(\gamma)\boldsymbol{U}(g_m)\boldsymbol{\sigma}^{-1}\boldsymbol{U}^*\boldsymbol{X}' \simeq \boldsymbol{A}(\gamma)\boldsymbol{U}(g_m)\boldsymbol{\sigma}^{-1}\boldsymbol{U}^*\boldsymbol{K}\boldsymbol{W}, \qquad (28)$$

by considering the indeterminacy of ICA solutions, it follows that

$$\boldsymbol{A}(\gamma)\boldsymbol{U}(g_m)\boldsymbol{\sigma}^{-1}\boldsymbol{U}^*\boldsymbol{K} = \pm\boldsymbol{I}, \mp i\boldsymbol{I}' \qquad (29)$$

Accordingly, letting $\boldsymbol{M}$ be a part which can be estimated, the relation to the original mixing matrix $\boldsymbol{K}$ can be expressed as follows:

$$\boldsymbol{M} \equiv \boldsymbol{U}\boldsymbol{\sigma}\boldsymbol{U}^*(g_m) = \pm\boldsymbol{K}\boldsymbol{A}^*(\gamma) \ \text{ or } \ \pm i\boldsymbol{K}\boldsymbol{I}'\boldsymbol{A}^*(\gamma) \qquad (30)$$

The left-hand-side of the above equation is known by PCA and ICA. Matrices $\pm\boldsymbol{I}, \pm\boldsymbol{I}'$ and $\boldsymbol{A}^*(\gamma)$ represent the indeterminacy in the estimation process. By substituting $\boldsymbol{K}$ into the above equation, the **estimated mixing matrix**$\boldsymbol{M}$ can be rewritten as:

$$\boldsymbol{M} \equiv \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \equiv \begin{bmatrix} |M_{11}|e^{i\phi_{11}} & |M_{12}|e^{i\phi_{12}} \\ |M_{21}|e^{i\phi_{21}} & |M_{22}|e^{i\phi_{22}} \end{bmatrix} \qquad (31)$$
$$= \pm \begin{bmatrix} P_\nu a e^{-i\Lambda_\nu p - i\gamma/2} & Q_\nu b e^{-i\Lambda_\nu q + i\gamma/2} \\ P_\nu c e^{-i\Lambda_\nu r - i\gamma/2} & Q_\nu d e^{-i\Lambda_\nu s + i\gamma/2} \end{bmatrix} \quad \text{or}$$
$$\pm\, i \begin{bmatrix} Q_\nu b e^{-i\Lambda_\nu q + i\gamma/2} & P_\nu a e^{-i\Lambda_\nu p - i\gamma/2} \\ Q_\nu d e^{-i\Lambda_\nu s + i\gamma/2} & P_\nu c e^{-i\Lambda_\nu r - i\gamma/2} \end{bmatrix} \qquad (32)$$

This is a mixing matrix for a narrow-banded signals of center frequency Λ_ν.

Estimation on ratios of a, b, c and d From entries of $\boldsymbol{M}$ in Eq.(32), the ratios of parameters in the mixing matrix $\boldsymbol{K}$ can be estimated, excerpt for the indeterminacy, as follows:

$$\frac{M_{11}}{M_{21}} \equiv \frac{|M_{11}|}{|M_{21}|} e^{i(\phi_{11} - \phi_{21})} = \frac{a}{c} e^{i(r-p)\Lambda_\nu}, \text{ or } \frac{b}{d} e^{i(s-q)\Lambda_\nu} \qquad (33)$$
$$\frac{M_{12}}{M_{22}} \equiv \frac{|M_{12}|}{|M_{22}|} e^{i(\phi_{12} - \phi_{22})} = \frac{b}{d} e^{i(s-q)\Lambda_\nu}, \text{ or } \frac{a}{c} e^{i(r-p)\Lambda_\nu} \qquad (34)$$

This means that the amplitude ratios a/c and b/d can be obtained from the absolute values of ratios on entries, and that the time delays $r - p$ and $s - q$ from the arguments of $e^{i(r-p)\Lambda_\nu}$ and $e^{i(r-p)\Lambda_\nu}$.

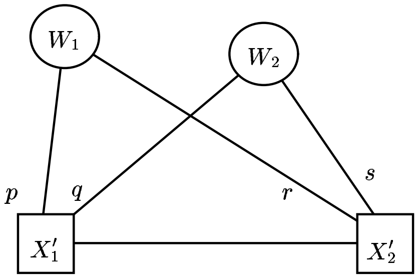

Fig. 2. Sound sources and observation points.

The amplitude ratios take constant values for point sources and are evaluated as an average of estimated values at each center frequency Λ_ν.

Estimation of time delays $p-r$ and $q-s$ Since generally a phase of matrix component M_{ij} has indeterminacy on the angle with 2π, the time delay cannot be evaluated even if the frequency Λ_ν is given. Now, the following linear relationship on the phase arg$M=\phi$ of the ratio on two matrix components can be utilized:

$$\phi = \Lambda_\nu P, \quad P = p-r,\ q-s, \tag{35}$$

and the time-delay parameters $P=p-r$ and $q-s$ corresponding to the time differences in mixtures can be determined by a slope of linear relationship between the frequency (Λ) and the phase of ratio on matrix components. More concretely, changing slightly the center frequency of bandpass filter (e.g., Λ_ν, $\Lambda_\nu+\delta\Lambda$, $\Lambda_\nu+2\delta\Lambda$, $\Lambda_\nu+3\delta\Lambda, \cdots$), it is possible to estimate the optimal value of slope from the linear relationship of phase.

Sound source localization(geometrical relationship between observation points and sound sources) Even if the time delays $r-p$ and $s-q$ are estimated, the geometrical position of 2 observation points and 2 sound sources (as shown in Fig.2) cannot be directly obtained. By considering the amplitude ratios a/c and b/d representing the magnitudes of sound sources, the geometrical relationship between observation points and sound sources can be evaluated as below.

Letting g_1 and g_2 be gains of sound sources, the decay inversely proportional to the distance yields:

$$a = \frac{1}{p}g_1,\ b=\frac{1}{q}g_2,\ c=\frac{1}{r}g_1,\ d=\frac{1}{s}g_2 \tag{36}$$

$$\frac{a}{c} = \frac{r}{p} = J_1 \text{ (const.)};\quad \frac{b}{d}=\frac{s}{q}=J_2 \text{ (const.)} \tag{37}$$

Thus, from the relations $p-r=K_1$, $q-s=K_2$, $r=J_1p$ and $s=J_2q$, we get

$$p=\frac{K_1}{1-J_1},\ r=\frac{K_1J_1}{1-J_1},\ q=\frac{K_2}{1-J_2},\ s=\frac{K_2J_2}{1-J_2} \tag{38}$$

and then the absolute values of p, q, r and s can be estimated. This method is applicable to a case when the observation points are located close to the sound sources. On the contrary, when the distance is very long, our method can estimate only the directions of sound sources.

3. INVERSE FILTER FOR TIME-SHIFTED MIXTURES

The response function of filter representing the time-shifted mixtures is expressed as follows:

$$\boldsymbol{A}(z) = \begin{bmatrix} az^p & bz^q \\ cz^r & dz^s \end{bmatrix} \tag{39}$$

The inverse of the above matrix can be written as:

$$\boldsymbol{H}(z) = \boldsymbol{A}^{-1}(z) = \frac{1}{\Delta(z)}\begin{bmatrix} dz^s & -bz^q \\ -cz^r & az^p \end{bmatrix} \tag{40}$$

$$\Delta(z) = \det \boldsymbol{A}(z) = adz^{p+s} - bcz^{q+r} \tag{41}$$

To derive a stable iterative algorithm for an inverse filter, it is necessary to locate the roots of $\Delta(z)$ outside a unit circle. There are two cases satisfying the above condition:

1. for a case where $\Delta = ad-bc>0$, $bc/ad<1$, $(r+q)-(p+s)>0$
2. for a case where $\Delta = ad-bc<0$, $bc/ad>1$, $(p+s)-(r+q)>0$

with $\Delta \equiv \Delta(1) = ad-bc$.

The relationship between input $\boldsymbol{X}'$ and output $\boldsymbol{V}$ of the inverse filter can be written as $\boldsymbol{V}(z) = \boldsymbol{H}(z)\boldsymbol{X}'(z)$. It can be rewritten as input-output relationship in time domain:

1. for a case where $\Delta = ad-bc>0$, $bc/ad<1$,

$$\begin{aligned} &V_1(m) - \frac{bc}{ad}V_1(m-(r+q)+(p+s)) \\ &= \frac{1}{ad}\left[dX_1'(m+p) - bX_2'(m+p+s-q)\right] \end{aligned} \tag{42}$$

$$\begin{aligned} &V_2(m) - \frac{bc}{ad}V_2(m-(r+q)+(p+s)) \\ &= \frac{1}{ad}\left[-cX_1'(m+p+s-r) + aX_2'(m+p)\right] \end{aligned} \tag{43}$$

2. for a case where $\Delta = ad-bc<0$, $bc/ad>1$,

$$\begin{aligned} &V_1(m) - \frac{ad}{bc}V_1(m-(r+q)+(p+s)) \\ &= -\frac{1}{bc}\left[dX_1'(m+r+q-s) - bX_2'(m+r)\right] \end{aligned} \tag{44}$$

$$\begin{aligned} &V_2(m) - \frac{ad}{bc}V_2(m-(r+q)+(p+s)) \\ &= -\frac{1}{bc}\left[-cX_1'(m+q) + aX_2'(m+r+q-p)\right] \end{aligned} \tag{45}$$

By using the above equations as recurrence formula with time-shifted $X_1'(m)$, $X_2'(m)$, and V_1, V_2 in the past time stage, we can produce iteratively $V_1(m)$ and $V_2(m)$.

4. SIMULATION WITH SPEECH DATA

4.1. Blind separation for time-shifted mixtures

To confirm the effectiveness of our method, it is applied to the actual speech data. As shown in Fig.3, male and female speeches (W_1, W_2) are utilized as sound sources (length of data: 70000, sampling frequency: 16kHz).

Mixing these source signals as:

$$\begin{cases} X_1'(m) = W_1(m-60) + W_2(m-40) \\ X_2'(m) = W_1(m-40) + W_2(m-50) \end{cases} \tag{46}$$

and setting the distance between two microphones to 40 steps (i.e., the true angles of sound sources are -30 ° and 14.5 °), the observations with time differences are given as X_1' and X_2' (See Fig.4). This mixing corresponds to the actual situation when two sound sources are located far from microphones.

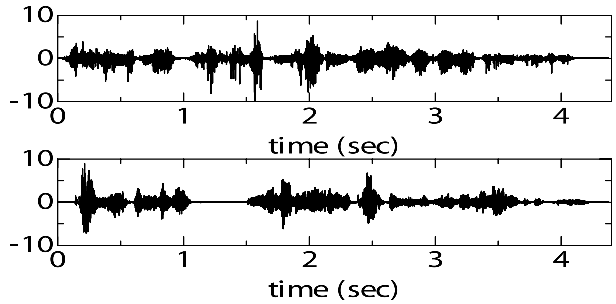

Fig. 3. Source signals.

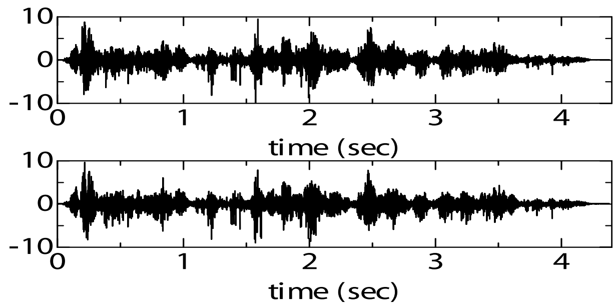

Fig. 4. Mixed signals.

First, let us transform the mixture $X_j'(m)$ $(j = 1, 2)$ into the complex signal $X_{j\nu}'(m)$, by using the complex filter of 1st order with the center angular frequency Λ_ν(rad) as in Eq.(3).

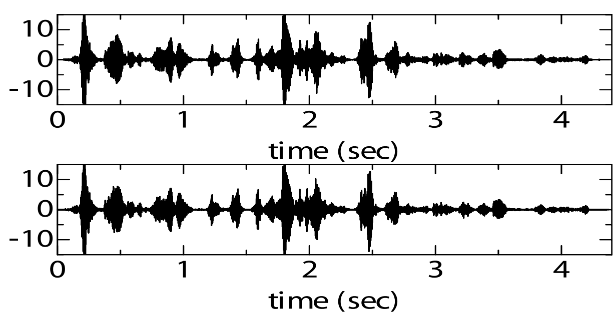

Fig. 5. Narrow-banded observations (f_0=1000Hz).

In this simulation, the complex ICA has been applied to cases with center frequency every 100Hz from 1,000Hz to 2,000Hz. Figure 5 shows a real part of mixtures narrow-banded by the filter with its center frequency f_0 =1,000Hz and the bandwidth parameter $K = 0.01$.

To standardize the narrow-banded observations, PCA has been adopted. As shown in Fig.6, the unknown parameters α and β are estimated by applying the gradient method to the evaluation function $\mathcal{Q}_{22}(g)$. Both of parameters α and β converge to constant values. As mentioned above, the parameter γ is not estimated, because it may take an arbitrary value.

By using these estimated parameters of unitary transformation, the separated signals are constructed from the mixed observations. Figure 7 shows the real part of separated signals. From this result, the previously proposed complex ICA demonstrates a good performance on the separation by comparing the waveforms between the original and separated signals. However, only through the complex ICA, we must further make the bandwidth of separated signals broad, e.g., introducing a number of bandpass filters.

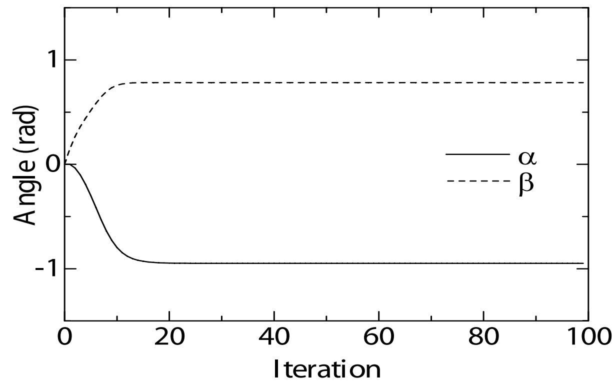

Fig. 6. Estimation process for unknown parameters α and β (f_0=1000Hz).

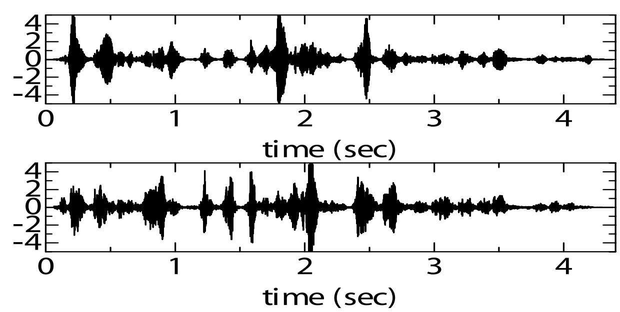

Fig. 7. Separated signals (f_0=1000Hz).

4.2. Parameter estimation

The complex ICA is applied to the data as an output of narrow-banded filter at each center frequency of 1,000Hz to 2,000Hz. The amplitude ratio and the phase difference are calculated using estimated parameters α and β as shown in Fig.8. The amplitude ratio almost takes a constant value for a plane wave though it must ideally be 1. The slopes of phase difference are evaluated by Least Squares Error Method as:

$$-7.5 \times 10^{-3}(\mathrm{rad/Hz}), \quad 4.0 \times 10^{-3}(\mathrm{rad/Hz})$$

These are respectively corresponding to the time steps as:

$$-7.5 \times 10^{-3} \times \frac{16000}{2\pi} = -19.86, \;\; 4.0 \times 10^{-3} \times \frac{16000}{2\pi} = 10.19$$

From these time steps corresponding to the phase differences of embedded signals, we can easily determine the arrival directions of sounds as -29.8 ° and 14.8 °. This method has an advantage such that it can estimate the directions of two wide-band sound sources by only two microphones. In the actual situation of acoustic measurement, however, there remains a future problem to be solved, e.g., the reverberant mixing, etc.

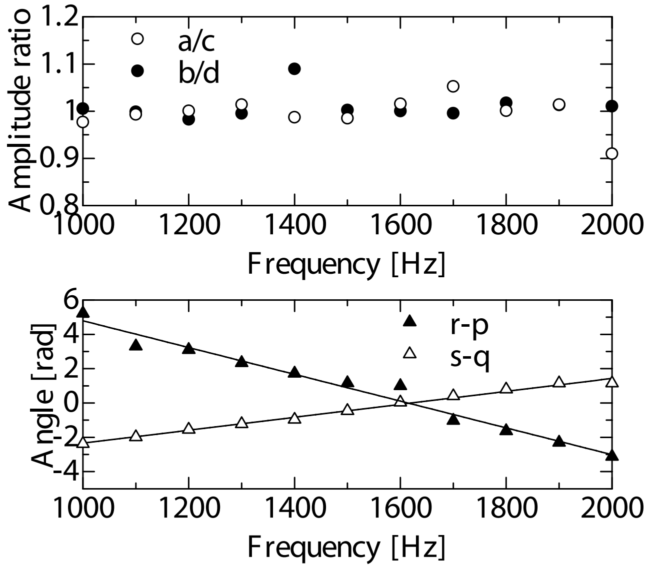

Fig. 8. Estimated results of amplitude ratio (Top) and phase difference (Bottom) for each center frequency of narrow bandwidth.

4.3. Inverse filtering

Using the estimated result of the phase difference for each separated signal, the mixed signals are separated through an inverse filter as shown in Fig.9. It should be noticed that though the usual complex ICA requires the synthesis of signals separated at each frequency, this method only needs the overall mixed signals. From these figures, the waveforms of inverse filtering voices seem similar to those of original voices.

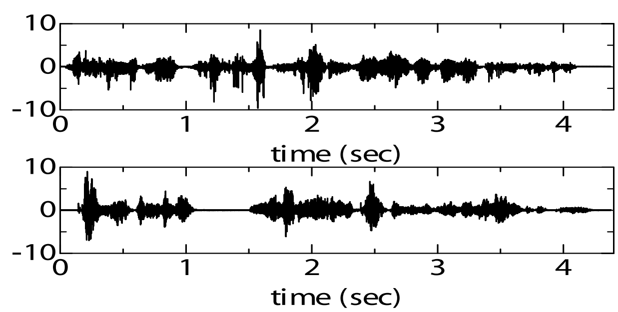

Fig. 9. Separated signals through an inverse filter with use of the estimated time-difference.

5. REFERENCES

[1] J.F.Cardoso; "Source separation using higher order moments", *Proc.IEEE ICASSP*, Vol.4, pp.2109-2112 (1989).

[2] C.Jutten and J.Herault, "blind separation of sources, part I," *Signal Processing'*, vol.24, pp.1-10 (1991).

[3] P.Comon; "Independent component analysis, A new concept ?", *Signal Processing* Vol.36, p.287-314 (1994).

[4] A.J.Bell abd T.J.Sejnowski, "A information maximization approach to blind separation and blind deconvolution", *Neural Computation*, vol.7, pp.1129-1159 (1995).

[5] H.Ogura and N.Nakasako; "Direct method of blind separation for two mixed signals based on the statistical independency from information theoretic approach", *Proc. ISNIC-98*, pp.215-220 (1998).

[6] N.Nakasako and H.Ogura; "Independent component analysis of three mixed signals based on information theory", *Proc. ISNIC-98*, pp.221-226 (1998).

[7] H.Ogura, N.Nakasako and N.Séo; "On information-theoretic blind separation of mixed signals - effect of probability distributions and additional noise-", *Proc. 6th ICSV*, pp.2757-2764 (1999).

[8] H.Ogura and N.Nakasako ; "Independent component analysis and its applications to signal processing", *Proc. 2001 Intl. Conf. CSCC*, vol.1, pp.656-659 (2001).

[9] A. Cichoki, S. Amari, and J. Cao, *Proc. 1996 Intl. Sympo. on Nonlinear Theory and Its Applications*, pp.229-232, Kochi, Oct. (1996).

[10] T.W. Lee, A.J. Bell, and R.H. Lambert, *Advances in Neural Information Processing Systems 9*, pp.758-764, MIT Press, Cambridge MA (1997).

[11] K. Torkkola, in IEEE Workshop on Neural Networks for Signal Processing VI, pp.423-432, Kyoto, Japan, Sept. (1996).

[12] D.C.B. Chans, S.J. Godsill, and P.J.W. Rayner, NIPS'96 Workshop on Blind Signal Processing and Their Applications, Denver, CO, U.S.A., Dec. (1996).

[13] F. Ehlers and H.G. Schuster, *IEEE Trans. Signal Processing*, vol.45, no.10, pp.2608-2612, oct (1997).

[14] M. Kawamoto, K. Matsuoka, and N. Ohnishi, *Proc. Intl. Sympo. on Nonlinear Theory and Its Applications*, pp.1001-1004, Hawaii, Dec. (1997).

[15] H.Ogura, N.Nakasako and N.Séo; "Blind separation of complex signals in terms of complex Hermite moments", *Proc. ISPACS 2000*, vol.I, pp.428-433 (2000).

[16] N.Nakasako and H.Ogura; "Blind separation of time sifted mixed speeches", *Proc. 2001 Intl. Conf. CSCC*, vol.1, pp.660-663 (2001)

[17] N.Nakasako, H.Ogura and H.Kuruuchi; "A complex signal ICA based on unitary transformation and complex Hermite moment for speech separation", *Proc. 8th ICSV*, pp.2453-2460 (2001).

[18] R.O. Schmidt; "Multiple emitter location and signal parameter estimation", Proc. RDAC Spectrum Estimation Workshop, pp.243-256 (1979).

[19] M. Wax, T.-J. Shan and T. Kailath; "Spatio-temporal spectral analysis by eigenstructure methods", *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-32, 817-827 (1984).

[20] G. Su and M. Morf; "The signal subspace approach for multiple wide-band emitter location", *IEEE trans. Acoust. Speech Signal Process.*, vol.ASSP-31, 1502-1522 (1983).

[21] J.A. Cadzow; "A high resolution of direction-of-arrival algorithm for narrow-band coherent and incoherent sources", *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-36, 965-979 (1988).

[22] K.M. Buckley and L. Griffi ths;"Broad-band signal subspace spatial-spectrum (BASS-ALE) estimation", *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-36, 953-964 (1988).

[23] H.Ogura, N.Nakasako and Y.Sakuma; "Blind separation of mixed speech with time delays – Estimation of source parameters and separation by an inverse fi lter–",*Proc. ISPACS 2002*, pp.123-127 (2002).

[24] K.Ito: Complex multiple Wiener integral; Japan Math. Soc. Japan, Vol.22, pp.63-86 (1952).

[25] H. Ogura: Theory of Stochastic Process (in Japanese), Chap.5, Corona Publ. (1978).

[26] N.Y.Vilenkin: Special Function and the Theory of Group Representations, Vol.22, Transaction of Mathematical Monographs, (translated from Russian) American Mathematical Society, Chap.IX (1968).