DIRECTION OF ARRIVAL ESTIMATION FOR MULTIPLE SOURCE SIGNALS USING INDEPENDENT COMPONENT ANALYSIS

Hiroshi Sawada *Ryo Mukai* *Shoji Makino*

NTT Communication Science Laboratories, NTT Corporation
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan
{sawada,ryo,maki}@cslab.kecl.ntt.co.jp

ABSTRACT

This paper presents a new method for estimating the directions of source signals. We assume a situation in which multiple source signals are mixed in a reverberant condition and observed at several sensors. The new method is based on independent component analysis, which separates mixed signals into original source signals. It can be applied where the number of sources is equal to the number of sensors, whereas the conventional methods based on subspace analysis, such as the MUSIC algorithm, are applicable where there are fewer sources than sensors. Even in cases where the MUSIC algorithm can be applied, the new method is better at estimating the directions of sources if they are closely placed.

1. INTRODUCTION

Direction of arrival (DOA) estimation is a basic and important technique in array signal processing not only for wireless communication systems, but also for audio/speech processing systems [1, 2]. MUSIC (MUltiple SIgnal Classification) [3] and its variants are popular methods for DOA estimation. They are based on subspace analysis, and can be applied even for mixed signals as long as there are fewer source signals than sensors.

This paper presents a new method for DOA estimation based on independent component analysis (ICA) [4], which identifies original source signals from their mixtures. The main advantage of the ICA-based method over the MUSIC algorithm is that it can be applied even when the number of sources is equal to the number of sensors.

The method for DOA estimation based on ICA is a by-product of research on blind source separation (BSS) for convolutive mixtures in a reverberant environment. If the mixture is instantaneous, the direct application of a simple ICA solves the BSS problem. However, ICA becomes difficult for convolutive mixtures because we need a matrix of filters, not just a matrix of scalars, to separate convolutively mixed signals. One major approach is frequency-domain BSS, where convolutive mixtures are reduced to instantaneous mixtures in every frequency bin so that a simple ICA can be applied. The problem occurred instead is a permutation problem [5]. We need to align the source signal order in each frequency bin since ICA is ambiguous with regard to the permutation of sources. One approach to the permutation problem is to estimate the DOA of each source, and sort the sources based on the directions.

A method has been proposed [6, 7] for estimating DOAs by plotting the directivity patterns of the separating system obtained by ICA. By analyzing the directivity patterns, we can estimate the directions of sources suppressed at each output, and therefore the direction of each source. However, it is difficult to perform such an analysis for more than two sources, as discussed in Sec. 3.2. In practice, this method can only be applied when there are two sources.

Instead of plotting directivity patterns, the new method described in this paper is based on estimating the mixing system from the separating system obtained by ICA. It can, in principle, be applied when there are more than two sources. Experimental results show that the new method succeeded in estimating the DOAs for a situation where three sources were mixed and observed at three sensors. Neither the MUSIC algorithm nor the method in [6, 7] is applicable for such a situation. Experimental results also revealed another advantage even when there are fewer sources than sensors. When sources were closely placed, the MUSIC algorithm failed to estimate their DOAs, whereas the proposed method succeeded.

2. PROBLEM FORMULATION

Suppose that P source signals $s_p(t)$ are mixed and observed at Q sensors $x_q(t) = \sum_{p=1}^{P} \sum_k h_{qp}(k) s_p(t-k)$, where $h_{qp}(k)$ represents the impulse response from source p to sensor q. Let d_q be the position of sensor q (we assume a linearly arranged array of sensors), and θ_p be the direction of source s_p (the direction orthogonal to the array is 90°). The task is estimating the directions $\theta_1, \ldots, \theta_P$ of source signals from the sensor observations $x_1(t), \ldots, x_Q(t)$.

Most DOA estimation techniques are performed in the frequency domain. By L-point short time DFT, time-domain signals $x_q(t)$ are converted into frequency-domain time-series signals $X_q(f, m)$, where $f = 0,\ f_s/L,\ \ldots,\ f_s(L-1)/L$ (f_s: sampling frequency), and m is the frame index. The mixing model above is represented as $X_q(f, m) = \sum_{p=1}^{P} H_{qp}(f) S_p(f, m)$ in the frequency domain. Although the signals are mixed in a reverberant condition, we can approximate the frequency response $H_{qp}(f)$ of an impulse re-

sponse $h_{qp}(k)$ as

$$H_{qp}(f) = e^{j2\pi f c^{-1} d_q \cos\theta_p}, \quad (1)$$

where c is the propagation velocity. If we regard θ_p as a variable θ, we have a steering vector

$$\mathbf{a}(f,\theta) = [e^{j2\pi f c^{-1} d_1 \cos\theta}, \ldots, e^{j2\pi f c^{-1} d_Q \cos\theta}]^T.$$

Then, the sensor observations can be modeled as $\mathbf{X}(f,m) = \sum_{p=1}^{P} \mathbf{a}(f,\theta_p) S_p(f,m)$, where $\mathbf{X}(f,m)$ is a Q-dimensional vector $\mathbf{X}(f,m) = [X_1(f,m), \ldots, X_Q(f,m)]^T$.

3. PREVIOUS WORK

3.1. The MUSIC algorithm

Let us review the procedure of the MUSIC algorithm [3]. The correlation matrix $\mathbf{R} = \langle \mathbf{X}(f,m) \cdot \mathbf{X}(f,m)^H \rangle_m$ of sensor observations $\mathbf{X}(f,m)$ is calculated, where $\cdot^H$ represents a conjugate transpose and $\langle \cdot \rangle_m$ denotes the averaging operator. Then, obtain the eigenvalue decomposition $\mathbf{R} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_Q]$, $\mathbf{\Lambda} = diag[\lambda_1, \ldots, \lambda_Q]$, where $\mathbf{v}_k$ is an eigenvector (Q-dimensional column vector) and λ_k is the eigenvalue of $\mathbf{v}_k$ sorted as $\lambda_1 \geq \cdots \geq \lambda_Q$. The P points where the function $U(\theta) = \sum_{k=P+1}^{Q} |\mathbf{v}_k^H \mathbf{a}(f,\theta)|^2$ approaches zero correspond to the directions $\theta_1, \ldots, \theta_P$ of the source signals.

Here we explain the mechanism of the MUSIC algorithm. If the number P of sources is smaller than the number Q of sensors, all the signal components are represented in the signal subspace spanned by the first P eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_P$, and the remaining $Q - P$ eigenvectors $\mathbf{v}_{P+1}$, $\ldots$, $\mathbf{v}_Q$ represents the noise subspace. The signal subspace and the noise subspace are orthogonal to each other since they are spanned by different eigenvectors. The subspace spanned by the P steering vectors $\mathbf{a}(f,\theta_1), \ldots, \mathbf{a}(f,\theta_P)$ is also the signal subspace because the sensor observation are represented as $\mathbf{X}(f,m) = \sum_{p=1}^{P} \mathbf{a}(f,\theta_p) S_p(f,m)$. When θ coincides with one of the source directions $\theta_1, \ldots, \theta_P$, the steering vector $\mathbf{a}(f,\theta)$ and the noise subspace $\mathbf{v}_{P+1}, \ldots, \mathbf{v}_Q$ are orthogonal, and therefore $U(\theta)$ approaches zero. This is why source directions can be estimated using $U(\theta)$. For the noise subspace to exist, the number Q of sensors should be larger than the number P of sources. Thus, the MUSIC algorithm is applicable for mixtures of up to $Q - 1$ signals.

3.2. Directivity Patterns formed by an ICA Solution

Next, we review the method [6, 7] for estimating DOAs by plotting the directivity patterns of the separating system in the context of frequency-domain BSS. The frequency responses $\mathbf{W}(f)$ of the separating system are obtained by solving the ICA problem $\mathbf{Y}(f,m) = \mathbf{W}(f)\mathbf{X}(f,m)$, where $\mathbf{Y}(f,m) = [Y_1(f,m), \ldots, Y_P(f,m)]^T$ is the vector of separated signals and $\mathbf{W}(f)$ is a $P \times Q$ separating matrix whose elements are $W_{rq}(f)$. Let the r-th row of $\mathbf{W}(f)$ be $\mathbf{w}_r(f) = [W_{r1}(f), \ldots, W_{rQ}(f)]$. Each $\mathbf{w}_r(f)$ produces $Y_r(f,m)$, which corresponds to one of the source signals $S_1(f,m), \ldots, S_P(f,m)$. Therefore, each $\mathbf{w}_r(f)$ extracts a source signal arriving from a specific direction and forms spatial nulls in the direction of the other sources to suppress them. If we plot the directivity pattern formed by $\mathbf{w}_r(f)$, we can see the directions of sources suppressed by $\mathbf{w}_r(f)$.

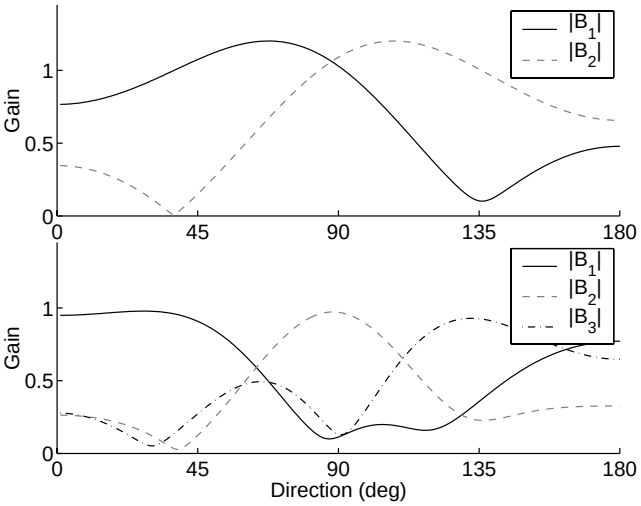

Figure 1: Directivity patterns for two sources (above) and three sources (below)

The directivity pattern formed by $\mathbf{w}_r(f)$ is calculated using the steering vector $\mathbf{a}(f,\theta)$ as $B_r(f,\theta) = \mathbf{w}_r(f)\mathbf{a}(f,\theta)$. Figure 1 shows the gain of directivity patterns at some frequency. The upper part shows the patterns for two sources. We see that $\mathbf{w}_1(f)$ extracts a source signal originating from around 37° and suppresses the other signal from around 135°. With a similar consideration regarding $\mathbf{w}_2(f)$, we estimate source signal directions of 37° and 135°.

Although this method works for two-source cases, it is difficult to use when there are more than two sources. The lower half of Fig. 1 shows directivity patterns for three sources. Here, each signal should be suppressed by two rows. However, the null directions of different rows may not coincide for a specific source signal, and it is difficult to decide the direction of a source signal in such cases. For example, the signal extracted by the first row $\mathbf{w}_1(f)$ is suppressed by the second $\mathbf{w}_2(f)$ and third $\mathbf{w}_3(f)$ rows at 30° and 38°, respectively. Moreover, if the difference between the null directions is large, we cannot decide which null direction corresponds to which source signal.

4. NEW METHOD

4.1. DOA estimation with $\mathbf{W}^{-1}(f)$ obtained by ICA

Instead of plotting directivity patterns and analyzing null directions, we propose a new DOA estimation method by estimating the frequency response $\mathbf{H}(f)$ of the mixing system, whose elements are $H_{qp}(f)$. If a separating matrix $\mathbf{W}(f)$ calculated by ICA successfully extracts source signals, $\mathbf{W}(f)\mathbf{H}(f) = \mathbf{I}$ holds. Thus, we can estimate the frequency response of the mixing system by $\mathbf{H}(f) = \mathbf{W}^{-1}(f)$.

Here, we should take account of the scaling and permutation ambiguities of ICA. Therefore, the $\mathbf{H}(f)$ columns can have arbitrary scales and be permuted arbitrarily compared with the real frequency response of the mixing system.

An element $H_{qp}(f)$ of the matrix $\mathbf{H}(f)$ may have an arbitrary amplitude. Since the approximation (1) of the mixing system does not suit this situation, we remodel the mixing system with attenuation A_{qp} (real-valued) and phase modulation $e^{j\varphi_p}$ at the origin:

$$H_{qp}(f) = A_{qp}\ e^{j\varphi_p} e^{j2\pi f c^{-1} d_q \cos\theta_p}.$$

The scaling ambiguity can be canceled out by calculating the ratio between two elements $H_{qp}(f)$ and $H_{q'p}(f)$ corresponding to the same source p:

$$H_{qp}/H_{q'p} = A_{qp}/A_{q'p}\ e^{j2\pi f c^{-1}(d_q - d_{q'})\cos\theta_p}.$$

Then, taking the angle yields a formula for estimating θ_p:

$$\theta_p = \cos^{-1} \frac{angle(H_{qp}/H_{q'p})}{2\pi f c^{-1}(d_q - d_{q'})}. \quad (2)$$

If the absolute value of the argument of $\cos^{-1}$ is larger than 1, θ_p becomes complex and no direction is obtained. In this case, formula (2) can be tested with another pair q and q'. In principle, this method can be applied for any number of source signals.

Because of the permutation ambiguity, θ_p may not correspond to s_p but to another source signal. However, by calculating θ_p for all $p = 1, \ldots, P$, we can obtain the directions of all source signals.

The new method offers an advantage in terms of computation cost. Estimated directions are provided by the closed form formula (2), whereas the minima of $U(\theta)$ and $|B_r(f,\theta)|$ should be searched for in the MUSIC algorithm and the previous method based on ICA, respectively.

4.2. Equivalence between θ_p and a null direction

For a two-source case, we prove that θ_p calculated by (2) is the same as a null direction that is the minimum of a directivity pattern. When $|B_r(f,\theta)|$ is minimized, θ corresponds to a null direction. Let $\alpha_q = 2\pi f c^{-1} d_q$ and f be omitted. The value to be minimized is

$$\begin{aligned} J(\theta) &= B_r(\theta) \cdot B_r(\theta)^* \\ &= (W_{r1}\, e^{j\alpha_1 \cos\theta} + W_{r2}\, e^{j\alpha_2 \cos\theta}) \cdot \\ &\quad (W_{r1}^*\, e^{-j\alpha_1 \cos\theta} + W_{r2}^*\, e^{-j\alpha_2 \cos\theta}). \end{aligned}$$

Let $\alpha = \alpha_2 - \alpha_1$. The first and second derivatives are

$$\begin{aligned} \tfrac{dJ}{d\theta} &= -\alpha \sin\theta \cdot 2\, \mathit{im}(W_{r1} W_{r2}^* e^{-j\alpha\cos\theta}), \\ \tfrac{d^2J}{d\theta^2} &= -\alpha\cos\theta \cdot 2\, \mathit{im}(W_{r1} W_{r2}^* e^{-j\alpha\cos\theta}) \\ &\quad -\alpha^2 \sin^2\theta \cdot 2\, \mathit{re}(W_{r1} W_{r2}^* e^{-j\alpha\cos\theta}) \end{aligned}$$

where *re* and *im* extract the real and imaginary parts of a complex, respectively. If $angle(W_{r1} W_{r2}^* e^{-j\alpha\cos\theta}) = \pi$, $\frac{dJ}{d\theta}$ is zero and $\frac{d^2J}{d\theta^2}$ is positive, and $J(\theta)$ is minimized. Thus, the null direction formed by the r-th row of $\mathbf{W}$ is given by

$$angle(-W_{r1} W_{r2}^*) = \alpha \cos\theta_r^{null} \quad \Leftrightarrow$$

$$\theta_r^{null} = \cos^{-1} \frac{angle(-W_{r1}/W_{r2})}{2\pi f c^{-1}(d_2 - d_1)}.$$

Table 1: Experimental conditions

Length of mixed signals	6 s
Positions of 3 sensors	0, 56.6, 113.2 mm
Sampling rate	f_s = 8 kHz
Maximum frequency for no-aliasing	3 kHz
Reverberation time	T_R = 190 ms
DFT frame size	$L = 1024$

Considering $H_{21} = -W_{21}/det(\mathbf{W})$ and $H_{11} = W_{22}/det(\mathbf{W})$, we see that θ_1 and θ_2^{null} are the same:

$$\theta_1 = \cos^{-1} \frac{angle(H_{21}/H_{11})}{2\pi f c^{-1}(d_2 - d_1)} = \cos^{-1} \frac{angle(-W_{21}/W_{22})}{2\pi f c^{-1}(d_2 - d_1)} = \theta_2^{null}.$$

5. EXPERIMENTAL RESULTS

We performed experiments to estimate the DOAs of speech signals mixed in a reverberant environment. We used sets of impulse responses in the "RWCP Sound Scene Database in Real Acoustic Environments" [8] to generate mixed signals. The experimental conditions are summarized in Table 1. We used three sensors with an equal spacing of 56.6 mm. Thus, the maximum frequency at which a spatial aliasing did not occur was $340/(0.0566 \times 2) \approx 3$ kHz. The ICA algorithm we used was the complex version of FastICA [4], although we can use any ICA algorithm, such as the information maximization approach [9].

Table 2 is a summary of the results shown in Figs. 2, 3 and 4. We used four sets of impulse responses corresponding to the positions of sources "$\diamond$", "$+$", "$\square$" and "$\times$". Although the RWCP database provides information about the directions of the positions, it is difficult to know the true directions. Thus, we estimated the DOA of each source with the MUSIC algorithm when there was only one source signal. They appear in the "MUSIC each" row of Table 2. We consider these directions to be the true directions.

Figure 2 shows the results when three source signals are mixed and observed at three sensors. Here, only the proposed method based on ICA is applicable. Each plot shows the direction of each source estimated in each frequency bin. The correspondence between a direction and a source signal is solved simply by sorting the three directions in each frequency. There are some frequency bins where the DOAs were not estimated because formula (2) produced complex numbers. The dotted lines as well as the numbers in Table 2 represent the average directions. Since spatial aliasing might occur at frequency bins over 3 kHz, we did not use these results for obtaining the average. We see that the average results were close to the results of "MUSIC each", and the DOAs were correctly estimated in most of the frequency bins.

To show the effectiveness of our proposed method even for a situation where the MUSIC algorithm can be applied, we performed experiments for two-source three-sensor cases with both methods. Figure 3 shows the results for two sources placed far apart. We see that both methods estimated the directions well enough. In contrast, Fig. 4 shows the results

Table 2: Result summary

	Figure 2			Figure 3		Figure 4	
Source signal	◇	+	□	◇	□	×	□
MUSIC each	48°	73°	119°	48°	119°	105°	119°
ICA mixed	45°	74°	123°	45°	123°	105°	124°
MUSIC mixed	—	—	—	45°	122°	94°	128°

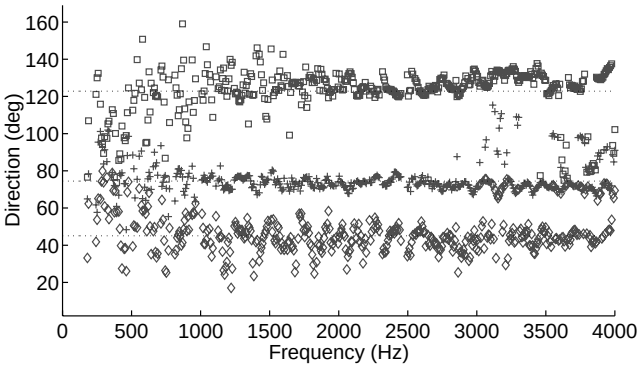

Figure 2: DOA estimations for three sources using ICA

for two closely placed sources. In this case, the MUSIC algorithm failed to estimate the directions in most frequency bins, whereas the proposed method based on ICA succeeded in estimating the directions in most cases. The average directions obtained with the ICA method were closer to the results of "MUSIC each" than those of the MUSIC algorithm. Therefore, we consider that the ICA method has an additional advantage over the MUSIC algorithm as regards its robustness for closely placed sources.

6. CONCLUSION

We proposed a new method for DOA estimation based on ICA. The advantages over the MUSIC algorithm are that it can be applied even if the number of sources is equal to the number of sensors, and also that it is robust for closely placed signals mixed in a reverberant condition. The proposed method solves the difficulty posed by the previous method based on ICA [6, 7] as regards its application when more than two sources are mixed. In terms of computational cost, estimated directions are provided by closed form formulas with the new method, whereas minima of a function must be searched for with the MUSIC algorithm or the previous method.

7. REFERENCES

[1] S. U. Pillai, *Array Signal Processing*, Springer-Verlag, 1989.

[2] M. Brandstein and D. Ward, Eds., *Microphone Arrays*, Springer, 2001.

[3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, pp. 276–280, Mar. 1986.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

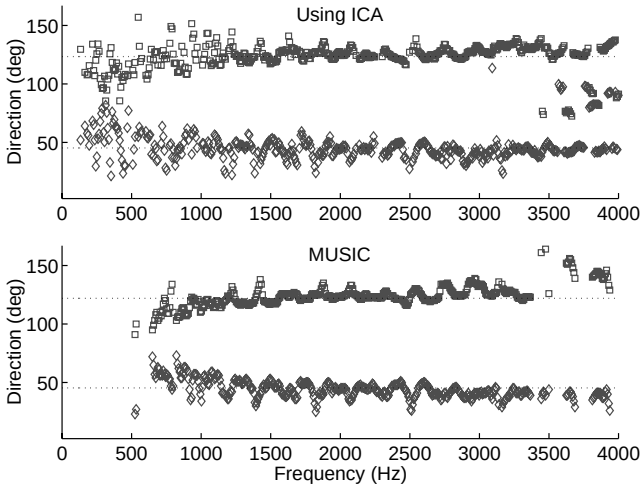

Figure 3: DOA estimations for two sources placed far apart

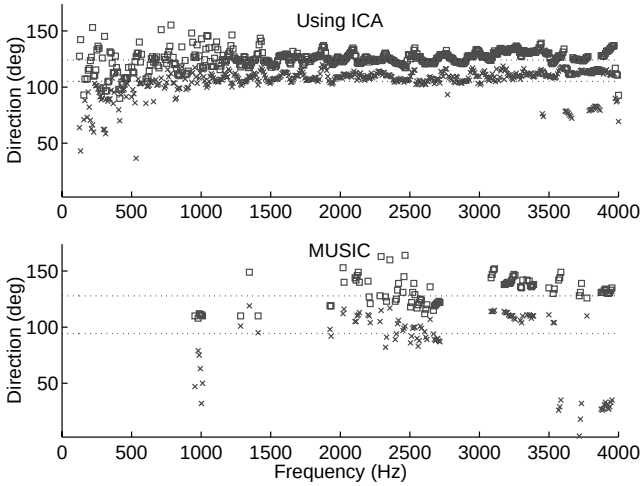

Figure 4: DOA estimations for two closely placed sources

[5] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," in *Proc. 4th International Symposium on Independent Component Analysis and Blind Signal Separation*, Apr. 2003, pp. 505–510.

[6] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura, "Evaluation of blind signal separation method using directivity pattern under reverberant conditions," in *Proc. ICASSP 2000*, June 2000, pp. 3140–3143.

[7] M. Z. Ikram and D. R. Morgan, "A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation," in *Proc. ICASSP 2002*, May 2002, pp. 881–884.

[8] Real World Computing Partnership, "Rwcp sound scene database in real acoustic environments," `http://tosa.mri.co.jp/sounddb/indexe.htm`.

[9] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.