DOA ESTIMATION FOR MULTIPLE SPARSE SOURCES WITH NORMALIZED OBSERVATION VECTOR CLUSTERING

Shoko Araki[†‡]   *Hiroshi Sawada*[†]   *Ryo Mukai*[†]   *Shoji Makino*[†‡]

† NTT Communication Science Laboratories, NTT Corporation
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan
‡ Graduate School of Information Science and Technology, Hokkaido University
Kita 14, Nishi 9, Kita-ku, Sapporo-shi, Hokkaido 060-0814, Japan
Email: shoko@cslab.kecl.ntt.co.jp

ABSTRACT

This paper presents a new method for estimating the direction of arrival (DOA) of source signals whose number N can exceed the number of sensors M. Subspace based methods, e.g., the MUSIC algorithm, have been widely studied, however, they are only applicable when $M > N$. Another conventional independent component analysis based method allows $M \geq N$, however, it cannot be applied when $M < N$. By contrast, our new method can be applied where the sources outnumber the sensors (i.e., an underdetermined case $M < N$) by assuming source sparseness. Our method can cope with 2- or 3-dimensionally distributed sources with a 2- or 3-dimensional sensor array. We obtained promising experimental results for 3×4, 3×5 and 4×5 (#sensors × #speech sources) in a room (RT$_{60}$= 120 ms).

1. INTRODUCTION

Direction of arrival (DOA) estimation is a basic and important technique in array signal processing [1, 2]. The DOA estimation of speech, on which we focus in this paper, has many applications such as teleconference systems and automatic conference minutes generators. Such applications usually face situations where more sources than sensors are active simultaneously. We propose a new method for estimating DOA that can be applied even when the sources outnumber the sensors.

The most widely used approaches for DOA estimation are subspace based methods, e.g., the MUSIC (MUltiple SIgnal Classification) algorithm [3], and its variants. Because these methods need a noise subspace, they require more sensors than sources $M \geq N + 1$, that is they can be applied only when $M > N$.

Recently, a DOA estimation method based on independent component analysis (ICA) has been proposed [4, 5]. It estimates DOAs directly from the separation matrix estimated with ICA by utilizing the fact that the separation matrix is related to the source mixing process. Because this method is based on ICA, it allows cases where $M \geq N$. However, it still cannot be applied when $M < N$.

In order to cope with underdetermined cases where $M < N$, we propose a new DOA estimation method by assuming source sparseness. If the signals are assumed to be sufficiently sparse in the time-frequency domain, we can assume that only one source is dominant at a time-frequency point. Therefore, we can estimate information about each source by collecting observation samples that appear to belong to one of the sources. With our method, first we normalize all the observations and cluster the normalized observation vectors.

As each cluster corresponds to an individual source, we can formulate the DOAs by using the cluster centroids and given sensor locations. We have already proposed a sparseness based blind source separation algorithm with normalized observation vector clustering [6]. In this paper, we show that we can also estimate the DOAs of more sources than sensors with the same normalization and clustering method.

Some DOA estimation methods for $M < N$ cases have already been proposed [7–9]. The method described in [7] is based on source sparseness. It clusters the phase difference of only two sensor observations, then estimates the DOA. In addition, [8] estimated the DOA with the cumulants of observations by assuming the non-gaussianity of sources, and utilizing a linear sensor array. These two methods limit the DOA estimation ability on a 2-dimensional half-plane. The authors of [9] utilized a triangular sensor array and clustered the phase differences of each sensor pair by assuming source sparseness. Their method expands the DOA estimation ability to an entire 2-dimensional plane. However, their approach still cannot handle 3-dimensionally distributed sources. Moreover, their formulation assumed a *regular*-triangle sensor array. They have to re-formulate their method to use an array with another arrangement.

On the other hand, our newly proposed formulation is more general: the DOA is given by the cluster centroids and the sensor coordinates. Therefore, we can employ an arbitrary sensor arrangement including a 3-dimensional arrangement, and cope with 3-dimensionally distributed sources.

Experimental results show that the new method succeeded in estimating the DOAs for $M \times N$ of 3×4, 3×5 and 4×5. Neither the MUSIC algorithm nor the ICA based method can be used in such situations. We also compare our method and MUSIC when $M = 3$ and $N = 2$. When sources were closely placed, the MUSIC algorithm failed to estimate their DOAs, whereas the proposed method still succeeded.

2. PROBLEM DESCRIPTION

2.1. Observation model

Suppose that sources $s_1, \ldots, s_N$ are convolutively mixed and observed at M sensors

$$x_j(t) = \sum_{i=1}^{N} \sum_{l} h_{ji}(l) s_i(t-l), \ j = 1, \ldots, M, \quad (1)$$

where $h_{ji}(l)$ represents the impulse response from source i to sensor j. In this paper, we especially consider a situation where the number of sources N can exceed the number of sensors M ($M < N$). Here we assume that the number of sources N is given or can be
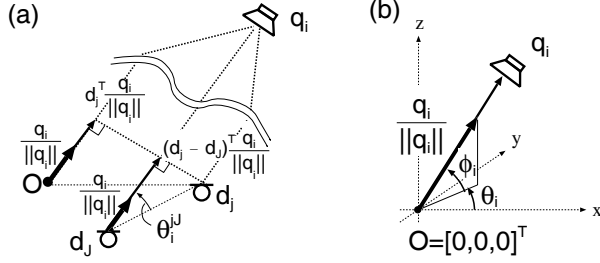
Fig. 1. (a) Far-field model, (b) definition of source location.

estimated. Our task is to estimate the directions of arrival (DOAs) of N sources from the sensor observations without information about sources s_i or impulse responses h_{ji}. We will formulate the DOA estimation problem in Sec. 2.2.

As with most DOA estimation techniques, this paper employs a time-frequency domain approach. Using a short-time Fourier transform (STFT), the convolutive mixtures (1) can be converted to instantaneous mixtures at each frequency f:

$$x_j(f,\tau) \approx \sum_{k=1}^{N} h_{ji}(f) s_i(f,\tau), \tag{2}$$

or in vector notation,

$$\mathbf{x}(f,\tau) \approx \sum_{i=1}^{N} \mathbf{h}_i(f) s_i(f,\tau), \tag{3}$$

where $h_{ji}(f)$ is the frequency response from source i to sensor j, $s_i(f,\tau)$ is the STFT of a source signal s_i, and τ is a time index. We call $\mathbf{x} = [x_1, \ldots, x_M]^T$ an observation vector and $\mathbf{h}_i = [h_{1i}, \ldots, h_{Mi}]^T$ is a vector of the frequency responses from source s_i to all sensors.

In the time-frequency domain, the sparseness of a source signal, which has been widely employed for solving the underdetermined problem [6, 7, 9–11], becomes prominent, if the source is colored and non-stationary such as speech. When the signals are sufficiently sparse, we can assume that the sources rarely overlap at each time-frequency point, and (3) can be approximated as

$$\mathbf{x}(f,\tau) \approx \mathbf{h}_k(f) s_k(f,\tau), \quad k \in \{1, \cdots, N\}, \tag{4}$$

where $s_k(f,\tau)$ is a dominant source at the time-frequency point (f,τ). For instance this is true for speech signals in the time-frequency domain [10, 11].

2.2. DOAs of source signals

Let us define the DOAs of sources more closely. Let $\mathbf{d}_j$ and $\mathbf{q}_i$ be 3-dimensional vectors representing the locations of sensor j and source i, respectively. We express the direction of source s_i as $\frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ with a unit-norm. Here, sensor locations $\mathbf{d}_j$ are given.

In order to estimate the DOAs of sources, we approximate the frequency response $h_{ji}(f)$ with a far-field model [Fig. 1(a)]:

$$h_{ji}(f) \approx \exp\left[j 2\pi f c^{-1} \mathbf{d}_j^T \frac{\mathbf{q}_i}{||\mathbf{q}_i||}\right], \tag{5}$$

where c is the propagation velocity of the signals. That is, we assume that the frequency response depends only on the path difference $\mathbf{d}_j^T \frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ between a source i to origin O and a source i to a sensor j [Fig. 1(a)]. When we consider the two sensors j and J, we obtain the following expressions:

$$\frac{h_{ji}(f)}{h_{Ji}(f)} \approx \exp\left[j 2\pi f c^{-1} (\mathbf{d}_j - \mathbf{d}_J)^T \frac{\mathbf{q}_i}{||\mathbf{q}_i||}\right] \tag{6}$$

$$= \exp\left[j 2\pi f c^{-1} ||\mathbf{d}_j - \mathbf{d}_J|| \cos \theta_i^{jJ}\right]. \tag{7}$$

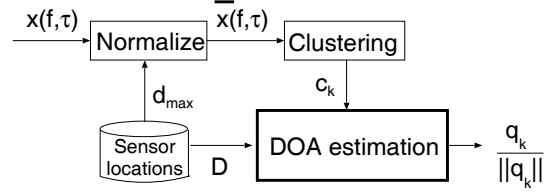These two equations show that we can estimate the DOA in two

ways: the DOA $\frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ with respect to a coordinate system, and the angle $\cos \theta_i^{jJ}$ with respect to a sensor pair j-J, [see Fig. 1(a)]. We term the former $\frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ an *absolute DOA*, and the latter $\cos \theta_i^{jJ}$ a *relative DOA*. The previous methods with two sensors [7] or a linear sensor array [8] can estimate only the relative DOAs. On the other hand, in this paper, our goal is to estimate the absolute DOAs $\frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ of sources from the all sensor observations $\mathbf{x}(f,\tau)$. Using the azimuth θ_i and elevation ϕ_i [see Fig. 1(b)], the absolute DOAs $\frac{\mathbf{q}_i}{||\mathbf{q}_i||}$ can be written as

$$\frac{\mathbf{q}_i}{||\mathbf{q}_i||} = [\cos \theta_i \cos \phi_i, \sin \theta_i \cos \phi_i, \sin \phi_i]^T. \tag{8}$$

3. PROPOSED METHOD

This section explains our proposed DOA estimation method, which is applicable when $M < N$. When the sparseness assumption (4) holds, the observation vector $\mathbf{x}(f,\tau)$ is a multiplication by an unknown scalar s_k of the frequency response $\mathbf{h}_k(f)$ about a source s_k which is dominant at (f,τ). Moreover, remember that the frequency response $\mathbf{h}_k(f)$ includes the DOA information $\frac{\mathbf{q}_k}{||\mathbf{q}_k||}$ of source s_k [see (5)]. Therefore, if we can collect the observation samples that appear to belong to one of the sources, we can estimate its DOA as an expectation value.

Based on this idea, we cluster the observation vector $\mathbf{x}(f,\tau)$ so that each cluster corresponds to each source. Figure 2 shows the flow of our method. To make $\mathbf{x}(f,\tau)$ form clusters, first we need normalizations of all the observations. Then we cluster the normalized observation vectors $\bar{\mathbf{x}}(f,\tau)$ to gather the observation samples that belong to one of the sources. Because each cluster corresponds to an individual source, finally, we estimate the DOAs by using the cluster centroids $\mathbf{c}_k$. We explain each step in detail in the following subsections.

3.1. Normalization

In this step, we normalize all observation vectors $\mathbf{x}(f,\tau)$ so that they form clusters, each of which corresponds to an individual source.

First, we remove the inconstancy of the scalar $s_k(f,\tau)$ in (4). This can be normalized by taking the ratio of two observation components: $x_j(f,\tau)/x_J(f,\tau) \approx h_{jk}(f,\tau)/h_{Jk}(f,\tau)$, which is modeled as (6). Moreover, if we can normalize the frequency dependence observed in the phase component of (6), we can handle all the frequency components together and obtain enough data samples even if we use short observations. The frequency dependence is normalized by dividing the phase of $x_j(f,\tau)/x_J(f,\tau)$ by f.

Including the above normalizations, our normalization is performed for the all components of $\mathbf{x}(f,\tau)$ by selecting one of the sensors J:

$$\bar{x}_j(f,\tau) \leftarrow |x_j(f,\tau)| \exp\left[j \frac{\arg[x_j(f,\tau)/x_J(f,\tau)]}{4 f c^{-1} d_{\max}}\right] \tag{9}$$

where c is the propagation velocity and $d_{\max}$ is the maximum distance between the reference sensor J and a sensor $^\forall j \in \{1, \ldots, M\}$.

Fig. 2. Flow of proposed method

The rationale for the frequency normalization $4fc^{-1}d_{\max}$ can be found in the Appendix. Here, we keep the amplitudes at all sensors, although this is not needed for DOA estimation. This information is useful when we perform source localization, i.e., estimate the positions $\mathbf{q}_i$ of sources, using a near-field model.

We also employ unit-norm normalization to handle the vectors on a unit-hypersphere,

$$\bar{\mathbf{x}}(f,\tau) \leftarrow \bar{\mathbf{x}}(f,\tau)\,/\,||\bar{\mathbf{x}}(f,\tau)|| \qquad (10)$$

for $\bar{\mathbf{x}}(f,\tau) = [\bar{x}_1(f,\tau), \ldots, \bar{x}_M(f,\tau)]^T$.

If sources are sparse, each component of the normalized observation vector is expressed as

$$\bar{x}_j(f,\tau) = \frac{1}{\sqrt{M}}\exp\left[\jmath\frac{\pi}{2d_{\max}}(\mathbf{d}_j - \mathbf{d}_J)^T\frac{\mathbf{q}_k}{||\mathbf{q}_k||}\right] \quad (11)$$

$$= \frac{1}{\sqrt{M}}\exp\left[\jmath\frac{\pi}{2d_{\max}}||\mathbf{d}_j - \mathbf{d}_J||\cos\theta_k^{jJ}\right], \quad (12)$$

by using (4), (5), (9), and (10). We can see that the normalized components $\bar{x}_j(f,\tau)$ keep the geometric information of sensors and a source s_k, which is dominant at (f,τ). Therefore, the normalized vectors $\bar{\mathbf{x}}(f,\tau)$ form clusters based on the source geometry in the clustering step.

3.2. Clustering

In this step, we find clusters $C_1, \ldots, C_N$ formed by all normalized vectors $\bar{\mathbf{x}}(f,\tau)$. The centroid $\mathbf{c}_k$ of a cluster C_k is calculated by

$$\mathbf{c}_k \leftarrow \boldsymbol{E}[\bar{\mathbf{x}}(f,\tau)]_k, \quad \mathbf{c}_k \leftarrow \mathbf{c}_k/||\mathbf{c}_k||, \qquad (13)$$

where $\boldsymbol{E}[\cdot]_k$ is a mean operator for the members of a cluster C_k. The clustering criterion is to minimize the total sum $\mathcal{J}$ of the squared distances between the cluster members and their centroid

$$\mathcal{J} = \sum_{k=1}^M \mathcal{J}_k, \quad \mathcal{J}_k = \sum_{\bar{\mathbf{x}}\in C_k}||\bar{\mathbf{x}} - \mathbf{c}_k||^2. \qquad (14)$$

This minimization can be performed efficiently with the k-means clustering algorithm [12].

3.3. DOA estimation

Because each cluster corresponds to an individual source, the centroid $\mathbf{c}_k$ represents the geometry of the source s_k. Note that the centroid gives an expectation value for the k-th source [see (13)]. From (11) and (12), the j-th component of the $\mathbf{c}_k$ is expressed as

$$\{\mathbf{c}_k\}_j \propto \boldsymbol{E}[\bar{x}_j(f,\tau)]_k$$

$$= \frac{1}{\sqrt{M}}\exp\left[\jmath\frac{\pi}{2d_{\max}}(\mathbf{d}_j - \mathbf{d}_J)^T\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||}\right] \qquad (15)$$

$$= \frac{1}{\sqrt{M}}\exp\left[\jmath\frac{\pi}{2d_{\max}}||\mathbf{d}_j - \mathbf{d}_J||\cos\tilde{\theta}_k^{jJ}\right]. \quad (16)$$

where $\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||}$ and $\tilde{\theta}_k^{jJ}$ are the estimated DOAs. When we use all the components of the centroid $\mathbf{c}_k$, we can derive the following relationship from (15),

$$\boldsymbol{D}\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||} = \frac{2d_{\max}}{\pi}\boldsymbol{r}_k \qquad (17)$$

where $\boldsymbol{D} = [\mathbf{d}_1 - \mathbf{d}_J, \cdots, \mathbf{d}_M - \mathbf{d}_J]^T$ is a matrix of the sensor locations and $\boldsymbol{r}_k = [\arg[\{\mathbf{c}_k\}_1], \cdots, \arg[\{\mathbf{c}_k\}_M]]$. Using the Moore-Penrose pseudo-inverse, the absolute DOA is obtained as

$$\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||} = \frac{2d_{\max}}{\pi}\boldsymbol{D}^+\boldsymbol{r}_k. \qquad (18)$$

If rank$(\boldsymbol{D}) \geq 3$, we can estimate the 3-dimensional absolute DOA. Note that if we use a *unit* sensor array system $[\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4] = \alpha[(0,0,0)^T, (1,0,0)^T, (0,1,0)^T, (0,0,1)^T]$ (α: a constant) and $J =$

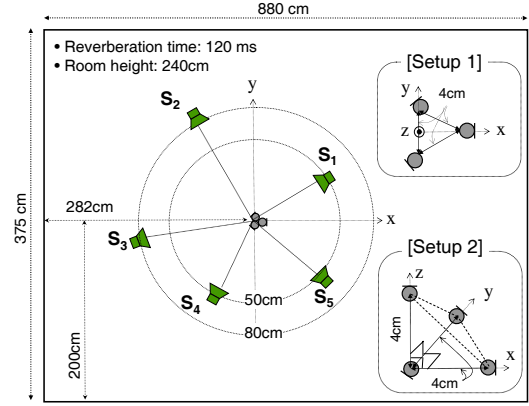**Fig. 3**. Experimental setups with non-linear sensor arrays.

Table 1. Experimental results for $M = 3, N = 4$ (Setup 1),

Source	s_1	s_2	s_3	s_4
MUSIC (each)	24°	117°	217°	311°
Proposed (5 sec.)	25°	114°	214°	313°
Proposed (1 sec.)	23°	112°	212°	318°

1, then $\arg[\{\mathbf{c}_k\}_1] \approx 0$ and $\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||} = \frac{\boldsymbol{r}_k'}{||\boldsymbol{r}_k'||}$, where $\boldsymbol{r}_k' = [\arg[\{\mathbf{c}_k\}_2], \arg[\{\mathbf{c}_k\}_3], \arg[\{\mathbf{c}_k\}_4]]^T$. That is, we do not need $\boldsymbol{D}^+$, and DOA $\frac{\tilde{\mathbf{q}}_k}{||\tilde{\mathbf{q}}_k||}$ can be obtained simply from $\boldsymbol{r}_k'$.

Using j-th component of the centroid $\mathbf{c}_k$, we can also estimate the relative DOA $\tilde{\theta}_k^{jJ}$ from (16) if needed:

$$\cos\tilde{\theta}_k^{jJ} = \frac{2d_{\max}}{\pi}\frac{\arg[\{\mathbf{c}_k\}_j]}{||\mathbf{d}_j - \mathbf{d}_J||}. \qquad (19)$$

4. EXPERIMENTAL RESULTS

We performed experiments in a reverberant condition. Observations were made by following (1) with the impulse responses $\mathrm{h}_{ji}(l)$ measured in a room (Fig. 3) and 5-second English speech sources $\mathrm{s}_i(t)$ sampled at 8 kHz. The frame size L for STFT was 512, and the frame shift was 256 $(= L/2)$.

We estimated the absolute DOAs $\frac{\mathbf{q}_k}{||\mathbf{q}_k||}$ with (18), and evaluated them with the azimuth θ_k and the elevation ϕ_k [see (8)]. As the *true* directions, we utilized the DOAs estimated by the MUSIC algorithm when there was only one source signal. They will appear in the tables with the label "MUSIC (each)". We investigated eight speaker combinations and averaged the results.

Table 1 shows the results for four sources with three sensors ($M = 3, N = 4$), that were arranged non-linearly (Fig. 3 [Setup 1]). Here, all source heights were the same as the height of the sensor array. Because all elevations ϕ_k are zero, only the results of θ_k are shown in the Table 1. We can see that the DOAs estimated with our proposed method were very close to the results of "MUSIC (each)". Even when we used only one second data for the DOA estimation, we still obtained reasonable results as shown in Table 1. In a more complicated situation where $M = 3$ and $N = 5$, the proposed method estimated the DOAs very accurately as shown in Table 2.

We also applied our method to a 3-dimensional sensor arrangement (Fig. 3 [Setup 2]). In this case, the sources had different heights, and therefore, we estimated both azimuths θ_k and elevations ϕ_k. Table 3 shows results for five sources with four sensors ($M = 4$, $N = 5$). As regards θ_k, although the estimation error was sometimes greater than the result for $M = 3$, $N = 4$ (Table 1), we still obtained reasonable results for such a complicated case. The

Table 2. Experimental results for $M = 3$, $N = 5$ (Setup 1).

Source	s_1	s_2	s_3	s_4	s_5
MUSIC (each)	24°	117°	176°	217°	311°
Proposed (5 sec.)	23°	112°	175°	218°	314°

Table 3. Experimental results for $M = 4$, $N = 5$ (Setup 2).

Source		s_1	s_2	s_3	s_4	s_5
MUSIC (each)	θ	31°	85°	133°	222°	302°
	ϕ	−26°	6°	30°	39°	−8°
Proposed (5 sec.)	θ	30°	79°	132°	221°	298°
	ϕ	−22°	7°	28°	35°	−9°

elevation values ϕ_k were also very close to the *true* values, "MUSIC (each)". We can say that our proposed method can be applied to such a 3-dimensional DOA estimation and that it gives us fairly precise DOAs.

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 (Setup 1) cases with both methods. Figure 4 shows the resolution of both methods, and Table 4 shows the estimated DOAs θ_k. Figure 4 (a) and (c) are example MUSIC spectra at a frequency $f = 1844$ Hz, and (b) and (d) are DOA histograms for members of each cluster (the DOA of each member can be calculated with (18) using $\bar{x}(f, \tau)$ instead of the centroid $\mathbf{c}_k$). When two sources were placed with a wide spacing [Fig. 4 (a) and (b), and Table 4 "wide"], both MUSIC and the proposed method estimated the directions well enough. In contrast, when the two sources were close to each other [Fig. 4 (c) and (d), and Table 4 "close"], MUSIC failed to estimate the two directions, whereas the proposed method was still successful. We consider that our proposed method with sparseness assumption has a high resolution, although this resolution depends on the sparseness of the source signals and is affected by the room reverberation condition.

5. CONCLUSION

We proposed a new DOA estimation method for cases where $M < N$ by assuming source sparseness. The method is based on the normalization and clustering of the observation vectors. We obtained promising experimental results for underdetermined cases in a reverberant condition. We also confirmed that our proposed method has higher resolution for estimating the directions of sources than the MUSIC algorithm.

We can also formulate the source localization, i.e., estimate the positions $\mathbf{q}_i$ of sources, using our proposed normalized observation vector and a near-field model. We plan to investigate its performance in the future.

Appendix

This appendix explains the rationale for the denominator of the phase normalization in (9). The normalized observation vector component can be written as (11) or (12). From the fact that

$$\left| (\mathbf{d}_j - \mathbf{d}_J)^T \frac{\mathbf{q}_k}{||\mathbf{q}_k||} \right| = \left| ||\mathbf{d}_j - \mathbf{d}_J|| \cos\theta_k^{jJ} \right| \leq d_{\max},$$

an inequality

$$-\pi/2 \leq \arg[\bar{x}_j(f, \tau)] \leq \pi/2$$

holds. Due to the frequency normalization with $4fc^{-1}d_{\max}$, the phase is distributed effectively from $-\pi/2$ to $\pi/2$. This property is also important for the distance measure (14), since $|\bar{x} - \bar{x}'|$ increases monotonically as $|\arg(\bar{x}) - \arg(\bar{x}')|$ increases.
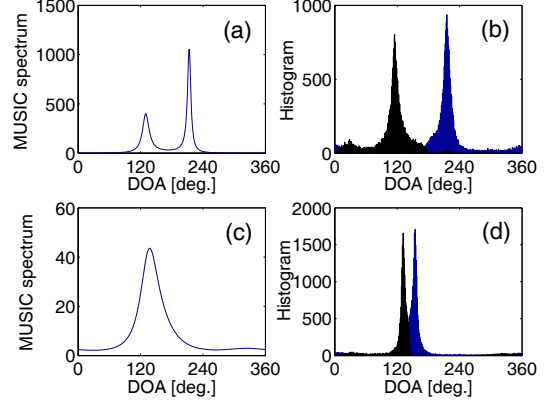
Fig. 4. Resolution of MUSIC [(a),(c)] and the proposed method [(b),(d)]. (a)(b): wide setting (θ_1=117°, θ_2=217°), (c)(d): close setting (θ_1=132°, θ_2=154°).

Table 4. Experimental results for $M = 3$, $N = 2$ (Setup 1).

	wide		close	
Source	s_1	s_2	s_1	s_2
MUSIC (each)	117°	217°	132°	154°
Proposed (5 sec.)	114°	217°	128°	156°
MUSIC (mix)	125°	217°	126°	345°

6. 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, no. 3, pp. 276–280, Mar. 1986.

[4] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Frequency-domain blind source separation," in *Speech Enhancement*, J. Benesty, S. Makino, and J. Chen, Eds., pp. 299–327. Springer, Mar. 2005.

[5] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Multiple source localization using independent component analysis," in *Proc. AP-S/URSI2005*, 2005.

[6] S. Araki, H. Sawada, R. Mukai, and S. Makino, "A novel blind source separation method with observation vector clustering," in *Proc. IWAENC 2005*, Sept. 2005, pp. 117–120.

[7] S. Rickard and F. Dietrich, "DOA estimation of many W-disjoint orthogonal sources from two mixtures using DUET," in *Proc. SSAP2000*, Aug. 2000, pp. 311–314.

[8] S. Shamsunder and G. B. Giannakis, "Modeling of non-Gaussian array data using cumulants: DOA estimation of more sources with less sensors," *Signal Processing*, , no. 30, pp. 279–297, 1993.

[9] M. Matsuo, Y. Hioka, and N. Hamada, "Estimating DOA of multiple speech signals by improved histogram mapping method," in *Proc. IWAENC2005*, Sept. 2005, pp. 129–132.

[10] Ö. Yılmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. SP*, vol. 52, no. 7, pp. 1830–1847, 2004.

[11] P. Bofill and M. Zibulevsky, "Blind separation of more sources than mixtures using sparsity of their short-time Fourier transform," in *Proc. ICA2000*, 2000, pp. 87–92.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley Interscience, 2nd edition, 2000.