Underdetermined Blind Separation of Convolutive Mixtures of Speech with Directivity Pattern Based Mask and ICA

Shoko Araki, Shoji Makino, Hiroshi Sawada, and Ryo Mukai

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

Abstract. We propose a method for separating N speech signals with M sensors where $N > M$. Some existing methods employ binary masks to extract the signals, and therefore, the extracted signals contain loud musical noise. To overcome this problem, we propose using a directivity pattern based continuous mask, which masks $N - M$ sources in the observations, and independent component analysis (ICA) to separate the remaining mixtures. We conducted experiments for $N = 3$ with $M = 2$ and $N = 4$ with $M = 2$, and obtained separated signals with little distortion.

1 Introduction

In this paper, we consider the blind source separation (BSS) of speech signals observed in a real environment, i.e., the BSS of convolutive mixtures of speech. Recently, many methods have been proposed to solve the BSS problem of convolutive mixtures [1]. However, most of these methods consider the determined or overdetermined case. In contrast, we focus on the underdetermined BSS problem where the N source signals outnumber M sensors.

Several methods have been proposed for underdetermined BSS [2–5]. There are two approaches, and both approaches rely on the sparseness of the source signals. One extracts each signals with time-frequency binary masks [2], and the other is based on ML estimation, where the sources are estimated after mixing matrix estimation [3–5]. In [2], the authors employ a time-frequency binary mask (BM) to extract each signal, and they have applied it to real speech mixtures. However, the use of binary masks causes too much discontinuous zero-padding to the extracted signals, and they contain loud musical noise.

To overcome this, we have proposed combining binary masks and ICA (BMICA) to solve the underdetermined BSS problem [6] especially for $N = 3$ and $M = 2$. This method consists of two stages: (1) one source removal with a binary mask and (2) separation of the remaining mixtures with ICA (for details see Sec. 3.3). As this one source removal extracts more time-frequency points than the BM method, it causes less zero-padding than the BM method, and therefore, we have been able to separate signals with less musical noise. However, the BMICA still employs a binary mask.

C.G. Puntonet and A. Prieto (Eds.): ICA 2004, LNCS 3195, pp. 898–905, 2004.

Therefore we have also proposed to utilize a directivity pattern based continuous mask (DCmask) instead of a binary mask at the source removal stage (DCmask and ICA: DCICA) [7]. The DCmask has a small gain for the DOAs of sources to be masked, and has a large gain for other directions. Because the DCmask is a non-binary mask, we can avoid the zero-padding. However, in [7], as we masked at most $M-1$ sources, we applied the DCICA only for $N \leq (M-1)+M$.

In this paper, to release this limit, we propose a method for masking $N-M$ sources for an arbitrary number of sources N. Our proposal is to utilize the directivity pattern of a null beamformer (NBF), which makes nulls towards given $N-M$ directions, formed by $V = N-M+1$ *virtual* microphones. We conducted experiments for $N = 3$ with $M = 2$ and $N = 4$ with $M = 2$, and the experimental results show that our method can separate signals with little distortion.

2 Problem Description

In real environments, N source signals s_i observed by M sensors are modeled as convolutive mixtures $x_j(n) = \sum_{i=1}^{N} \sum_{l=1}^{L} h_{ji}(l)\ s_i(n-l+1)\ \ (j = 1, \cdots, M)$, where h_{ji} is the L-taps impulse response from a source i to a sensor j. Our goal is to obtain separated signals $y_k(n)\ \ (k = 1, \cdots, N)$ using only the information provided by observations $x_j(n)$. Here, we consider the case of $N > M$.

This paper employs a time-frequency domain approach because speech signals are more sparse in the time-frequency domain than in the time-domain [5] and convolutive mixture problems can be converted into instantaneous mixture problems at each frequency. In the time-frequency domain, mixtures are modeled as $\mathbf{X}(\omega, m) = \mathbf{H}(\omega)\mathbf{S}(\omega, m)$, where $\mathbf{H}(\omega)$ is an $M \times N$ mixing matrix whose ji component is a transfer function from a source i to a sensor j, $\mathbf{S}(\omega, m) = [S_1(\omega, m), \cdots, S_N(\omega, m)]^T$ and $\mathbf{X}(\omega, m) = [X_1(\omega, m), \cdots, X_M(\omega, m)]^T$ denote short-time Fourier transformed sources and observed signals, respectively. ω is the frequency and m is the time-dependence of the short-time Fourier transformation (STFT). We assume that sources are mutually independent and that each source has a sparse distribution in a time-frequency domain. These assumptions are approximately true for speech signals. Moreover, $\mathbf{Y}(\omega, m) = [Y_1(\omega, m), \cdots, Y_N(\omega, m)]^T$ denotes the STFT of separated signals.

3 Conventional Methods

3.1 Classification of Time-Frequency Points with Sparseness

Several methods have been proposed [2–8] for solving the underdetermined BSS problem, and they all utilize source sparseness. When signals are sufficiently sparse, it can be assumed that sources do not overlap very often. Therefore, a histogram of $(\frac{|X_i(\omega,m)|}{|X_j(\omega,m)|}, \angle\frac{X_i(\omega,m)}{X_j(\omega,m)})\ \ (i \neq j)$ for example, contains N peaks. Furthermore, we can classify the observation sample points $X_j(\omega, m)$ into N classes according to the histogram, which is what the BM method does (see Sec. 3.2).

In this paper, we utilize omnidirectional microphones, therefore we use the phase difference $\varphi(\omega,m) = \angle\frac{X_i(\omega,m)}{X_j(\omega,m)}$ $(i \neq j)$ between two observations. A histogram of the direction of arrival (DOA) $\theta(\omega,m) = \cos^{-1}\frac{\varphi(\omega,m)c}{\omega d}$ (d: the microphone space, c: the speed of sound) has N peaks (Fig. 1). Each peak corresponds to each source. Let these peaks be $\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_N$ where $\tilde{\theta}_1 \leq \tilde{\theta}_2 \leq \cdots \leq \tilde{\theta}_N$ (Fig. 1), and the signal from $\tilde{\theta}_\xi$ be $\tilde{S}_\xi$ $(\xi = 1, \cdots, N)$.

3.2 Conventional Method 1: With Only Binary Masks (BM)

As alluded to in Sec. 3.1, we can extract each signal using time-frequency binary masks (e.g., [2]). We can extract each signal with a binary mask

$$\text{[BM]} \qquad M^{\xi}_{\text{BM}}(\omega,m) = \begin{cases} 1 & \tilde{\theta}_\xi - \Delta \leq \theta(\omega,m) \leq \tilde{\theta}_\xi + \Delta \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

by calculating $Y_\xi(\omega,m) = M^{\xi}_{\text{BM}}(\omega,m)X_j(\omega,m)$ where Δ is an extraction range parameter.

Although we can obtain separated signals with binary masks (1), the signals are discontinuously zero-padded by binary masks, and therefore, we hear musical noise in the outputs. Moreover, the performance depends on the parameter, Δ.

3.3 Conventional Method 2: With Binary Mask and ICA (BMICA)

To overcome the musical noise problem, we have proposed using both a binary mask and ICA (BMICA) [6]. The BMICA has two stages. At the first stage, using the sparseness assumption, we *remove* the $N - M$ sources from the observations with a binary mask. Then in the second stage, we apply ICA to the remaining mixtures to obtain M separated signals.

Let $\Theta_S = \{\tilde{\theta}_{\mathsf{s}(1)}, \cdots, \tilde{\theta}_{\mathsf{s}(M)}\}$ be the set of DOAs of M signals to be separated and $\Theta_R = \{\tilde{\theta}_{\mathsf{r}(1)}, \cdots, \tilde{\theta}_{\mathsf{r}(N-M)}\}$ be the set of DOAs of $N - M$ signals to be removed (Fig. 1). To define the masks, let $\mathcal{I}_S = \{\mathsf{s}(1), \cdots, \mathsf{s}(M)\}$ be the set of indexes of Θ_S and $\mathcal{I}_R = \{\mathsf{r}(1), \cdots, \mathsf{r}(N-M)\}$ be the set of indexes of Θ_R.

For an index set $\mathcal{I}$, we define an area $\mathbb{A}$ by the following procedure:

1. $\mathbb{A} \leftarrow \emptyset$
2. if $1 \in \mathcal{I}$, $\mathbb{A} \leftarrow \mathbb{A} \cup [0°, \tilde{\theta}_1]$
3. if $N \in \mathcal{I}$, $\mathbb{A} \leftarrow \mathbb{A} \cup [\tilde{\theta}_N, 180°]$
4. for every index i such that $i \in \mathcal{I}$ and $i+1 \in \mathcal{I}$, $\mathbb{A} \leftarrow \mathbb{A} \cup [\tilde{\theta}_i, \tilde{\theta}_{i+1}]$

We define the separation area $\mathbb{A}_S$ by using $\mathcal{I}_S$, and the removal area $\mathbb{A}_R$ by using $\mathcal{I}_R$. We also define the transition area $\mathbb{A}_T = \bar{\mathbb{A}}_S \cap \bar{\mathbb{A}}_R$ (Fig. 1).

In the first stage, unlike the BM method where each source is extracted, we attempt to *remove* $N - M$ sources from Θ_R using a binary mask

$$\text{[BMICA]} \qquad M_{\text{BMICA}}(\omega,m) = \begin{cases} 1 & \theta(\omega,m) \in \mathbb{A}'_S \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

by calculating $\hat{\mathbf{X}}(\omega,m) = M_{\text{BMICA}}(\omega,m)\mathbf{X}(\omega,m)$, where $\mathbb{A}'_S = \mathbb{A}' \cup \mathbb{A}_S$, $\mathbb{A}' = \bigcup_{1 \leq i \leq M} [\tilde{\theta}_{\mathsf{s}(i)} - \Delta, \tilde{\theta}_{\mathsf{s}(i)} + \Delta]$ and Δ is an extraction range parameter. Here,

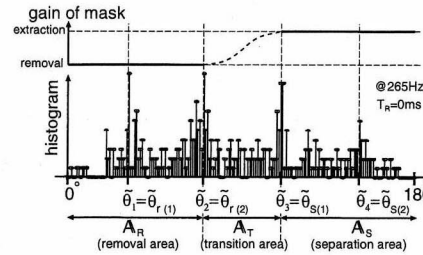

Fig. 1. Example histogram. ($N = 4$. Two male and two female combination with STFT frame size $T = 512$. $T_R = 0$ ms). An example of the area definition is also drawn for N=4, M=2. Here $\Theta_S = \{\tilde{\theta}_3, \tilde{\theta}_4\}$ and $\Theta_R = \{\tilde{\theta}_1, \tilde{\theta}_2\}$. Signals from $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are masked in the 1st stage, and signals from $\tilde{\theta}_3$ and $\tilde{\theta}_4$ are separated in the 2nd stage

$\hat{\mathbf{X}}(\omega, m)$ are expected to be mixtures of M signals from Θ_S. Therefore, in the second stage, we apply a standard ICA to these remaining mixtures.

We expect the zero-padding of the separated signals to cause less trouble because we extract more time-frequency points at the 1st stage than with the BM method. However, as BMICA still employed a binary mask for the source removal, the zero-padding to the separated signals still remained. Moreover, we have to find a reasonable Δ. This is not an easy problem and we relied on manual setting.

4 Proposed Method: Directivity Pattern Based Continuous Mask and ICA (DCICA)

Although the basic scheme (Fig. 2) of our proposed method is the same as that of BMICA, here we utilize non-binary masks at the 1st stage.

[1st Stage] $N - M$ Source Removal with New DC Mask: Here, we utilize a directivity pattern based continuous mask (DCmask) instead of a binary mask M_{BMICA}. When we have M microphones, we can utilize $M \times M$ ICA at the 2nd stage if we can mask $N - M$ signals. This can be realized by applying a mask that has $N - M$ nulls towards the DOAs Θ_R of the signals to be removed.

One way to obtain such a mask is to utilize the directivity pattern of a null beamformer (NBF), which makes nulls towards given $N - M$ directions Θ_R, formed by $V = N - M + 1$ (virtual) microphones. Here, V is not necessarily equal to M because *a mask is determined only by the number of signals to be removed at the 1st stage*: remember that we do not need information on the microphone number M when designing the masks for BM and BMICA methods.

Here, we assume that the number of sources N is known or estimated beforehand, e.g., from a histogram such as that shown in Fig. 1. First we form a $(V \times V)$ matrix $\mathbf{H}_{\text{NBF}}(\omega)$ whose ji element $H_{\text{NBF}ji}(\omega) = \exp(j\omega\tau_{ji})$, where

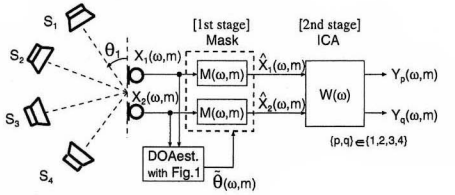

Fig. 2. Block diagram of proposed method. N =4 and M =2 case is drawn for example

$\tau_{ji} = \frac{d_j}{c}\cos\tilde{\theta}_i$, d_j is the position of the j-th virtual microphone, c is the speed of sound, $\{\tilde{\theta}_i\ (i = 2, \cdots, V)\} = \Theta_R$, and $\tilde{\theta}_1 = \theta_c \notin \mathbb{A}_R$ from which the signal's gain and phase are constrained at a constant value. By making a $(V \times V)$ matrix $\mathbf{H}_{\mathrm{NBF}}(\omega)$, we can remove $N - M$ signals even if $N >$ (the number of nulls formed by M sensors) + (the number of outputs of a standard ICA)$= (M-1) + M$.

Then one of the directivity patterns of the NBF, $\mathbf{W}(\omega) = \mathbf{H}_{\mathrm{NBF}}^{-1}(\omega)$, is

$$F(\omega, \theta) = \sum_{k=1}^{V} W_{1k}(\omega) \exp\left(j\omega d_k \cos\theta / c\right). \tag{3}$$

In this paper, we use the directivity pattern of the NBF as our mask,

$$\text{[DCICA 1]} \qquad M_{\mathrm{DC1}}(\omega, m) = F(\omega, \theta(\omega, m)). \tag{4}$$

This is our new mask, the DCmask. Figure 3 shows an example of the gain pattern of a DCmask.

We can also use a modified directivity pattern, for example,

$$\text{[DCICA 2]} \qquad M_{\mathrm{DC2}}(\omega, m) = \begin{cases} c_s & \theta(\omega, m) \in \mathbb{A}_S \\ F(\omega, \theta(\omega, m)) & \theta(\omega, m) \in \mathbb{A}_T \\ c_r & \theta(\omega, m) \in \mathbb{A}_R \end{cases} \tag{5}$$

where c_s is a constant (e.g., $\min_{\tilde{\theta}_i \in \Theta_s} |F(\omega, \tilde{\theta}_i)|$) and c_r is a small constant (e.g., the minimum value of the directivity pattern). By the mask M_{DC2}, the constant gain c_s is given to the M signals in the area $\mathbb{A}_S$. Moreover, this M_{DC2} changes smoothly in the transition area $\mathbb{A}_T$.

The source removal is achieved by $\hat{\mathbf{X}}(\omega, m) = M_{\mathrm{DC}k}(\omega, m)\mathbf{X}(\omega, m)$ (k=1 or 2). It should be noted that the DCmask is applied to all channels (Fig. 2), because ICA in the 2nd stage needs M inputs that maintain the mixing matrix information.

Because M_{DC1} and M_{DC2} are spatially smooth in the transition area $\mathbb{A}_T$, it is expected that the discontinuity of the extracted signals by these DCmasks is less serious than that by a mask M_{BM} in the BMICA.

[2nd Stage] Separation of Remaining Sources by ICA: Because the remaining signals $\hat{\mathbf{X}}$ are expected to be mixtures of M signals, we separate the signals using $M \times M$ ICA. The separation process is formulated as

$$\mathbf{Y}(\omega, m) = \mathbf{W}(\omega)\hat{\mathbf{X}}(\omega, m), \tag{6}$$

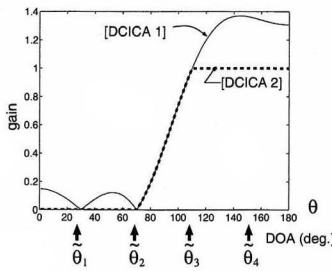

Fig. 3. Example mask pattern

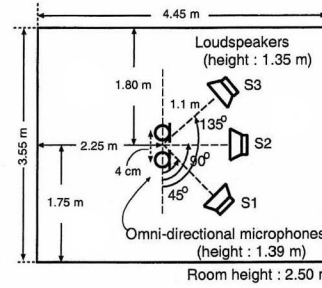

Fig. 4. Room for reverberant tests. T_R = 130 ms

where $\hat{\mathbf{X}}$ is the masked observed signal, $\mathbf{Y}(\omega, m) = [Y_1(\omega, m), \cdots, Y_M(\omega, m)]^T$ is the separated output signal, and $\mathbf{W}(\omega)$ represents an $(M \times M)$ separation matrix. $\mathbf{W}(\omega)$ is determined so that the output signals become mutually independent.

Note that we need several masks with nulls towards different directions to obtain all N separated signals because our system has only M outputs.

5 Experiments

5.1 Experimental Conditions

We conducted anechoic tests and reverberant tests. For the anechoic tests (T_R = 0 ms), we mixed speech signals using the mixing matrix $H_{ji}(\omega) = \exp(j\omega\tau_{ji})$, where $\tau_{ji} = \frac{d_j}{c}\cos\theta_i$, d_j is the position of the j-th microphone, and θ_i is the direction of the i-th source. The source directions were 45°, 90° and 135° (N=3), and 30°, 70°, 90° and 150° (N=4). For the reverberant tests, the speech data was convolved with impulse responses recorded in a real room (Fig. 4) whose reverberation time was T_R = 130 ms. As the original speech, we used Japanese sentences spoken by male and female speakers. We investigated three combinations of speakers.

The STFT frame size T was 512 and the frame shift was 256 at a sampling rate of 8 kHz. The Δ value for the conventional methods was 15° in DOA (N = 3) 10° in DOA (N = 4).

The adaptation rule of ICA we used was $\mathbf{W}_{i+1}(\omega) = \mathbf{W}_i(\omega) + \eta\left[\mathbf{I} - \langle\Phi(\mathbf{Y})\mathbf{Y}^{\mathrm{H}}\rangle\right] \cdot \mathbf{W}_i(\omega)$, where $\Phi(\mathbf{y}) = \phi(|\mathbf{y}|) \cdot e^{j\cdot\angle(\mathbf{y})}$, $\phi(x) = \mathrm{sign}(x)$. To solve the permutation problem of frequency domain ICA, we employed the DOA and correlation approach [9], and to solve the scaling problem of frequency domain ICA, we used the minimum distortion principle [10].

5.2 Performance Measures

We used the signal to interference ratio (SIR) and the signal to distortion ratio (SDR) as measures of separation performance and sound quality, respectively:

Table 1. Results of $N = 3$, $M = 2$ simulations. (a) T_R=0ms, (b) T_R=130ms

(a)

	pq	SIR1	SIR2	SIR3	SDR1	SDR2	SDR3
BM		18.0	8.9	18.4	7.9	11.5	8.3
BMICA	12	12.6	5.9		18.1	15.2	
	23		6.1	13.0		13.6	17.4
	13	16.9		16.4	11.7		11.7
DCICA1	12	16.2	4.9		15.2	13.1	
	23		4.6	16.3		13.2	15.6
	13	18.2		18.7	11.3		11.9
DCICA2	12	12.7	5.8		19.0	16.3	
	23		5.6	13.0		15.9	18.0

pq: Θ_S=$\{\tilde{\theta}_p, \tilde{\theta}_q\}$ [dB]

(b)

	pq	SIR1	SIR2	SIR3	SDR1	SDR2	SDR3
BM		12.3	6.3	11.0	5.0	13.9	5.8
BMICA	12	9.8	5.5		7.8	15.9	
	23		5.5	9.2		14.5	9.3
	13	11.9		12.5	6.9		7.2
DCICA1	12	13.6	4.1		7.0	11.2	
	23		3.9	11.7		14.4	8.6
	13	10.0		11.3	5.6		8.0
DCICA2	12	10.9	5.1		8.3	13.9	
	23		4.5	8.7		16.3	9.2

pq: Θ_S=$\{\tilde{\theta}_p, \tilde{\theta}_q\}$ [dB]

$\mathrm{SIR}_i = 10 \log \frac{\sum_n y_{is_i}^2(n)}{\sum_n (\sum_{i \neq j} y_{is_j}(n))^2}$ and $\mathrm{SDR}_i = 10 \log \frac{\sum_n x_{ks_i}^2(n)}{\sum_n (x_{ks_i}(n) - \alpha y_{is_i}(n-D))^2}$, where y_i is the estimation of s_i, and y_{is_j} is the output of the whole separating system at y_i when only s_j is active, and $x_{ks_i} = h_{ki} * s_i$ ($*$ is a convolution operator). α and D are parameters to compensate for the amplitude and phase difference between x_{ks_i} and y_{is_i}.

The SIR and SDR values were averaged over three speaker combinations.

5.3 Experimental Results

Applicability of ICA at the 2nd Stage Before trying to separate signals with our method, we investigated the masking performance. The percentage of each signal power extracted by M_{DC1} was S_1:S_2:S_3:S_4 = 78:20:1:1, and by M_{DC2} was 50:47:2:1 (N=4 (all female), M=2, T_R= 0ms, $\Theta_S = \{\tilde{\theta}_1, \tilde{\theta}_2\}$), for example. Two signals are dominant and other two signals are small. Therefore, we can use (2×2) ICA at the 2nd stage.

Separation results Table 1 (a) shows the experimental results for $T_R = 0$ ms and $N = 3, M = 2$. With BM method, the SDR values were unsatisfactory, and a large musical noise was heard. In contrast, with our proposed method (DCICA), we were able to obtain high SDR values without any serious deterioration in the separation performance SIR. Although the SDR values were slightly degraded compared with those by BMICA, we heard no musical noise with DCICA. Some sound samples can be found at our web site [11].

In DCICA1, SIR2 was degraded. This is because the gain for $\tilde{\theta}_2$ was less than the gain for $\tilde{\theta}_1$ or $\tilde{\theta}_3$. In DCICA2, which had constant gains for $\tilde{\theta}_2$ and $\tilde{\theta}_1$ or $\tilde{\theta}_3$, the SIR2 was improved and we obtained high SDR values.

Tables 2 shows the results for $N = 4$ and $M = 2$. We can apply our method for $N = 4$.

Table 1 (b) shows the results of reverberant tests for $T_R = 130$ ms ($N = 3, M = 2$). In the reverberant case, due to the decline of sparseness, the performance with all methods was worse than when $T_R = 0$ ms. However, we were

Table 2. Results of $N = 4$, $M = 2$ simulations. T_R=0ms

	pq	SIR1	SIR2	SIR3	SIR4	SDR1	SDR2	SDR3	SDR4
BM		16.7	9.6	7.7	16.7	4.4	7.1	7.5	4.7
BMICA	12	11.3	6.7			8.9	9.2		
	34			5.4	10.5			9.3	10.1
DCICA1	12	14.1	3.4			9.2	7.7		
	34			3.6	14.3			8.8	9.7
DCICA2	12	10.9	5.4			10.7	11.3		
	34			4.4	9.8			11.2	12.2

pq: $\Theta_s=\{\tilde{\theta}_p,\tilde{\theta}_q\}$ [dB]

able to obtain higher SDR values with DCICA than with the BM method even in a reverberant environment without musical noise.

It should be noted that it remains difficult to separate signals at the center position with any method.

6 Conclusion

We proposed utilizing a directivity pattern based continuous mask and ICA for BSS when speech signals outnumber sensors. Our method avoids discontinuous zero-padding, and therefore, can separate the signals with no musical noise.

References

1. Haykin, S.: Unsupervised adaptive filtering. John Wiley & Sons (2000)
2. Rickard, S., Yilmaz, O.: On the W-disjoint orthogonality of speech. In: Proc. ICASSP2002. (2002) 529–532
3. Theis, F.J., Puntonet, C.G., Lang, E.W.: A histogram-based overcomplete ICA algorithm. In: Proc. ICA2003. (2003) 1071–1076
4. Vielva, L., Erdogmus, D., Pantaleon, C., Santamaria, I., Pereda, J., Principe, J.C.: Underdetermined blind source separation in a time-varying environment. In: Proc. ICASSP2002. (2002) 3049–3052
5. Bofill, P., Zibulevsky, M.: Blind separation of more sources than mixtures using sparsity of their short-time Fourier transform. In: Proc. ICA2000. (2000) 87–92
6. Araki, S., Makino, S., Blin, A., Mukai, R., Sawada, H.: Blind separation of more speech than sensors with less distortion by combining sparseness and ICA. In: Proc. IWAENC2003. (2003) 271–274
7. Araki, S., Makino, S., Sawada, H., Mukai, R.: Underdetermined blind speech separation with directivity pattern based continuous mask and ICA. In: EUSIPCO2004. (2004)
8. Blin, A., Araki, S., Makino, S.: Blind source separation when speech signals outnumber sensors using a sparseness-mixing matrix combination. In: Proc. IWAENC2003. (2003) 211–214
9. Sawada, H., Mukai, R., Araki, S., Makino, S.: Convolutive blind source separation for more than two sources in the frequency domain. In: Proc. ICASSP2004. (2004)
10. Matsuoka, K., Nakashima, S.: A robust algorithm for blind separation of convolutive mixture of sources. In: Proc. ICA2003. (2003) 927–932
11. http://www.kecl.ntt.co.jp/icl/signal/araki/dcica.html