EVALUATION OF A SELF-GENERATOR METHOD FOR INITIAL FILTERS OF SIMO-ICA APPLIED TO BLIND SEPARATION OF BINAURAL SOUND MIXTURES

*Tomoya Takatani, Satoshi Ukai, Tsuyoki Nishikawa, Hiroshi Saruwatari, Kiyohiro Shikano**

Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama-cho, Ikoma-shi, Nara, 630-0192, JAPAN
{tomoya-t, sawatari, shikano}@is.naist.jp

ABSTRACT

Blind separation of binaural mixed sounds using Single-Input Multiple-Output (SIMO)-model-based Independent Component Analysis (SIMO-ICA) with self-generator for initial filter (SIMO-ICA-SG) is now being studied by the authors. This method contains frequency-domain ICA (FDICA-PB), single-talk detection, direction of arrival (DOA) estimation, head related transfer function (HRTF) matrix bank, and SIMO-ICA. This paper describes robustness of SIMO-ICA-SG against the mismatch of HRTF matrix bank. To evaluate it, the sound decomposition experiments are carried out under the real acoustic conditions The experimental results reveal that the decomposition performance of the proposed method with mismatched HRTF matrix bank is superior to those of the conventional methods, and almost the same as those of the proposed method with matched one.

1. INTRODUCTION

Blind source separation (BSS) is the approach taken to estimate original source signals using only the information of the mixed signals observed in each input channel. In recent works of BSS based on ICA,various methods have been proposed to deal with a means of separation of acoustic sounds which corresponds to the convolutive mixture case [1]. However, the conventional ICA-based BSS approaches are basically means of extracting each of the independent sound sources as a *monaural* signal, and consequently they have a serious drawback in that the separated sounds cannot maintain information about the directivity, localization, or spatial qualities of each sound source. This prevents any BSS methods from being applied to, e.g., binaural signal processing [2].

In order to solve the above-mentioned fundamental problems, several high-fidelity BSSs using the ICA-based algorithm have been proposed, in which the convolutive mixtures of acoustic signals are decomposed into the Single-Input Multiple-Output (SIMO) components. Here the term "SIMO" represents the specific transmission system in which the input is a single source signal and the outputs are its transmitted signals observed at multiple microphones. Murata et al. have proposed FDICA-PB [3]. In this algorithm, first, the source signals are estimated as a monaural signal by FDICA, and then projection-back processing projects the source signals estimated by FDICA onto SIMO components of the observed signals using the inverse filter of the separation filter in FDICA. However, this algorithm has some disadvantages [4] as follows. First, the invertibility of the separation filter cannot be guaranteed. Thus, the inversion of the separation filter often fails and yields harmful results. Secondly, the circular convolution effect in FDICA is likely to cause the deterioration of separation performance. To solve these problems, the authors have proposed SIMO-model-based ICA (SIMO-ICA) [5]. The SIMO-ICA consists of multiple time-domain ICA (TDICA) parts and a fidelity controller. Since SIMO-ICA estimates SIMO components of the observed signals directly, inversion problem does not arise. Also, since SIMO-ICA is constructed of TDICA, it is free from the circular convolution problem. However, the convergence of SIMO-ICA is very slow, and the sensitivity to the initial settings of the separation filter is very high. In order to improve the decomposition performance, SIMO-ICA-SG is now studied [6]. This method consists of FDICA-PB, DOA estimation, and SIMO-ICA. First, we perform FDICA-PB to decompose the observed signals to some extent. After the FDICA-PB, we estimate the DOAs of sources using outputs of FDICA-PB. Then the proposed method resets the separation filter to the valid initial filter and re-optimizes the filter using both FDICA-PB and SIMO-ICA. In this procedure, a filter bank of previously measured head related transfer functions (HRTFs) for multiple DOAs is supplied to generate the valid initial filter.

As a preliminary study [6] on the proposed SIMO-ICA-SG, we carried out the SIMO separation experiment using a specific HRTF matrix bank which is matched with the mixing system. However, this only corresponds to the closed-data test, and the further open-data test remained as an open problem. In this paper, to evaluate robustness of the proposed method, decomposition experiments are carried out using the mismatched HRTF under the reverberant condition. The experimental results reveal that the proposed method can run robustly even if a mismatch arises between the mixing process and the HRTF matrix bank in the proposed method.

2. MIXING PROCESS

In this study, the number of microphones is $K = 2$ and the number of multiple sound sources is $L = 2$. In general, the observed signals in which multiple source signals are mixed linearly are expressed as $\boldsymbol{x}(t) = \sum_{n=0}^{N-1} \boldsymbol{a}(n)\boldsymbol{s}(t-n) = \boldsymbol{A}(z)\boldsymbol{s}(t)$, $\boldsymbol{s}(t) = [s_1(t), s_2(t)]^{\mathrm{T}}$ is the source signal vector and $\boldsymbol{x}(t) = [x_1(t), x_2(t)]^{\mathrm{T}}$ is the observed signal vector. Also, $\boldsymbol{a}(n) = [a_{kl}(n)]_{kl}$ is the mixing filter matrix with the length of N, and $\boldsymbol{A}(z) = [A_{kl}(z)]_{kl} = [\sum_{n=0}^{N-1} a_{kl}(n)z^{-n}]_{kl}$ is the z-transform of $\boldsymbol{a}(n)$, where z^{-1} is used as the unit-delay operator, i.e., $z^{-n} \cdot x(t) = x(t-n)$, a_{kl} is the impulse response between the k-th microphone and the l-th sound source, and $[X]_{ij}$ denotes the matrix which includes the element X in the i-th row and the j-th column.

* This research is partly supported by CREST program "Advanced Media Technology for Everyday Living" of Japan Science and Technology Agency.

0-7803-9154-3/05/$20.00 ©2005 IEEE

3. CONVENTIONAL SIMO SEPARATION METHODS

For the extraction of the SIMO components in the mixed signals, several SIMO separation methods using the ICA-based algorithm have been proposed, The aim of SIMO separation method is to decompose the mixed observed signals $\boldsymbol{x}(t)$ into the SIMO components of each independent sound source, that is, we estimate $A_{kl}(z)s_l(t)$ for all k and l (up to the permissible time delay in the separation filtering).

3.1. Frequency-Domain ICA with Projection Back (FDICA-PB)[3]

Murata et al. have proposed an FDICA-PB method which can estimate the SIMO components of the observed signals on the basis of the monaural outputs of FDICA. In this method, first, the separation filter matrix $\boldsymbol{W}_{(\mathrm{FD})}(f)$ in the frequency domain is optimized to the separate source signals to obtain the monaural signals. The separated signals $\boldsymbol{Y}_{(\mathrm{FD})}(f,t)$ in the time-frequency domain are expressed as

$$\boldsymbol{Y}_{(\mathrm{FD})}(f,t) = \boldsymbol{W}_{(\mathrm{FD})}(f)\boldsymbol{X}(f,t), \tag{1}$$

where $\boldsymbol{X}(f,t)$ is the observed signal vector which is calculated by means of a frame-by-frame discrete Fourier transform (DFT). The iterative learning algorithm is expressed as

$$\boldsymbol{W}^{[i+1]}_{(\mathrm{FD})}(f) = \eta\left\{\boldsymbol{I} - \left\langle \Phi(\boldsymbol{Y}^{[i]}_{(\mathrm{FD})}(f,t))\boldsymbol{Y}^{[i]}_{(\mathrm{FD})}(f,t)^{\mathrm{H}}\right\rangle_t\right\} \boldsymbol{W}^{[i]}_{(\mathrm{FD})}(f) + \boldsymbol{W}^{[i]}_{(\mathrm{FD})}(f), \tag{2}$$

where the initial value of $\boldsymbol{W}_{(\mathrm{FD})}(f)$ is given by Eq. (15). However, the output signals of FDICA given by Eq. (1) are monaural signals with respect to the sound sources, not SIMO-model-based signals. Thus, using the following equations, we project the separation filter onto the SIMO separation filters.

$$\boldsymbol{w}_{(\mathrm{PB1})}(n) = \mathrm{IDFT}\Big[\mathrm{diag}\left\{\boldsymbol{W}^{-1}_{(\mathrm{FD})}(f)\right\}\boldsymbol{W}_{(\mathrm{FD})}(f)\Big], \tag{3}$$

$$\boldsymbol{w}_{(\mathrm{PB2})}(n) = \mathrm{IDFT}\Big[\text{off-diag}\left\{\boldsymbol{W}^{-1}_{(\mathrm{FD})}(f)\right\}\boldsymbol{W}_{(\mathrm{FD})}(f)\Big], \tag{4}$$

where IDFT[$\cdot$] represents an inverse DFT with the time shift of the $D/2$ samples. The separated signals of FDICA-PB in the time domain are expressed as

$$\boldsymbol{y}_{(\mathrm{PB}i)}(t) = \sum_{n=0}^{D-1} \boldsymbol{w}_{(\mathrm{PB}i)}(n)\boldsymbol{x}(t-n). \tag{5}$$

3.2. SIMO-model-based ICA (SIMO-ICA)[5]

SIMO-ICA consists of the TDICA part and a *fidelity controller*, and the TDICA runs in parallel under the fidelity control of the entire separation system. The output signals of the TDICA part in SIMO-ICA are defined by

$$\boldsymbol{y}_{(\mathrm{TD})}(t) = [y_k^{(\mathrm{TD})}(t)]_{k1} = \sum_{n=0}^{D-1} \boldsymbol{w}_{(\mathrm{TD})}(n)\boldsymbol{x}(t-n), \tag{6}$$

where $\boldsymbol{w}_{(\mathrm{TD})}(n)$ is the separation filter matrix of the TDICA. Regarding the fidelity controller, the following signal vector is calculated, in which all of the elements are to be mutually independent,

$$\boldsymbol{y}_{(\mathrm{FC})}(t) = [y_k^{(\mathrm{FC})}(t)]_{k1} = \boldsymbol{x}(t-\frac{D}{2}) - \boldsymbol{y}_{(\mathrm{TD})}(t). \tag{7}$$

Hereafter, we regard $\boldsymbol{y}_{(\mathrm{FC})}(t)$ as an output of a *virtual* ICA, and define its virtual separation filter matrix as

$$\boldsymbol{w}_{(\mathrm{FC})}(n) = \boldsymbol{I}\delta(n-\frac{D}{2}) - \boldsymbol{w}_{(\mathrm{TD})}(n), \tag{8}$$

where $\delta(n)$ is a delta function, where $\delta(0)=1$ and $\delta(n)=0$ $(n \neq 0)$. From (8), we can rewrite (7) as

$$\boldsymbol{y}_{(\mathrm{FC})}(t) = \sum_{n=0}^{D-1} \boldsymbol{w}_{(\mathrm{FC})}(n)\boldsymbol{x}(t-n). \tag{9}$$

The reason why we use the word "virtual" here is that fidelity controller does not have own separation filters unlike the TDICA, and $\boldsymbol{w}_{(\mathrm{FC})}(n)$ is subject to $\boldsymbol{w}_{(\mathrm{TD})}(n)$. To explicitly show the meaning of the fidelity controller, (7) is rewritten as

$$\boldsymbol{y}_{(\mathrm{TD})}(t) + \boldsymbol{y}_{(\mathrm{FC})}(t) - \boldsymbol{x}(t-D/2) = [0]_{k1}. \tag{10}$$

Equation (10) means a constraint to force the sum of the all of output vectors $\boldsymbol{y}_{(\mathrm{TD})}(t) + \boldsymbol{y}_{(\mathrm{FC})}(t)$ to be the sum of all of the SIMO components $[\sum_{l=1}^{L} A_{kl}(z)s_l(t-D/2)]_{k1} (= \boldsymbol{x}(t-D/2))$. Here the delay of $D/2$ is used as to deal with nonminimum phase systems.

If the independent sound sources are separated by (6), and simultaneously the signals obtained by (7) are also mutually independent, then the output signals converge on unique solutions,

$$\boldsymbol{y}_{(\mathrm{TD})}(t) = [A_{11}(z)s_1(t-D/2), A_{22}(z)s_2(t-D/2)]^{\mathrm{T}}, \tag{11}$$

$$\boldsymbol{y}_{(\mathrm{FC})}(t) = [A_{12}(z)s_2(t-D/2), A_{21}(z)s_1(t-D/2)]^{\mathrm{T}}, \tag{12}$$

where diag$\{\boldsymbol{X}\}$ and off-diag$\{\boldsymbol{X}\}$ are the operation for setting every nondiagonal and diagonal elements of the matrix $\boldsymbol{X}$ to be zero. The proof of theorem and more details are given in [5]. Equations (11) and (12) represent necessary and sufficient SIMO components of all source signals.

In order to obtain the above-mentioned solutions, the natural gradient [1] of Kullback-Leibler divergence of (7) with respect to $\boldsymbol{w}_{(\mathrm{TD})}(n)$ should be added to the iterative learning rule of the separation filter in the TDICA. The iterative algorithm of the TDICA part in SIMO-ICA is given as

$$\begin{aligned}\boldsymbol{w}^{[j+1]}_{(\mathrm{TD})}(n) = \boldsymbol{w}^{[j]}_{(\mathrm{TD})}(n) - \alpha \sum_{d=0}^{D-1} \Big[&\text{off-diag}\left\{\left\langle \boldsymbol{\varphi}(\boldsymbol{y}^{[j]}_{(\mathrm{TD})}(t)) \boldsymbol{y}^{[j]}_{(\mathrm{TD})}(t-n+d)^{\mathrm{T}}\right\rangle_t\right\}\boldsymbol{w}^{[j]}_{(\mathrm{TD})}(d) \\ &- \text{off-diag}\left\{\left\langle \boldsymbol{\varphi}(\boldsymbol{y}^{[j]}_{(\mathrm{FC})}(t))\boldsymbol{y}^{[j]}_{(\mathrm{FC})}(t-n+d)^{\mathrm{T}}\right\rangle_t\right\} \Big(\boldsymbol{I}\delta(d-\frac{D}{2}) - \boldsymbol{w}^{[j]}_{(\mathrm{TD})}(d)\Big)\Big],\end{aligned} \tag{13}$$

where α is the step-size parameter, the superscript $[j]$ is used to express the value of the j-th step in the iterations, and $\langle\cdot\rangle_t$ denotes the time-averaging operator. In (13), the initial values of $\boldsymbol{w}_{(\mathrm{TD})}(n)$ and $\boldsymbol{w}_{(\mathrm{FC})}(n)$ are arbitrary, but should be different each other.

4. PROPOSED SIMO-ICA WITH SELF-GENERATOR FOR INITIAL FILTER [6]

The proposed algorithm is conducted by the following steps.

[Step 0: Early Initialization] Set DOAs of sources $\hat{\theta}_i$ to early initial (arbitrary) values, $\hat{\theta}_{\mathrm{init}i}$.

[Step 1: HRTF Matrix Bank] The HRTF matrix bank consists of multiple HRTF matrices. The single HRTF matrix for θ_1 and θ_2 is given as

$$\boldsymbol{H}(\theta_1,\theta_2,f) = \begin{bmatrix} H_L(\theta_1,f) & H_L(\theta_2,f) \\ H_R(\theta_1,f) & H_R(\theta_2,f) \end{bmatrix}, \tag{14}$$

where $H_L(\theta,f)$ (or $H_R(\theta,f)$) is the HRTF between the left (right) ear and the source whose direction is θ. To construct the HRTF matrix bank, we prepare the multiple HRTF matrices in advance by changing θ_1 and θ_2. Using the HRTF matrix bank and the DOAs of sources, we can automatically generate the valid initial value for FDICA as follows:

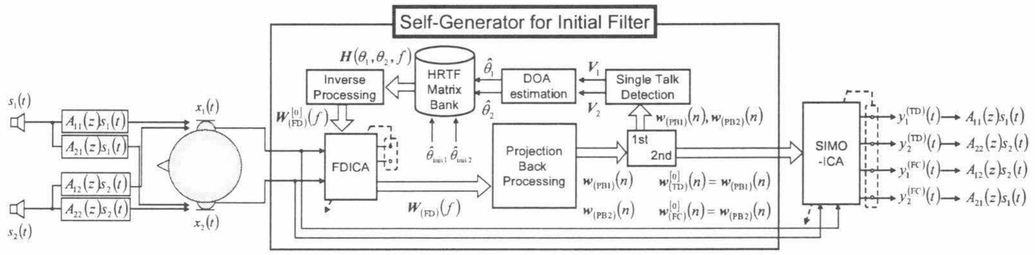

Figure 1: Example of input and output relations in the proposed method.

$$W^{[0]}_{(\mathrm{FD})}(f) = H^{-1}(\hat{\theta}_1, \hat{\theta}_2, f). \tag{15}$$

Note that the initial value is not an optimal separation filter matrix under a reverberant condition, but the separation filter matrix can be finally optimized through ICA iterations.

[Step 2: FDICA-PB [3]] Optimize the monaural separation filter $W_{(\mathrm{FD})}(f)$ using Eq. (2), and project it onto the SIMO separation filters $w_{(\mathrm{PB}i)}(n)$ using Eqs. (3) and (4).

[Step 3: Single Talk Detection] In order to detect the single talk segments of the observed signals, we divide the observed signals and output signals of FDICA-PB into multiple frames. Each frame of these signals is expressed as

$$x(u, v) = x(u + (v-1) \times U), \tag{16}$$

$$y_{(\mathrm{PB}i)}(u, v) = y_{(\mathrm{PB}i)}(u + (v-1) \times U), \tag{17}$$

where u is the time index in a frame, U is the number of samples in a frame, v is the frame index. Each single talk segment V_i of the observed signals is detected on the basis of the following criteria:

$$V_1 = \Big\{ v | Q_1^{(\mathrm{PB1})}(v) > T; Q_2^{(\mathrm{PB2})}(v) > T; \quad Q_2^{(\mathrm{PB1})}(v) < T; Q_1^{(\mathrm{PB2})}(v) < T \Big\}, \tag{18}$$

$$V_2 = \Big\{ v | Q_1^{(\mathrm{PB1})}(v) < T; Q_2^{(\mathrm{PB2})}(v) < T; \quad Q_2^{(\mathrm{PB1})}(v) > T; Q_1^{(\mathrm{PB2})}(v) > T \Big\}, \tag{19}$$

$$Q_k^{(\mathrm{PB}i)}(v) = 10 \log_{10} \frac{\sum_{u=1}^{U} |y_k^{(\mathrm{PB}i)}(u, v)|^2}{\max\limits_{v} \left\{ \sum_{u=1}^{U} |y_k^{(\mathrm{PB}i)}(u, v)|^2 \right\}}, \tag{20}$$

where T is a threshold which is experimentally determined.

[Step 4: DOA Estimation Using Single Talk Segments] We can obtain the DOAs $\hat{\theta}_i$ of sources by using the single talk segments. The estimated angle $\hat{\theta}_i$ is given as

$$\hat{\theta}_i = \arg\max_{\theta} \Big\{ \Big\langle \sum_f X_1(f, v) X_2(f, v)^{\mathrm{H}} e^{-\frac{j2\pi f d \sin\theta}{c}} \Big\rangle_{v \in V_i} \Big\}, \tag{21}$$

where $\langle \cdot \rangle_{v \in V}$ is the frame-averaging operator which is composed of elements v in single talk segments V. Thus, we can obtain the valid initial value of the separation filter matrix using these estimated values, $\hat{\theta}_1$ and $\hat{\theta}_2$.

[Step 5: Re-Optimization] Using the DOAs of the sources estimated with Eq. (21), we reset the separation filter to the initial value of the separation filter, and re-optimize that in FDICA-PB (execute [**Steps 1 and 2**] again).

[Step 6: SIMO-ICA] Optimize the separation filter matrices $w_{(\mathrm{TD})}(n)$ and $w_{(\mathrm{FC})}(n)$ in the time domain, by using Eq. (13) to enhance the target components further. The separation filter matrices (3) and (4) are used as the initial values of the separation filter matrix $w_{(\mathrm{TD})}(n)$ and $w_{(\mathrm{FC})}(n)$ in SIMO-ICA.

If the early initialization, HRTF matrix bank, FDICA-PB, and SIMO-ICA (**[Steps 0–2, 6]**) are executed without single talk detection, DOA estimation, nor re-optimization (**[Steps 3–5]**), this algorithm corresponds to the multistage SIMO-ICA (MS-SIMO-ICA) algorithm [7] which one of the authors have previously proposed.

5. EXPERIMENTS AND RESULTS

5.1. Conditions for Experiments

We carried out binaural-sound-separation experiments using source signals which are convolved with impulse responses recorded with a head and torso simulator (HATS) (Brüel & Kjær) in the experimental room. The reverberation time in this room is 200 ms. Two speech signals are assumed to arrive from different directions, θ_1 and θ_2; $\theta_1 = \{-90°, -75°, -60°, -45°, -30°, -15°, 0°\}$ and $\theta_2 = \{0°, 15°, 30°, 45°, 60°, 75°, 90°\}$. The distance between HATS and the sound source is 1.5 m. Two kinds of sentences, spoken by two male and two female speakers, are used as the original speech samples. Using these sentences, we obtain 12 combinations. The sampling frequency is 8 kHz and the length of speech is limited to 3 seconds. The length of $w(n)$ in each method is 1024, and the initial values are inverse filters of HRTFs whose directions of sources, $\hat{\theta}_{\mathrm{init1}}$ and $\hat{\theta}_{\mathrm{init2}}$, are $-60°$ and $60°$. The step-size parameters η and α are 5.0×10^{-2} and 1.0×10^{-6}. SIMO-model accuracy (SA) [8] is used as an evaluation score. The SA indicates the degree of similarity between the outputs of SIMO-ICA and the real SIMO-model-based signals.

5.2. Results and Discussion

We evaluate robustness of the proposed method against *mismatch* between the mixing process and the HRTF matrix bank. Our measured transfer function which contains both HRTF of the HATS and the room reverberation is used as the mixing system in this experiment. With respect to the HRTFs used in Step 2 of the proposed method, we replace the previous HRTF database measured by ourselves (hereafter we call this the *NAIST HRTF database*) with the alternative *MIT HRTF database* [9] which is recorded via KEMAR dummy head. The NAIST HRTF database is matched with the mixing system, but the MIT HRTF database is a mismatched one. Figure 2 shows the results of SA for different θ_1 and θ_2, where the following methods are compared; (a) the conventional FDICA-PB, for reference, (b) the conventional MS-SIMO-ICA, (c) the proposed method with the NAIST HRTF database (matched case), (d) the proposed method with the MIT HRTF database (mismatched case). As shown in Figures 2 (a)–(d), the performances of the proposed method with the mismatched HRTF database are still superior to those of the conventional FDICA-PB

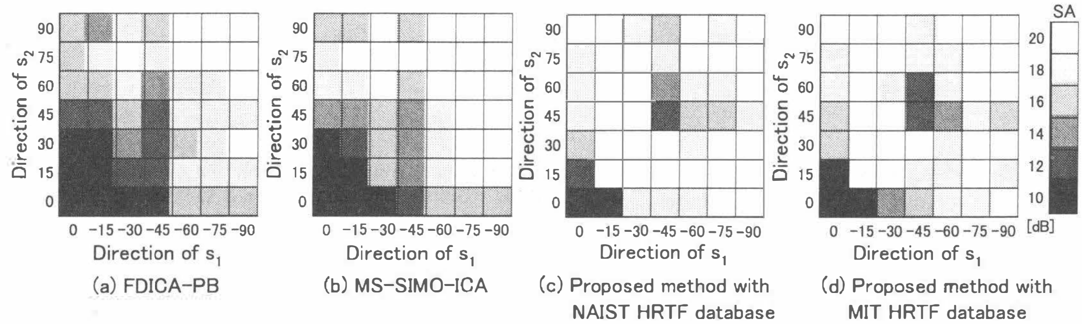

Figure 2: Experimental results of SA using different HRTF databases, where the mixing system is our measured transfer function which contains both HRTF of HATS and room reverberation. (a) FDICA-PB, (b) MS-SIMO-ICA, (c) proposed method with NAIST HRTF database (matched case), (d) proposed method with MIT HRTF database (mismatched case).

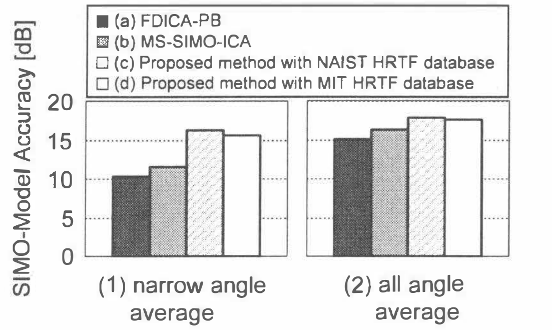

Figure 3: Experimental results of averaged SIMO-model accuracy in FDICA-PB, MS-SIMO-ICA, proposed method with NAIST HRTF database, and proposed method with MIT database when (1) the angle between the sources is narrow $(-45 \leq \theta_1 \leq 0, 0 \leq \theta_2 \leq 45)$ and (2) the angle is not limited.

and MS-SIMO-ICA. In composition to Figures 2 (c) and (d), the performances of the proposed method with the mismatched HRTF database are almost the same as those of the proposed method with the matched HRTF database. These results imply that the proposed method has sufficient robustness against the HRTF mismatch.

Figure 3 shows the results of averaged SA in conventional and proposed methods when (1) the angle between the sources is narrow and (2) the angle is not limited. From Figure 3 (1), compared with the conventional MS-SIMO-ICA, the improvements of SA in proposed method with NAIST and MIT database are 4.7 dB and 4.1 dB. From Figure 3 (2), compared with MS-SIMO-ICA, the improvements are 1.5. dB and 1.3 dB. Thus the proposed method's improvement is mainly dominant especially when the source angle is narrow, regardless the mismatch issue of the HRTF matrix bank.

Overall, the results indicate the proposed method can function robustly even if a mismatch arises between the mixing process and the HRTF matrix bank in the proposed method.

6. CONCLUSION

We discuss the SIMO separation problem which decompose the observed signals into the SIMO-model-based signals, not separate it into monaural source signals. In order to obtain the accurate SIMO-model-based signals, authors have proposed SIMO-ICA-SG which combines FDICA-PB, single-talk detection, DOA estimation, HRTF matrix bank, and SIMO-ICA. In this paper, in order to evaluate robustness of this method against the mismatch of HRTF matrix bank, we carried out the decomposition experiments using recorded in reverberant condition. The experimental results reveal that the decomposition performance of the proposed method with mismatched HRTF matrix bank is superior to those of the conventional methods, and almost the same as those of the proposed method with matched one.

7. REFERENCES

[1] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, John Wiley & sons, Ltd, West Sussex, 2002.

[2] J. Blauert, *Spatial Hearing (revised edition)*, Cambridge, MA: The MIT Press, 1997.

[3] N. Murata and S. Ikeda, "An on-line algorithm for blind source separation on speech signals," *Proc. NOLTA '98*, vol. 3, pp.923–926, 1998.

[4] T. Nishikawa, H. Saruwatari, K. Shikano, "Stable leaning algorithm for blind separation of temporally correlated acoustic signals combining multistage ICA and linear prediction," *IEICE Trans. Fundamentals*, vol.E86-A, no.4, pp.846–858, 2003.

[5] T. Takatani, T. Nishikawa, H. Saruwatari, K. Shikano, "High-fidelity blind source separation of acoustic signals using SIMO-model-based ICA with information-geometric learning," *Proc. IWAENC 2003*, pp.251–254, 2003.

[6] T. Takatani, S. Ukai, T. Nishikawa, H. Saruwatari, K. Shikano, "Blind Separation of Binaural Sound Mixtures Using SIMO-ICA with Self-Generator for Initial Filter," *Proc. EUSIPCO 2005* (accepted).

[7] S. Ukai, H. Saruwatari, T. Takatani, R. Mukai, H. Sawada, "Multistage SIMO-Model-Based Blind Source Separation Combining Frequency-Domain ICA and Time-Domain ICA," *Proc. ICASSP 2004*, pp.IV-109–IV-112, 2004.

[8] H. Yamajo, H. Saruwatari, T. Takatani, T. Nishikawa, and K. Shikano, "Evaluation of blind separation and deconvolution for convolutive speech mixture using SIMO-model-based ICA," *Proc. IWAENC 2003*, pp.299–302, 2003.

[9] http://sound.media.mit.edu/KEMAR.html