Evaluation of SIMO-model-based ICA with Self-Generator for Initial Filter for Blind Source Separation

June 23, 2005

T. Takatani

(Speech and Acoustics Lab.,
Graduate School of Information Science,
Nara Institute of Science and Technology)

tomoya-t@is.naist.jp
http://www.aist-nara.ac.jp/~tomoya-t/

Outline

Background & Research Goal
Mixing Process
Our Approach
Conventional Method
- Frequency-Domain Independent Component Analysis (ICA) with Projection Back Processing (FDICA-PB)
- Single-Input Multiple-Output (SIMO)-Model-Based ICA (SIMO-ICA)
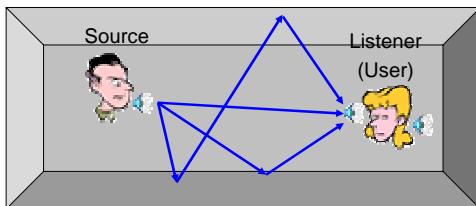- Problems of the conventional methods

Proposed Method
Experiments
Discussion
Conclusion

# Single-Input Multiple-Output (SIMO)-model-based signals

Generally speaking, human being do not hear the source signal itself. Because the signals we hear involve not only the information about source signal but also the information about spatial qualities.


The multiple signals from single source are defined as Single-Input Multiple-Output (SIMO)-model-based signals.
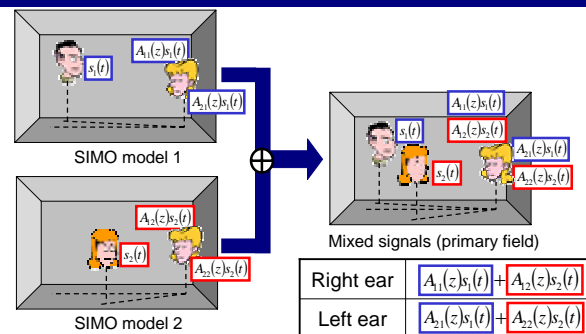

Transmission of acoustic signals
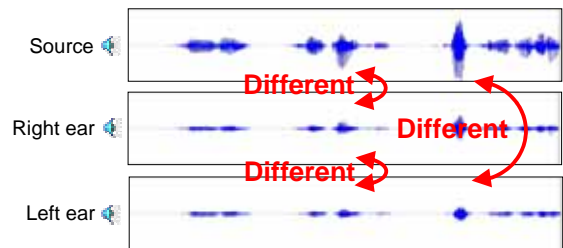
The SIMO-model-based signals in both ears include the information about the directivity, localization, or spatial qualities of each sound source.
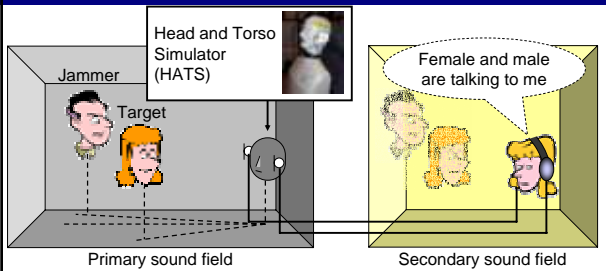

Mixing process (1/2)

Mixing process (2/2)

Right ear	$A_{11}(z)s_1(t) + A_{12}(z)s_2(t)$
Left ear	$A_{21}(z)s_1(t) + A_{22}(z)s_2(t)$
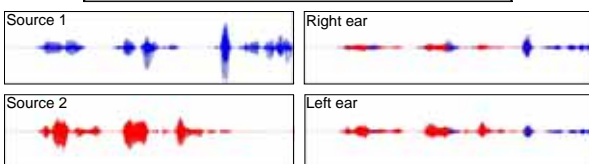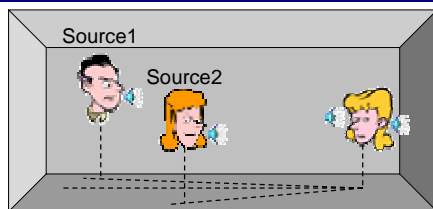
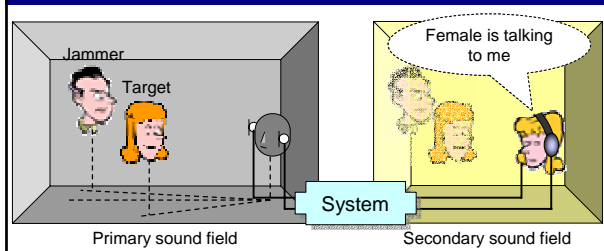**The mixed signals are sum of the SIMO-model-based signals.**

Binaural sound reproduction system

Binaural sound reproduction system can copy the primary sound field.

This system can not reproduce only the target sound.

Research goal

Realization of audio augmented reality system which can enhance and reproduce the target sound

Application

Tele-operation system using vision and audio modality


robot   rescue personnel (operator)

In order to represent only target sound to an operator, the followings are required.

High-fidelity sound (Sound reproduction)
Noise suppression (Enhancement of the target sound)

Conventional Blind Decomposition Method

- Conventional Method 1:
  Frequency-Domain Independent Component Analysis (ICA) with Projection Back (FDICA-PB)
  [Murata, *NOLTA1998*]

- Conventional Method 2:
  Multistage SIMO-Model-Based ICA (MS-SIMO-ICA)
  [Ukai, *ICASSP2004*]

- Proposed Method:
 SIMO-ICA with Self-Generator for Initial Filter

Frequency-Domain ICA with Projection Back (FDICA-PB)
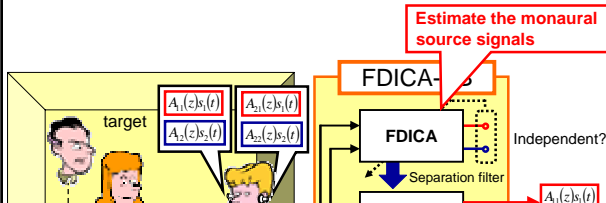
Disadvantage
- Invertibility of the separation filter cannot be guaranteed.
- Circular convolution effect causes the deterioration of separation performance.
- The good initial filter is needed, but the valid filter setting is difficult in advance.

Multistage SIMO-Model-Based ICA (MS-SIMO-ICA)
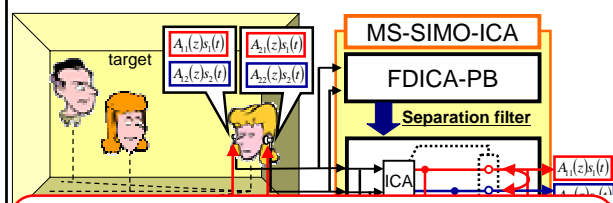
Advantage
- Even if circular convolution effect causes the deterioration of separation performance in FDICA-PB, SIMO-ICA can correct it.

Disadvantage
- The good initial filter is needed, but the valid filter setting is difficult in advance.

Problems of Conventional Methods

	FDICA-PB	SIMO-ICA
Inversion problem	Arise	Not arise
Circular convolution problem	Arise	Not arise
Does it have the strategy for the initial filter?	No	No

The development of the self-generation of good initial filters is a problem demanding attention.

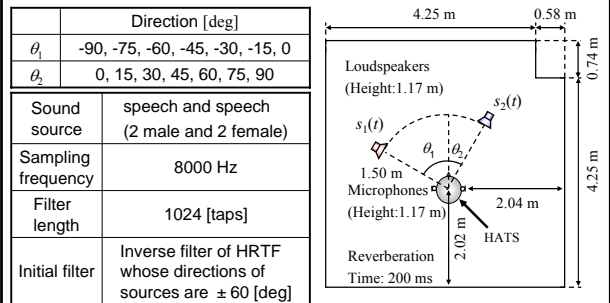## SIMO-ICA with Self-Generator for Initial Filter (SIMO-ICA-SG)
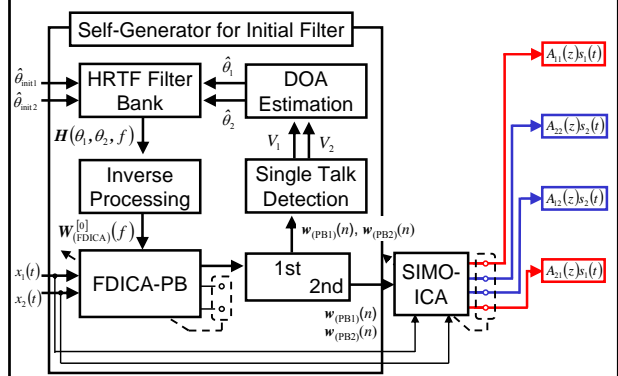
Comparison Among the Conventional and Proposed Methods

	FDICA-PB	SIMO-ICA	Proposed Method
Inversion problem	Arise	Not arise	Not arise
Circular convolution problem	Arise	Not arise	Not arise
Does it have the strategy for the initial filter?	No	No	Yes

Experiment

In order to evaluate its effectiveness, the decomposition experiments of the binaural mixed signals are carried out for the sources at the different directions.

	Direction [deg]
θ_1	-90, -75, -60, -45, -30, -15, 0
θ_2	0, 15, 30, 45, 60, 75, 90
Sound source	speech and speech (2 male and 2 female)
Sampling frequency	8000 Hz
Filter length	1024 [taps]
Initial filter	Inverse filter of HRTF whose directions of sources are ± 60 [deg]

Evaluation Score: SIMO-Model Accuracy (SA)

Similarity between the output signals of the decomposition processing and the real SIMO-model-based signals

$$\text{SA [dB]} = 10\log_{10} \frac{\left\langle \left| \text{Real SIMO - model - based signal} \right|^2 \right\rangle_t}{\left\langle \left| \text{Real SIMO - model - based signal} - \text{Output signal} \right|^2 \right\rangle_t}$$

Experimental Results

- **The white area in proposed method is much more than those in conventional methods.**
- **The performance of the proposed method can be remarkably improved, especially when the angle between the speakers is narrow, e.g., θ_1 = - 45 ~0 and θ_2 = 0 ~45 .**
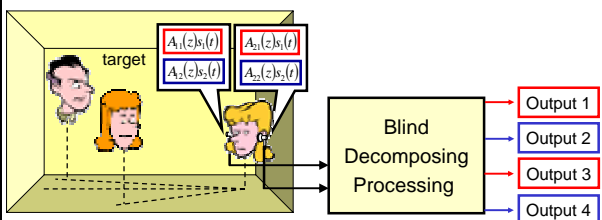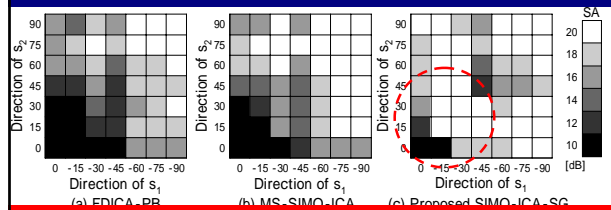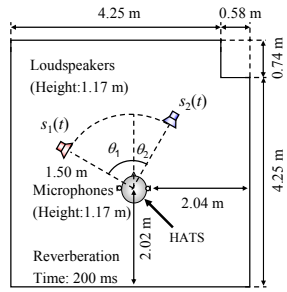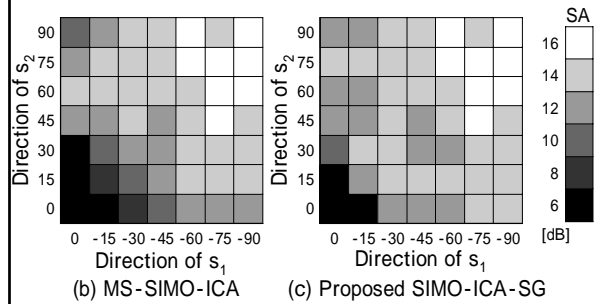
Experiment

In order to evaluate its effectiveness, the decomposition experiments of the binaural mixed signals are carried out for the sources at the different directions.

	Direction [deg]
θ_1	-90, -75, -60, -45, -30, -15, 0
θ_2	0, 15, 30, 45, 60, 75, 90

Sound source	speech and stationary noise
Sampling frequency	8000 Hz
Filter length	1024 [taps]
Initial filter	Inverse filter of HRTF whose directions of sources are ± 60 [deg]

4.25 m   0.58 m

Loudspeakers (Height:1.17 m)

$s_2(t)$

$s_1(t)$

θ_1 θ_2

1.50 m

Microphones (Height:1.17 m)

2.04 m

HATS

Reverberation Time: 200 ms

0.74 m

4.25 m

2.02 m

Experimental Results

(b) MS-SIMO-ICA   (c) Proposed SIMO-ICA-SG

Direction of s_2 (vertical axis): 0, 15, 30, 45, 60, 75, 90
Direction of s_1 (horizontal axis): 0, -15, -30, -45, -60, -75, -90

SA: 16, 14, 12, 10, 8, 6 [dB]

• The performance of the proposed method is almost the same as that of the conventional method.

Conclusion

- We propose the novel blind decomposition algorithm which combines FDICA-PB, DOA estimation, and SIMO-ICA.
- This proposed method have the advantage which can generate the valid initial value of separation filter **blindly**.
- In order to evaluate its effectiveness, separation experiments are carried out under the reverberant condition.
  - ◆ Speech and Speech
    The performance of proposed method is superior to those of conventional methods.
  - ◆ Stationary noise and Speech