Blind Signal Separation of Moving Sound Sources

Masanori Ito, Yoshinori Takeuchi, Tetsuya Matsumoto,
Hiroaki Kudo, Noboru Ohnishi

Graduate School of Information Science, Nagoya University
{ito,takeuchi,matumoto,kudo,ohnishi}@ohnishi.nuie.nagoya-u.ac.jp

Many separation algorithms have been proposed in the field of source separation. However, most of them cannot separate sources in the real environment. Furthermore, few methods can deal with cases involving source movement. When a source is moving, the mixing matrix also varies with time, therefore divide the mixed signals into blocks and compute the separation matrix for each block.

To estimate the separating matrix, we adopted a Second-Order Statistics (SOS) algorithm based on nonstationarity of source signals, which works in time domain. When a source moves, the mixing matrix varies, so we divide the mixed signals into short time blocks and assume that, in each of the blocks, movement is small enough to be negligible. We then separate each block of the mixed signals. Finally, we combine the separated signals of each block together.

We propose using directional microphones and close microphone arrangement to separate moving sources, setting the tips of directional microphones at the same point but directed toward different directions. We call this arrangement close arrangement, while we call the conventional method nonclose arrangement. As a source moves, the mixing matrix also varies. To obtain separated signals we must compute the separating matrix according to the variation of mixing matrix. Using nonclose arrangement, there are delays from a source to microphones, that is, microphones observe the same signal at different times, and these delays affect the convergence of the separating matrix. When a source moves, delays vary as well, thus if the variation of a delay varies greatly, then so must the separating matrix. For example even if the source moves only 10 cm, the separation matrix before move might be completely different to the matrix after the move. The reason why we use close arrangement is that we want to make all the delay times be zero. Therefore, the variation of separating matrix should be small.

Close arrangement carries another advantage: can reduce the filter size of the separating matrix and lessen computation time.

We conducted an experiment of separation of two moving sources in a reverberant environment. Source 1 moved at 10 cm/sec and source 2 moved between 3 and 6 sec. We evaluated the separation performance with respect to the Signal-to-Interference Ratio (SIR), finding that peformances with close arrangement were better than those with nonclose arrangement. Close arrangement made the drop in performance small.

For blind separation we do not extract a target signal but suppress interference signals, that is, it is difficult to suppress the moving source.

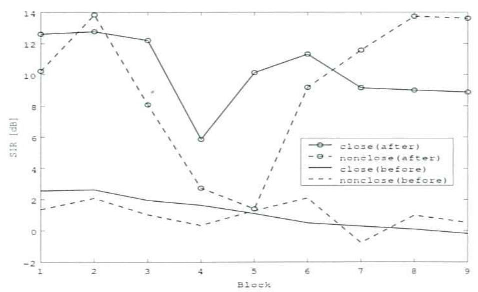


Figure 1: Performance of source 1

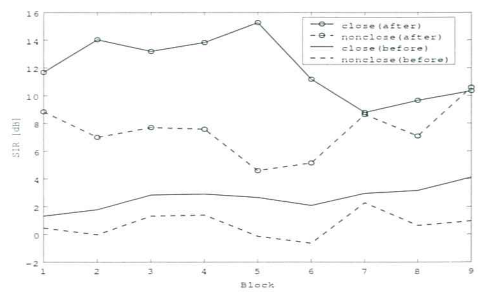


Figure 2: Performance of source 2

Blind Signal Separation of Moving Sound Sources

Masanori Ito, Yoshinori Takeuchi , Tetsuya Matsumoto, Hiroaki Kudo, Noboru Ohnishi

1

Background

- Source separation is the technique reconstructing source signals using only observed signals.
- Various source separation methods have been proposed for convolutive mixture.
  - M. Kawamoto et al. (1998), N. Murata et al. (1998)
- Applications of source separation
  - Pre-processing for speech recognition
  - Hands-free telephone system
  - Auditory scene analysis
  - Robot ears

2

Objective

- Most algorithms assume that sources' positions are fixed. (Invariant System)

⇩

- Separating mixed signals from moving sources. (Time-variant system)
  - The number of sources is 2
  - In the real environment

3

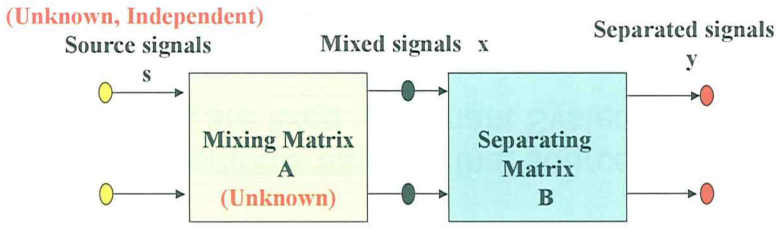
Moving-source Separation

- Mixing matrix and Separating Matrix are time variant.
- We divide the mixed signals and assume that sources are fixed in the divided periods.
- We separate each divided signal.
- Division is small ⇨ Convergence is difficult
- Division is long ⇨ Move distance is too long

4

Separation Algorithm

- Minimizing the correlation of output signals based on the non-stationarity of source signals. (Kawamoto et al. 1998)

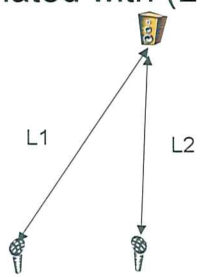


5

Separating Parameters

- The convergence of separating parameters is related with (L1-L2).

(Ito et al. 2002
Proc. of ICONIP '02)

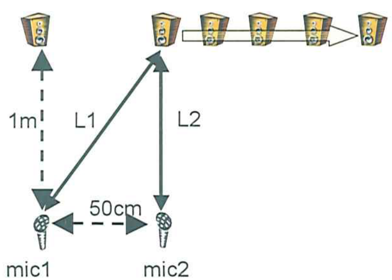


Depend on the delayed time

6

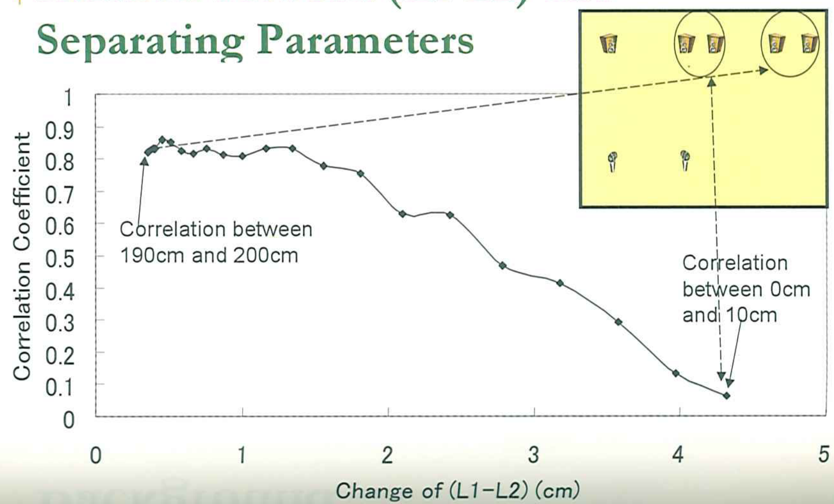
Change of Separating Parameters

- We moved source2 in every 10cm, observed mixed signals and separated them.



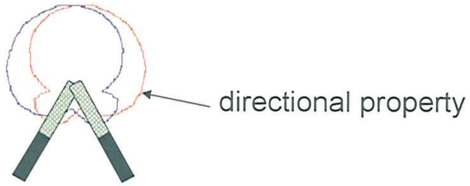
- We examined the relation between (L1-L2) and separating parameters.

Relation between (L1-L2) and Separating Parameters



Close Mic Arrangement

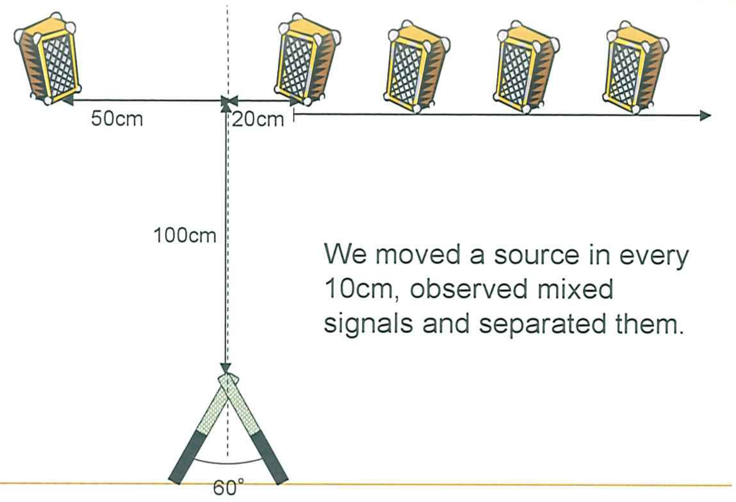
- Tips of two microphones are set at the same point.



- L1=L2 is always satisfied.
- Because of the directional property, two microphones observe different signals.

9

Experiment in close arrangement



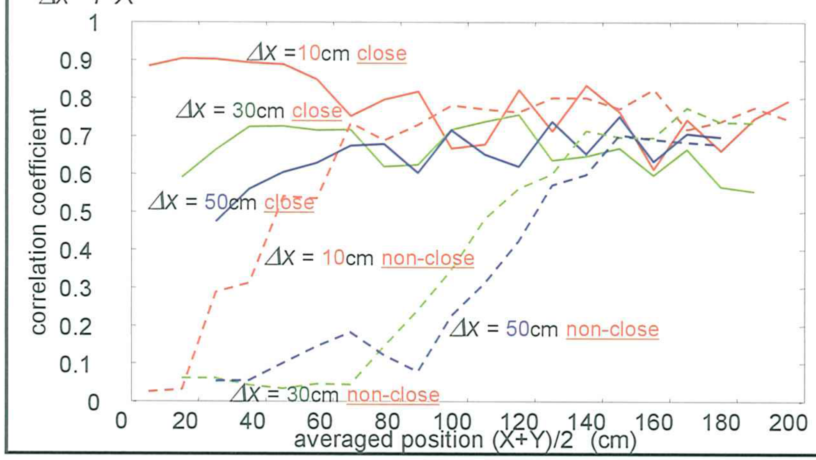
We moved a source in every 10cm, observed mixed signals and separated them.

10

Change of separating parameters in close mic arrangement

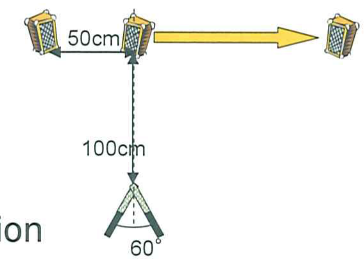
Correlation of parameters at *position X* with those at *position Y*.

$\Delta x = Y - X$

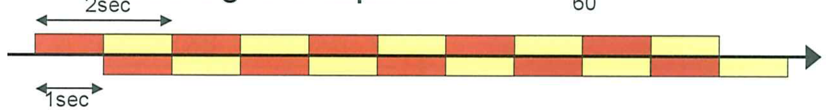


Separation Experiment

- Source2 moves at 10cm/sec
  - close, non-close mic arrangement



- divided signals separation



- all-time signals separation

12

Experimental Setup

- Experiment Room

  8.3m * 7.7m * 2.2m

  Real Environment

  Reverberation time 0.81sec
- Sampling Frequency 16kHz
- Source1 Fixed
- Source2 Moved (on Robot)

13

Separation Performance

Performance using all 20sec (dB)		Close		Non-Close	
		Fixed	Moving	Fixed	Moving
Input SNR		6.1	-3.1	6.6	-4.3
Output SNR	divide	10.9	6.8	8.8	5.3
	all	7.0	6.2	7.4	5.4

Performance using first 5sec (dB)		Close		Non-Close	
		Fixed	Moving	Fixed	Moving
Input SNR		2.1	-0.3	3.0	-1.0
Output SNR	divide	7.5	10.2	3.6	9.0
	all	1.1	10.5	2.8	10.0

Why is extracting fixed source difficult?

- When source2 is moving,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}(z) & a_{12}(z,t) \\ a_{21}(z) & a_{22}(z,t) \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$$

- When the separation is done, the product BA becomes diagonal;

$$a_{12}(z,t) + b_{12} a_{22}(z,t) = 0$$

$$b_{21} a_{11}(z) + a_{21}(z) = 0$$

- b_{12} is time-variant and used to extract the fixed source.

15

Conclusions

- Explained the method to separate mixed signals from moving sources.
- Proposed using directional microphones and "Close mic arrangement"
- Separated mixed signals from moving sources in the reverberant environment.

16