BLIND SOURCE SEPARATION AND DOA ESTIMATION USING SMALL 3-D MICROPHONE ARRAY

Ryo Mukai *Hiroshi Sawada* *Shoko Araki* *Shoji Makino*

NTT Communication Science Laboratories, NTT Corporation
2–4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619–0237, Japan
Email: {ryo, sawada, shoko, maki}@cslab.kecl.ntt.co.jp
URL: http://www.kecl.ntt.co.jp/icl/signal/mukai/

We present a prototype system for Blind Source Separation (BSS) of many speech signals. Our system uses 8 microphones located at the vertexes of a 4cm×4cm×4cm cube and has the ability to separate signals distributed in three-dimensional space (Fig. 1). The mixed signals observed by the microphone array are processed by Independent Component Analysis (ICA)[1] in the frequency domain and separated into a given number of signals (up to 8). The system estimates the directions of arrival (DOA) of the source signals as a by-product of the separation process. In our previous work [2], we described the separation of 6 source signals of simulation data, i.e. signals made by convolving impulse responses. In contrast, this prototype system performs an on-the-spot BSS of live recorded signals. The system specifications are summarized in Table 1. The key technologies used in our system are as follows.

Frequency Domain BSS using ICA

ICA is one of the main statistical methods used for BSS. In a reverberant environment, the signals are mixed in a convolutive manner with reverberations, and the separation system is a matrix of filters. The frequency domain approach decomposes a convolutive mixture into multiple instantaneous mixtures. Then, we apply an ordinary (instantaneous) ICA algorithm to each frequency bin and calculate the separation matrices. The time domain filters are obtained by the inverse discrete Fourier transform of frequency domain separation matrices. The computation cost of the frequency domain approach is much less than that of the time domain approach, where ICA is applied directly to the convolutive mixture model.

DOA estimation using ICA solution

We can estimate the DOA of a frequency component by using the separation matrix obtained by ICA. The inverse of the separation matrix can be assumed to be an estimation of the mixing system up to the permutation and scaling ambiguity. An interpretation of the estimated mixing system by the farfield model yields DOAs of separated signals relative to the

Figure 1: Prototype system (8 microphones and 6 loudspeakers)

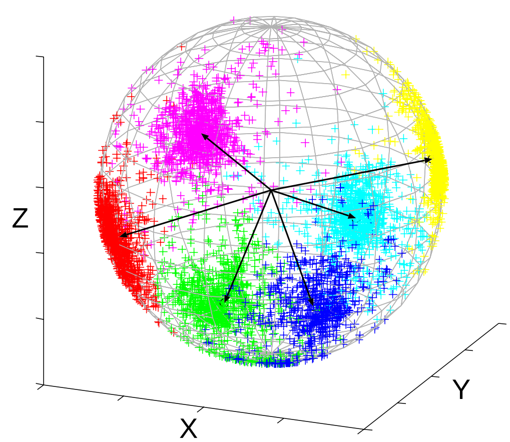

Figure 2: Estimated DOAs and clustered frequency components

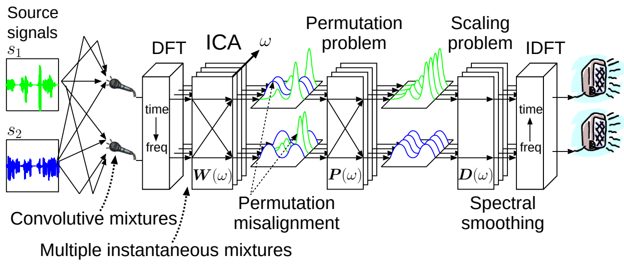

Figure 3: Flow of frequency domain BSS

Table 1: Specifications of prototype system

Microphone	8 omni-directional microphones
Sampling rate	8 kHz
Frame length	2048 point (256 ms)
Frame shift	512 point (64 ms)
ICA algorithm	FastICA + Infomax (complex valued)
CPU	AMD Athlon (2.4 GHz)
Coding	MATLAB
Computation time	46 s for 8 s, 6 sources

microphone axis [3]. Actually, the relative direction determines a cone on which a source signal exists. By combining multiple cones estimated by using multiple microphone pairs, we can estimate the absolute direction of a source signal in three-dimensional space. The procedure is described in detail in [2]. Figure 2 shows an example of a DOA estimation result. Each point plotted on a unit sphere denotes the estimated DOA of a frequency component in one frequency bin. The points are clustered by the k-means method, then the DOAs of source signals are given as the centroids of the clusters. This information is useful for solving the permutation problem described below.

Permutation solver using DOA and correlation

The permutation problem is the most critical issue as regards frequency domain BSS. Figure 3 shows the flow of frequency domain BSS and the permutation problem. Swapping output signals in the frequency domain does not affect the independence, thus there is a permutation ambiguity in the ICA solution. Before constructing output signals in the time domain, we have to align the permutation so that each channel contains frequency components from one source signal. There are two major approaches for solving this problem: the DOA based method and the correlation based method. The estimated DOA is useful for solving the permutation problem, however the estimation suffers from errors in a reverberant environment and the classification according to the DOA is inconsistent in some frequency bins. Thus we employ the correlation based method for such frequency bins. The combination of these two methods provides a good solution [4].

Spectral smoothing

Frequency domain BSS is influenced by the circularity of the discrete frequency representation. This causes a problem when we convert separation matrices in the frequency domain into a separation filter in the time domain. This problem is not apparent when there are two sources, however it is crucial when the number of source signals exceeds two. Our technique for solving this problem involves spectral smoothing of separation filters by using a window that tapers smoothly to zero at each end. The direct application of windowing changes the frequency responses for separation obtained by ICA and causes an error. Therefore, we adjust the frequency responses before windowing so that the error is minimized. The procedure is presented in detail in [5].

We carried out experiments in an ordinary office room and obtained good separation performance. Some sound examples can be found on our web site (http://www.kecl.ntt.co.jp/icl/signal/mukai/demo/hscma2005/).

References

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[2] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Frequency domain blind source separation for many speech signals," in *Proc. ICA2004 (Lecture Notes in Computer Science 3195)*, pp. 461–469. Springer-Verlag, 2004.

[3] H. Sawada, R. Mukai, and S. Makino, "Direction of arrival estimation for multiple source signals using independent component analysis," in *Proc. ISSPA 2003*, 2003, vol. 2, pp. 411–414.

[4] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech Audio Processing*, vol. 12, no. 5, pp. 530–538, 2004.

[5] H. Sawada, R. Mukai, S. de la Kethulle, S. Araki, and S. Makino, "Spectral smoothing for frequency-domain blind source separation," in *Proc. IWAENC 2003*, 2003, pp. 311–314.