Blind Separation Combined Frequency Invariant Beamforming and ICA for Far-field Broadband Acoustic Signals

Qi Lv[1], Xianda Zhang[1,⋆], and Ying Jia[2]

[1] Dept. of Automation, Tsinghua University, Beijing 100084, China
[2] Intel China Research Center, Beijing, China

Abstract. Many famous blind source separation (BSS) in frequency domain have been developed while they can still not avoid the permutation problem. We propose a new BSS approach for far-field broadband acoustic signals via combining the frequency invariant bemforming (FIB) technique and complex-valued independent component analysis (ICA). Compared with other frequency methods, our method can avoid the permutation problem and has much faster convergency rate. We also present a new performance measure to evaluate the separation. Finally, the simulation is given to verify the efficiency of the proposed method.

1 Introduction

Blind source separation (BSS) is a method for recovering independent source signals from their mixtures without any prior knowledge of signals and mixing process besides some statistical features [1]. Since the pioneering work by Jutten and Herault [2], BSS has drawn lots of attention in signal processing community and neural networks community [1], [3] ,[4] ,[5] ,[6].

Early BSS studies dealt with an instantaneously mixing process [5] ,[6], while recent reports are mainly concerned with convolutive mixtures [4] ,[7] ,[8] ,[9] which is much more difficult from theoretical and computational points of view. Roughly speaking, BSS methods for convolutive mixtures can be classified into the two types: the time domain approach and frequency domain one.

In time domain, BSS problem can be solved by applying independent component analysis (ICA) directly to the convolutive mixtures model [7]. This type of BSS methods can avoid the permutation indeterminacy which can hardly be avoided in frequency domain and can achieve good separation once the used algorithm converges. Its disadvantage is that ICA for convolutive mixtures is not as simple as ICA for instantaneous mixtures and computationally expensive for long FIR filters because it includes convolution operations [4].

In frequency domain, the convolutive mixtures problem of time domain is converted into instantaneous mixtures problem at each frequency. Hence, the

⋆ This work was supported by the major programm of the National Natural Science Foundation of China under Grant 60496311 and the Intel Sponsored Research Foundation.

J. Wang, X. Liao, and Z. Yi (Eds.): ISNN 2005, LNCS 3497, pp. 538–543, 2005.

complex-valued ICA can be applied at each frequency bin [4]. The merit of these approaches is that the ICA algorithm becomes simple and can be performed separately at each frequency [4]. However, the indeterminacy of permutation and gain of the ICA solution becomes a serious problem. Luckily, many permutation correction approaches have been proposed [8] ,[9]. Recently, Sawada et. al. have proposed a robust and precise method for solving the permutation problem [4]. But it is a pity that their approach computationally expensive for permutation correction.

In this paper, we propose a frequency domain approach for blind separation of mixtures of acoustic signals in far field. Our approach consists of two steps: 1) filter using the frequency invariant beamforming (FIB); 2) separation using the complex-valued ICA algorithm in frequency domain. The merit of our approach is that it avoid the permutation problem and has good separated results. We also proposed a new performance measure to evaluate the quality of separation.

The rest of this paper is organized as follows. In section 2 we propose our separation method. After that, a new performance measure is discussed in section 3. The simulation result is given in section 4. Finally, this paper is concluded in Section 5.

2 New BSS Method for Acoustic Signals

In this section, we present our new BSS architecture which has two parts: FIB design and separation matrix design.

We first assume that there are q sensors in a linear array and p unknown sources in far field emitting acoustic signals from direction $\mathbf{\Theta} = [\theta_1, \cdots, \theta_p]$, where θ_i is the direction to the ith source measured relative to the array axis. And we also assume $p < q$. The discrete time signal received at jth sensor is given by

$$x_j[k] = \sum_{i=1}^{p} s_i[k - \tau_j(\theta_i)] + v_j[k] \tag{1}$$

where $v_j[k]$ is the addition white noise, $s_i[k]$ is the ith source signal, and $\tau_j(\theta_i) = d_j \sin\theta_i / c$ is the propagation delay of ith source to the jth sensor, d_j is the position of the jth sensor, c is the propagation velocity of the signals. Define the q-dimensional vector of stacked array data as

$$\mathbf{x}[k] = [x_1[k], \cdots, x_q[k]] \tag{2}$$

with a frequency response given by

$$\mathbf{X}(f) = \mathbf{A}(\mathbf{\Theta}, f)\mathbf{S}(f) + \mathbf{V}(f) \tag{3}$$

where $\mathbf{S}(f) = [s_1(f), \cdots, s_p(f)]$ is the source signal vector in frequency, $\mathbf{V}(f) = [v_1(f), \cdots, v_q(f)]$ is the additive noise vector, and $\mathbf{A}(\mathbf{\Theta}, f)$ is the $q \times p$ source direction matrix with its element $a_{ij} = e^{-j2\pi f \tau_i(\theta_j)}$. The BSS problem is to recover $\mathbf{s}[k] = [s_1[k], \cdots, s_p[k]]$ using only $\mathbf{x}[k]$. Equation (3) indicates that the mixing process at each frequency is instantaneous. Then, the complex-valued ICA approach can be applied at each frequency to solve the BSS problem [4]

$$\mathbf{Y}(f) = \mathbf{W}(f)\mathbf{X}(f) \tag{4}$$

where $\mathbf{Y}(f) = [y_1(f), \cdots, y_p(f)]$ is the separated result in frequency domain, $\mathbf{W}(f)$ is a $p \times q$ separation matrix. The learning rule of the separation matrix is given by [4]

$$\Delta\mathbf{W} = \mu\left[\mathbf{I} - \left\langle \mathbf{\Phi}(\mathbf{Y})\mathbf{Y}^H \right\rangle_t\right]\mathbf{W} \tag{5}$$

where μ is a step-size parameter, $\langle\cdot\rangle_t$ denotes the averaging operator overtime, and $\mathbf{\Phi}(\cdot)$ is some nonlinear function.

For the mixing matrix is different at different frequency, the ICA approach can not avoid the permutation problem (see [4]). If the the mixing matrix is essentially identical for all frequencies, we can serially update the separation matrix from one frequency to another. And the permutation problem need not be considered. Luckily, we have the FIB technique to realize that. The main idea of FIB is to design a filter $\mathbf{b}(f)$ such that the response of this beamformer may be made approximately constant with respect to frequency over the design bandwidth $f_L \sim f_U$, i.e.,

$$r(\theta, f) = \mathbf{b}^H(f)\mathbf{a}(\theta, f) \approx r_{FI}(\theta), \quad \forall\theta, \forall f \in [f_L, f_U] \tag{6}$$

where $\mathbf{a}(\theta, f) = [e^{-j2\pi f\tau_1(\theta)}, \cdots, e^{-j2\pi f\tau_q(\theta)}]^T$, and $r(\theta, f)$ is the response of beamformer. Several methods of designing a FIB have been proposed [10] ,[11]

After we apply an FIB to the received array data, the beamformer output in frequency is

$$\mathbf{Z}(f) = \mathbf{B}^H(f)\mathbf{X}(f) \tag{7}$$

where $\mathbf{Z}(f)$ is referred as the frequency invariant beamspace (FIBS) data observation vector, $\mathbf{B}(f) = [\mathbf{b}_1(f), \cdots, \mathbf{b}_p(f)]$ is $q \times p$ filter response matrix, and $\mathbf{b}_i(f)$ is the ith set of beam shaping filter response vector. By using (3), the FIBS data vector can be rewritten as

$$\begin{aligned}\mathbf{Z}(f) &= \mathbf{B}^H(f)\mathbf{X}(f)\\ &= \mathbf{B}^H(f)\mathbf{A}(\mathbf{\Theta}, f)\mathbf{S}(f) + \mathbf{B}^H(f)\mathbf{V}(f)\\ &= \mathbf{A}_B(\mathbf{\Theta}, f)\mathbf{S}(f) + \mathbf{V}_B(f)\end{aligned} \tag{8}$$

where $\mathbf{A}_B(\mathbf{\Theta}, f) = \mathbf{B}^H(f)\mathbf{A}(\mathbf{\Theta}, f)$ is the $p \times p$ FIBS source direction matrix, and $\mathbf{V}_B(f) = \mathbf{B}^H(f)\mathbf{V}(f)$ is the $p \times 1$ FIBS noise vector.

Because the beamformers are designed to satisfy the frequency invariant property (6), the FIBS source direction matrix is approximately constant for all frequencies within the designed band, i.e., $\mathbf{A}_B(\theta, f) \approx \mathbf{A}_B(\theta)$, $\forall f \in [f_L, f_U]$. Hence, the mixing process of acoustic signals is completely characterized by a single beamspace source direction matrix $\mathbf{A}_B(\theta)$ which is independent of frequency and only decided by the direction-of-arrival (DOA) of the source signals.

After FIB preprocessing, we separate source signals from FIBS data observation vector $\mathbf{Z}(f)$ which can be rewritten as the following mixing model:

$$\mathbf{Z}(f) \approx \mathbf{A}_B(\theta)\mathbf{S}(f) + \mathbf{V}_B(f) \tag{9}$$

Then, we use method (5) to update the separation matrix $\mathbf{W}$ serially from frequency f_L to f_U.

Our proposed method, compared with general frequency domain BSS methods [4] ,[8] ,[9], can avoid the indeterminacy of permutation and scaling because the mixing process with FIB applied is invariant through a broad frequency band. On the other hand, our approach has much faster convergence rate since it uses the information of all the frequency data to update the separation matrix serially.

3 Performance Measure

In this section, we discuss how to measure the quality of separation. We first define

$$k_i(f) \triangleq \arg\max_k |\mathbf{C}_{ik}(f)| \tag{10}$$

$$m_i(f) \triangleq \arg\max_m |\mathbf{C}_{mi}(f)| \tag{11}$$

where $\mathbf{C}(f) = \mathbf{W}(f)\mathbf{A}_B(\theta, f)$. Then we define k_i and m_i as the most frequently occurring numbers of group $k_i(f)$ and $m_i(f)$ respectively. A new performance measure ($\mathbf{P_m}$) of separation is given by

$$\mathbf{P_m} = \frac{\sum_{f=1}^{f_s}\sum_{i=1}^{p}\left(\left|\mathbf{C}_{ik_i}(f)\right|^2 \left|s_{k_i}(f)\right|^2 + \left|\mathbf{C}_{m_i i}(f)\right|^2 \left|s_i(f)\right|^2\right)}{\sum_{f=1}^{fs}\sum_{i=1}^{p}\left(\sum_{j=1,j\neq k_i}^{p}\left|\mathbf{C}_{ij}(f)\right|^2 \left|s_j(f)\right|^2 + \sum_{j=1,j\neq m_i}^{p}\left|\mathbf{C}_{ji}(f)\right|^2 \left|s_i(f)\right|^2\right)} \tag{12}$$

where f_s is the sampling frequency. The proposed performance measure of separation, in a sense, describes the average ratio of the total separated signal power to the total interference power.

4 Simulation Result

To demonstrate the efficiency of the proposed approach, we consider two speech signals (see Fig. 1) of 30s with the sampling frequency $f_s = 48$k Hz emitting from $-5°$ and $20°$ respectively. The source signals can be downloaded from the internet address http://medi.uni-oldenburg.de/demo/demo_separation.html.

The frame size of short time Fourier transform (STFT) is 1600, and the STFT overlap is 1200. Two FIB's were designed according to [11] to be frequency invariant over the frequency band $[300, 3400]$ Hz. The aperture size is 5 half-wavelengths. Then, 17 sensors are needed at least and the array is approximately 2.8m long. We compared our method with the one in [4]. The step size in equation (5) is 0.0005 for our approach while it is 0.01 for the method in [4]. The nonlinear function is selected as $\mathbf{\Phi}(\cdot) = e^{j\cdot\arg(\cdot)}$. The separation performance defined by equation (12) is plotted in Fig. 2.

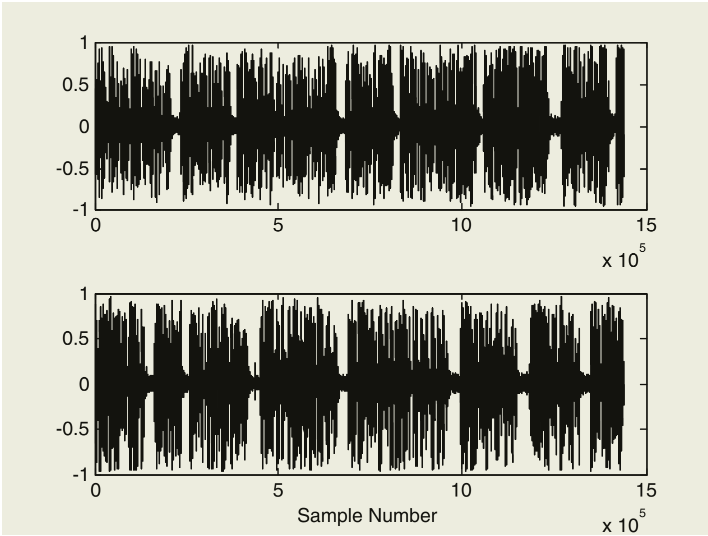

Fig. 1. Two source signals.

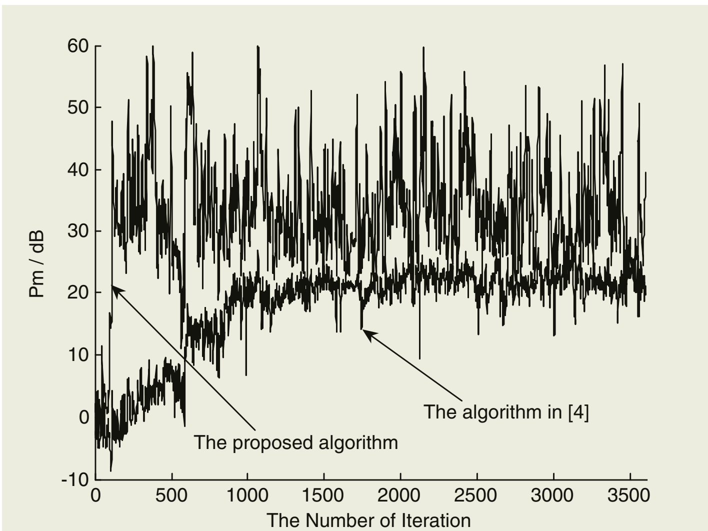

Fig. 2. The performance measure of separation.

5 Conclusion

We have proposed a new BSS approach by combining the FIB technique and complex-valued ICA for far-field broadband acoustic signals. The application of FIB makes our approach avoid the permutation problem, and the updating of separation matrix can be realized in the frequency serially which brings much faster convergence rate than other frequency-domain methods. We also proposed a new performance measure for the ability of separation. At last, the simulation result is given to verify the efficiency of the proposed approach.

References

1. Comon, P.: Independent Component Analysis, a New Concept Signal Processing **36** (1994) 287–314
2. Jutten, C., Herault, J.: Blind Separation of Sources Part I: An Adaptive Algorithm Based on Neuromimetic Architecture. Signal Processing **24** (1991) 1–10
3. Amari, S.: Natural Gradient Works Efficiently in Learning. Neural Computation **10** (1998) 251–276
4. Sawada, H., Mukai, R., Araki, S., Makino, S.: A Robust and Precise Method for Solving the Permutation Problem of Frequency-domain Blind Source Separation. IEEE Trans. on Speech and Audio Processing, **12** (2004) 530–538
5. Cardoso, J.F., Laheld, B.: Equivariant Adaptive Source Separation. IEEE Trans. Signal Processing, **44** (1996) 3017–3029
6. Zhu, X.L., Zhang, X.D.: Adaptive RLS Algorithm for Blind Source Separation Using a Natural Gradient. IEEE Signal Processing Letters, **9** (2002) 432–435
7. Douglas, S.C., Sun, L., Ser, X.: Convolutive Blind Separation of Speech Mixtures Using the Natural Gradient. Speech Commun, **39** (2003) 65–78
8. Kurita, S., Saruwatari, H., Kajita, S., Takeda, K., Itakura, F.: Evalutation of Blind Signal Separation Method Using Directivity Pattern Under Reverbrant Conditions. In Proc. ICASSP (2000) 3140–3143
9. Ikram, M.Z., Morgan, D.R.: A Beamforming Approach to Permutation Alignment for Multichannel Frequency-domain Blind Speech Separation. In Proc. ICASSP (2002) 881–884
10. Ward, D.B., Kennedy, R.A., Williamson, R.C.: Fir Filter Design for Frequencyinvariant Beamformers. IEEE Signal Processing Lett. **3** (1996) 69–71
11. Ward, D.B., Kennedy, R.A., Williamson, R.C.: Theory and Design of Broadband Sensor Arrays with Frequency Invariant Far-field Beam Patterns. J. Acoust. Soc. Amer. **97** (1995) 1023–1034