

Multiple Source Localization using Independent Component Analysis

Hiroshi Sawada*, Ryo Mukai, Shoko Araki, and Shoji Makino
NTT Communication Science Laboratories, Kyoto 619-0237, Japan

This paper presents a method for estimating location information about multiple sources. The proposed method uses independent component analysis (ICA) as a main statistical tool. The nearfield model as well as the farfield model can be assumed in this method. As an application of the method, we show experimental results for the direction-of-arrival (DOA) estimation of three sources that were positioned 3-dimensionally.

Introduction

Source localization is a basic and important technique in array signal processing not only for wireless communication systems, but also for audio/speech processing systems [1, 2]. Many source localization methods have been proposed. A well-known method is the MUSIC (MUltiple Signal Classification) algorithm [3]. It identifies the noise subspace with second order statistics and searches for location parameters that orthogonalize the steering vector and the noise subspace. The proposed ICA-based method, on the other hand, employs higher order statistics and directly identifies basis vectors, each of which contains the location information about a source. Consequently, the resolution of source localization for a multiple source case is superior with the ICA-based method [4], and it is easy to adapt it to the nearfield model as well as the farfield model.

The ICA-based source localization technique is a by-product of research on frequency-domain blind source separation (BSS), where the permutation ambiguity of the ICA solution in each frequency bin should be aligned in order to reconstruct time-domain signals properly [5]. One of the approaches for aligning this permutation ambiguity involves clustering estimated source locations [6–9]. Earlier works [6, 7] calculated directivity patterns from the separation matrix \mathbf{W} obtained by ICA, and then searched for null directions, which correspond to the directions of sources [10]. However, it is simpler and more effective to estimate the directions directly from the basis vectors \mathbf{a}_i , which are given by the pseudoinverse of \mathbf{W} . The method presented in this paper is based on this idea.

Independent Component Analysis

We consider a situation where N sources are convolutively mixed and observed at M sensors. Convulsive mixtures in the time domain can be approximated as multiple instantaneous mixtures in the frequency domain:

$$\mathbf{x}(f, \tau) = \sum_{i=1}^N \mathbf{h}_i(f) s_i(f, \tau), \quad (1)$$

where $\mathbf{x}(f, \tau) = [x_1(f, \tau), \dots, x_M(f, \tau)]^T$ is a sensor vector at frequency f and frame index τ , and $\mathbf{h}_i(f) = [h_{1i}(f), \dots, h_{Mi}(f)]^T$ is the vector of the frequency responses from source $s_i(f, \tau)$ to all M sensors. We obtain only $\mathbf{x}(f, \tau)$ by applying the short-time Fourier transform (STFT) to time-domain sensor observations.

The first step of the proposed method is to solve complex-valued ICA

$$\mathbf{y}(f, \tau) = \mathbf{W}(f) \mathbf{x}(f, \tau), \quad (2)$$

where $\mathbf{y}(f, \tau) = [y_1(f, \tau), \dots, y_N(f, \tau)]^T$ is a vector of separated signals and $\mathbf{W}(f)$ is an $N \times M$ separation matrix. There are various algorithms available for ICA as summarized in

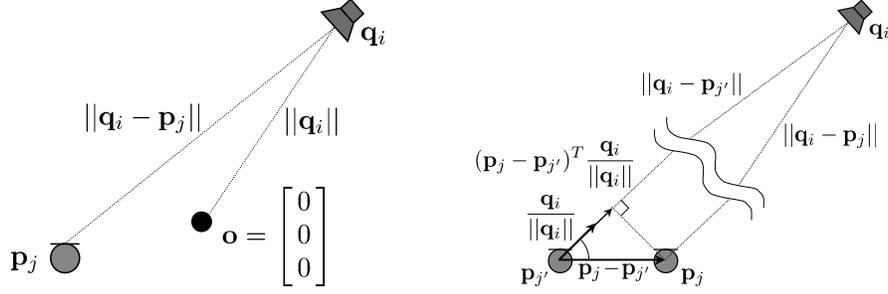


Figure 1: Nearfield (left) and farfield (right) model

e.g. [11]. Then, we calculate the Moore-Penrose pseudoinverse \mathbf{W}^+ (which is equivalent to the inverse \mathbf{W}^{-1} if $N = M$) of \mathbf{W}

$$[\mathbf{a}_1, \dots, \mathbf{a}_N] = \mathbf{W}^+, \quad \mathbf{a}_i = [a_{1i}, \dots, a_{Mi}]^T. \quad (3)$$

It is not difficult to make \mathbf{W} invertible by using an appropriate ICA procedure, such as whitening followed by unitary transformation (e.g. FastICA [11]). By multiplying both sides of (2) by \mathbf{W}^+ , the sensor vector $\mathbf{x}(f, \tau)$ is represented by a linear combination of basis vectors $\mathbf{a}_1, \dots, \mathbf{a}_N$:

$$\mathbf{x}(f, \tau) = \sum_{i=1}^N \mathbf{a}_i(f) y_i(f, \tau). \quad (4)$$

By comparing (1) and (4), we observe the following fact. If the ICA algorithm works well and the outputs y_1, \dots, y_N are the estimation of the sources s_1, \dots, s_N , then the basis vectors $\mathbf{a}_1, \dots, \mathbf{a}_N$ are also estimations of the mixing vectors $\mathbf{h}_1, \dots, \mathbf{h}_N$ up to permutation and scaling ambiguity: $\mathbf{a}_i = \alpha_i \mathbf{h}_{\Pi(i)}$. Here, scalar α_i represents scaling ambiguity and $\Pi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$ represents permutation ambiguity.

Source Localization with Nearfield Model

Let \mathbf{p}_j and \mathbf{q}_i be 3-dimensional vectors representing the locations of sensor j and source i , respectively. To estimate source location \mathbf{q}_i from basis vectors \mathbf{a}_i and sensor positions \mathbf{p}_j , we approximate the frequency response $h_{ji}(f)$ with a nearfield model (Fig. 1):

$$h_{ji}(f) = \frac{1}{\|\mathbf{q}_i - \mathbf{p}_j\|} e^{j2\pi f c^{-1} (\|\mathbf{q}_i - \mathbf{p}_j\| - \|\mathbf{q}_i\|)}, \quad (5)$$

where c is the propagation velocity of the signals. We assume that the phase depends on the difference between the distances $\|\mathbf{q}_i - \mathbf{p}_j\| - \|\mathbf{q}_i\|$ from the source to the sensor and to the origin $\mathbf{o} = [0, 0, 0]^T$. This makes the phase zero at the origin. In accordance with the model (5), the ratio between two elements $a_{ji}, a_{j'i}$ of the same basis vector \mathbf{a}_i provides the key equation for source localization:

$$\frac{a_{ji}}{a_{j'i}} = \frac{\alpha_i h_{ji}}{\alpha_i h_{j'i}} = \frac{\|\mathbf{q}_i - \mathbf{p}_{j'}\|}{\|\mathbf{q}_i - \mathbf{p}_j\|} e^{j2\pi f c^{-1} (\|\mathbf{q}_i - \mathbf{p}_j\| - \|\mathbf{q}_i - \mathbf{p}_{j'}\|)}, \quad (6)$$

where the scaling ambiguity α_i is canceled out by calculating the ratio. The permutation ambiguity still remains. However, if we estimate the location \mathbf{q}_i for all $i = 1, \dots, N$, the set of all estimated locations does not depend on the permutation.

The set of vectors \mathbf{q}_i in the argument of (6)

$$\|\mathbf{q}_i - \mathbf{p}_j\| - \|\mathbf{q}_i - \mathbf{p}_{j'}\| = \frac{\arg(a_{ji}/a_{j'i})}{2\pi f c^{-1}}. \quad (7)$$

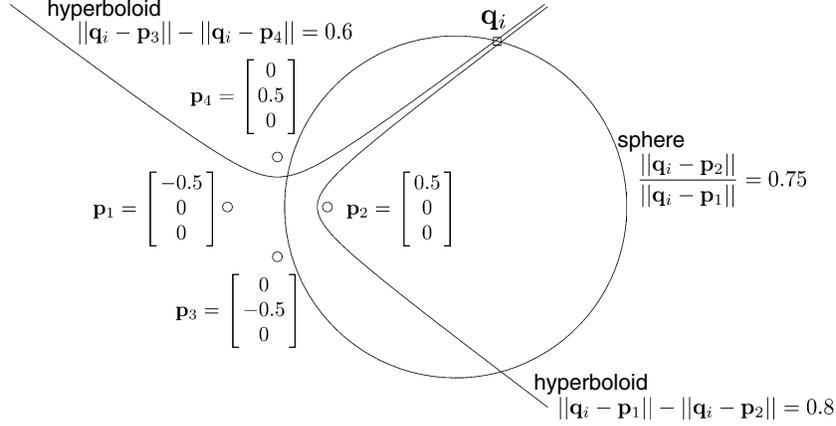


Figure 2: Source localization by the intersection of two hyperboloids and a sphere

defines a surface where the difference between the distances from \mathbf{p}_j and $\mathbf{p}_{j'}$ is constant. The surface is one sheet of a two-sheeted hyperboloid. Alternatively, the modulus of (6)

$$\frac{\|\mathbf{q}_i - \mathbf{p}_{j'}\|}{\|\mathbf{q}_i - \mathbf{p}_j\|} = \left| \frac{a_{ji}}{a_{j'i}} \right| \quad (8)$$

defines a sphere where the ratio of the distances from \mathbf{p}_j and $\mathbf{p}_{j'}$ is constant. With these two equations (7) and (8), we can estimate the possible location \mathbf{q}_i of source s_i . Such a hyperboloid and sphere are defined for a pair of sensors j and j' . If we select another pair, a different hyperboloid and sphere are obtained. In this way, the location \mathbf{q}_i is estimated as the intersection of several hyperboloids and spheres. An example is shown in Fig. 2.

DOA Estimation with Farfield Model

Although it is useful to estimate a 3-dimensional location, calculating the intersections of hyperboloids and spheres is computationally demanding. There are many cases where it is sufficient to estimate just the DOA of source s_i . If we assume that the source location \mathbf{q}_i is far from sensors \mathbf{p}_j and $\mathbf{p}_{j'}$, (7) can be approximated as a farfield model (Fig. 1)

$$(\mathbf{p}_j - \mathbf{p}_{j'})^T \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|} = \frac{\arg(a_{ji}/a_{j'i})}{2\pi f c^{-1}}. \quad (9)$$

The set of vectors \mathbf{q}_i satisfying (9) represents a cone [8], which is the asymptotic surface of the corresponding hyperboloid (7). The intersections of several cones specify the DOA of a source. Let us assume that we select u cones whose corresponding sensor pairs are $(j_1, j'_1), \dots, (j_u, j'_u)$. Then, the set of equations (9) for the u pairs is represented as

$$\mathbf{D} \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|} = \frac{\mathbf{r}_i}{2\pi f c^{-1}}, \quad (10)$$

where $\mathbf{D} = [\mathbf{p}_{j_1} - \mathbf{p}_{j'_1}, \dots, \mathbf{p}_{j_u} - \mathbf{p}_{j'_u}]^T$, $\mathbf{r}_i = [\arg(a_{j_1 i}/a_{j'_1 i}), \dots, \arg(a_{j_u i}/a_{j'_u i})]^T$. In practical situations, there is no exact solution for (10) because the u conditions do not coincide exactly. We typically solve it in the least-square sense by using the Moore-Penrose pseudoinverse [9]:

$$\frac{\mathbf{q}_i}{\|\mathbf{q}_i\|} = \frac{\mathbf{D}^+ \mathbf{r}_i}{2\pi f c^{-1}}. \quad (11)$$

If $\text{rank}(\mathbf{D}) \geq 3$, the set of vectors \mathbf{q}_i that satisfy (11) represents a line in 3-dimensional

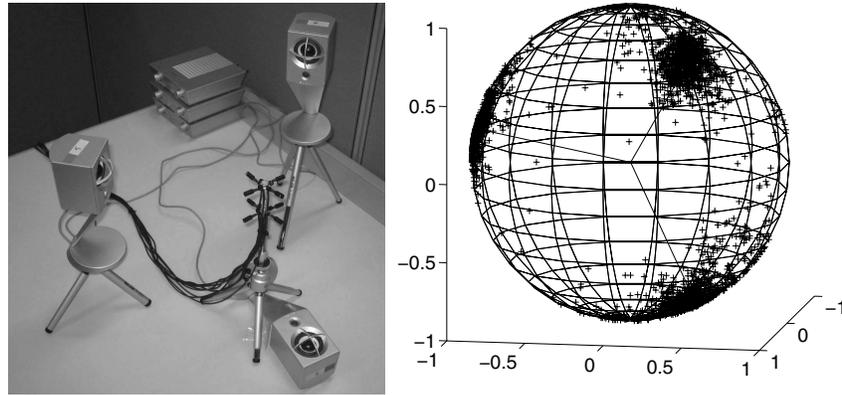


Figure 3: 3-dimensional arrangement of eight microphones and three loudspeakers (left) and DOA estimation results for this case (right)

space, and thus the DOA of a source i is estimated as $\frac{\mathbf{q}_i(f)}{\|\mathbf{q}_i(f)\|}$.

The left photograph in Fig. 3 shows a case where eight microphones and three loudspeakers are arranged 3-dimensionally, and the right plot shows the DOA estimations for this case. Each point shows a location vector normalized to unit norm $\mathbf{q}_i(f) \leftarrow \frac{\mathbf{q}_i(f)}{\|\mathbf{q}_i(f)\|}$. The estimations are obtained for all frequencies f and all output indexes i . As shown in the plot, they form clusters, each of which corresponds to the location of each source.

References

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, July 1996.
- [2] M. Brandstein and D. Ward, Eds., *Microphone Arrays*, Springer, 2001.
- [3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. 34, pp. 276–280, Mar. 1986.
- [4] H. Sawada, R. Mukai, and S. Makino, "Direction of arrival estimation for multiple source signals using independent component analysis," in *Proc. International Symposium on Signal Processing and its Applications*, July 2003, pp. 411–414.
- [5] 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, pp. 530–538, Sept. 2004.
- [6] H. Saruwatari, S. Kurita, K. Takeda, F. Itakura, T. Nishikawa, and K. Shikano, "Blind source separation combining independent component analysis and beamforming," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 11, pp. 1135–1146, 2003.
- [7] M. Z. Ikram and D. R. Morgan, "A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation," in *Proc. ICASSP 2002*, May 2002, pp. 881–884.
- [8] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Frequency domain blind source separation using small and large spacing sensor pairs," in *Proc. ISCAS 2004*, May 2004, vol. V, pp. 1–4.
- [9] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Frequency domain blind source separation for many speech signals," in *Proc. ICA 2004 (LNCS 3195)*, Sept. 2004, pp. 461–469.
- [10] S. Araki, S. Makino, Y. Hinamoto, R. Mukai, T. Nishikawa, and H. Saruwatari, "Equivalence between frequency domain blind source separation and frequency domain adaptive beamforming for convolutive mixtures," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 11, pp. 1157–1166, 2003.
- [11] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.