Forum Acusticum 2005

Direction of Arrival Estimates using Matching Pursuit

Yasuhiro Oikawa and Yoshio Yamasaki
Global Information and Telecommunication Institute, Waseda University, 1011 Okuboyama, Nishi-Tomida, Honjo-shi, Saitama 367-0035, Japan,
e-mail: **{yoikawa,y-yamasaki}@waseda.jp**

Many topics on microphone array processing, e.g. signal enhancement, noise reduction and direction of arrival (DOA) estimation, has recently been researched. Many methods has been suggested for DOA estimation and the number of estimated DOAs is generally limited by the number of microphones and we has spatial aliasing problem based on relation between the microphones distance and wave length. We has proposed new DOA estimation method using matching pursuit algorithm, which can estimate more DOAs than the microphones number. The DOA estimation we did consists of the following steps. First, we calculate the normalized power of the array output, $P(\theta)$, for each frequency bin using the Delay-and-Sum method. We then average $P(\theta)$ over all frequency bins. Finally, we perform peak-picking using a matching pursuit algorithm to estimate the DOA over all frequency bands. The matching pursuit algorithm includes, after each iteration step, a reoptimization of all DOAs found thus far. Its main characteristics is that it is possible to find true DOAs when the number of sources exceeds that of microphones. In this paper we report DOA estimation in real-world using this method. We could estimate more DOAs than the number of microphones. Our experiments yielded a better DOAs estimation with the new method than with the conventional method.

1 Introduction

Many topics on microphone array processing, e.g., signal enhancement, noise reduction and direction of arrival (DOA) estimation, have recently been researched, and many methods have been suggested for DOA estimation. The number of estimated DOAs is generally limited by the number of microphones and we have spatial aliasing problem on the basis of the relationship between microphone distance and wave length [1].

We have proposed a new DOA estimation method using a matching pursuit algorithm, which can estimate more DOAs than microphones [2]. First, we obtain the array output of the Delay-and-Sum method and calculate the average array output in the frequency range considered. Next, we estimate DOAs using a matching pursuit algorithm. In this study, we performed DOA estimation in a real world using this method and estimated more DOAs than microphones.

2 DOA estimation

Our DOA estimation consists of the following steps. We first separately calculate the normalized power of the array output, $P(\theta)$, for each frequency bin using the Delay-and-Sum method [3]. We then average $P(\theta)$ over all frequency bins. Finally, we perform peak picking using a matching pursuit algorithm to estimate the DOA over all frequency bands. The matching pursuit algorithm includes, after each iteration step, a reoptimization of all DOAs found thus far. Its main characteristic is that it is possible to find true DOAs when the number of sources exceeds that of microphones. We will discuss these steps in more detail in the following subsections.

2.1 Calculation of power of array output

The power of the Delay-and-Sum array output is calculated as

$$P(\theta) = \mathbf{d}(\theta)^H \mathbf{R}\mathbf{d}(\theta), \tag{1}$$

where $\mathbf{d}(\theta)$ is the steering vector:

$$\mathbf{d}(\theta) = [1, \exp(-j\omega\tau), \cdots, \exp(-j\omega(M-1)\tau)]^T. \tag{2}$$

Here, ω is the angular frequency, M is the number of microphones, $\tau = \frac{d \sin\theta}{c}$, d is the distance between microphones, c is the velocity of sound, and $\mathbf{R}$ is the covariance matrix of the array outputs $\mathbf{x}(t)$ i.e.:

$$\mathbf{R} = E[\mathbf{x}(t)\mathbf{x}(t)^H]. \tag{3}$$

For K sounds (i.e., K different DOAs) and two microphones, the observed signals are

$$\mathbf{X}(\omega, t) = \begin{bmatrix} X_1(\omega,t) \\ X_2(\omega,t) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{K} H_{1k}(\omega) S_k(\omega,t) \\ \sum_{k=1}^{K} H_{2k}(\omega) S_k(\omega,t) \end{bmatrix}, \tag{4}$$

where t is the time, H_{1k} and H_{2k} are the respective transfer functions between the kth sound and each microphone, and $S_k(\omega, t)$ is the kth sound. The covariance matrix is

$$\mathbf{R}(\omega) = E[\mathbf{X}(\omega,t)\mathbf{X}(\omega,t)^H] = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \quad (5)$$

$$r_{11} = E\big[|\sum_{k=1}^{K} H_{1k}S_k|^2\big], \quad (6)$$

$$r_{12} = E\big[\sum_{k=1}^{K} H_{1k}H_{2k}^*|S_k|^2 + \sum_{l\neq k} H_{1k}H_{2l}^* S_k S_l^*\big], \quad (7)$$

$$r_{21} = r_{12}^*, \text{ and} \quad (8)$$

$$r_{22} = E\big[|\sum_{k=1}^{K} H_{2k}S_k|^2\big], \quad (9)$$

and the steering vector is

$$\mathbf{d}(\theta,\omega) = [1, \exp(-j\omega\tau)]^T. \quad (10)$$

The power of array output is then

$$\begin{aligned} P(\theta,\omega) &= \mathbf{d}(\theta,\omega)^H\mathbf{R}(\omega)\mathbf{d}(\theta,\omega) \\ &= r_{11} + r_{22} + 2\cdot\Re\big\{r_{12}\exp(-j\omega\tau)\big\}, \quad (11) \end{aligned}$$

where $\Re$ indicates the real component.

The first and second terms of Equation (11) do not depend on θ and we only need to consider the third term. The third term of Equation (11), $\hat{P}(\theta,\omega)$, is

$$\begin{aligned} \hat{P}(\theta,\omega) =& P(\theta,\omega) - E[|X_1|^2] - E[|X_2|^2] \\ =& 2\Re\Big[E\big[\sum_{k=1}^{K} H_{1k}H_{2k}^*|S_k|^2 \\ &+ \sum_{l\neq k} H_{1k}H_{2l}^* S_k S_l^*\big]\exp(-j\omega\tau)\Big]. \quad (12) \end{aligned}$$

Therefore, the average of $\hat{P}(\theta,\omega)$ over the frequency bins is

$$\begin{aligned} \hat{P}_{avrg}(\theta) =& \frac{1}{N}\sum_{i=1}^{N}\hat{P}(\theta,\omega_i) \\ =& \frac{1}{N}\sum_{i=1}^{N}\Big[\sum_{k=1}^{K} 2\cdot E\big[|S_k(\omega_i)|^2\big] \\ &\cdot\Re\big\{H_{1k}(\omega_i)H_{2k}(\omega_i)^*\exp(-j\omega_i\tau)\big\}\Big] \\ &+ \frac{1}{N}\sum_{i=1}^{N}\Big[\sum_{l\neq k} 2\cdot\Re\big\{H_{1k}(\omega_i)H_{2l}(\omega_i)^* \\ &\cdot E\big[S_k(\omega_i)S_l(\omega_i)^*\big]\exp(-j\omega_i\tau)\big\}\Big], \quad (13) \end{aligned}$$

where N is the number of frequency bins. Since $E[S_k(\omega)S_l(\omega)^*]$ is generally smaller for $k \neq l$ than for $k = l$, we have assumed that the second term in Equation (13) can be set to zero. We can then rewrite Equation (13) as

$$\hat{P}_{avrg}(\theta) \approx \sum_{k=1}^{K}\hat{P}_{avrg}(\theta|\theta_k), \quad (14)$$

where $\hat{P}_{avrg}(\theta|\theta_k)$ is the frequency average of the θ-dependent component of the array output power from the kth sound, i.e., the θ-dependent component of the Delay-and-Sum array output is approximately the sum of that for each source.

As we are only interested in finding the DOAs at this point, we let $H_{1k}(\omega) = \exp(-j\omega\tau_{1k})$ and $H_{2k}(\omega) = \exp(-j\omega\tau_{2k})$. Thus, the frequency average from the kth sound, $\hat{P}_{avrg}(\theta|\theta_k)$, becomes

$$\begin{aligned} \hat{P}_{avrg}(\theta|\theta_k) =& \frac{2E[|S_k(\omega_i)|^2]}{N} \\ &\cdot\sum_{i=1}^{N}\Re\big\{\exp(-j\omega_i(\tau_{1k}-\tau_{2k})) \\ &\cdot\exp(-j\omega_i\tau)\big\}, \quad (15) \end{aligned}$$

$$\tau_{1k} - \tau_{2k} = \frac{d\sin\theta_k}{c}, \text{ and} \quad (16)$$

$$\tau = \frac{d\sin\theta}{c}, \quad (17)$$

where θ_k is the true direction of the kth sound position and θ is the steering direction.

2.2 Matching Pursuit to Estimate DOA

A matching pursuit algorithm was introduced to decompose any signal into a linear expansion of waveforms [4]. We used a modified matching pursuit algorithm that includes a reoptimization step [5] to decompose the signal into a set of direct and reflected sounds. We define the vector of the angles of i DOAs that is estimated during i iterations as

$$\Theta_i = [\hat{\theta}_1, \cdots, \hat{\theta}_i]^T, \quad (18)$$

where Θ_0 is a vector without any elements. The matching pursuit algorithm for DOA estimation consists of the following steps.

Step 1) Define a dictionary as

$$\mathcal{D} = \{\hat{P}_{avrgn}(\theta|\theta_k)\}_{-\pi/2<\theta_k<\pi/2}, \quad (19)$$

i.e., an element of family $\mathcal{D}$ is defined as Equation (15) normalized by its norm:

$$\hat{P}_{avrgn}(\theta|\theta_k) = \frac{\hat{P}_{avrg}(\theta|\theta_k)}{\sqrt{\frac{1}{\pi}\int_{-\pi/2}^{\pi/2}|\hat{P}_{avrg}(\theta|\theta_k)|^2 d\theta}}. \quad (20)$$

Step 2) Initialization:

$$e_0(\theta) = \hat{P}_{observed}(\theta) \text{ and} \tag{21}$$
$$i = 1. \tag{22}$$

Step 3) Calculate the residual for all θ_k:

$$e_i(\theta|\theta_k) = e_{i-1}(\theta) - a_{i-1}(\theta_k)\hat{P}_{avrgn}(\theta|\theta_k), \tag{23}$$

where $a_{i-1}(\theta_k)$ denotes the inner product of $e_{i-1}(\theta)$ and $\hat{P}_{avrgn}(\theta|\theta_k)$.

Step 4) Select θ_k (estimate DOA $\hat{\theta}_i$):

$$\hat{\theta}_i = \underset{\theta_k}{\text{argmin}} \sum |e_i(\theta|\theta_k)|^2. \tag{24}$$

Step 5) Reoptimize Θ_i (all DOAs) and calculate the residual $e_i(\theta)$:

$$e_i(\theta) = e_0(\theta) - \sum_{l=1}^{i} \hat{a}(\hat{\theta}_l)\hat{P}_{avrgn}(\theta|\hat{\theta}_l), \tag{25}$$

where $\hat{a}(\hat{\theta}_l)$ is computed using Equation (32).

Step 6) If

$$10 \log \frac{\int e_0^2(\theta)d\theta}{\int e_i^2(\theta)d\theta} < \delta, \tag{26}$$

where δ is the stopping criterion, and

$$i = i + 1, \tag{27}$$

return to Step 3), or else end the procedure.

2.3 Reoptimization of DOAs

A high-quality, consistent analysis-synthesis method with reoptimization of amplitude and frequency parameters in sinusoidal coding has been described by Vos et al. [5]. They presented techniques for the optimization of sinusoidal parameters based on the squared difference between the input signal and reconstruction. Here, we use a similar method in order to reoptimize the DOAs with a gradient algorithm. We describe optimization techniques of DOAs based on the squared difference between the array output and reconstruction of components for estimated DOAs.

We define the vector of the angles of L DOAs as

$$\Theta = [\theta_1, \cdots, \theta_L]^T. \tag{28}$$

The basis vectors and the observed vector are defined as

$$\hat{\mathbf{p}}_{avrgn}(\theta_k) = [\hat{P}_{avrgn}(-\frac{\pi}{2}|\theta_k), \cdots, \hat{P}_{avrgn}(\frac{\pi}{2}|\theta_k)]^T \text{ and} \tag{29}$$

$$\mathbf{e}_0 = [e_0(-\frac{\pi}{2}), \cdots, e_0(\frac{\pi}{2})]^T, \tag{30}$$

where we discretized the normalized frequency average of the power of array output of the kth sound as a function of the continuous steering direction variable θ.

For a given set of DOAs, the analysis matrix containing the basis vectors is constructed according to

$$\hat{\mathbf{P}}_{avrgn\Theta} = [\hat{\mathbf{p}}_{avrgn}(\theta_1), \cdots, \hat{\mathbf{p}}_{avrgn}(\theta_L)]. \tag{31}$$

The projection of $\mathbf{e}_0$ onto a space that is defined by bases $\hat{\mathbf{p}}_{avrgn}(\theta_1), \cdots,$ and $\hat{\mathbf{p}}_{avrgn}(\theta_L)$ is

$$\hat{\mathbf{a}} = (\hat{\mathbf{P}}^T_{avrgn\Theta} \cdot \hat{\mathbf{P}}_{avrgn\Theta})^{-1} \cdot \hat{\mathbf{P}}^T_{avrgn\Theta} \cdot \mathbf{e}_0, \tag{32}$$

which is from the least-squares residual. Optimum DOAs are those for which the energy of the projection of $\mathbf{e}_0$ onto the column space of $\hat{\mathbf{P}}_{avrgn\Theta}$ is maximized:

$$\begin{aligned} &\underset{\Theta}{\text{argmax}}\{\hat{\mathbf{P}}_{avrgn\Theta} \cdot \hat{\mathbf{a}}\}^T \cdot \{\hat{\mathbf{P}}_{avrgn\Theta} \cdot \hat{\mathbf{a}}\} \\ &= \underset{\Theta}{\text{argmax}}\, \mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0, \end{aligned} \tag{33}$$

where we define the projection matrix as

$$\mathbf{P}_\Theta = \hat{\mathbf{P}}_{avrgn\Theta} \cdot (\hat{\mathbf{P}}^T_{avrgn\Theta} \cdot \hat{\mathbf{P}}_{avrgn\Theta})^{-1} \cdot \hat{\mathbf{P}}^T_{avrgn\Theta}. \tag{34}$$

We used a gradient search based on Newton's method to find the local maximum of $\mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0$ in the neighborhood of a set of initial DOAs. At iteration m, Newton's method defines the updated DOAs as

$$\Theta^{(m)} = \Theta^{(m-1)} + \mathbf{H}^{-1}_{\Theta^{(m-1)}} \mathbf{g}_{\Theta^{(m-1)}}, \tag{35}$$

where $\mathbf{g}_\Theta$ and $\mathbf{H}_\Theta$ represent the gradient vector and the Hessian matrix of the energy of the projection:

$$\mathbf{g}_\Theta = \frac{\partial}{\partial \Theta} \mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0 \text{ and} \tag{36}$$
$$\mathbf{H}_\Theta = \frac{\partial^2}{\partial \Theta \partial \Theta^T} \mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0. \tag{37}$$

We now present the equations for $\mathbf{g}_\Theta$ and $\mathbf{H}_\Theta$. They can be obtained from the formulas of the first and second derivatives of a projection matrix $\mathbf{P}_\Theta$ with respect to the elements in Θ [6]. Since $\hat{\mathbf{P}}_{avrgn\Theta}$ is of full rank,

$$\begin{aligned} \mathbf{P}_\Theta &= \hat{\mathbf{P}}_{avrgn\Theta} \cdot (\hat{\mathbf{P}}^*_{avrgn\Theta} \cdot \hat{\mathbf{P}}_{avrgn\Theta})^{-1} \cdot \hat{\mathbf{P}}^*_{avrgn\Theta} \\ &= \hat{\mathbf{P}}_{avrgn\Theta} \cdot \hat{\mathbf{P}}^+_{avrgn\Theta}, \end{aligned} \tag{38}$$

where $(\cdot)^*$ means the Hermitian transpose and $(\cdot)^+$ means the pseudoinverse. The element of $\mathbf{g}_\Theta$ is

$$\begin{aligned} \frac{\partial}{\partial \theta_\eta} \mathbf{e}_0^T \mathbf{P}_\Theta \mathbf{e}_0 &= \mathbf{e}_0^T \frac{\partial \mathbf{P}_\Theta}{\partial \theta_\eta} \mathbf{e}_0 \\ &= \mathbf{e}_0^T \mathbf{P}_{\Theta\eta} \mathbf{e}_0 \\ &= \mathbf{e}_0^T (\hat{\mathbf{P}}_\eta \hat{\mathbf{P}}^+ + \hat{\mathbf{P}} \hat{\mathbf{P}}^+_\eta) \mathbf{e}_0^T, \end{aligned} \tag{39}$$

where we write $\hat{\mathbf{P}}$ instead of $\hat{\mathbf{P}}_{avrgn\Theta}$ for the ease of notation. After some algebraic manipulations, the following is obtained for the pseudoinverse:

$$\hat{\mathbf{P}}_\eta^+ = (\hat{\mathbf{P}}^*\hat{\mathbf{P}})^{-1}\hat{\mathbf{P}}_\eta^*\mathbf{P}_\Theta^\perp - \hat{\mathbf{P}}^+\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+, \quad (40)$$

where $\mathbf{P}_\Theta^\perp = \mathbf{I} - \mathbf{P}_\Theta$. Combining Eqs. (39) and (40) gives

$$\frac{\partial}{\partial\theta_\eta}\mathbf{e}_0^T\mathbf{P}_\Theta\mathbf{e}_0 = \mathbf{e}_0^T\{\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+ + (\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+)^*\}\mathbf{e}_0. \quad (41)$$

The second derivative is given by

$$\begin{aligned}&\frac{\partial^2}{\partial\theta_\eta\partial\theta_\xi}\mathbf{e}_0^T\mathbf{P}_\Theta\mathbf{e}_0\\ &= \mathbf{e}_0^T\{\mathbf{P}_{\Theta\xi}^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_{\eta\xi}\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}_\xi^+\\ &+ (\mathbf{P}_{\Theta\xi}^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_{\eta\xi}\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}_\xi^+)^*\}\mathbf{e}_0. \quad (42)\end{aligned}$$

Using Equation (40) and $\mathbf{P}_{\Theta\xi}^\perp = -\mathbf{P}_{\Theta\xi}$ gives

$$\begin{aligned}&\frac{\partial^2}{\partial\theta_\eta\partial\theta_\xi}\mathbf{e}_0^T\mathbf{P}_\Theta\mathbf{e}_0\\ &= \mathbf{e}_0^T\{-\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\xi\hat{\mathbf{P}}^+\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+ - \hat{\mathbf{P}}^{+*}\hat{\mathbf{P}}_\xi^*\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+\\ &+ \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_{\eta\xi}\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta(\hat{\mathbf{P}}^*\hat{\mathbf{P}})^{-1}\hat{\mathbf{P}}_\xi^*\mathbf{P}_\Theta^\perp\\ &- \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+\hat{\mathbf{P}}_\xi\hat{\mathbf{P}}^+\\ &+ (-\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\xi\hat{\mathbf{P}}^+\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+ - \hat{\mathbf{P}}^{+*}\hat{\mathbf{P}}_\xi^*\mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+\\ &+ \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_{\eta\xi}\hat{\mathbf{P}}^+ + \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta(\hat{\mathbf{P}}^*\hat{\mathbf{P}})^{-1}\hat{\mathbf{P}}_\xi^*\mathbf{P}_\Theta^\perp\\ &- \mathbf{P}_\Theta^\perp\hat{\mathbf{P}}_\eta\hat{\mathbf{P}}^+\hat{\mathbf{P}}_\xi\hat{\mathbf{P}}^+)^*\}\mathbf{e}_0. \quad (43)\end{aligned}$$

3 Experiments

We estimated DOAs under both artificial and real-world experimental conditions.

3.1 Artificial Experiments

We compared the DOAs estimation performance without and with reoptimization of DOA for artificial data. The experiment conditions are shown in Figure 1. The two speech sources and two walls were arranged as shown in Figure 1. The speech sounds were assumed to arrive at microphones from each direct and each first-reflected sound direction. This is equivalent to a mixture with one true source and one image source for each independent signal.

We used 64 mixtures with 8 English and 8 French speeches, which were spoken by one female and one male for each language. The average length of speech is 5 s. We estimated DOAs for each sentence and averaged 64 results. The distance between microphones was 5 cm. We used a data sampling frequency of 44.1 kHz, a frame length of 46 ms, and a frame update of 23 ms.

Figure 2 shows the array output. Figure 3 shows the estimated DOA without reoptimization. Figure 4 shows that with reoptimization.

Table 1 lists DOA estimation results. When the reoptimization process was included, it had much better performance of DOA estimation than that without reoptimization.

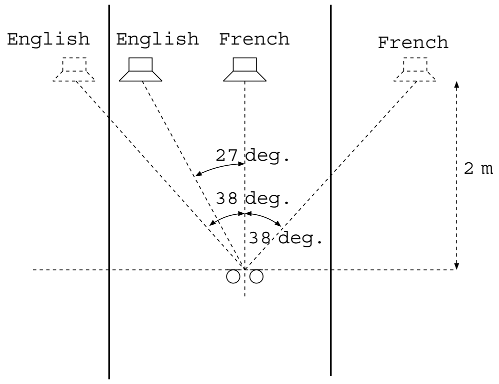


Figure 1: Artificial mixture conditions.

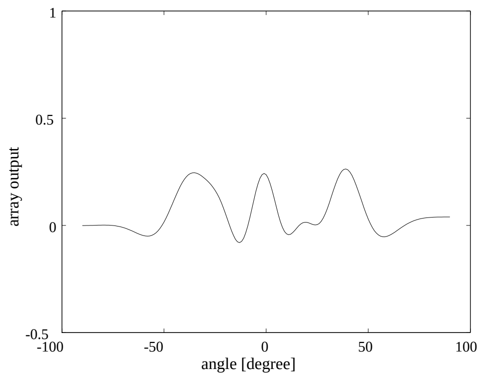


Figure 2: Observed power.

Table 1: DOA estimation.

Method	DOA [degree]
True	-38.0, -27.0, 0.0, 38.0
Without reoptimization	-34.0, -23.0, -1.0, 39.0
With reoptimization	-38.5, -27.7, -0.3, 38.2

3.2 Real-World Experiments

The real-world data were recorded in two rooms. A reverberation time was 0.36 s in room 1 and 0.94 s in room 2. Figure 5 shows the real-world experiment conditions.

Sources were located at -30 degrees and 45 degrees. The sources are the same as those described in Sect. 3.1

Figures 6 and 7 show the matching pursuit iterations in room 1 and room 2, respectively. The x-axis is the direction of arrival, and the y-axis is the power of the arriving sound. The top curve is the observed power e_0. The second curve is the residual e_1 after the first DOA is estimated. The third, fourth, and last curves are the residual e_2, e_3, and e_4. We found ten sounds in the real-world data that reached the stopping criterion of 30 dB. The residual e_{10} was almost flat.

4 Summary

In this study, we estimated more DOAs than microphones using a matching pursuit algorithm under reverberant conditions. Source separation has recently been studied by many researchers to separate sources effectively under real world conditions. Thus, one of the research topics is to separate sources for convoluted mixtures consisting of long impulse responses in acoustics. We expect that additional improvement can be achieved in actual cases by considering the spatial information from this method.

References

[1] J. Ohga, Y. Yamasaki and Y. Kaneda, *Acoustic system and digital processing* (Korona Publishing Co., Ltd., Tokyo, 1995).

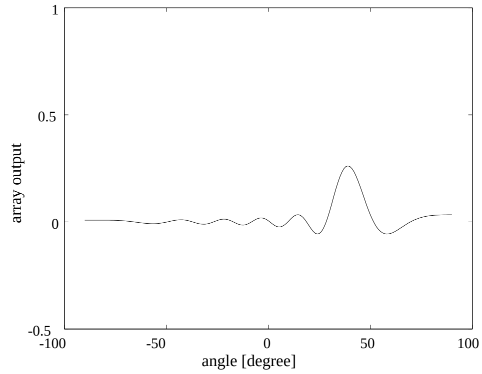
[2] Y. Oikawa and Y.Yamasaki, "Direction of arrival estimates by two microphones using matching pursuit," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 501–502 (2003).

[3] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Prentice-Hall Inc., New Jersey, 1993).

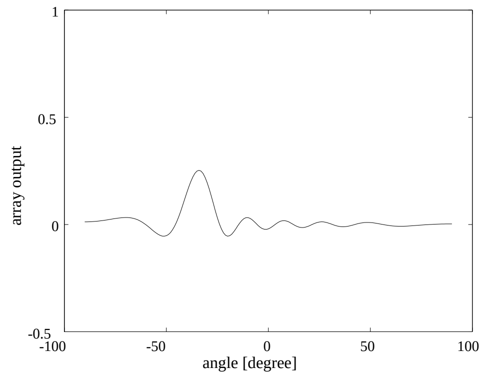
[4] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, **41**(12), 3397–3415 (1993).

[5] K. Vos, R. Vafin, R. Heusdens and W. B. Kleijn, "High-quality consistent analysis-synthesis in sinusoidal coding," *Proc. 1999 Audio Eng. Soc. 17th Conf. "High Quality Audio Coding"*, pp. 244–250 (1999).

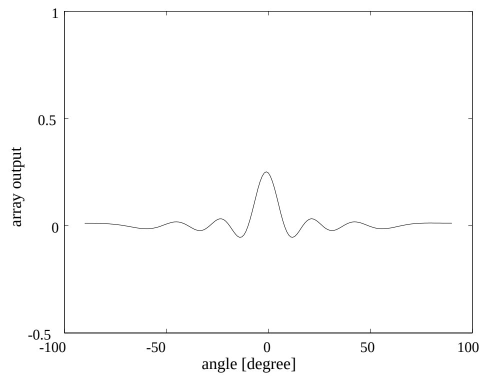
[6] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Process.*, **39**(5), 1110–1121 (1991).



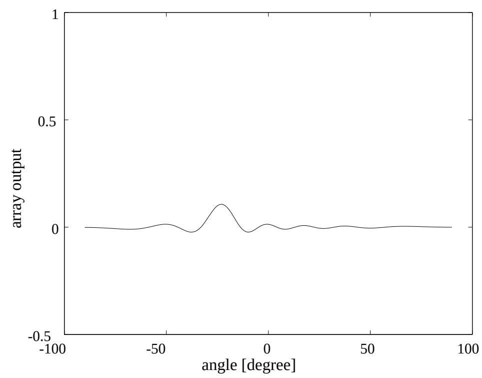
(a) First DOA.



(b) Second DOA.

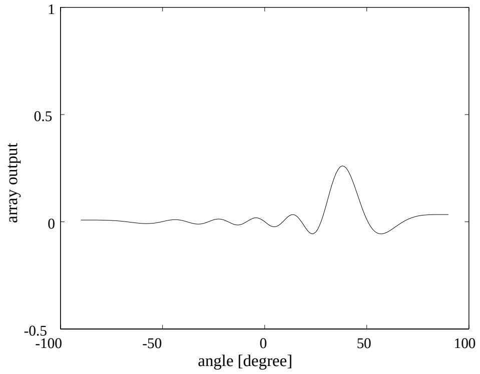


(c) Third DOA.

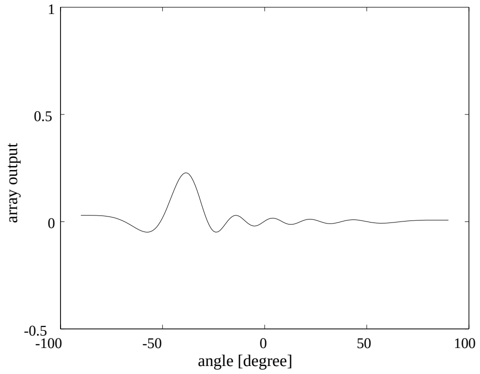


(d) Fourth DOA.

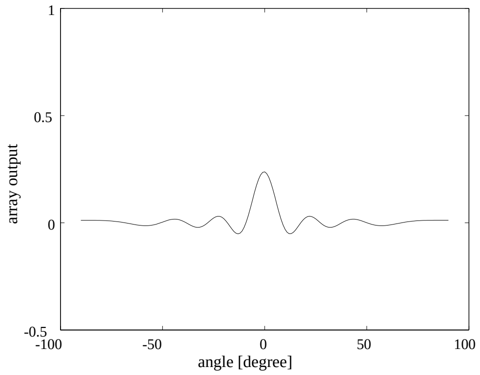
Figure 3: Estimated DOAs without reoptimization.



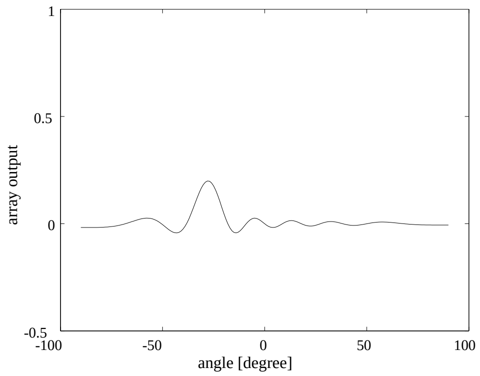
(a) First DOA.



(b) Second DOA.



(c) Third DOA.



(d) Fourth DOA.

Figure 4: Estimated DOAs with reoptimization.

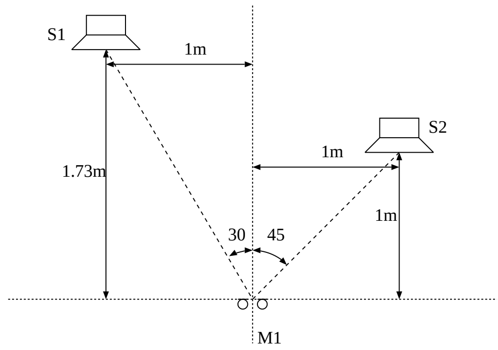


Figure 5: Real-world mixture conditions.

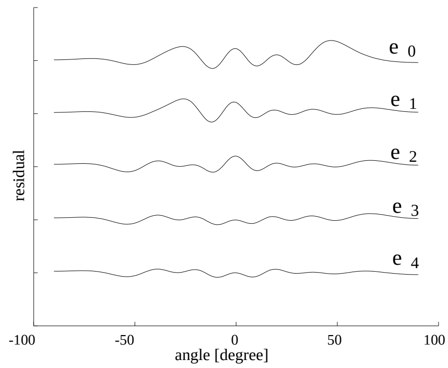


Figure 6: Matching pursuit iterations in room 1.

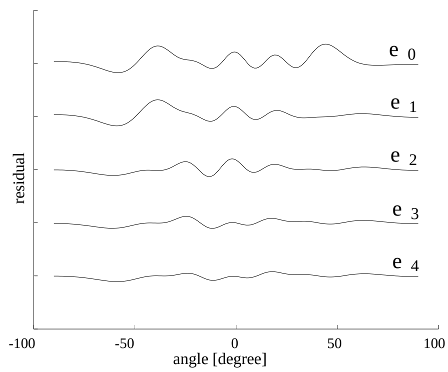


Figure 7: Matching pursuit iterations in room 2.