Simultaneous Localization and Communication Methods Using Short-Time and Narrow-Band Acoustic Signals

Masanari Nakamura*, Hiromichi Hashizume**, Masanori Sugimoto*

* Hokkaido University

** National Institute of Informatics

Contact : masanari@ist.hokudai.ac.jp

0. Short Resume

- MASANARI NAKAMURA received the B.E. degree, and M.S., D.S. degree in information science and technology from the Hokkaido University in 2014, 2016, 2018, respectively.

- He is currently an assistant professor with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Japan.

- His research interests include acoustic engineering, signal processing, indoor localization, tracking.

1. Introduction

- Mobile devices, such as smartphones, tablets, and smart glasses are widely used nowadays. Complementing these is the location data of mobile devices, which are employed in various services.
- In this paper, we describe simultaneous localization and communication methods using acoustic signals transmitted simultaneously from multiple speakers.
- In the case where short-time and narrow-band acoustic signals are used, the interference of signals could cause a systematic error of localization and communication depending on the modulation value.
- To reduce this systematic error, a noninterference region was constituted in the received signal, and localization and demodulation were processed using this region.
- Through simulation and real-environment experiments, it was confirmed that the proposed method can reduce the systematic error.
- To concisely discuss the influence of signal interference on localization and communication, we evaluated the azimuth estimation performance by the TDoA of two speakers as one-dimensional localization.

2. Related Work

- We describe the problem while performing azimuth estimation using two speakers and communication by DPSK simultaneously using short-time and narrow-band signals.
- To describe these processes of transmitting and receiving acoustic signals, we use the chirp signals as an example.

Transmission Signal

$$s_R^i(t) = \sin\left(2\pi(f_1 t + (1/2)\alpha_{12} t^2) + \phi_i^R\right) \qquad (0 \le t \le T)$$
$$s_L^i(t) = \sin\left(2\pi(f_3 t + (1/2)\alpha_{34} t^2) + \phi_i^L\right)$$

where ϕ_i^R and ϕ_i^L are DPSK modulation values and $a_{12} = (f_2 - f_1)T, a_{34} = (f_4 - f_3)/T$

Received Signal

$$r^i(t) = a_R s_R^i(t - t_R) + a_L s_L^i(t - t_L)$$

where t_R and t_L represent the reception times of the left and right signals of the speaker.

Analytic signal

$$e_R(t) = \exp\left(j2\pi(f_1 t + (1/2)\alpha_{12} t^2)\right) \qquad (0 \le t \le T)$$
$$e_L(t) = \exp\left(j2\pi(f_3 t + (1/2)\alpha_{34} t^2)\right)$$

where $a_{12} = (f_2 - f_1)T, a_{34} = (f_4 - f_3)/T$

2. Related Work

- The received signal is processed using the matched filter with analytic signals as follows:

$$c_R^i(t) = \int_t^{t+T} r^i(t) e_R(\tau - t) d\tau$$
$$= \int_t^{t+T} a_R s_R^i(\tau - t_R) e_R(\tau - t) d\tau + \int_t^{t+T} a_L s_L^i(\tau - t_L)\, e_R(\tau - t) d\tau \quad \cdots (1)$$

- The reception time of speaker R is estimated as follows:

$$\hat{t}_R = \arg\max_t \left| c_R^i(t) \right|$$

- The second term in (1) that represents interference decreases as the length and bandwidth of the signal increase.
- Therefore, in conventional methods, such as chirp signal-based, OFDM-based, and CDMA-based methods, and random signal methods, interference was avoided by setting the length and bandwidth of signals to sufficiently large values.
- In this paper, we propose a method that can reduce the above systematic errors with short-time and narrow-band signals.

3. Proposed Method

Transmission Signal

- The speakers R and L transmit the following signals simultaneously,

$$s_R^i(t) = \sin\left(2\pi f_1 t + \phi_i^R\right) + \sin\left(2\pi f_2 + \phi_i^R + \pi\right)$$
$$s_L^i(t) = \sin\left(2\pi f_3 t + \phi_i^L\right) + \sin\left(2\pi f_4 + \phi_i^L + \pi\right) \qquad (0 \le t \le T)$$

where $f_n = f_1 + \left((n-1)/(T/2)\right)$.

Analytic signal

$$e_R(t) = \exp(j2\pi f_1 t) + \exp(j2\pi f_2 + \pi)$$
$$e_L(t) = \exp(j2\pi f_3 t) + \exp(j2\pi f_4 + \pi) \qquad (0 \le t \le T)$$

where $f_n = f_1 + \left((n-1)/(T/2)\right)$.

Estimation of temporary reception time

- The signals transmitted from the speakers are received as follows:

$$r^i(t) = a_R s_R^i(t - t_R) + a_L s_L^i(t - t_L)$$

- Calculate the temporary reception time ${t'}_R$, ${t'}_L$ as follows.

$$t'_R = \arg\max_t \left|c_R^i(t)\right|, \;\; t'_L = \arg\max_t \left|c_L^i(t)\right|$$

where $c_R^i(t), c_L^i(t)$ are

$$c_R^i(t) = \int_t^{t+T} r^i(t) e_R(\tau - t) d\tau$$
$$c_L^i(t) = \int_t^{t+T} r^i(t) e_L(\tau - t) d\tau.$$

3. Proposed Method

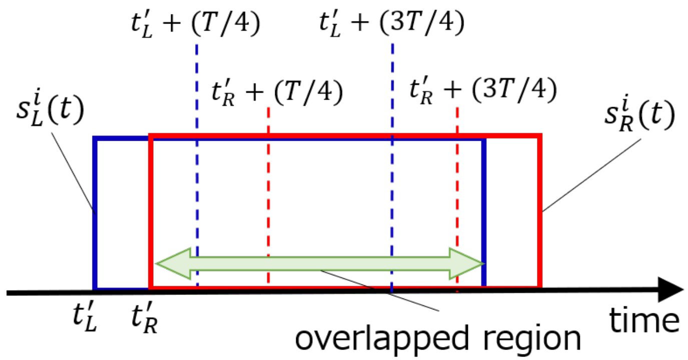


Figure 1. Illustration of the received signals.

TDoA estimation

- Figure 1 indicates an illustration of received signal. Here, suppose the absolute value of TDoA $|\Delta t| = |t_L - t_R|$ is less than $T/4$.
- From the definition of f_n, if the signal of length $T/2$ is cut out of the overlapped region of the received signal, sine waves that compose the cut-out signal are mutually orthogonal.
- The signal of length T/2 is obtained by cutting out a region from the overlapped region in Figure 1. In the following, t_{max} is the reference time of cutting out.
- To avoid the error due to the influence of $c_R^i(t)$, $c_L^i(t)$, t_{max} is set to t_R' or t_L' as follows.
  - If $|c_R^i(t_R')|$ is larger than or equal to $|c_L^i(t_L')|$, t_{max} is set to t_R'.
  - If $|c_R^i(t_R')|$ is less than $|c_L^i(t_L')|$, t_{max} is set to t_L'.
- In this section, we take $t_{max} = t_R'$ as an example.

3. Proposed Method

- The cut-out signal $r_c^i(t)$ is represented as follows:

$$r_c^i(t) = a_R \sin\left(2\pi f_1(t + t_{max} + T/4) + \phi_i^R\right) + a_R \sin\left(2\pi f_2(t + t_{max} + T/4) + \phi_i^R\right) \\ + a_L \sin\left(2\pi f_3(t + t_{max} + T/4 + \Delta t) + \phi_i^L\right) + a_L \sin\left(2\pi f_4(t + t_{max} + T/4 + \Delta t) + \phi_i^L\right)$$

- The phase values of each frequency can be calculated by quadrature detection.

$$\phi_1^i = 2\pi f_1(t_{max} + T/4) + \phi_i^R$$
$$\phi_2^i = 2\pi f_2(t_{max} + T/4) + \phi_i^R$$
$$\phi_3^i = 2\pi f_3(t_{max} + T/4 + \Delta t) + \phi_i^L$$
$$\phi_4^i = 2\pi f_4(t_{max} + T/4 + \Delta t) + \phi_i^L$$

- Therefore, TDoA Δt can be calculated as follows:

$$\Delta t = \frac{\left(\phi_4^i - \phi_3^i\right) - \left(\phi_2^i - \phi_1^i\right)}{2\pi}(T/2)$$

- The modulation values ϕ_i^R and ϕ_i^L are canceled out. Therefore, our proposed method can obtain TDoA without systematic error due to modulation values.

Demodulation

- In our proposed method, signals are demodulated by using phase difference between successive symbols.
- As the reference time of demodulation, we utilized the first symbol's t_{max}.

4. Simulation Experiments

Settings

- Figure 2 shows the location of the speakers and a microphone.
- The amplitudes of the received signals were set to the same value for all locations
- To avoid the effect of aliasing, the sampling rate was set to 960 kHz.
- The values of symbol sequence were set as shown in Table I.
  - To cover all combinations of ϕ_i^R and ϕ_i^L.
  - The number of symbol values was set to 4.

Transmission signal of conventional method

Chirp method:

$$s_R^i(t) = \sin\left(2\pi\left(f_1 t + (1/4T)(f_2 - f_1)\, t^2\right) + \phi_i^R\right)$$
$$s_L^i(t) = \sin\left(2\pi\left(f_3 t + (1/4T)(f_4 - f_3) t^2\right) + \phi_i^L\right)$$

$f_1 = 2.5$ kHz, $f_2 = 4.5$ kHz, $f_3 = 4.5$ kHz, $f_4 = 6.5$ kHz, $T = 2$ ms

OFDM (and proposed) method:

$$s_R^i(t) = \sin\left(2\pi f_1 t + \phi_i^R\right) + \sin\left(2\pi f_2 + \phi_i^R + \pi\right)$$
$$s_L^i(t) = \sin\left(2\pi f_3 t + \phi_i^L\right) + \sin\left(2\pi f_4 + \phi_i^L + \pi\right)$$

$f_1 = 3$ kHz, $f_2 = 4$ kHz, $f_3 = 5$ kHz, $f_4 = 6$ kHz, $T = 2$ ms

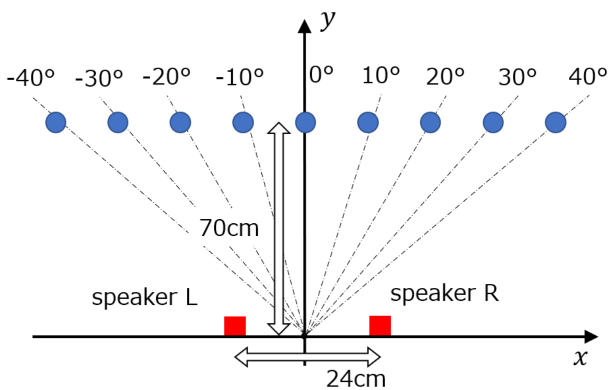


Figure 2. Speaker and microphone arrangement.

TABLE I. ENCODED VALUES.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Φ_i^R	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$
Φ_i^L	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{3\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$

4. Results

Chirp method

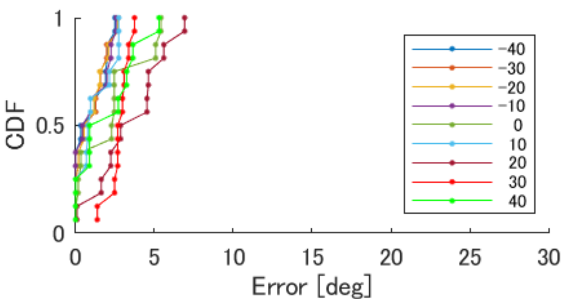
- Systematic errors occur due to interference at all locations.

OFDM method

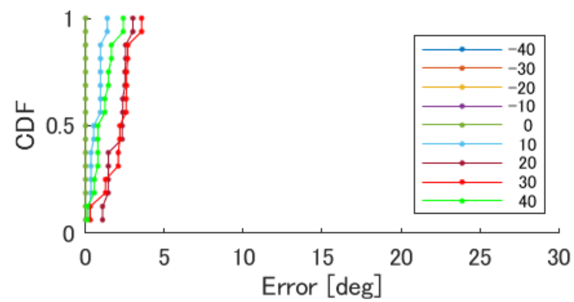
- Systematic errors occurs due to interference without the location where TDoA is zero ($\theta = 0$).

Proposed method

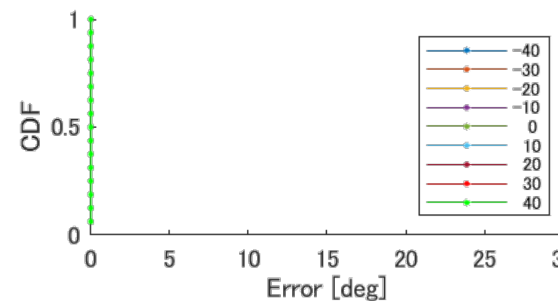
- No systematic error due to interference at all locations.



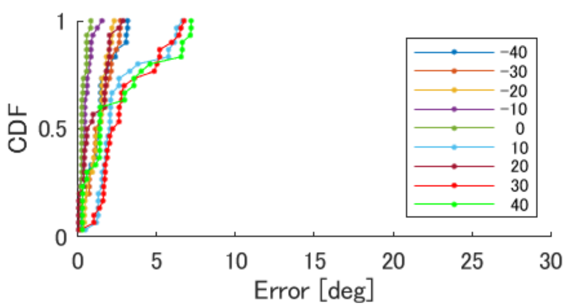
Azimuth Estimation: chirp method



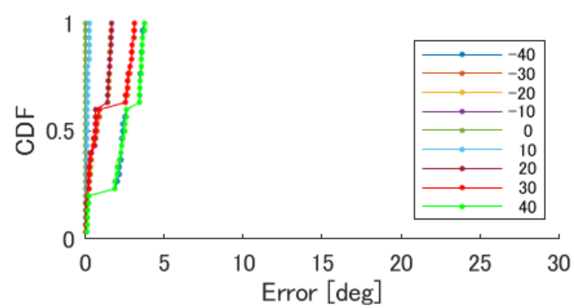
Azimuth Estimation: OFDM method



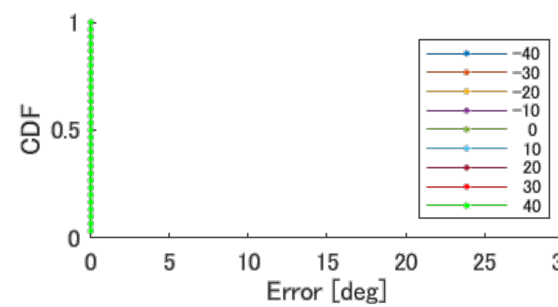
Azimuth Estimation: proposed method



Demodulation: chirp method



Demodulation: OFDM method



Demodulation: proposed method

Figure 3. Simulation results (Legends indicate the azimuth θ of the microphone's location).

5. Real Environmental Experiments

- The signals and locations of the speakers and a microphone are the same as in the simulation experiments.
- The symbol sequence having the encoded values in Table I were measured 15 times at each location.
  - The number of localizations was 15 × 16 = 240.
  - The number of demodulation values was 15 × 15 = 225.
- The transmission interval of symbols was set to 100 ms.
- Sampling rate was set to 48 kHz.

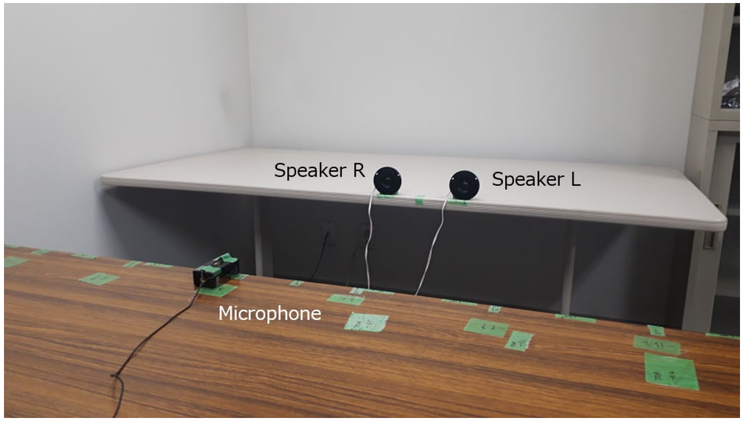


Figure 4. Experimental environment.

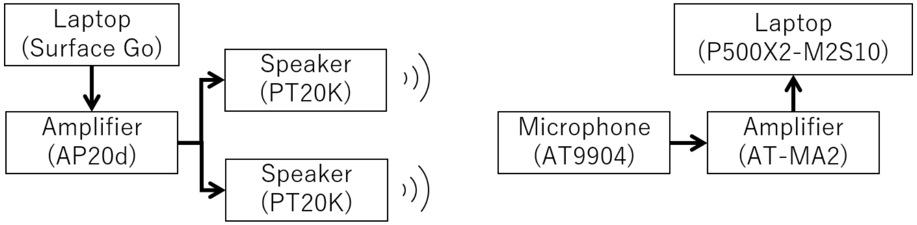


Figure 5. configuration of the measurement system.

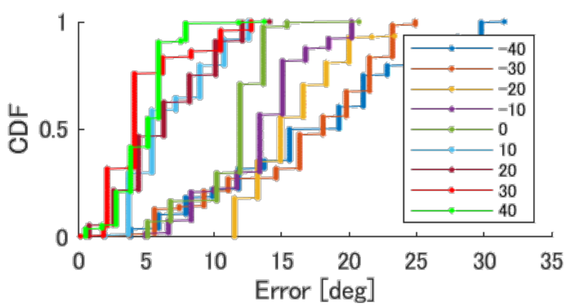
5. Results

<u>Azimuth estimation</u>

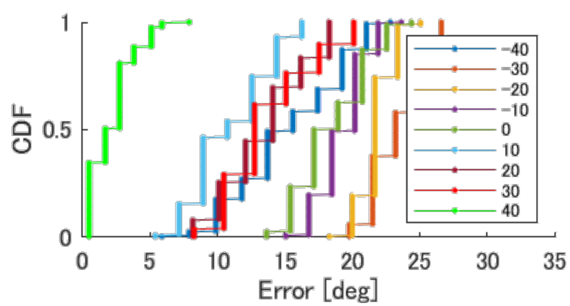
- Our proposed method shows the best estimation performance.
  - The standard deviation by the chirp, OFDM, and proposed methods were within 7.36, 3.93, and 1.15 degrees, respectively.

<u>DPSK</u>

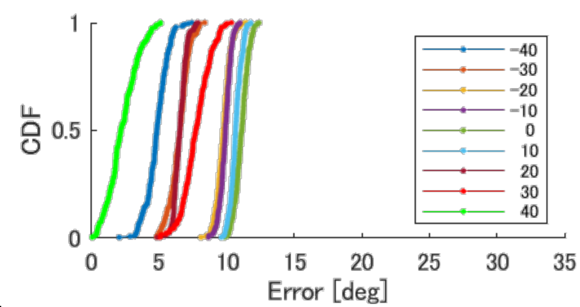
- the demodulation performance of each method were similar.
  - The standard deviation of demodulation by the chirp, OFDM, and proposed methods were within 3.73, 2.27, and, 2.95 degrees, respectively.



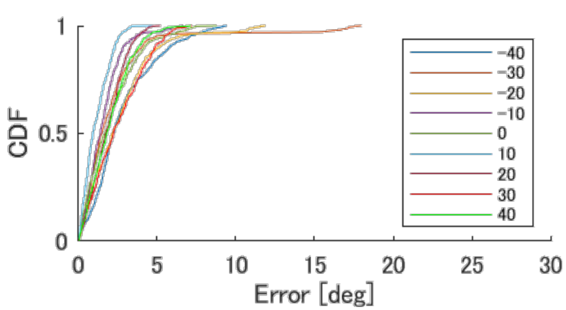
Azimuth Estimation: chirp method



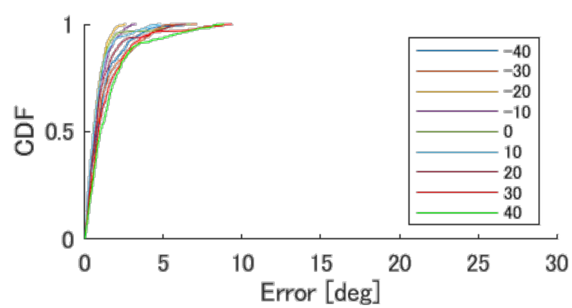
Azimuth Estimation: OFDM method



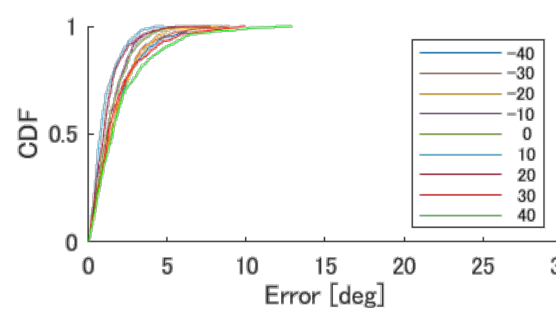
Azimuth Estimation: proposed method



Demodulation: chirp method



Demodulation: OFDM method



Demodulation: proposed method

Figure 6. Measurements results (Legends indicate the azimuth θ of the microphone's location)

6. Discussion: Systematic Error of Azimuth Estimation

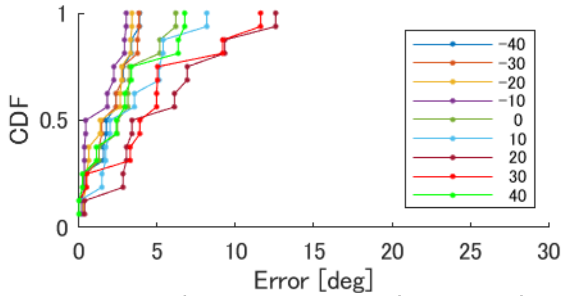
- There are clear differences between simulation and real-environment experiments, and the received noises are not enough to explain these differences. Herein, we discuss the causes for this mismatch.

Proposed method

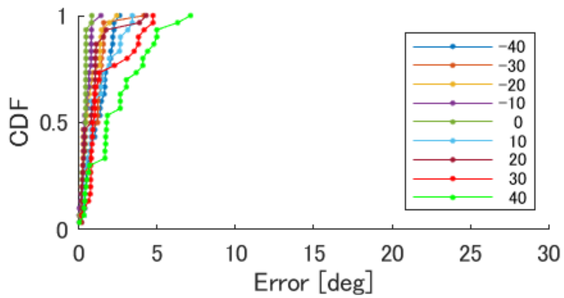
- Although systematic errors depending on the modulation value do not occur, there are some systematic errors at each location.
- We consider that these errors are caused by multipath signals reflected from the walls or ceilings in the room as there is no law on the relation between systematic errors and locations.

Conventional method

- In addition to the above multipath signals, the amplitude of the received signal that depends on the location could be one of the reasons.
- Therefore, we conducted simulations that consider the received amplitude of real environment experiments.
- Figure 7 shows the results of these simulations of the chirp method. These figures show that the systematic errors change in response to the received amplitude.
- Thus, in the conventional methods, the received amplitude is a factor in the difference of the systematic error between simulation and real environment experiments.



Azimuth Estimation: chirp method

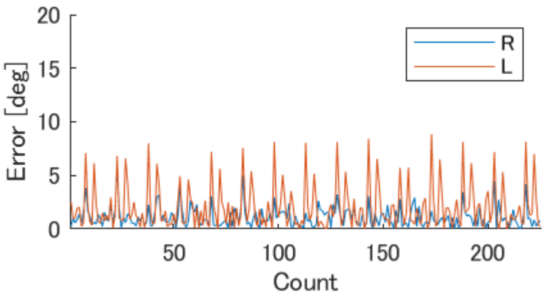


Demodulation: chirp method

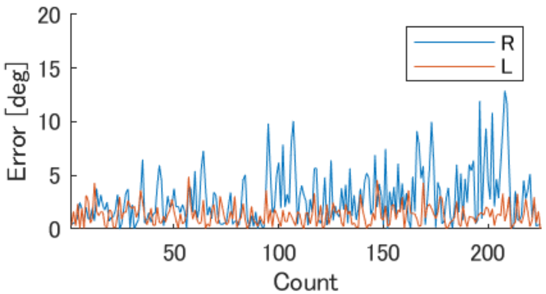
Figure 7. Simulation results that reflect the amplitude of received signals.

6. Discussion: Random Error of Demodulation

- Herein, we discuss the reason why the standard deviation of the proposed method tends to be larger than the OFDM method.
- Figure 8 shows the error of each symbol for the 16-symbol sequence (Table I) repeated 15 times at the received location with maximum standard deviation.
- Although the errors of the chirp method contains periodic systematic errors and random errors, its random error is smaller than that of the proposed method as the signal length used in the OFDM method is twice as long as in our proposed method.
- Therefore, our proposed method has a disadvantage in terms of SNR, and its performance might be worse than conventional methods when systematic errors are small.



Chirp method



Proposed method

Figure 8. Demodulation error of each symbol (Legends indicate speaker)

7. Conclusion

- In this paper, we proposed highly precise localization and communication methods using short-time and narrow-band acoustic signals.
- The simulation and real-environment experiments showed that our proposed method could reduce systematic errors compared to the conventional methods as interference between the acoustic signals could be avoided.