Accurate and Precise Acoustic Positioning Using Simultaneous Transmission of Phase-shifted Pulses

Masanari Nakamura[1,*], Konatsu Abe[1], Hiromichi Hashizume[2] and Masanori Sugimoto[1]

[1] *Graduate School of Information Science and Technology, Hokkaido University, Kita 14, Nishi 9, Kita-ku, Sapporo, Hokkaido, 060-0814 Japan*

[2] *National Institution for Academic Degrees and Quality Enhancement of Higher Education, 1-29-1 Gakuen-nishimachi Kodaira-shi, Tokyo 187-8587 Japan*

Abstract

In a multipath environment, to obtain accurate received times for multiple speakers using a single receiver, it is necessary to use broadband transmission signals and avoid interference between the signals. For such a case, we propose a transmission/reception method to improve the precision while maintaining the accuracy. In our proposed method, all speakers simultaneously transmit phase-shifted pulses multiple times and the overlapped received signals are separated without interference. Through a real environmental experiment, we confirmed the effectiveness of this method.

Keywords

Acoustic signal, phase shift keying, indoor positioning, reverberation

1. Introduction

With the spread of mobile devices, such as smartphones, various location-based services are being developed. In indoor environments, global navigation satellite system (GNSS) could have significant errors, owing to the shielding of radio waves. Therefore, indoor positioning methods using sensors embedded in mobile devices have been widely studied [1].

Almost all mobile devices have a built-in microphone for calls and voice recognition. With multiple speakers installed in an indoor environment, the positions of microphones of multiple mobile devices can be estimated simultaneously. In such a method, the microphone receives signals transmitted from each speaker, and their received times are estimated. The microphone's position is then calculated using the differences between these received times.

In indoor environments, the microphone receives a direct wave and reflected waves. The reflected waves are received with a delay, relative to the direct wave, and could cause systematic errors by interfering it. To avoid this error, the time resolution must be improved. This can be realized by increasing the bandwidth allocated to each speaker. In practice, however, the available bandwidth is limited. Some techniques, such as code-division multiplexing (CDM), assign a broadband signal to each speaker. However, this could cause interference between signals and result in systematic errors.

To avoid these problems, time-division multiplexing (TDM) is commonly used. It allocates the entire bandwidth signal to each speaker and transmits the signals in sequence, as shown in Figure 1. In TDM transmission schemes, if signals are transmitted, even in empty slots, the signal-to-noise ratio (SNR) can be improved by increasing the signal energy. However, interference between signals could reduce the accuracy. Therefore, it is necessary to be able to separate the signals for each speaker without interference.

In this paper, we propose a method that simultaneously transmits signals from N speakers N times and separates them without interference after reception. This method can increase the SNR by N times while maintaining accuracy as high as TDM. The proposed method improves the SNR by N times

IPIN-WCAL 2025: Workshop for Computing & Advanced Localization at the Fifteenth International Conference on Indoor Positioning and Indoor Navigation, September 15–18, 2025, Tampere, Finland

*Corresponding author.

masanari@ist.hokudai.ac.jp (M. Nakamura)

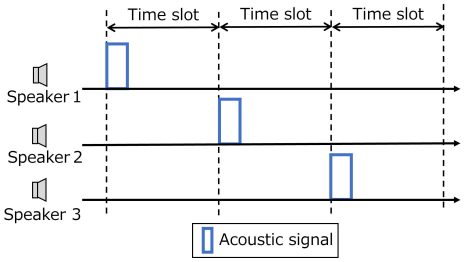

Figure 1: TDM transmission scheme.

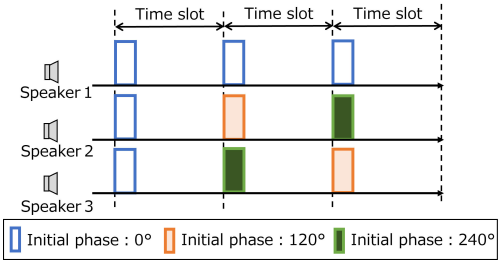

Figure 2: PDM transmission scheme.

because N signals are added for each speaker in the separation process. The proposed method uses phase-shift keying modulated chirp signals. By setting the appropriate phase value for each signal, the received signal can be separated into signals from each speaker without interference.

The main contributions of this study are as follows:

- We devised a phase-division multiplexing (PDM) transmission/reception method that improves the SNR by N times while maintaining accuracy, using the same number of slots as TDM.
- Through real-environment experiments, we confirmed the effectiveness of the proposed method. Specifically, we showed that the precision was improved with the same level of systematic error as TDM.

This paper is organized as follows: Section 2 shows the related work for this paper. In Section 3, the details of the proposed method are described. In Section 4, the effectiveness of the proposed method is shown through comparative experiments with the conventional method (TDM) in a real environment. In Section 5, we discuss the experimental results and limitations. We present the conclusion in Section 6.

2. Related work

In a positioning method called time difference of arrival (TDoA), acoustic signals are transmitted from multiple speakers, and the microphone position is determined using the difference between the received times of these signals. When there are three speakers, let δ_1, δ_2, and δ_3 be the transmission times of the acoustic signals and t_1, t_2, and t_3 be the received times of each signal. The relationship between these times and the positions of the speakers and microphone is represented as follows:

$$\|\boldsymbol{p}_2 - \boldsymbol{p}_m\| - \|\boldsymbol{p}_1 - \boldsymbol{p}_m\| = c((t_2 - \delta_2) - (t_1 - \delta_1)) \tag{1}$$

$$\|\boldsymbol{p}_3 - \boldsymbol{p}_m\| - \|\boldsymbol{p}_2 - \boldsymbol{p}_m\| = c((t_3 - \delta_3) - (t_2 - \delta_2)) \tag{2}$$

where the 2D positions of the three speakers are $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, and $\boldsymbol{p}_3$. $\boldsymbol{p}_m$ indicates the 2D position of the microphone. $\|\cdot\|$ is the Euclidean norm. Solving this for $\boldsymbol{p}_m$ gives the position of the microphone.

Next, we describe the transmission/reception method to obtain the received times. There are two main types of transmission schemes: simultaneous transmissions and sequential transmissions from each speaker. First, we explain the case of simultaneous transmission ($\delta_1 = \delta_2 = \delta_3 = 0$). This method requires only one slot to obtain all the received times. Therefore, the positioning time is short. Let $s_i(t)$ $(0 \leq t \leq T_S)$ be the signal from speaker i and $h_i(t)$ be the impulse response representing the propagation characteristics between speaker i and the microphone. The received signal $r(t)$ is

$$r(t) = \sum_{i=1}^{3} (h_i * s_i)(t - t_i) + n(t) \tag{3}$$

where $*$ indicates a convolution. $n(t)$ is noise following a white Gaussian distribution.

Let $\boldsymbol{r}$ be the discretized form of $r(t)$ and $\boldsymbol{e}^1$, $\boldsymbol{e}^2$, and $\boldsymbol{e}^3$ be the vector of discretized analytic signals corresponding to transmission signals $s_1(t)$, $s_2(t)$, and $s_3(t)$. These are convolved into $\boldsymbol{r}$ to enhance the signal.

$$\boldsymbol{c}_i = [c_i^1, c_i^2, \dots], \qquad c_i^k = \sum_{q=1}^{S} r_{k+q-1} e_q^i \tag{4}$$

where r_d and e_d^i are the dth elements of $\boldsymbol{r}$ and $\boldsymbol{e}^i$, respectively. S is the number of samples when $s_i(t)$ is discretized. This process is known as a matched filter [2]. The received time t_i is obtained as the time when the absolute value of $\boldsymbol{c}_i$ is the maximum.

For positioning, $s_1(t)$, $s_2(t)$, and $s_3(t)$ must be different to identify which speaker transmitted it. When calculating $\boldsymbol{c}_i$, $\boldsymbol{e}^i$ could respond to $s_l(t)$ $(l \neq i)$. This causes an error in the estimated t_i. In particular, when the microphone position is close to speaker l and far from speaker i, the response of $s_l(t)$ during the convolution of $\boldsymbol{e}^i$ is larger. This is because the received signal of speaker l is larger, causing a systematic error. This is called the near–far problem.

Various signal-multiplexing methods have been proposed to reduce such interference, such as frequency-division multiplexing (FDM) [3, 4, 5], orthogonal frequency-division multiplexing (OFDM) [6] and orthogonal chirp-based method [7]. However, when pulse signals are superimposed, the interference between these signals cannot be reduced to zero at an arbitrary position, in principle[1].

When reverberation is present, as in Equation (3), the reverberation responds to $\boldsymbol{e}^i$ and could cause systematic errors in estimating the received time. This error can be reduced by increasing the bandwidth of the acoustic signal and improving the time resolution of the matched filter. For this reason, code-division multiplexing (CDM) and spread-spectrum methods, such as [8, 9, 10, 11, 12, 13, 14], have been proposed to assign differently modulated signals using the same wide band to each speaker. However, interference of these techniques cannot also be reduced to zero at an arbitrary position[1].

By transmitting signals sequentially from each speaker, the interference between the signals of each speaker can be zero. This method is known as TDM. TDM enables us to assign the entire bandwidth to all speakers, thus maximizing the suppression of reverberation. Therefore, TDM is suitable for situations where accurate and precise positioning is required, such as robot arm control [15, 16]. However, because the transmissions from each speaker are sequential, the time required for positioning increases compared to that of simultaneous transmissions.

In TDM, transmitting signals in empty slots can improve the precision because the signal energy increases. However, this can cause interference, which reduces accuracy. Therefore, the signals must be separated without interference to improve the SNR while maintaining accuracy.

3. Proposed method

In this section, we propose a transmission/reception method that improves the SNR by N times while maintaining the accuracy of TDM. This is realized by transmitting signals N times from N speakers simultaneously and separating them without interference at the receiver.

In the proposed method, the following chirp signal is used for the transmitted signal $s_i(t)$ of speaker i $(i = 1, 2, \dots, N)$.

$$s_i(t) = \sum_{m=1}^{N} s_i^m(t), \quad s_i^m(t) = \begin{cases} f(t,i,m) & (m-1)T_M \le t \le (m-1)T_M + T_S \\ 0 & (m-1)T_M + T_S < t < mT_M \end{cases}, \tag{5}$$

$$f(t,i,m) = w(t) \sin\left(2\pi\left(f_0 t + \frac{F}{2}t^2\right) + \phi_i^m\right), \tag{6}$$

[1]The convolution of pulses can be expressed as the product of a matrix and a vector. Because the upper part of this matrix is a lower triangular matrix, the columns of this matrix are linearly independent. Therefore, the only vector that makes the result of this product a zero vector is the zero vector. This indicates that pulses that do not interfere with each other cannot be constructed.

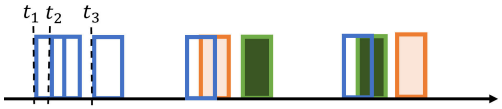

Figure 3: Illustration of received signals.

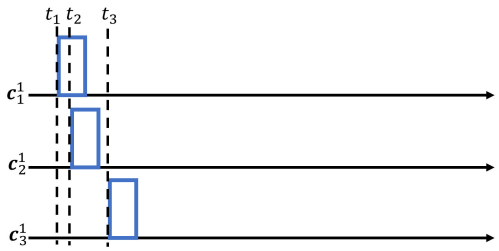

Figure 4: Illustration of the filter output of each speaker after separation.

$$\phi_i^m = \frac{2\pi \times (m-1) \times (i-1)}{N}, \quad F = \frac{f_1 - f_0}{T_S} \tag{7}$$

where m is an index representing the transmission order ($m = 1, 2, \ldots, N$). f_0 and f_1 are the start and end frequencies of the sweep. These values are set to occupy the entire available frequency band. $w(t)$ is a Tukey window with window length T_S and the parameter set to 0.5. T_M is the transmission interval and should be set to a value greater than the reverberation time.

As can be seen from these equations, the transmitted signals are phase-shift modulated chirp signals. Because the proposed method uses this phase-shift modulation for multiplexing, we call the proposed method PDM. Figure 2 shows the transmission scheme.

The received signal is represented as follows:

$$r(t) + n(t), \quad r(t) = \sum_{m=1}^{N} \sum_{i=1}^{N} r_i^m(t), \tag{8}$$

$$r_i^m(t) = (h_i * s_i^m)(t - t_i). \tag{9}$$

Figure 3 shows an example of the received signal when there are three speakers. In this figure, all received signals are represented as rectangles for convenience, but the actual received signals are different for each speaker, owing to reverberation.

We discretize this signal and select $MN + S - 1$ samples[2]. Here, at least the first S samples should not contain any signal components. M and S represent the number of points when discretizing the transmission period T_M and signal length T_S, respectively. Let $\boldsymbol{g}$ and $\boldsymbol{n}$ respectively be the signal and noise components of the Hilbert transform of this vector.

Here, in a preparation for the following explanation, we show that the part of $\boldsymbol{g}$ corresponding to the mth transmission signal of speaker i can be represented as the product of the part corresponding to the first transmission signal of speaker i and $\exp(j\phi_i^m)$, as shown in Equation (14). Let $\boldsymbol{g}_i$ be the result of the above process with signals transmitted from only speaker i under ideal conditions without noise. In this case, $\boldsymbol{g}$ can be represented as $\boldsymbol{g} = \sum_{i=1}^{N} \boldsymbol{g}_i$.

The division of $\boldsymbol{g}_i$ into M samples can be represented as follows:

$$\boldsymbol{g}_i = [\boldsymbol{g}_i^1, \boldsymbol{g}_i^2, \ldots, \boldsymbol{g}_i^N, \mathbf{0}_{S-1}]. \tag{10}$$

$\mathbf{0}_{S-1}$ is a zero vector of length $S - 1$. In this case, the non-zero component of the Hilbert transform of $r_i^m(t)$ is contained only in the mth frame $\boldsymbol{g}_i^m$.

Here, the pulse $f(t, i, m)$ of the transmission signals can be represented as follows:

$$\frac{w(t)}{2j} \exp\left(j\left(2\pi\left(f_0 t + \frac{F}{2}t^2\right) + \phi_i^m\right)\right) - \frac{w(t)}{2j} \exp\left(-j\left(2\pi\left(f_0 t + \frac{F}{2}t^2\right) + \phi_i^m\right)\right). \tag{11}$$

Thus, the result of the Hilbert transform of Equation (9) is

$$\mathcal{H}(r_i^m(t)) = (h_i * v_i^m)(t - t_i) \tag{12}$$

[2]The $S - 1$ sample points are required to bring the matched-filter output to MN points.

$$v_i^m(t) = \begin{cases} \frac{w(t)\exp\left(j\left(2\pi\left(f_0 t + \frac{F}{2}t^2\right) + \phi_i^m\right)\right)}{2j} & (m-1)T_M \le t \le (m-1)T_M + T_S \\ 0 & (m-1)T_M + T_S < t < mT_M. \end{cases} \tag{13}$$

Note that

$$\exp\left(j\left(2\pi\left(f_0 t + \frac{F}{2}t^2\right) + \phi_i^m\right)\right) = \exp(j\phi_i^m)\exp\left(j2\pi\left(f_0 t + \frac{F}{2}t^2\right)\right), \quad \exp(j\phi_i^1) = 1.$$

Therefore, Equation (10) is

$$\boldsymbol{g}_i = [\boldsymbol{g}_i^1,\ \exp(j\phi_i^2)\boldsymbol{g}_i^1,\ \ldots,\ \exp(j\phi_i^N)\boldsymbol{g}_i^1,\ \boldsymbol{0}_{S-1}]. \tag{14}$$

We apply the matched filter to the Hilbert transform output $\boldsymbol{g} + \boldsymbol{n}$. Letting $\boldsymbol{y}$ denote this output, its kth element can be represented as follows:

$$y_k = \sum_{q=1}^{S}(g_{k+q-1} + n_{k+q-1})e_q. \tag{15}$$

e_d is the dth element of the discretized complex chirp signal:

$$e(t) = w(t)\exp\left(j2\pi\left(f_0 t + \frac{F}{2}t^2\right)\right), \ \ 0 \le t \le T_S. \tag{16}$$

Let $\boldsymbol{c}$ be the signal component of $\boldsymbol{y}$ and $\boldsymbol{n}'$ be the noise component of $\boldsymbol{y}$. Let $\boldsymbol{c}_i$ be the component of $\boldsymbol{c}$ transmitted by speaker i; then, these relations are

$$\boldsymbol{c} = \sum_{i=1}^{N}\boldsymbol{c}_i. \tag{17}$$

We divide $\boldsymbol{c}_i$ into M-sample.

$$\boldsymbol{c}_i = [\boldsymbol{c}_i^1,\ \boldsymbol{c}_i^2,\ \ldots,\ \boldsymbol{c}_i^N] \tag{18}$$

Because the matched filter is a linear operation, this can be represented as

$$\boldsymbol{c}_i = [\boldsymbol{c}_i^1,\ \exp(j\phi_i^2)\boldsymbol{c}_i^1,\ \ldots,\ \exp(j\phi_i^N)\boldsymbol{c}_i^1], \tag{19}$$

as in Equation (14). Therefore, based on Equations (17) and (19), $\boldsymbol{y}$ can be represented as follows:

$$\boldsymbol{y} = A\boldsymbol{c}' + \boldsymbol{n}' \tag{20}$$

$$A = \begin{pmatrix} I_M & I_M & \cdots & I_M \\ I_M & e^{j2\pi/N}I_M & \cdots & e^{j2\pi(N-1)/N}I_M \\ \vdots & \vdots & \ddots & \vdots \\ I_M & e^{j2\pi(N-1)/N}I_M & \cdots & e^{j2\pi(N-1)(N-1)/N}I_M \end{pmatrix} \tag{21}$$

$$\boldsymbol{c}' = (\boldsymbol{c}_1^1, \boldsymbol{c}_2^1, \cdots, \boldsymbol{c}_N^1)^*, \tag{22}$$

where * indicates a conjugate transpose. I_M indicates the $M \times M$ identity matrix. The kth element of $\boldsymbol{n}'$ is

$$n_k' = a + jb, \ \ a,\ b \sim \mathcal{N}(0, \sigma^2). \tag{23}$$

We estimate $\boldsymbol{c}'$ using $\boldsymbol{y}$. Because A is a unitary matrix, the inverse matrix of A exists.

$$A^{-1} = \frac{1}{N}A^* \tag{24}$$

By multiplying $\boldsymbol{y}$ by A^{-1}, we obtain $\boldsymbol{c}'$ as follows:

$$A^{-1}\boldsymbol{y} = \boldsymbol{c}' + A^{-1}\boldsymbol{n}' = \boldsymbol{c}' + \frac{1}{N}A^{*}\boldsymbol{n}'. \tag{25}$$

Figure 4 shows an example of this result. From Equation (21), the real and imaginary parts of the kth element of $(1/N)A^{*}\boldsymbol{n}'$ follow a Gaussian distribution[3] with mean 0 and variance σ^2/N.

Here, we discuss the SNR of TDM and PDM. The signal component of the filter output of speaker i is the same for TDM and PDM, $\boldsymbol{c}_i^1$. The noise component is $\boldsymbol{n}'$ for TDM and $(1/N)A^{*}\boldsymbol{n}'$ for PDM. The power of each element is $2\sigma^2$ for TDM and $(2/N)\sigma^2$ for PDM. This shows that the SNR of PDM is N times higher than that of TDM.

The estimated $\boldsymbol{c}_1^1,\ \boldsymbol{c}_2^1,\ \ldots,\ \boldsymbol{c}_N^1$ are independent matched-filter outputs for each speaker. Therefore, t_i can be obtained by finding the time at which the absolute value of $\boldsymbol{c}_i^1$ is maximum.

4. Evaluation

The proposed method assumes that indoor acoustic-wave propagation is a linear time-invariant system, such that Equation (9) holds. However, there is no guarantee that the reflected waves that generate reverberation have a linear response. If these nonlinearities are large, the effectiveness of the proposed method is degraded. Therefore, we verified the effectiveness of PDM through experiments in a real environment. Specifically, we evaluated whether the proposed method has the same level of systematic errors as TDM and whether it improves the precision.

4.1. Experimental setting

4.1.1. Measurement setting

The experimental scenario is the 2D positioning of a single microphone using three speakers. The speaker and its amplifier were a Fostex FT28D and Fostex AP20d, respectively. The microphone, pre-amplifier, and amplifier were a RION UC-31, RION NH-05A, and RION UN-14, respectively. We used a Roland Rubix24 as the audio interface for speakers and a microphone. In Figure 5, $\triangle$ indicates the position of the speakers. $\bigcirc$ shows the six positions of the microphone. The measurements were conducted 100 times at each position. The heights of the speakers and microphone were 1.2 m. We placed the speakers and microphone to face the $+y$ and $-y$ directions, respectively. To obtain the true values, we used the MotionAnalysis Mac3D System. The error of this system is less than 1 mm. The room temperature was 21.2°C when this experiment was conducted. The experimental environment is shown in Figure 6.

For the transmission signal, we set $f_0 = 15$ kHz, $f_1 = 23$ kHz, and $T_S = 10$ ms. These signals can be measured with microphones embedded in mobile devices such as smartphones [3, 10]. We set the transmission interval T_M to 210 ms. This value is sufficient to remove the reverberation [3, 10]. The sampling rate was set to 48 kHz. In estimating the received time, we upsampled $\boldsymbol{c}_i^1$ 10 times.

The audio interface used in this experiment had a systematic time error of 2.0417 ms between channel three and the other channels. Therefore, in this experiment, we computed the positions using the received times calibrated by this value.

4.1.2. Position calculation

In this experiment, we obtained the position by finding $\boldsymbol{p}_m$ that minimizes the following equation.

$$L(\boldsymbol{p}_m) = (\|\boldsymbol{p}_2 - \boldsymbol{p}_m\| - \|\boldsymbol{p}_1 - \boldsymbol{p}_m\| - c(t_2 - t_1))^2 + (\|\boldsymbol{p}_3 - \boldsymbol{p}_m\| - \|\boldsymbol{p}_2 - \boldsymbol{p}_m\| - c(t_3 - t_2))^2 \tag{26}$$

[3]Because $\boldsymbol{n}'$ is the matched-filter output, its kth element is correlated with the previous and next S samples. Thus, when $M < S$, the variance is not σ^2/N.

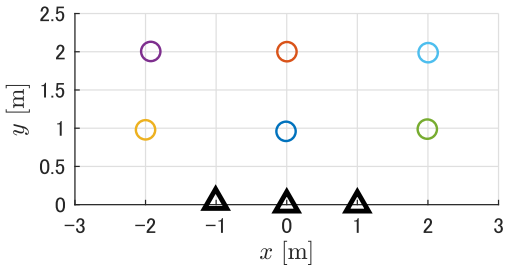

Figure 5: Positions of speakers and microphone (△: speaker, ◯: microphone).

Figure 6: Experimental environment.

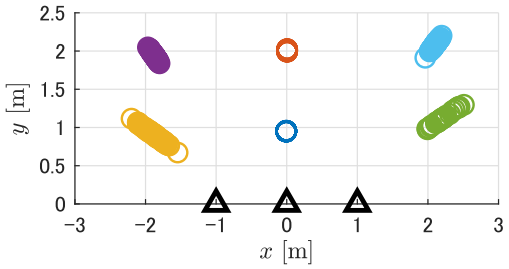

Figure 7: 2D plot of TDM (◯: result, △: speaker. Each color corresponds to the microphone position in Figure 5.).

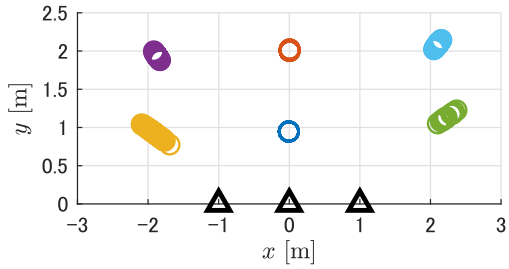

Figure 8: 2D plot of PDM (◯: result, △: speaker. Each color corresponds to the microphone position in Figure 5.).

Specifically, we first obtained the intersection of the lines that are asymptotes of the hyperbolas in Equations (1) and (2).

$$y = \alpha_{21}x - 0.5\alpha_{21}, \quad y = \alpha_{32}x + 0.5\alpha_{32} \tag{27}$$

α_{21} and α_{32} are the gradients of the asymptotes of each hyperbola and are obtained as follows:

$$\alpha_{21} = \frac{\sqrt{\|\boldsymbol{p}_2 - \boldsymbol{p}_1\|^2 - v_c^2(t_1 - t_2)^2}}{v_c(t_1 - t_2)}, \quad \alpha_{32} = \frac{\sqrt{\|\boldsymbol{p}_2 - \boldsymbol{p}_3\|^2 - v_c^2(t_2 - t_3)^2}}{v_c(t_2 - t_3)} \tag{28}$$

where v_c is the speed of sound.

For Equation (26), we conducted a grid search in 1-cm increments in the range ±1 m centered at this intersection point. Then, another grid search was conducted in 1-mm increments within ±10 cm centered at the position of the first grid search result.

4.2. Measurement results

We show 2D plots of the TDM and PDM positioning results and the cumulative distribution function (CDF) in Figures 7, 8, and 9, respectively. Table 1 shows the ratio of variances between TDM and PDM on each axis and systematic errors of each method in order. Only at position (2, 1), 67 and 75 outliers occurred in TDM and PDM, respectively. We excluded these outliers in the figures and tables.

Figures 7, 8 and 9, and Table 1 show that PDM improved the precision and its systematic error was comparable to that of TDM. From Table 1, we can see that the variance on each axis is approximately one-third compared to that of TDM. These indicate the effectiveness of the proposed method.

For both TDM and PDM, the spread of the positioning results increases as the distance from the coordinate origin increases. This depends on the relationship between the speakers and microphone positions.

Table 1
Ratio of variances on each axis and systematic error

Position [m]	(-2,1)	(-2,2)	(0,1)	(0,2)	(2,1)	(2,2)
$\frac{\sigma^2_{PDM(x)}}{\sigma^2_{TDM(x)}}$	0.483	0.263	0.429	0.364	0.217	0.328
$\frac{\sigma^2_{PDM(y)}}{\sigma^2_{TDM(y)}}$	0.493	0.282	0.346	0.555	0.217	0.314
TDM [m]	0.142	0.089	0.008	0.006	0.261	0.138
PDM [m]	0.120	0.084	0.012	0.009	0.276	0.142

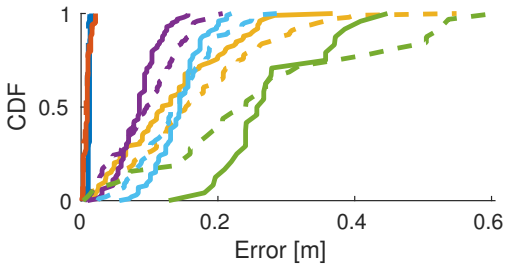

Figure 9: CDF at each position (dotted line: TDM, solid line: PDM. Each color corresponds to the microphone position in Figure 5.).

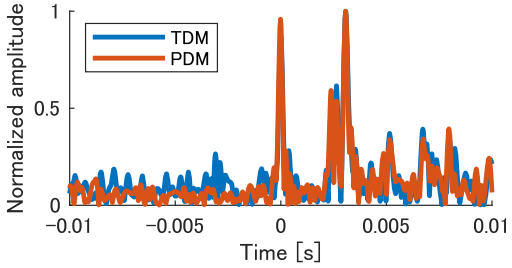

Figure 10: Filter-output example of a signal transmitted by a speaker at position (2, 1).

5. Discussion

5.1. Outlier

Figure 10 shows a filter-output example of a signal transmitted by a speaker when an outlier occurs at position (2, 1). This figure shows that a reflected wave from objects is larger than the direct wave. Such reflected waves may be suppressed by post-processing such as the constant false alarm rate (CFAR) [2]. Note that this figure shows that PDM can suppress noise in the real data, as the theory suggests.

5.2. Computational complexity

Compared to TDM, the additional calculation in PDM is Equation (25). Because there are only N nonzero components in each column of A^{-1} and the unit-matrix part of A^{-1} does not require multiplication, the required number of multiplications is $(N-1) \times M$. For Equation (25), as c_i^1 contains a direct-wave component in its front, the only range requiring computation is the front of each c_i^1. This can further reduce the number of product operations.

5.3. Microphone movement

In our experiments, measurements were conducted with a stationary microphone. If the microphone moves during the measurements, the reverberation will change. In this case, there is a difference between the observed data and the observation model (Equation (20)), and an error may occur in the filter output $\boldsymbol{c}'$ after separation. We would like to address this point in future work.

6. Conclusion

In this paper, we proposed a method to improve the precision without degrading the accuracy of TDM by simultaneously transmitting pulse sequences with the same bandwidth and different initial phases and separating them at the receiver side. Through real-world measurements, we confirmed that our method can improve the positioning precision without worsening the accuracy compared to the conventional method (TDM).

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers JP21K17797.

Declaration on Generative AI

During the preparation of this work, the authors used Gemini in order to: Grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] J. Xiao, Z. Zhou, Y. Yi, L. M. Ni, A survey on wireless indoor localization from the device perspectivew, *ACM Computing Surveys* 49 (2017) 1–31.

[2] M. A. Richards, J. A. Scheer, W. A. Holm (Eds.), Principles of Modern Radar: Basic principles, 1st ed., Scitech publishing, 2010.

[3] X. Chen, Y. Chen, S. Cao, L. Zhang, X. Zhang, X. Chen, Acoustic indoor localization system integrating tdma+fdma transmission scheme and positioning correction technique, *Sensors* 19 (2019) 1–18.

[4] M. Nakamura, H. Hashizume, M. Sugimoto, Simultaneous localization and communication method using short-time and narrow-band dual-carrier acoustic signals, *IEEE Sensors Journal* 22 (2022) 5163–5172.

[5] M. Nakamura, H. Kameda, Indoor localization using multiple stereo speakers for smartphones, in: *IPIN-WiP 2019*, volume 2498, 2019, pp. 235–242.

[6] M. O. Khyam, M. J. Alam, A. J. Lambert, M. A. Garratt, M. R. Pickering, High-precision OFDM-based multiple ultrasonic transducer positioning using a robust optimization approach, *IEEE Sensors Journal* 16 (2016) 5325–5336.

[7] M. O. Khyam, S. S. Ge, X. Li, M. Pickering, Orthogonal chirp-based ultrasonic positioning, *Sensors* 17 (2017) 1–14.

[8] M. O. Khyam, S. S. Ge, X. Li, M. R. Pickering, Pseudo-orthogonal chirp-based multiple ultrasonic transducer positioning, *IEEE Sensors Journal* 17 (2017) 3832–3843.

[9] M. O. Khyam, M. Noor-A-Rahim, X. Li, C. Ritz, Y. L. Guan, S. S. Ge, Design of chirp waveforms for multiple-access ultrasonic indoor positioning, *IEEE Sensors Journal* 18 (2018) 6375–6390.

[10] P. Lazik, A. Rowe, Indoor pseudo-ranging of mobile devices using ultrasonic chirps, in: *ACM Sensys 2012*, 2012, pp. 99–112.

[11] F. J. Álvarez, T. Aguilera, R. López-Valcarce, CDMA-based acoustic local positioning system for portable devices with multipath cancellation, *Digital Signal Processing* 62 (2017) 38–51.

[12] C. Sertatıl, M. A. Altınkaya, K. Raoof, A novel acoustic indoor localization system employing CDMA, *Digital Signal Processing* 22 (2012) 506–517.

[13] S. Murano, C. Pérez-Rubio, D. Gualda, F. J. Álvarez, T. Aguilera, C. D. Marziani, Evaluation of zadoff–chu, kasami, and chirp-based encoding schemes for acoustic local positioning systems, *IEEE Transactions on Instrumentation and Measurement* 69 (2020) 5356–5368.

[14] I. Rishabh, D. Kimber, J. Adcock, Indoor localization using controlled ambient sounds, in: *IPIN 2012*, 2012, pp. 1–10.

[15] H. Schweinzer, M. Syafrudin, Losnus: An ultrasonic system enabling high accuracy and secure tdoa locating of numerous devices, in: *IPIN 2012*, 2010, pp. 1–8.

[16] J. C. Prieto, A. R. Jimenez, J. Guevara, J. L. Ealo, F. Seco, J. O. Roa, F. Ramos, Performance evaluation of 3d-locus advanced acoustic lps, *IEEE Transactions on Instrumentation and Measurement* 58 (2009) 2385–2395.