

Short-Time Noise Suppression Method with Multiple Signals for Received Time Estimation

Masanari Nakamura¹, Akira Tanaka¹, Hiromichi Hashizume² and
Masanori Sugimoto¹

¹Hokkaido University, Kita 8, Nishi 5, Kita-ku, Sapporo, Hokkaido, 060-0808 Japan

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

Abstract

The detection of the leading edge of a signal using a threshold value can be used as a method for estimating the received time, which is not easily affected by reverberation or the sensor characteristics of the speaker and microphone. When the noise level is relatively high compared with the signal, the threshold must be set high to prevent false alarms, causing a systematic error in the received time estimation. The noise can be suppressed by transmitting the signal multiple times and averaging the results; however, it is necessary to wait for the reverberation to disappear after each transmission. Therefore, this measurement is time consuming. In this study, we propose a noise suppression method that can be applied to cases in which signals are transmitted multiple times without waiting for the disappearance of reverberation. For the proposed method, an analysis of noise suppression performance was conducted, and the limits of performance were provided under the given conditions. In addition, numerical experiments on the performance of the received time estimation in a reverberant environment were conducted to compare the proposed method with the conventional method and confirm the effectiveness of the proposed method.

Keywords

received time estimation, threshold, false alarm, reverberation, sensor characteristics

1. Introduction

With the recent spread of mobile devices, indoor positioning techniques for these devices have been extensively studied[1]. In this study, we discuss a positioning method that uses the propagation time of acoustic signals. When using the propagation time, it is important to estimate the received time of the acoustic signal. In indoor environments, reverberation and the physical characteristics of speakers and microphones can cause large systematic errors in the estimation of received time.


The detection of the leading edge of the signal with a threshold value can be used as a method that is less sensitive to reverberation and sensor characteristics. In detection using thresholds, it is necessary to set a threshold based on the noise level to avoid false noise detection. When the signal-to-noise ratio is low, the threshold must be set high, the parts of the leading edge of the signal are delayed, and a systematic error is caused in the estimation of the received

IPIN 2022 WiP Proceedings, September 5 - 7, 2022, Beijing, China

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



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

time. Therefore, to reduce systematic errors with the desired probability of a false alarm, the signal-to-noise ratio must be improved.

If the noise follows a white normal distribution, it can be suppressed by transmitting the signal multiple times and averaging them. However, this method requires waiting for the disappearance of the reverberation after each transmission. Therefore, it takes a long time to send all the signals. This could be a problem, especially when positioning is required at many locations.

In this study, we propose a noise suppression method that can be applied even when the transmission time is set shorter than the reverberation time to reduce the time required for signal transmission. By applying an orthogonal projection matrix that preserves the transmitted signal component of the received signal, only the noise component is suppressed.

The contributions of this study are as follows:

- We devised a noise suppression method that can be applied when the transmission period is set to less than the reverberation time.
- The relationship between the transmission period and the number of transmissions and the noise suppression performance of the proposed method is discussed. The limitations of the noise suppression performance of the proposed method under given conditions are also provided.
- We conducted numerical experiments on the performance of received time estimation under multiple reverberation conditions, and demonstrated the effectiveness of the proposed method over the conventional method. In addition, we conducted numerical experiments regarding the performance of received time estimation when the required transmission time is changed, and discussed their relationship.

The remainder of this study is structured as follows. In Section 2, we describe conventional methods for estimating the received time. In Section 3, we describe the proposed method and discuss the relationship between the transmission period, number of transmissions, and noise suppression performance. In Section 4, we discuss the results of the experiments comparing the proposed method with conventional methods. Finally, in Section 5, we present our conclusions.

2. Related work

2.1. Observation model

First, we described the observational model discussed in this study. The start times of the transmission and the transmission signal are set to 0 and $s(t)$, respectively. The signal length was T . A signal transmitted through a speaker is nonlinearly distorted, depending on the physical properties of the speaker. This is expressed as follows:

$$f(t) = \Theta_s(s(t)) \quad (1)$$

$f(t)$ and $\Theta_s(\cdot)$ represent the signal distorted by the speaker and the characteristics of the speaker that cause distortion, respectively.

In indoor environments, reverberation is caused by multiple delayed waves reflected on walls, ceilings, and other surfaces. When this signal is observed using a microphone, nonlinear distortion occurs at the microphone. Let $h(t)$ and $\Theta_r(\cdot)$ be the indoor impulse response and microphone characteristics that cause distortions, respectively. The reverberation signal can be expressed as follows:

$$g(t) = \Theta_r((h * f)(t)) \quad (2)$$

* is a convolution.

Let Δt be the time required for propagation between the speaker and microphone. The received signal can be expressed as:

$$r(t) = g(t - \Delta t - \delta) + n(t) \quad (3)$$

δ represents the clock offset between the transmitter and receiver. $n(t)$ is white normal noise.

If δ is known, the propagation time Δt can be estimated from the received time. The distance can be estimated by multiplying Δt by the propagation velocity. Even if δ is not known, δ can be eliminated if the transmission times of multiple speakers are known. Therefore, the distance between the speaker and the microphone can be estimated.

2.2. Received time estimation by maximum peak detection

The matched filter is a well-known method for estimating the received time in noise[2]. In the following, we explain the matched filter in the case where the transmission signal $s(t)$ is a sinusoidal wave, as an example.

$$s(t) = \begin{cases} \sin(2\pi ft) & 0 \leq t \leq T \\ 0 & otherwise \end{cases} \quad (4)$$

f is frequency, which is a constant multiple of $1/T$. The matched filter output is expressed as

$$c(t) = \frac{1}{T} \int_{-\infty}^{\infty} r(\tau) e(\tau - t) d\tau, \quad (5)$$

$$e(t) = \begin{cases} \exp(j2\pi ft) & 0 \leq t \leq T \\ 0 & otherwise \end{cases} \quad (6)$$

where $e(t)$ denotes the reference signal. For simplicity, δ is assumed as zero in this study. If there is no influence of reverberation or sensor characteristics, the absolute value of $c(t)$ reaches its maximum value, where $t = \Delta t$. Therefore, in this case, the received time must be the time when the absolute value of $c(t)$ is largest. In this paper, this method is referred to as the maximum peak method. The maximum peak method is optimal when there is only white normal noise and no reverberation or sensor characteristic influence[2], and it is used in various acoustic positioning methods[3, 4, 5, 6, 7, 8, 9].

In a reverberant environment, as expressed in Eq. (5), the reference signal also responds to the delayed waves. Therefore, there is an error in the received time estimation with the

maximum peak method when the delayed wave is superimposed on the direct wave or when the delayed wave is larger than the direct wave.

The maximum peak method assumes that the direct waves of the received and reference signals are coincident. However, in practice, because distortions occur depending on the sensor (Eq. (1)), it has been reported that systematic errors can occur [10].

One way to avoid both the reverberation and distortion of the received signal is to measure $g(t)$ in advance and use this signal as a reference. However, because the delayed waves that constitute reverberation and the sensor characteristics vary depending on the location of the speaker and microphone, a large amount of premeasurement is required.

2.3. Received time estimation by leading-edge detection

2.3.1. False alarm probability and threshold

To avoid the effects of reverberation and sensor characteristics, detection of the leading edge of the signal can be used to estimate the received time. Specifically, a threshold was set for the absolute value of the matched filter output, and the first point at which it exceeded the threshold was detected. From the matched filter equation (Eq. (5)), the time of the detected point is T earlier than the received time. Therefore, the received time was the sum of the detected time and T . Hereafter, this method is referred to as the leading-edge method.

In this section, we discuss the probability of false noise detection. This is called the false-alarm probability. When a matched filter is applied to white normal noise, the probability density distribution of the absolute value of its output is a Rayleigh distribution. We describe the Rayleigh distribution when a discretized matched filter is used because the process with discrete data is discussed in Section 3 and later. Let N denote the signal length T converted to the number of samples. Let T_s denote the sampling period, and let T be a consist multiple of T_s . In this case, $N = T/T_s$. Let τ be the threshold; then, the false alarm probability P_{FA} is calculated as follows:

$$P_{FA} = \int_{\tau}^{\infty} \frac{x}{E^2} \exp\left(-\frac{x^2}{2E^2}\right) dx \quad (7)$$

The Rayleigh parameter E is computed using the power spectrum σ^2 of the white normal noise as $E^2 = \sigma^2/2$. Using the cumulative distribution function of the Rayleigh distribution, the above equation can be expressed as $P_{FA} = \exp(-\tau^2/(2E^2))$. Solving this for τ provides a threshold that allows signal detection at a specified false-alarm probability as $\tau = \sqrt{-2E^2 \log(P_{FA})}$.

As can be observed from these equations, setting a large threshold τ reduces the false-alarm probability. However, a larger threshold value caused a greater delay in the estimated received time. Therefore, to reduce the systematic error in the received time with a sufficiently small false-alarm probability, the signal-to-noise ratio of the received signal must be improved.

2.3.2. Improvement of signal-to-noise ratio by averaging

Let n_1, n_2, \dots, n_Z be samples that follow a white normal distribution with mean 0 and variance σ^2 , and the variance of the average $\sum_{i=1}^Z n_i/Z$ is σ^2/Z . Therefore, if the noise of the

received signal follows a white normal distribution, it can be suppressed by averaging multiple transmitted signals. Hereafter, this method will be referred to as the averaging method.

In reverberant environments, if the averaging method is used with the transmission period set to less than the reverberation time, the transmitted signal components will be corrupted during averaging by the delay waves of reverberation. In this case, the received time cannot be properly estimated. Therefore, to improve the signal-to-noise ratio, the signal transmission period should be longer than the reverberation time. However, this requires a significant amount of time. In the next section, we propose a noise suppression method that can improve the signal-to-noise ratio even when the transmission period is set below the reverberation time to reduce the required transmission time.

3. Proposed method

3.1. Transmission and reception model

In the proposed method, $s(t)$ is transmitted L times in period T_d . Let N_d denote the signal transmission period T_d converted to the number of samples. For simplicity, let T_d be constant multiples of T_s . In this case, $N_d = T_d/T_s$.

Let \mathbf{r} be the discrete data of the signal that is received and sampled and let N_r be its length. Let N_r be sufficiently large. Subsequently, \mathbf{r} can be expressed as follows:

$$\mathbf{r} = \sum_{l=1}^L \begin{bmatrix} O_{\Delta n-1} \\ O_{N_d(l-1)} \\ I_M \\ O_{N_d(L-l)} \\ O_{N_r-M-\Delta n+1-N_d(L-1)} \end{bmatrix} \mathbf{g} + \mathbf{n}_r. \quad (8)$$

From the explanation in Section 2.3.1, Δn is the index of the sample closest to $\Delta t - T$ and rounded to the nearest $(\Delta t - T)/T_s$. Let $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_M]^T$ be the sampling of Eq. (2) and let M be its length. T in the upper-right corner of the symbol indicates transposition. I_M is the identity matrix of degree M and O_i denotes the zero matrix whose size is $i \times M$. \mathbf{n}_r is a vector of the noise components. Each element is assumed to follow a white normal distribution with a mean of zero and a variance σ^2 . In the following, it is assumed that σ^2 is known from the prior measurements.

The noise suppression method detailed in Section 3.3 assumes that g_1 is contained in the first N_d elements of the vector to which the orthogonal projection matrix is applied. To satisfy this assumption, the signal was cut out of \mathbf{r} . Let \mathbf{y} denote the signal. Hereafter, the range from the first to the N_d -th sample of \mathbf{y} is called the decision interval.

The signal cutout procedure is described in detail. First, a discrete matched filter is applied to the received signal to determine the starting point of the signal cutout as $c_i = \mathbf{r}_{i,i+N-1}^T \mathbf{e} / \sqrt{N}$. Here, $\mathbf{r}_{i,i+N-1}$ is the cutout of \mathbf{r} from the i -th to the $(i + N - 1)$ -th.

Let $\Delta n^{(1)}$ be the index that exceeds threshold $\tau^{(1)}$ for the first time in $|c_i|$. Here, $\tau^{(1)}$ is set such that the false-alarm probability $P_{FA}^{(1)}$ is sufficiently small. Therefore, $\Delta n^{(1)} \geq \Delta n$. The starting point of the signal cutout is $\Delta n^{(1)} - N_d + 1$, which allows g_1 to be included in the

decision interval. In this case, the cut-out signal \mathbf{y} can be expressed as follows:

$$\mathbf{y} = A\mathbf{x} + \mathbf{n} \quad (9)$$

$$A = \sum_{l=1}^L \begin{bmatrix} O'_{N_d(l-1)} \\ I_{M'} \\ O'_{N_d(L-l)} \end{bmatrix} \quad (10)$$

$$\mathbf{x} = [\mathbf{0}_{z_1} \quad \mathbf{g}^T \quad \mathbf{0}_{z_2}]^T \quad (11)$$

\mathbf{x} denotes a vector with zeros before and after \mathbf{g} . M' denotes the length of \mathbf{x} . O'_i is the zero matrix whose size is $i \times M'$. z_1, z_2 are $N_d - 1 + \Delta n^{(1)} - \Delta n$ and $M' - M - N_d + 1 - \Delta n^{(1)} + \Delta n$, respectively. The cutout length is $M' + N_d(L - 1)$, which includes all transmitted signals and their reverberations. This equals the number of rows in A , and let R . \mathbf{n} is a part of \mathbf{n}_r that is cut out of \mathbf{r} with the signal component.

Here, we discuss how to set the parameter M' . The length M of \mathbf{g} contained in \mathbf{x} is determined by $\Theta_s(\cdot)$, $\Theta_r(\cdot)$, and $h(t)$, and varies depending on the sensor characteristics and environment. Therefore, M is assumed to be unknown. In addition, because the starting point of the cutout is $\Delta n^{(1)} - N_d + 1$, the beginning of \mathbf{x} can contain fewer than N_d zeros. Considering the above, M' should be set to a value with some margin.

3.2. Noise suppression in the decision interval

In the proposed method, the orthogonal projection matrix P_A

$$P_A = A(A^T A)^{-1} A^T \quad (12)$$

was applied to \mathbf{y} . To show that $(A^T A)^{-1}$ exists, it is sufficient to show that $A^T A$ has full rank, which is proved in the Appendix. The result \mathbf{y}' can be expressed as follows:

$$\begin{aligned} \mathbf{y}' &= P_A \mathbf{y} \\ &= A(A^T A)^{-1} A^T A \mathbf{x} + P_A \mathbf{n} \\ &= A \mathbf{x} + P_A \mathbf{n} \end{aligned} \quad (13)$$

This equation shows that signal component $A\mathbf{x}$ can be retained without corruption. For the noise component, let $P_A \mathbf{n}$ be \mathbf{n}' , then, the i -th element of \mathbf{n}' can be expressed as $n'_i = \mathbf{p}_i^T \mathbf{n}$ where \mathbf{p}_i refers to the i -th columns of P_A . Because \mathbf{n}' is a linear sum of white normal noise, n'_i follows a normal distribution, and its variance $(\sigma'_i)^2$ is $(\sigma'_i)^2 = \mathbf{p}_i^T \mathbf{p}_i \sigma^2$. This implies that $\mathbf{p}_i^T \mathbf{p}_i$ represents the noise suppression performance; if this value is less than one, a smaller threshold value can be set with the same false alarm probability. In this case, as discussed in section 2.3.1, the systematic error in the received time estimation can be reduced. The detailed process is as follows: The threshold value $\tau_i^{(2)}$ is calculated from a given false-alarm probability $P_{FA}^{(2)}$ and $(\sigma'_i)^2$. The absolute value of the discrete matched filter output $c'_i = (\mathbf{y}'_{i,i+N-1})^T \mathbf{e} / \sqrt{N}$ is calculated, and detection is performed using the threshold $\tau_i^{(2)}$. $\mathbf{y}'_{i,i+N-1}$ is the cutout of \mathbf{y}' from the i -th to the $(i + N - 1)$ -th.

Let $\Delta n^{(2)}$ be the index of \mathbf{y}' first detected by the threshold $\tau_2^{(2)}$, then, the estimate of Δn is expressed as $\Delta n^{(1)} + \Delta n^{(2)} - N_d$.

3.3. Relationship between transmission period, number of transmissions, and noise suppression performance

In this section, we first show that to study the noise suppression performance of P_A in the decision interval, it is sufficient to investigate the diagonal components of $(A^T A)^{-1}$. Next, we show that the noise suppression performance can be expressed in a simple form using L , N_d , and M' under the given conditions, and we discuss the relationship between each parameter and the noise suppression performance. In addition, the performance limits of the proposed method under these conditions are discussed.

3.3.1. Analysis of P_A

$\mathbf{p}_i^T \mathbf{p}_i$ representing the noise suppression performance, is the i -th diagonal component of $P_A^T P_A$. Here, P_A is a projective matrix and is symmetric. Hence, $P_A^T P_A = P_A P_A = P_A$. Therefore, $\mathbf{p}_i^T \mathbf{p}_i$ is identical to the i -th diagonal component of P_A . Here, A can be expressed as $A = (I_{1,N_d}^T \ K_0^T)^T$ where I_{1,N_d} is the first to N_d -th rows of the identity matrix of degree M' , and K_0 is the $(N_d + 1)$ -th and the later rows of A . Therefore, P_A can be written as:

$$\begin{aligned} P_A &= \begin{pmatrix} I_{1,N_d} \\ K_0 \end{pmatrix} (A^T A)^{-1} \begin{pmatrix} I_{1,N_d}^T & K_0^T \end{pmatrix} \\ &= \begin{pmatrix} (A^T A)_{1,N_d}^{-1} \\ K_0 (A^T A)^{-1} \end{pmatrix} \begin{pmatrix} I_{1,N_d}^T & K_0^T \end{pmatrix}. \end{aligned} \quad (14)$$

Thus, each element in the top-left N_d -by- N_d matrix of P_A and $(A^T A)^{-1}$ is identical. This indicates that the noise suppression performance in the decision interval can be evaluated by investigating the first to N_d -th diagonal components of $(A^T A)^{-1}$.

$A^T A$ is a symmetric Toeplitz matrix, as follows:

$$A^T A = \begin{pmatrix} v_1 & v_2 & v_3 & \cdots & \cdots & v_{M'} \\ v_2 & v_1 & v_2 & \ddots & & \vdots \\ v_3 & v_2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & v_2 & v_3 \\ \vdots & & \ddots & v_2 & v_1 & v_2 \\ v_{M'} & \cdots & \cdots & v_3 & v_2 & v_1 \end{pmatrix} \quad (15)$$

All the elements of this matrix are integers and are represented as follows:

$$v_m = \begin{cases} L - \frac{m-1}{N_d} & ((m-1) \bmod N_d) = 0 \wedge L > \frac{m-1}{N_d} \\ 0 & otherwise \end{cases}. \quad (16)$$

Note that the inverse matrix of symmetric Toeplitz matrix can be obtained by the Trench algorithm.

We then examine the diagonal components of inverse matrix $(A^T A)^{-1}$. Let each of these elements be defined as follows:

$$(A^T A)^{-1} = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,M'} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,M'} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M',1} & w_{M',2} & \cdots & w_{M',M'} \end{pmatrix}. \quad (17)$$

This matrix satisfies $(A^T A)(A^T A)^{-1} = I$.

We consider the first diagonal component $w_{1,1}$ of $(A^T A)^{-1}$. For the first column of $(A^T A)^{-1}$, we assume that $1 + N_d(i-1)$ ($i = 1, 2, \dots, \lceil \frac{M'}{N_d} \rceil$)-th rows are real numbers, and all other rows are zero. $\lceil \cdot \rceil$ denotes the ceiling function. Here, in m -th rows of $A^T A$, all columns except the $((m-1) \bmod N_d) + 1 + N_d(i-1)$ -th column are zero, and thus the inner product of the first column of $(A^T A)^{-1}$ and all rows except the $(1 + N_d(i-1))$ -th row of $A^T A$ are zero. From the above discussion, the product of $A^T A$ and the first column of $(A^T A)^{-1}$ can be expressed as:

$$B_Q^L \begin{pmatrix} w_{1,1} \\ w_{1+N_d,1} \\ w_{1+2N_d,1} \\ \vdots \\ w_{1+(Q-1)N_d,1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (18)$$

$$B_Q^L = \begin{pmatrix} v_1 & v_{(1+N_d)} & \cdots & \cdots & \cdots & v_{(1+(Q-1)N_d)} \\ v_{(1+N_d)} & v_1 & v_{(1+N_d)} & \cdots & \cdots & v_{(1+(Q-2)N_d)} \\ v_{(1+2N_d)} & v_{(1+N_d)} & v_1 & \cdots & \cdots & v_{(1+(Q-3)N_d)} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ v_{(1+(Q-1)N_d)} & \cdots & \cdots & v_{(1+2N_d)} & v_{(1+N_d)} & v_1 \end{pmatrix} \quad (19)$$

where $Q = \lceil M'/N_d \rceil$.

When $Q = 1$, $w_{1,1} = 1/L$ according to Eq. (18). Q becomes one only when the transmission period N_d is greater than M' , which corresponds to the averaging method.

When $Q > 1$, $w_{1,1}$ is the (1,1) component of $(B_Q^L)^{-1}$. Therefore, we investigate $(B_Q^L)^{-1}$. In the following, we impose the condition $L \geq Q - 1 \geq 1$ for L and Q . Then each element of B_Q^L can be expressed as follows:

$$B_Q^L = \begin{pmatrix} L & L-1 & \cdots & \cdots & L-Q+1 \\ L-1 & L & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & L & L-1 \\ L-Q+1 & \cdots & \cdots & L-1 & L \end{pmatrix} \quad (20)$$

Theorem 1. When $L \geq Q - 1 \geq 1$, the inverse matrix of B_Q^L is as follows.

$$(B_Q^L)^{-1} = \begin{pmatrix} \beta_1 & -0.5 & 0 & \cdots & \cdots & \cdots & 0 & \beta_2 \\ -0.5 & 1 & -0.5 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -0.5 & 1 & -0.5 & 0 & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & 0 & -0.5 & 1 & -0.5 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & -0.5 & 1 & -0.5 \\ \beta_2 & 0 & \cdots & \cdots & \cdots & 0 & -0.5 & \beta_1 \end{pmatrix}. \quad (21)$$

$$\beta_1 = \frac{2L - Q + 2}{4L - 2Q + 2} \quad (22)$$

$$\beta_2 = \frac{1}{4L - 2Q + 2}. \quad (23)$$

Proof. We show that Eq. (21) is $B_Q^L(B_Q^L)^{-1} = I$. For the first column of $(B_Q^L)^{-1}$, the inner product of the first row of B_Q^L is

$$\frac{(2L - Q + 2)L}{4L - 2Q + 2} - 0.5(L - 1) + \frac{L - Q + 1}{4L - 2Q + 2} = 1. \quad (24)$$

The inner product of B_Q^L with the i ($i = 2, 3, \dots, Q$)-th row is:

$$\frac{(L - i + 1)(2L - Q + 2)}{4L - 2Q + 2} - 0.5(L - i + 2) + \frac{L - Q + i}{4L - 2Q + 2} = 0. \quad (25)$$

For the i ($i = 2, 3, \dots, Q - 1$)-th column of $(B_Q^L)^{-1}$, the inner product of the i -th row of B_Q^L is

$$-0.5(L - 1) + L - 0.5(L - 1) = 1. \quad (26)$$

For the inner product of the i ($i = 2, 3, \dots, Q - 1$)-th column of $(B_Q^L)^{-1}$ and j ($j \neq i$)-th row of B_Q^L , if $i > j$,

$$-0.5(L - 1 + i - j) + (L - 2 + i - j) - 0.5(L - 3 + i - j) = 0. \quad (27)$$

If $i < j$, then

$$-0.5(L - 1 + i - j) + (L + i - j) - 0.5(L + 1 + i - j) = 0. \quad (28)$$

We considered the Q -th column of $(B_Q^L)^{-1}$. Each of B_Q^L and $(B_Q^L)^{-1}$ is equal to the matrix that inverts the row and column orders. Therefore, by the same reasoning as in the first column, the inner product of the Q -th column of $(B_Q^L)^{-1}$ with the Q -th row of B_Q^L is 1, and the other rows of B_Q^L are 0, which concludes the proof. \square

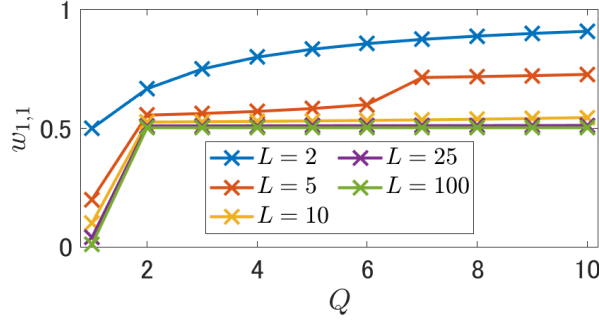


Figure 1: Examples of noise suppression performance

This indicates that the noise suppression performance $w_{1,1}$ can be expressed in the form β_1 (Eq. (22)).

The above discussion is for the first column of $(A^T A)^{-1}$, but holds for the second and subsequent columns, and the value of each element can be calculated using $(B_Q^L)^{-1}$. Therefore, we can represent the noise suppression performance in the decision interval as β_1 (Eq. (22)).

3.3.2. Relationship between L , Q , and noise suppression performance

Figure 1 shows the results of calculating the (1,1) components of $(B_Q^L)^{-1}$ for L and Q . First, as mentioned in the previous section, $Q = 1$ corresponds to the averaging method, and the noise suppression performance is $1/L$. For Q , the noise suppression performance decreased as Q increased. For L , the noise suppression performance increases as L increases but appears to converge to 0.5, when Q is not one.

In the following, we show that the above properties are maintained under the condition $L \geq Q - 1 \geq 1$. We will discuss the case $L < Q - 1$ in future work.

First, we investigate the noise suppression performance when Q increases. If $L \geq Q - 1 \geq 1$ is satisfied, the noise suppression performance can be expressed by Eq. (22), and thus the difference between the noise suppression performance when Q' and $Q' + 1$ ($L \geq Q' + 1 > Q' \geq Q - 1 \geq 1$) is

$$\frac{2L - (Q' + 1) + 2}{4L - 2(Q' + 1) + 2} - \frac{2L - Q' + 2}{4L - 2Q' + 2} \quad (29)$$

This expression can be simplified as

$$\frac{1}{2(2L - Q')(2L - Q' + 1)}. \quad (30)$$

The equation is positive because $L \geq Q' \geq 1$. In other words, as Q increased, the noise suppression performance decreased.

The value of Q increased as the transmission period N_d decreased. Therefore, the noise suppression performance decreased as the transmission period was shortened.

Next, we investigate the noise suppression performance as the number of transmissions L increases. The difference between β_1 when the number of transmissions L' ($L' \geq L \geq Q - 1 \geq 1$) and $L' + 1$ are

$$\frac{2(L' + 1) - Q + 2}{4(L' + 1) - 2Q + 2} - \frac{2L' - Q + 2}{4L' - 2Q + 2} = \frac{-1}{(2L' - Q + 1)(2L' - Q + 3)}. \quad (31)$$

The equation is negative because $L' \geq Q \geq 1$. Therefore, the noise suppression performance improves when the number of transmissions is increased. In addition, from Eq. (22), it can be observed that β_1 converges to 0.5 as L is increased. To summarize the above, it can be said that under the condition that $L \geq Q - 1 \geq 1$, the noise suppression performance can be improved by increasing the number of transmissions L , but there is an lower limit of 0.5.

The lower limit of β_1 0.5 is the same as that of the noise suppression performance of the averaging method when the transmission period is set to more than M and the number of transmissions is set to two, as described in Section 2.3.2. In other words, when the proposed method was used with a transmission period of M or less, its noise suppression performance was less than or equal to that of the averaging method under the aforementioned conditions. In contrast, the proposed method has the advantage of reducing the time required for the signal transmission. Specific numerical examples and performance comparisons are presented in Section 4.3.

3.4. Relationship between the starting point of the cutout and the error caused by the projection

In this section, we consider the case in which \mathbf{y} does not contain a part of \mathbf{g} . There are two cases in which the cutout is made either too early or too late, and we discuss them sequentially.

If the starting point of the cutout is too early, and \mathbf{y} does not contain the last i components of \mathbf{g} , \mathbf{y} is represented as follows:

$$\mathbf{y} = \begin{pmatrix} O'_i \\ A_{1,R-i} \end{pmatrix} \mathbf{x} + \mathbf{n} \quad (32)$$

$A_{1,R-i}$ is the first to $(R - i)$ -th rows of A . Let \mathbf{a}_i be the i -th column of matrix A , then, the matrix on the right-hand side of Eq. (32) can be expressed as

$$(\mathbf{a}_{1+i} \quad \mathbf{a}_{2+i} \quad \cdots \quad \mathbf{a}_{M'} \quad \mathbf{a}_{M'+1} \quad \cdots \quad \mathbf{a}_{M'+i}) \quad (33)$$

Here, \mathbf{a}_j ($j > M'$) is a vector consisting of a i -by-1 zero vector, and the first to $(M' - i)$ -th rows of \mathbf{a}_{j-i} . The product of this and A is:

$$A^T \begin{pmatrix} O'_i \\ A_{1,R-i} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_{M'}^T \end{pmatrix} (\mathbf{a}_{1+i} \quad \mathbf{a}_{2+i} \quad \cdots \quad \mathbf{a}_{M'+i}). \quad (34)$$

This indicates that the first to $(M' - i)$ -th columns are identical to the $(i - 1)$ -th to M' -th columns of $A^T A$. Thus, \mathbf{y}' can be represented as:

$$\begin{aligned}\mathbf{y}' &= A(A^T A)^{-1} A^T \begin{pmatrix} O'_i \\ A_{1,R-i} \end{pmatrix} \mathbf{x} + P_A \mathbf{n} \\ &= A \begin{pmatrix} O''_i & K_1 \\ I_{M'-i} & K_2 \end{pmatrix} \mathbf{x} + P_A \mathbf{n}.\end{aligned}\quad (35)$$

O''_i is the i -by- $(M' - i)$ zero matrix. K_1 and K_2 are i -by- i and $(M' - i)$ -by- i matrices, respectively. From Eq. (34), K_1 and K_2 are not zero matrices. Therefore, it can be observed that \mathbf{y}' contains an error that is the product of the last i components of \mathbf{x} and K_1 , K_2 . Note that this error does not occur when $z_2 \geq i$ because the last z_2 components of \mathbf{x} are zero (Eq. (11)).

If the starting point of the cutout is too late, and \mathbf{y} does not contain the first i components of \mathbf{g} , then \mathbf{y} is represented as follows:

$$\mathbf{y} = \begin{pmatrix} A_{i+1,R} \\ O'_i \end{pmatrix} \mathbf{x} + \mathbf{n}.\quad (36)$$

The matrix on the right-hand side of this equation can be written as

$$(\mathbf{a}_{1-i} \quad \mathbf{a}_{2-i} \quad \cdots \quad \mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{M'-i}).\quad (37)$$

Here, \mathbf{a}_j ($j < 1$) is a vector comprising the i -th to M' -th rows of \mathbf{a}_{j+i} and i -by-1 zero vectors. The product of this and A is:

$$A^T \begin{pmatrix} A_{i+1,R} \\ O'_i \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_{M'}^T \end{pmatrix} (\mathbf{a}_{1-i} \quad \mathbf{a}_{2-i} \quad \cdots \quad \mathbf{a}_{M'-i}).\quad (38)$$

It can be observed that $(i + 1)$ -th to M' -th columns of this matrix are identical to first to $(M' - i)$ -th columns of $A^T A$. Thus, \mathbf{y}' can be represented as:

$$\begin{aligned}\mathbf{y}' &= A(A^T A)^{-1} A^T \begin{pmatrix} A_{i+1,R} \\ O'_i \end{pmatrix} \mathbf{x} + P_A \mathbf{n} \\ &= A \begin{pmatrix} K_3 & I_{M'-i} \\ K_4 & O''_i \end{pmatrix} \mathbf{x} + P_A \mathbf{n}.\end{aligned}\quad (39)$$

K_3 and K_4 are $(M' - i)$ -by- i and i -by- i matrices, respectively. From Eq. (38), K_3 and K_4 are not zero matrices. Therefore, this indicates that \mathbf{y}' contains errors owing to the product of the first i components of \mathbf{x} and K_3 , K_4 .

4. Numerical Experiment

In this section, numerical experiments were conducted to verify the effectiveness of the proposed method. In Experiment 1, we compared the performances of the conventional and proposed methods when the number of delay waves was varied. Experiment 2 was a comparative evaluation of the required transmission time and noise suppression performance.

Table 1
Parameter

Sampling frequency $f_s = 1/T_s$	48 kHz
Frequency of transmission signal f	18 kHz
Length of signal T	0.001 s
Signal-to-noise ratio	20 dB
False alarm probability $P_{FA}^{(1)}$	10^{-10}
False alarm probability $P_{FA}^{(2)}$	10^{-6}
Transmission period T_d	0.04 s
M'	$0.13 \times f_s$
Trial count	1000

4.1. Experimental setting

This section describes the experimental setup common to Experiments 1 and 2. The list of parameters is presented in Table 1. If the parameters are not specified, the values listed in Table 1 are used.

The transmitted signal was a sine wave in Eq. (2). The characteristics of the speaker and microphone were ignored here to evaluate only the influence of the delay waves. The number of samples in the received signal r was set to $0.54 \times f_s = 25,920$. For direct waves, the received time is not necessarily a constant multiple of N_s when sampled. To simulate this, a uniform random number in the range of 0.15 s to $0.15 + N_s$ s was generated for each trial and was set as the received time. The amplitude of the direct wave was set as 1.

For the simulated delayed waves, the amplitudes of each delayed wave were uniformly random numbers, ranging from 0 to 0.8. The received time of each delayed wave was set to a uniform random number in a range from the received time of the direct wave to 0.08 s.

The threshold $\tau^{(1)}$ was given values calculated from the false alarm probabilities $P_{FA}^{(1)}$ and the noise variance. M' is set to a value of 0.13 s with a margin with respect to M based on the discussion in Section 3.1.

4.2. Experiment 1: Comparison with conventional methods

In this section, we compare the performance of each method when the number of delay waves is varied. In the proposed method, the number of signal transmissions is set to four. The following three methods were evaluated using conventional methods.

The first is the maximum peak method described in Section 2. Note that for transmission, the signal length $s(t)$ is set to $T = 0.004$ s, and the number of transmissions is set to one. The reason for this was to make the energy identical to that of the signal transmitted using the proposed method.

The second is the leading-edge method described in Section 2. Here, τ in Section 2.3.1 is calculated from the false-alarm probabilities $P_{FA}^{(2)}$ and noise variance.

The third is the leading-edge method using the threshold $\tau_i^{(2)}$. Hereafter, this method will be referred to as the low-threshold leading-edge method. The difference between this method and the proposed method is that detection is performed using the threshold $\tau_i^{(2)}$ without the noise

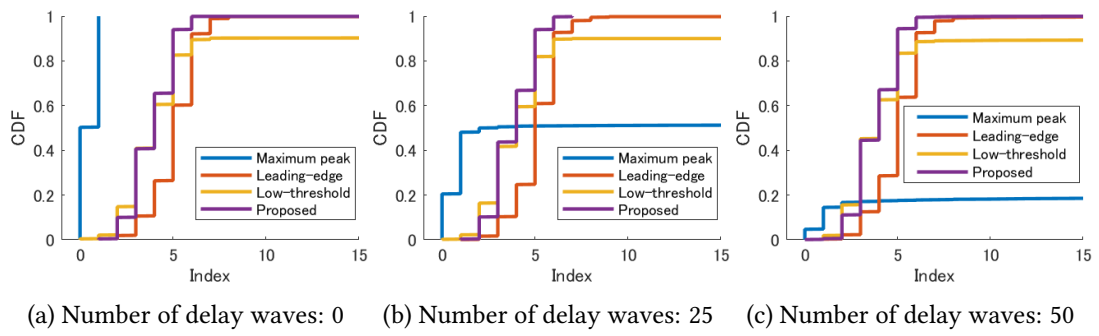


Figure 2: Results of Experiment 1

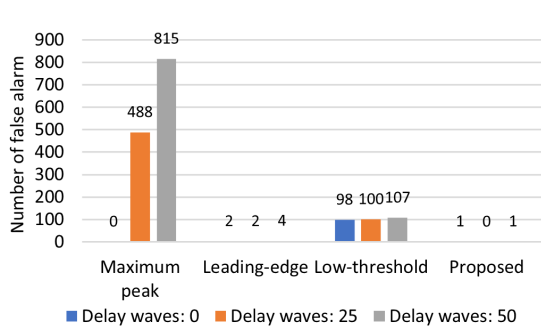


Figure 3: Number of false alarm

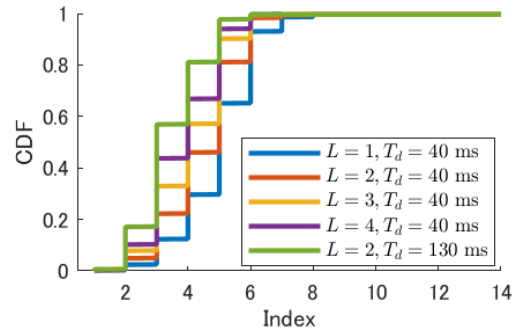


Figure 4: Results of Experiment 2

suppression process described in Section 3. This method was evaluated to show that the number of false alarms increased when the threshold was lowered, without the noise suppression process of the proposed method.

The evaluation results of each method are shown in Figures 2a, 2b, and 2c. In these figures, the horizontal axis shows the error of the estimated index and the vertical axis shows the cumulative distribution function (CDF). The number of false alarms for each condition and method is shown in Figure 3. False alarms were defined as errors greater than 15.

Figure 2a shows that, in the absence of delay waves, the maximum peak method yields the best performance, as described in Section 2. However, in the maximum peak method, the false alarm rate increases as the number of delay waves increases (Figures 2b, 2c, and 3).

For the methods based on leading-edge detection (leading-edge, low-threshold leading-edge, and proposed methods), the performance was not significantly degraded even when the number of delay waves increased, as shown in Figures 2a, 2b, 2c, and 3. Moreover, Figures 2a, 2b, and 2c show that the proposed method reduces the error around the true value compared to the leading-edge method. Comparing the low-threshold leading-edge method and the proposed method, although the errors around the true value are not significantly different, the number of false alarms increases in the low-threshold leading-edge method compared to the proposed method, as shown in Figure 3.

4.3. Experiment 2: Required transmission time and noise suppression performance

This section presents a comparative study of the relationship between the required transmission time and noise suppression performance. In this experiment, the number of delay waves is set to 25.

For the proposed method, the number of transmissions was set to one, two, three, and four. Their required transmission times are 0.13, 0.17, 0.21, and 0.25 s, respectively. For the averaging method, because M is unknown, the transmission period T_d is set to $M' \times T_s = 0.13$ s, and the number of transmissions is set to two. In this case, the required transmission time was 0.26 s. As described in Section 2.3.2, under these conditions, the noise suppression performance of the averaging method was 0.5, which is identical to the performance limit of the proposed method shown in the 3.3.2 section.

The results of the experiment are shown in Figure. 4. As aforementioned, for the proposed method, the required transmission time increases as the number of transmissions increases. However, the performance of the estimation of the received time improves as the number of transmissions increases. The averaging method had the best performance in estimating the received time, but the required transmission time was the largest. These results confirm that the proposed method can provide noise suppression with reduced transmission time if the required noise suppression performance is greater than 0.5.

5. Conclusion

In this study, we proposed a noise suppression method that can be applied to multiple signal transmissions without waiting for the disappearance of reverberation. Numerical experiments were conducted to compare the effectiveness of the proposed method with that of conventional methods in estimating the received time under multiple reverberation conditions.

The relationship between noise suppression performance and L and Q is discussed under the conditions $Q = 1$ or $L \geq Q - 1 \geq 1$. The case in which $L < Q - 1$ will be discussed in our future work.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers 20K23313 and 21K17797.

References

- [1] J. Xiao, Z. Zhou, Y. Yi, L. M. Ni, A survey on wireless indoor localization from the device perspective, *ACM Comput. Surv.* 49 (2016).
- [2] M. A. Richards, J. A. Scheer, W. A. Holm (Eds.), *Principles of Modern Radar: Basic principles*, 1st ed., Scitech publishing, 2010.
- [3] M. O. Khyam, S. S. Ge, X. Li, M. Pickering, Orthogonal chirp-based ultrasonic positioning, *Sensors* 17 (2017).

- [4] 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.
- [5] M. Hazas, A. Hopper, Broadband ultrasonic location systems for improved indoor positioning, *IEEE Trans. on Mobile Computing* 5 (2006) 536–547.
- [6] F. Höflinger, R. Zhang, J. Hoppe, A. Bannoura, L. M. Reindl, J. Wendeberg, M. Bühner, C. Schindelbauer, Acoustic self-calibrating system for indoor smartphone tracking (AS-SIST), in: *IPIN 2012*, 2012, pp. 1–9.
- [7] P. Lazik, A. Rowe, Indoor pseudo-ranging of mobile devices using ultrasonic chirps, in: *ACM Sensys 2012*, 2012, pp. 99–112.
- [8] M. Nakamura, H. Kameda, Indoor localization using multiple stereo speakers for smartphones, in: *IPIN 2019*, volume 2498 of *CEUR Workshop Proceedings*, 2019, pp. 235–242.
- [9] 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.
- [10] T. Suzaki, M. Nakamura, H. Murakami, H. Watanabe, H. Hashizume, M. Sugimoto, Expanding the positioning area for acoustic localization using cots mobile devices, in: *MobiQuitous2021*, 2021, pp. 422–437.

Appendix

We show that $A^T A$ is full rank. Because $A^T A$ is a Gram matrix, it is sufficient to show that all the column vectors of A are linearly independent. First, for the i ($1 \leq i \leq N_d$)-th column vector, the i -th row is one, but all i -th rows of column vectors except i -th column vector are zero, and thus they are linearly independent.

For the i ($i > N_d$)-th column vector, let Λ be the index set of the column vector whose i -th row is one, except for the i -th column as follows:

$$\left\{ ((i-1) \bmod N_d) + 1 + N_d(m-1) \mid m = 1, 2, \dots, \left\lceil \frac{i}{N_d} \right\rceil - 1 \right\} \quad (40)$$

In the following, let $\lambda_1, \lambda_2, \lambda_3 \in \Lambda$: For λ_1 -th rows, the value is zero in i -th columns but one only in λ_2 ($\lambda_1 \geq \lambda_2$)-th columns. Therefore, if the i -th column vector is linearly dependent, then the linear sum of $\lceil i/N_d \rceil - 1$ column vectors must satisfy the $\lceil i/N_d \rceil$ equality constraint, where λ_3 -th rows are zero and i -th row is one. Because this is impossible, it can be said that all i -th column vectors are linearly independent. Thus, $(A^T A)$ is shown to have full rank.