

Underwater acoustic localization based on IR-GCC-PHAT in reverberant environments

Shingo Yoshizawa

Kitami Institute of Technology, Kitami, Hokkaido 0908507, Japan

Received: June 23, 2020. Revised: February 9, 2021. Accepted: March 8, 2021. Published: March 12, 2021.

Abstract—This paper presents a method of underwater acoustic localization in reverberant environments. Time difference of arrival (TDOA) measurement algorithm is a key technology for estimating direction of arrival (DOA) of an underwater sound source. In strong multipath interference, the pseudo-peaks in a correlation function disturb the detection of a correct time position and lead to a large TDOA measurement error. The proposed algorithm computes a time difference by taking cross-correlation of two impulse responses and improves robustness to multipath interference. The comparison of TDOA algorithms is done by evaluating the position accuracy of underwater sound source in both simulation and experiment.

Keywords—direction of arrival, multipath interference, underwater acoustic localization, time difference of arrival

I. INTRODUCTION

UNDERWATER acoustic localization plays an important role in determining the positions of underwater vehicles such as remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) because the global positioning system (GPS) signal does not propagate underwater. In ultra-short baseline (USBL) systems, a receiver unit estimates the direction of arrival (DOA) of a sound source. The methods of DOA estimation for acoustic signals are classified into beamforming [1][2] and time difference of arrival (TDOA) measurement [3]. We focus on the latter. TDOA measurement is suitable for estimating the DOA with a simple array having two receiving elements [4]. For the receiving elements underwater, hydrophones are used instead of microphones.

The TDOA measurement algorithm computes an arrival time difference between received signals by a correlation function. One of the most commonly used algorithm is the generalized cross-correlation with phase transform (GCC-PHAT) [3]. GCC-PHAT calculates the cross-spectrum with normalizing magnitude of each frequency at 1.

GCC-PHAT is strong in reverberation, however, is easily affected by wideband additive noise [5]. Methods of combining PHAT with noise countermeasures have been investigated in

recent studies. The impact of pre-filtering method on GCC-PHAT in various types of noise environments was addressed in [6]. Grondin proposed a noise masking method in [5]. The received signal is separated into voice and non-voice sections and the weighting factor of 0 is applied in the non-voice section. This noise masking method is applied to both the time and frequency axes. Wang optimized the weighting factor by using the result of learning ambient noise signal with deep neural network (DNN) [7]. Lee presented a method of calculating the weighting factor based on the coherent-to-diffuse power ratio (CDR) [4].

Since the above-mentioned studies target applications in voice operation of equipment and voice dialogue, the unknown sound source signal is assumed. On the other hand, the underwater acoustic positioning system can use an artificially generated signal such as a pseudo noise (PN) code sequence as a sound source. The source signal can be designed arbitrarily by developers. When the sound source signal is known, it is possible to take the cross-correlation between the received signal and the reference signal given by a replica of the transmitted signal. This algorithm is known as the matched filter (MF) [8], which is used for DOA estimation in underwater acoustics.

MF calculates an arrival time difference by detecting the highest peak position in the cross-correlation function in the two channels [9]. By taking the cross-correlation between the received signal and the reference signal, uncorrelated noise components can be suppressed. MF is superior to GCC-PHAT with respect to noise robustness.

DOA estimation is strongly influenced by the reflection of sound waves. In underwater acoustics, there are many reflected waves caused by the reflection on water surface, bottom and obstacles. It causes the pseudo-peaks are seen in the correlation function, which can be known as the phenomenon of multipath interference. If the wrong time position is measured from the pseudo peaks, the TDOA measurement presents a significant error in the estimated angle of DOA, known as an outlier.

As a countermeasure against the above-mentioned outliers, methods of detecting and removing outliers by applying a Kalman filter or Bayesian estimation to the TDOA measurement data series have been presented in [9-10]. Another method detects and removes outliers in combination with the

position information of the inertial navigation system (INS) mounted on the AUV [11]. However, these studies do not refer to improve the TDOA measurement algorithm.

We present an improved TDOA measurement algorithm for underwater acoustic localization in reverberant environments. The proposed method is called the impulse response based GCC-PHAT (IR-GCC-PHAT), which computes a time difference by taking a cross-correlation between two impulse responses. IR-GCC-PHAT prevents false detection caused by the pseudo peaks in correlation function and has the robustness of noise interference as well as MF. The comparison of TDOA algorithms is carried out by evaluating the accuracy of underwater acoustic localization in simulation and experiment.

II. TDOA MEASUREMENT ALGORITHM

A. Signal Model

Two received signals $y_1(k)$ and $y_2(k)$ can be modeled by using a transmitted signal $x(k)$ and impulse responses $h_1(k)$ and $h_2(k)$ that express a propagation path from a transmitter to a receiver as

$$\begin{aligned} y_1(k) &= h_1(k) * x(k) + n_1(k) \\ y_2(k) &= h_2(k) * x(k) + n_2(k), \end{aligned} \quad (1)$$

where k indicates a discrete time index and $*$ shows a convolution operation. $n_1(k)$ and $n_2(k)$ are noise component uncorrelated with the transmitted signal. The received signals can be expressed in frequency domain as

$$\begin{aligned} Y_1(l) &= \text{DFT}_N[y_1(k)] = H_1(l)X(l) + N_1(l) \\ Y_2(l) &= \text{DFT}_N[y_2(k)] = H_2(l)X(l) + N_2(l). \end{aligned} \quad (2)$$

$\text{DFT}_N[\cdot]$ indicates the discrete Fourier transform for N samples and l denotes a discrete frequency index.

When the arrival time difference is obtained by the TDOA algorithm, an angle of DOA is computed as

$$\theta = \arcsin\left(\frac{c\tau}{d}\right), \quad (3)$$

where the arrival time difference for the two received signals is expressed by τ . An array space between receiver elements is given by d and a sound velocity is given by c .

B. GCC-PHAT

The cross correlation (CC) is a basic approach to measure the arrival time difference [12]. The CC correlation function is given by

$$\Phi_{\text{CC}}(k) = \text{IDFT}_N[Y_1(l)Y_2^*(l)], \quad (4)$$

where $\text{IDFT}_N[\cdot]$ indicates the inverse discrete Fourier transform for N samples. GCC-PHAT normalizes the magnitude of each frequency to 1. The GCC-PHAT correlation function is modified by the following equation:

$$\Phi_{\text{GCC-PHAT}}(k) = \text{IDFT}_N\left[\frac{Y_1(l)Y_2^*(l)}{|Y_1(l)Y_2^*(l)|}\right]. \quad (5)$$

The time difference is detected by the highest peak detection as

$$\tau_{\text{GCC-PHAT}} = \underset{k}{\text{argmax}} \Phi_{\text{GCC-PHAT}}(k). \quad (6)$$

GCC-PHAT can improve temporal resolution in the correlation function by whitening the spectrum and is said to be resistant to reverberation [3].

C. MF

MF is a method for detecting the arrival time difference from the correlation function between the received signal and the reference signal. The cross spectrum calculation of MF is computed as

$$\begin{aligned} \Phi_{\text{MF1}}(k) &= \text{IDFT}_N[Y_1(l)X^*(l)] \\ \Phi_{\text{MF2}}(k) &= \text{IDFT}_N[Y_2(l)X^*(l)] \end{aligned} \quad (7)$$

The time difference is detected from the two time positions with the highest peaks, expressed as

$$\tau_{\text{MF}} = \underset{k}{\text{argmax}} |\Phi_{\text{MF1}}(k)| - \underset{k}{\text{argmax}} |\Phi_{\text{MF2}}(k)|. \quad (8)$$

When the spectrum of transmitted signal is given by $|X(l)| = 1$ and the assumptions of $|H_1(l)| \gg |N_1(l)|$ and $|H_2(l)| \gg |N_2(l)|$ are used, (7) is transformed as

$$\begin{aligned} \Phi_{\text{MF1}}(k) &= \text{IDFT}_N[H_1(l)|X(l)|^2 + N_1(l)X^*(l)] \\ &= \text{IDFT}_N[H_1(l) + N_1(l)X^*(l)] \\ &\approx \text{IDFT}_N[H_1(l)] \\ &\approx h_1(k) \\ \Phi_{\text{MF2}}(k) &\approx h_2(k). \end{aligned} \quad (9)$$

The above equations indicate that MF estimates the impulse responses when the influence of noise is trivial. (8) can be approximated as

$$\tau_{\text{MF}} \approx \underset{k}{\text{argmax}} |h_1(k)| - \underset{k}{\text{argmax}} |h_2(k)|. \quad (10)$$

III. PROPOSED ALGORITHM

A. IR-GCC-PHAT

IR-GCC-PHAT directly computes the two impulse responses by the frequency-domain division that is expressed as

$$\begin{aligned} h'_1(k) &= \text{IDFT}_N[Y_1(l)/X(l)] \\ &= \text{IDFT}_N[H_1(l) + N_1(l)/X(l)] \\ h'_2(k) &= \text{IDFT}_N[Y_2(l)/X(l)] \\ &= \text{IDFT}_N[H_2(l) + N_2(l)/X(l)]. \end{aligned} \quad (11)$$

This procedure is known as the cross spectral method [13]. The term of $N_1(l)/X(l)$ is transformed as

$$\frac{N_1(l)}{X(l)} = \frac{N_1(l)X^*(l)}{|X(l)|^2}. \quad (12)$$

(9) and (11) show the same characteristic when $|X(l)| = 1$. The time difference is detected by the cross correlation after taking absolute values for the two impulse responses:

$$G_1(l) = \text{DFT}_N[|h'_1(k)|] \quad (13)$$

$$G_2(l) = \text{DFT}_N[|h'_2(k)|] \quad (14)$$

$$\Phi_{\text{IR}}(k) = \text{IDFT}_N \begin{bmatrix} G_1(l)G_2^*(l) \\ G_1(l)G_2^*(l) \end{bmatrix} \quad (14)$$

$$\tau_{\text{IR}} = \underset{k}{\text{argmax}} \Phi_{\text{IR}}(k). \quad (15)$$

The whole procedure of IR-GCC-PHAT is illustrated in Fig. 1. It corresponds to the combination of MF and GCC-PHAT.

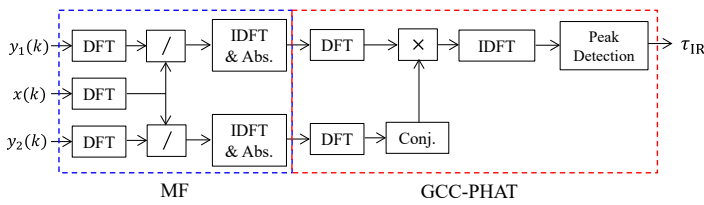


Fig. 1 Procedure of IR-GCC-PHAT algorithm.

B. Robustness to Reverberant Environments

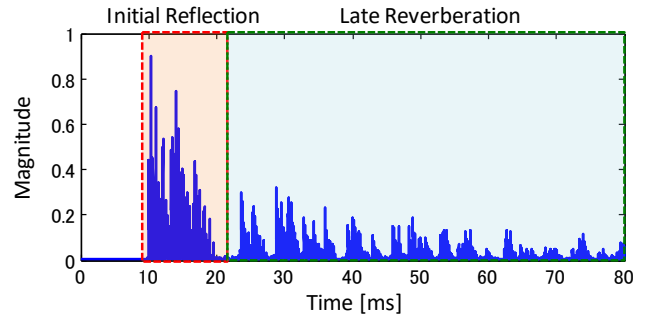
Figure 2 shows examples of the impulse responses $|h_1(k)|$, $|h_2(k)|$ for the two receiver elements, which are generated by the sound wave propagation simulator. From this graph, we can read the magnitude and time of arrival of for direct and reflected waves.

The reflected waves are classified into initial reflection and late reverberation from the viewpoint of reverberation. In the initial reflection, the sound wave is reflected several times. In the late reverberation, the sound wave is reflected many times and the direction in which the sound wave moves and the signal phase become random and exponentially attenuated.

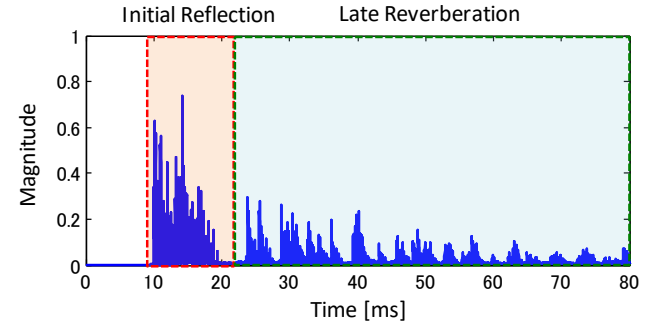
Figure 3 shows acoustic paths for the direct and reflected waves from the sound source (TX) to the receiver elements (RX1 and RX2). Although there are a lot of reflected waves as shown in Fig.2, we pick up only several acoustic paths in the schematic of Fig. 3.

We explain the influence of the reflected waves belonging to the initial reflection (W_1 , W_2 , W'_1 , and W'_2 in Fig.3). When the path length of one reflected wave is the same as that of and the other reflected wave, their peaks are combined at the same time position on the impulse response. The magnitude of the combined peak is comparable to or higher than the peak of the direct wave. This multipath interference directly affects the accuracy of the TDOA algorithms.

In computing the correlation function of TDOA algorithms, the pseudo peaks caused by the initial reflection disrupt the correct detection of the arrival time difference. Figure 4 shows the detection of the arrival time difference when the pseudo peaks occur. The true time difference is given by τ^{true} in the figure.



(a) Impulse response $|h_1(k)|$



(b) Impulse response $|h_2(k)|$

Fig. 2 Example of impulse response.

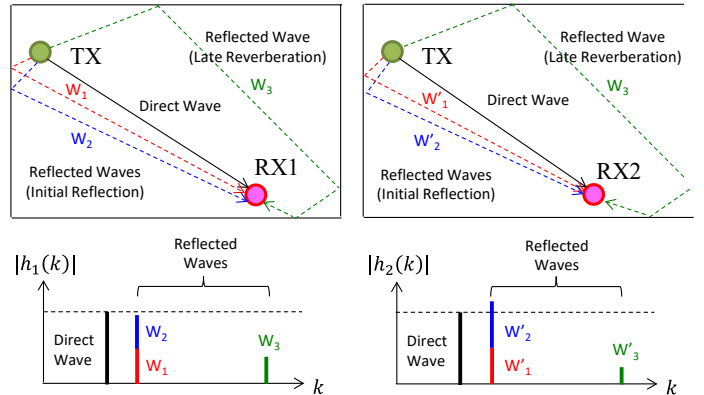


Fig. 3 Acoustic paths for direct and reflected waves.

In the MF algorithm, the time difference is detected from the two time positions with the highest peaks. Although the targeted peaks correspond to P_1 and P'_1 derived from the direct waves, P_1 and P'_2 derived from the reflected wave are detected in this example. These peaks cannot detect the correct time difference ($\tau_{\text{MF}} \neq \tau^{\text{true}}$). This is because the magnitude of the reflected wave is greater than that of the direct wave due to multipath interference.

IR-GCC-PHAT treats the peaks derived from the direct and reflected wave as one group. When comparing the time positions of P_1 to P_3 and P'_1 to P'_3 , the time difference of their groups is near to the true time difference. When the direct and reflected waves arrive at the receiver in the same direction, the time differences between the same types of waves are almost the same. The cross correlation of the impulse responses emphasizes this group time difference ($\tau_{\text{IR}} \approx \tau^{\text{true}}$).

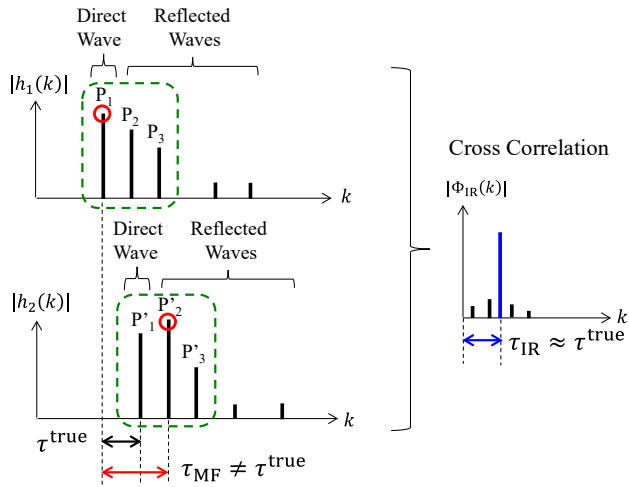


Fig. 4 Detection of the arrival time difference when the pseudo peaks occur.

For late reverberation, the reflected waves belonging to late reverberation are marked by W_3 and W'_3 in Fig.3. Their reflected waves are considerably attenuated due to the long path length. The magnitude of the reflected waves is much smaller than that of the direct wave. The false detection of time difference by the MF algorithm (as explained in Fig. 4) would not occur as for the late reverberation. It should be noted that there are a lot of reflected waves in the late reverberation as well, as observed in Fig. 2. These reflected waves could be treated as a noise component because they have a lower correlation (compared to the initial reflection) with the amplitude and phase between the received signals. As for the robustness to noise interference, MF and IR-GCC-PHAT are stronger than GCC-PHAT by taking the cross-correlation between the received signal and the reference signal.

IV. POSITION ESTIMATION

This paper focuses on the accuracy comparison of the TDOA algorithms and evaluates the following simple two-dimensional localization. When the coordinates of one of the receiver elements are represented by $\mathbf{p}_r = [x_r, y_r]$, the transmitter position $\mathbf{p}_t = [x_t, y_t]$ is computed by the estimated angle θ :

$$\begin{aligned} x_t &= x_r + D \cos \theta \\ y_t &= y_r + D \sin \theta. \end{aligned} \quad (16)$$

The distance between the receiver and the transmitter is given by D .

For the distance measurement, we use the time synchronization scheme that a common pulse signal is entered into the transmitter and receiver units, as shown in Fig. 5. The transmitter unit starts transmitting a signal at the timing of the pulse signal. The receiver unit measures the time delay T as a sound propagates underwater using the timing of the pulse signal. The distance can be measured by $D = cT$. This time synchronization is used in the pool localization experiment described in Section VI.

The time delay is simply computed by taking the cross correlation between the received signal and the reference signal,

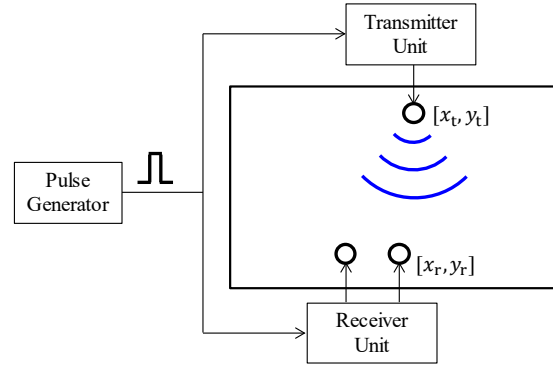


Fig. 5 Time synchronization scheme for distance measurement.

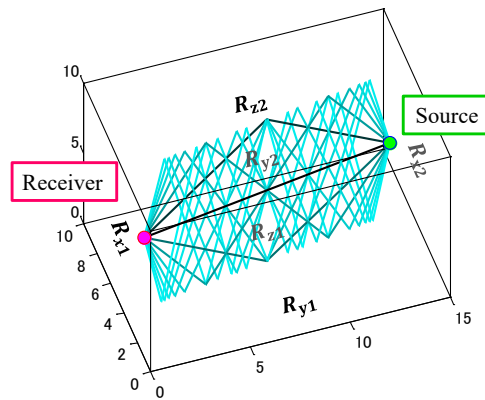


Fig. 6 Sound wave propagation simulator.

which is commonly applied regardless of the TDOA algorithms. The time difference τ should be exactly in the order of several microseconds for DOA estimation. On the other hand, the time delay T requires accuracy on the order of several hundred microseconds and is not as severe as the time difference.

V. SIMULATION

A. Acoustic Simulator

In order to reproduce the sound wave reflection on a simulation, a sound wave propagation simulator that tracks sound waves by the mirror image method [14] is used. The sound wave propagation simulator generates an impulse response from the path through which the sound wave reaches a receiver from a sound source after a user specifies the sound field space, sound pressure reflectance, sound source and receiver positions [15].

Figure 6 shows an example of sound ray tracing analysis using a sound wave propagation simulator. The parameters of $\mathbf{R} = [R_{x1}, R_{x2}, R_{y1}, R_{y2}, R_{z1}, R_{z2}]$ indicate the sound pressure reflectance on the six surfaces that are the wall surfaces in each axis direction. In this analysis, the reflectance ratios are set to $[0, 0, 0, 0, 0.7, 1]$ and the waves reflected many times on the top and bottom reach the receiver.

B. Simulation Conditions

Table 1 presents the specifications of the transmitted signal and the simulation conditions. The pseudo noise signal

Table 1 Specifications of transmitted signal and simulation conditions.

| | |
|-------------------------|------------------------------|
| Sampling frequency | 200 kHz |
| Frequency band | 12 kHz - 32 kHz |
| Measurement time | 250 ms |
| Transmitted signal | Pseudo-noise (PN) sequence |
| Signal length | 163.8 ms |
| Number of signal points | 32768 |
| Number of receivers | 2 |
| Receiver interval | 0.3 m |
| TDOA measurement period | 81.9 ms |
| DFT size | 16384 |
| Sound field | 25 × 15 × 1.35 m |
| Reflectance ratios | [0.7, 0.7, 0.7, 0.7, 0.7, 1] |

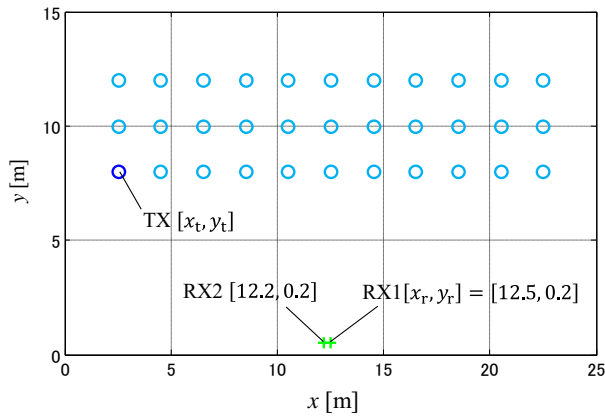
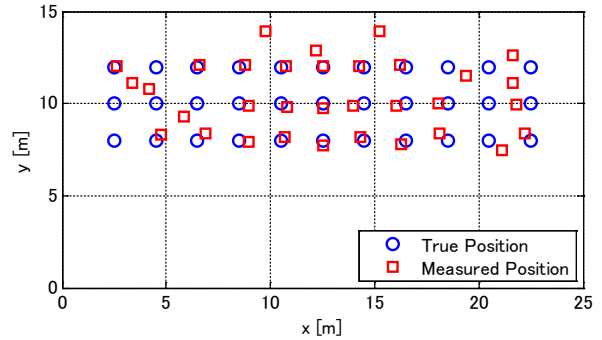


Fig. 7 Locations of transmitter and receiver elements.

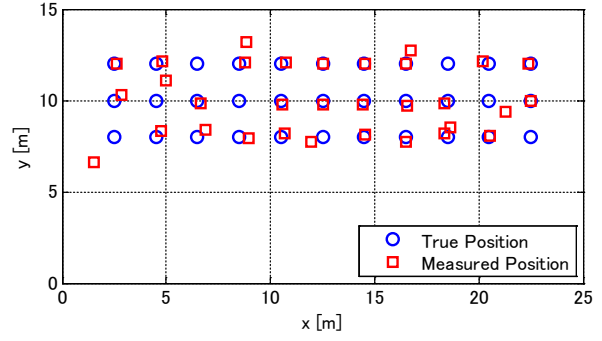
generated by the PN code sequence is used as the transmitted signal. The frequency band of the transmitted signal is 12 kHz to 32 kHz, and it is a flat spectrum with approximately $|X(f)| = 1$ within the band. The acoustic field size is $25 \times 15 \times 1.35$ m (length, width, and height), and the reflectance ratios are set to 1 for water surface and 0.7 for surrounding wall according to the experimental environment.

The locations of the transmitter (TX) and receiver elements (RX1 and RX2) are shown in Fig. 7. TX is moved every 2 m along the x-axis (2.5 to 22.5 m) and the y-axis (8 to 12 m). RX1 is fixed at $x=12.5$ m and $y=0.2$ m with an interval of 0.3 m between the receiver elements. The height of the transmitter and receiver elements is set to the same 0.8 m.

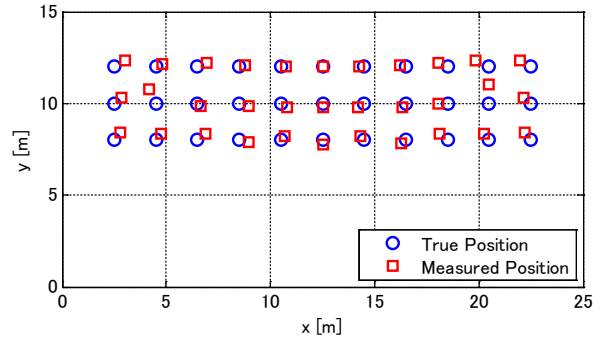
In the signal model of (1), the impulse response depends on the size of acoustic field, the reflectance ratios, and the positions of transmitter and receiver elements in the acoustic simulator. Regarding the uncorrelated noise, we use a signal-to-noise ratio (SNR) given by the ratio of the average signal power of $h_1(k) * \chi(k)$ and the average noise power of $n_1(k)$. We adjust the amplitude of additive white Gaussian noise (AWGN) according to the SNR setting value. We evaluate the position errors for each TDOA algorithm where those errors are calculated by the Euclidean distance between true and measured positions.



(a) GCC-PHAT



(b) MF



(c) IR-GCC-PHAT

Fig. 8 Simulation results.

C. Simulation Results

The simulation results for an SNR of 23 dB are shown in Fig. 8. This SNR corresponds to the value observed in the experience described in Section VI. The true and measured positions are compared in each TDOA algorithm.

GCC-PHAT and MF have a tendency to cause the position errors when the sound source is near the wall. The sound waves reflected from the side walls interfere strongly in this case. GCC-PHAT is influenced by late reverberation. MF induces the errors due to multipath interference caused by the initial reflection. IR-GCC-PHAT can keep high position accuracy (less than about 0.5 m) for most source positions. The summary of simulation results is shown in Table 2. The average distance error is 0.15 m, obtained by the time delay estimation in Section IV. The impact of distance errors is smaller than those of DOA estimation.

Table 2 Summary of simulation results.

| SNR 23 dB | GCC-PHAT | MF | IR-GCC-PHAT |
|----------------------------|----------|------|-------------|
| Average position error [m] | 1.63 | 1.84 | 0.42 |
| Average angle error [deg] | 7.0 | 11.8 | 1.7 |
| Average distance error [m] | 0.15 | | |

Table 3 Simulation results for other conditions.

(a) Average position error for low SNR conditions

| $R_{z22}=1.0$, others=0.7 | GCC-PHAT | MF | IR-GCC-PHAT |
|----------------------------|----------|------|-------------|
| SNR 10 dB | 1.34 | 1.80 | 0.43 |
| SNR 5 dB | 1.50 | 1.80 | 0.44 |
| SNR 0 dB | 2.35 | 3.42 | 0.42 |
| SNR -5 dB | 5.67 | 4.86 | 0.53 |
| SNR -10 dB | 9.82 | 7.49 | 4.90 |

(b) Average position error for non-reflective condition

| $all=0$ | GCC-PHAT | MF | IR-GCC-PHAT |
|------------|----------|------|-------------|
| SNR 20 dB | 0.35 | 0.69 | 0.35 |
| SNR 10 dB | 0.40 | 0.54 | 0.35 |
| SNR 5 dB | 0.42 | 0.54 | 0.35 |
| SNR 0 dB | 0.58 | 0.75 | 0.43 |
| SNR -5 dB | 0.86 | 1.06 | 0.45 |
| SNR -10 dB | 4.59 | 1.12 | 0.76 |
| SNR -15 dB | 6.52 | 1.27 | 1.76 |

The simulation results for the other conditions are reported in Table 3. Table 3(a) shows the average position errors when decreasing a SNR. As the SNR is lowered, the position errors display larger values. IR-GCC-PHAT maintains high position accuracy up to -5 dB. MF has higher position accuracy than GCC-PHAT for the SNR conditions of -5 dB and -10 dB. These results could be explained by the fact that MF and IR-GCC-PHAT are stronger than GCC-PHAT in terms of noise interference.

Table 3(b) gives the average position errors for the non-reflective condition, where the reflectance ratios are all zeros. All TDOA algorithms exhibit comparable position accuracy in case of high SNR conditions (more than 5 dB). The phenomenon that the magnitude of the reflected waves surpasses that of the direct wave (as illustrated in Fig. 3) would not occur in this condition. Under low SNR conditions (less than -5 dB), MF and IR-GCC-PHAT show higher noise resistance than GCC-PHAT.

VI. EXPERIMENT

A. Experimental Conditions

The underwater acoustic location experiment was carried out in the swimming pool, where the experimental scenery is depicted in Fig. 9. The acoustic field size and the positions of the transmitter and receiver elements are the same as in the simulation. The transmitted signal is generated by computer software and transmitted via a DA converter, amplifier, and transducer. The specifications of the transmitted signal are the same as in Table 1. The received signal was recorded from the

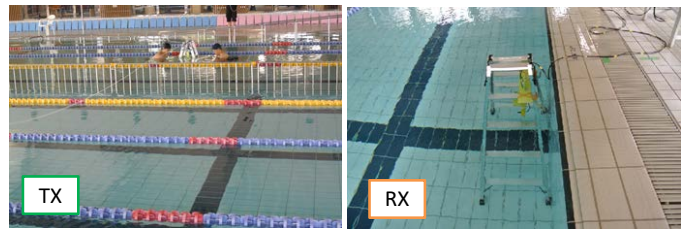
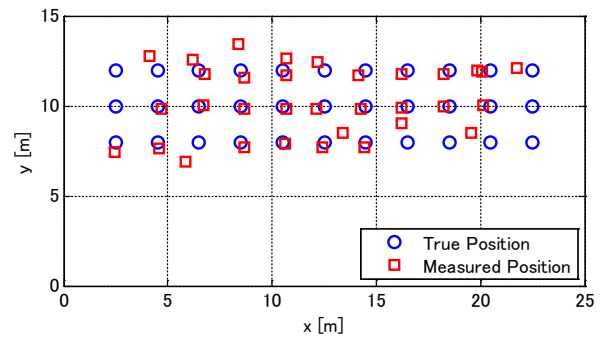
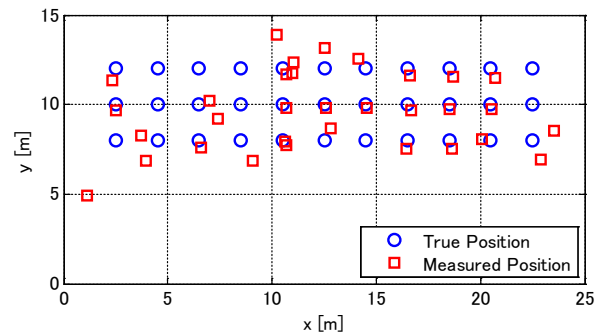


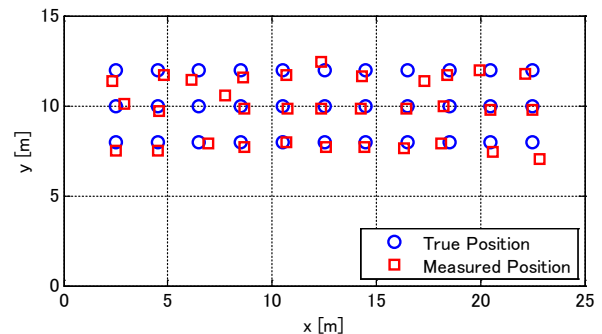
Fig. 9 Experimental scenery.



(a) GCC-PHAT



(b) MF



(c) IR-GCC-PHAT

Fig. 10 Experimental results.

output of the AD converter and analyzed using each TDOA algorithm. The SNR was measured from the ratio of the received power when the transmitted signal was being transmitted and when the transmission was stopped. The average SNR was 23 dB.

B. Experimental Results

The experimental results are shown in Fig. 10. Similar to the simulation, GCC-PHAT and MF have a tendency to cause the

Table 4 Summary of experimental results.

| Average SNR 23 dB | GCC-PHAT | MF | IR-GCC-PHAT |
|----------------------------|----------|------|-------------|
| Average position error [m] | 1.30 | 2.56 | 0.45 |
| Average angle error [deg] | 5.8 | 12.5 | 1.4 |
| Average distance error [m] | 0.28 | | |

Table 5 Average position errors for other SNR conditions.

| | GCC-PHAT | MF | IR-GCC-PHAT |
|------------|----------|------|-------------|
| SNR 10 dB | 1.81 | 2.27 | 0.50 |
| SNR 5 dB | 2.31 | 2.65 | 0.50 |
| SNR 0 dB | 3.31 | 2.95 | 0.71 |
| SNR -5 dB | 5.72 | 2.10 | 1.10 |
| SNR -10 dB | 6.07 | 4.81 | 4.24 |

position errors when the sound source is near the wall. IR-GCC-PHAT displays stable localization performance at all sound source positions. The summary of experimental results is shown in Table 4. The average position errors exhibit characteristics similar to the simulation results. The superiority of IR-GCC-PHAT has been demonstrated even in the actual experiment.

Table 5 shows the average position errors when decreasing a SNR. The SNR adjustment was performed by artificially adding AWGN signals to the original recorded signals. IR-GCC-PHAT shows the highest position accuracy for all SNR conditions. MF shows higher position accuracy than GCC-PHAT for low SNR conditions (less than 5 dB). It provides the similar tendency reported by the simulation results in Table 3(a).

VII. DISCUSSION

Position accuracy of underwater acoustic localization based on USBL system depends on the performance of DOA estimation. In TDOA measurement algorithms, GCC-PHAT [10] and MF [9] are widely used for DOA estimation in underwater acoustic localization. GCC-PHAT is said to be strong in reverberation, however, is sensitive to noise interference including the late reverberation. MF has a noise resistance by using the reference signal, however, happens to cause a significant error by the initial reflection. This paper has presented IR-GCC-PHAT that is strong with both noise and multipath interferences. The effectiveness of the proposed algorithm was confirmed through the simulation and experimental results of Sections V and VI.

The proposed algorithm is effective in an underwater sound wave propagation environment where the water depth is shallow and the sound field is surrounded by walls. As for realistic applications, we assume acoustic positioning of underwater vehicles in a harbor.

IR-GCC-PHAT has the limitation that the source signal should be known at the receiving to measure impulse responses. If the blind estimation of impulse responses for unknown signals was achieved, the proposed algorithm might be applied in wider fields such as room acoustics and others applications [16][17].

VIII. CONCLUSION

This paper has presented a method of underwater acoustic localization in reverberant environments. The proposed TDOA algorithm showed superior position accuracy in both simulation and experiment. In the simulation results, IR-GCC-PHAT maintained high position accuracy less than about 0.5 m for most source positions while the conventional algorithms had larger position errors of more than 1.5 m. The experimental results in the swimming pool had characteristics similar to the simulation results.

In future research, we are trying to investigate how to apply the proposed method to three-dimensional localization and conduct an experimental test using an underwater vehicle.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 20K04477.

References

- [1] M. Jian, A. C. Kot, and M. H. Er, "DOA estimation of speech source with microphone arrays," IEEE International Symposium on Circuits and Systems (ISCAS), pp. 293-296, May 1998.
- [2] J. Zhuang, T. Tan, D. Chen, and J. Kang, "DOA tracking via signal subspace projector update," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4905-4909, May 2020.
- [3] C. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 24, Issue 4, pp. 320-327, Aug. 1976.
- [4] R. Lee, M. Kang, B. Kim, and K. Park, S. Q. Lee and H. Park, "Sound source localization based on GCC-PHAT with diffuseness mask in noisy and reverberant environments," IEEE Access, Vol. 8, pp. 7373-7382, Jan. 2020.
- [5] F. Grondin and F. Michaud, "Noise mask for TDOA sound source localization of speech on mobile robots in noisy environments," IEEE International Conference on Robotics and Automation (ICRA), pp. 1-5, May 2016.
- [6] H. Kang, M. Graczyk, and J. Skoglund, "On pre-filtering strategies for the GCC-PHAT algorithm," IEEE International Workshop on Acoustic Signal Enhancement (IWAENC), pp. 1-5, Sep. 2016.
- [7] Z. Q. Wang, X. Zhang, and D. Wang, "Robust TDOA estimation based on time frequency masking and deep neural networks," 19th Annual Conference of the International Speech Communication Association (Interspeech 2018), pp. 322-326, Aug. 2018.
- [8] G. L. Turin, "An introduction to matched filters," IRE Transactions on Information Theory, Vol. 6, Issue 3, pp. 311-329, June 1960.
- [9] B. Kouzoundjian, F. Beaubois, S. Reboul, J. B. Choquel, and J. Noyer, "A TDOA underwater localization approach for shallow water environment," MTS/IEEE OCEANS 2017 - Aberdeen, pp. 1-4, June 2017.
- [10] J. Choi, J. Park, Y. Lee, J. Jung, and H. Choi, "Robust directional angle estimation of underwater acoustic sources

- using a marine vehicle,” MDPI Sensors, Vol. 18, Issue 9, pp. 1-14, Sep. 2018.
- [11] T. Zhang, H. Shi, L. Chen, Y. Li, and J. Tong, “AUV positioning method based on tightly coupled SINS/LBL for underwater acoustic multipath propagation,” MDPI Sensors, Vol. 16, Issue 3, pp. 1-16, Mar. 2016.
- [12] J. Chen, J. Benesty, and Y. Huang, “Time delay estimation in room acoustic environments: an overview,” EURASIP Journal Advanced Signal Processing, pp. 1-19, Dec. 2006.
- [13] A. Bultan and R. A. Haddad, “Channel estimation in noisy conditions using time-frequency domain filtering,” IEEE Thirty-Third Asilomar Conference on Signals, Systems, and Computers, pp. 1642-1646, Oct. 1999.
- [14] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” Journal of the Acoustical Society of America, Vol. 65 Issue 4, pp. 943-950, Apr. 1979.
- [15] D. R. Campbell, K. J. Palomaki, and G. J. Brown, “Roomsim, a MATLAB simulation of shoebox room acoustics for use in teaching and research,” Computing and Information Systems Journal, Vol. 9, Issue 3, pp. 1-4, Oct. 2005.
- [16] K. Dudzik, "The possibility of applying acoustic emission method to optimize determination of milling parameters, WSEAS Transactions on Systems and Control," Vol. 15, pp. 302-310, 2020.
- [17] A. Ticherahine, S. Bourebia, A. Makhlof, and P. Wira, "Consumption temporary density for the detection of water leakages in real-time," Engineering World, Vol. 1, pp. 125-130, 2019.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US