IV. Evaluation

A. Position Estimation

The comparison of the TDOA algorithms is performed by and the following simple three-dimensional position estimation. When the coordinates of one of the receiver elements are represented by $\boldsymbol{p}_{\mathrm{r}} = [x_{\mathrm{r}}, y_{\mathrm{r}}, z_{\mathrm{r}}]$, those of the transmitter $\boldsymbol{p}_{\mathrm{t}} = [x_{\mathrm{t}}, y_{\mathrm{t}}, z_{\mathrm{t}}]$ are computed by the estimated angle θ:

$$x_{\mathrm{t}} = x_{\mathrm{r}} + D\cos\theta$$
$$y_{\mathrm{t}} = y_{\mathrm{r}} + D\cos\phi$$
$$z_{\mathrm{t}} = z_{\mathrm{r}} - V$$
$$\cos\phi = \sqrt{1 - \cos^2\theta - \left(\frac{V}{D}\right)^2}, \qquad (16)$$

where D indicates the distance between the receiver and the transmitter positions and V is the vertical position difference. We assume that the parameters of D and V are ideally known for simplify in the evaluation. We briefly describe how to measure the distance and the vertical position below.

In the distance measurement, we assume the time synchronization scheme that transmitter and receiver units have the same clock timing. After the transmitter unit starts sending a signal, the receiver unit measures the time delay T as a sound propagates underwater. The distance can be measured by $D = cT$. In [12], the time synchronization has been achieved by integrating a chip-scale atomic clock (CSAC) into an acoustic modem.

The vertical position can be measured by the depth sensor. The depth sensor converts an underwater pressure value into a position from water surface. In [13], a pressure sensor (e.g., Blue Robotics Bar30) was used in a sensor suite for underwater reconstruction. As long as the depth sensor is connected to the receiver unit by a wired communication cable, the depth information is easily obtained. If not, we should consider how to send the depth information in underwater wireless communication. This solution would be discussed in our future paper.

B. Simulation Conditions

Table 1 presents the specifications of the transmitted signal and the simulation conditions. The pseudo noise signal 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, given by a flat spectrum with approximately $|X(l)| = 1$ within the band. The acoustic field size is 25 × 15 × 5 m (length, width, and height), and the reflectance ratios are set to 1 for water surface and 0.8 for water bottom and all surrounding walls.

The locations of the transmitter (TX) and receiver elements (RX1 and RX2) are shown in Fig. 5. TX is moved every 2 m along the x-axis (2.5 to 22.5 m) and the y-axis (8 to 12 m) and 0.5 m along the z-axis (0.5 m to 4 m). RX1 and RX2 are fixed at [12.5, 0.05, 4.5] and [12.2, 0.05, 4.5], respectively. The array space between receiver elements is 0.3 m. The sound velocity is 1480 m/s.

The signal modeling in (1) is done by using a sound wave propagation simulator [14]. The generated impulse response

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 array space	0.3 m
TDOA measurement period	81.9 ms
DFT size	16384
Acoustic field	25 × 15 × 5 m

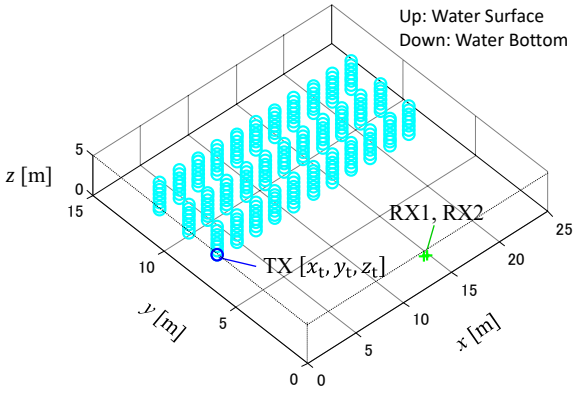


Fig. 5 Locations of transmitter and receiver elements.

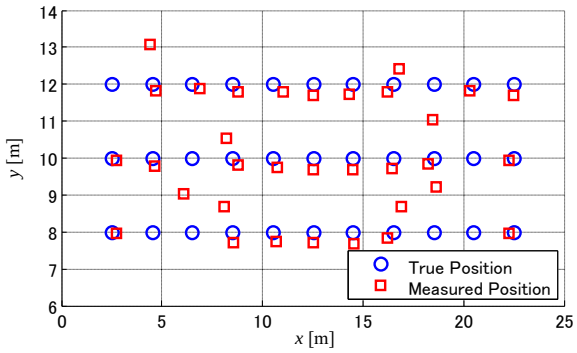
depends on the size of acoustic field, the reflectance ratios, and the positions of transmitter and receiver elements. The magnitude of uncorrelated noise is adjusted by a signal-to-noise ratio (SNR). The SNR is given by the ratio of the average signal power of $h_1(k) * x(k)$ and the average noise power of $n_1(k)$. The amplitude of additive white Gaussian noise (AWGN) is adjusted by the SNR setting value. We evaluate the position errors for the TDOA algorithms where those errors are calculated by the Euclidean distance between true and measured positions. The measured position is computed from (16) where the position of RX1 is used as the coordinates of $\boldsymbol{p}_{\mathrm{r}}$. We apply 326 samples for s_{I} as for the impulse response shortening in (15). See Section V.C about how to determine the value of s_{I}.

C. Simulation Results

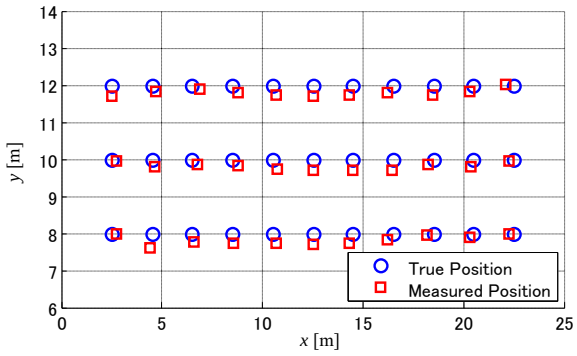
The simulation results for a 30 dB SNR are reported in Table 2. The average position errors for the TDOA algorithms are compared for each TX height. IR-GCC-PHAT is superior to GCC-PHAT and MF in a highly reflective environment as the similar results were previously presented in [7]. SIR-GCC-PHAT provides smaller position errors than IR-GCC-PHAT. Average position errors are 3.16 m, 2.30 m, 0.64 m, and 0.37 m for GCC-PHAT, MF, IR-GCC-PHAT and SIR-GCC-PHAT, respectively. IR-GCC-PHAT shows the best performance in underwater acoustic localization under reflective environment.

Table 2 Average position errors for a 30 dB SNR in meter.

TX location in z-axis [m]	GCC-PHAT	MF	IR-GCC-PHAT	SIR-GCC-PHAT
0.5	1.47	0.42	0.54	0.37
1.0	2.14	3.01	0.52	0.36
1.5	2.32	2.67	0.49	0.41
2.0	2.74	2.89	0.51	0.31
2.5	2.61	3.03	0.49	0.31
3.0	4.01	2.23	0.95	0.58
3.5	3.25	3.14	0.89	0.30
4.0	3.49	1.02	0.70	0.30
Average for all	**3.16**	**2.30**	**0.64**	**0.37**



(a) IR-GCC-PHAT



(b) SIR-GCC-PHAT

Fig. 6 Measured positions when TX is located at 4.0 m in *z*-axis.

Figure 6 illustrates the measured positions when the TX is located at 4.0 m in *z*-axis. IR-GCC-PHAT has large position errors more than 1 m in some TX locations. On the other hand, SIR-GCC-PHAT keeps high position accuracy for most TX positions. This difference is linked to the results of the average position errors.

The simulation results for the other conditions are summarized in Table 3. Table 3(a) shows the average position errors when decreasing a SNR. The tendency of position accuracy for the TDOA algorithms is the same as Table 2. SIR-GCC-PHAT can keep a small position error less than 0.5 m even in a −5 dB SNR.

Table 3(b) gives the average positions for the non-reflective condition, where the reflectance ratios are all zeros. As long as there is no reflection in the acoustic field, MF shows the highest position accuracy. SIR-GCC-PHAT provides higher accuracy

Table 3 Simulation results for other conditions.

(a) Average position errors for other SNR conditions

SNR [dB]	GCC-PHAT	MF	IR-GCC-PHAT	SIR-GCC-PHAT
20	2.70	2.38	0.63	0.37
10	3.21	2.56	0.62	0.38
0	5.32	3.02	0.62	0.31
−5	5.75	3.21	0.76	0.49
−10	5.73	4.40	3.92	1.99

(b) Average position errors for non-reflective condition

SNR [dB]	GCC-PHAT	MF	IR-GCC-PHAT	SIR-GCC-PHAT
5	0.19	0.17	0.19	0.19
0	0.25	0.17	0.19	0.19
−5	1.88	0.17	0.19	0.19
−10	5.73	0.17	0.20	0.19
−15	5.76	0.17	2.92	0.30
−20	5.76	2.16	5.80	3.53

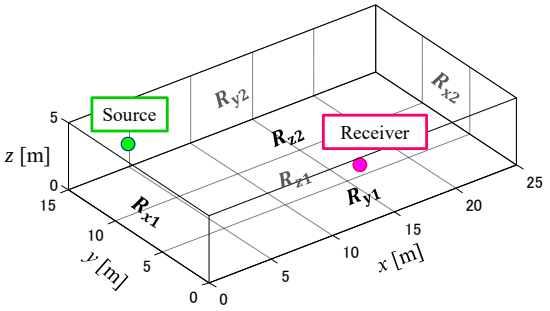


Fig. 7 Settings on sound wave propagation simulator.

than IR-GCC-PHAT in a −15 dB SNR. The impulse response shortening can mitigate the influence of noise components when taking the cross correlation function. The noise robustness of SIR-GCC-PHAT is comparable to MF.

V. Analysis of Acoustic Reflection Paths

A. Sound wave propagation simulator

We explain the effectiveness of SIR-GCC-PHAT by analyzing acoustic reflection paths on a sound wave propagation simulator. Figure 7 illustrates the settings on the sound wave propagation simulator. The parameters of $\boldsymbol{R} = [R_{x1}, R_{x2}, R_{y1}, R_{y2}, R_{z1}, R_{z2}]$ give reflectance ratios on the six surfaces. R_{z1} and R_{z2} are reflectance ratios on the water bottom and surface. The other reflectance ratios on the side walls are expressed as R_{x1}, R_{x2}, R_{y1}, and R_{y2}. Given the coordinates of a sound source and a receiver, acoustic paths and an impulse response from a source to a receiver are analyzed.

Figure 8 shows an example of acoustic paths and impulse responses for the non-reflective condition. The locations of TX, RX1, and RX2 have been used in the simulation of Section IV. Although the impulse responses in RX1 and RX2 seems to be the same, their arrival times of direct waves are slightly different. The time difference of 0.125 ms is converted to a DOA angle by (3).

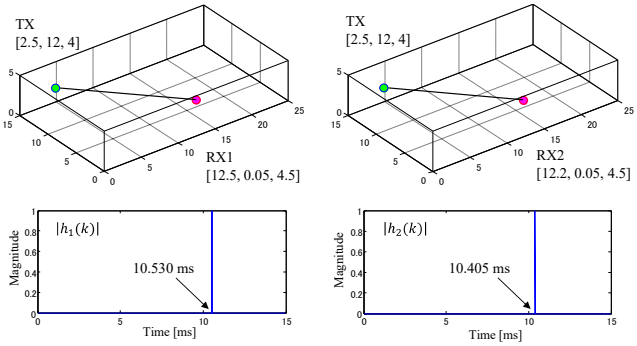


Fig. 8 Example of acoustic paths and impulse responses.

B. *TDOA measurement under Acoustic Reflection*

Figure 9 illustrates the acoustic paths and the impulse responses when the reflections on water surface and bottom are valid. The reflection ratio is set to $\boldsymbol{R} = [0, 0, 0, 0, 0.8, 1]$. The direct wave is observed as the peak of P_1 (P_1' in case of RX2) in the impulse response. The reflected waves are observed as the peaks of P_2, P_3, P_4, ... (P_2', P_3', P_4', ...).

Table 4 shows the magnitude and arrive time for the peaks. In the magnitude, these values are normalized by the maximum value. The peak value of P_2 is almost the same as P_1. The acoustic path of this reflected wave is highlighted in Fig. 9. Since the path length difference between direct and reflected waves is slight, the peak difference becomes only 0.004. As long as impulse response estimation is ideally done, this peak difference would be correctly detected. However, we should consider the uncertainty related to noise components in (9) and the limitation of time resolution in terms of FFT size and sampling frequency. It is hard to distinguish between P_1 and P_2.

In the MF algorithm, the time difference is detected from the two time positions with the highest peaks in RX1 and RX2. When the correct time difference cannot be detected when the magnitude relationship between P_1 and P_2 (P_1' and P_2') is reversed.

The time differences for the peak combinations between RX1 and RX2 are shown in Table 5. Interestingly, the time difference between P_1 and P_1' is very similar to that between P_2 and P_2'. It indicates that the DOA angles for the direct waves and the reflected waves are similar in the horizontal direction. IR-GCC-PHAT emphasizes only this time difference by taking the cross correlation function for the two impulse responses. We can detect the correct time difference from the highest peak in the cross correlation function (see Fig. 3), independent of the magnitude relationship between the direct and reflected waves.

Figure 10 shows the acoustic paths and the impulse responses under a more reflective condition where the reflection ratio is set to $\boldsymbol{R}$=[0.8, 0.8, 0.8, 0.8, 0.8, 1]. There are many acoustic paths caused by the reflected waves due to the reflection on the side walls. A lot of peaks are observed in the impulse responses. The reflected waves that arrive at the receiver element after being reflected on the wall have different DOA angles compared to the direct wave. Their peaks does not help to emphasize the correct time difference even for taking the cross correlation function by IR-GCC-PHAT.

The reasonable way is to eliminate the unnecessary peaks in the impulse responses. It is easily done by replacing a part of

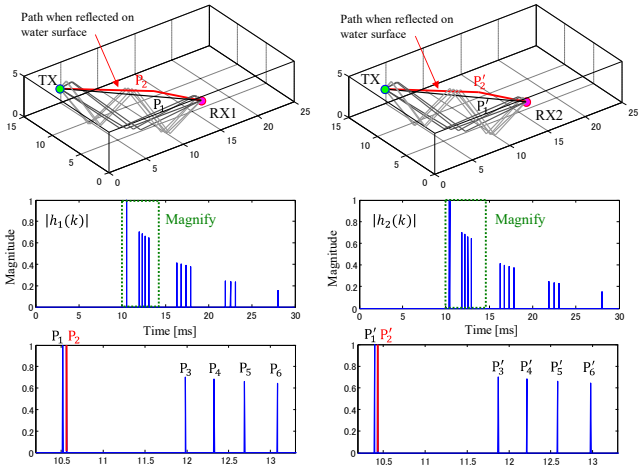


Fig. 9 Acoustic paths and impulse responses for surface and bottom reflection.

Table 4 Magnitude and arrival time for peaks.

(a) Acoustic paths for RX1

Peak	Magnitude	Arrival Time [ms]
P_1	**1.000**	10.530
P_2	**0.996**	10.575
P_3	0.702	11.990
P_4	0.683	12.330
P_5	0.664	12.695
P_6	0.644	13.085

(b) Acoustic paths for RX2

Peak	Magnitude	Arrival Time [ms]
P_1'	1.000	10.405
P_2'	0.996	10.445
P_3'	0.701	11.880
P_4'	0.681	12.220
P_5'	0.661	12.585
P_6'	0.641	12.980

Table 5 Time differences for peak combinations.

Combinations	Time Difference [ms]
P_1 and P_1'	**0.125**
P_2 and P_2'	**0.130**
P_3 and P_3'	0.110
P_4 and P_4'	0.110
P_5 and P_5'	0.110
P_6 and P_6'	0.105

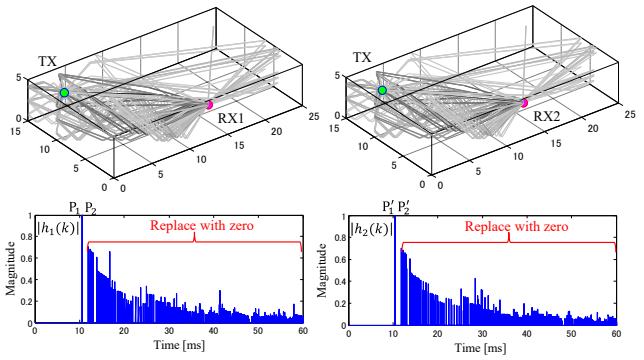


Fig. 10 Acoustic paths and impulse responses for all reflections.

impulse data with zero values, i.e., the impulse response shortening by SIR-GCC-PHAT. We extract only the peaks of P_1 and P_2 (P_1' and P_2').

We compare the cross correlation functions between IR-GCC-PHAT and SIR-GCC-PHAT in Fig. 11. IR-GCC-PHAT has a high noise floor caused by the reflected waves having individual arrival times. It degrades the sharpness of the cross correlation peak and prevents the detection of the correct time difference. On the other hand, SIR-GCC-PHAT maintains a sharp correlation peak and can find a time difference that is close to the true value.

C. Determination of length of shortened impulse response

We explain how to determine an appropriate value of s_I in (15). Since we want to keep the peaks of P_1 and P_2 in the impulse response, the arrival time difference between P_1 and P_2 becomes a clue. Figure 12 shows the vertical view of the acoustic paths to calculate the maximum path length difference between P_1 and P_2. Using the parameters related to the TX and RX1 locations and the water depth, the maximum path length difference is computed as

$$l_{\text{diff}} = 2\sqrt{\left(\frac{l_{\max}}{2}\right)^2 + h^2} - l_{\max}, \tag{17}$$

where $l_{\max}$ denotes the maximum horizontal distance between TX and RX1 and h is the maximum vertical distance between RX1 and water bottom (or water surface). The value of s_I in sample unit can be calculated by

$$s_I = \frac{l_{\text{diff}}}{c} \cdot f_s, \tag{18}$$

where f_s is sampling frequency.

In the simulation evaluation, the value of $l_{\max}$ becomes 15.58 m when the TX is at [2.5, 12] as the furthest position and the RX1 is fixed to [12.5, 0.05] in the *x*-*y* axis. We apply h = 4.5 m from the vertical distance between the RX1 and the water bottom. The value of s_I becomes 326 samples (1.63 ms).

VI. Discussion

TDOA measurement is affected by the reflection of sound waves and noise interference. In TDOA measurement algorithms, GCC-PHAT [5] and MF [6] are widely used in underwater acoustic localization. IR-GCC-PHAT has been reported in our previous work [7].

We have proposed SIR-GCC-PHAT based on the impulse response shortening. By analyzing the acoustic paths under reflective environments, we explained that the unnecessary peaks that does not help to emphasize the correct time difference can be removed in the impulse responses. In the simulation evaluation, SIR-GCC-PHAT has shown smaller position errors that IR-GCC-PHAT and others in the highly reflective condition. As reported in Table 2, average position errors are 3.16 m, 2.30 m, 0.64 m, and 0.37 m for GCC-PHAT, MF, IR-GCC-PHAT and SIR-GCC-PHAT, respectively.

SIR-GCC-PHAT has provided comparable noise robustness

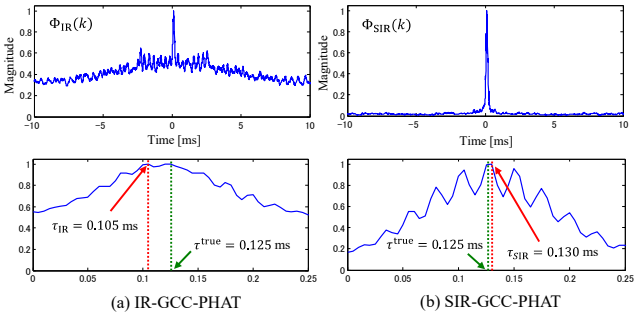


Fig. 11 Comparison of cross correlation functions between IR-GCC-PHAT and SIR-GCC-PHAT.

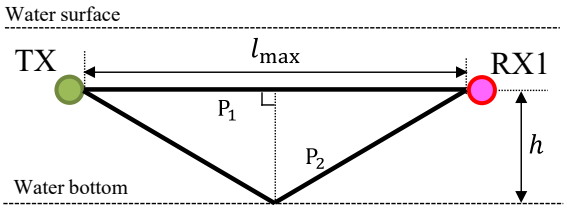


Fig. 12 Vertical view of acoustic paths in direct waves and reflected waves.

to MF. As shown in Table 3(b), MF and SIR-GCC-IR-PHAT can keep smaller position errors less than of 0.5 m up to a −15 dB SNR for the non-reflective condition. Although IR-GCC-PHAT can keep high position accuracy up to a −10 dB SNR, it is slightly inferior to MF and SIR-GCC-PHAT in terms of noise tolerance.

The proposed algorithm is effective when the water depth is shallow and the sound field is surrounded by walls. We assume acoustic positioning of underwater vehicles in a harbor and a dam as a use case.

VII. Conclusion

This paper studied the effect of impulse response shortening in TDOA measurement for underwater acoustic localization. From the acoustic path analysis, we explained that the suppression of unnecessary peaks in impulse responses help to generate a sharp peak with a correct time difference in the cross correlation function. The proposed algorithm showed superior position accuracy compared to other algorithms in the simulation.

In future research, we are planning to present how to apply the proposed algorithm in three-dimensional localization when the depth information is sent in wireless communication.

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. Chen, J. Benesty, and Y. Huang, "Time delay estimation in room acoustic environments: an overview," EURASIP Journal on Advances in Signal Processing, Vol. 2006, No. 1, pp. 1-19, Dec. 2006.
[3] J. Fischer, C. Doolan, M. Rowan, D. Lamos, J. Seers, O. Vargas, S. Lam, and A. Skvortsov, "Acoustic localization of a buoyancy driven model using a beamforming hydrophone array," Applied Acoustics, Vol. 174, No. 107798, Mar. 2021.
[4] G. Pham, V. Baron, A. Finez, J. I. MARS, and B. Nicolas, "High resolution source localization in underwater acoustics for deep sea mining monitoring," MTS/IEEE OCEANS 2019 - Marseille, pp. 1-7, June 2019.
[5] 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.
[6] 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.
[7] S. Yoshizawa, "Underwater acoustic localization based on IR-GCC-PHAT in reverberant environments," International Journal of Circuits, Systems and Signal Processing, Vol. 15, pp.164-171, Mar. 2021.
[8] 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.
[9] 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.
[10] A. Mertins, T. Mei and M. Kallinger, "Room impulse response shortening/reshaping with infinity- and p-norm optimization," IEEE Transactions on Audio, Speech, and Language Processing, Vol. 18, No. 2, pp. 249-259, Feb. 2010.
[11] P. J. W. Melsa, R. C. Younce and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," IEEE Transactions on Communications, Vol. 44, No. 12, pp. 1662-1672, Dec. 1996.
[12] K. G. Kebkal, O. G. Kebkal, I. Glushko, V. Kebkal et.al, "Underwater acoustic modems with integrated atomic clocks for one-way travel-time underwater vehicle positioning," 4th Underwater Acoustics Conference and Exhibition, pp. 315-323, Sep. 2017.
[13] S. Rahman, N. Karapetyan, A. Q. Li, and I. Rekleitis, "A modular sensor suite for underwater reconstruction," MTS/IEEE OCEANS - Charleston, pp.1-6, Oct. 2018.
[14] 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.