INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING
DOI: 10.46300/9106.2021.15.18 Volume 15, 2021

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

Shingo Yoshizawa
Kitami Institute of Technology, Kitami, Hokkaido 0908507, Japan

Received: January 28, 2021. Revised: February 27, 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

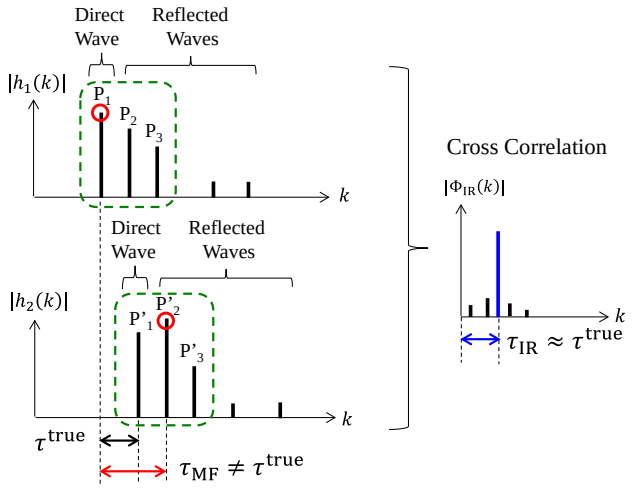


Fig. 4 Detection of the arrival time difference when the pseudo peaks occur.

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 $\boldsymbol{p}_{\text{r}} = [x_{\text{r}}, y_{\text{r}}]$, the transmitter position $\boldsymbol{p}_{\text{t}} = [x_{\text{t}}, y_{\text{t}}]$ is computed by the estimated angle θ:

$$x_{\text{t}} = x_{\text{r}} + D\cos\theta$$
$$y_{\text{t}} = y_{\text{r}} + D\sin\theta. \quad (16)$$

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,

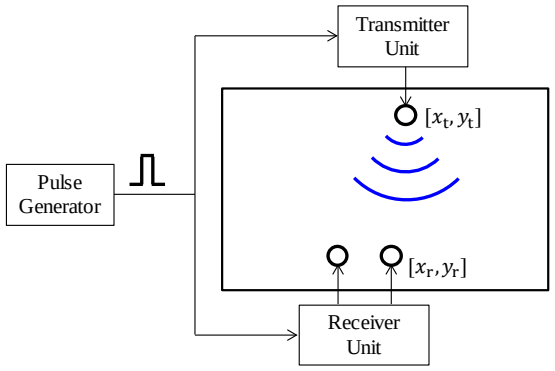


Fig. 5 Time synchronization scheme for distance measurement.

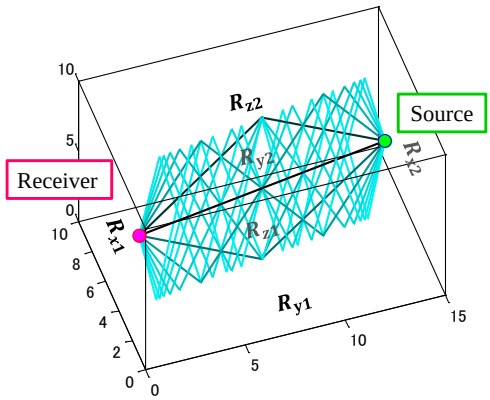


Fig. 6 Sound wave propagation simulator.

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 $\boldsymbol{R} = [R_{\text{x1}}, R_{\text{x2}}, R_{\text{y1}}, R_{\text{y2}}, R_{\text{z1}}, R_{\text{z2}}]$ 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]

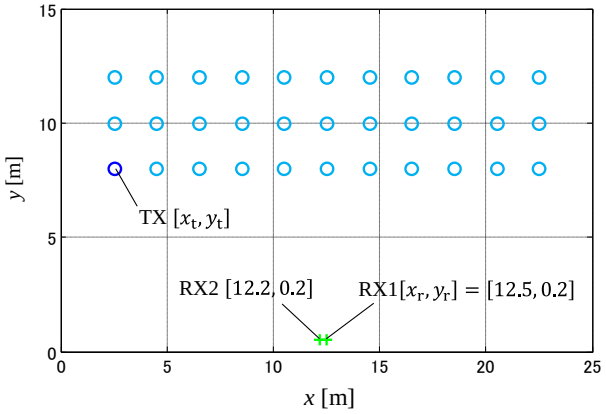
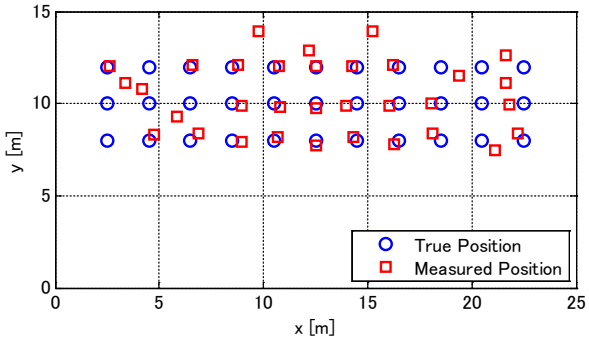


Fig. 7 Locations of transmitter and receiver elements.

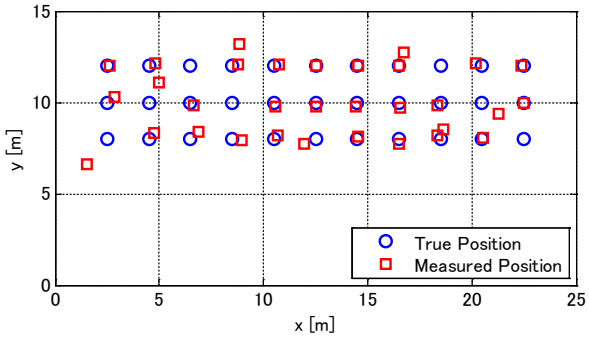
generated by the PN code sequence is used as the transmitted signal. The frequency band of the transmitted signal is 12 kHz to 32 kHz, and it is a flat spectrum with approximately $|X(l)| = 1$ within the band. The acoustic field size is 25 × 15 × 1.35 m (length, width, and height), and the reflectance 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.

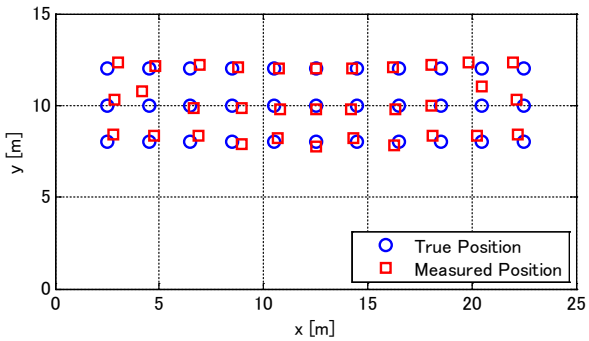
In the signal model of (1), the impulse response depends on the size of acoustic field, the reflectance ratios, and the positions of transmitter and receiver elements in the acoustic simulator. Regarding the uncorrelated noise, we use a signal-to-noise ratio (SNR) given by the ratio of the average signal power of $h_1(k) * x(k)$ 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_{z2}=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.

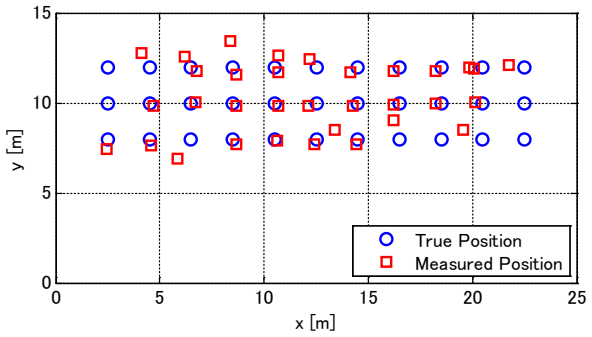
Table 3(b) gives the average position errors for the non-reflective condition, where the reflectance ratios are all zeros. All TDOA algorithms exhibit comparable position accuracy in case of high SNR conditions (more than 5 dB). The phenomenon that the magnitude of the reflected waves surpasses that 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

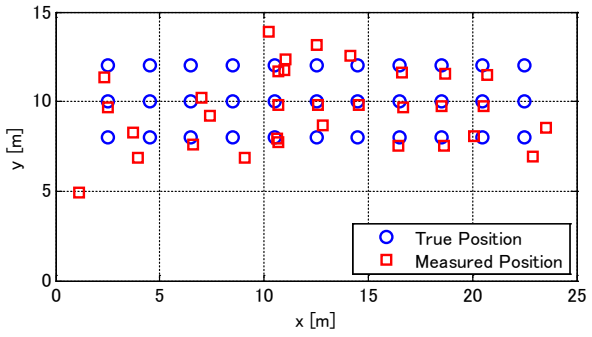
A. Experimental Conditions

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

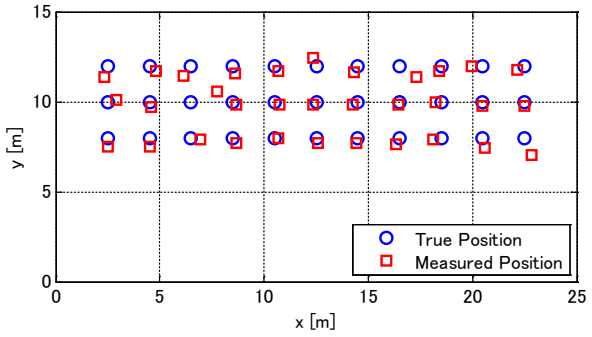
Fig. 9 Experimental scenery.



(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. Makhlouf, and P. Wira, "Consumption temporary density for the detection of water leakages in real-time," Engineering World, Vol. 1, pp. 125-130, 2019.