Estimation of bubble size distribution using deep ensemble physics-informed neural network

딥앙상블 물리 정보 신경망을 이용한 기포 크기 분포 추정

Sunyoung Ko,[1] Geunhwan Kim,[1] Jaehyuk Lee,[2] Hongju Gu,[2] Kwangho Moon,[3] and Youngmin Choo[1†]

(고선영,[1] 김근환,[1] 이재혁,[2] 구홍주,[2] 문광호,[3] 추영민[1†])

[1]Sejong University, [2]Hanwha Ocean Co, [3]LIG Nex1

ABSTRACT: Physics-Informed Neural Network (PINN) is used to invert bubble size distributions from attenuation losses. By considering a linear system for the bubble population inversion, Adaptive Learned Iterative Shrinkage Thresholding Algorithm (Ada-LISTA), which has been solved linear systems in image processing, is used as a neural network architecture in PINN. Furthermore, a regularization based on the linear system is added to a loss function of PINN and it makes a PINN have better generalization by a solution satisfying the bubble physics. To evaluate an uncertainty of bubble estimation, deep ensemble is adopted. 20 Ada-LISTAs with different initial values are trained using the same training dataset. During test with attenuation losses different from those in the training dataset, the bubble size distribution and corresponding uncertainty are indicated by average and variance of 20 estimations, respectively. Deep ensemble Ada-LISTA demonstrate superior performance in inverting bubble size distributions than the conventional convex optimization solver of CVX.

Keywords: Deep ensemble, Physics-Informed Neural Network (PINN), Adaptive Learned Iterative Shrinkage Thresholding Algorithm (Ada-LISTA), Attenuation loss, Bubble size distribution, Uncertainty

PACS numbers: 43.30.Es, 43.60.Pt

초   록: 기포 크기 분포를 음파 감쇄 손실을 이용하여 역산하기 위해 Physics-Informed Neural Network(PINN)을 사용하였다. 역산에 사용되는 선형시스템을 풀기 위해 이미지 처리 분야에서 선형시스템 문제를 해결한 Adaptive Learned Iterative Shrinkage Thresholding Algorithm(Ada-LISTA)를 PINN의 신경망 구조로 이용하였다. 더 나아가, PINN의 손실함수에 선형시스템 기반의 정규항을 포함함으로써 PINN의 해가 기포 물리 법칙을 만족하여 더 높은 일반화 성능을 가지도록 하였다. 그리고 기포 추정값의 불확실성을 계산하기 위해 딥앙상블 기법을 이용하였다. 서로 다른 초기값을 갖는 20개의 Ada-LISTA는 같은 훈련데이터를 이용하여 학습되었다. 이 후 테스트시 훈련데이터와 다른 경향의 감쇄 손실을 입력으로 사용하여 기포 크기 분포를 추정하였고, 추정값과 이에 대한 불확실성을 20개 추정값의 평균과 분산으로 각각 구하였다. 그 결과 딥앙상블이 적용된 Ada-LISTA는 기존 볼록 최적화 기법인 CVX보다 기포 크기 분포를 역산하는데 더 우수한 성능을 보였다.

핵심용어: 딥앙상블, Physics-Informed Neural Network (PINN), Adaptive Learned Iterative Shrinkage Thresholding Algorithm (Ada-LISTA), 감쇄 손실, 기포 크기 분포, 불확실성

I. Introduction

The population of bubbles in water is one of the key factors affecting the transmission of underwater sound.

The acoustic characteristics of bubbles, such as resonance, scattering, and attenuation depend on the size and amount of the bubbles. In various fields, many work have been performed to measure the bubble size distribution (or

†**Corresponding author:** Youngmin Choo (ychoo@sejong.ac.kr)
Department of Defense Systems Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea
(Tel: 82-2-6935-2532, Fax: 82-2-3408-3300)

bubble population).[1] As it is challenging to observe the bubble population directly, several studies have focused on predicting the bubble size distribution by utilizing their properties.[2]

Based on the Medwin's approach, measurable attenuation loss according to frequencies and bubble size distribution can be expressed as a simple linear relationship[2] and its inversion yields the bubble size distribution from the attenuation loss. In general, however, it is an underdetermined problem as the dimension of the attenuation loss (the number of given conditions) is smaller than that of the bubble size distribution (the number of unknowns).[3]

One of approaches for solving the ill-posed problem is a convex optimization, which estimates the optimal solution that minimizes objective function subject to constraints.[4] However, the convex optimization solver such as CVX requires a huge computational burden and cannot assess the uncertainty of the solutions.[5]

To estimate the uncertainty, we employed the concept of homogeneous deep ensemble learning, which is the method of combining multiple models based on Neural Networks (NNs) with the same architecture, but different initial values. Also, it yielded better performance than any of the individual models.[6] For the deep ensemble, we chose Physics-Informed Neural Network (PINN) with Adaptive Learned Iterative Shrinkage Thresholding Algorithm (Ada-LISTA), which considers characteristics of the linear system during training to obtain an effective solution and enhance generalization.[7,8]

This paper is organized as follows. Sec. II explains the system model for relationship between the attenuation loss and bubble size distribution. Sec. III introduces PINN with Ada-LISTA for deep ensemble. In Sec. IV, simulations are carried out to examine performance of proposed scheme for estimating bubble size distribution using attenuation loss. Sec. V concludes the present study.

II. System model for estimating bubble size distribution using attenuation loss

As sound wave propagates through bubbly water, bubbles oscillate at resonant frequency, which is determined by bubble size. This oscillation causing absorption and scattering results in power loss of the incident wave. The total loss in power is denoted as attenuation loss, which is the ratio of incident wave intensity to attenuated wave intensity in decibel scale.[2] It can be expressed as follows:

$$\alpha_b(f_i) = \sum_{j=1}^{N} 4.34\sigma_e(f_i, a_j)n(a_j)\triangle a, \qquad (1)$$

where N is the number of the interval of bubble size distribution, σ_e is extinction cross sections related to the absorption and scattering of bubble, $n(a_j)$ is the number of bubble between a_j and $a_j + \triangle a$ per unit volume (also referred as to bubble density), and $\triangle a$ is a radius spacing between the interval.

From Eq. (1), the attenuation loss has a linear relationship with the bubble size distribution as follows:

$$y = \widehat{A}\hat{x}, \qquad (2)$$

where elements of y, $\widehat{A}$ and $\hat{x}$ are attenuation loss $\alpha_b(f_i)$, $4.34\sigma_e(f_i, a_j)\triangle a$, and bubble density $n(a_j)$, respectively.

To prevent a bias causing from uneven l_2 norms of columns in matrix $\widehat{A}$ during the solving the linear system of Eq. (2), we conduct a normalization to make the columns of matrix $\widehat{A}$ have ones as follows:

$$y = Ax. \qquad (3)$$

Elements of A and x are $4.34\sigma_e(f_i, a_j)\dfrac{\triangle a}{\| \hat{a}_j \|}$ and $\| \hat{a}_j \| n(a_j)$, respectively. $\hat{a}_j$ is the jth-column of matrix

$\widehat{A}$ and $\parallel \widehat{a_j} \parallel$ is the corresponding l_2 norm.

In this problem, the inverse problem to obtain x from y has an underdetermined linear system because the number of the attenuation loss is generally smaller than that of the bubble size distribution.

III. PINN for estimating bubble size distribution

PINN is a NN used for numerical solutions of Partial Differential Equation (PDE). It employs prior information such as PDE (i.e., physical laws) as a regularization term to constrain the solution space and enhance generalization.[7]

We propose a modified version of PINN that utilizes the linear system of Eq. (3) as prior knowledge to obtain solutions effectively. First, we replace the conventional NN architecture of PINN with Ada-LISTA considering the linear system model to enhance the performance of estimating the bubble size distribution and generalization (Sec III. 3.1); when using the conventional NN architecture of PINN, estimated bubble size distributions are significantly deviated from true values during test (not seen here). Next, loss functions of Ada-LISTA involve the physics-informed loss function from the linear system to satisfy the relationship between the attenuation loss and bubble size distribution (Sec III. 3.2). At the end, we apply the deep ensemble to increase the performance of modified model, and to assess the uncertainty of solutions (Sec III. 3.3).

3.1 NN architecture for linear system

Ada-LISTA, which is derived from Iterative Soft Thresholding Algorithm (ISTA), has been introduced in image processing. ISTA has been used to infer an optimal solution for linear system in image denoising by iteratively updating the solution through a series of loops. However, ISTA has limitations, such as low convergence rates.[9]

To produce approximate estimations fast by training, Learned-ISTA (LISTA), which interprets ISTA from a NN perspective, was proposed.[9] LISTA is a data-driven
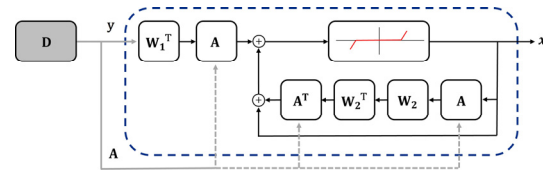

Fig. 1. (Color available online) Ada-LISTA architecture as an iterative model. Ada-LISTA obtains a solution x for a linear system with measurement y and dictionary matrix A by updating the weight matrix W_1, W_2 over iterations (or layers).

algorithm that uses measurement data and matrix of system to train the model. However, it has lack of generalization performance. To address these limitations, Ada-LISTA was introduced.[8]

Ada-LISTA is a learned solver that adapts to the matrix of the system model during training. The output is iteratively reconstructed to be close to label by updating parameters along layers (Fig. 1). Using y, A and x_k (the output of the previous k^{th} layer), Ada-LISTA produces x_{k+1} with the following iteration:

$$x_{k+1} = S_{\theta_{k+1}} \left\{ (I - \lambda_{k+1} A^T W_2^T W_2 A) x_k \right. \tag{4}$$
$$\left. + \lambda_{k+1} A^T W_1^T y, \right.$$

where k is the layer index from 1 to K, $S_{\theta_k}(x) = sign(x)(|x| - \theta_k)$ is a soft thresholding function with threshold value θ_k, and λ_k is the parameter which supports Ada-LISTA to convergence in linear rate. $W_1 = L^{-1} A^T$ and $W_2 = I - L^{-1} A^T A$ are weight matrices. L^{-1} is the step size. Learned parameters, which are collected in Θ, are W_1, W_2, θ_k, and λ_k. Here, the output of Kth layer becomes final prediction x in this problem.

3.2 Loss function in PINN

The loss function in PINN for estimating the bubble size distribution is defined as follows:[7]

$$\min_\Theta \sum_{i=1}^{N_{tr}} \gamma \parallel x_{pred}^{(i)} - x^{(i)} \parallel_2^2, \tag{5}$$
$$+ \parallel y^{(i)} - A x_{pred}^{(i)} \parallel_2$$

where N_{tr} is the number of training data, x_{pred} is the Kth output of Ada-LISTA, and x is label. Each term in the loss function means as follows: 1) $\| x_{pred}^{(i)} - x^{(i)} \|_2^2$ trains the Ada-LISTA in supervised learning frame which minimizes the error between $x_{pred}^{(i)}$ from $y^{(i)}$ and the corresponding label $x^{(i)}$. γ is a hyperparameter that controls the impact of the supervised learning loss function in the training. 2) $\| y^{(i)} - A x_{pred}^{(i)} \|_2$ trains the Ada-LISTA to follow the physics in the form of linear system, which is different from the conventional PINN using a PDE for loss function.

3.3 Deep ensemble for assessing the uncertainty of estimation

The deep ensemble is a method of combining the predictions from several machine learning models to increase a prediction accuracy, which is better than any of the individual models. Especially, homogenous deep ensemble is a technique that averages predictions from NNs having the same architecture but having different initialization. Since different initial values in NNs lead to different local minima, these diverse solutions can yield a better solution with their average and offer uncertainty of the solutions with their variance.[6]

We train many Ada-LISTAs with different initial values using Eq. (5) and obtain the final solution and uncertainty with the average and variance, respectively.

IV. Performance analysis of deep ensemble Ada-LISTA in estimating bubble size distribution

4.1 Data setup

Owing to difficulty in conducting experiments and labeling data for inverting the bubble size distribution using attenuation loss, we generated training and test dataset using the system model in Sec. II.

The training dataset consists of 5 000 input-output pairs.

Elements of input and output are attenuation loss at a specific frequency and normalized bubble density at a specific bubble radius. For the normalization, $n(a_j)$ is randomly generated between zero and one and is scaled to be a low void fraction regime, where the linear system is valid, as follows:

$$ n(a_j) \leftarrow \left\{ n(a_j) / \sum_{i=1}^{N} \frac{4}{3} \pi a_j^3 n(a_j) \triangle a \right\} V , \qquad (6) $$

where V is void fraction, which is less than 10^{-4}.

Input data y consists of $\alpha_b(f_i)$ at 781 frequencies spaced 10 Hz apart, ranging from 200 Hz to 8k Hz and has weak noise; signal-to-noise ratio is about 10 dB. The frequency range is determined by considering data from the relevant experiment conducted by Hanwha Ocean Co (not seen here). Output data x consists of $n(a_j)$ at 1 500 bubble radii from 580 µm to 15 000 µm. The smallest radii of 580 µm corresponds to a resonant frequency 8 kHz. However, the largest bubble size is set as 15 000 µm because of the difficulty to observe bubble radii above the limit in the relevant experiments.[5]

In the training phase, the data pair is made with randomly generated bubble size distribution x and the corresponding attenuation loss $y = Ax$. Since the attenuation loss becomes different according to activating region in the bubble size distribution, the training dataset involves a data pair of $(x,\ y)$, where components in a specific region of x have non-zero values, as shown in Fig. 2 (the first 2 500 instances for x); the remaining data pairs are obtained by activating all components of x. By using the training dataset as above, deep ensemble Ada-LISTA experiences dynamic patterns during training and it is advantage to better generalization.

During test, the bubble size distribution is estimated using the trained deep ensemble Ada-LISTA for a given y. Unlike x in the training phase, normal distributions with various standard deviations are used for x to generate y deviated from that in the training dataset.[10] With this setting, the generalization can be evaluated by analyzing
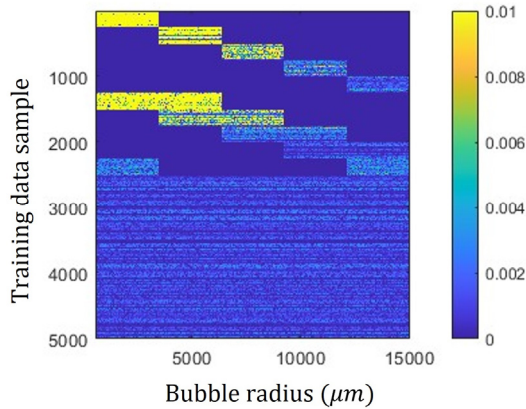
Fig. 2. (Color available online) Data samples for x in training dataset. For the visualization, original (or unnormalized) bubble size distribution is displayed in the range of 0 to 0.01. To generate dynamic attenuation loss patterns during training, components of specific region in x have non-zero values (the first 2 500 samples for x).

Table 1. Summary of dataset.

Dataset		Training / Validation	Test
Number of dataset		5 000 / 1 000	500
Input y $\left(\alpha_b(f_i)\right)$	Number of frequencies (Range)	781 (200 Hz to 8 000 Hz)	
	Distribution	Activation distribution in specific region (Fig. 2.)	Gaussian distribution
Output x $\left(n(\alpha_j)\right)$	Number of bubble radii (Range)	1 500 (580 μm to 15 000 μm)	

bubble estimation performance in the test phase. The datasets are summarized in Table 1.

4.2 Inversion of bubble size distribution using PINN of deep ensemble Ada-LISTA

Layer number, epoch number, and learning rate are 6 (K = 6), 1 500 and 10^{-4}, respectively, which are determined empirically. The hyperparameter γ of the loss function is set as 0.05, which yields similar results in the range of 0.01 to 0.1. We train 20 Ada-LISTAs with different initial values and apply 500 attenuation losses from the test dataset,
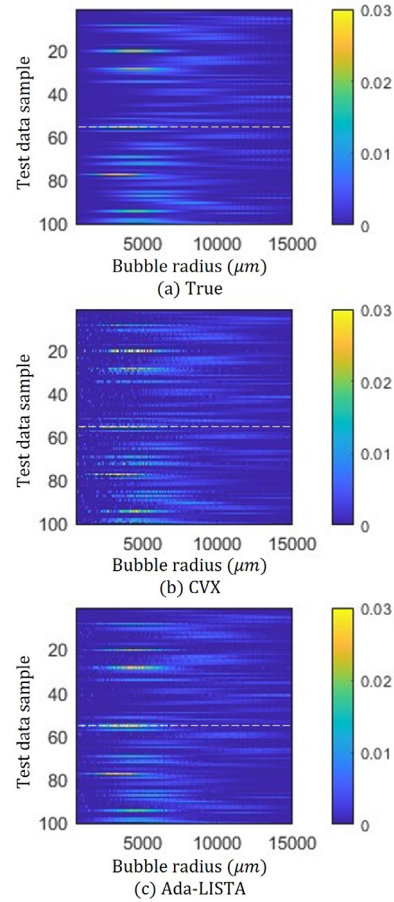

Fig. 3. (Color available online) Bubble size distributions estimated using CVX and deep ensemble Ada-LISTA: true values (a) and estimated values from CVX (b) and deep ensemble Ada-LISTA (c). 100 out of 500 results are displayed for a clear performance comparison. The bubble size distribution marked with dashed line (the 55th test sample) will be used for a detailed performance investigation.

which are different from those in the training dataset.

On the other hand, the convex optimization solver of CVX is applied to the same attenuation losses for a comparison. The objective function in the convex optimization is as follows:

$$\min_{x_c} \| y - Ax_c \|_2 + \gamma_c \| x_c \|_2 \tag{7}$$

$$subject\ to\ x_c \geq 0.$$

The first l_2 norm find a solution to satisfy the linear system and the second l_2 norm prevents the solution from CVX to
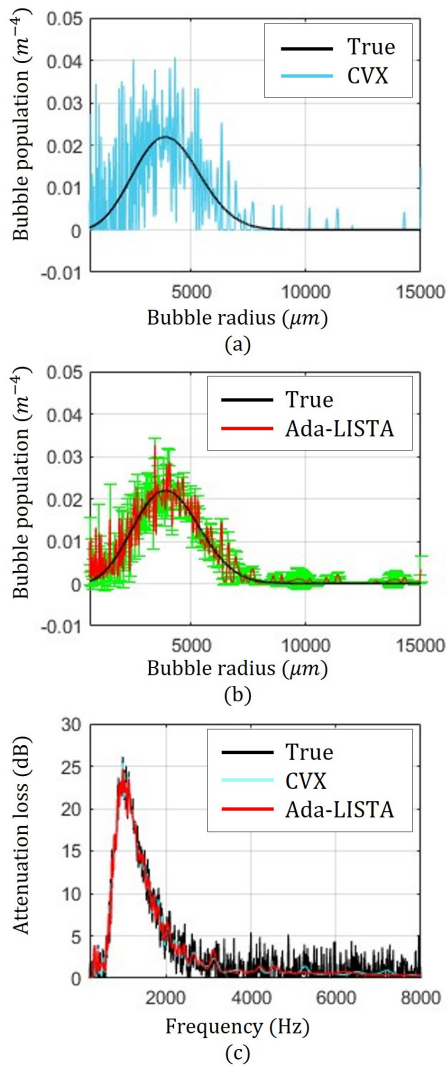
Fig. 4. (Color available online) Bubble size distri-butions estimated using CVX (a), deep ensemble Ada-LISTA (b) for a specific attenuation loss, and reconstructed attenuation loss using the estimated bubble size distributions (c). The CVX results are deviated from true values by significant fluctuations at small bubble size region, which are reduced by deep ensemble Ada-LISTA displaying an estimation uncertainty with variance. The variances are dis-played in every four bubble radius indices for clear representation.

have too-large components. γ_c, which controls the balance between two l_2 norms, is empirically determined as one. The constraint makes all components of solution have non-negative values as in the actual bubble population.

As shown in Fig. 3(b) and (c), estimated results from deep ensemble Ada-LISTA are in better agreement with true values than CVX. For a quantitative analysis, mean square errors calculated using 500 test samples for deep ensemble Ada-LISTA and CVX, which correspond to 6.2 × 10^{-4}, 2.4 × 10^{-3}, respectively.

For a detailed performance investigation, the specific bubble size distribution of the 55th test sample is used and the inversion results using CVX and deep ensemble Ada-LISTA are shown in Fig. 4(a) and (b), respectively; for a clearer comparison, two results are displayed sepa-rately. The CVX result is in good agreement in terms of overall pattern. However, significant fluctuations are observed at the small bubble size region, which make the CVX result deviated from the true values. While the deep ensemble Ada-LISTA estimation also has a gap from the true values at the small bubble size region as in the CVX result, the fluctuations from CVX are considerably reduced. Furthermore, by using the variance of 20 Ada-LISTA estimations for deep ensemble, the estimation uncertainty can be evaluated [error bars in Fig. 4(b)]. The error bars are smaller in the large bubble size region, where the deep ensemble Ada-LISTA result is in good agreement with the true values. The opposite happens in the small bubble size region. From the estimated bubble size distributions, attenuation losses can be reconstrcted as shown in Fig. 4(c). The attenuation losses are very close to true value owing to physics-based constraint ($\| y - Ax_c \|_2$) in the objective or loss function.

V. Conclusion

In this paper, we propose the estimation of the bubble size distribution using deep ensemble of PINN with Ada-LISTA, which can consider the system's characteristics.

This approach improves the bubble inversion perfor-mance and quantifies the uncertainty of the bubble inversion result.

In future work, we will estimate bubble size distributions using attenuation losses from water tank experiments to further investigate feasibility and utility of the proposed model.

Acknowledgement

References

1. H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic Press, Cambridge, 1998), pp. 287-333.

2. H. Medwin "Acoustical determination of bubble-size spectra," J. Acoust. Soc. Am. **62**, 1041-1044 (1977).

3. R. Duraiswami, S. Prabhukumar, and G. L. Chahine, "Bubble counting using an inverse acoustic scattering method," J. Acoust. Soc. Am. **104**, 2699-2717 (1998).

4. S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004), pp. 16-23.

5. C. Park, S. W. Jeong, G. D. Kim, I. Moon, and G. Yim, "A study on the estimation of bubble size distribution using an acoustic inversion method" (in Korean), J. Acoust. Soc. Kr. **39**, 151-162 (2020).

6. M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," Eng. Appl. Artif. lntell. **115**, 105151 (2022).

7. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," J. Comput. Phys. **378**, 686-707 (2019).

8. A. Aberdam, A. Golts, and M. Elad. "Ada-lista: Learned solvers adaptive to varying models," IEEE TPAMI/ PAMI, **44**, 9222-9235 (2021).

9. K. Gregor and Y. L. cun, Learning fast approximations of sparse coding. Proc. 27th International Conf. Machine Learning, 399-406 (2010).

10. K. Lee, C. Lee, and C. Park. "Insertion loss by bubble layer surrounding a spherical elastic shell submerged in water" (in Korean), J. Acoust. Soc. Kr. **41**, 174-183 (2022).

Profile

‣ Sunyoung KO (고 선 영)

Sunyoung KO received the B.S. degree in weapon systems engineering from R.O.K Naval Academy in 2017. Currently, she is a Master's student at Sejong University, where she is focused on signal processing and deep learning.

‣ Geunhwan KIM (김 근 환)

Geunhwan KIM received the B.S. and M.S. degrees in electronics engineering and the Ph.D. degree in electrical and electronics engineering from Kyungpook National University, Daegu, in 2015, 2017, and 2022, respectively. Since 2022, he has been a Postdoctoral Researcher with the Department of Ocean Systems Engineering, Sejong University, Seoul, Republic of Korea. His research interests include underwater acoustic signal processing, high-duty cycle sonar systems, and deep learning applications for sonar systems.

‣ Jaehyuk LEE (이 재 혁)

Jaehyuk LEE received the B.S. degree in International Relationship from R.O.K Naval Academy in 1999, the M.S. degree in Marine Science and Convergence Engineering from Hanyang University in 2005, and the Ph.D. degree in Shipbuilding and Marine Engineering from Seoul National University in 2012. After completing his military service until 2018, he has been a researcher with the Dept. of Vibration Noise R&D Team, Hanwha Ocean, Geoje, R.O.K. His research interests include underwater acoustic signal processing and noise vibration.

‣ Hongju GU (구 홍 주)

Hongju GU received the B.S. and M.S. degree in Ocean Engineering from Korea Maritime & Ocean University in 2016 and 2018, respectively. Since 2018, she has been a researcher with the Dept. of Vibration Noise R&D Team, Hanwha Ocean, Geoje, Republic of Korea. Her research interests include underwater acoustic signal processing and noise vibration.

▸ **Kwangho MOON (문 광 호)**

Kwangho MOON received the B.S. degrees in computer science and the M.S. degree in information and communication engineering from Inha University, Incheon, in 2009, 2011, respectively. Since 2011, he has been a Research Engineer with LIG Nex1, Seongnam-si, Gyeonggi-do, Republic of Korea. His research interests include Target Classification and Applications for sonar systems.

▸ **Youngmin CHOO (추 영 민)**

Youngmin CHOO received the B.A. degree in engineering and the Ph.D. degree in underwater acoustics from Seoul National University, Republic of Korea, in 2005 and 2012, respectively. He was a Postdoctoral Research Fellow at Seoul National University and the Scripps Institution of Oceano-graphy, from 2012 to 2016. Since 2016, he has been an Assistant Professor at the Department of Defense Systems Engineering, Sejong University, Republic of Korea, where he is currently an Associate Professor. He has particular research interest in the areas of signal processing and underwater acoustics.