Development of Acoustic Camera-Imaging Simulator Based on Novel Model

Seungchul Kwak, Yonghoon Ji, Atsushi Yamashita, and Hajime Asama
Department of Precision Engineering
The University of Tokyo
Tokyo, Japan

Abstract—**Acquisition of accurate image information from underwater environments is essential to operate underwater tasks such as maintenance, inspection, trajectory estimation, target recognition and SLAM (simultaneous localization and mapping). In this respect, an acoustic camera is superior device because it can provide images with more accurate details than what the optical cameras provide even in turbid water. However, the structure of the acoustic image is very dissimilar to that of optical image because the acoustic camera provides sequences of distance and azimuth angle data, which elevation angle data are missing. Consequently, it is difficult to generate simulated acoustic images although simulating realistic images is vital for verification for performance of proposed algorithms and saving costs and time. This paper describes the principles of acoustic natures and imaging geometry model of the acoustic camera which greatly contribute to developing simulator for the acoustic images. When compared to a state-of-the-art technique, our approach demonstrates surpassing performance in representing simulated acoustic images in complete gray-scale. We evaluate the simulated images from developed simulator by comparing with real acoustic images. The results indicate that the proposed simulator can generate realistic virtual acoustic images based on the principles.**

Keywords—acoustic image; acoustic imaging simulator; acoustic camera

I. INTRODUCTION

Automated target recognition in an underwater environment is an essential task for ship hull inspection, harbor surveillance, environmental monitoring, and SLAM (simultaneous localization and mapping). Several researches based on optical images have been proposed for SLAM [1] or exploration [2]. Optical vision provides high resolution images which enable us to recognize underwater environments with high accuracy in case of clear waters. The optical vision, however, has a limitation in their visibilities, especially in turbid or deep waters. In contrast, acoustic devices whose visibilities seldom depend on the depth of water and turbidness have high reliability in utilizing in the underwater environment. Forward-looking sonar system is proposed for automatic detection in the underwater environment [3]. Mechanically scanned imaging sonar system is also developed to solve SLAM problem in the underwater environment [4], [5], [6]. In recent years, 2-D high-frequency acoustic cameras such as DIDSON (*D*ual-Frequency *ID*entification *SON*ar) [7], ARIS

(Adaptive Resolution Imaging Sonar) [8] and BlueView [9] have become common sonar devices because they provide acoustic images with high resolution which allow target recognition for human even in the turbid water. A method of generating 3-D views with DIDSON is developed by virtue of its high resolution [10], [11]. As the acoustic cameras provide acoustic images with high definition, sonar-based 3-D reconstruction methods are developed [12], [13]. The acoustic cameras, however, have some problems to conquer for practical use. Due to the signal processing and displaying mechanism of the acoustic camera, acoustic image is very dissimilar to optical image (Fig. 1). As a result, vision-based models cannot be applied to description of the acoustic phenomenon, which means that it is difficult to acquire simulated acoustic images. However, simulating realistic images is important to verify the performance of proposed algorithms and to save costs and time. To this end, a simulator for the acoustic image has been developed in [14]. However, this approach describes the acoustic image by only two colors: black and white, but rather than expressing complete gray-scale. Here, the principles of acoustic natures and imaging geometry model can greatly improve solutions to the approach. Therefore, to obtain the realistic simulated acoustic images, we propose a novel approach using these principles.
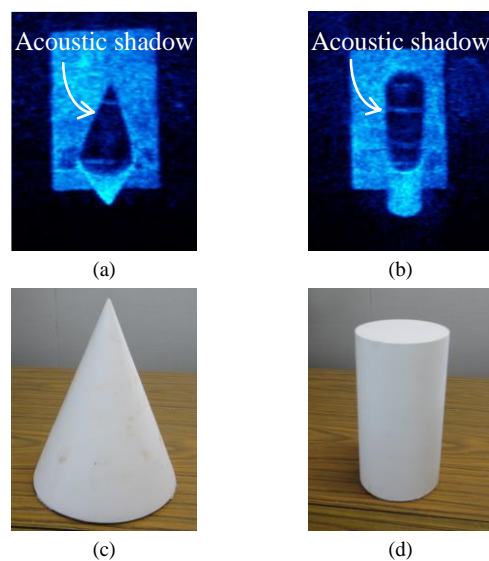

Fig. 1. Comparison of acoustic images and optical images: (a) acoustic image of cone, (b) acoustic image of cylinder, (c) optical image of cone, and (d) optical image of cylinder.

The contribution of this research is as follow. Previous study which could only express the simulated acoustic images in black and white has a significant limitation on the description of the realistic acoustic images. On the other hand, because we figured out the principles of acoustic natures and imaging geometry model, better acoustic-imaging model is established. Our approach shows that it is possible to construct the realistic acoustic images which could be represented in complete gray-scale.

The remainder of this paper is organized as follows. Section 2 provides preliminaries on acoustic projection model and imaging geometry model. In Section 3, the proposed acoustic-imaging model is presented. Section 4 deals with the contents of the proposed simulator. The results of the experiments are described in Section 5 together with the corresponding simulation results. The final section concluded the paper and points out future work prospects.

II. PRELIMINARIES

In this section, acoustic projection model and imaging geometry model are described. These models are depicted based on the characteristics of the acoustic camera. Understanding the acoustic camera's principles can contribute to establishing an acoustic-imaging model which is represented in Section 3.

A. Acoustic Projection Model

To form an acoustic image, an acoustic camera insonifies an acoustic wave in forward direction spanning its field of view in azimuth θ and elevation ϕ direction within the scope of maximum range r. Traveling in forward direction, an insonified acoustic wave collides with an object, which is leading to reflection. The reflected acoustic wave is processed by an array of transducers as a function of distance and azimuth angle, resulting in an image with the attenuated intensities at each point (r, θ). Although the elevation angle ϕ is missing, in the sonar system, 3-D scene point x_{car} represented by Cartesian coordinates is more suitable to be expressed in x_{shp} represented by spherical coordinates. The point x detected by an acoustic wave is defined in the local coordinate system (Fig. 2). Relationship between the Cartesian and spherical coordinates is represented as follow:

$$x_{car} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r\cos\phi\sin\theta \\ r\cos\phi\cos\theta \\ r\sin\phi \end{bmatrix} \quad (1)$$

$$x_{sph} = \begin{bmatrix} r \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2 + y^2 + z^2} \\ \tan^{-1}(x/y) \\ \tan^{-1}\left(z/\sqrt{x^2 + y^2}\right) \end{bmatrix} \quad (2)$$
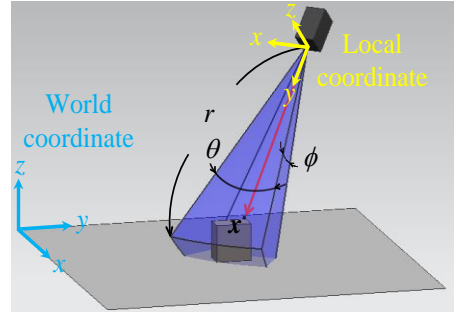

Fig. 2. Acoustic projection model.

where (x, y, z) denotes the coordinates of the point x in Cartesian coordinate system and (r, θ, ϕ) in spherical coordinate system.

B. Imaging Geometry Model

The imaging geometry model is also the most essential to establish an acoustic-imaging process. An acoustic image is represented by three parameters (r, θ, I) where I denotes the intensity value of sound pressure which is usually expressed by the brightness of one single color in the acoustic image. In other words, pixel coordinate in an acoustic image is determined by (r, θ), and each pixel is colored resulting from intensity value. By imaging geometry model, intensity at 3-D point $x=(r, \theta, \phi)$ is projected into a point x' on the image plane, and the point x' exists along the arc defined by the elevation angle as shown in Fig. 3. The point x' which has three information (r, θ, I) is mapped as gray-scale at pixel (r, θ) in an acoustic image (in regard of intensity of sound pressure, it is described in Section 3). In this way, the area whose intensities have no values indicates the acoustic shadow (black-colored area in an acoustic image as shown in Fig. 1). According to the imaging geometry model, in the same beam slice, if multiple acoustic waves travel the same distance, they overlap at same point on the image plane. For example, in Fig. 3, some part of the upper side and the right side of the object are located at the same distance from the sound source, resulting in overlapping on the acoustic image. When the data overlap, the pixels that are displayed are the aggregate of each point's intensity.

III. ACOUSTIC-IMAGING MODEL

In this section, the acoustic-imaging model is proposed based on acoustic natures and the models described in Section 2. The acoustic-imaging model is crucial for development of simulator. In this paper, we propose a novel acoustic-imaging model which makes it possible for simulator to generate acoustic images in complete gray-scale. The ultimate goal of making acoustic-imaging model is to represent a simulated acoustic image more realistic. Thus, we constructed 128×281 direction vectors, 128 vectors for azimuth direction with 0.25 degree intervals and 281 vectors for elevation direction with 0.05 degree intervals. As these direction vectors are considered as the acoustic wave, they are originated from the acoustic camera. Each direction vector can be represented as spherical coordinates as follow:
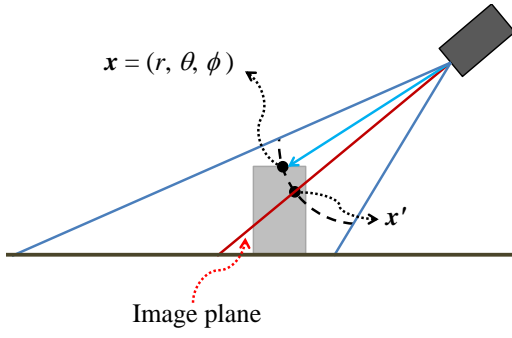
$$x = (r, \theta, \phi)$$

Image plane

Fig. 3. Imaging geometry model.

$$\hat{e}_{(i,j)} = \begin{bmatrix} r_e \\ \theta_i \\ \phi_j \end{bmatrix} = \begin{bmatrix} 1 \\ 106 - \left(\dfrac{0.25}{2} \right) \cdot (2i+1) \\ 7 - 0.05 \cdot j \end{bmatrix} \quad (3)$$

where $\hat{e}_{(i,j)}$ denotes a direction vector. r_e, θ_i and ϕ_j are the norm of unit vector, azimuth angle, and elevation angle respectively. $i=1, \dots, 128$ indicates the indices of azimuth direction vectors and $j=1, \dots, 281$ indicates the indices of elevation direction vectors. These vectors can be described in Cartesian coordinate system by using (1). The global coordinates of these vectors can be transformed by homogeneous transform as follow:

$$\begin{bmatrix} x_w & y_w & z_w & 1 \end{bmatrix}^{\mathrm{T}} = $$
$$\begin{bmatrix} \cos\theta\cos\varphi & -\cos\psi\sin\varphi - \sin\psi\sin\theta\cos\varphi & \sin\psi\sin\varphi - \cos\psi\sin\theta\cos\varphi & T_x \\ \cos\theta\sin\varphi & \cos\psi\cos\varphi - \sin\psi\sin\theta\sin\varphi & -\sin\psi\cos\varphi - \cos\psi\sin\theta\sin\varphi & T_y \\ \sin\theta & \cos\theta\sin\psi & \cos\psi\cos\theta & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix} \quad (4)$$

where $(T_x, T_y, T_z, \psi, \theta, \varphi)$ indicates 6DOF poses of the acoustic camera. (x_w, y_w, z_w) and (x_l, y_l, z_l) indicate world coordinates, and local coordinates of direction vectors respectively.

Each direction vector advances inch by inch until the vectors collide with an object. The vectors reflected from the object return to acoustic camera with having information about which orientation they traveled on, how far they traveled, and how much sound pressure is attenuated (in regard of sound pressure attenuation, it is depicted in next paragraph). Therefore, returned vectors have enough information to represent an acoustic image which is necessary distance data, azimuth angle data and intensity data to be mapped.

A. Sound Pressure Attenuation

It is essential to consider the sound pressure attenuation because it plays a great role to express the intensity in acoustic images. The sound pressure attenuation occurs because of absorptive attenuation and reflection on an object. A formula of absorptive attenuation in seawater is established in [15], however, we only consider freshwater contribution in this paper because the experiment that we performed was executed in freshwater. Considering freshwater contribution only, absorptive attenuation coefficient α can be represented as

$$\alpha = 0.00049 f^2 e^{-(T/27 + D/17)} \quad (5)$$

where α is measured in decibels per kilometer. f, T, and D mean frequency of an acoustic wave in kilohertz, temperature of water in Celsius degrees, and depth of water in kilometers, respectively. The unit of absorption coefficient α can be converted to pascals per kilometer by using

$$n = 20 \log_{10} \left(\frac{p}{p_0} \right) \quad (6)$$

where p is the amount of sound pressure being considered, p_0 is a reference sound pressure, and n is the number of decibels denoting the ratio p/p_0.

A formula of reflection attenuation which is called Rayleigh reflection coefficient formula is established in [16]. It is represented by four parameters: two acoustic impedances z_1 and z_2, incidence angle θ_i and refraction angle θ_t (Fig. 4). Acoustic impedance is the ratio of acoustic pressure p to acoustic volume flow u. If ρ is density of a medium and c is speed of sound, acoustic impedance z can be represented as product of ρ and c. Acoustic impedance is a characteristic value determined by medium. It has great effect on reflection of sound wave because the bigger the gap of bordering two media's impedance, the more sound wave is reflected. The formula of Rayleigh reflection coefficient is described as follow:

$$R_p = \frac{z_2 \cos\theta_i - z_1 \cos\theta_t}{z_2 \cos\theta_i + z_1 \cos\theta_t} \quad (7)$$

here, subscript on acoustic impedance z indicates bordering two media respectively. Thus, by substituting (5), (6), and (7), whole sound pressure attenuation can be represented as follow:

$$\alpha_{\text{whole}} = (\eta l)^2 R_p \quad (8)$$

where α_{whole} indicates how much the sound pressure attenuates respecting to a reference sound pressure, η denotes constant value calculated by (5) and (6), and l indicates half traveled distance of an acoustic wave in meters.

B. Object

As the acoustic-imaging model is constructed by the number of the direction vectors, an object is recognized as a group of points. These points act as giver to present their intensity data to acoustic image. We figured out a formula of whole sound pressure attenuation, however, an incidence angle is still unknown parameter in (7). An incidence angle can be calculated based on the general wave reflection's law (θ_i is equal to θ_r which is reflection angle as shown in Fig. 4) and Snell's law. Furthermore, to calculate the incidence angles corresponding to each direction vector, it is necessary to build a tangent plane at point where each direction vector collides with
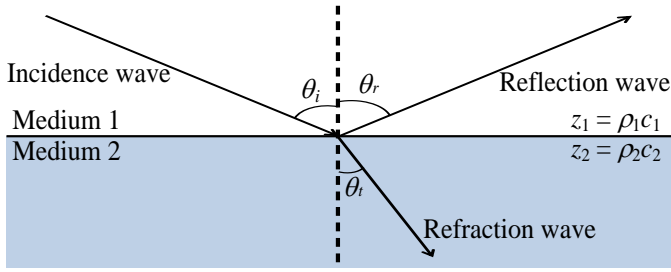
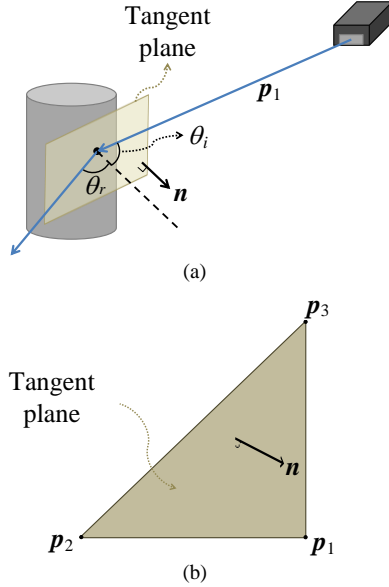Fig. 4. Reflection and refraction of acoustic wave.

(a)

(b)

Fig. 5. Incidence angle of direction vector p_1: (a) tangent plane is generated at point where direction vector p_1 collides with object, (b) normal vector of tangent plane is represented as n. p_2 and p_3 indicate direction vectors which are adjacent to p_1.

an object. The normal vector of the tangent plane n is expressed as follow (Fig. 5):

$$n = (p_3 - p_1) \times (p_2 - p_1) \qquad (9)$$

where p_1 is a direction vector which collides with an object, and p_2 is a direction vector which is located just to the left of p_1 with the same index of the elevation direction order. p_3 is a direction vector which is located just above of p_1 with the same index of the azimuth direction order. Therefore, an incidence angle θ_i of direction vector p_1 can be represented as follow:

$$\theta_i = \cos^{-1}\left(\frac{(-p_1) \cdot n}{|-p_1||n|} \right) \qquad (10)$$

Because all the parameters in (8) become known, the attenuation of sound pressure owing to reflection can be evaluated.

C. Acoustic Image

We described the acoustic image as gray-scale colors based on intensity values. First, we generated a raw acoustic image which has 1,000 rows in height and 128 columns in width. Each column line indicates the beam slice of an acoustic wave and each row line indicates the distance from the acoustic camera. As the direction vectors have their own values for azimuth angle and distance to an object, corresponding pixels can be defined based on these values. Then, the pixels can be mapped as gray-scale colors because the direction vectors have their own intensity values. Moreover, by setting maximum range value R_{max} for the most bottom row of the raw acoustic image and minimum range value R_{min} for the most upper row of that, length scale in the raw acoustic image can be converted. The raw acoustic image with gray-scale-colored is then converted to sector-formed acoustic image whose form is the same as real output image from acoustic cameras (Fig. 6).

IV. SIMULATION

The acoustic camera is a sensor which can be attached to ships, AUVs (autonomous underwater vehicles) or ROVs (remotely operated vehicles), which enables it can operate underwater tasks such as search, inspection, trajectory estimation, target recognition, SLAM keep moving. This means that it requires us to determine another coordinate system to describe the pose of them. Thus, in our simulation, we introduce two coordinate systems: world coordinate system and local coordinate system (Fig. 2). As represented as (4), coordinates given in local coordinate system can be converted to coordinates given in world coordinate system by using transformation matrix. In our simulator, because virtual space model including an object is represented in global coordinate system, all direction vectors considered as an acoustic wave are redescribed in world coordinate system.

We performed a simulation that generates an acoustic image by applying the model depicted in Section 2 and Section 3. On the basis of intensity values that the direction vectors have, corresponding pixels in sector-formed acoustic image are mapped in gray-scale, and the area whose intensities have no values is mapped in black, whereas the ground on which an object lies is mapped in gray.

In our simulation, we used the specifications of ARIS EXPLORER 3000 for the acoustic camera because we performed experiments with the instrument. The adopted specifications of ARIS EXPLORER 3000 in the experiments are shown in Table I [17]. In the acoustic-imaging model, the number of azimuth direction vectors was constructed based on the number of transducer beams which is 128. Its field of view in azimuth angle (32°) is resulting from the number of transducer beams and their beam width. It is applied to generate the central angle of the sector-formed acoustic image (Fig. 6 (c)). The values of parameters used in the experiments are shown in Table II. For calculating absorptive attenuation (5), we adopted f for 3.0×10^3 (3.0 megahertz), T for 20, D for 1.0×10^{-3} (1.0 meter). Thus, absorptive attenuation coefficient α is calculated as 2102.4 in decibels per kilometer. By applying this value to (6), α is calculated as 0.79 in p_0 per meter, which means the sound pressure attenuates to 0.79 times of p_0 while
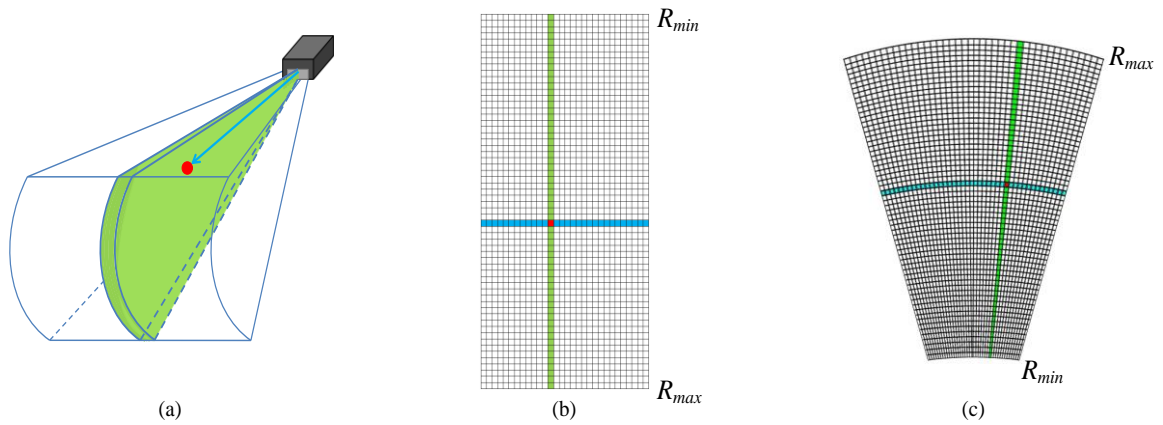
Fig. 6. Generating of a sector-formed acoustic image: (a) acoustic projection model, (b) raw acoustic image, and (c) sector-formed acoustic image.

TABLE I.   SPECIFICATIONS OF ARIS EXPLORER 3000 [17]

Specifications of ARIS EXPLORER 3000	
Identification frequency	3.0 MHz
Identification range	5 m
Number of transducer beams	128
Beam width	0.25°
Field-of-view	32°×14°

TABLE II.   VALUES OF PARAMETERS USED IN EXPERIMENTS

Equation	Parameter	Value	Unit
(4)	f	3.0×10^3	kHz
	T	20	°C
	D	1.0×10^{-3}	km
	α	2102.4	dB/km
		0.79	0.79 p_0/meter (p_0: reference sound pressure)
(5)	η	0.79	none

traveling one meter. As a result, the constant value η in (8) is 0.79. In the proposed simulator, the initial value of sound pressure (p_0) which is originated from the acoustic camera is set to arbitrary value. Each direction vector after it is reflected from an abject has lower intensity value than the initial value. Then, minimum value and maximum value of intensity data are extracted, and the coverage of intensity values is converted to gray-scale which is used to map the acoustic image. This method is based on the way the acoustic cameras actually display the output acoustic images.

V. EXPERIMENTAL RESULTS

We performed experiments to prove the validity of the proposed acoustic-imaging model. We used cylinder (10.5D × 22H centimeters), cone (16D × 21H centimeters), and cuboid (7D × 12W × 22H centimeters) plaster figures. Fig. 7 shows the plaster figures used in the experiments and Fig. 8 shows the experimental environment. The plaster figure was hung upside down with being fixed to aluminum board, and the acoustic camera was located lower than plaster figure so that it could insonify an acoustic wave to the plaster figure diagonally. Fixing an object to aluminum board is effective to confirm the appearance of acoustic shadow.

The experimental results are shown in Fig. 9. Figs. 9 (a) and (b) show the real acoustic images and simulated images generated from the proposed simulator, respectively. By comparing the simulated images with the real acoustic images, it is confirmed that the proposed simulator generates realistic acoustic images represented as complete gray-scale. The optical images of used plaster figures are arranged in Fig. 7 and it is possible to clarify the shape of targets used in experiments and simulation.

VI. CONCLUSION

In this paper, we developed a novel simulator to generate the realistic acoustic images. The contribution of our research is that we dealt successfully with the problem on state-of-the-art technique which described acoustic images by two colors: white for reflected acoustic wave from an object and black for acoustic wave which is not reflected. In this paper, by applying the proposed models, it is confirmed that our simulator generates the realistic acoustic images in complete gray-scale colors. We evaluated the proposed simulation by comparing the simulated acoustic images with real acoustic images and qualitative comparison indicates that the proposed simulation generates the acoustic images as complete gray-scale colors.

However, we only considered an absorptive attenuation and reflection as acoustic nature, which has a limitation on describing real environments. In real environments, there are a plenty of physical characteristics related to sound transmission such as noise from water surface, bubbles and marine organisms. To describe acoustic images as if it is generated in real environments such as turbid waters, noise factors should also be considered in acoustic-imaging model. This remains an unsolved problem. For the next study, acoustic-imaging model will be examined considering noise factors.
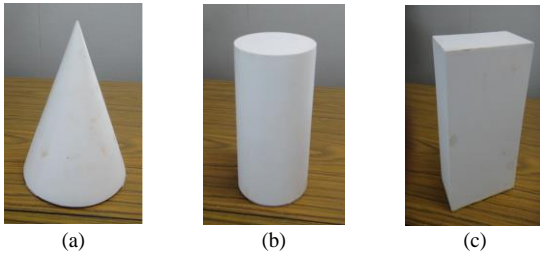
Fig. 7. Plaster figures used in the experiment: (a) cone (16D × 21H cm), (b) cylinder (10.5D × 22H cm), and (c) cuboid (7D × 12W × 22H cm).
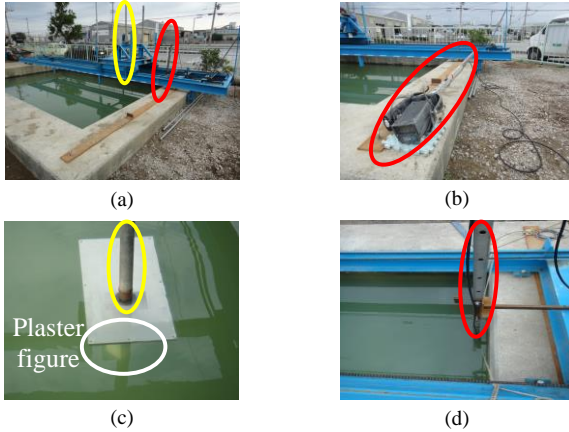

Fig. 8. Experimental environment: (a) experimental tank and robotic positioning system used for experiment, (b) ARIS EXPLORER 3000 with being fixed to steel bar, (c) plaster figure was hung upside down with being fixed to aluminum board, and (d) steel bar was fastened to edge of the tank.

Fig. 9. Experimental results: (a) experimental acoustic images and (b) simulated acoustic images.

REFERENCES

[1] R. Eustice, H. Singh, J. Leonard, M. Walter, and R. Ballard, "Visually navigating the RMS titanic with SLAM information filters," *Proceedings of the Robotics: Science and Systems*, pp. 57-64, 2005.

[2] O. Pizarro, R. Eustice, and H. Singh, "Large area 3D reconstructions from underwater surveys," *Proceeding of the 2004 MTS/IEEE OCEANS Conference and Exhibition*, Vol. 2, pp. 678-687, 2004.

[3] N. Hurtos, N. Palomeras, S. Nagappa, and J. Salvi, "Automatic detection of underwater chain links using a forward-looking sonar," in *OCEANS'13 MTS/IEEE*, 2013.

[4] D. Ribas, P. Ridao, J. D. Tardos, and J. Neira, "Underwater SLAM in man-made structured environments," *Journal of Field Robotics*, Vol. 25, No. 8, pp.898-921, 2008.

[5] E. Hernandez, P. Ridao, D. Ribas, and A. Mallios, "Probabilistic sonar scan matching for an AUV," *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 255-260, 2009.
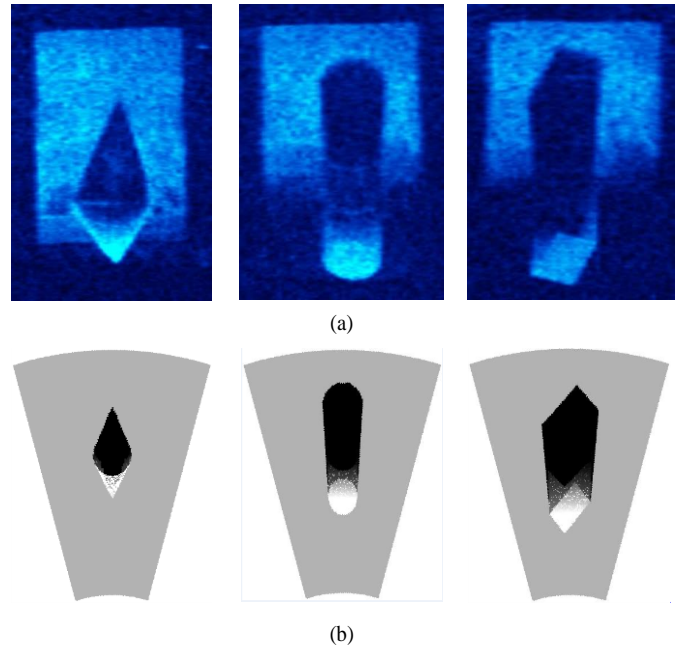
[6] A. Mallios, P. Ridao, E. Hernandez, D. Ribas, F. Maurelli, and Y. Petillot, "Pose-based SLAM with probabilistic scan matching algorithm using a mechanical scanned imaging sonar," in *IEEE OCEANS-EUROPE*, 2009.

[7] E. Belcher, W. Hanot, and J. Burch, "Dual-frequency identification sonar (DIDSON)," *Proceedings of the 2002 IEEE International Symposium on Underwater Technology*, pp. 187-192, 2002.

[8] "Aris," 2015, retrieved March 10, 2015, from http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000.

[9] "Blueview," 2015, retrieved March 10, 2015, from http://www.blueview.com/products/2d-imaging-sonar/p900-2250-dual-frequency.

[10] C. Xu, A. Asada, and K. Abukawa, "A method of generating 3D views of aquatic plants with DIDSON," *2011 IEEE Symposium on Underwater Technology and 2011 IEEE Workshop on Scientific Use of Submarine Cables and Related Technologies*, 2011.

[11] S. C. Yu, D. J. Kang, J. W. Choi, and M. K. Park, "Development of 3D visual underwater landmark recognition method for AUVs," *Proceedings of the 2009 International Offshore and Polar Engineering Conference*, pp. 677-682, 2009.

[12] N. Brahim, D. Gueriot, S. Daniel, and B. Solaiman, "3D reconstruction of underwater scenes using image sequences from acoustic camera," in *IEEE OCEANS*, 2010.

[13] N. Brahim, D. Gueriot, S. Daniel, and B. Solaiman, "3D reconstruction of underwater scenes using DIDSON acoustic sonar image sequnces through evolutionary algorithms," *2011 IEEE OCEANS*, 2011.

[14] J. H. Gu, H. G. Joe, and S. C. Yu, "Development of image sonar simulator for underwater object recognition," *2013 MTS/IEEE OCEANS*, 2013.

[15] M. A. Ainslie, and J. G. McColm, "A simplified formula for viscous and chemical absorption in sea water," *The Journal of the Acoustical Society of America*, Vol. 103, No. 3, pp.1671-1672, 1998.

[16] J. W. S. B. Rayleigh, "The theory of sound," *Dover Publications Inc., New York*, Vol. 2, 1945.

[17] "Specifications of ARIS EXPLORER 3000," 2015, retrieved April 24, 2015, from http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000/ARIS-3000-Product-Specs-English.