

Acoustic Image Simulator Based on Active Sonar Model in Underwater Environment*

Ngoc Trung Mai¹, Yonghoon Ji², Hanwool Woo¹, Yusuke Tamura¹, Atsushi Yamashita¹, and Hajime Asama¹

Abstract—Underwater tasks such as maintenance, inspection, target recognition, or simultaneous localization and mapping (SLAM) require accurate underwater information. Acoustic cameras are outstanding sensors for acquiring underwater information because, even in turbid water, they can provide acoustic images with more accurate detail than what other sensors provide. In this paper, we propose a novel acoustic image simulator based on active sonar model analyzing the correlation between signal processing and an image display mechanism that have not yet been clarified. The results demonstrate that our proposed simulator can successfully generate realistic virtual acoustic images from arbitrary viewpoints.

I. INTRODUCTION

Acoustic cameras have outstanding ability in their visibility even in dark or turbid water. Hence, important contributions in underwater exploration can be made by equipping an underwater robot with an acoustic camera. However, due to the high cost of acoustic cameras, research on sensing of a 3D environment and estimation of a robot pose, using this sensor are still in their infancy [1], [2]. Moreover, underwater experiments for validation are extremely difficult. Therefore, it is hard to start research on underwater, and even if a good theory is proposed, it is very difficult to validate. Therefore, to construct a simulation environment is important in order to develop research environment on underwater measurement using an acoustic camera.

A simulator for acoustic camera has been developed by Gu et al. [5]. However, the previous simulators can only represent acoustic image with binary values for each pixel (i.e., black and white). This is because that the models presenting the principles of acoustic cameras were not accurate. A simulator with the gray-scale from 0 to 255 has been developed by Kwak et al. [6]. However, the models developed to represent the principles of acoustic cameras were not accurate for the case of acoustic cameras and the unit used in the calculation formula were not right for the normal sonar systems. Firstly, about the proposed reflection model in this paper is the specular reflection, in which the incident acoustic wave is reflected along an angle symmetrical to the incidence angle and away from the direction of the arriving

signal. However, this model is only true when the object has ideal smooth surface. Obviously, the projector and receiver must be physically separated in order to receive this specular echo or only when the sound direction is perpendicular to the object. This is not true for the case of acoustic cameras, which has the receiver collocated with the source. Secondly, the unit used in the calculation formula for calculating the intensity of the of sound is not in decibels (dB), which is usually used to express the quantities related to the level of sound. In addition, Sac et al. assumed that target objects have only continuous surfaces [7] and Cerqueira et al. applied GPU-based approach to develop each 2D sonar simulator [8]. However, physical models applied for these simulators are also incomplete.

In order to fill the research gaps and the needs of previous work, this paper presents a novel model which employs acoustic camera images and uses physical properties of underwater waves. The remainder of this paper is organized as follows. Sections II and III introduce properties of an acoustic camera and physical properties of sonar, respectively. Then, process of acoustic image generation is presented in Section IV. The validity of the proposed acoustic image simulator is evaluated with the simulation results in Section V. Finally, Section VI gives conclusions of this paper.

II. PROPERTIES OF ACOUSTIC CAMERA

A. Acoustic Projection Model

An acoustic camera generates an acoustic image by transmitting acoustic waves in a 3D area space. The acoustic waves generated by the acoustic camera travel through water to the object, and returned as acoustic echoes. The acoustic camera receives reflected waves from the object, then calculates the power of the reflected waves, and reports

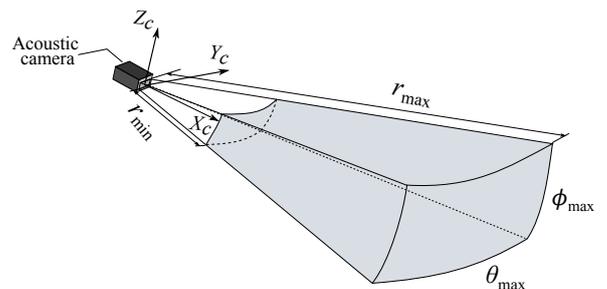


Fig. 1. Geometrical model of the acoustic camera image: Sensing range of the acoustic camera is determined by the maximum measuring range r_{\max} , minimum measuring range r_{\min} , azimuth angle θ_{\max} , and elevation angle ϕ_{\max} .

*This work was in part supported by Tough Robotics Challenge, ImpACT Program (Impulsing Paradigm Change through Disruptive Technologies Program).

¹N. T. Mai, H. Woo, Y. Tamura, A. Yamashita, and H. Asama are with the Department of Precision Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan {mai, woo, tamura, yamashita, asama}@robot.t.u-tokyo.ac.jp

²Y. Ji is with the Department of Precision Mechanics, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan ji@mech.chuo-u.ac.jp

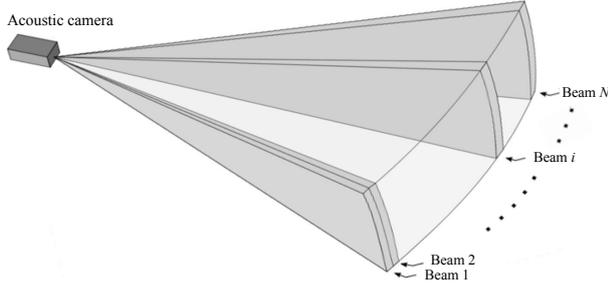


Fig. 2. Illustration of beam slices in the acoustic camera. The acoustic wave is processed by N array of transducers as a function of azimuth angle.

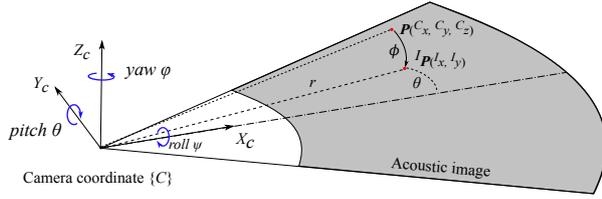


Fig. 3. Imaging sonar geometry: The values of the range r and azimuth angle θ can be derived from the pixel coordinate of acoustic image, while the elevation angle ϕ is missing.

it at the pixel corresponding to the direction of the reflected waves on a power map.

The field of view (FoV) of the acoustic camera is determined by the maximum measuring range r_{\max} , minimum measuring range r_{\min} , azimuth angle θ_{\max} , and elevation angle ϕ_{\max} , as shown in Fig. 1. An acoustic camera has N transmitter elements side by side. A thin acoustic beam is oscillated from each of N transmitter elements, as shown in Fig. 2. The returned echoes after reflected at the target are received by N wave receiving elements. The acoustic waves reflected from the object are processed as a function of only the range r and the azimuth angle θ , not related to the elevation angle ϕ . In other words, different points in the 3D sensing area with same range r and same azimuth angle θ are mapped at same pixel on the 2D acoustic image.

B. Imaging Acoustic Geometry

Although an acoustic camera senses a 3D area, the output of the sensing process is a 2D acoustic image. As shown in Fig. 3, the acoustic image provides the range r and the azimuth angle θ , while the elevation angle ϕ is missing. Thus, each point (r, θ, ϕ) which is representation for polar coordinate system of (x, y, z) in the 3D sensing area is mapped to (r, θ) in the 2D image. For example, a 3D point $P^{(C_x, C_y, C_z)}$ in the camera FoV is mapped at $I^P(r, \theta)$ in the 2D acoustic image, as shown in Fig. 3. Here, superscripts C and I respectively indicate camera coordinate system and image coordinate system. The conversion of Cartesian coordinates and spherical coordinates is given as follows:

$$\begin{bmatrix} C_x \\ C_y \\ C_z \end{bmatrix} = \begin{bmatrix} r \cos \phi \cos \theta \\ r \cos \phi \sin \theta \\ r \sin \phi \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} r \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{C_x^2 + C_y^2 + C_z^2} \\ \tan^{-1}(C_y/C_x) \\ \tan^{-1}(C_z/\sqrt{(C_y^2 + C_x^2)}) \end{bmatrix}. \quad (2)$$

Next, the conversion of image coordinates and spherical coordinates is given as follows:

$$\begin{bmatrix} I_x \\ I_y \end{bmatrix} = \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix}. \quad (3)$$

Figure 4 shows a real acoustic image. From the acoustic image, the range r and the azimuth angle θ are obtained on the polar coordinate system; however, unfortunately, the elevation angle ϕ is lost. Therefore, it is generally impossible to retrieve the 3D information of objects by using a single acoustic image. To solve this problem, our research group have developed a theoretical methodology to recover 3D information by using multi-acoustic images from different viewpoints [2], [3], [4]. In order to manage such studies more efficiently, it is essential to develop a simulator that works correctly based on the principle of an acoustic camera.

III. PHYSICAL PROPERTIES OF SONAR

There are two types of the sonar systems, passive sonar system and active sonar system. In a passive sonar system, sound propagates directly from a source to a receiver. In an active sonar system, the sound is transmitted from a source, reflects off by an object, and returns to a receiver as shown in Fig. 5. An acoustic camera is an active sonar system in which the receiver is collocated with the source. This section introduces the active sonar equation for the active sonar system.

In an active sonar system, the sound is generated by one of the system components called the projector. The sound waves generated by the projector travel through the water to the target and return as sonar echoes to the hydrophone, which converts sound into electricity. The electric output of the hydrophone is amplified and processed in various ways, and finally, the output data of the system is generated. In the case of acoustic cameras, the output is an acoustic image.

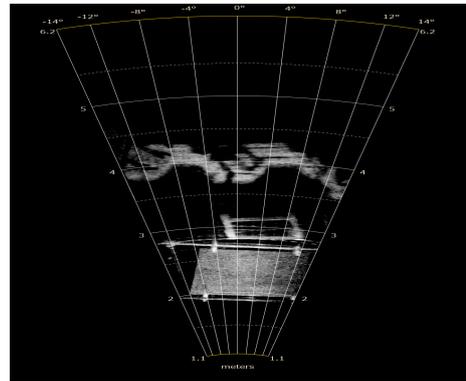


Fig. 4. Acoustic image taken by a real acoustic camera. The range r and the azimuth angle θ are obtained; however, unfortunately, the elevation angle ϕ is lost.

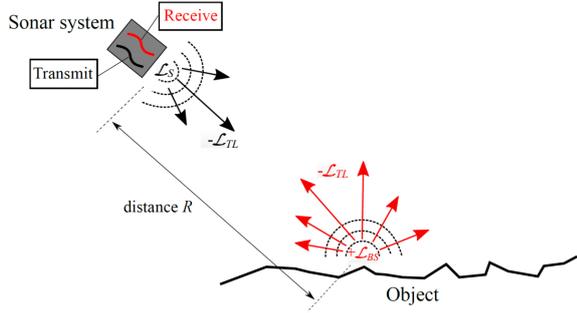


Fig. 5. An active sonar system.

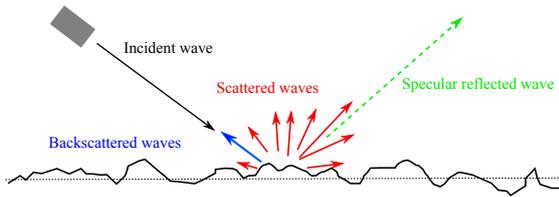


Fig. 6. Reflection model.

A. Sonar Equation

The equation computes the received echo-to-noise ratio from the transmitted signal level, taking consideration of transmission loss, noise level, sensor directivity, and target strength [10], [11]. The active sonar equation is

$$\mathcal{L}_{EN} = \mathcal{L}_S - 2\mathcal{L}_{TL} - \mathcal{L}_N + \mathcal{L}_{TS}, \quad (4)$$

where:

- The echo-to-noise ratio \mathcal{L}_{EN} is the received echo-to-noise ratio in dB.
- The source level \mathcal{L}_S is the ratio of the transmitted intensity from the source to a reference intensity, converted to dB.
- Transmission loss \mathcal{L}_{TL} is the attenuation of sound intensity as the sound propagates through the underwater, converted to dB.
- Target strength \mathcal{L}_{TS} is the ratio of the intensity of a reflected signal at 1 m from a target to the incident intensity, converted to dB.
- Noise level \mathcal{L}_N is the ratio of the noise intensity at the receiver to the same reference intensity used for source level.

The \mathcal{L}_{EN} is calculated by subtracting the round trip propagation loss $2\mathcal{L}_{TL}$ from the \mathcal{L}_S to the object and the \mathcal{L}_N and further subtracting the seafloor scattering intensity \mathcal{L}_{TS} which is often negative value. Here, the unit used in the sonar equation is dB because it is common to express certain of the above quantities as levels using dB in acoustics. A level is a method of expressing the magnitude of a quantity as a logarithmic ratio to a reference value. The decibel uses logarithms to base 10.

As mentioned above, the sonar equation is conventionally written using the dB (logarithmic) scale in terms of sound

levels. The most common form of these (i.e., each element in (4)) is the intensity level \mathcal{L} , defined in dB as:

$$\mathcal{L} = 10 \log_{10} \left(\frac{I}{I_{\text{ref}}} \right), \quad (5)$$

where I_{ref} is a reference intensity. In many cases the reference value for acoustic pressure is $I_{\text{ref}} = 1 \mu\text{Pa}$. In this case, the pressure values is also can be expressed in (dB re. $1 \mu\text{Pa}$). The decibel unit provides a convenient way to handle large changes in variables. In addition, it permits quantities to be multiplied together simply by adding their decibel equivalents in the sonar equations.

B. Transmission Loss

As the acoustic wave leaves the projector and travels out into the water, the amplitude of the pressure signal attenuates for two reasons: geometrical spreading of the sound wavefront and absorption of the sound during propagation. Firstly, the acoustic wave expands as a spherical wave in a homogeneous medium centered on the sound source. The acoustic intensity I decreases with range R in inverse proportion to the diffusion surface of the sphere $4\pi R^2$ in homogeneous media. Therefore, the attenuation is proportional to the square of the distance R^2 . Secondly, attenuation is due to the conversion of acoustic energy into thermal energy as a result of the chemical reaction that sound pressure has on the ionic structure of boric acid and magnesium sulfate in seawater. The amount of attenuation by sound waves depends on distance, acoustic frequency, water temperature, and pressure. Hence, the total one-way propagation attenuation is as follows:

$$\begin{aligned} \mathcal{L}_{TL} &= 10 \log_{10} R^2 + \alpha R \\ &= 20 \log_{10} R + \alpha R, \end{aligned} \quad (6)$$

where, α defined in dB/m is absorptive attenuation coefficient, as follows:

$$\alpha = 0.49 f^2 \exp \left(-\frac{T}{27} + \frac{D}{17} \right). \quad (7)$$

The absorptive attenuation coefficient α is determined by the acoustic frequency f in kHz, the water temperature T in Celsius, and water depth D in km (i.e., the pressure depends on the water depth) [9]. For two way propagation, the acoustic wave expands as a spherical wave to the reflector. Then, the reflector spreads the signal in all directions and the reflected field expands as a spherical wave back to the receiver. In a homogeneous medium, the two-way loss becomes $2\mathcal{L}_{TL}$.

C. Reflection Model

When the sonar signal comes to the surface of the target, it is reflected in many directions as shown in Fig. 6. The level of each reflection depends on the characteristics of the object surface and the frequency of the sound waves. The echo sounder is mainly related to the sound intensity in the backscatter direction. This backscatter gives rise to the

acoustic echo, whose presence is the basic working principle for an acoustic camera, which has a co-located projector and receiver. Although backscatter is often low compared to the incident intensity, this returned echo proves to be still detectable and measurable.

The \mathcal{L}_{TS} , defined by the ratio of the intensity of a reflected signal at 1 m from a target to the incident intensity, is used to indicate how much of the sound wave is scattered toward the sonar receiver. In general, the backscattered intensity depends on the target properties as well as the sound wave itself. The \mathcal{L}_{TS} is defined as follows:

$$\begin{aligned}\mathcal{L}_{TS} &= 10 \log_{10}(S_{BS}A) \\ &= \mathcal{L}_{BS} + 10 \log_{10} A,\end{aligned}\quad (8)$$

where $\mathcal{L}_{BS} = 10 \log_{10} S_{BS}$ is the backscattering strength in dB and A is the insonified area on the target [10], [12].

In the case of low roughness surface, the majority of the scattered waves are still concentrated around the specular direction, and small parts are generated in the direction returning to the projector. Conversely, in the case of high roughness surface, the intensity of the incident wave is significantly scattered in all directions. As roughness increases, it becomes more evenly scattered in all directions.

When pictured by a sonar system, the interface roughness has to be considered in relation to the acoustic wavelength. Hence, acoustic backscatter may be interpreted in terms of the acoustic roughness, which is defined as the ratio of the geometrical roughness to the acoustic wavelength. Here, the generic seafloor acoustic backscatter (GSAB) model is introduced to represent the \mathcal{L}_{BS} . The GSAB is affordable-approach and has been used in various studies [13], [14], [15]. The GSAB model expresses the \mathcal{L}_{BS} as a combination of Gaussian law for specular angle and Lambert law for the incident angle. The model is given by:

$$\mathcal{L}_{BS} = 10 \log_{10} \left[\mu \exp\left(-\frac{\vartheta^2}{2h^2}\right) + \nu \cos^q \vartheta \right], \quad (9)$$

where ϑ denotes the incident angle. In this model, four parameters μ , h , ν , and q are defined depending on the characteristics of the object surface and the frequency of the sound waves. The first term $\mu \exp(-\vartheta^2/2h^2)$ represents the specular reflection effect. μ is related to the coherent reflection coefficient at the interface. It is high for smooth interfaces. h quantifies the angular range of the specular area. Note that this term can be omitted for the development of the simulator because it has little effect compared to the second term. The second term $\nu \cos^q \vartheta$ is a generalized variant of Lambert's law. ν and q are the offsets associated with Lambert's law which describes the \mathcal{L}_{BS} at the mid-angle of the rough interface. ν is proportional to the frequency of the sonar wave, the roughness of the interface, and the impedance. q is the backscatter angular decrement, commanding the fall-off at grazing angles. It is high for soft and smooth interfaces. Typically, $10 \log_{10} \nu$ ranges from -20 to -50 dB and q is normally equal to 1 or 2 [13]. Consequently,

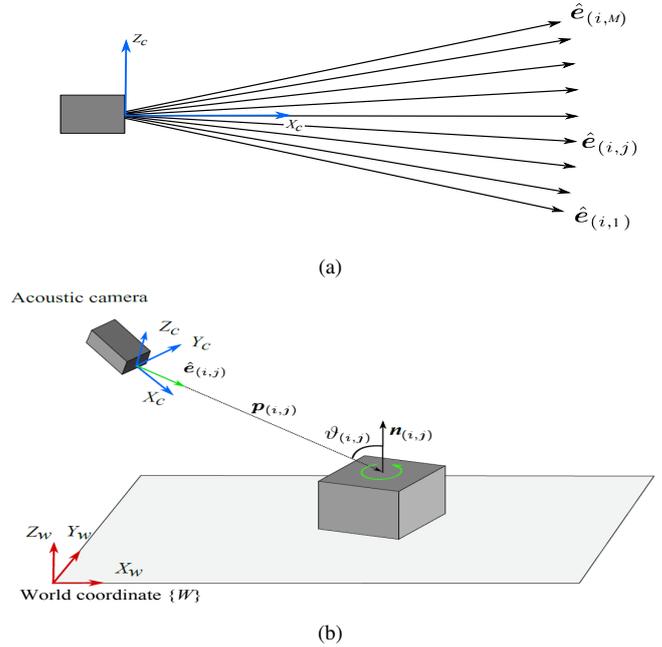


Fig. 7. Acoustic projection model and imaging sonar geometry: (a) M vectors to present the sound waves in each acoustic beam and (b) the relation between the direction vector and the object.

when the values of μ , h , ν , and q are fixed, the value \mathcal{L}_{BS} is only dependent on the incidence angle ϑ .

IV. ACOUSTIC IMAGE GENERATION

This section presents the overview of acoustic image generation process. Based on the novel model which employs acoustic cameras images, physical properties of underwater waves introduced in Sections II and III, and uses the acoustic camera-imaging model proposed in Kwak et al. [6], a 2D imaging sonar simulator was developed.

A. Acoustic Camera-Imaging Model

The main idea is using vectors to present the sound waves. Each vector represents for a sound wave, travel to an object, reflected and returns to the acoustic camera. Then, the value of the returned acoustic echo is calculated by using the sonar equation (4) introduced in previous section. Next, the 2D acoustic image is generated based on the calculated value of the returned acoustic echo.

We generate N rows of vectors to describe N acoustic beams, as shown in Fig. 2. In each beam, M unit direction vectors \hat{e} are also generated to describe acoustic waves, as shown in Fig. 7(a). Totally, there are $N \times M$ direction vectors are generated. In the azimuth direction, there are N rows of vectors with n radian intervals corresponding with N acoustic beams. In each row, there are M vectors for the elevation direction with m radian intervals. Each unit direction vector can be represented in camera spherical coordinates as follows:

$$\hat{e}_{(i,j)} = \begin{bmatrix} r_e \\ \theta_i \\ \phi_j \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\theta_{\max} - n(2i+1)}{2} \\ \frac{\phi_{\max}}{2} - mj \end{bmatrix}, \quad (10)$$

where $\hat{e}_{(i,j)}$ consisting of normalized range r_e , azimuth angle θ_i , and elevation angle ϕ_j denotes the j^{th} direction vector of the i^{th} beam. Sensing vectors \mathbf{p} corresponding to the direction vectors \hat{e} are also generated. As shown in Fig. 7(b), each sensing vector $\mathbf{p}_{(i,j)}$ is emitted from the acoustic camera and proceeds until it hits the object. In order to present pose relations between these directional vectors and underwater objects, it is necessary to describe the vectors in the same world Cartesian coordinate as the objects. Each direction vector can be converted to the world Cartesian coordinates in two steps. First, the coordinate of each direction vector $\hat{e}_{(i,j)}$ can be converted from the camera spherical coordinates to the camera Cartesian coordinates as follows:

$$\begin{bmatrix} Cx_{(i,j)} \\ Cy_{(i,j)} \\ Cz_{(i,j)} \end{bmatrix} = \begin{bmatrix} r_e \cos \phi_j \sin \theta_i \\ r_e \cos \phi_j \cos \theta_i \\ r_e \sin \phi_j \end{bmatrix}, \quad (11)$$

where $(Cx_{(i,j)}, Cy_{(i,j)}, Cz_{(i,j)})$ is the coordinate of the direction vector $\hat{e}_{(i,j)}$ in the camera Cartesian coordinate. Next, from the camera Cartesian coordinate, the coordinate of each direction vector can be transformed to the world Cartesian coordinate as follows:

$$\begin{bmatrix} x_{(i,j)} & y_{(i,j)} & z_{(i,j)} & 1 \end{bmatrix}^T = \begin{bmatrix} c_\theta c_\varphi & -c_\psi s_\varphi + c_\varphi s_\theta s_\psi & s_\psi s_\varphi + c_\psi c_\varphi s_\theta & x \\ c_\theta s_\varphi & c_\psi c_\varphi + s_\theta s_\psi s_\varphi & -c_\varphi s_\psi + c_\psi s_\theta s_\varphi & y \\ -s_\theta & c_\theta s_\psi & c_\theta c_\psi & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cx_{(i,j)} \\ Cy_{(i,j)} \\ Cz_{(i,j)} \\ 1 \end{bmatrix}, \quad (12)$$

where $(x, y, z, \psi, \theta, \varphi)$ indicates the six degrees of freedom (6-DOF) pose of the acoustic camera, and $(x_{(i,j)}, y_{(i,j)}, z_{(i,j)})$ is the coordinate of the direction vector $\hat{e}_{(i,j)}$ in the world Cartesian coordinate. Here, s and c represent sine and cosine functions, respectively.

As described in (9), the intensity of reflected signal is a function on the incidence angle ϑ . Therefore, to calculate the backscatter intensity of the reflected waves, it is necessary to calculate the incidence angles of each direction vector. The incident angle corresponding to each direction vector is calculated by a tangent plane at the measurement point. The normal vector \mathbf{n} of the tangent plane as shown in Fig. 7 is represented as follows:

$$\mathbf{n}_{(i,j)} = (\mathbf{p}_{(i-1,j)} - \mathbf{p}_{(i,j)}) \times (\mathbf{p}_{(i,j-1)} - \mathbf{p}_{(i,j)}), \quad (13)$$

where $\mathbf{p}_{(i,j)}$, $\mathbf{p}_{(i-1,j)}$, and $\mathbf{p}_{(i,j-1)}$ are sensing vectors which hits underwater object corresponding with the direction vector $\hat{e}_{(i,j)}$, $\hat{e}_{(i-1,j)}$, and $\hat{e}_{(i,j-1)}$, respectively. Then, the incidence angle $\vartheta_{(i,j)}$ of the sensing vector $\mathbf{p}_{(i,j)}$ is calculated as follows:

$$\vartheta_{(i,j)} = \arccos \left(\frac{-\mathbf{p}_{(i,j)} \cdot \mathbf{n}_{(i,j)}}{\|\mathbf{p}_{(i,j)}\| \|\mathbf{n}_{(i,j)}\|} \right). \quad (14)$$

Since the incidence angle $\vartheta_{(i,j)}$ can be calculated by (14), the backscattering intensity of the wave due to the reflection can be evaluated. The sensing vectors reflected from the underwater object return to the acoustic camera with having information on which orientation they travelled, how far they travelled, and acoustic pressure. Therefore, returned direction

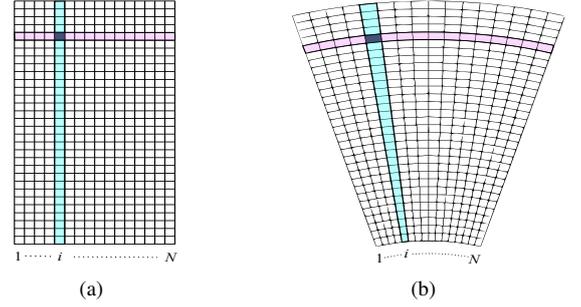


Fig. 8. Generation process of a sector-formed acoustic image: (a) a raw acoustic image and (b) a sector-formed acoustic image.

vectors have enough information to represent the acoustic image given that it needs range, azimuth angle, and acoustic pressure values to be mapped.

B. Acoustic Image Structures

Since elements necessary for generating an acoustic image have been prepared, the value of each returned acoustic echo \mathcal{L}_{EN} can be calculated by using the sonar equation (4) introduced in Section III. These values are used to determine the gray-scale to be mapped in each pixel of simulated acoustic images. Here, we calculate each pixel value to be proportional to the \mathcal{L}_{EN} for mapping.

Figure 8 shows the image generation process. First, a quadrangular form raw acoustic image is generated based on the data derived from the camera-imaging model and physical properties described above as shown in Fig. 8(a). Next, the raw acoustic image mapped in gray-scale is converted to a sector-formed acoustic image having the same shape as the real acoustic image, as shown in Fig. 8(b). Each column line indicates the beam slice of an acoustic wave and each row line indicates the distance from the acoustic camera to the measuring point. As the direction vectors have their own azimuth angle and travelled distance values, corresponding pixels are determined according to these values and mapped in scale-converted gray colors in reference to their corresponding diminished acoustic pressure values.

V. SIMULATION OF ACOUSTIC IMAGES

We developed a 2D imaging sonar simulator based on the novel model which employs properties of the acoustic camera and physical properties of underwater waves introduced in Sections II and III. In order to implement the realistic sonar simulator, the specifications of ARIS EXPLORER 3000 which is a state-of-the-art acoustic camera are applied, as listed in Table I. Based on these specifications, parameters shown in Table II are used to implement the simulator. Note that detailed descriptions of these parameters can be found in Sections III and IV.

In order to verify the validity of the simulator, first of all, we took acoustic images from ARIS EXPLORER 3000 in a real underwater environment, as shown in Fig. 9(a). A square prism (70 D \times 120 W \times 220 H mm³) was used as a measurement object. Next, as shown in Fig. 9(b),

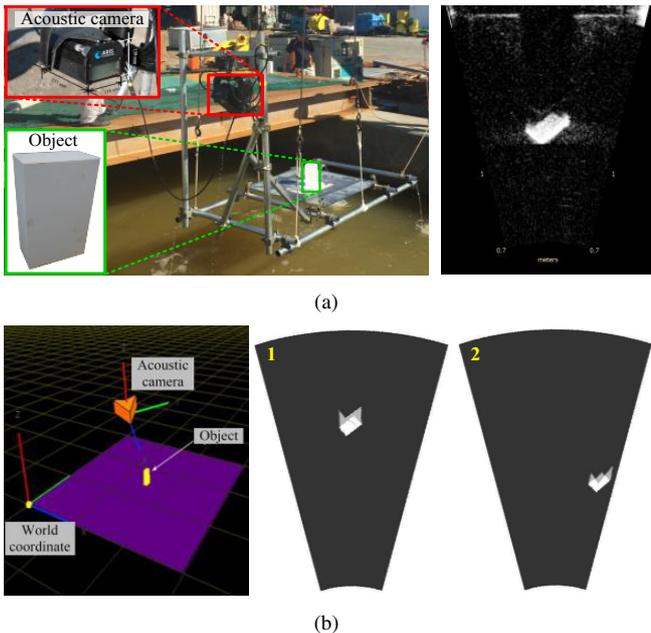


Fig. 9. Experiments for verification: (a) example of a captured real acoustic image using an acoustic camera ARIS EXPLORER 3000 in an underwater environment and (b) a simulation environment with a measurement object located about 1.5 m from the X_w and Y_w axes, respectively and simulated acoustic images from camera pose 1 (0.34 m, 2.42 m, 0.90 m, -0.97 deg, 29.79 deg, -38.96 deg) and camera pose 2 (0.84 m, 2.42 m, 0.90 m, 6.88 deg, 28.07 deg, -41.25 deg).

we simulated several acoustic images from different camera poses by constructing a simulation environment with almost similar conditions. The results showed that the proposed simulator was able to simulate acoustic images almost similar to actual images from arbitrary camera poses. The time required to generate the acoustic images was about 0.8–1.3 s.

VI. CONCLUSION

In this paper, we developed a novel sonar simulator to generate the realistic acoustic image from arbitrary view-points of camera poses. The simulator was implemented based on the active sonar model that is most suitable model for the principle of the acoustic camera. The validity of the simulated images were investigated through comparison with images captured in a real underwater environment.

ACKNOWLEDGMENT

The authors would like to thank S. Fuchiyama, A. Ueyama, K. Okada and their colleagues at KYOKUTO Inc. for full access to their facilities for the real experiment. We would also like to thank Y. Yamamura, F. Maeda, S. Imanaga at TOYO Corp. who helped in this research project over a day for the equipment set up and the acquisition of the sonar data.

REFERENCES

[1] T. A. Huang and M. Kaess, "Towards acoustic structure from motion for imaging sonar," in *Proc. 2015 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2015, pp. 758–765.

TABLE I
SPECIFICATIONS OF ARIS EXPLORER 3000

Specification	Value	Unit
Identification frequency	3,000	kHz
Identification range	5	m
Azimuth angle θ_{\max}	32	deg
Elevation angle ϕ_{\max}	14	deg
Field of view (FoV)	32×14	deg \times deg
Number of transducer beams	128	-
Beam width	0.25	deg

TABLE II
IMPLEMENTATION DETAILS OF THE SIMULATOR

Parameter	Value	Unit
\mathcal{L}_S	230	dB
f	3,000	kHz
T	10	$^{\circ}\text{C}$
D	3	m
α	3.04	dB/m
\mathcal{L}_N	10	dB
$10 \log_{10} \nu$	-45	dB
q	2	-
N	128	-
M	281	-

[2] Y. Ji, S. Kwak, A. Yamashita, and H. Asama, "Acoustic camera-based 3D measurement of underwater objects through automated extraction and association of feature points," in *Proc. 2016 IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, 2016, pp. 224–230.

[3] N. T. Mai, H. Woo, Y. Ji, Y. Tamura, A. Yamashita, and H. Asama, "3-D reconstruction of underwater object based on extended Kalman filter by using acoustic camera images," *IFAC PaperOnLine*, vol. 50, no. 1, pp. 1043–1049, Jul. 2017.

[4] N. T. Mai, H. Woo, Y. Ji, Y. Tamura, A. Yamashita, and H. Asama, "3D reconstruction of line features using multi-view acoustic images in underwater environment," in *Proc. 2017 IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, 2017, pp. 312–317.

[5] J. H. Gu, H. G. Joe, and S. C. Yu, "Development of image sonar simulator for underwater object recognition," in *Proc. 2013 MTS/IEEE OCEANS*, 2013, pp. 1–6.

[6] S. Kwak, Y. Ji, A. Yamashita, and H. Asama, "Development of acoustic camera-imaging simulator based on novel model," in *Proc. 2015 IEEE 15th Int. Conf. Environment and Electrical Engineering*, 2015, pp. 1719–1724.

[7] H. Sac, K. Leblebicioglu, and G. Bozdagi, "2D high-frequency forward-looking sonar simulator based on continuous surfaces approach," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 23, pp. 2289–2303, Jan. 2015.

[8] R. Cerqueira, T. Trocoli, G. Neves, S. Joyeux, J. Albiez, and L. Oliveira, "A novel GPU-based sonar simulator for real-time applications," *Computers Graphics*, vol. 68, pp. 66–76, Jan. 2017.

[9] M. A. Ainslie and J. G. McColm, "A simplified formula for viscous and chemical absorption in sea water," *J. Acoustical Society of America*, vol. 103, no. 3, pp. 1671–1672, Jun. 1998.

[10] R. J. Urick, *Principles of Underwater Sound*, McGraw-Hill, 1983.

[11] R. E. Hansen, "Introduction to Sonar," Course Material to INF-GEO4310, University of Oslo, Oct. 2009.

[12] X. Lurton, G. Lamarche, C. Brown, V. L. Lucieer, G. Rice, A. Schimel, and T. Weber, "Backscatter measurements by seafloor mapping sonars: guidelines and recommendations," *A collective report by members of the GeoHab Backscatter Working Group*, pp. 1–200, May 2015.

[13] G. Lamarche, X. Lurton, A. L. Verdier, and J. M. Augustin, "Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—application to cook strait," *Continental Shelf Research*, vol. 31, pp. S93–S109, Feb. 2011.

[14] J. M. Augustin and X. Lurton, "Image amplitude calibration and processing for seafloor mapping sonars," in *Proc. OCEANS 2005-Europe*, 2005, pp. 698–701.

[15] L. Hellequin, X. Lurton, and J. M. Augustin, "Postprocessing and signal corrections for multibeam echosounder images," in *Proc. 1997 MTS/IEEE OCEANS*, 1997, pp. 23–26.