An Optical Image Transmission System for Deep Sea Creature Sampling Missions Using Autonomous Underwater Vehicle

An Optical Image Transmission System for Deep Sea Creature Sampling Missions Using Autonomous Underwater Vehicle

Jonghyun Ahn [ID], *Member, IEEE*, Shinsuke Yasukawa, Takashi Sonoda, Yuya Nishida, *Member, IEEE*, Kazuo Ishii, *Member, IEEE*, and Tamaki Ura, *Fellow, IEEE*

Abstract—The exploration of oceans using autonomous underwater vehicles (AUVs) is necessary for activities, such as the sustainable management of fishery resources, extraction of seafloor minerals and energy resources, and inspection of underwater infrastructure. As the next step in ocean exploration, AUVs are expected to employ end-effectors to make physical contact with seafloor creatures and materials. We propose a scenario for realizing a sampling mission using an AUV that is equipped to sample marine life. In this scenario, the sampling AUV observes the seafloor while concurrently transmitting the observed images to a surface vessel for inspection by the AUV operators. If the received images show an object of interest, the object is selected as a candidate of sampling target by the operators, who send a sampling command to the AUV. After receiving the command, the AUV returns to the target area and attempts to sample it. In this paper, we propose a system for transmitting images of the seafloor as part of the sampling-mission scenario. The proposed image transmission system includes a process for enhancing images of the deep seafloor, a process for selecting interesting images, and processes for compressing and reconstructing images. The image enhancement process resolves imaging problems resulting from light attenuation, such as color attenuation and uneven illumination. The process for selecting interesting images selects those that contain interesting objects, such as marine life. The selection process prevents the transmission of meaningless images that contain only flat sand on the seafloor. The proposed image compression method, which is based on color depth compression, reduces the amount of data. The combined process of selecting an interesting image and compressing it reduces various problems in acoustic communication, such as low information density and data loss. Instead of an overall image, part of an overall image is transmitted by a set of data packet, and each received data packet is reconstructed onboard the vessel. Because of image compression, the colors of a reconstructed image differ from those of an enhanced image. However, the reconstructed image contains similar colors, and the structural similarity index was found to be 91.4% by evaluating images that were subjected to a 4-b color compression. The proposed image transmission system was tested in the Sea of Okhotsk, and these tests were performed four times in different sea areas (minimum depth 380 m, maximum depth 590 m). The results show that the size of the data for a single image was reduced by a factor of 18 using the proposed image compression process, with each image taking 3.7 s to be transmitted via an acoustic modem (20 kb/s). Of the automatically selected images, 63% contained marine life, and the total transmission success rate was 22%.

Index Terms—Acoustic transmission system, AUV mission, deep sea AUV experiment, image compression, interesting image selection, seafloor optical imaging.

I. Introduction

THE many discoveries of minerals, energy resources, and marine species in the deep sea have attracted considerable interest from researchers in various fields. Underwater robotics is an important method of observing the deep sea ecosystem, with autonomous underwater vehicles (AUVs) having become useful tools for ocean exploration. An objective of most AUV missions is the optical imaging of parts of the seafloor that contain marine biotic communities. Such seafloor images contribute to research on deep sea ecosystems [1]–[5]. As the next step in ocean exploration, AUVs are expected to employ end-effectors to make physical contact with deep sea creatures and materials on the seafloor to sample them.

Conventional missions to sample deep sea creatures using underwater robotics have employed either human occupied vehicles (HOVs) or remotely operated vehicles (ROVs) [6]. However, using an HOV introduces the necessity of securing the safety of the crew, limiting the mission duration, and making it difficult to explore a wide area. ROVs also have a limited exploration range, owing to the umbilical cable that is necessary for communication, control, and the supply of electric power. Therefore, to sample deep sea creatures over a wide area for a long time, AUVs are required. Such AUV sampling missions are expected to contribute to marine research.

J. Ahn is with the Department of Human Intelligence Systems, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu 808-0196, Japan (e-mail: ahn@brain.kyutech.ac.jp).

S. Yasukawa, Y. Nishida, and K. Ishii are with the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu 808-0196, Japan (e-mail: s-yasukawa@brain.kyutech.ac.jp; ynishida@lsse.kyutech.ac.jp; ishii@brain.kyutech.ac.jp).

T. Sonoda is with the Department of Integrated System Engineering, Nishinippon Institute of Technology, Miyako-gun 800-0394, Japan (e-mail: sonoda@nishitech.ac.jp).

T. Ura is with the Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, Kitakyushu 808-0196, Japan (e-mail: ura@lsse.kyutech.ac.jp).
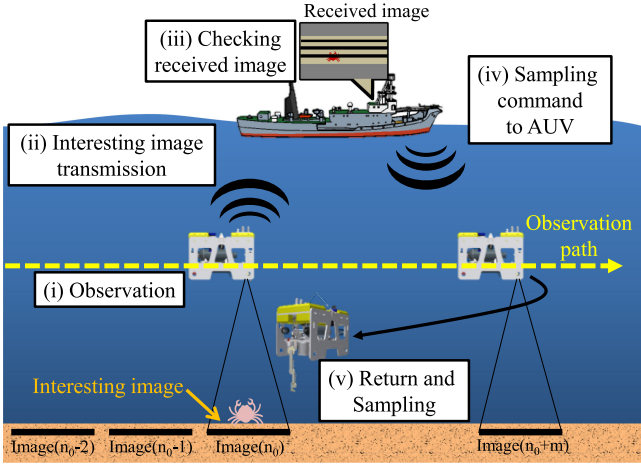
Fig. 1.   Overview of the proposed sampling mission scenario.

Fig. 2.   Process of the proposed image transmission system.

We refer to an AUV that is suitably equipped to perform sampling missions as a sampling-AUV (SAUV) [7], and we propose the SAUV mission scenario illustrated in Fig. 1, comprising the following steps.

i) The SAUV captures optical seafloor images chronologically along a preprogrammed observation path.

ii) If the present obtained image "Image(n_0)" contains an interesting object, then the SAUV transmits it.

iii) The operator on the vessel checks the received image.

iv) If the received image shows valuable marine life, then the operator sends sampling commands to the SAUV.

v) Finally, the SAUV returns to the specified area and attempts to sample the marine life there.

In this paper, we focus on steps i)–iii) in Fig. 1, and present a transmission system to realize the proposed mission scenario. We also present results from experiments in which the proposed transmission system was tested four times in different areas (minimum depth 380 m, maximum depth 590 m) in the Sea of Okhotsk to the north of Japan.

II. PROPOSED IMAGE TRANSMISSION SYSTEM

To realize a transmission system, we must resolve certain problems relating to imaging and acoustic communication. The imaging problems are a result of light attenuation (specifically, different attenuation rates for each color) and uneven illumination. Meanwhile, acoustic communication has a low information density, and data are lost during communication because of the positional relationship between the SAUV and the vessel, thruster noise, and other acoustic signals.

The process of the proposed image transmission system is illustrated in Fig. 2. The SAUV acquires optical images of the seafloor, enhances them, selects interesting ones, compresses these, and transmits them to the vessel. On the vessel, the received data are reconstructed as images and are checked by researchers.

We resolve the imaging problems using image enhancement and the acoustic communication problems by only selecting interesting images and then compressing them.
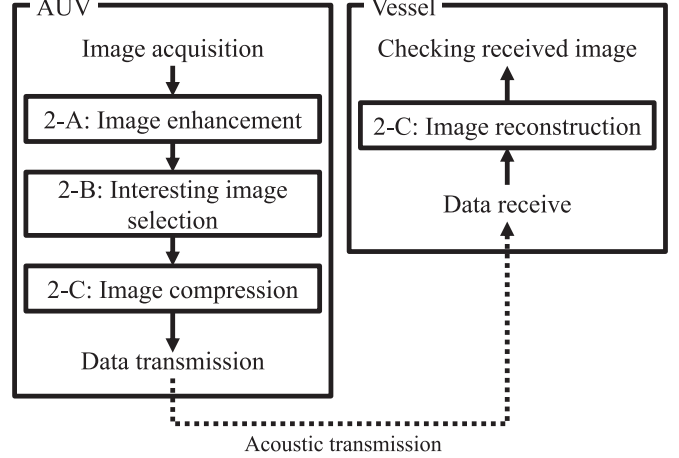
A. Image Enhancement

Because of light attenuation, the properties of images of the deep seafloor differ from those of ground images taken in air [8]. Each color has a different attenuation rate, and images of the deep seafloor captured by an AUV suffer from varying and uneven illumination. Recently, Retinex theory [9], [10] has been applied to solve such underwater imaging problems [11]–[13]. Retinex is a popular method for improving the visibility of an image using color constancy. In Retinex theory, an image comprises an illumination that comes from a light source and a reflectance, which is the color of object without influence of illumination. The relationship between the illumination and the reflectance in an image is supposed to be

$$I\ (x,y) =\ L\left(x,y\right)\ R\left(x,y\right) \tag{1}$$

where I represents the obtained image, (x,y) is pixel location in Cartesian coordinates, L is the illumination, and R is the reflectance.

In the Retinex model, the illumination is assumed to change smoothly in the image, that is, changes smoothly both across the physical dimension and the depth direction in the air. In water, light attenuates exponentially as a function of distance propagated. Therefore, the illumination in deep seafloor image changes smoothly on the image plane like strong lowpass filtered images. In typical deep seafloor images, the center area in images has high illumination values, and the surrounding area has low illumination values. Consequently, the illumination values smoothly change approaching the center on the image plane. Although the illumination in deep seafloor images experiences higher attenuation than in-air images, Retinex processing is still effective for enhancing images of the deep sea floor. For the image enhancement, we employ Retinex processing as follows:

$$_e I\ (x,y) =\ \alpha\ \{\log I\left(x,y\right) - \log\left(I\left(x,y\right) * G\left(x,y\right)\right)\} + \beta. \tag{2}$$

The symbol "$*$" represents convolution, G is a Gaussian filter, α is a scaling gain parameter, β is a scaling offset parameter, and $_e I$ denotes the enhanced image based on reflectance R.
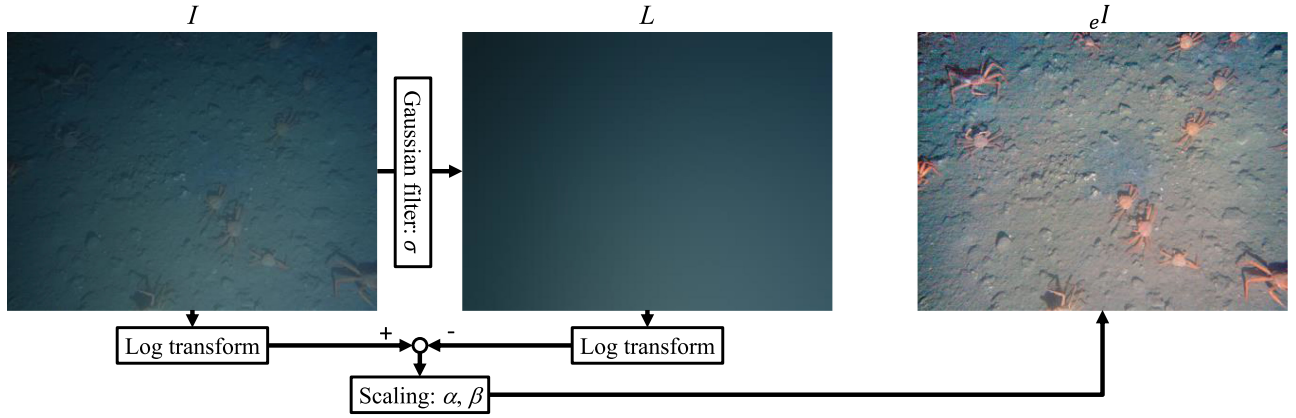
Fig. 3. Process of image enhancement using Retinex and the processed result. I: Obtained image by AUV. L: Approximated illumination using a Gaussian filter. $_e I$: Enhanced image. σ: Standard deviation of the distribution for a Gaussian filter. α: Scaling gain parameter. β: Scaling offset parameter.

A Gaussian filter can be defined as follows:

$$G\ (x,y) = \frac{1}{2\pi\sigma^2}\exp\left(\frac{-\left(x^2+y^2\right)}{2\sigma^2}\right), \iint G\left(x,y\right)dxdy = 1$$

$$\tag{3}$$

where σ is the standard deviation of the distribution.

The image enhancement process and a result of Retinex processing are illustrated in Fig. 3. The image obtained by the AUV, which shows red snow crabs, contains attenuated colors and uneven illumination (the left-hand side contains low color intensities, owing to factors such as the altitude of the AUV and the positional relationship between the imaging system and the seafloor). The approximated illumination is generated by applying a Gaussian filter to the image obtained by the AUV [image size $(niw, nih) = (640, 480)$; $\sigma = 128$]. We log transform both the obtained image and the approximated illumination, and then we subtract the log-transformed illumination from the log-transformed image. Finally, scaling by α and β is necessary to enhance visibility for human recognition. The enhanced image presented in Fig. 3 shows an improved visibility: The red snow crabs located on the left-hand side can be seen far more easily than in the image obtained by the AUV. In the image enhancement process, we employed a fixed set of parameters (σ, α, and β), which were decided through the numerical simulations for 640×480 pixels images, such as shown in Fig. 3 (obtained seafloor image by AUV, altitudes 1.5~2.5 m). We set $\sigma = 128$ and $\alpha = 512$ heuristically considering the calculation speed, the target resolution and the contrast of images. We set $\beta = 127$, the median value in 8-b format. Although this fixed set of parameters would not work effectively in all environments, they are effective in enhancing the obtained images, which were photographed by AUV at altitudes between 1.5 and 2.5 m.

B. Interesting Image Selection

While observing the deep seafloor, an AUV captures many seafloor images; however, images without creatures or objects are low priority for a sampling mission. To carry out an efficient sampling mission, the AUV should select only interesting or important images.

For interesting image selection, automatic recognition of marine life is the first subject, and the methods for recognizing marine life have been presented [11], [14] based on scale and rotationally invariant features [15], as has a highly accurate method for detecting fish based on deep learning [16], [17]. These methods require many marine life data sets for learning, whereas the methods are hard to apply especially in the first mission, because it is difficult to prepare a lot of marine life images taken by the similar color property camera with a certain distance to the target marine life under the similar lighting condition in the target observation area.

In this research, the saliency map [18], [19] is employed to select interesting images as the recognition of marine life is not needed. The saliency map is a model of human visual attention proposed by Itti *et al.* [18], and encodes the saliency or conspicuity of objects in a two-dimensional map based on the simple features of color, intensity, and orientation (see Fig. 4). The "$_e I$" is an enhanced image by Retinex processing, and the three features: "Color," "Intensity," and "Orientation" compose the Saliency map "SM", and the "Saliency areas in the image" indicates the high saliency areas. The individual pixel values in "SM" represent the saliency intensity. The light circles indicate areas of high saliency, and mostly coincide with the locations of crabs. The seafloor, for which the changes in pixels within an area are small, has lower saliency values compared with objects such as crabs or stones. The interesting image selection based on the saliency map is shown in Fig. 5.

Here, k is a running index and *ns* is an interesting image number. The acquired raw image, $I_{(k)}$, is obtained by the AUV, $_e I_{(k)}$ represents the enhanced image by Retinex process using $I_{(k)}$, $\mathrm{SV}_{(k)}$ is the maximum saliency value from the saliency map created using $_e I_{(k)}$, and the number of images for comparison is *ni*. "$k \geq ni$" performs to collect number of images more than *ni* before running the image selection algorithm. For example, if *ni* is 4 then the image selection process begins when k is 4, and $\mathrm{SV}_{(1)}$, $\mathrm{SV}_{(2)}$, $\mathrm{SV}_{(3)}$, and $\mathrm{SV}_{(4)}$ are compared to select interesting images. An image obtained by the AUV is selected as interesting when certain conditions are satisfied, such as $k \geq ni$, $\mathrm{SV}_{(k)}$ is greater than *th*, which is the seafloor threshold, and $\mathrm{SV}_{(k)}$ is the largest saliency value. When these conditions
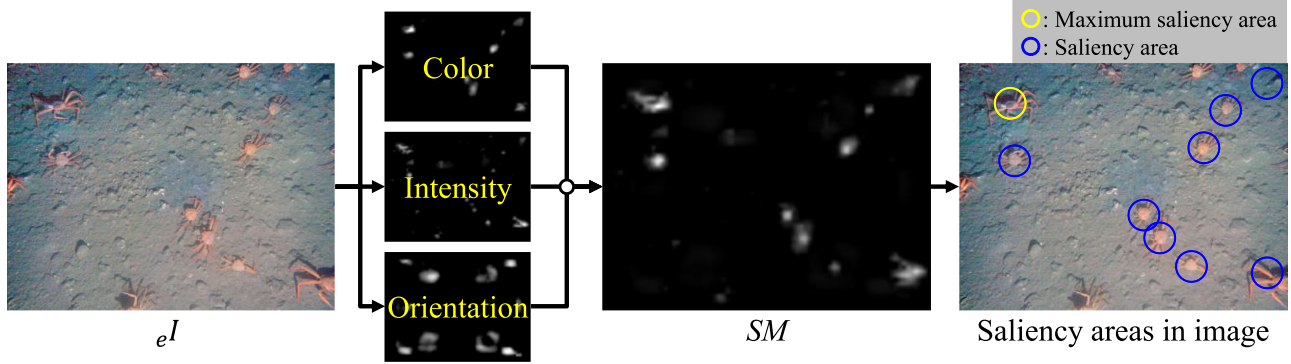
Fig. 4. Example of a Saliency map and the result. $_eI$: Enhanced image. SM: Saliency map.
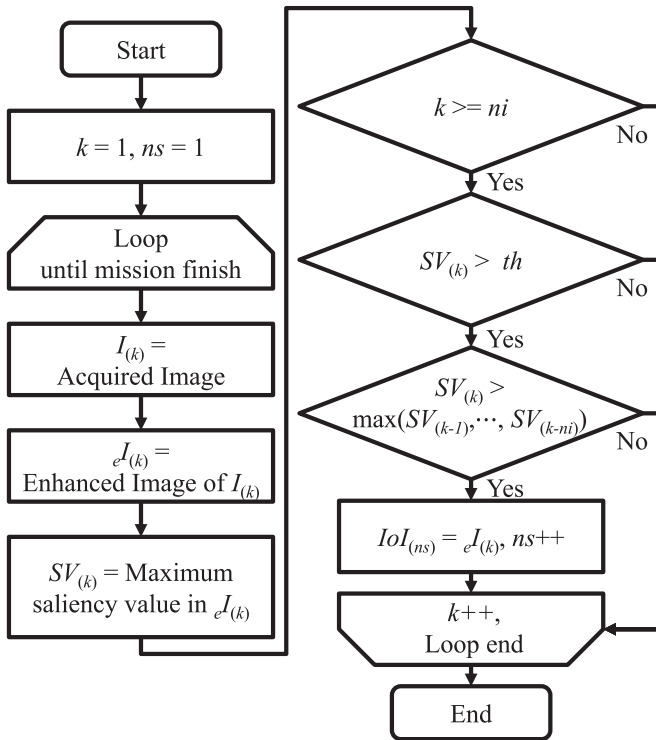

Fig. 5. Flow chart of proposed interesting image selection method. k: Running index. ns: Interesting image number. ni: Number of images for comparison to select interesting image. th: Seafloor threshold. IoI: Interesting image.

are met, $_eI_{(k)}$ is selected as an interesting image which named $IoI_{(ns)}$. Then, $IoI_{(ns)}$ is transferred to the next process for image compression. The loop in the interesting image selection process is repeated as long as the imaging by AUV is continued.

C. Image Color Compression and Reconstruction

In an underwater environment, acoustic communication is employed for the interaction between the AUV and the surface vessel during exploration. Acoustic communication has a lower information density than radio communication, and sometimes, data are lost during communication.

The well-known JPEG format image [20], which has a high data compression rate, is not suitable to transmit acoustically, because data cannot be restored if data loss occurs. In addition,

other data formats cannot be reconstructed if some of the format header data are missed. Hoag *et al.* [21] presented a method for video transmission by AUVs using wavelet compression. Wavelet compression is a solution when the loss of data is small. However, the AUV will move continually in our sampling mission, resulting in a large loss of data during communication. If data loss occurs often, then the operator cannot check the marine life in detail. Therefore, for our sampling strategy we strongly require an image compression method that is robust against data loss.

In this paper, we present a method for compression and reconstruction to check for marine life in a received image. The proposed method is based on color depth compression, which maintains pixel position information. The concept of the proposed method is presented in Fig. 6.

The left-hand side part within the red line represents the steps for the AUV, and the part on the right-hand side within the blue line represents the steps for the vessel. In the "Input image (24 b)", the area with the largest saliency values is selected as an "Area image of Maximum Saliency". The "Set of color palettes" was prepared before the mission, and comprises several colors extracted from a seafloor image data set. Then, AUV and vessel have the same "Set of color palettes" in the compression and reconstruction process. The horizontal direction represents each color of a palette and vertical direction represents each palette. We present the details of this color palette composition in Section III. The color palette that is most similar to that of the interesting object is selected using the Euclidean distance as follows:

$$nsp = \underset{i=1}{\overset{np}{\arg\min}} \left(\sum_{v=1}^{nwh} \sum_{u=1}^{nww} \min_{j=1}^{nc} \left|_w I\left(u, v\right) - P_{(i)}\left(j\right)\right| \right) \quad (4)$$

where $_wI$ is the image of an "Area image of Maximum Saliency" in Fig. 6, (u, v) is pixel location of $_wI$ in Cartesian coordinates, P is a "Set of color palettes," i is a palette number in a "Set of color palettes," j is a color number in that palette, nc is the total color number in that palette, (nww, nwh) denotes pixel numbers of width and height of $_wI$. np and nsp are the numbers of total palettes and the selected palette, respectively.

The "Area for transmission" in Fig. 6 consists of horizontal lines with the interesting object(s), and pixels in "Area for transmission" are replaced using the most close color numbers in
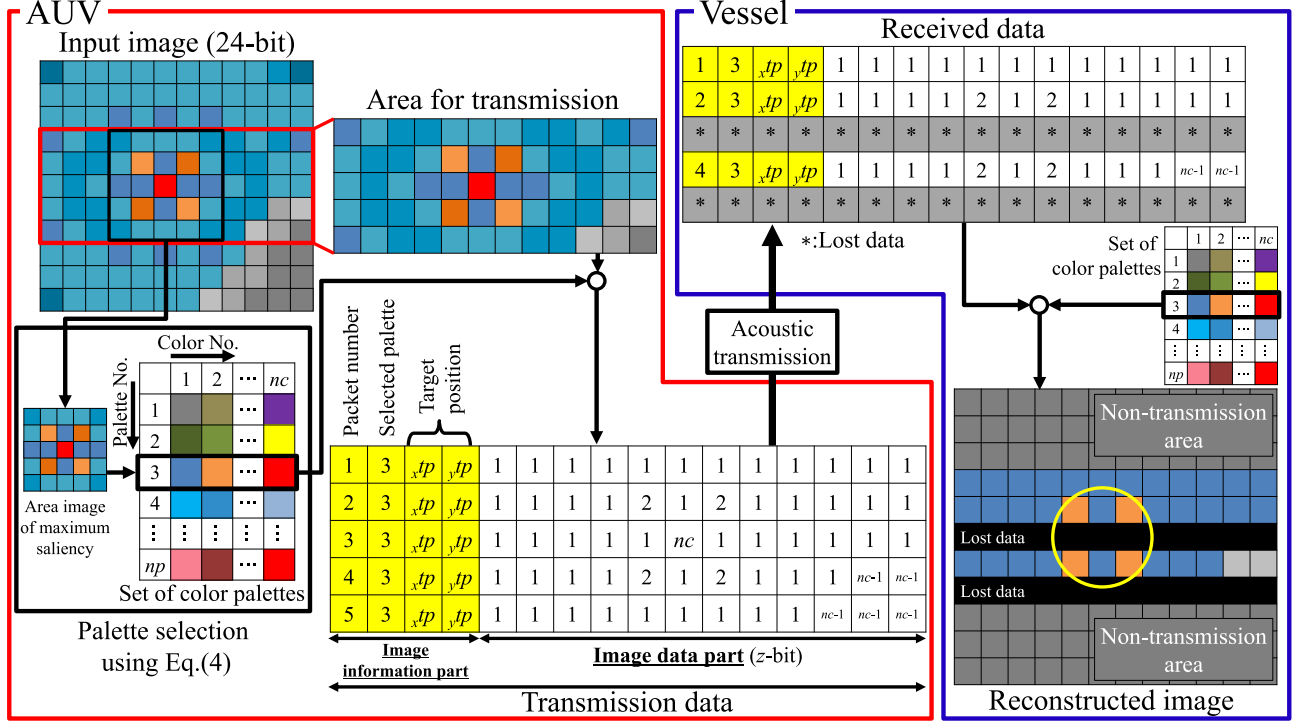
Fig. 6. Concept of proposed image color compression and reconstruction process. nc: Total color number of color palettes. np: Total palette number in a set of color palette. $(_xtp,\ _ytp)$: Center point of area image of maximum saliency. Circle in Reconstructed image: Circle is drawn using $(_xtp,\ _ytp)$ to represent maximum saliency area.

the selected color palette. The "Transmission data" comprises "Image information part" and "Image data part." The "Image information part" contains the numbers of the packet and the selected palette, and the position $(_xtp,\ _ytp)$ which is the center point of "area image of maximum saliency" image. One line of "data part" is the corresponding line of "Area for transmission", which is transmitted to the vessel by one packet using the acoustic transmitter. In the vessel, an image is reconstructed from a set of received data using the "set of color palettes; however, some of packets are lost so that the lost area is colored by black. Finally, the operator checks the "reconstructed image" with circle whose center is the position $(_xtp,\ _ytp)$.

D. Color Palettes Composition Method

The color palettes are the most important factor in determining the similarities of colors between the transmitted area and the reconstructed image. The proposed method for composing a set of color palette is illustrated in Fig. 7.

In the flow chart shown in Fig. 7, a data set of enhanced seafloor images $_pD$ is employed to compose a set of color palettes, the saliency maps $_pSM$ are made from $_pD$, and k and i are running indexes. The area image of maximum saliency $_wI_{(k)}$, which have the maximum saliency values in images, are classified into groups $IG_{(1)}, IG_{(2)}, \dots, IG_{(ng)}$, such as fish, crab manually. ng is total number of the image groups. All pixels in $IG_{(i)}$ are expressed in a RGB-color space, and then representative colors are selected using a color quantization method. A palette $P_{(i)}$ is composed by the representative colors, and a
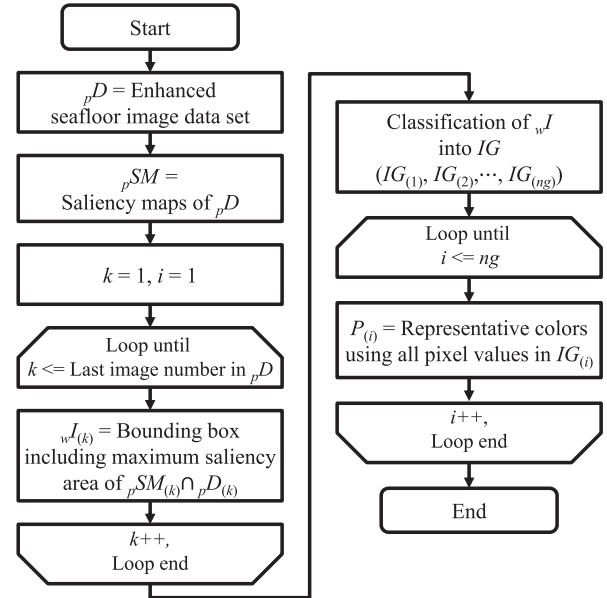

Fig. 7. Flowchart of color palette composition. k and i: Running indexes. IG: Image group. ng: Total number of image groups.

set of color palettes P is composed by several palettes. For the selection of representative colors in $IG_{(i)}$, minimum variance quantization [22]–[25] is employed, and an example of minimum variance quantization in the RGB space is presented in Fig. 8.
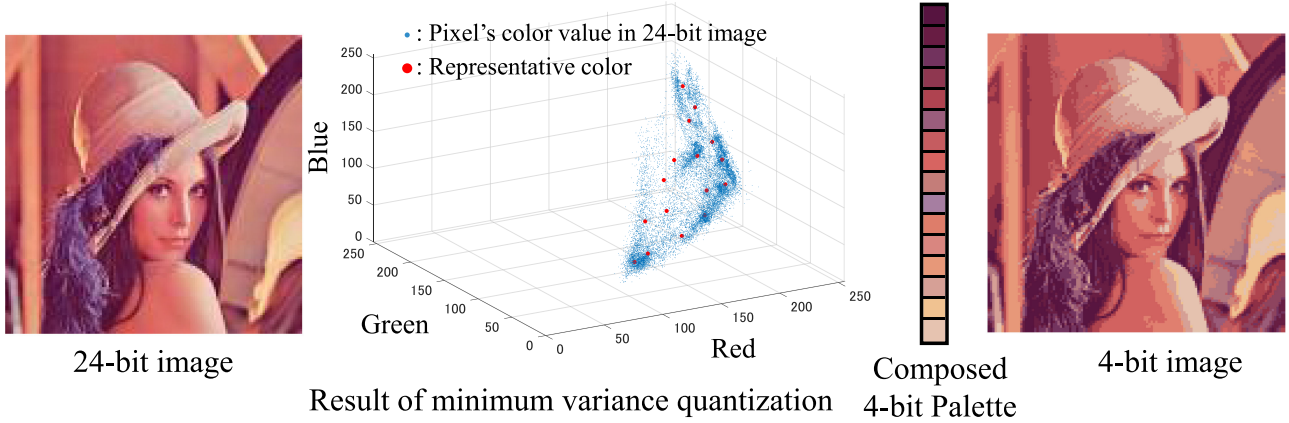
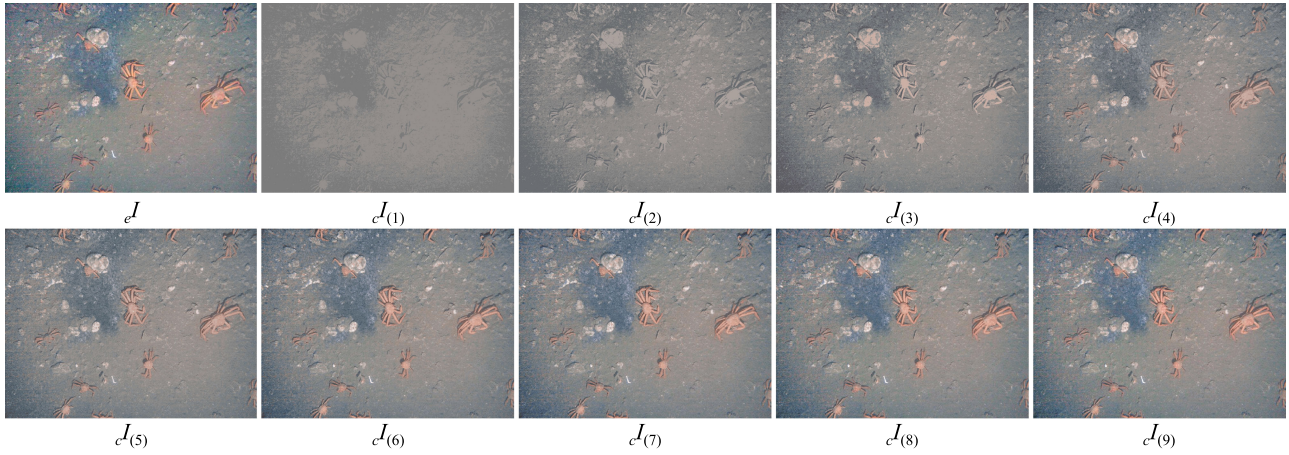Fig. 8. Example of minimum variance quantization (4-b color compression).

Fig. 9. Example of images compressed by different color bit palettes. $_e I$: Enhanced seafloor image (24 b). $_c I_{(k)}$: Compressed image by k-bit color palette.

In Fig. 8, dot circles represent pixel in a 24-b image and big circles indicate the representative colors in a palette. The 4-b image in the left consists of the 16 representative colors and the data size is 1/6 of the raw image, and it can be seen that the main features still remain. The evaluation of the process is given in Section III.

III. EVALUATIONS OF COMPRESSION RATE AND SIMILARITIES OF COLOR COMPRESSED IMAGE

The relationship between the image color quality and the compression rate is evaluated by changing bit-size of the color palettes, and the minimum bit-size is selected for the deep sea experiments using the similarities of images.

A. Evaluation of Compression Bits for Color Palettes

In the image compression method, the color bit-size of a color palette and the image color quality exhibit a tradeoff relationship. To evaluate the balance between color quality and the color bit-size, the peak signal-to-noise ratio (PSNR) [26]–[28], e_{PSNR}, is employed, which is used to measure the quality of reconstruction in lossy compression. The PSNR is defined as follows:

$$e_{\text{PSNR}}\ (I_A, I_B) = \ 10 \log \frac{255^2}{\text{mse}\ (I_A, I_B)} \qquad (5)$$

where I_A and I_B are images to compare, and mse denotes the mean squared error as follows:

$$\text{mse}\ (I_A, I_B) = \frac{1}{niw \cdot nih} \sum_{x=1}^{niw} \sum_{y=1}^{nih} [I_A\ (x, y) - I_B\ (x, y)]^2 \qquad (6)$$

where (niw, nih) is image size.

The proposed evaluation method compares enhanced image and compressed images which is used by different color bits palette. The k-bit color palette composed using $_p D$ is denoted as $P_{(k)}$. The compressed image $_c I_{(k)}$ made from $P_{(k)}$ is compared with the original enhanced image $_e I$ using (5) with changing k from 1 to 23 b. The results of k-bit compression are shown in Fig. 9.

In Fig. 9, it can be seen that higher image qualities correspond to higher numbers of color palette bits until 6 b. For further details concerning the image color quality and compression
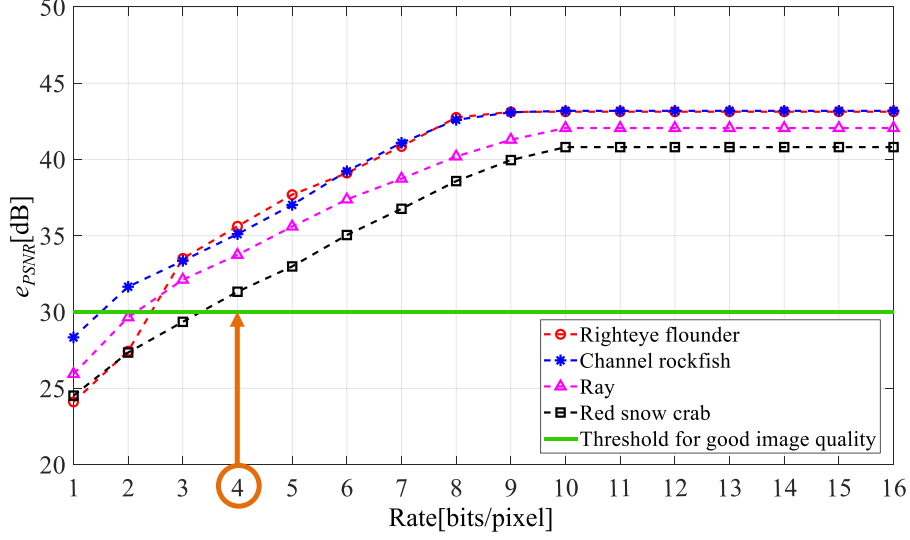
Fig. 10. Rate-distortion curve of compressed seafloor images.

bits, the rate-distortion curve of compressed seafloor images is shown in Fig. 10.

In Fig. 10, the images of different species, such as the right eye flounder, channel rockfish, ray, and red snow crab, are selected. Then, the images are processed using proposed evaluation method. The *x*-axis represents the rate [bits/pixel], which corresponds to the color palette bit number, and the *y*-axis represents the e_{PSNR} [dB]. In the result of Fig. 10, e_{PSNR} is increase until 10 b, then converges if it is larger than 10 b.

In the PSNR, 30 dB commonly represents a good image quality, and 4-b compression yields a value of over 30 dB for all images. Therefore, we select the 4-b compression for evaluation of similarity between 24-b image and compressed 4-b image.

B. Evaluation of Structural Similarity of Compressed Images

For the evaluation of 4-b color compressed images using the proposed compression method, the Structural SIMilarity (SSIM) [29] is employed. The SSIM index is a method used to evaluate the similarity between two images. The luminance, contrast, and structure are the three factors considered in the SSIM, which is calculated as follows:

$$e_{\text{SSIM}}\,(A,B) = \frac{(2\mu_A\,\mu_B + c_1)\,(s_{AB} + c_2)}{(\mu_A^2 + \mu_B^2 + c_1)\,(s_A^2 + s_B^2 + c_2)} \quad (7)$$

where A and B are the windows to measure the input and compressed images, respectively. The parameters μ_A and μ_B represent the averages of A and B, respectively, s_A^2 and s_B^2 are the variances of A and B, respectively, and s_{AB} is the covariance between A and B. The variables c_1 and c_2 are employed to stabilize the division with the weak denominator, and SSIM(A,B) represents the structural similarity between A and B.
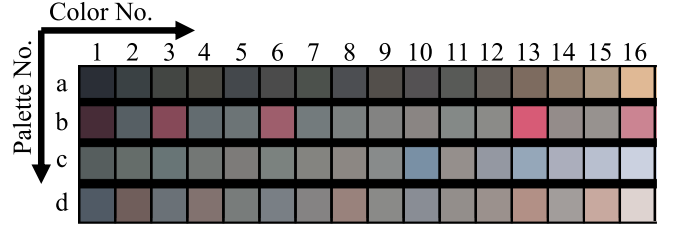

Fig. 11. Set of color palettes P for the evaluation using SSIM. a palette: Composed palette using ten right eye flounder images. b palette: Composed palette using ten channel rockfish images. c palette: Composed palette using ten ray images. d palette: Composed palette using ten red snow crab images.

A set of color palettes P, which is consisted of four palettes, is employed for the evaluation process based on the SSIM. The used P is shown in Fig. 11. P is composed using a seafloor image data set which contains ten right eye flounder images, ten channel rockfish images, ten ray images and ten red snow crab images. five images of each species, which are randomly selected for evaluation, are compressed by P, then the similarity between enhanced image (24 b) and compressed image (4 b) is calculated by (7). The evaluation result is presented in Table I. In Table I, the numbers "1–5" label the images for each species, "SSIM" is the structural similarity between the enhanced image (24 b) and compressed image (4 b), and "Selected palette" indicates the palette selected from Fig. 11 using the proposed color palette selection method. The average SSIM value for the evaluation is 91.4%, with a maximum of 93.8% and a minimum of 88.0%. The examples of the images are shown in Fig. 12 in the form of enhanced and color-compressed images. The example images in Fig. 12 illustrate the differences between enhanced and compressed images. Although all of the compressed images are composed of different colors, the colors are similar, and the operator can recognize species. The SSIM evaluation proved that 4-b compression is available. The deep sea experiment based on 4-b color palettes is given in Section IV.

TABLE I
EVALUATION OF THE COMPRESSION METHOD USING SSIM

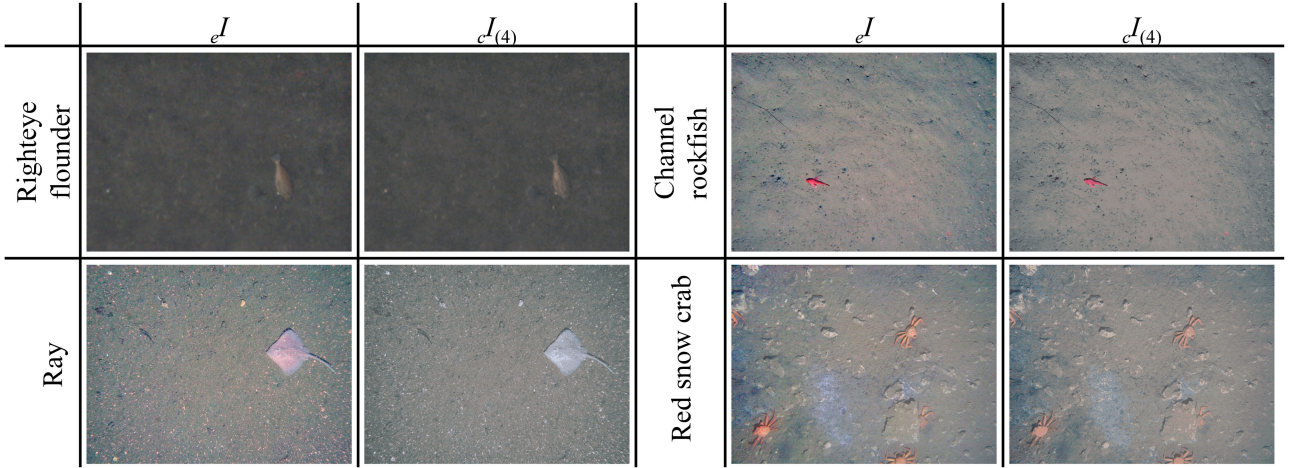| Species in image	1 SSIM	1 Selected palette	2 SSIM	2 Selected palette	3 SSIM	3 Selected palette	4 SSIM	4 Selected palette	5 SSIM	5 Selected palette
Righteye flounder	93.3%	a	93.5%	a	93.4%	a	93.7%	a	93.8%	a
Channel rockfish	92.2%	b	91.9%	b	92.3%	b	92.1%	b	92.1%	b
Ray	92.0%	c	91.7%	c	91.9%	c	91.6%	c	92.0%	c
Red snow crab	88.2%	d	88.0%	d	88.0%	d	89.2%	d	88.6%	d

Fig. 12. Examples of enhanced and compressed images for evaluation. $_eI$: Enhanced image (24 b). $_cI_{(4)}$: Compressed image by 4-b color palette.
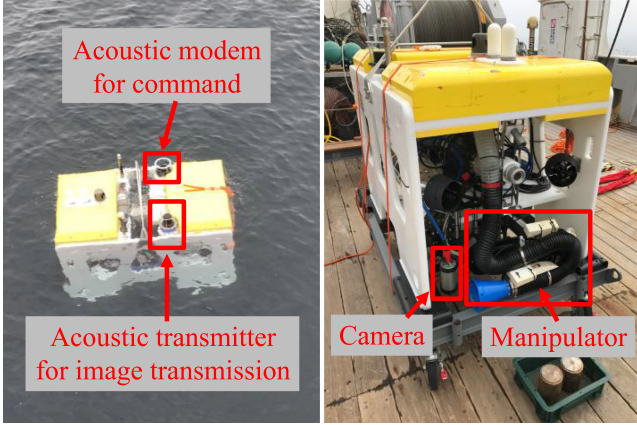

Fig. 13. Appearance of the SAUV:TUNA-SAND2.

IV. DEEP SEA EXPERIMENT AND RESULT

We tested the proposed image transmission method using the SAUV: TUNA-SAND2 [30], which was developed by our research group for a sampling mission in the Sea of Okhotsk. The appearance of TUNA-SAND2 is shown in Fig. 13. The dimensions of TUNA-SAND2 are 1.4 m × 1.2 m × 1.3 m, the weight is 380 kg, and six thrusters were mounted. In addition, a manipulator for sampling and acoustic transmitter for image transmission was mounted on TUNA-SAND2.

In the experiments, we employed an acoustic transmitter with a data transmission speed of 20 kb/s. Details of the acoustic transmitter and receiver are provided in Table II, and the experimental setup of the seafloor image transmission system is described in Table III. In Table III, "Compression bit" represents the number of color bits of the compressed images, with 4 b being used, as selected in the evaluation. The "Image size" is that of the image resized for faster compression and transmission. The "Area size for transmission" represents the part of the image that contains the object of most interest, and the "ni" represents a number of images for comparison to select the interesting image which is presented in Fig. 5. The camera frame rate is 1/5, theoretical required time for one transmission is 3.7 s, and the AUV traveling speed for seafloor observation is 0.1 m/s. The AUV altitude was set as either 10 or 2 m. We used employed 10 m altitude for fast cruising toward the starting point for observations, to prevent a collision with the seafloor, whereas we used the 2 m altitude for the observations.

In the experiment, TUNA-SAND2 undertook transmission missions in various sea areas and at different depths, and the "compressed data (z bit)," which is explained in Fig. 6, is composed of 60 packets (4 b). Each packet consists of a 24-B header part and 128-B data part. The 24-B header part is automatically assigned packet information by the transmitter. The 128-B data part is sorted into an 8-B image information part, which contains a 1-B packet number, 1-B selected palette number, 2-B

Transmitter		Receiver	
Beam angle	±35°	Range	±35°
Power	DC24V, 50W	Power	DC24V, 5W
Operating Depth	2000m	Operating Depth	100m
Acoustic Connection	20 kbit/s (in ideal condition)	Acoustic Connection	20 kbit/s (in ideal condition)
Data packet	152-byte (Header: 24-byte, Data: 128-byte)	Data packet	152-byte (Header: 24-byte, Data: 128-byte)

TABLE III
DETAILS OF PARAMETERS IN EXPERIMENT

Category	Value	Category	Value
Compression bit	4-bit	Frame rate	1/5
Image size	240×180	Theoretical required time for one transmission	3.7 s
Area size for transmission	240×60	AUV traveling speed for observation	0.1m/s
ni	4	AUV altitudes	10m, 2m

image number, 2-B target position (coordinates of the center point), and 2-B position of area for transmission (coordinates of starting point); and a 120-B image data part. 1-B of image data contains the data for two pixels compressed into a 4-b format.

Fig. 14 presents examples of images that were selected and received as part of the transmission experiments, during which various creatures were photographed, such as octopuses, starfish, sea cucumbers, and channel rockfish. The photographs taken by TUNA-SAND2 from the 2 m altitude show a rectangular area of the seafloor of size 1.1 m by 1.5 m. If the transmission success rate was more than 50%, we could easily identify the species, such as starfish and octopus, from most received images. The results of the transmission experiments are summarized in Table IV.

In Table IV, "Mission Dive number," "Mission Depth," and "Mission Date" provide information about the transmission mission. "No. of Obtained images" is the total number of images obtained by the AUV during one mission. "No. of Selected images" is the number of images selected by the proposed method as interesting images. "No. of Selected images when moving to start point" is the number of images taken from the 10 m altitude, and "No. of Selected images during observation" is the number captured from the 2 m altitude. "No. of Selected marine life images during observation" is the number of images that included marine life. "No. of Received images" is the number of received images, "No. of Transmitted packets" is the number of packets sent from the AUV, and "No. of Received packets" is the number of packets that reached the surface vessel. In these experiments, the total rate of selected marine life images was 63.3%, and the total transmission success rate was 21.8%.

V. DISCUSSION

In this paper, we proposed an image transmission system to be employed in SAUV missions, and we presented the results of experiments in which an AUV transmitted images of a deep sea environment. Raw images of the deep seafloor, which suffer from attenuated colors and uneven illumination, are enhanced using the Retinex process. The colors in the enhanced images are not the same as those of the same object in air, because the Retinex approach does not consider light attenuation under water. The approximated illumination in a deep seafloor image contains high green and blue values, because red values are highly attenuated. Therefore, the enhanced images contain high red values. Thus, the use of Retinex with a hue adjustment has been proposed for restoring the colors in underwater images [11]. However, we employed basic Retinex in our system, because the entire image transmission process must be completed within the imaging interval (5 s). Furthermore, because the received images are checked by a human, some color changes can be accommodated.

The proposed process for selecting interesting images, which is based on a saliency map, selects the enhanced images that contain marine life, and exhibited an average accuracy of 63.3% in the experiments. The minimum accuracy was 37.2% on dive 2, and the maximum accuracy was 100% on dives 3 and 4. This differing accuracy was a result of the sea area. The accuracy is higher when more creatures are present in a sea area. This is because the saliency map, which is a model of human visual attention, is employed in the selection of interesting images. Although this process is prone to large differences in accuracy, it remains suitable for sampling missions, because sampling targets are ultimately determined by the operator on the surface vessel.
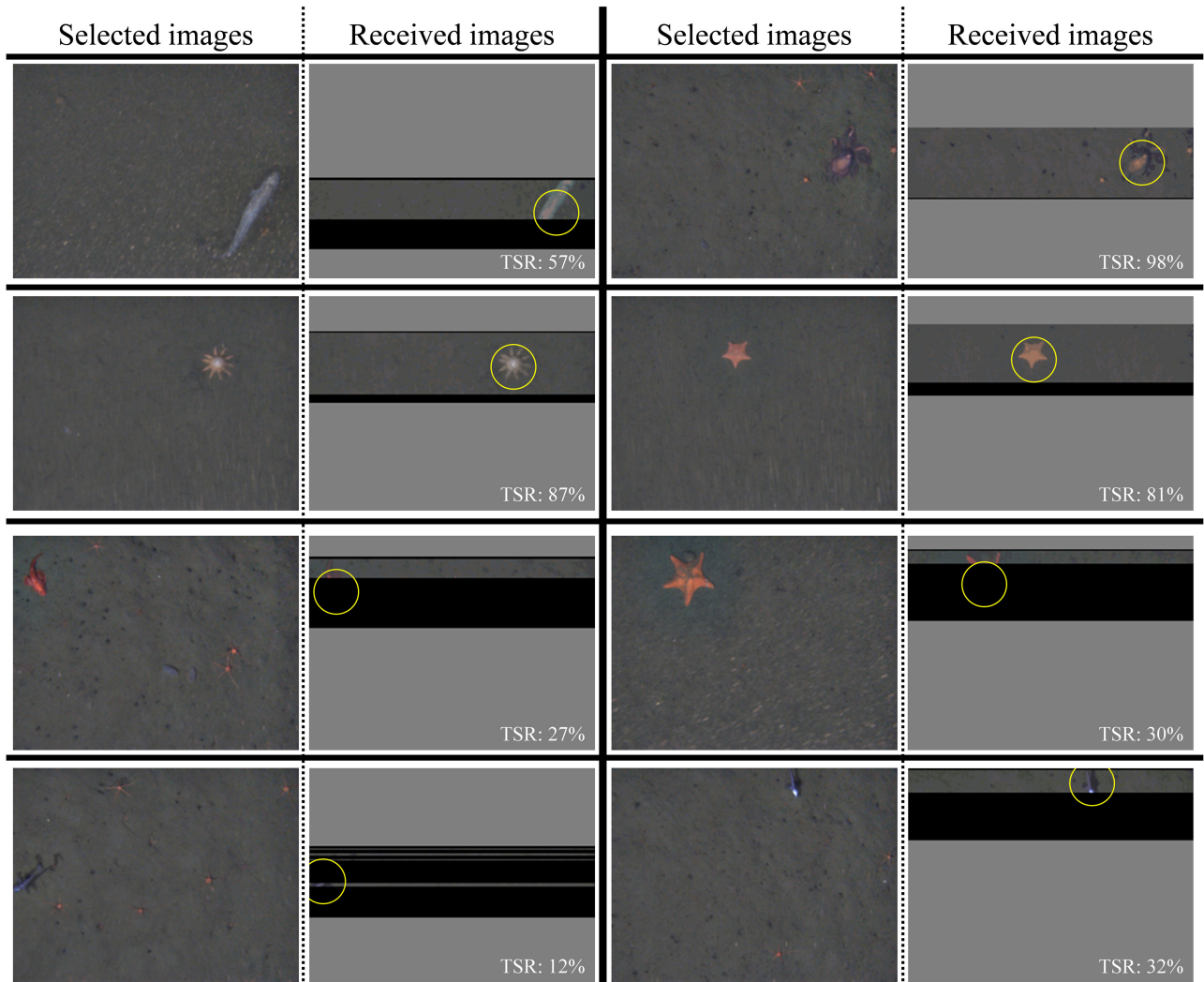
Fig. 14. Examples of selected and received images during the image transmission experiments in the Sea of Okhotsk.

TABLE IV
RESULTS OF TRANSMISSION EXPERIMENTS IN THE SEA OF OKHOTSK

Mission			No. of							
Dive No.	Depth	Date	Obtained images	Selected images	Selected image when moving to start point	Selected images during observation	Selected marine live images during observation	Received images	Trasmitted packets	Recived packets
1	350m	30-May-17	819	203	6	197	122	65	12180	1873
2	510m	31-May-17	1081	175	38	137	51	97	10500	2982
3	580m	1-Jun-17	891	127	51	76	76	60	7620	1546
4	590m	2-Jun-17	214	41	12	29	29	27	2460	736
Total			3005	546	107	439	278	249	32760	7137

The proposed image compression and reconstruction method based on color depth compression exhibited a 91.4% similarity in terms of the SSIM, and the size if the data required for information delivery was decreased by a factor of 18 (Only 1/3 of the image, which contains the target, is selected as the area for transmission, and the data size is reduced by a factor of 6 using color compression). The compression data was received by the vessel within 3.7 s (Theoretical required time), and the transmission success rate was 21.8% during the four missions. This 21.8% transmission success rate reflects the difficulty of using

acoustic transmission for AUVs in a deep sea environment. The fact that many packets did not reach their intended destination only because of the positional relationship and signal attenuation shows that the proposed color compression method, which allows reconstruction even if data loss occurs, is feasible for real-time deep sea image transmission missions.

VI. Conclusion

The use of AUVs on missions to sample deep sea creatures is expected to constitute the next step in ocean exploration. Therefore, SAUVs are being developed according to such requirements [7], [30]. In this paper, we presented an image transmission system as part of an SAUV sampling system, and we demonstrated the effectiveness of this transmission system in deep sea experiments. In conventional AUV observations of the deep seafloor, the AUV cruises a set observation area while imaging [1], even if photographing rare creatures or interesting scenes that are worthy of observation. This is a major disadvantage in the field of marine biology. If researchers could check the images obtained by the AUV in real time and concurrently command it to track by recording video or taking samples, they could collect more biological information. We expect the proposed system to facilitate the aforementioned type of deep sea exploration.

References

[1] T. Ura, "Observation of deep seafloor by autonomous underwater vehicle," *Indian J. Geo-Mar. Sci.*, vol. 42, pp. 1028–1033, 2013.

[2] T. Maki, A. Kume, and T. Ura, "Volumetric mapping of tubeworm colonies in Kagoshima Bay through autonomous robotic surveys," *Deep Sea Res. Part I, Oceanogr. Res. Papers*, vol. 58, no. 7, pp. 757–767, 2011.

[3] T. Maki, Y. Sato, T. Matsuda, R.-T. Shiroku, and T. Sakamaki, "AUV Tri-TON 2: An intelligent platform for detailed survey of hydrothermal vent fields," in *Proc. IEEE/OES Auton. Underwater Vehicles*, 2014, pp. 1–5.

[4] Y. Nishida *et al.*, "Development of an autonomous underwater vehicle for survey of cobalt-rich manganese crust," in *Proc. MTS/IEEE OCEANS Conf.*, Washington, DC, USA, 2015, pp. 1–5.

[5] T. Nakatani *et al.*, "AUV TUNA-SAND and its exploration of hydrothermal vents at Kagoshima Bay," in *Proc. MTS/IEEE OCEANS Conf.*, Kobe, Japan, 2008, pp. 1–5.

[6] M. R. Clark, M. Consalvey, and A. A. Rowden, *Biological Sampling in the Deep Sea*. Hoboken, NJ, USA: Wiley, 2016, pp. 285–305.

[7] K. Ishii, T. Sonoda, Y. Nishida, S. Yasukawa, T. Ura, and K. Watanabe, "A new tool to access deep-sea floor "Sampling-AUV"," in *Proc. Int. Conf. Artif. Life Robot. Conf.*, 2017, pp. 301–303.

[8] H. R. Gorden, "Can the Lambert-Beer law be applied to the diffuse attenuation coefficient of ocean water?," *Limnol. Oceanogr.*, vol. 34, no. 8, pp. 1389–1409, 1989.

[9] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Amer.*, vol. 61, pp. 1–11, 1971.

[10] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," in *Proc. Int. Soc.Opt. Photon.*, 2002, pp. 390–401.

[11] J. Ahn, S. Yasukawa, T. Sonoda, T. Ura, K. Ishii, "Enhancement of deep-sea floor images obtained by an underwater vehicle and its evaluation by crab recognition," *J. Mar. Sci. Technol.*, vol. 22, no. 4, pp. 758–770, 2017.

[12] X. Fu, P. Zhuang, Y. Huang, Y. Liao, X.-P. Zhang, and X. Ding, "A Retinex-based enhancing approach for single underwater image," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 4572–4576.

[13] S. M. Alex Raj and M. H. Supriya, "Underwater image enhancement using single scale Retinex on a reconfigurable hardware," in *Proc. Int. Symp. Ocean Electron.*, 2015, pp. 1–5.

[14] Y. Nishida, T. Ura, T. Hamatsu, K. Nagahashi, S. Inaba, and T. Nakatani, "Resource investigation for Kichiji rockfish by autonomous underwater vehicle in Kitami-Yamato bank off Northern Japan," *ROBOMECH J.*, vol. 1, no. 2, pp. 1–6, 2014.

[15] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. Comput. Vis. Conf.*, 1999, pp. 1–8.

[16] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, "DeepFish: Accurate underwater live fish recognition with a deep architecture," *Neurocomputing*, vol. 187, pp. 49–58, 2016.

[17] M. Sung , S.-C. Yu, and Y. Girdhar, "Vision based real-time fish detection using convolutional neural network," in *Proc. MTS/IEEE OCEANS Conf.*, Aberdeen, Scotland, 2017, pp. 1–6.

[18] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[19] D. Walther and C. Koch, "Modeling attention to salient proto-objects," *Neural Netw.*, vol. 19, no. 9, pp. 1395–1407, 2006.

[20] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. 18–34, Feb. 1992.

[21] D. F. Hoag, V. K. Ingle, and R. J. Gaudette, "Low-bit-rate coding of underwater video using wavelet-based compression algorithms," *IEEE J. Ocean. Eng.*, vol. 22, no. 2, pp. 393–400, Apr. 1997.

[22] S. J. Wan, P. Prusinkiewicz, and S. K. M. Wong, "Variance-based color image quantization for frame buffer display," *Color Res. Appl.*, vol. 15, no. 1, pp. 52–58, 1990.

[23] T. W. Spencer, "Efficient inverse color map computation," in *Graphics Gems II*. Boston, MA, USA: Academic, pp. 116–125, 1991.

[24] Z. Xiang, "Color image quantization by minimizing the maximum inter-cluster distance," *ACM Trans. Graph.*, vol. 16, no. 3, pp. 260–276, 1997.

[25] C. Ozturk, E. Hancer, and D. Karaboga, "Color image quantization: A short review and an application with artificial bee colony algorithm," *Informatica*, vol. 25, no. 3, pp. 485–503, 2014.

[26] B. Girod, "What's wrong with mean-squared error," in *Digital Images and Human Vision,* A. B. Watson, Ed. Cambridge, MA, USA: MIT Press, 1993, pp. 207–220.

[27] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," *Signal Process.*, vol. 70, pp. 177–200, 1998.

[28] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, 2010, pp. 2366–2369.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[30] Y. Nishida *et al.*, "Development of an autonomous underwater vehicle with human-aware robot navigation," in *Proc. MTS/IEEE OCEANS Conf.*, Monterey, CA, USA, 2016, pp. 1–4.

Jonghyun Ahn (S'16–M'18) received the Ph.D. degree in underwater robotics from the Kyushu Institute of Technology, Kitakyushu, Japan, in 2017.

From April to September 2017, he was a Researcher with the Center for Socio-Robotic Synthesis, Kyushu Institute of Technology. Since October 2017, he has been an Assistant Professor with the Department of Human Intelligence Systems, Kyushu Institute of Technology. His research interests include underwater imaging system, intelligent sensing, and underwater communication system.

Shinsuke Yasukawa received the Ph.D. degree in engineering from the Division of Electrical, Electronic and Information Engineering, Osaka University, Osaka, Japan, in January 2017.

He is currently an Associate Professor with the Department of Human Intelligence Systems, Kyushu Institute of Technology, Kitakyushu, Japan. His research interests include information processing in biological systems and their applications in robotics.

Takashi Sonoda received the Ph.D. degree in robotics from the Kyushu Institute of Technology, Kitakyushu, Japan, in 2010.

From April 2009 to March 2012, he was a Researcher with Fukuoka Industry, Science and Technology Foundation, Kitakyushu, Japan, and from April 2012 to March 2015, he was a Researcher with the Department of Brain Science and Engineering, Kyushu Institute of Technology. He was also an Associate Professor with the Department of Human Intelligence Systems, Kyushu Institute of Technology, from April 2015 to March 2018. Since April 2018, he has been an Associate Professor with the Department of Integrated System Engineering, Nishinippon Institute of Technology, Kitakyushu, Japan. His research interests include underwater robotics and robot manipulator system.

Kazuo Ishii (S'94–M'96) received the Ph.D. degree from the Department of Naval Architecture and Ocean Engineering, University of Tokyo, Tokyo, Japan, in 1996.

He is currently a Professor with the Department of Human Intelligence Systems, and the Director of the Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, Kitakyushu, Japan. His research interests include underwater robots, agricultural robots, roboCup soccer robots, and intelligent systems.

Yuya Nishida (M'16) received the Ph.D. degree in engineering from the Kyushu Institute of Technology, Kitakyushu, Japan, in 2011.

After spending one year at the university, he continued research related to underwater vehicle until here joined the Kyushu Institute of Technology in 2015 in his current role. He has been working on the development of autonomous underwater vehicles and underwater instruments for survey the marine resources. His research interests include the automation of marine survey and inspection conducted by humans.

Tamaki Ura (M'91–SM'02–F'07) graduated from the Faculty of Engineering, University of Tokyo, Tokyo, Japan, in 1972, and the Doctor of Engineering degree from the University of Tokyo in 1977.

He is currently a Project Professor with the Kyushu Institute of Technology and a Professor Emeritus with the University of Tokyo. He is one of the Leader in R/D of autonomous underwater vehicle in Japan.

Prof. Ura was the recipient of the IEEE Oceanic Engineering Society Distinguished Technical Achievement Award in 2010.