

東京大学大学院新領域創成科学研究科
海洋技術環境学専攻

平成29年度

修士論文

**New method of fish classification by using high-
resolution acoustic video camera-ARIS and local
invariant feature descriptor**

(音響ビデオカメラ ARIS と局所特徴量を用いた新しい魚類分類手法の開発に関する研究)

2016年2月15日提出
指導教員：浅田 昭 教授

張 宇
Zhang Yu

Table of Contents

ACKNOWLEDGEMENT	1
CHAPTER 1: INTRODUCTION	2
1.1. Research background	2
1.1.1. Fish survey and fish classification nowadays.....	2
1.1.2. Izumuma Lake.....	4
1.2. Significance and motivation of the study	5
CHAPTER 2: EQUIPMENT AND OBSERVATION METHOD	6
2.1. ARIS and acoustic lens	6
2.2. Observation method for survey	9
2.3. Theory analysis of two observation methods	11
CHAPTER 3: FIELD EXPERIMENT I	13
3.1. Experiment site and settings	13
3.2. Data processing method	15
3.2.1. Pre-processing	15
3.2.2. NCC method.....	17
3.2.3. SIFT method.....	18
3.2.3.1 Scaling and detection of key point	20
3.2.3.2 Localization of key point.....	22
3.2.3.3 Calculation of orientation at key point	23
3.2.3.4 Description of feature at key point	24
3.3. Result and discussion	27
3.4. Potential problems and future work	32

CHAPTER 4: FIELD EXPERIMENT II	33
4.1. Objectives of experiment II	33
4.2. Experiment site and settings	34
4.3. Data processing	38
4.4. Result and discussion	39
4.5. Potential problems and future work	45
CHAPTER 5: SIMULATION	46
5.1. Objective of simulation	46
5.2. Theory introduction and parameter settings	46
5.3. Construction of 3-D fish model	47
5.4. Coordinate settings	48
5.5. Backscattering strength	50
5.6. Comparison between simulated images and sonar images	51
5.7. Discussion of result and potential problems	60
CHAPTER 6: CONCLUSION AND FUTURE WORK	61
CHAPTER 7: REFERENCE	62

ACKNOWLEDGEMENT

First of all, I would like to express my gratitude to my adviser Professor Akira Asada for his assistance to me not only in research but also in the daily life for 2 years. My adviser professor is full of passion and enthusiasm to the research, which encouraged me to work hard all the time. I leant much knowledge about acoustics and the insight to analyze problems at every seminar from him. And patient guidance and invaluable on my master thesis paper benefit me a lot.

I also would like to thank my vice-adviser Professor Shigeru Tabeta for his kindness and patient guidance. His advice from different academic perspective always gave me lots of inspiration and helped me improve and develop my research.

I still want to give my sincere appreciation to research associate Katsunori Mizuno. I was touched by his knowledge, kindness and patience to students, and enthusiasm for research, from whom I have leant a lot that would affect my whole life. Without him, I could not have completed my master research.

And then I want to thank researcher Tetsuo Shimada and Yasufumi Fujimoto of the Miyagi Prefectural Izunuma-Uchinuma Environmental Foundation for their kindness and patient support in the field experiment.

Last but not least, I also thank all of my fellow lab-mates in Asada Lab because they gave me lots of support and help in research and daily life.

CHAPTER 1: INTRODUCTION

1.1. Research background

1.1.1. Fish survey and fish classification nowadays

Nowadays, monitoring of fish state and behavior during cultivation has increasingly received attention by not only fish producers but also lots of scientific researchers as it will help to improve profitability for producers and also reduce the threat of disease and incidents. Conventionally, direct sampling method by sampling net (Fig. 1) or electric shocker (Fig. 2) has been used to assess the fish species and amount. However, such methods are usually time-consuming, laborious and invasive. Furthermore, conventional sampling is always affected by the transparency of the water and the observational skills of the persons. Therefore, using efficient and noninvasive method is important and expected.



Fig. 1 Fish survey by net sampling



Fig. 2 Fish survey by electric shocker

In the past three decades, optical sensors and machine vision system (MVS) inventions were widely applied into the fish monitoring and they are becoming more sensitive, powerful and cheaper day by day [1]. However, in the turbid water, the acoustic video camera can “see” more clearly than optical sensors. Acoustic systems can penetrate the aquatic environment over long distance. Echo-sounders transmit acoustic pulses into water through a transducer. When pulses meets a target with a different density from the propagation environment, they will reflect, penetrate and backscatter from the interface. And then fish or objects underwater will be detected. Hydro-acoustic technologies have developed from single-beam sonar, dual-beam sonar, split-beam sonar to multi-beam sonar. [2]

As we know, initially designed for military purposes, dual-frequency identification sonar (DIDSON, SOUND METRICS CORP., Lake Forest Park, WA, USA, as Fig. 3 shows) has been used in environmental management for a decade. This acoustic camera uses higher frequencies

and more sub-beams than common hydro acoustic tools, which improves image resolution and then enables observation of fish morphology and swimming behavior. The ability to subtract static echoes from echograms and directly measure fish length improve the species-identification process. Fig. 4 is a sample acoustic image obtained by DIDSON in Ganges river of India by Asada lab.



Fig. 3 Acoustic video camera DIDSON [3]

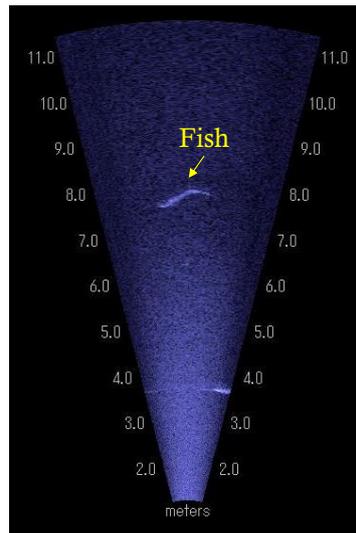


Fig. 4 Acoustic image obtained with DIDSON in Ganges river of India by Asada lab

Besides, acoustic video camera is widely applied into environmental management, migratory fish monitoring and so on. Our team has succeeded in classifying the aquatic plant species using acoustic video camera and image processing technique. [4,5] However, the frequency of DIDSON is not high enough to get images of clear shape of fish, which leads to the fact that fish classification using acoustic video camera still remains a difficult topic of scientific researchers. Nowadays, fish classification is mainly carried out in some indirect ways, such as by fish's acoustic shadow [6] or frequency of fish's caudal fin [7].

Therefore, it is key point to get high-quality acoustic images for fish classification.

1.1.2. Izunuma Lake

Izunuma Lake is a fresh water lake that located in the north of Miyagi Prefecture ($38^{\circ}43'N$, $141^{\circ}07'E$, Fig. 5), famous as an inner marsh and also for its abundant living creature research resources.

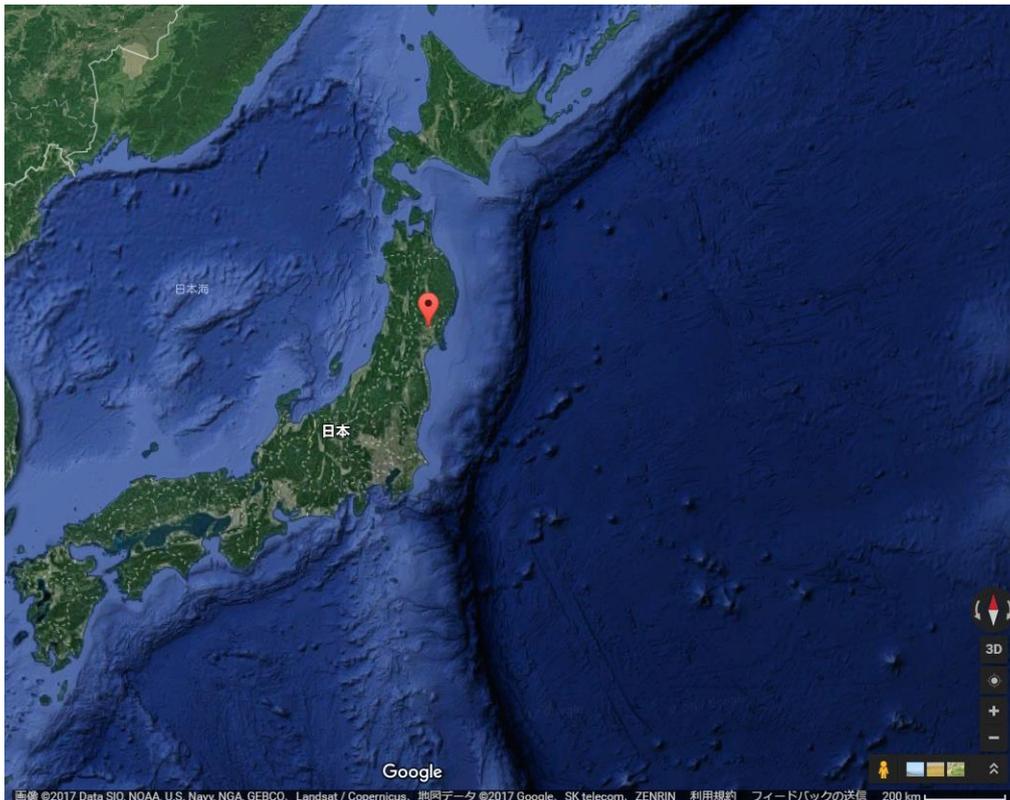


Fig. 5 Location of Izunuma Lake (Google map)

There are many kinds of aquatic plants, fish and birds living in Izunuma Lake. Talking about fish, there are carp, Japanese crucian carp, Japanese barbel, black bass, snakehead mullet and so on. However, in the recent decade, the water was becoming turbid and muddy because of eutrophication. In addition, recently with the coming of unknown species of exotic fish, local fish's lives became to be threaten. Therefore, the statistics of fish' species and amount in the lake has become a problem now. Instead of the time-consuming, laborious and invasive conventional fish survey method, more efficient and noninvasive method is expected by the administrator of Izunuma Lake.

1.2. Significance and motivation of the study

Many scientists and researchers show their interest in the topic of fish classification. Effective fish classification provides fish farmers with better marketing in polyculture fish farming system, feeding strategies and stock assessment, because they can sort harvested fish according to species and size. In addition, when it comes to the fish disease, fish species should be properly classified before the disease can be identified because it will manifest different symptoms in different species of fish. Therefore fish classification is helpful to fish disease diagnosis. What's more, in environmental management, fish classification can also contribute to assess the species and amount of some endangered migratory fish through long-time monitoring.

As mentioned in 1.1.1, fish classification by acoustic camera is mainly studied in some indirect ways. Firstly, it is due to the low quality of the sonar image of fish. With the conventional method by acoustic sonar, it is the truth that researchers can hardly obtain clear and complete fish images.

Therefore, this study proposes a new observation method in order to get high-quality sonar images of fish. Besides, this study aims to raise the accuracy of fish classification by making use of the local invariant feature of acoustic images and the result will be compared with another template matching method called NCC. What's more, this study focuses on the survey of effect which fish's swimming angle causes on matching rate. Last but not least, the study will use simulation method to compare simulated images with sonar images in order to solve the problem of low matching accuracy with imperfect sonar images of fish (images with only part body of fish).

CHAPTER 2: EQUIPMENT AND OBSERVATION METHOD

2.1. ARIS and acoustic lens

Acoustic video camera is multi-beam, high-frequency sonar with a unique acoustic lens system designed to focus the beam to create high-resolution images, transmitting sound pulses and convert the returning echoes into digital images. ARIS (Sound Metrics, Bellevue, WA, USA) is the next generation product of DIDSON (Dual-Frequency Identification Sonar) which are both produced by Sound Metrics Corporation. ARIS has higher resolution than that of DIDSON, but the observation range is shorter. In this study, ARIS was used for observation of fish. Fig. 6 show appearance of ARIS and its accessories, the pictures were taken in Asada lab.

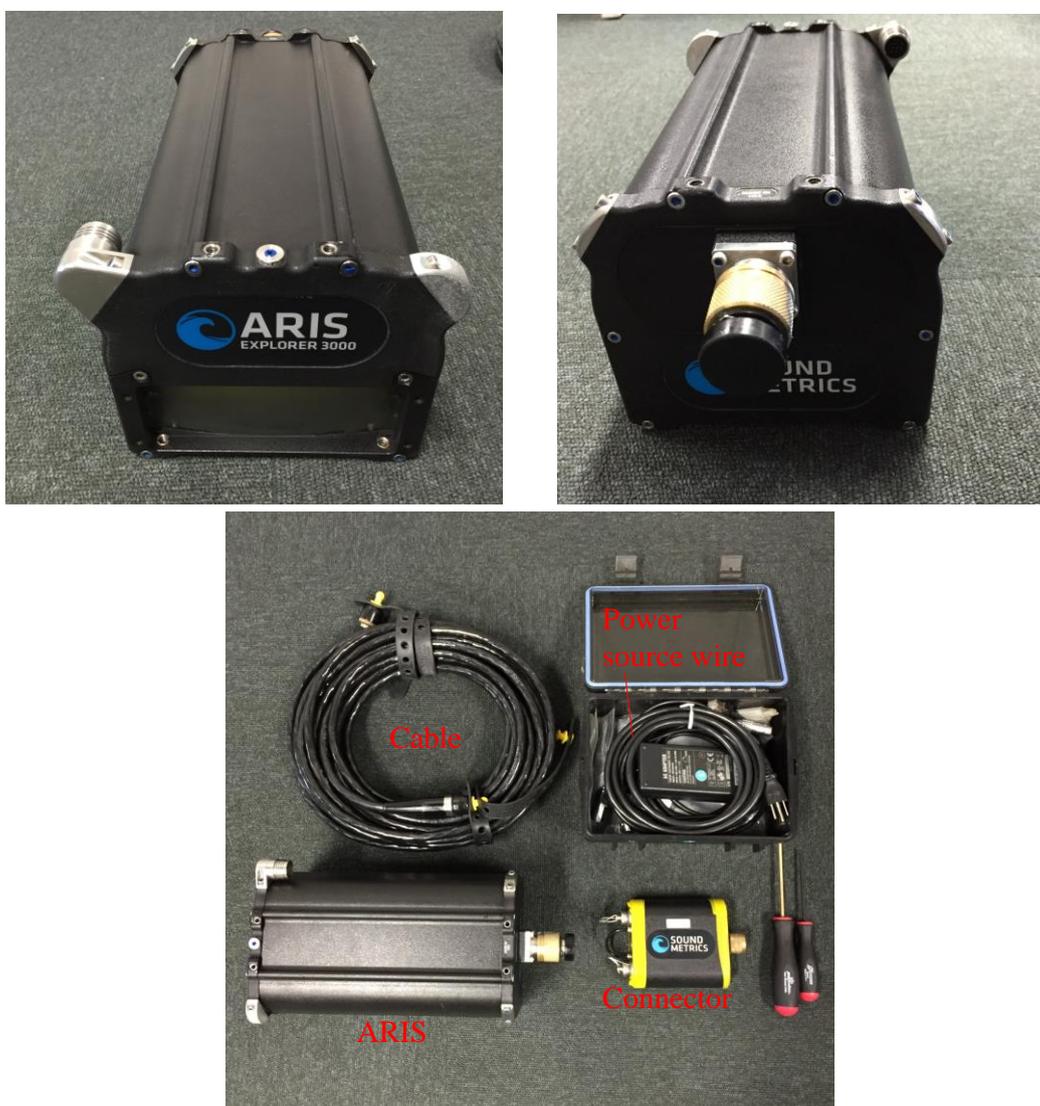


Fig. 6 Appearance of ARIS and its accessories (at Asada lab)

The default lens makes ARIS work with 14° vertical acoustic beam width. As Fig. 7 shows, 3° concentrator lens help ARIS narrow its beam for better observation of fish, which is connected to the proposed observation method of this study.



Fig. 7 Acoustic concentrator lens of ARIS [3]

With 128 beams operating at 3MHz, the ARIS Explorer 3000 collects more data than any other imaging sonar in its class. Two frequencies can be used: identification frequency 3.0 MHz, observation range 5 meters; detection frequency 1.8 MHz, observation range 15 meters. The specifications of ARIS is as Table. 1 shows.

Table. 1 The specifications of ARIS [8]

Depth Rating	300m
Number of Transducer Beams	128
Beam Width	0.25°
Field-of-view	30° x 14°
Frame Rate	Up to 15 frames/sec
Range Resolution	Down to 3mm
Weight in Air	5.17 kg
Weight in Water	1.06 kg
Dimensions	26cm x 16cm x 14cm
Cable Length	Up to 150m

POWER REQUIREMENTS	
For Supply Input	48 Volts
ARIS Explorer	80 Watts
ARIS Explorer with AR2	150 Watts

POWER CONSUMPTION	
ARIS Explorer	20 Watts typical
ARIS Explorer with AR2	35 Watts typical

COMPUTER REQUIREMENTS
Recommended PC Configuration
Windows 7 (32 or 64-bit)
DirectX 11 compatible graphics
Multi-core CPU
4GB RAM
512 MB Video RAM
PC Interface: 100BaseT Ethernet

2.2. Proposed observation method for survey

Conventional observation method is as Fig. 8 shows, ARIS provides $30^\circ \times 14^\circ$ field-of-view for observation. However, in this case, it is hard to get high-quality acoustic images. Fig. 9 is an acoustic image taken by conventional observation method with acoustic video camera ARIS. From the image, we can know that fish classification on the image of such quality will be very difficult.

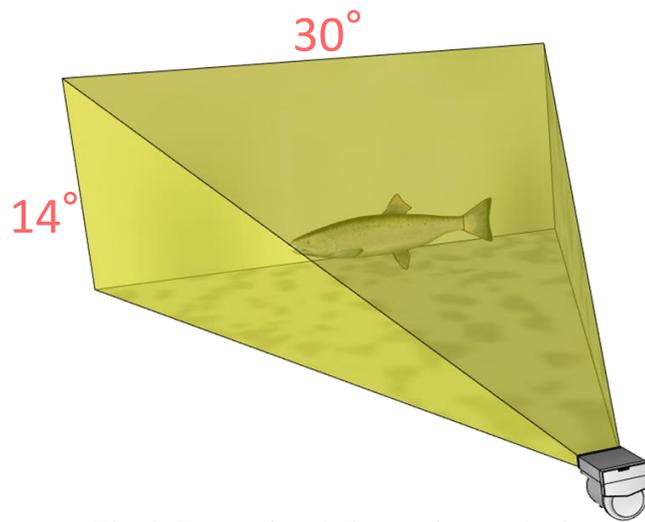


Fig. 8 Conventional observation method

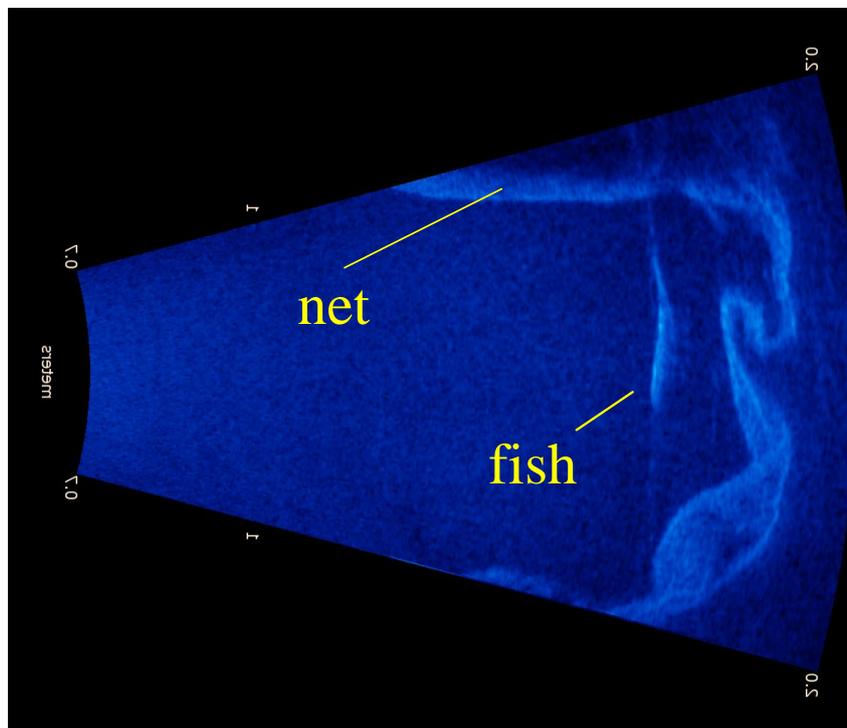


Fig. 9 Acoustic image obtained by conventional method

In this condition, a new observation method was proposed. That is, ARIS was rotated by 90° , and mounted with a 3° concentrator lens. In proposed observation method, ARIS offers $3^\circ \times 30^\circ$ field-of-view beam for observation as Fig. 10 shows, which matches well with fish's shape.

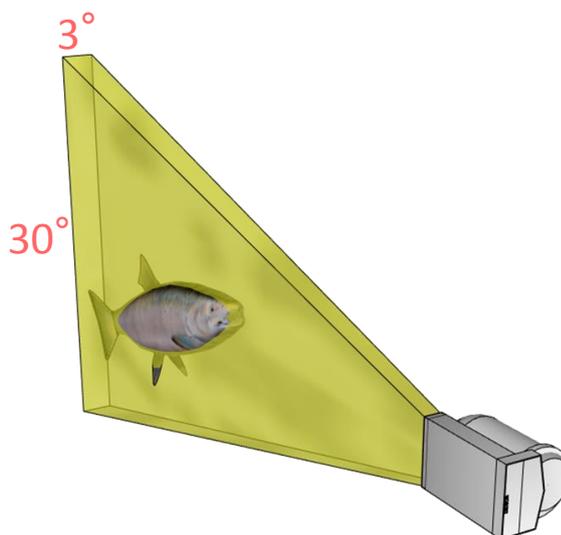


Fig. 10 Proposed observation method

And then, by proposed observation method, high-quality acoustic image of fish could be obtained as Fig. 11 shows. In this image, the head, fin and tail's shape of fish can be seen clearly.

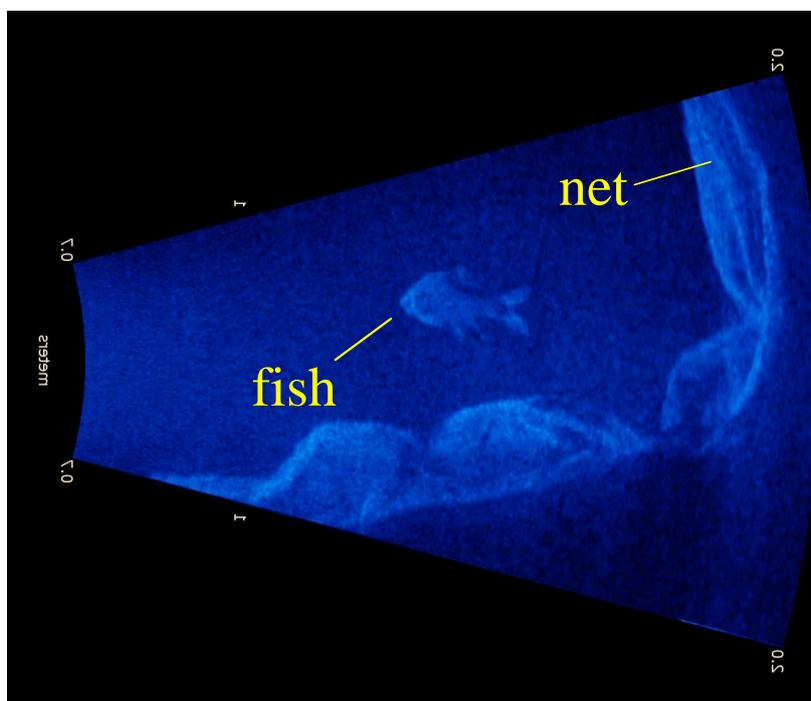


Fig. 11 Acoustic image obtained by proposed method

2.3 Theory analysis of two observation methods

By conventional observation method, the two-dimensional frame consisted of 128 horizontal beams, and the object's echo strength can be detected by receivers of sonar from the different time which backscatter echoes use. In this condition, as figure. 12 shows, from top view divided beams will help to distinguish the head and caudal fin of fish; left or right of fish will be made clear by the different time back to sonar. However, in vertical direction, from side view the top and bottom of fish cannot be distinguished clearly because backscatter echoes of either top or bottom of fish take same time to go back to sonar, and in horizontal direction there are no divided beams which can help to distinguish. Hence, we cannot get clear acoustic images of fish by conventional method.

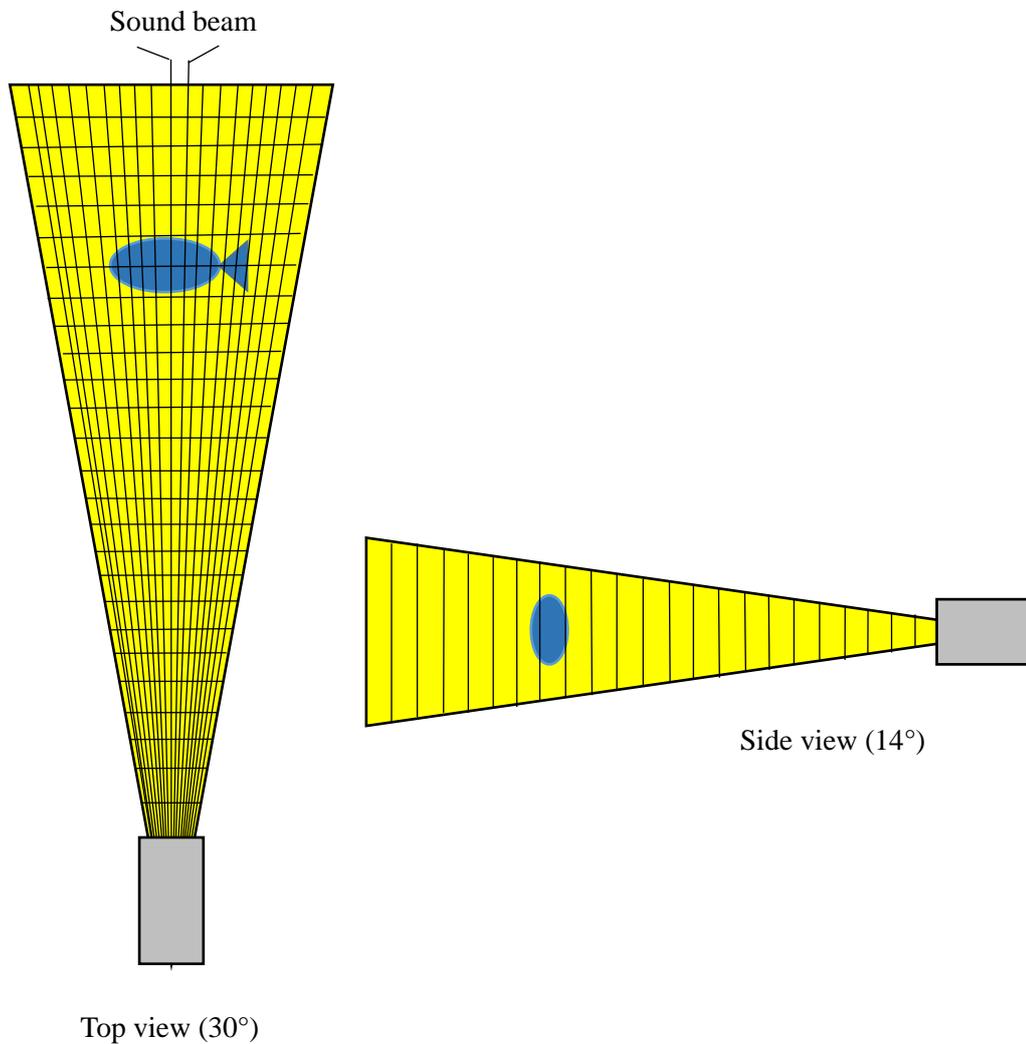


Fig. 12 Top view and side view of sonar by conventional observation method

After ARIS was rotated by 90° and mounted with a 3° acoustic concentrator lens, when fish come into the sight of sonar, as figure. 13 shows, from side view divided beams will help to distinguish the top and bottom of fish; head or caudal of fish will be made clear by the echo's different time back to sonar. In addition, because of the acoustic lens, from top view the angle was concentrated to 3° . Therefore, from left or right side of fish, only one side can be detect by ARIS clearly, the misunderstanding of left side or right side was deleted, leading to the fact that fish can be distinguished from three directions by proposed method.

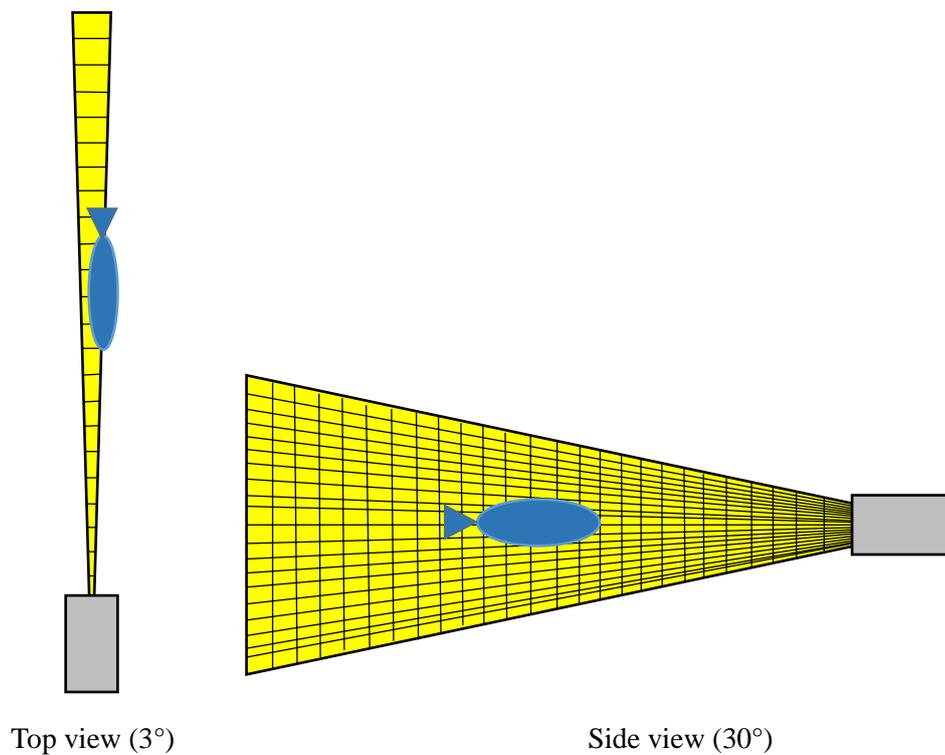


Fig. 13 Top view and side view of sonar by proposed observation method

From the theory analysis, the reason why high-quality acoustic image can be obtained by proposed observation method is clear to us. Both of the field experiments used this proposed observation method.

CHAPTER 3: FIELD EXPERIMENT I

3.1. Experiment site and settings

As shown in Fig.14, ARIS was put into fish tank and set up by proposed observation method. The size of fish tank is as Fig. 14 shows. The observation frequency of ARIS was set as 3.0 MHz (high-frequency mode), and the observation range was about 1.8 m from the imaging sonar. The two-dimensional frame consisted of 128 horizontal beams and 432 range samples. The frame rate was 8 fps. The field test was performed on 22 October 2014.

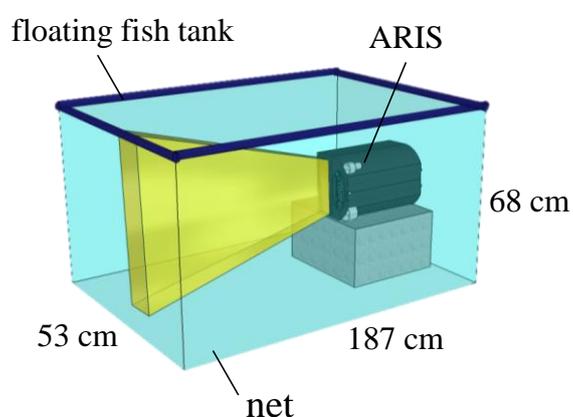


Fig. 14 Floating fish tank used for observation

As Fig.5 and Fig. 15 show, the site of field experiment was a small eutrophic pond adjacent to Izunuma Lake, where latitude is 38.722 degrees north and longitude is 141.092 degrees east, and the water depth was 0.7 meter. The environmental conditions in the pond were similar to those in Izunuma Lake and the observation equipment was convenient to set there. Therefore, the pond was chosen as the field site for testing the observation methods.



Fig. 15 Field site for experiment- a pond adjacent to Izunuma Lake

For experiment I, 4 species of fish were prepared as Fig. 16 shows. The alive fish samples were measured and then put into fish-tank respectively for observation by proposed method for about ten minutes. The sizes of fish samples are as the Table. 2 shows.



(a) Japanese crucian carp



(b) carp



(c) Japanese barbel



(d) Black bass

Fig. 16 4 species of fish used for experiment

Table. 2 Sizes of fish samples in experiment I

Species	Length[mm]	Height[mm]
Japanese crucian carp	210	80
carp	300	90
Japanese barbel	380	65
black bass1	180	48
black bass2	180	46
black bass3	180	46

3.2. Data processing method

3.2.1 Pre-processing

As Fig. 17 shows, acoustic images of very high quality were obtained by proposed observation method. Because backscattering strength from head and fin of fish were strong, characteristics of different species can be distinguished clearly. And then the complete and clear sonar images were taken as the template images to match with the object images afterwards. After that, 10 frames of fish images were chosen as object images for matching as Fig. 18.

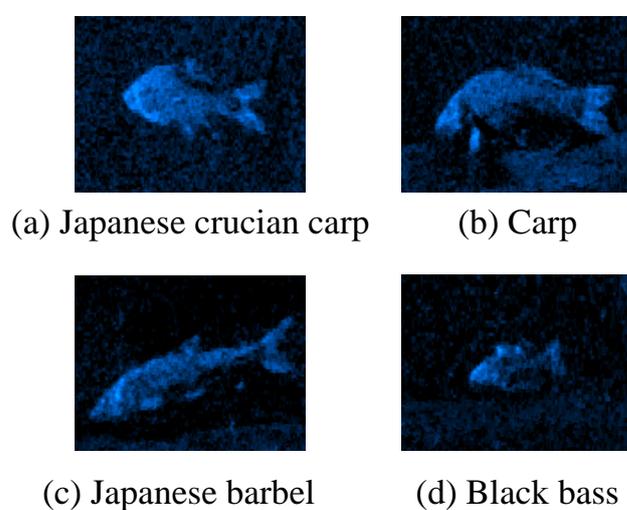


Fig. 17 Template images for template matching

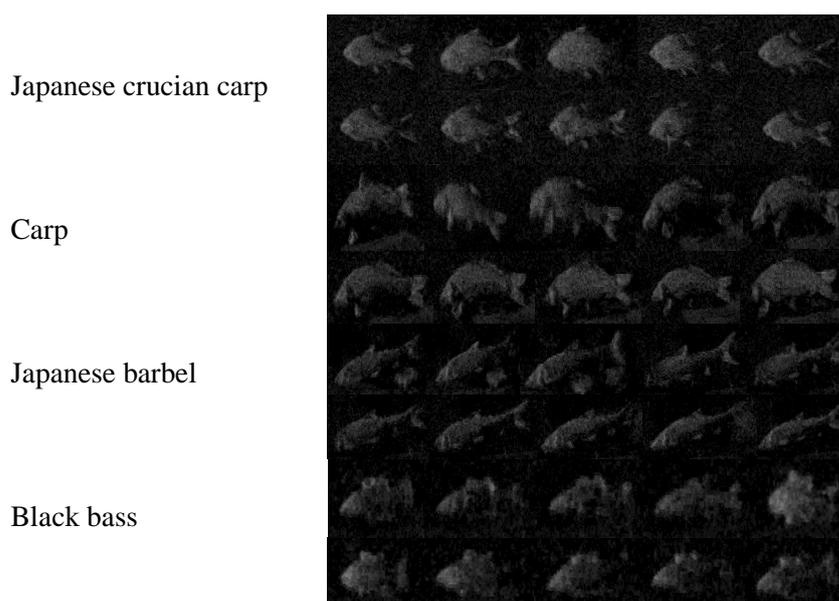


Fig. 18 Object images for template matching

For experiment I, ARIS raw data is consisted from 128 horizontal beams and 432 range samples. The raw data were restored in 128×432 matrix like the shape of rectangular. In order to expand the rectangular to sector shape, here bilinear pixel interpolation method [9] was used for the reconstruction of two dimensional acoustic image. The value of pixel between the spilt beams were calculated by 4 neighboring pixels' distance and value along beams.

Next, the fixed pattern of the background (image without fish) was subtracted from every frame of raw data image on ratio. The backscatter echoes from fish tank were removed and the fish in subtracted images are clearer than raw images as Fig. 19 and Fig. 20 shows.

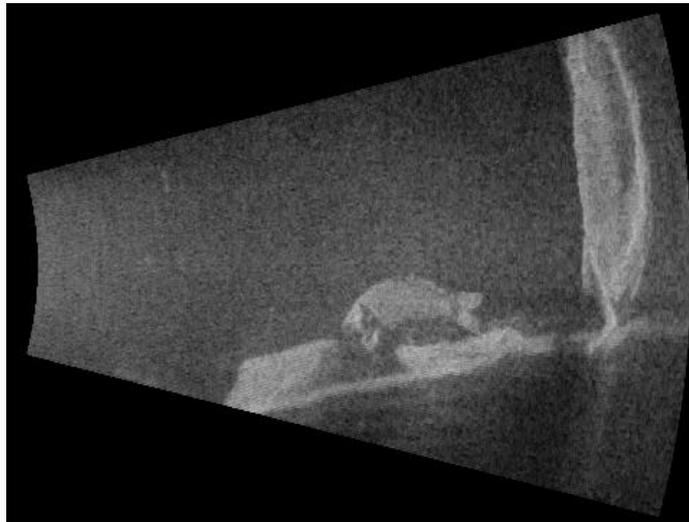


Fig. 19 Sample sonar image before background subtraction (carp)



Fig. 20 Sample sonar image after background subtraction (carp)

After background subtraction, every pixel whose intensity is below a threshold will be treated as noise and set to zero (black level). Here the threshold can be adjusted and it would be determined by the background noise level. And then the image of fish will be cut from processed image which is used for template matching, fish object images which were cut are as Fig. 18 shows.

3.2.2 NCC method

NCC (Normalized cross correlation) method on fish's template matching has been researched by Asada lab. [10] This method is usually used on template matching for detection or recognition of target in image or picture. Flow chart of ARIS data processing by NCC method is as Fig. 21 shows.

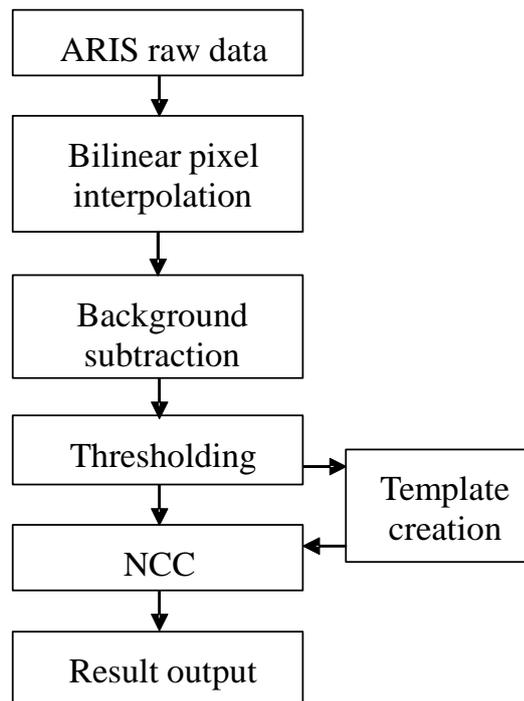


Fig. 21 Flow chart of ARIS data processing by NCC method

The result of NCC (R_{NCC}) is not affected by the variation of image brightness and easy to be evaluated the results respectively. The result of NCC is represented as degree of similarity R_{NCC} and calculated as Formula. 1,

$$R_{NCC} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j)T(i, j)}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i, j)^2} \sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} T(i, j)^2}} \quad \text{Formula. 1}$$

Here i and j stand for the pixel positions in the image, M and N stand for the numbers of pixels of template image in horizontal and vertical direction, I and T stand for object image and template image, respectively. When the template image perfectly matches with object image, R_{NCC} will shows a maximum value of 1.0. As Fig. 22 shows, template image was moved along the arrow direction, when it matched with the right pattern, R_{NCC} showed the highest value.

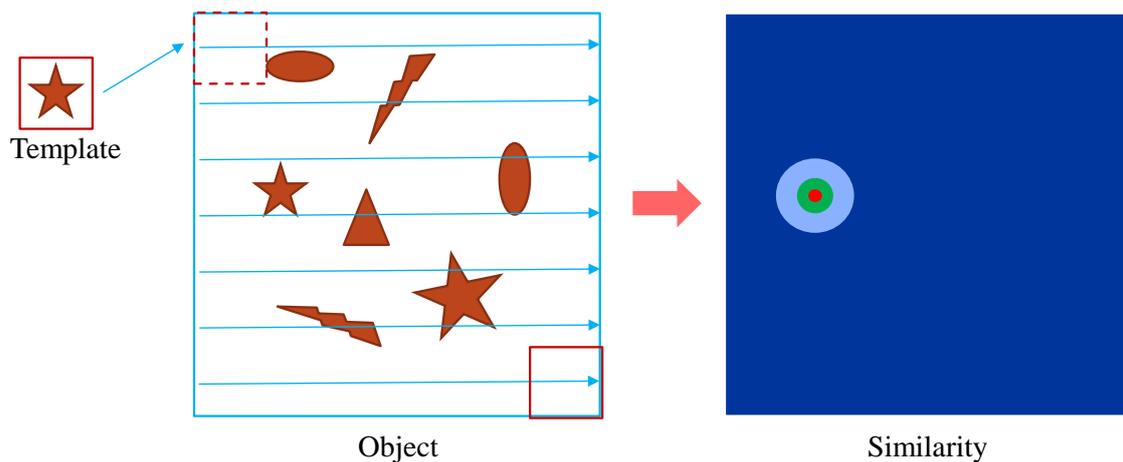


Fig. 22 Test of template matching by NCC method

3.2.3 SIFT method

SIFT [11] (Scale-Invariant Feature Transform) method has been developed for optical images' matching for many years. This algorithm has the advantages of being robust to pattern's scale and angle change, which is suitable to template matching between fish. Therefore, this study applied SIFT method into acoustic images for the first time. The flow chart of image processing method based on SIFT is as Fig. 23 shows.

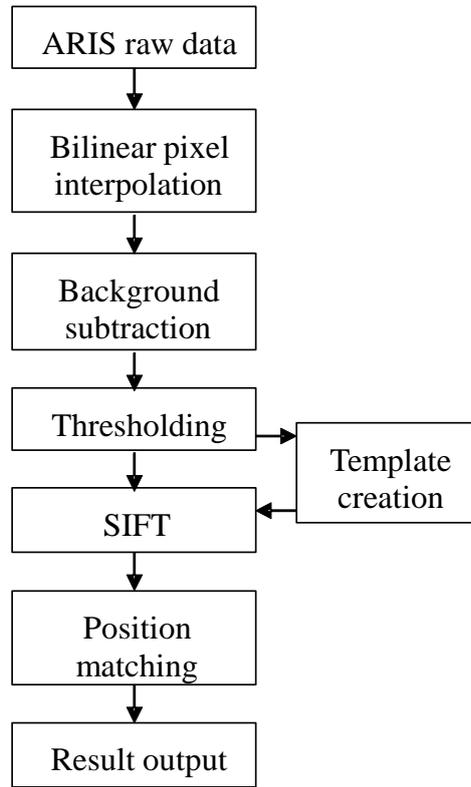


Fig. 23 Flow chart of image processing method based on SIFT

After the pre-processing, SIFT method mainly includes 4 steps as follows, scaling and detection of key point, localization of key point, calculation of orientation at key point and description of feature at key point. Firstly, a testing experiment was done by SIFT method with the sample picture as Fig. 24 shows.



Fig. 24 Sample picture tested by SIFT method

3.2.3.1 Scaling and detection of key point

Firstly, DoG (Difference of Gaussian) images were made by changing the value of σ as Fig. 25 shows according to the Formula. 2 and Formula. 3.

$$D(x, y, \sigma) = ((G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)) \quad \text{Formula. 2}$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad \text{Formula. 3}$$

Here x and y stand for pixel positions, σ is the factor of dispersion, k is the scaling factor, and I stands for object and template images.

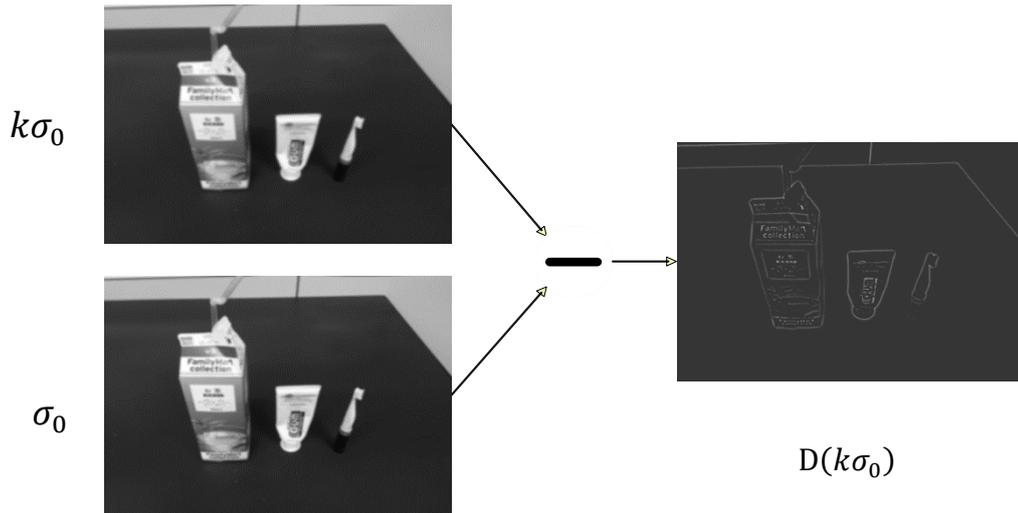


Fig. 25 DoG (Difference of Gaussian) images by changing the value of σ

And then, I checked value of pixel of DoG images. The value was compared with adjacent 26 pixels (Fig. 26). When target pixel shows extremal value, it was treated as candidate of key point.

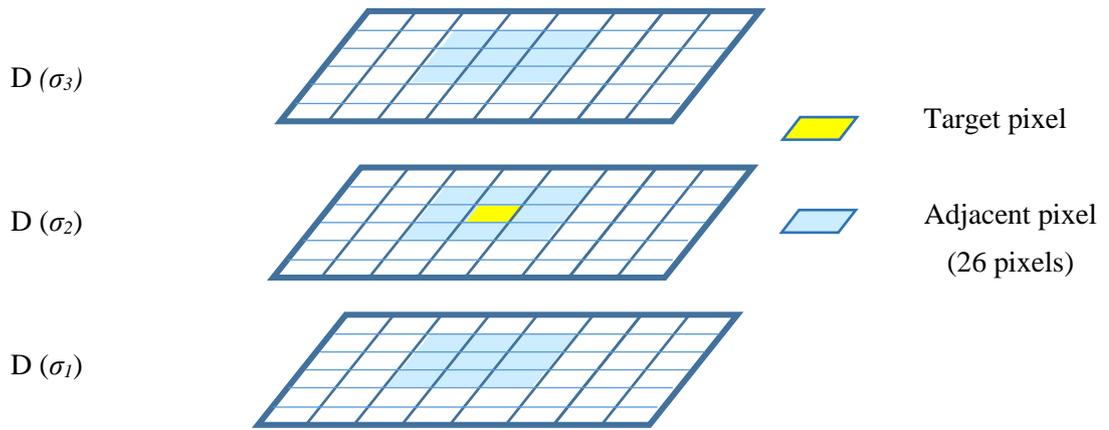


Fig. 26 Target pixel with its 26 adjacent pixels

To detect the key points in various spatial scales, dispersion factor σ was changed and many DoG images were created. In order to reduce the expanding of Gaussian window, image was down sampled when $2\sigma_0$ was instead of the initial value as Fig. 27 shows.

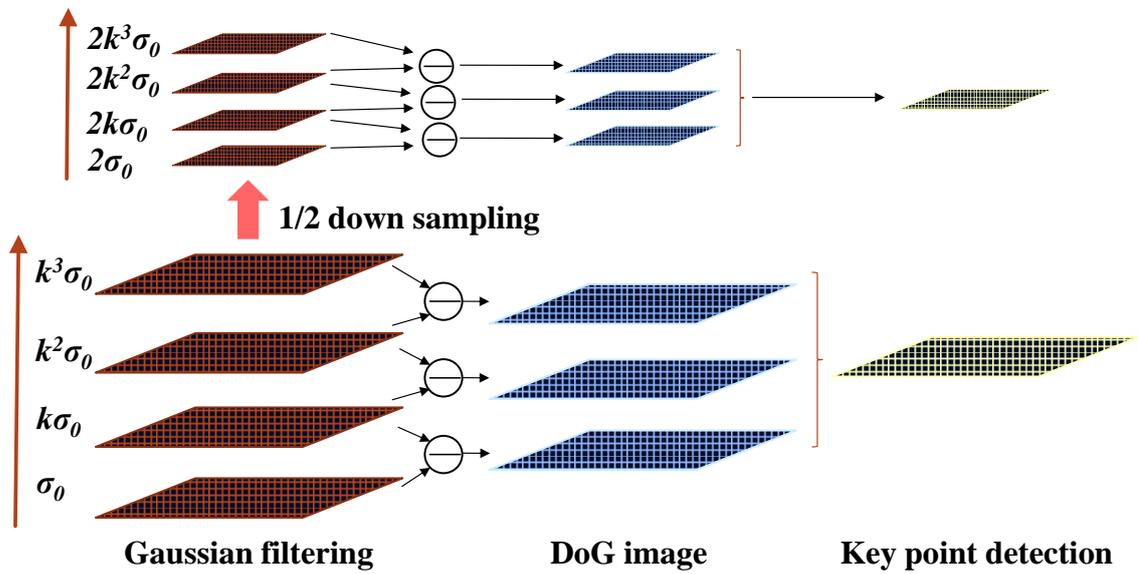


Fig. 27 DoG images was down sampled by changing σ

And then many key points were detected as Fig. 28 shows.

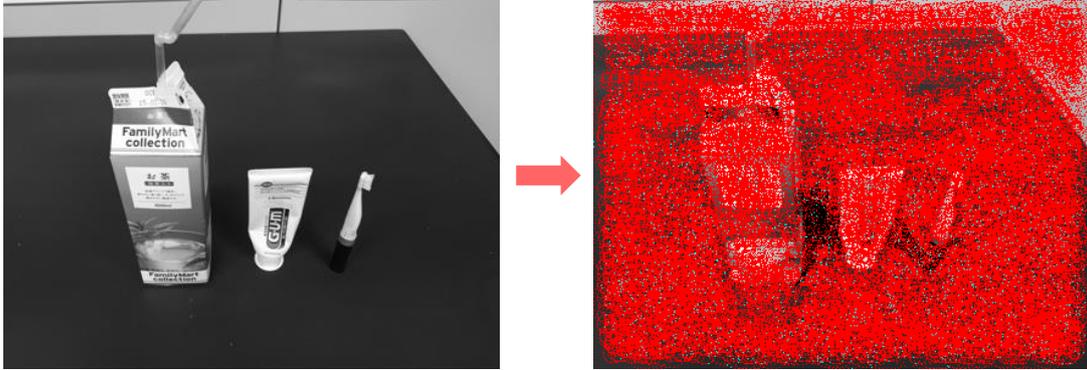


Fig. 28 Key points were detected

3.2.3.2 Localization of key point

To localize the key points, the key points on edge were removed based on the Hessian matrix (Formula. 4) of DoG image. Hessian matrix was often used to check the pixel feature in corner or edge of images. Key points were deleted by proportion of eigenvalues with a threshold as Formula.5 shows. After that, the key points on edge were removed as Fig. 29.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

Formula. 4 Hessian matrix

$$\frac{\text{Trace}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma\beta + \beta)^2}{\lambda\beta^2} = \frac{(\gamma + 1)^2}{\gamma}$$

$$\frac{\text{Trace}(H)^2}{\text{Det}(H)} < \frac{(\gamma_{th} + 1)^2}{\gamma_{th}}$$

Formula. 5 Deleting condition of key points

Here D in Hessian matrix stands for second derivatives, α and β eigenvalue of matrix, γ is the proportion of α and β , γ_{th} is a threshold value for judgement.

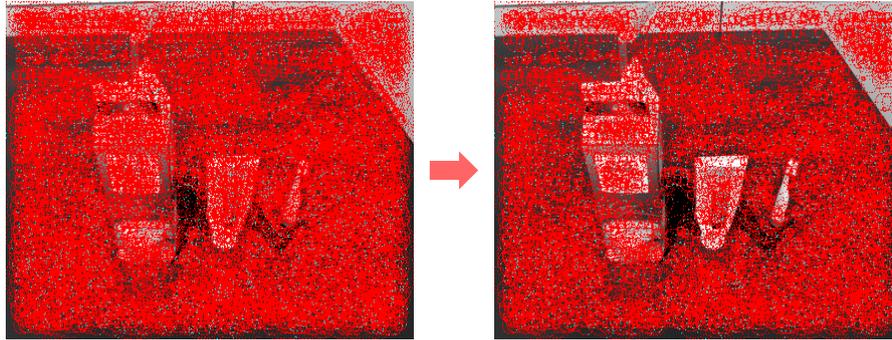


Fig. 29 Key points on edge were removed

And then small value of DoG image was removed by threshold and high contrast key points were remained as Fig. 30 shows.

$$D(x, y, \sigma) < th_D$$

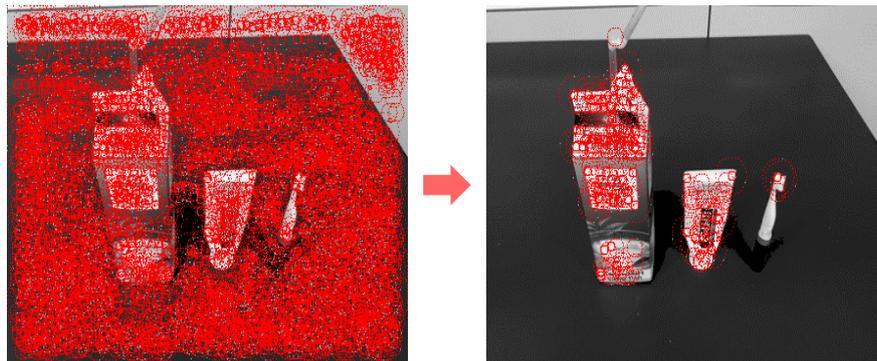


Fig. 30 High contrast key points were remained

3.2.3.3 Calculation of orientation at key point

In order to obtain the feature of key points without the effects of rotation angle, main orientation at the key points was calculated. The main orientation was determined by histogram of gradient vectors in the calculation window as Fig. 31 shows.

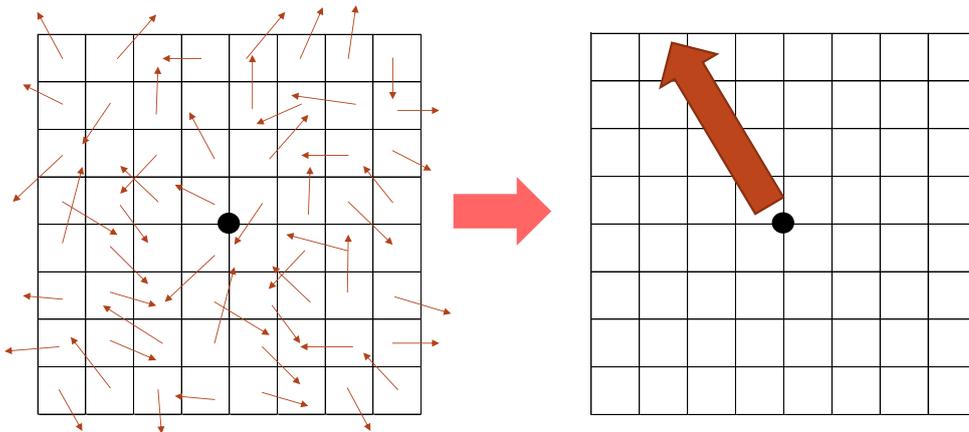


Fig. 31 Calculation of main orientation at key point

3.2.3.4 Description of feature at key point

To describe the feature at key point, calculation area was rotated following the main orientation as Fig. 32 shows.

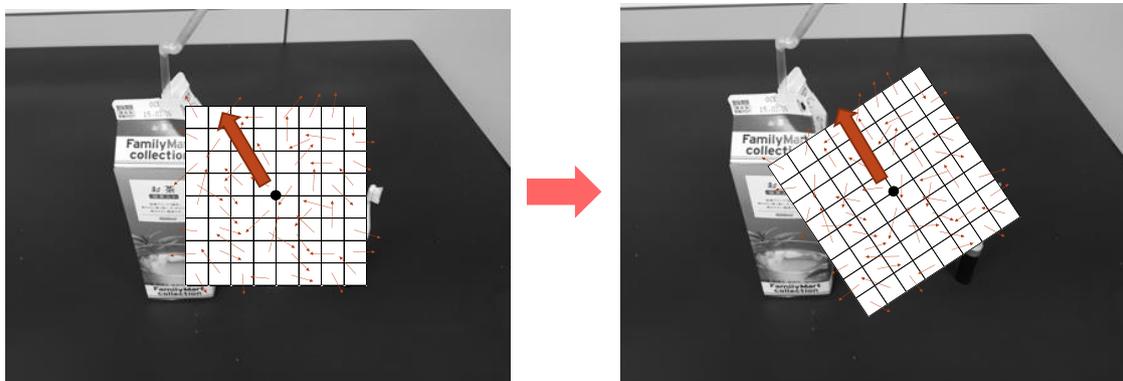


Fig. 32 Calculation area was rotated

Gradient vectors were calculated in new calculation area again. Finally, calculation area was divided to 16 blocks and calculated 8 gradient vectors at each block, thus totally 128 features were described at each key point as Fig. 33 shows.

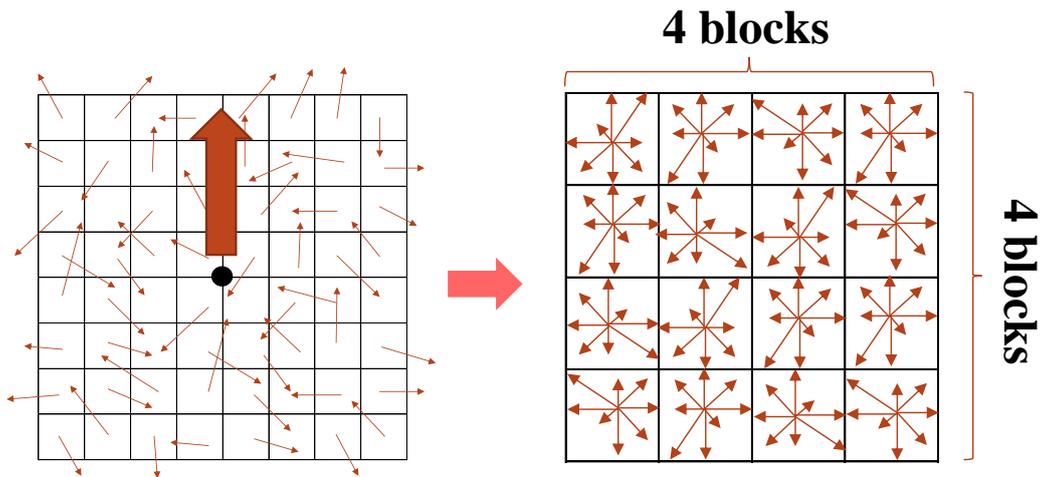


Fig. 33 128 features at each key point

Position matching between each key point was calculated by using Euclidean distance (Formula. 6). As shown in Fig. 34, same positions in both pictures were connected.

$$d = \sqrt{\sum_{i=1}^{128} (v_i^1 - v_i^2)^2}$$

Formula. 6

Here d stands for the Euclidean distance, and v stands for the calculated features.

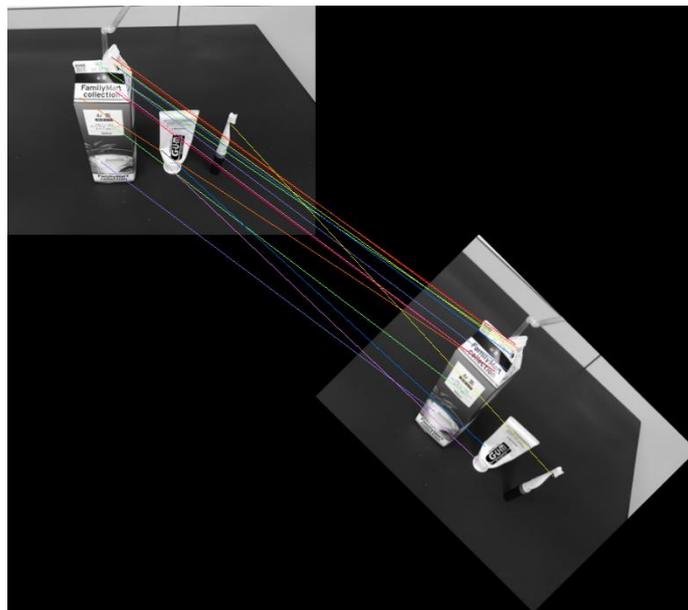


Fig. 34 Key points are connected correctly

And then I utilized the method of SIFT into the acoustic images of fish. By position matching method, this study counted how many features were connected correctly between the template images and object images in the divided areas and then calculated the matching rate. Fig. 35 is example image of matching between same species of Japanese barbel. And Fig. 36 is example image of matching between different species of Japanese barbel (template) and Japanese crucian carp (object).

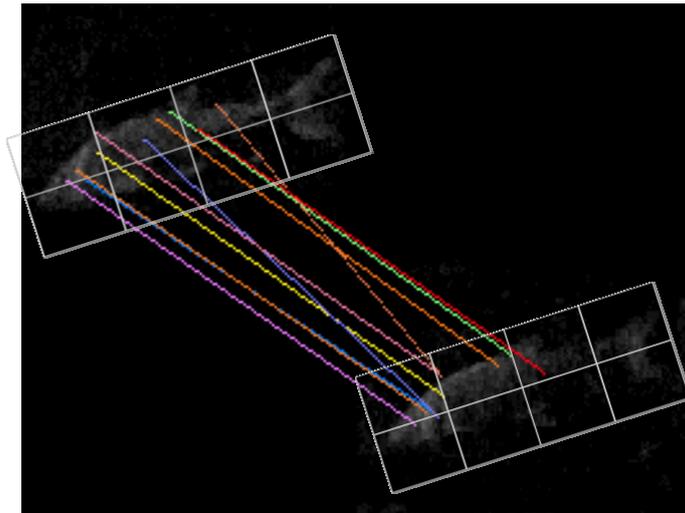


Fig. 35 Position matching between same species of Japanese barbel (matching rate 0.7)

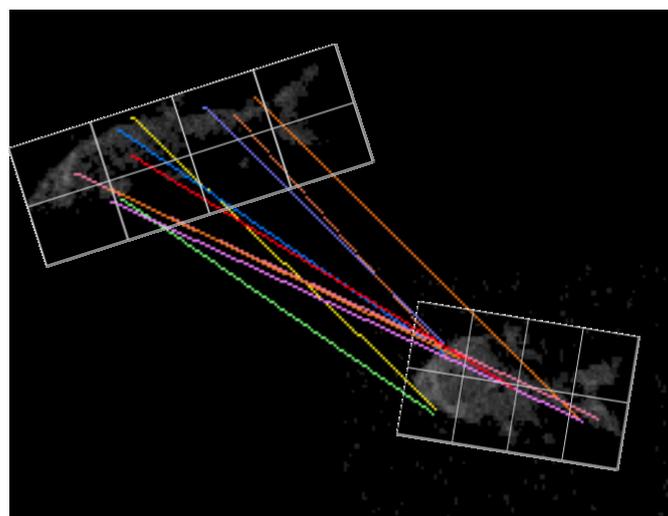


Fig. 36 Position matching between different species of Japanese barbel (template) and Japanese crucian carp (object) (matching rate 0.1)

3.3. Result and discussion

The result of NCC method is as Table. 3 shows. In the table, JCC, C, JB and BB stand for Japanese crucian carp, carp, Japanese barbel and black bass, respectively. O means object image and T means template images.

Table. 3 The result of R_{NCC} of template matching by NCC method

	JCCO	CO	JBO	BBO
JCCT	0.75	0.64	0.50	0.50
CT	0.60	0.65	0.58	0.47
JBT	0.54	0.53	0.70	0.44
BBT	0.61	0.59	0.60	0.68

From the result, it can be concluded that the value of R_{NCC} is higher when template images matching with object images of same species, which means the image processing by NCC method can detect the image of same species from the object images prepared. However, as Table.3 shows, the difference of R_{NCC} ' value is not significant between matching of same species and different species.

What's more, when the angle and scale changes, the value of R_{NCC} will change significantly even for the same species (Fig. 37).

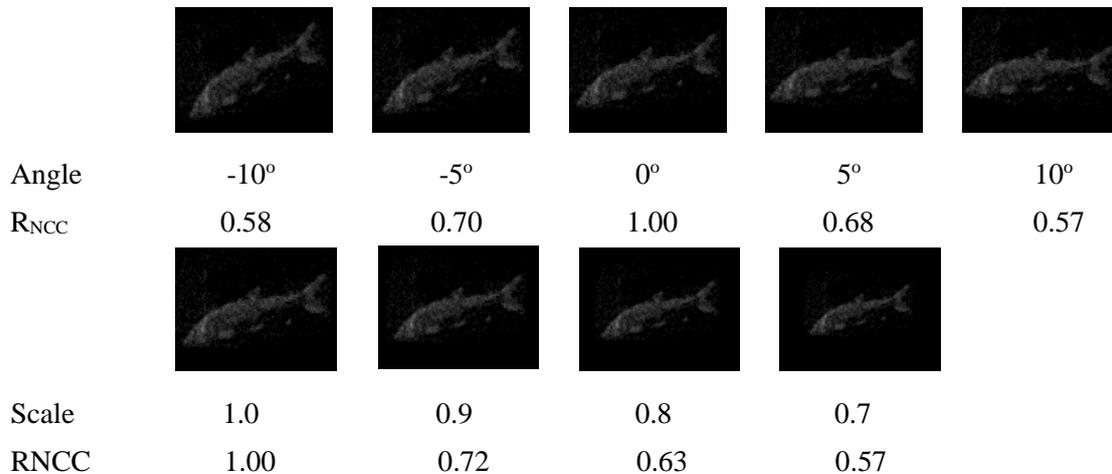


Fig. 37 The change of R_{NCC} with the change of angle and scale (Japanese barbel)

Therefore, it is known to us all that the value of R_{NCC} is easy to change with the angle of fish to acoustic video camera and the scale of fish, which needs to be improved.

The result of SIFT method is as Table. 4 shows.

Table. 4 Matching rate by SIFT method between four fish species

	JCCO	CO	JBO	BBO
JCCT	0.81	0.32	0.23	0.16
CT	0.30	0.71	0.22	0.17
JBT	0.24	0.17	0.74	0.15
BBT	0.23	0.18	0.15	0.70

And then program was improved. In the SIFT algorithm, every key points of templates matched with those of objects to find the shortest Euclidean distance, and matched key points were connected by line when the Euclidean distance is the shortest ten values of all key points. After improving algorithm, when the ratio of shortest Euclidean distance and second shortest Euclidean distance of the key point is higher than a threshold (here it is set as 0.8), that key point will be deleted because strong and stable key point is expected.

After improving, compared with Table. 5, the matching rate between same species were raised on Japanese crucian carp, carp, Japanese barbel and black bass, but the matching rate between different species almost did not change.

Table. 5 SIFT matching rate by improved program

	JCCO	CO	JBO	BBO
JCCT	0.86	0.30	0.24	0.15
CT	0.28	0.75	0.20	0.18
JBT	0.22	0.18	0.78	0.14
BBT	0.20	0.17	0.16	0.82

And then, I compared the template matching rate between NCC and SIFT method. All matching rates are shown in Fig. 38. For same species, the matching rate of NCC method is 0.70 ± 0.04 and that of SIFT method is 0.80 ± 0.06 ; for different species, the matching rate of NCC method is 0.55 ± 0.05 and that of SIFT is 0.22 ± 0.08 .

From the comparison, we can know that the difference of matching rate between correct and other ones on NCC method is 0.15, and that on SIFT method is 0.58. So matching rate's difference between correct and other ones was clearly improved by using SIFT-based algorithm, and therefore it is easier to distinguish same species of fish from others by SIFT method.

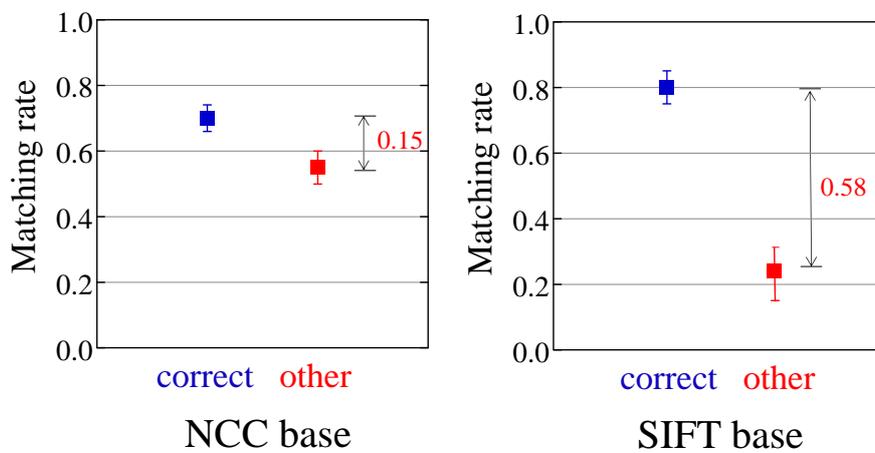


Fig. 38 Comparison of matching rate between NCC and SIFT method

Fig. 39 is sample template matching images of Japanese barbel by SIFT method when there are angle and scale change, which shows that SIFT method is robust to angle and scale changes.

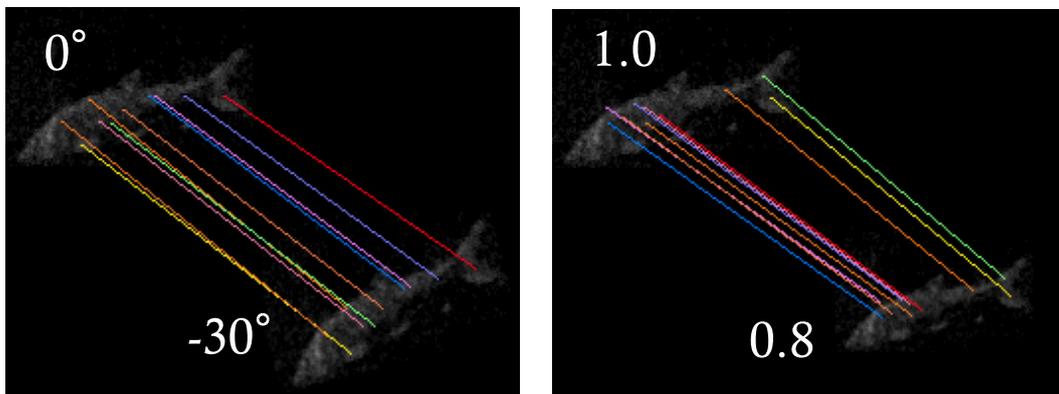


Fig. 39 sample template matching images of Japanese barbel by SIFT method with scale and angle change

And then, the matching rate of NCC method and SIFT method were compared when there are scale and angle changes in object images. As Fig. 40 shows, SIFT method is more robust to scale and angle change than NCC method.

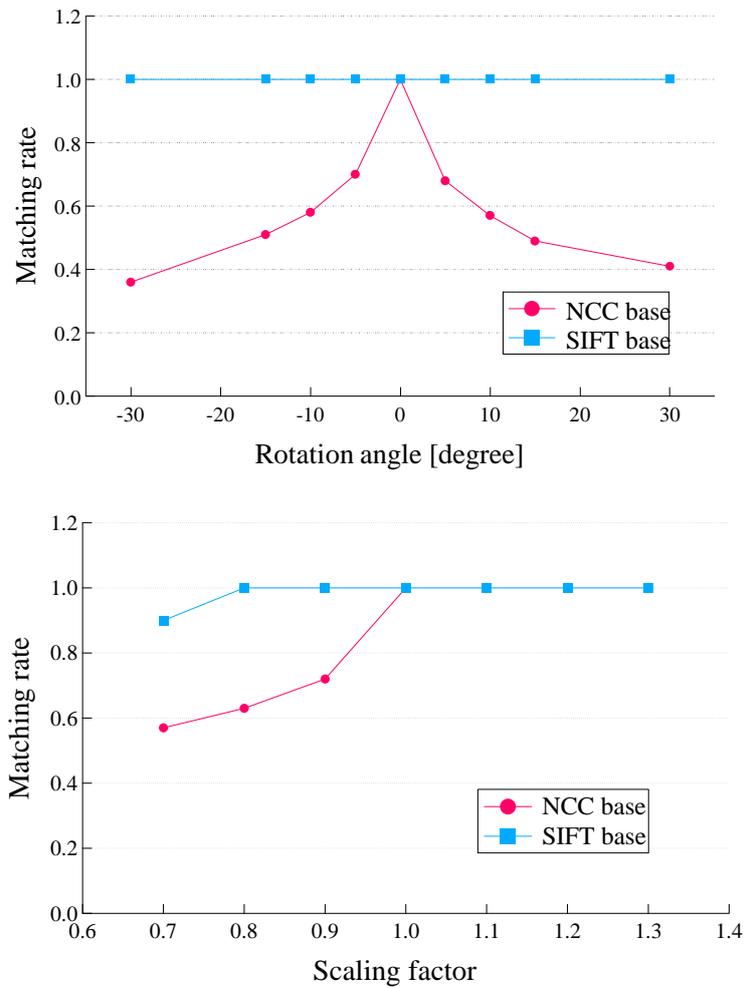


Fig. 40 Comparison of NCC and SIFT method to scale and angle change

And then, a threshold (average matching rate of same and different species) was set to help for the fish classification with the object images chosen in this experiment (Fig. 41). Here, when matching rate is higher than 0.51, the object will be distinguished as same species; when matching rate is lower than 0.51, the object will be distinguished as different species. And then the following accuracy was obtained: Japanese crucial carp 100%, carp 92.5%, Japanese barbel 92.5% and black bass 95%.

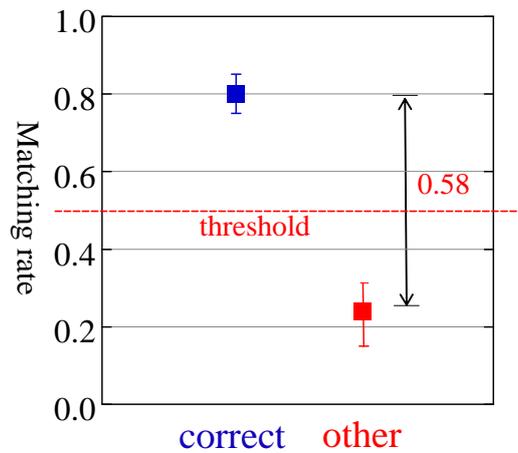


Fig. 41 A threshold was set to help for the fish classification by SIFT method

3.4. Potential problems and future work

It is the first time to apply the SIFT algorithm into acoustic image processing. High-quality images were obtained by high-resolution acoustic video camera with proposed observation method. We have known that SIFT based image processing method is more robust to angle and scale changes than NCC method, and it is easier for SIFT method than NCC to distinguish fish of same species from others.

However, the data amount was small in experiment I because the observation time for each fish was about 10 min. And the effect of individual differences of same species on fish classification has not made clear in this experiment. In addition, as Fig. 42 shows, even in the improved algorithm, when template matches with object images which are without head or caudal fin, the matching rate will be unstable. Because of the narrow beam of ARIS, it is easy to obtain imperfect images by ARIS. Although high accuracy has been got with the object images chosen in experiment I, in the future, it is important to make clear the effect of individual difference and raise the matching rate on imperfect images.

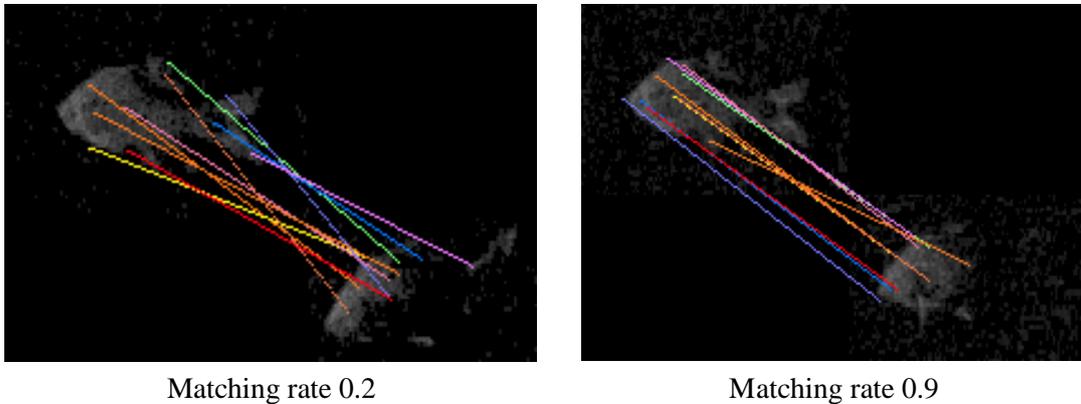


Fig. 42 matching rate on imperfect object images of same species (Japanese crucian carp)

CHAPTER 4: FIELD EXPERIMENT II

4.1. Objective of the experiment II

From the result and discussion of experiment I, we can know that SIFT method is robust to scale and angle change of fish, which is a good property for fish classification. However, as Fig.43 shows, because the 3° acoustic beam is narrow, when fish comes into the sight of ARIS with an angle, imperfect images (images with only part of fish body) will be obtained by sonar. When the template matches with such imperfect images, the matching rate will be unstable. In order to acquire the specific connection between fish's angle with sonar's beam and corresponding sonar image, furtherly raise the accuracy of fish classification, field experiment II was carried out. Besides, experiment II would test the effect of individual difference on SIFT method which was not noticed in experiment I, important to fish classification of large data amount.

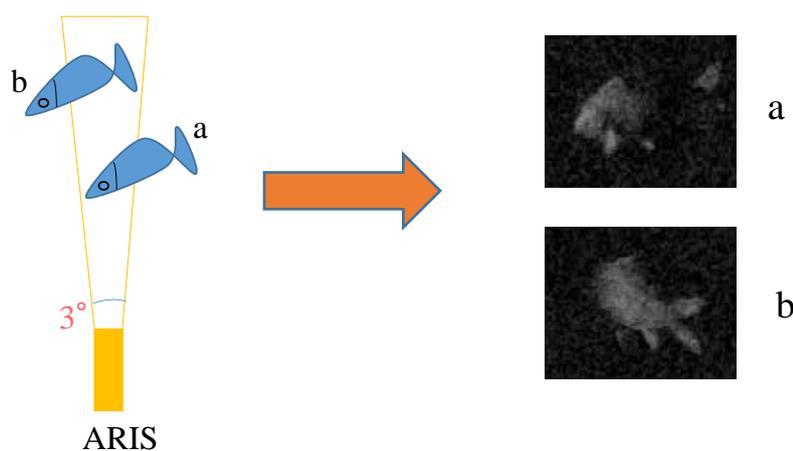


Fig. 43 Imperfect acoustic images

4.2. Experiment site and settings

The experiment can be divided into two parts.

The first part of it was carried out in the Miyagi Prefectural Izunuma-Uchinuma Environmental Foundation on 10th, November of 2016, near to Izunuma Lake. As Fig. 44 shows, it was an outdoor experiment in fish tank. The size of fish tank is about 200cm x 80cm x 60cm. Optical digital camera (OLYMPUS TG-3, OLYMPUS IMAGING CORP.) was put above the tank to observe the angle between fish and sonar' center beam.

ARIS was put in one side of the tank bottom as Fig. 45 shows, using proposed observation method same to experiment I (ARIS was rotated by 90° and mounted with a 30° acoustic concentrator lens). A calibration line was set to be parallel with center beam of ARIS in order to help watch the angle between fish and ARIS. The net was put in front of ARIS to restrict fish's swimming area for more effective observation.

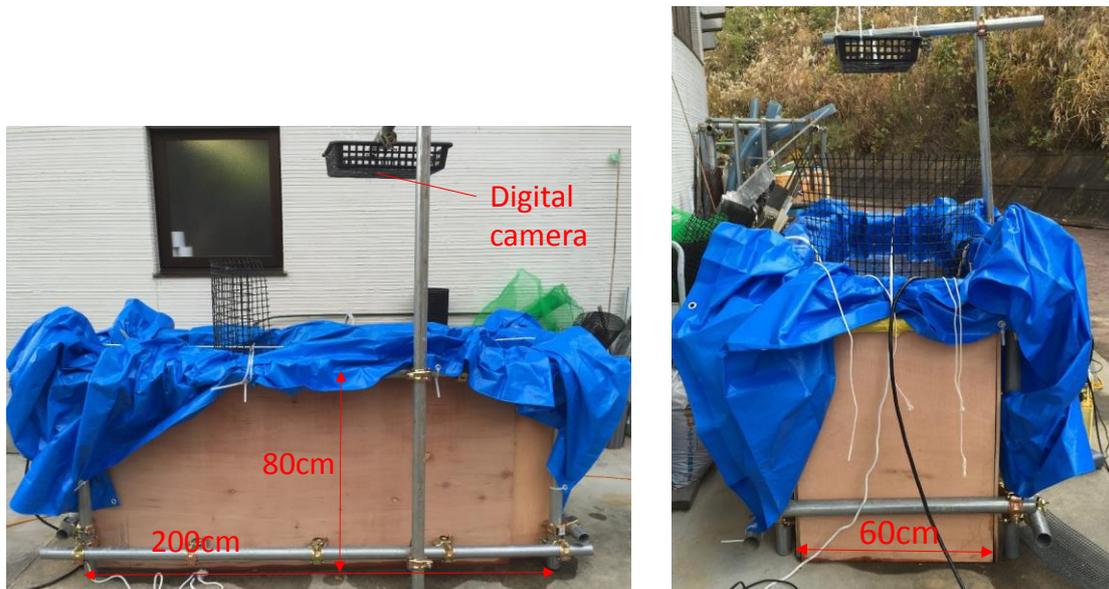


Fig. 44 Fish tank for experiment II



Fig. 45 ARIS and calibration line

In the first part of experiment II, 5 species of fish were prepared for observation and every fish was observed for about 30 mins. Optical images are as Fig.46 shows. Sizes of fish are as Table. 6 shows.



Fig. 46 Optical images of fish for the observation

Table. 6 Sizes of fish for observation

Species	Length[mm]	Height[mm]
Carp	310	100
Japanese crucian carp1	410	140
Japanese crucian carp2	360	120
Japanese crucian carp3	370	110
Japanese crucian carp4	190	60
Japanese crucian carp5	170	60
Japanese crucian carp6	330	120
Japanese barbel 1	500	100
Japanese barbel 2	400	80
Snakehead mullet 1	630	90
Snakehead mullet 2	620	100
Black bass1	150	40
Black bass2	100	50
Black bass3	180	50

The second part of the experiment was done at the same site on 11th, November of 2016, aiming to obtain optical images of fish sample in good shape (frozen fish), and observe the fish sample underwater in different angle by ARIS. The frozen fish samples' sizes are as Table. 7 shows. The pictures for measuring size are as Fig. 47 shows.

Table. 7 Sizes of frozen fish samples for observation

Species	Length[mm]	Height[mm]
Black bass1	450	130
Black bass2	430	120
Japanese crucian carp	340	110

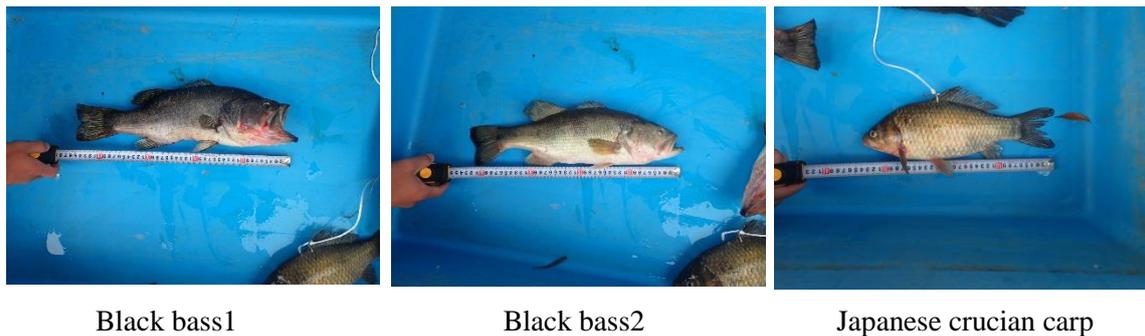


Fig. 47 Optical images of frozen fish samples

Therefore, firstly I took digital pictures of the fish sample from many space angles by digital acoustic camera (OLYMPUS TG-3) in order to build a 3-D model of it. Some example pictures are as Fig. 48 shows.

After that, fish samples were put underwater for observation hanging with a rope by ARIS for about 15 mins, respectively



Fig. 48 Optical digital pictures of fish for construction of 3-D model (Japanese crucian carp)

4.3. Data processing

In the first part of experiment II, because of the cold water temperature, fish always moved on the bottom of the fish tank, which may cause some effect on the matching rate. Because fish swam freely and the water temperature was low, all angles' image data of fish couldn't be obtained. I combined the sonar images with the optical video above at the same time, summarized the sonar images of fish at certain angles. And then complete and clear images of fish were chosen as template images, 10 object images were chosen at every angle where data were obtained. After that, template matching at those angles were carried out by SIFT method (as 3.2.3).

4.4. Result and discussion

Fig. 49 shows the corresponding connection between fish's angles (from 0° to 180°) and sonar images (Japanese barbel for example).

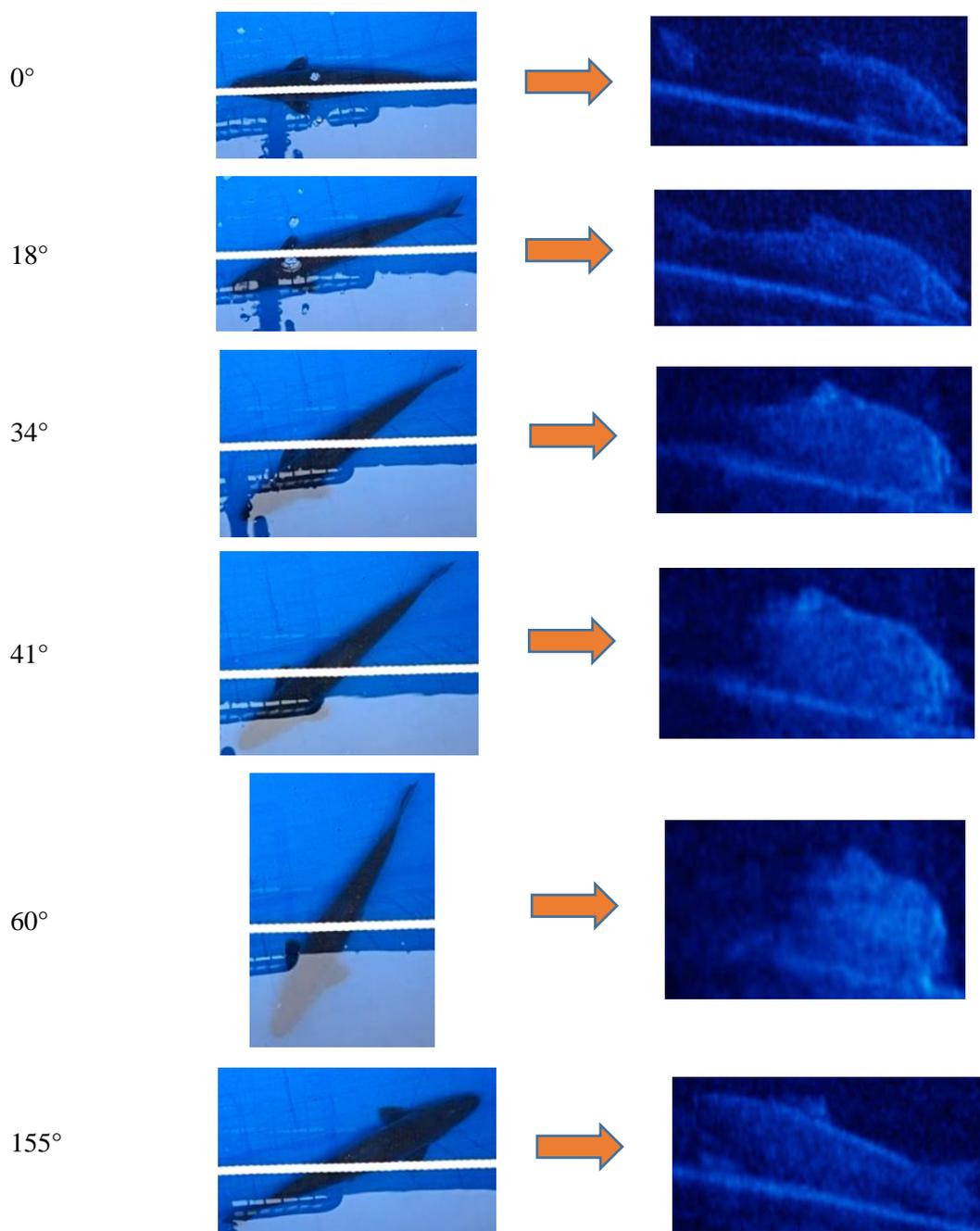




Fig. 49 Corresponding connection between fish's angles and sonar images

From the corresponding angle connections we can conclude as follows:

At the angle (between fish body and center beam of ARIS) of 0° , sound beams almost cannot reach fish's back caudal fin part of body because of their streamlined body shape, hence fish usually display only front part body in sonar images.

With the angle's increasing, sonar images gradually display the whole body of fish. At the angle of 15° - 30° , the sonar images of fish are complete and clear.

When the angle exceed 45° , because of the narrow beam of ARIS, fish cannot display a whole body in sonar images in most occasions.

When fish stay at angle between 75° and 105° , the acoustic backscatter intensity becomes so strong that people can hardly judge if the pattern in sonar image is a fish or not.

As Fig.49 shows, the sonar images when fish stay at 0° - 90° is opposite to images when fish stay at 91° to 180° , which means that the position of head and caudal fin will change. Besides, between 105° and 180° , fish's head in the sonar image is not clear because the energy loss in longer transmitting distance and weaker backscattering strength.

However, the angle between fish body and center beam of ARIS is just one factor of sonar images forming. The extent that how much fish body is inside the beam range will also affect the sonar images heavily. For instance, even at the angle of 15° , fish may display part body in the sonar images because just part of fish body is inside the beam range at that time.

And then, template matching of different angles were carried out on the fish prepared by SIFT method, which was used in experiment I. Complete and clear images were chosen as template images, and then 10 frames of images were chosen as object images at every angle where data were obtained. The results of average matching rate are as Table.8 shows. Fig. 50 shows the result in Table. 8.

Table. 8 Average matching rate at different angles

Species	Angle ($\pm 5^\circ$)	Matching rate
Japanese crucian carp1	0°	0.41
Japanese crucian carp1	15°	0.58
Japanese crucian carp1	30°	0.78
Japanese crucian carp1	45°	0.54
Japanese crucian carp1	60°	0.34
Japanese crucian carp1	75°	0.10
Japanese crucian carp1	90°	0.05
Japanese crucian carp1	180°	0.30
Japanese crucian carp2	0°	0.44
Japanese crucian carp2	15°	0.70
Japanese crucian carp2	30°	0.55
Japanese crucian carp2	45°	0.40
Japanese crucian carp2	60°	0.33
Carp	0°	0.50
Carp	15°	0.55
Carp	30°	0.77
Japanese barbel 1	0°	0.53
Japanese barbel 1	15°	0.60
Japanese barbel 1	30°	0.90
Japanese barbel 1	45°	0.44
Japanese barbel 1	60°	0.29
Japanese barbel 1	90°	0.05
Japanese barbel 1	150°	0.33

Japanese barbel 1	180°	0.32
Snakehead mullet2	0°	0.41
Snakehead mullet2	15°	0.86
Snakehead mullet2	30°	0.73
Snakehead mullet2	45°	0.58
Black bass3	0°	0.62
Black bass3	15°	0.70
Black bass3	30°	0.69
Black bass3	150°	0.28

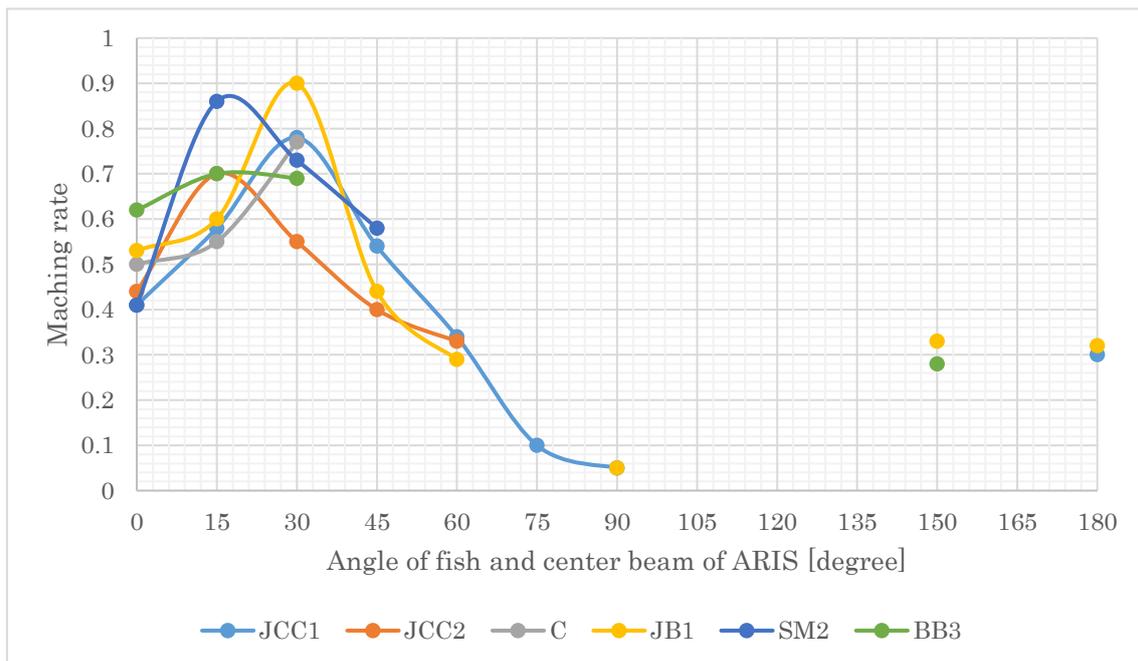


Fig. 50 Matching rate of different fish individuals on certain angels

During the observation, fish moved freely but not actively because of the low water temperature, therefore all angles' data could not be obtained in that situation. In addition, some fish was observed by ARIS but the angle data was not recorded because of the camera's power problem, such as Japanese crucian carp3 and Japanese barbel 2. Snakehead mullet1 always swim with "S" shape, it is hard to judge its angle with ARIS's center beam from above. Some other fish just did

not or hardly appear in the sight of ARIS. Therefore, the data of matching rate and corresponding angle were summarized as Table.8 and Fig.50 shows.

Compared to experiment I, in experiment II, belly part and fins were not seen very clearly as the fish were always staying in the bottom of fish tank because of the low water temperature. And it affected the matching rate to some extent.

From the results of matching rate, it can be concluded that when fish stay at the angle of near 0° , the matching rate is not very high because the sonar images usually show only front part of the fish body (imperfect images), owing to the fact that its streamlined body shape make the acoustic beam cannot reach the back part of body.

During from 15° to 30° , the template matching rate shows the highest value, which is identical to the fact that fish shows the clearest shape at 15° - 30° in Fig.49. At this range of angle, fish's body can stay in the beam of sonar completely and sonar beam can reach every part of fish. But it is also unstable as we can see from Table.8 that the matching rate of Japanese crucian carp1 at 15° is 0.58; the matching rate of Japanese crucian carp2 at 30° is 0.55; the matching rate of carp at 15° is 0.55; the matching rate of Japanese barbell at 15° is 0.60.

When the angle increase from 45° to 60° , the matching rate is not very high. Because of the narrow beam of ARIS, fish can usually show only part body in sonar images for that fish's body exceeds the range of beam (except for little fish).

From 75° to 105° , the backscattering of echo is so strong that we even cannot distinguish if the sonar image is fish or not. So the fish classification by SIFT method during this angle range is very difficult.

From 150° to 180° , the position of head and caudal fin change. And fish's head in the sonar image is not clear because the energy loss in longer transmitting distance and weaker backscattering strength. The method of counting matching rate change to counting how many lines are connected to the right body part in eight divided areas, such as head to head, caudal fin to caudal fin as Fig.35 shows. But the matching rate is low as Table. 8 shows.

And then, the template image of every individual was matched with each other and calculated the matching rate to help consider the effect that individual difference causes in fish classification by SIFT method. Table. 9 shows the matching rate between template images of same species.

Table. 9 shows the matching rate between template images of same species.

Species	Matching rate
Black bass2 and Black bass3	0.60
Japanese crucian carp1 and Japanese crucian carp2	0.60
Japanese crucian carp1 and Japanese crucian carp3	0.70
Japanese crucian carp1 and Japanese crucian carp6	0.20
Japanese crucian carp2 and Japanese crucian carp3	0.60
Japanese crucian carp2 and Japanese crucian carp6	0.70
Japanese crucian carp3 and Japanese crucian carp6	0.90
Japanese barbel 1 and Japanese barbel 2	0.60
Snakehead mullet1 and Snakehead mullet2	0.60

From the result, it can be concluded that except for the matching rate between Japanese crucian carp1 and 6, the matching rate between template images of individuals of same species is higher than 0.60, which means that individual difference will actually affect the matching rate of SIFT method, because even individuals of same species have their own sizes, density, hardness of body and so on.

Here Japanese crucian carp 6 hardly came to the sight of ARIS, so the chosen template of it was not clear enough, it was acceptable that matching rate is low with 1 and 6, although 6 has high matching rate with 2 and 3. Therefore, the result of SIFT method is still worth expectation in the future and more data of individual difference need to be tested.

4.5. Potential problems and future work

In experiment II, I have watched the corresponding connection between sonar images of fish and its angle with ARIS's center beam, which supports a good hint to the fish classification by acoustic video camera in the future. From the result of matching rate, I have found that the sonar images of fish will be affected by the angle between fish and acoustic video camera's center beam and the extent how much fish's body is inside sonar's beam.

If fish's body is not inside sonar's beam completely, the sonar images of fish are very likely to be only part of fish body (imperfect images). Fish staying at angle of 15° to 30° usually show complete and clear sonar images.

However, only one template image for matching was used until now. And it can be found that only when fish stay at good angles and show complete sonar images can the matching rate be high as expected. And the matching rate of SIFT method remains low in most occasions. It seems that we cannot get satisfying results with SIFT method by matching with only one template because it is usual for ARIS to obtain imperfect images. In the next step, I will consider raising the matching rate by a series of template images at every angle.

CHAPTER 5: SIMULATION

5.1. Objective of simulation

In experiment II, the matching rate cannot be raised by one template images with SIFT method, and both the angle with sonar's center beam and fish's position in water will affect the forming of sonar image. If we use imperfect images to match with imperfect images at same angle, the matching rate is likely to rise. In this situation, simulation can provide simulated images at different angles as template images which is convenient, and possibly raise the template matching rate of SIFT method.

5.2. Theory introduction and parameter settings

Firstly, fish will be considered as a 3-D model of closed curved surface constituted by multiple polygon triangle meshes, which will be fixed in the origin of coordinate system. Acoustic video camera will be consider as a point, which can emit straight lines of every space angle. Here, the straight lines are consider as model of the beam of acoustic video camera.

In this condition, the straight lines emitted by acoustic video camera will cross with the triangle meshes of 3-D fish model as Fig. 51 shows. Assuming that straight lines will cross the meshes in the barycenter, through calculating the backscattering strength of every mesh of fish surface, the backscattering strength will be displayed by brightness. And then, simulated images will be compared to the real sonar images.

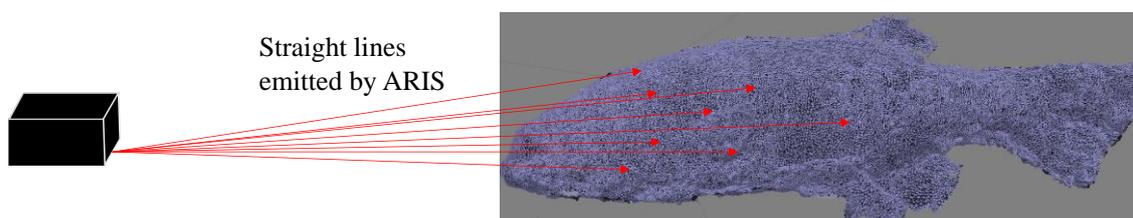


Fig. 51 Straight lines emitted by acoustic cross with triangle meshes of fish model

The parameters of model of acoustic video camera are its coordinate in space, deciding its distance and space angle to fish model. And the parameters of 3-D fish model include the number of meshes, the coordinates of every triangle mesh's 3 vertices and corresponding normal vector's coordinate.

5.3. Construction of 3-D fish model

In the second part of experiment II, optical digital pictures of the fish samples from multiple space angles were taken by digital camera as Fig.48 shows. And then through the software of PHOTOSCAN, dense cloud points' model of fish was constructed as Fig. 52.



Fig. 52 Dense cloud points' model of fish (Japanese crucian carp)

The coordinates of all dense cloud points was exported as .txt file. And then the .txt file was imported by the software of MESHLAB, after the mesh processing, a 3-D model of fish which was constituted by triangle meshes was obtained as Fig. 53 shows. The resolution of the 3-D model is about 2mm. Finally, the normal vector and coordinates of 3 vertices of every mesh and the number of meshes were exported as .stl file.

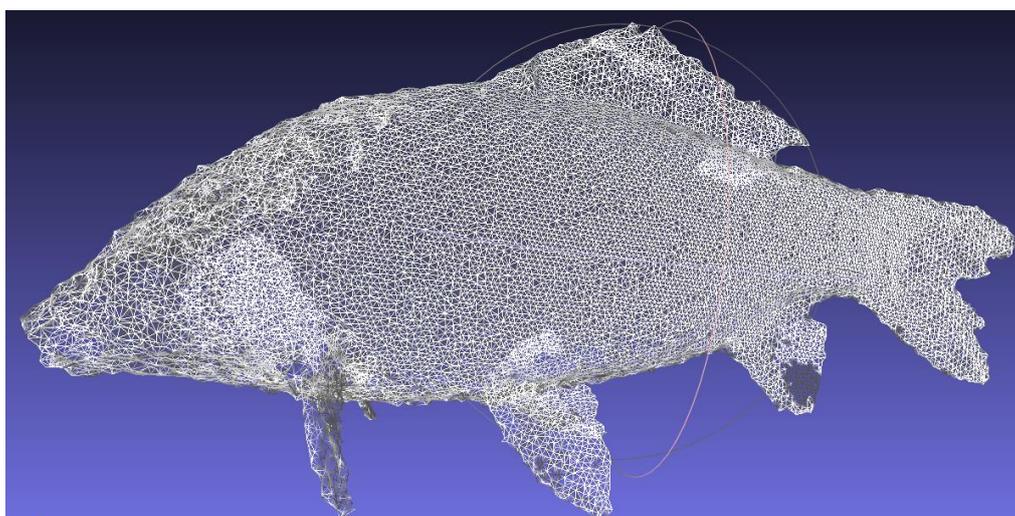


Fig. 53 3-D fish model of triangle meshes (Japanese crucian carp)

5.4. Coordinate settings

Firstly, the center of fish model was set to the origin of coordinate system. I compared the x, y, z coordinates of the all vertices, and get the x_{min} , y_{min} , z_{min} , x_{max} , y_{max} , z_{max} . So the coordinate of the center of fish model x_{center} , y_{center} , z_{center} are as Formula. 7.

$$\begin{aligned}x_{center} &= (x_{min} + x_{max})/2 \\y_{center} &= (y_{min} + y_{max})/2 \\z_{center} &= (z_{min} + z_{max})/2\end{aligned}\quad \textbf{Formula. 7}$$

I would take new coordinates x'_i , y'_i , z'_i of fish model's all vertices instead of old coordinates x_i , y_i , z_i . Therefore, through the following Formula. 8, the fish model' center would be at the origin.

$$\begin{aligned}x'_i &= x_i - x_{center} \\y'_i &= y_i - y_{center} \\z'_i &= z_i - z_{center}\end{aligned}\quad \textbf{Formula. 8}$$

And then, the fish model need to be parallel with x axis (take x axis for example) for convenient operation and calculation. According to the following Formula. 9 [12], calculation should be from right to left, X, Y, Z stand for coordinates after transformation. α , β , γ stand for the angle that the model is rotated by x, y, z axis. I adjusted the value of α , β , γ , and fix their values when fish was parallel with x axis.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{pmatrix} * \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Formula. 9

As Fig. 54 shows, P stands for the acoustic video camera. It will be moved around the fish model by adjusting the value of φ and θ . The coordinates of P are as Formula. 10 shows, r stands for the distance to origin. By changing the value of φ and θ , the position of acoustic video camera can be controlled.

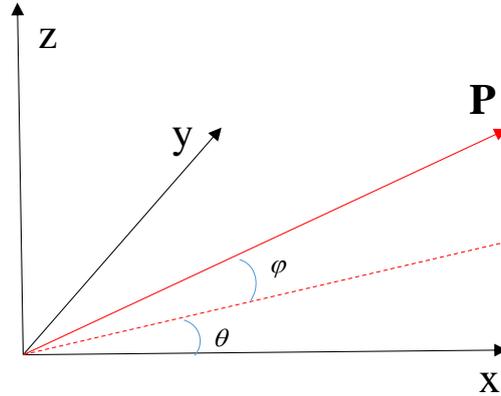


Fig. 54 The position of acoustic video camera in space

$$\begin{aligned}
 x_p &= r * \cos \varphi * \cos \theta \\
 y_p &= r * \cos \varphi * \sin \theta \\
 z_p &= r * \sin \varphi
 \end{aligned}$$

Formula. 10

In the next step, as Fig. 55 shows, **P** stands for the vector from mesh's barycenter to the acoustic video camera. **N** is the normal vector of the mesh. So the angle between the two vectors δ can be calculated as Formula. 11. Assume the coordinates of **P** is (P_x, P_y, P_z) , coordinate of **N** is (N_x, N_y, N_z) , then

$$\delta = \cos^{-1} \frac{P_x * N_x + P_y * N_y + P_z * N_z}{\sqrt{(P_x^2 + P_y^2 + P_z^2)} * \sqrt{(N_x^2 + N_y^2 + N_z^2)}}$$

Formula. 11

When $\delta > 90^\circ$, the backscattering strength at that mesh will not be calculated, because it means the fish surface opposite to acoustic video camera, where the sonar beam cannot reach.

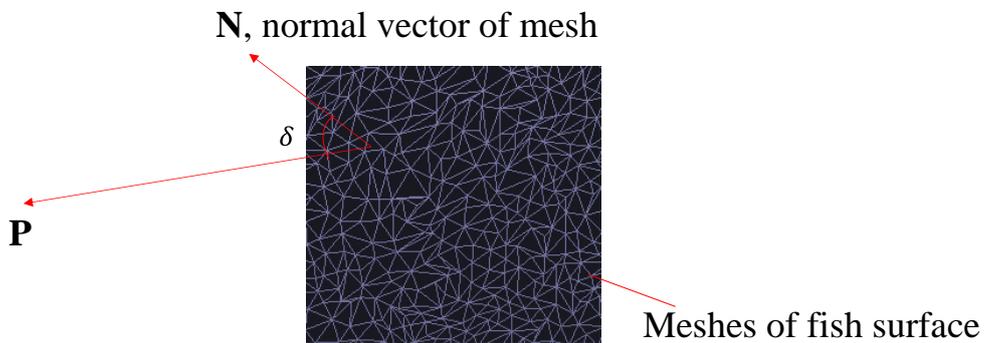


Fig. 55 Normal vector **N** and the vector **P** from mesh's barycenter to the acoustic video camera

5.5. Backscattering strength

In physics, backscatter (or backscattering) is the reflection of waves, particles, or signals back to the direction from which they came. It is a diffuse reflection due to scattering, as opposed to specular reflection like a mirror.

In the water, when sound wave meets with objects of different acoustic impedance, such as fish in this study, backscatter will happen as Fig. 56 shows. As we know, the surface of fish skin is rough and this irregular property in skin is less than the wave length. If the factor of fish body's acoustic impedance is not considered, in this situation according to Lambert rule, strength of backscattering is as Formula. 12 shows. [13]

$$S_B = 10 \log \mu + 10 \log(\sin \varepsilon)^2 \quad \text{Formula. 12}$$

Here μ is ratio constant, assuming the sound wave does not penetrate fish body, $\mu = 1/\pi$.

As Fig. 55 shows, $\delta = \pi/2 - \varepsilon$. Therefore, in this study, backscattering strength is calculated as Formula. 13,

$$S_B = 10 \log \mu + 10 \log(\cos \delta)^2 \quad \text{Formula. 13}$$

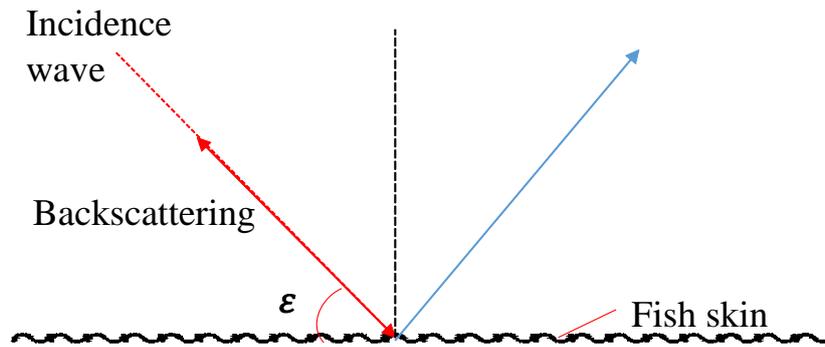


Fig. 56 Backscattering when sound wave meets with fish surface

5.6. Comparison between simulated images and sonar images

By the calculation method above, the backscattering strength of 3-D fish model surface's every mesh was calculated according to the angle δ , and the backscattering strength was displayed as brightness of the mesh. The higher backscattering strength is, the brighter the mesh becomes. Through adjusting the position of acoustic video camera by changing the value of φ and θ (Fig.54), the backscattering strength of fish model's surface will change.

And I compared the simulated images with the sonar images at same space angle. In the observation of fish sample by ARIS in the second part of experiment II, as Fig. 57 shows, because of bubble inside fish, head was always higher than caudal fin. Therefore, when θ was within 0° to 90° , φ was nearly -15° ; when θ was within 90° to 180° , φ was nearly 15° .

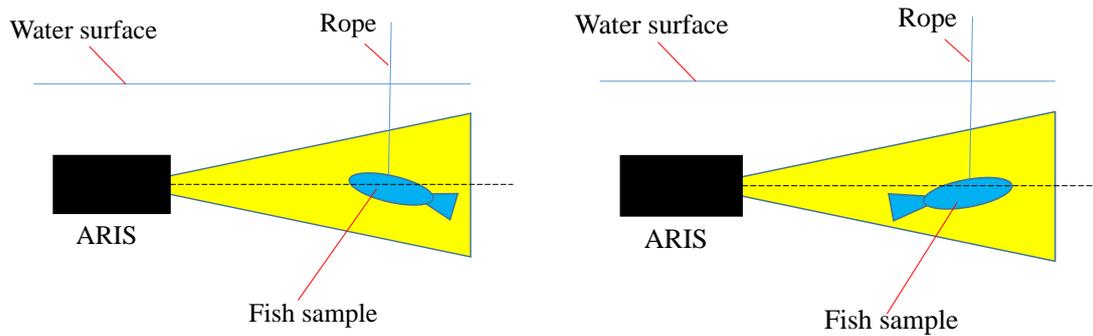
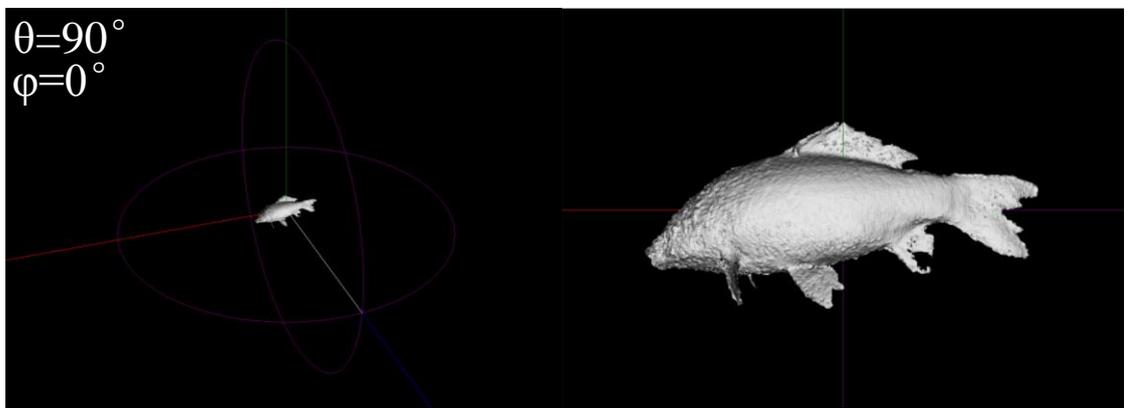
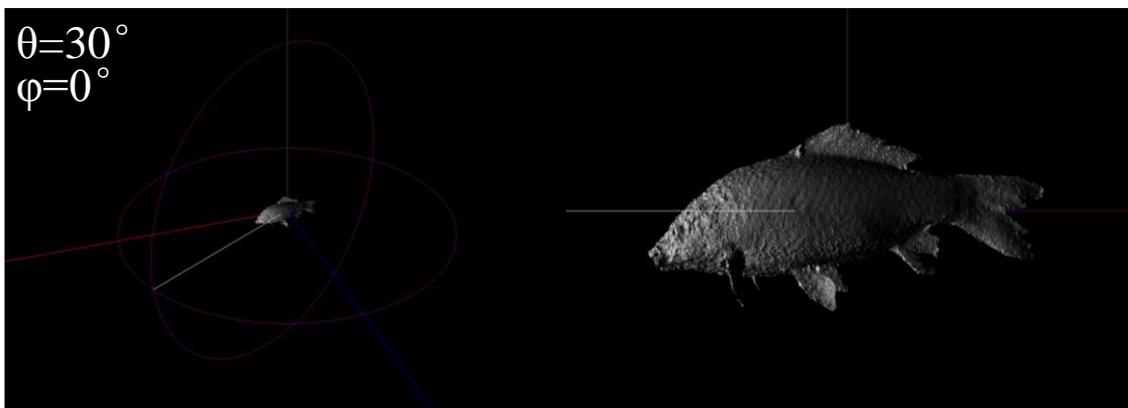
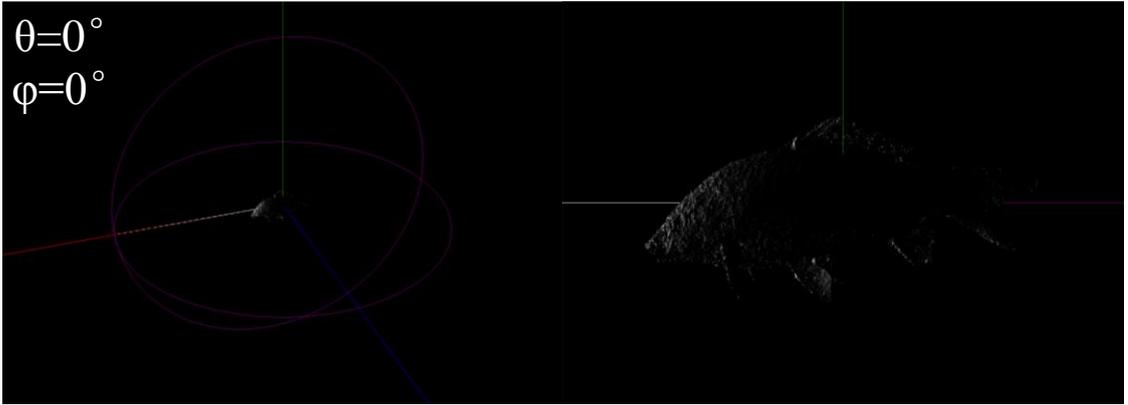


Fig. 57 Observation of frozen fish sample by ARIS

Therefore firstly the simulated images are shown as Fig. 58, Fig. 59 and Fig. 60, changing with θ when $\varphi = -15^\circ, 0^\circ, 15^\circ$. And then, simulated images are compared with sonar images with the same space angle θ and φ as Fig. 61 shows. (Here θ and φ can refer to Fig. 54)



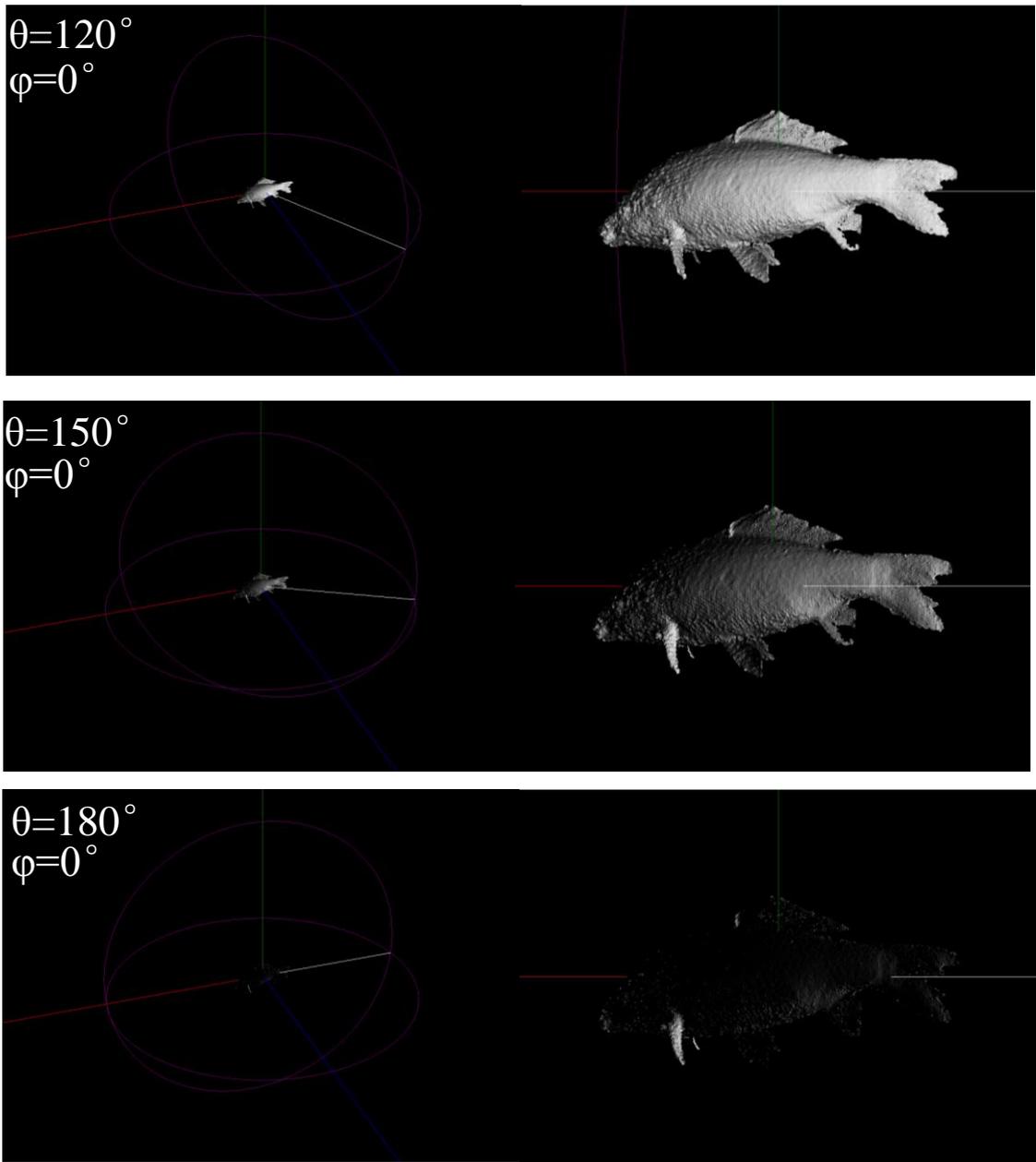
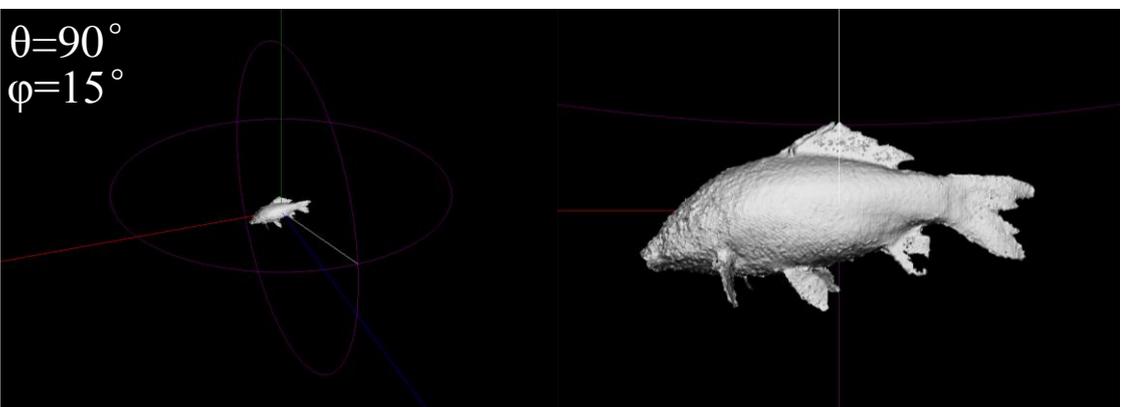
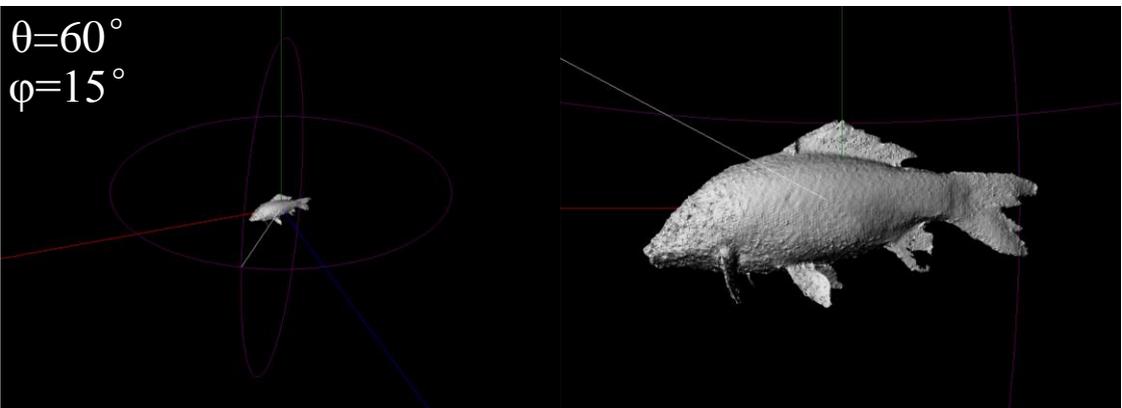
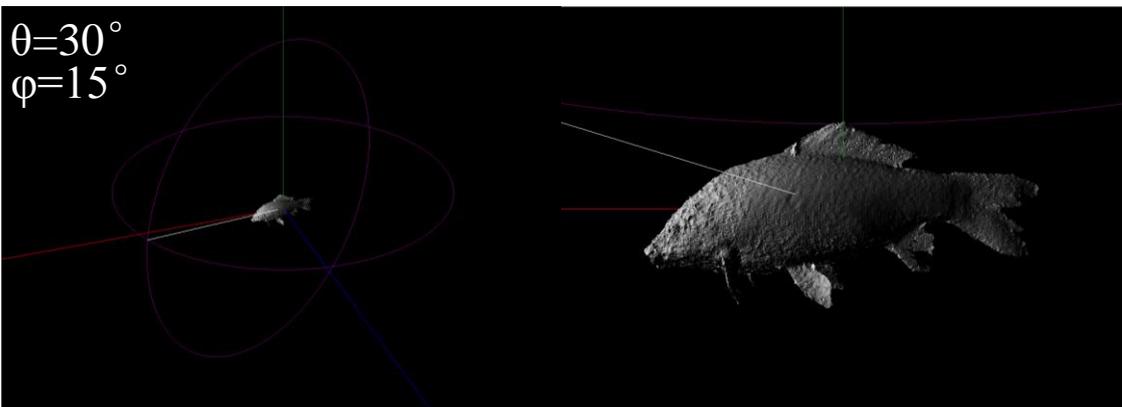
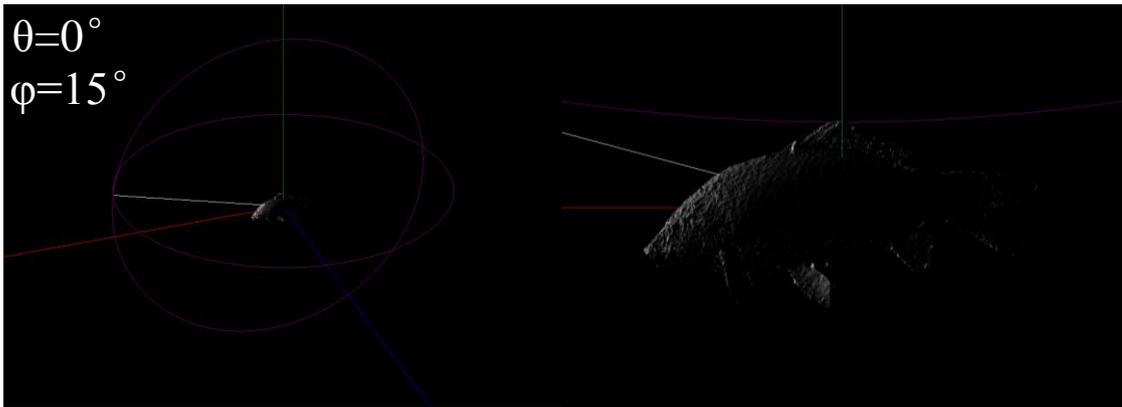


Fig. 58 Change of backscattering strength of fish surface with the change of θ when $\varphi = 0^\circ$.



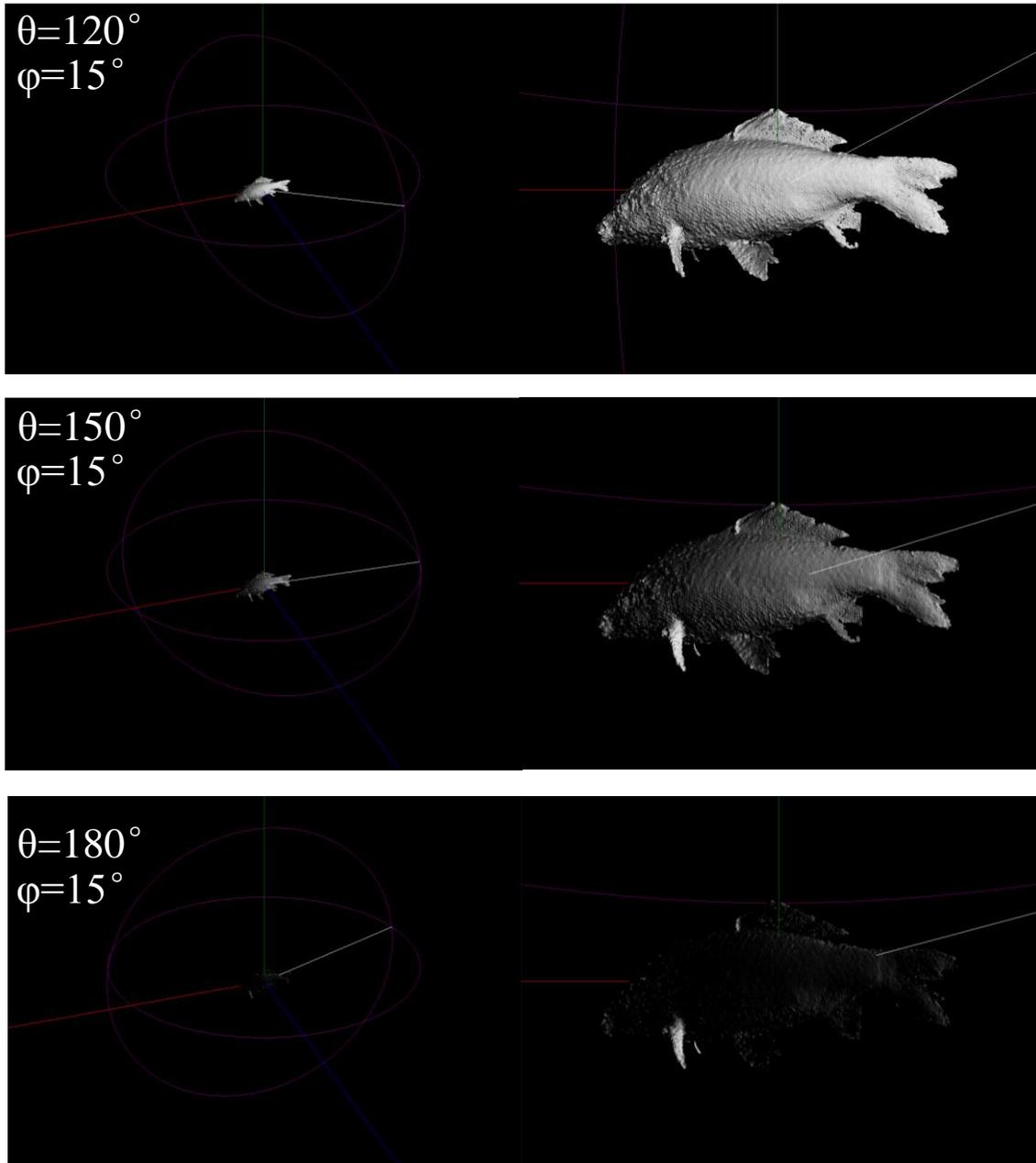
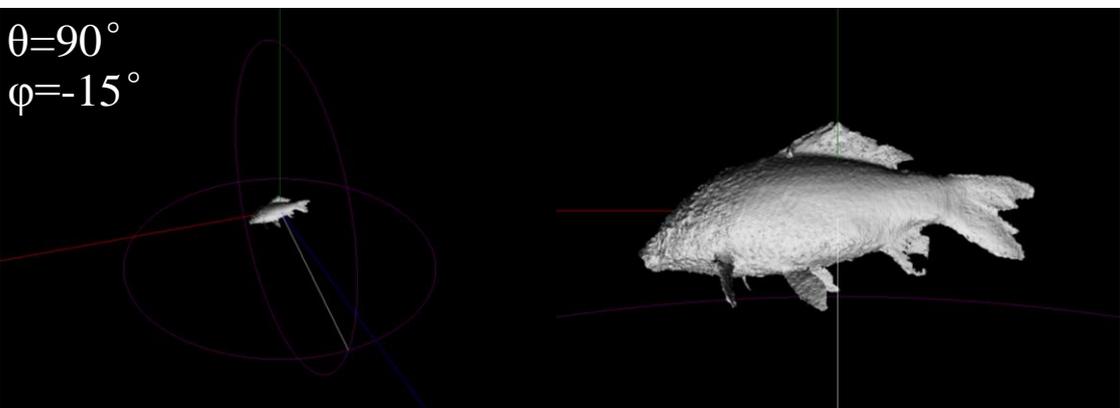
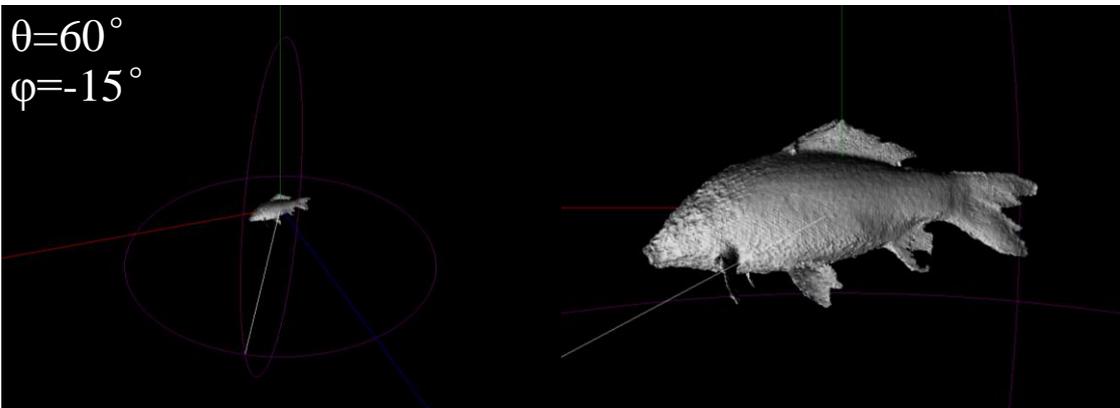
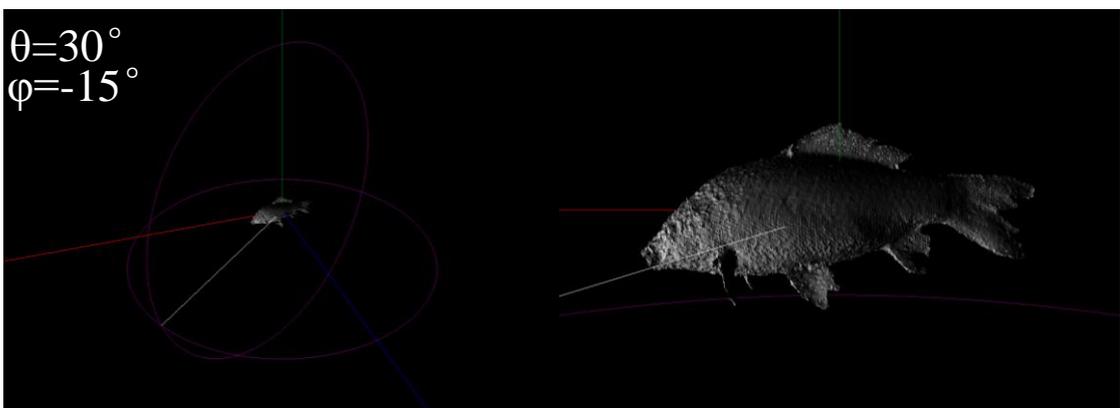
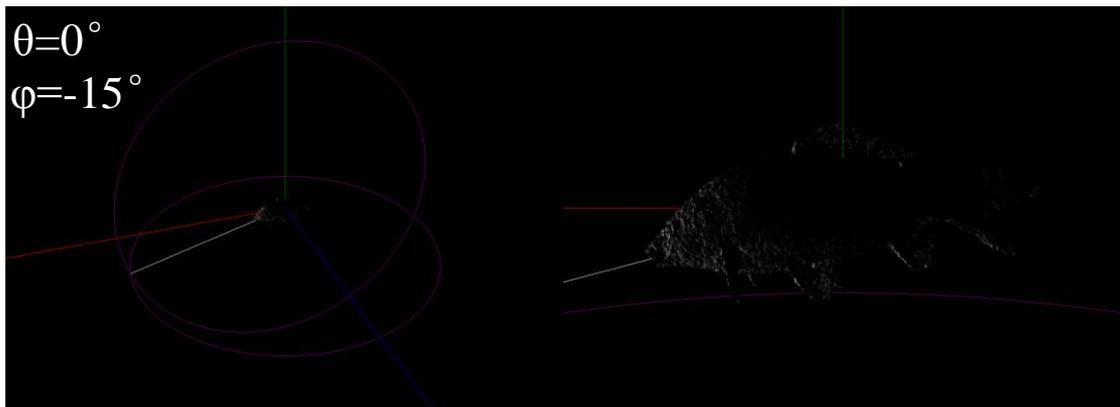


Fig. 59 Change of backscattering strength of fish surface with the change of θ when $\varphi = 15^\circ$.



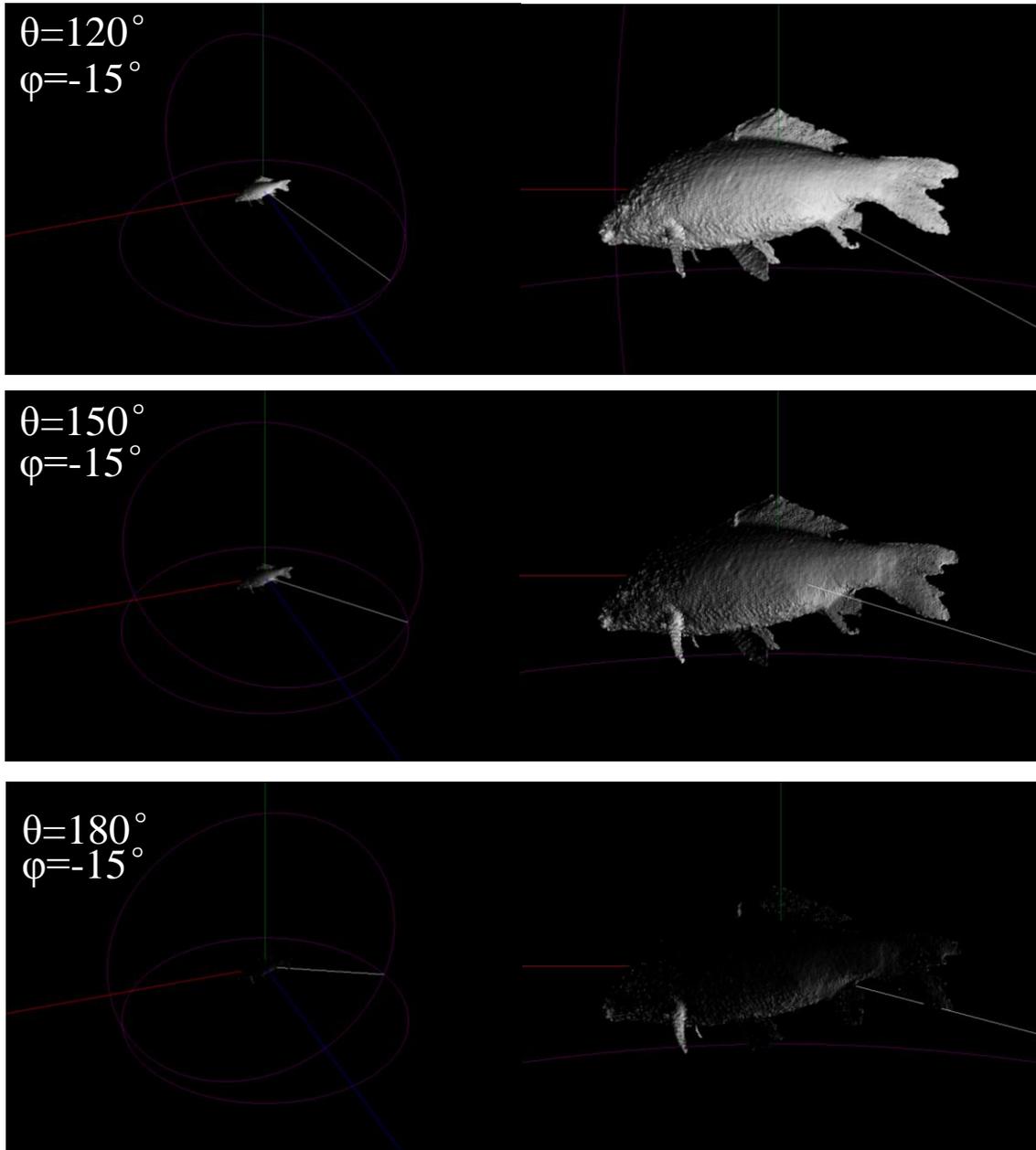
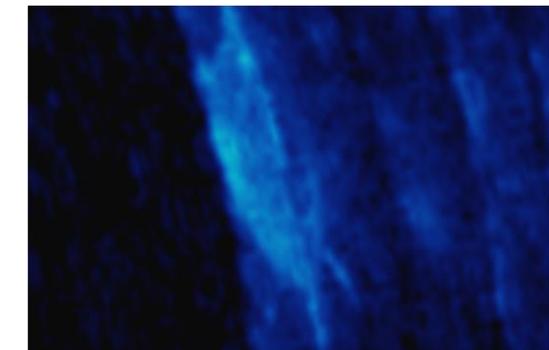
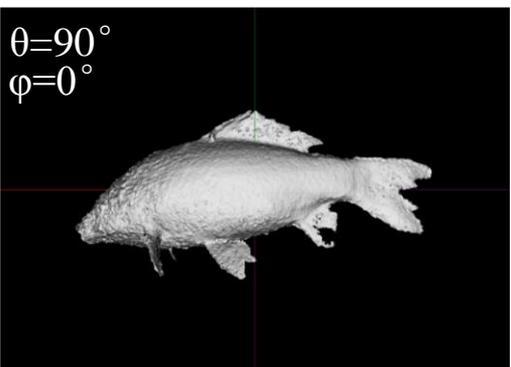
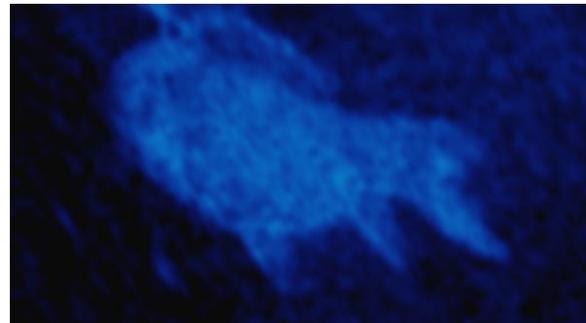
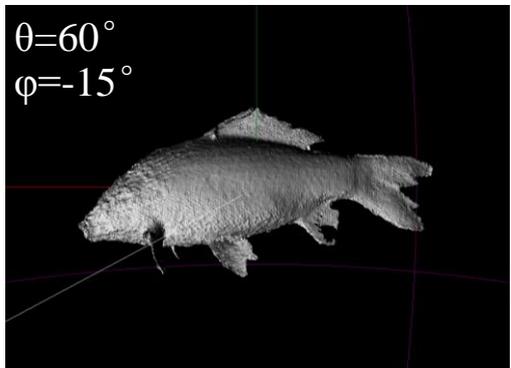
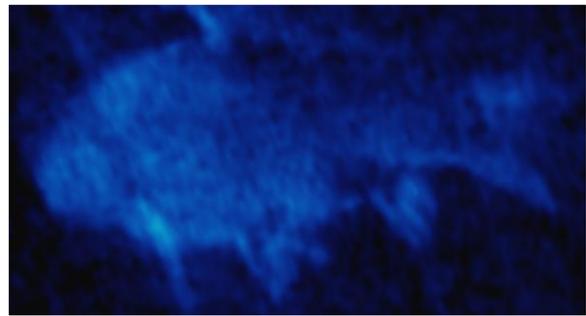
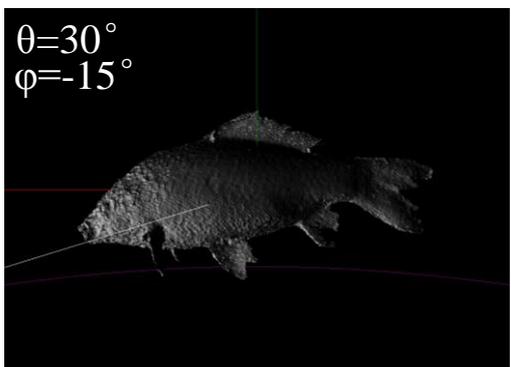
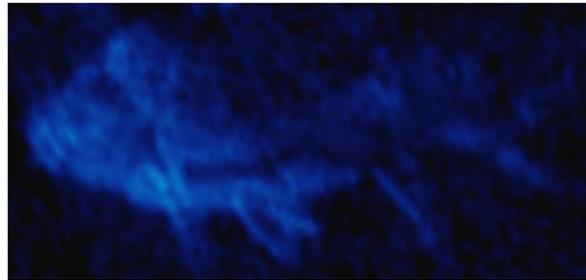
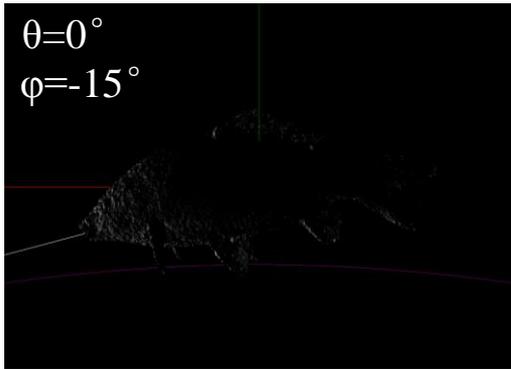


Fig. 60 Change of backscattering strength of fish surface with the change of θ when $\varphi = -15^\circ$.



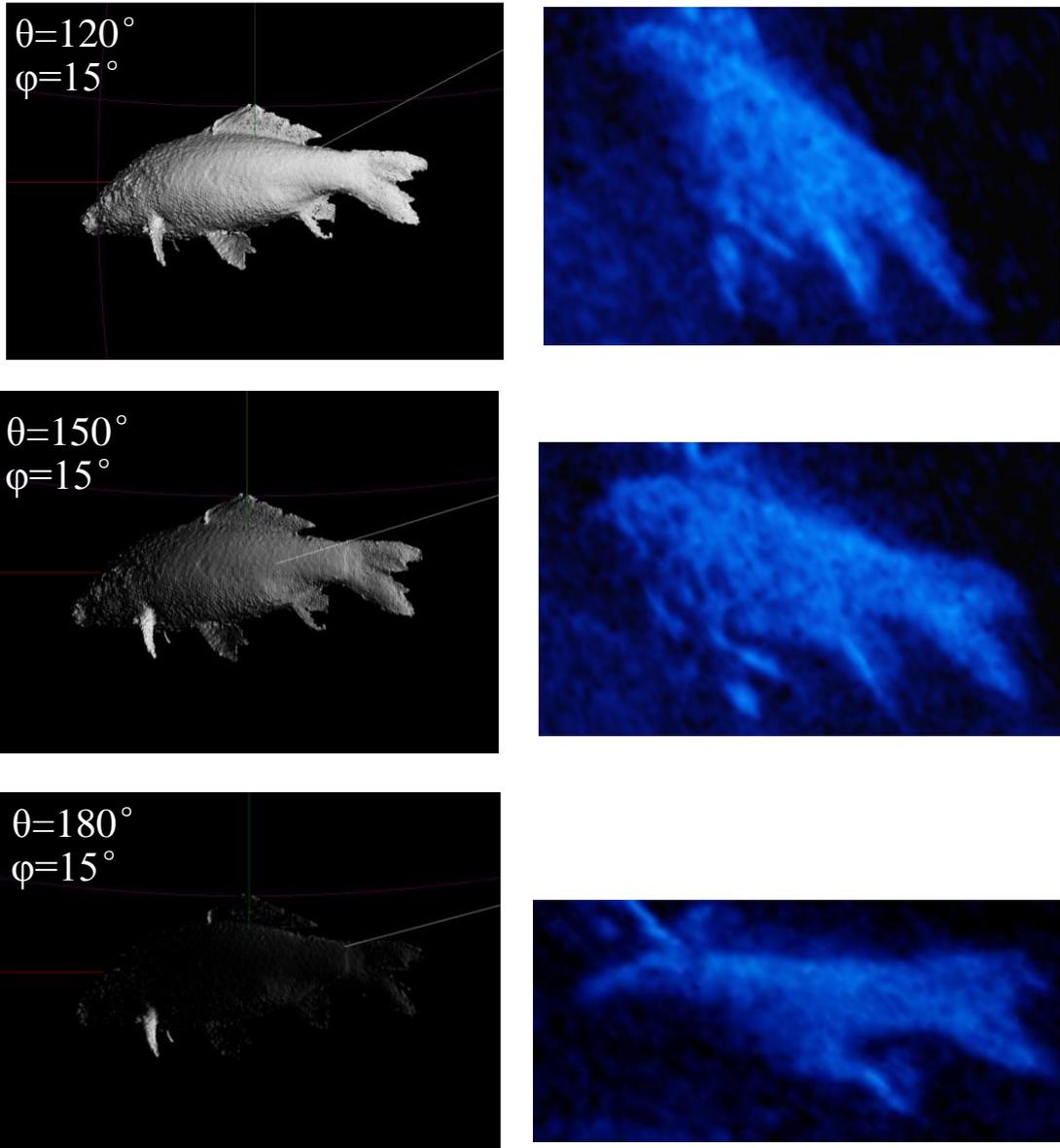


Fig. 61 Simulated images compare with sonar images with the same space angle θ and φ

5.7. Discussion of result and potential problems

Firstly, Fig. 58, 59 and 60 show the backscattering strength of fish model's surface by the brightness of every triangle mesh. From the simulated images, we can clearly see the change of brightness when acoustic video camera's position varies. When $\varphi = -15^\circ$, the bottom part of fish is brighter, which means the backscattering strength from bottom part is higher than top part of fish; When $\varphi = 15^\circ$, the situation is opposite.

Simulation supplied a possibility to compare simulated images with the real sonar images. Through comparison, it helps us to know the theory how sonar images form. As Fig.61 shows, when $\varphi = -15^\circ$, $\theta = 0^\circ$, only the bottom of fish's head and belly fin shows in the sonar image, which is consistent to the simulated image. On occasion of $\varphi = -15^\circ$, $\theta = 30^\circ$, fish shows whole body clearly in both simulated image and sonar image. When $\theta = 60^\circ$ and 90° , although the backscattering strength is shown in the simulated image, two images do not match very well because simulation did not consider the narrow beam of ARIS and the calculation is towards all meshes of fish model. When $\varphi = 15^\circ$; $\theta = 120^\circ$, 150° and 180° , the head of fish hardly can be seen from the images. As we can see, simulated images and corresponding sonar images matched well.

As discussed above, simulated images match well with sonar images in most occasions. However, the simulation does not take the factors of ARIS's narrow beam and fish's acoustic impedance into consideration, and it is still difficult to compared 2-D sonar images with 3-D fish model, quantitative results has not been obtained.

CHAPTER 6: CONCLUSION AND FUTURE WORK

This study researched on fish classification by using high-resolution video camera ARIS and local invariant feature descriptor. Firstly, in order to obtain high-quality acoustic images, new observation method was proposed, ARIS was rotated by 90° and mounted with a 3° concentrator lens. And then through the new observation method, 2 field experiments were carried out near Izunuma Lake.

In experiment I, SIFT algorithm was applied into acoustic image processing for the first time. High-quality images were obtained by high-resolution acoustic video camera with proposed observation method. And it is known to us that SIFT based image processing method is more robust to angle and scale changes than NCC method, and it is easier for SIFT method than NCC to distinguish fish of same species from others. However, the data amount was small in experiment I and the effect of individual differences of same species on fish classification has not made clear. In addition, the matching rate imperfect images was unstable.

In experiment II, corresponding connection between sonar images of fish and its angle with ARIS's center beam were observed, which supports a good hint to the fish classification by acoustic video camera in the future. From the result of matching rate, it is found that the sonar images of fish will be affected by the angle between fish and acoustic video camera's center beam and the extent how much fish's body is inside sonar's beam. However, satisfying results can't be obtained with SIFT method by matching with only one template because it is usual for ARIS to obtain imperfect images.

And then simulation was done in order to produce simulated images at different angles as template images which is convenient, and to compared with the sonar images. As a result, simulated images match well with sonar images in most occasions. However, the simulation does not take the factors of ARIS's narrow beam and fish's acoustic impedance into consideration, and it is still difficult to compared 2-D sonar images with 3-D fish model, quantitative results has not been obtained.

In the future, I plan to develop and revise simulation for now. The simulation should be able to generate simulated 2-D sonar images considering the fundamental theory with which acoustic video camera forms images. The narrow beam of ARIS, acoustic impedance of fish will be take into consideration, too. Thousands of template images for each species at any angle will be prepared for the template matching, which is expected to connect with the technology of deep learning. And I think multiple template images will be necessary for fish classification in the future.

CHAPTER 7: REFERENCE

- [1] Franc_ois Martignac, Aur_elie Daroux, Jean-Luc Bagliniere, *et. al.*, “The use of acoustic cameras in shallow waters: new hydroacoustic tools for monitoring migratory fish population. A review of DIDSON technology”, *FISH and FISHERIES*, (2015), 16, 486–510
- [2] Mohammadmehdi Saberioon, Asa Gholizadeh, Petr Cisar, Aliaksandr Pautsina, *et. al.*, “Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues”, *Reviews in Aquaculture* (2016) 0, 1–19
- [3] Homepage of sound metrics
<http://www.soundmetrics.com/Products/DIDSON-Sonars/DIDSON-300m>
- [4] Chunhui XU, Mizuno K, Asada A, Abukawa K, *et. al.*, 3D Views generation and Species Classification Methods of Aquatic Plants Using Acoustic Images. *J Marine Acoust Soc Jpn* (2013)40:14-26
- [5] Mizuno K, Asada A, Three dimensional mapping of aquatic plants at shallow lakes using 1.8 MHz high-resolution acoustic imaging sonar and image processing technology. *Proceedings of 2014 IEEE Ultrasonics Symposium*, (2014)1384-1387
- [6] Langkau, M.C., Balk. H., Schmidt, M.B., *et. al.*, Can acoustic shadows identify fish species? A novel application of imaging sonar data. *Fisheries Management and Ecology* (2012)19, 313–322.
- [7] Mueller, A.M., Burwen, D.L., Boswell, K.M., *et. al.*, Tail-beat patterns in dual-frequency identification Sonar Echograms and their potential use for species identification and bioenergetics studies. *Transactions of the American Fisheries Society* (2010)139, 900–910.
- [8] ARIS-3000-Product-Specs-English
<http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000/ARIS-3000-Product-Specs-English>
- [9] Richard WD, Arthur RM, Real-time ultrasonic scan conversion via linear interpolation of oversampled vectors. *Ultrasound Imaging*. (1994) 16:109-123
- [10] K. Mizuno, A. Asada, J. Ashizawa, *et. al.*, Application of a high-resolution acoustic video camera to fish classification: an experimental study. *Proceedings of Underwater Technology 2015, IEEE international*, DOI information: 0.1109/UT. (2015).7108250
- [11] D. G. Lowe, “Object recognition from local scale-invariant features”, *Computer vision, The Proceedings of the Seventh IEEE International Conference vol. 2*, (1999),1150-1157
- [12] http://fnorio.com/0098spherical_trigonometry1/spherical_trigonometry1.html
- [13] Urick, Robert J., <Principles of underwater sound/ 3d edition>, ISBN 0-07-066087-5, pp 271-280