Robust Sensing against Bubble Noises in Aquatic Environments with a Stereo Vision System

Atsushi Yamashita, Susumu Kato and Toru Kaneko
Department of Mechanical Engineering, Shizuoka University
3–5–1 Johoku, Hamamatsu-shi, Shizuoka 432–8561, Japan
Email: yamashita@ieee.org

***Abstract*— In this paper, we propose robust sensing method against bubble noises in aquatic environments with a stereo vision system. Usually, three-dimensional (3-D) measurement by robot vision techniques is executed under the assumptions that cameras and objects are in aerial environments. However, an image distortion occurs when vision sensors measure objects in liquid. It is caused by the refraction of the light on the boundary between the air and the liquid, and the distorted image brings errors in a triangulation for the range measurement. Additionally, it is often the case that there exist air bubbles in the field of view when we observe aquatic environments. Therefore, it becomes difficult to acquire clear images because of these view-disturbing noises. As to the former problem, accurate 3-D coordinates of objects' surfaces in liquid are measured by taking for calculating the refraction effect. As to the latter problem, bubble noises are eliminated from a moving image to divide objects in the image into still backgrounds, moving objects, and bubble noise by an image processing technique. Experimental results showed the effectiveness of our proposed method.**

(a) Bubble noise. (b) Refraction effects.

Fig. 1. Examples of aquatic images.

I. Introduction

In this paper, we propose robust sensing method against bubble noises in aquatic environments with a stereo vision system.

In recent years, demands for underwater tasks, such as digging of ocean bottom resources, exploration of aquatic environments, rescues, and salvages, have increased. Therefore, underwater robots or underwater sensing systems that work instead of human become important, and technologies for observing underwater situations correctly and robustly from cameras of these systems are needed. However, it is very difficult to observe underwater environments with cameras, because of the following two big problems.

1) Bubble noises (Fig. 1(a))
2) Refraction effects (Fig. 1(b))

The former problem is about suspended matters, such as bubble noises, small fishes, and small creatures. They may disturb camera's field of view (Fig. 1(a)).

The latter problem is about the refraction effects of light. If cameras and objects are in the different condition where the refraction index differs from each other, several problems occur and a precise measurement cannot be achieved. For example, Fig. 1(b) shows an image of a single rectangular object when water is filled to the middle. In this case, the size and the shape of the object look different between above and below the water surface. These problems occur not only when a vision sensor is set outside the liquid but also when it is set inside, because in the latter case we should usually place a protecting glass plate in front of viewing lens. Therefore, it becomes difficult to measure precise positions and shapes of objects when water exists because of the image distortion by the refraction of the light.

As to the former problem about view-disturbing noises, there are several methods that can remove noises from images or can detect moving objects other than moving noises, such as snowfall noises, waterdrop noises, and so on [1]–[7]. These techniques remove moving objects or noises by taking the difference between the initial background scene and a current scene, or taking the difference between temporarily adjacent two frames. These methods are robust against the change of background [2] or the change of the lighting condition [3], and can also remove snowfall noises [4] or waterdrop noises [5]–[7]. Image interpolation method [8] that adopts computer graphics technique can remove noises by using only one image. However, these methods are not suitable for removing bubble noises in aquatic environments. This is because they have uncertain outlines, or are not automatic method in which human operator must indicate the positions of noises.

As to the latter problem about refraction effects, three-dimensional (3-D) measurement methods in aquatic environments are also proposed [9]–[12]. However, techniques that do not consider the influence of the refraction effects [9]–[11] may have the problems of accuracy. Accurate 3-D measurement methods of objects in liquid [13]–[17] with a laser range finder by considering the refraction effects are also proposed. However, it is difficult to measure moving objects with a laser range finder. A stereo camera system is suitable for measuring moving objects, though the methods by using a stereo camera system [12] have the problem that the

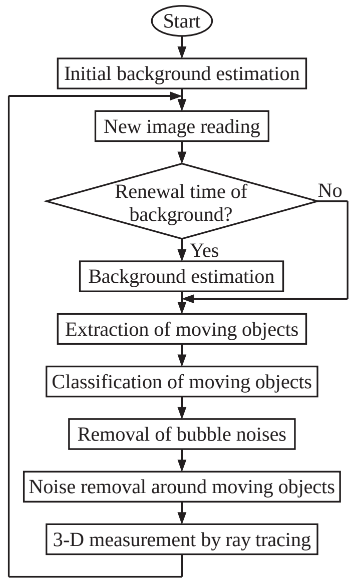


Fig. 2. Outline of the proposed method.

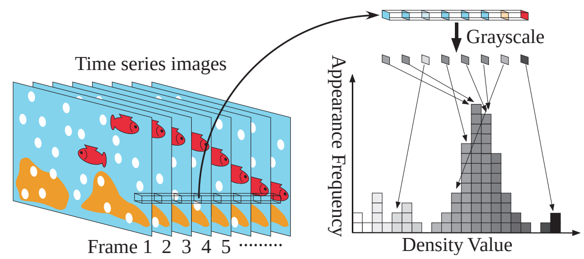
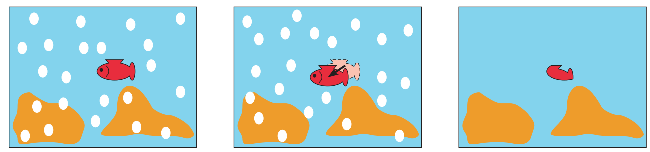


Fig. 3. Background estimation and generation.



(a) n-th frame. (b) $(n+m)$-th frame. (c) Background.

Fig. 4. Quasi-still moving object.

corresponding points are difficult to detect when the texture of the object's surface is simple in particular when there is the refraction on the boundary between the air and the liquid. The method by the use of motion stereo images obtained with a moving camera [18] also has the problem that the relationship between the camera and the object is difficult to estimate because the camera moves. The surface shape reconstruction method of objects by using an optical flow [19] is not suitable for the accurate measurement, too.

In this paper, we aim at solving above two problems.

The first aim is to reduce the influence of bubble noises in aquatic images by using image processing techniques. In this paper, we use a clustering method that can distinguish moving objects from bubble noises with characteristics such as object's color, its size, and so on. After that, only bubble noises are removed.

The second aim is to measure moving objects in aquatic environments accurately. Our method uses a stereo vision system. We construct accurate 3-D measurement method by considering the refraction effects in aquatic environments. Robust detection of corresponding points of stereo image pairs can be also realized under epipolar constraints by considering the refraction effects.

Finally, we combine these two techniques to realize robust observation of moving objects against bubble noises in aquatic environments.

II. Extraction of Moving Objects

Overview of bubble noise removal method is shown in Fig. 2. In our method, bubble noises are distinguished from extracted moving objects after clustering them in acquired images. Finally, noise-free images can be generated by removing bubble noises from images.

The method extracts moving objects from a color image sequence by the following procedure.

1) Estimate a background image periodically.
2) Extract moving objects by the background subtraction.
3) Classify moving objects and distinguish bubble noises from other moving objects by referring to their colors, sizes and so on.

A. Background Estimation

After acquiring time-series stereo images, moving objects are extracted from left and right images respectively by using background subtraction method.

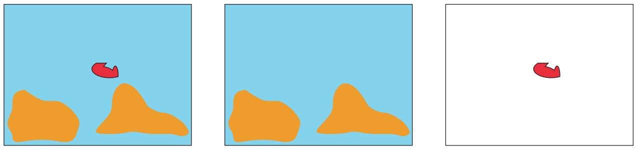
Background images must be prepared in advance, or must be generated in the process of measurement. Our method generates background images by using the histogram of time-series pixel values (brightness of each pixel).

The process of background generation is as follows. At first, image data is accumulated during a certain fixed period as color data which consists of R, G, and B color components. Then, these images are converted to grayscale images, and histograms whose horizontal axes are pixel values and vertical axes are frequency of appearance of pixel values are generated (Fig. 3). Background pixel value of each pixel is regard as the pixel value with the highest frequency of appearance in a certain period. After that, grayscale data is converted back to RGB values by calculating the average RGB color values in the times of the highest frequency.

In order to respond to scene change owing to time progression, background scenes are generated periodically.

However, quasi-still moving objects that moves very slowly and whose positions hardly change during a certain fixed period of background generation may exist in the generated background images (Fig. 4). In this case, quasi-still moving objects are regarded as background scenery and cannot detected as moving objects.

It is assumed that the generated first background contains no moving or quasi-still moving objects. If the second background contains quasi-still moving objects, they can be detected by calculating the difference between the generated present



(a) Present one. (b) Previous one. (c) Interpolation area.

Fig. 5. Background correction.

background (second background) and the previous one (first background). Therefore, the regions where quasi-still moving objects exist in the present background can be interpolated by using the previous background (Fig. 5).

In this way, we can detect and measure quasi-still moving objects.

B. Extraction of Moving Objects

Moving objects in the image sequence are detected by background subtraction. Background subtraction is a method for extracting the moving objects in the image by taking the difference between the background image generated beforehand and the present image. We use the background image shown in Section II-A for background subtraction.

C. Classification of Moving Objects

In order to distinguish moving objects from floating air bubbles, ISODATA (Iterative Self Organizing Data Analysis techniques A) clustering method [20] is employed.

This technique is based on randomly choosing initial cluster centers, or means. These initial cluster centers are updated in such a way that after a number of cycles they represent the clusters in the data as much as possible. The ISODATA algorithm circumvents the problem by removing redundant clusters. Whenever a cluster centre is not assigned enough samples, it may be removed. In this way one is left with a more or less optimal number of clusters. We adopt the variety of ISODATA method proposed in [21].

Our ISODATA method classifies objects by the following procedure (Fig. 6).

1) Set parameters (number of final clusters, divergence condition of splitting and merging and so on).
2) Assign primary cluster centers initially.
3) Assign samples to the nearest cluster.
4) Recalculate the cluster centers and reassign samples (Repeat until reassignment settles).
5) Discard samples in clusters with few members.
6) Calculate the distance between clusters and the variance within each cluster.
7) Split the cluster with the largest variance or merge clusters with smallest distance between them according to divergence condition.

The IOSDATA method can flexibly classify objects by learning the optimal number of classes according to merging and splitting of the classes. Various features such as an object size, color component values and a vector of motion can be used as data for clustering. In this paper, we use the size

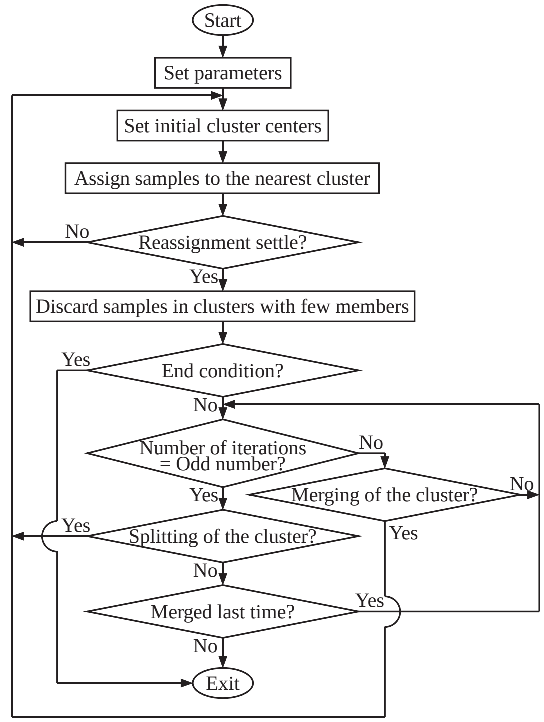


Fig. 6. Outline of ISODATA clustering method.

of each moving object and RGB color component values as feature values.

D. Removal of Bubble Noises

In this paper, we assume that a view of the camera is interrupted by air bubble noises and that the number of air bubble noises is the largest among all moving objects. Therefore, the class with the largest number of elements becomes the class of the air bubble noises and the others become the class of the measurement objects[1]. Finally, we use the latter objects for the following procedure and remove the former objects by replacing them with pixel values of the present background image.

E. Noise Removal around Moving Objects

It becomes difficult to measure moving objects accurately when a part of air bubbles remains in surroundings of them. We solve this problem by matching the position of the moving object between the previous frame and present frame, and calculating the logical product (Fig. 7).

III. 3-D Measurement by Ray Tracing

First of all, it is necessary to search for corresponding points from a right and left image to measure the object by using the stereo vision system. The corresponding points can be detected by using the epipolar constraints (Section III-B).

After detecting corresponding points, an accurate 3-D measurement can be executed by considering the refraction effects

[1]It will be also possible to distinguish bubble noises from other moving objects by using the difference of color, shape, and trajectory.

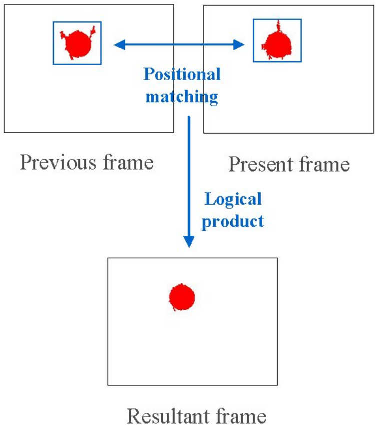


Fig. 7. Shape correction of moving object.

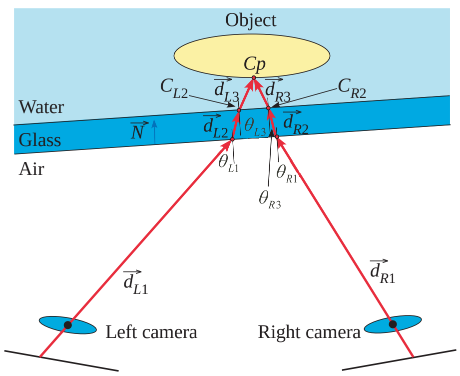


Fig. 8. Ray tracing from each camera.

of light in aquatic environments. At first, we explain about the principle of 3-D measurement that is based on the ray tracing technique.

A. Ray Tracing from Each Camera

Refractive angles at the boundary surfaces among air, glass and water can be determined by using Snell's law, if location and orientation of the glass surface are known (Fig. 8).

We assume the refractive index of air and the glass to be n_1 and n_2, respectively, and the incidence angle from air to the glass to be θ_1. A unit ray vector $\vec{d_2} = (\alpha_2, \beta_2, \gamma_2)^T$ (T denotes transposition) traveling in the glass is shown by Equation (1).

$$\begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} + \left(\sqrt{1 - \frac{n_1^2}{n_2^2} \sin^2 \theta_1} - \frac{n_1}{n_2} \cos \theta_1 \right) \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}, \quad (1)$$

where $\vec{d_1} = (\alpha_1, \beta_1, \gamma_1)^T$ is the unit ray vector of the camera in air and $\vec{N} = (\lambda, \mu, \nu)^T$ is a normal vector of the glass plane.

A unit ray vector $\vec{d_3} = (\alpha_3, \beta_3, \gamma_3)^T$ traveling in water is

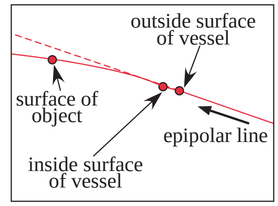


Fig. 9. Epipolar constraints in aquatic environments.

shown by Equation (2).

$$\begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} + \left(\sqrt{1 - \frac{n_2^2}{n_3^2} \sin^2 \theta_3} - \frac{n_2}{n_3} \cos \theta_3 \right) \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}, \quad (2)$$

where n_3 is the refractive index of water and θ_3 is the angle of incidence from the glass to water.

An arbitrary point $\vec{C_p} = (x_p, y_p, z_p)^T$ on the ray vector is shown by Equation (3).

$$\begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} = c \begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}, \quad (3)$$

where $\vec{C_2} = (x_2, y_2, z_2)^T$ is the point on the water tank plane and c is a constant.

Two rays are calculated by ray tracing from the left and the right cameras, and the intersection of the two rays gives the 3-D coordinates of the target point in water.

B. Detection of Corresponding Points

We use noise-free stereo image pairs that is explained in Section II-E for detecting corresponding points.

The relationship between corresponding points of the left and the right images is formulated with epipolar constraints, and the corresponding points exist on the epipolar lines [22]. In aerial environments, the epiplor line is straight after correcting the image distortion derived from the effect of the lens distortion [23]. However, the epipolar lines are not straight in aquatic environments because of the refraction of light (Fig. 9). Therefore, we calculate the epipolar lines with the ray tracing technique in the same way of Section III-A [14]–[16].

Corresponding points on epipolar lines are searched for with template matching by using the normalized cross correlation (NCC) method.

After corresponding point and disparity of each pixel are acquired, the 3-D position of each corresponding point can be measured with triangulation.

IV. Experiment

We constructed an underwater environment by using a water tank and an air shower (device to generate air bubbles) (Fig. 10). It is an equivalent optical system to sinking the waterproof camera in underwater. We used two digital video cameras for

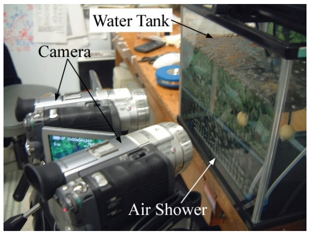
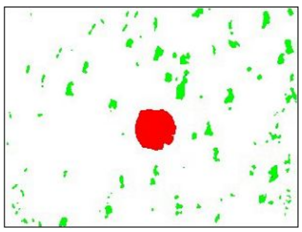


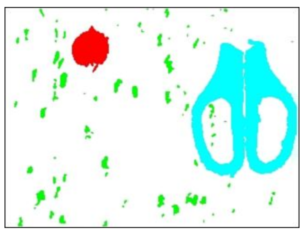
Fig. 10. Experimental setup.

(a) Original image 1.

(b) Original image 2.



(c) Moving object 1.



(d) Moving object 2.

(e) Noise-free image 1.

(f) Noise-free image 2.

Fig. 11. Experimental result 1.

taking 30fps moving pictures whose sizes are 720×480pixels. Background images were generated from 30 frames, and the update cycle of the background was 3fps.

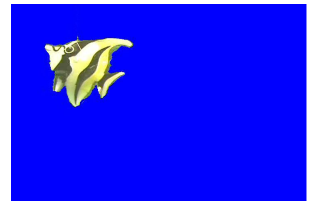
A. Extraction of Moving Objects

We used a wooden ball and scissors for experiment. Figure 11 shows an example of original image, classified moving objects, and noise-free images, respectively. Green, red, blue colors mean bubble noises, wooden ball, scissors, respectively in Fig. 11(c) and (d). These results show that each object is classified as a different object and measurement objects are well extracted.

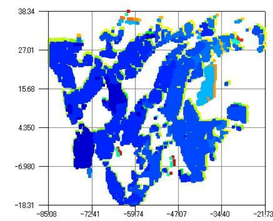
B. 3-D Measurement Result

We measured 3-D position and shape of a fish-like object moving in the water tank (Fig. 12). The result of 3-D measurement (front view) is shown in Fig. 12(c). Color density means the depth of the object in this figure. The bird eye view result

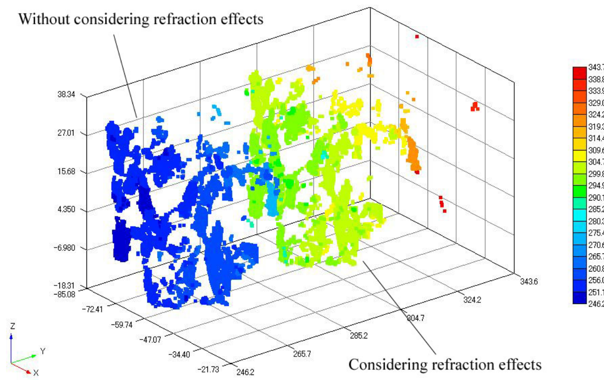
(a) Original image.



(b) Moving object.



(c) 3-D measurement (front view).



(d) 3-D measurement (bird eye view).

Fig. 12. Experimental result 2.

is shown in Fig. 12(d), and it is verified that 3-D position and shape can be measured with high accuracy by considering the refraction effects.

We also measured the 3-D position of the wooden block moving horizontally in the tank (Fig. 13(a)–(f)). We used a robot manipulator to move the wooden block, and estimated measurement result by comparing it with the actual path that was made by the manipulator. The measured trajectory of the object is shown in Fig. 13(g). This shows that the target was measured nearer to the cameras by 20−30mm compared with the actual path when refraction is not considered. The target was measured near to the actual path when the refractive index was taken into consideration.

V. Conclusion

We proposed a method for extracting measurement objects from noisy images with air bubbles in aquatic environments, and a method for 3-D measurement of the extracted objects using a stereo camera system. We confirmed the validity of the proposed method by experiments.

However, misclassification sometimes happens when different moving objects come into the field of view. It is also a problem that the processing cost is rather high. Therefore,

(a) Left image 1.

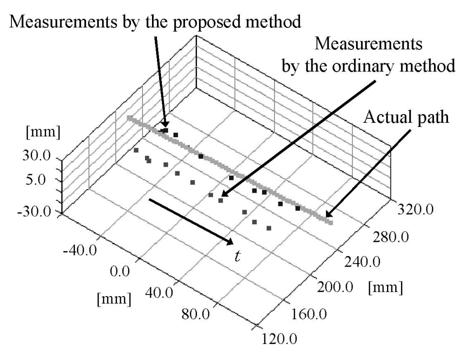
(b) Right image 1.

(c) Left image 2.

(d) Right image 2.

(e) Left image 3.

(f) Right image 3.



(g) Trajectory of the moving object.

Fig. 13. Experimental result 3.

it will be necessary to improve the extraction method of measurement objects in the future (*e.g.* [24], [25]). It will be necessary to examine the problem of measurement accuracy such as the corresponding point search and camera calibration for 3-D measurement (*e.g.* [11]).

References

[1] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald and P. J. W. Rayner: "Detection of Missing Data in Image Sequences," *IEEE Transactions on Image Processing*, Vol.4, No.11, pp.1496–1508, 1995.

[2] A. Nagai, Y. Kuno and Y. Shirai: "Surveillance System Based on Spatio-Temporal Information," *Proceedings of the 1996 IEEE International Conference on Image Processing*, Vol.2, pp.593–596, 1996.

[3] T. Matsuyama, T. Ohya and H. Habe: "Background Subtraction for Non-Stationary Scenes," *Proceedings of the 4th Asian Conference on Computer Vision*, pp.662–667, 2002.

[4] H. Hase, K. Miyake and M. Yoneda: "Real-time Snowfall Noise Elimination," *Proceedings of the 1999 IEEE International Conference on Image Processing*, Vol.2, pp.406–409, 1999.

[5] A. Yamashita, T. Kaneko and K. T. Miura: "A Virtual Wiper –Restoration of Deteriorated Images by Using a Pan-Tilt Camera–," *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, pp.4724–4729, 2004.

[6] K. Garg, S. K. Nayar: "Detection and Removal of Rain from Videos," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.1, pp.528–535, 2004.

[7] A. Yamashita, Y. Tanaka and T. Kaneko: "Removal of Adherent Waterdrops from Images Acquired with Stereo Camera," *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.953–958, 2005.

[8] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: "Image Inpainting," *Proceedings of SIGGRAPH 2000*, pp.417–424, 2000.

[9] B. W. Coles: "Recent Developments in Underwater Laser Scanning Systems," *SPIE Vol.980 Underwater Imaging*, pp.42–52, 1988.

[10] R. F. Tusting and D. L. Davis: "Laser Systems and Structured Illumination for Quantitative Undersea Imaging," *Marine Technology Society Journal*, Vol.26, No.4, pp.5–12, 1992.

[11] N. Pessel, J. Opderbecke, M.-J. Aldon: "Camera Self-Calibration in Underwater Environment," *Proceedings of the 11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, pp.104–110, 2003.

[12] R. Li, H. Li, W. Zou, R. G. Smith and T. A. Curran: "Quantitive Photogrammetric Analysis of Digital Underwater Video Imagery," *IEEE Journal of Oceanic Engineering*, Vol.22, No.2, pp.364–375, 1997.

[13] H. Kondo and T. Ura: "Detailed Object Observation by Autonomous Underwater Vehicle with Localization Involving Uncertainty of Magnetic Bearings," *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, pp.412–419, 2002.

[14] A. Yamashita, E. Hayashimoto, T. Kaneko and Y. Kawata: "3-D Measurement of Objects in a Cylindrical Glass Water Tank with a Laser Range Finder," *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1578–1583, 2003.

[15] A. Yamashita, H. Higuchi, T. Kaneko and Y. Kawata: "Three Dimensional Measurement of Object's Surface in Water Using the Light Stripe Projection Method," *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, pp.2736–2741, 2004.

[16] A. Yamashita, S. Ikeda and T. Kaneko: "3-D Measurement of Objects in Unknown Aquatic Environments with a Laser Range Finder," *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pp.3923–3928, 2005.

[17] H. Kondo, T. Maki, T. Ura, Y. Nose, T. Sakamaki and M. Inaishi: "Relative Navigation of an Autonomous Underwater Vehicle Using a Light-Section Profiling System," *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1103–1108, 2004.

[18] H. Saito, H. Kawamura and M. Nakajima: "3D Shape Measurement of Underwater Objects Using Motion Stereo," *Proceedings of 21th International Conference on Industrial Electronics, Control, and Instrumentation*, pp.1231–1235, 1995.

[19] H. Murase: "Surface Shape Reconstruction of a Nonrigid Transparent Object Using Refraction and Motion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.1045–1052, 1992.

[20] G. H. Ball and D. J. Hall: *ISODATA – Novel Method of Data Analysis and Pattern Classification*, Stanford Research Institute, 1965.

[21] K. Takahashi, T. Sugiyama and K. Abe: "Region Segmentation of Color Images Based on ISODATA Clustering Method with Image Space Information," *Proceedings of the 4th Asian Conference on Computer Vision*, Vol.1, pp.216–221, 2000.

[22] B. K. P. Horn: *Robot Vision*, The MIT Press, 1986.

[23] J. Weng, P. Cohen and M. Herniou: "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.965–980, 1992.

[24] M. Harville, G. Gordon and J. Woodfill: "Foreground Segmentation Using Adaptive Mixture Models in Color and Depth," *Proceedings of the IEEE Workshop on Detection and Recognition of Events in Video*, pp.3–11, 2001.

[25] D.-S. Lee: "Improved Adaptive Mixture Learning for Robust Video Background Modeling," *Proceedings of IAPR Workshop on Machine Vision Applications*, pp.443–446, 2002.