

Monocular Underwater Stereo – 3D Measurement Using Difference of Appearance Depending on Optical Paths –

Atsushi Yamashita, Yudai Shirane and Toru Kaneko

Abstract—Sensing in aquatic environments meets the difficulty that, when a camera is set in air behind a watertight glass plate, image distortion occurs by refraction of light at the boundary surface between air and water. This paper proposes an aquatic sensing method to execute three dimensional measurement of objects in water by taking refraction of light into account. The proposed method is based on a monocular stereo technique using a difference of appearance depending on optical paths. Optimization of the angles of refracting surfaces which are key elements of the sensing device is given to realize the accurate measurement. Experimental results show the effectiveness of the proposed method.

I. INTRODUCTION

In this paper, we propose a monocular stereo vision system for underwater sensing. The proposed stereo vision system can measure objects in water by using a difference of appearance depending on optical paths.

Underwater imagery is essential to marine environment investigation, and three dimensional (3D) measurement of objects in water gives important information. Photographic 3D measurement methods have the advantage that it is possible to measure objects without contact, and they are well used in atmospheric environments. However, measurement of objects in water by these methods has a problem of light refraction. Image distortion occurs by refraction of light at the boundary surface between air and water when a camera is set in air behind a watertight glass plate.

Figure 1 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. Therefore, it becomes difficult to measure precise positions and shapes of objects when water exists because of the image distortion by light refraction.

Ray tracing solves the problem of light refraction. Accurate results considering the influence of light refraction can be obtained by calculating the direction of rays from the camera before and after light refraction.

There have been proposed a variety of underwater sensing methods such as using a stereo camera [1]–[6], using a single camera [7]–[11], and using a laser light projection [12]–[16].

However, the methods using a stereo camera need two power sources for two cameras. The synchronization mechanisms of two cameras are also needed. Those using a laser

This research was in part supported by MEXT KAKENHI, Grant-in-Aid for Young Scientist (A), 22680017, and the Asahi-Glass Foundation.

A. Yamashita, Y. Shirane and T. Kaneko are with Department of Mechanical Engineering, Shizuoka University, 3–5–1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka 432–8561, Japan yamashita@ieee.org

also need two power sources for a camera and a laser. Additionally, the whole sizes of these sensors become large. These problems of power supply and size are critical in underwater sensing because underwater robots have several limitations. The methods by the use of single camera images obtained based on a motion stereo or a photometric stereo also have the problem that the relationship between the camera and the object is difficult to estimate with high accuracy.

There are studies about stereo measurement by using a single camera in air [17]–[20]. These studies can obtain a stereo pair of images by using a special hardware such as a biprism because it generates multiple optical paths.

In underwater environments, multiple optical paths can easily occur with simple equipment (not with special prism) thanks to the light refraction effects.

In this paper, we propose a 3D measurement method by using a monocular stereo camera that consists of a normal camera and a waterproof case (Fig. 2). The waterproof case has two refractive surfaces that generate a difference of appearance depending on optical paths.

Figure 3(a) shows an image that is acquired by using the monocular stereo camera when it places in air medium (not in water). Only a single image of a box with checkerboard pattern is observed in Fig. 3(a). On the other hand, Fig. 3(b) shows an image that is acquired in water. Two images of the box can be observed in Fig. 3(b) like a stereo camera system.

II. MEASUREMENT METHOD

A. Image Acquisition

The equivalent of a stereo pair images is formed as the left and right halves of a single image by using the monocular

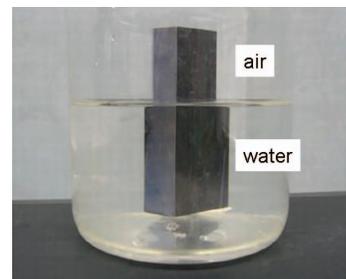


Fig. 1. Light refraction effect. The size and the shape of the object look different between above and below the water surface.

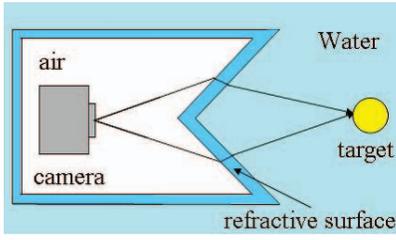
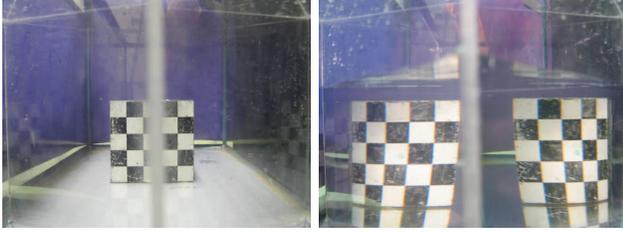


Fig. 2. Two optical paths. The waterproof case has two refractive surfaces that generate a difference of appearance depending on optical paths.



(a) From air to air. (b) From air to water.

Fig. 3. Example of acquired images. (a) Image that is acquired by using the monocular stereo camera when it places in air medium (not in water). (b) Image that is acquired in water. Two images of the box can be observed like a stereo camera system.

stereo camera system (Fig. 3(b)). The left half of a single image is regarded as a left image, and the right half is regarded as a right image, respectively, in a similar way of a stereo camera.

B. Optical Ray Tracing

The ray from the left image (left optical path) is refracted at the boundary of air and waterproof container, and then is refracted at the boundary of waterproof container and water. Finally, the ray projects onto the object in water, and then the ray reflected by the object is refracted again at the boundary of water, container, and air to project onto the right image plane of the camera (right optical path). This phenomenon can be analyzed by ray tracing [1].

Figure 4 shows light refraction effects from air to watertight container and from watertight container to water.

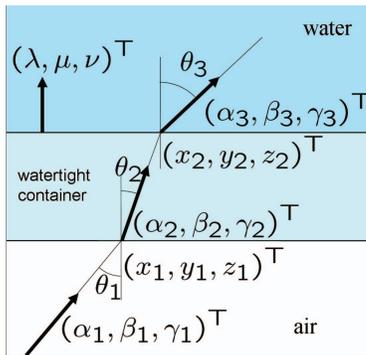


Fig. 4. Light refraction effect from air to water. Ray tracing can be done by considering Snell's law of refraction.

Here, let refractive indices of air and waterproof case be n_1 and n_2 , incident and refractive angles from air to waterproof case be θ_1 and θ_2 , respectively. A unit vector of ray in waterproof case $(\alpha_2, \beta_2, \gamma_2)^T$ can be calculated by using a unit vector of ray from air $(\alpha_1, \beta_1, \gamma_1)^T$ and a unit normal vector of waterproof case $(\lambda, \mu, \nu)^T$ as follows.

$$\begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} = \frac{n_1}{n_2} \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} + c_1, \quad (1)$$

where

$$c_1 = \left\{ \sqrt{1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2 \theta_1} - \frac{n_1}{n_2} \cos \theta_1 \right\} \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}.$$

A unit vector in water $(\alpha_3, \beta_3, \gamma_3)^T$ is also calculated by using the refractive index of water n_3 and the refractive angle of water θ_3 .

$$\begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix} = \frac{n_2}{n_3} \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix} + c_2, \quad (2)$$

where

$$c_2 = \left\{ \sqrt{1 - \left(\frac{n_2}{n_3}\right)^2 \sin^2 \theta_2} - \frac{n_2}{n_3} \cos \theta_2 \right\} \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}.$$

When Snell's law of refraction is applied, the following equation is obtained:

$$\theta_2 = \sin^{-1} \left(\frac{n_1}{n_2} \sin \theta_1 \right). \quad (3)$$

The ray from the camera finally reaches on the surface of the underwater object at the point $P(x_P, y_P, z_P)^T$.

$$\begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix} = s \begin{pmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}, \quad (4)$$

where s is a constant and $(x_2, y_2, z_2)^T$ is the intersection point between the ray from waterproof case and the refraction boundary, respectively.

C. 3D Measurement

Two rays are calculated by ray tracing from the left and the right images thanks to the monocular stereo camera, and the intersection of the two rays gives the 3D coordinates of the target point in water.

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. In aerial environments, the epipolar line is straight. However, the epipolar lines are not straight in aquatic environments because of the refraction of light. Therefore, we calculate the epipolar lines with the ray tracing technique in the same way of Section II-B.

Corresponding points on epipolar lines are searched for with template matching by using the normalized cross correlation (NCC) method (Fig. 5).

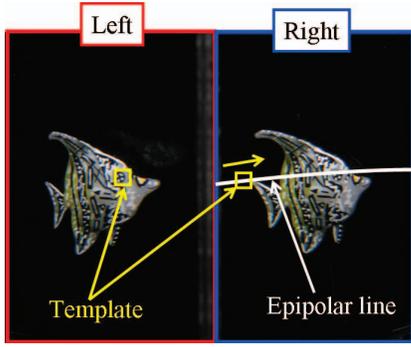


Fig. 5. Stereo matching. Corresponding points on epipolar lines are searched for with template matching.

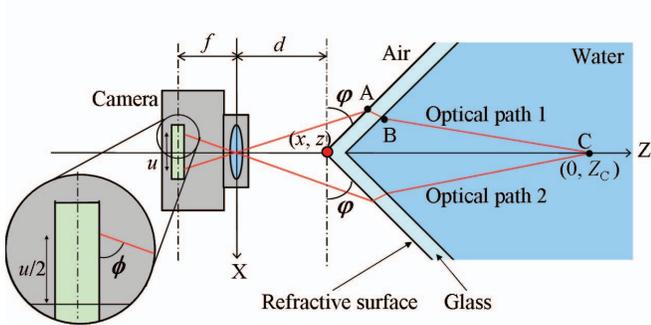


Fig. 6. Parameters for optimal design of refractive surface.

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

III. OPTIMAL DESIGN OF REFRACTIVE SURFACE

The accuracy of 3D measurement depends on the angle of the refractive surface. Therefore, we design a monocular stereo vision system that has an optimal angle of the refractive surface.

The accuracy of stereo measurement becomes high when the rate of change of disparity between left and right images becomes large in the case there is tiny variation of position along depth direction. Therefore, we can optimize the angle of the refractive surface φ based on the relationship among φ , the rate of the change of disparity Δu , and the position change of depth direction Δz .

Figure 6 shows parameters for the optimal design of the refractive surface. In Fig. 6, f is an image distance (focal length), d is the distance between the lens center and the position of the intersection point of two refractive surfaces, z_C is the length about depth direction of the intersection point of two optical paths, and u is the disparity, respectively.

As to the optical path 1, the ray from the camera in air is on the line expressed as Eq.(5).

$$Z = \tan \phi \cdot X, \quad (5)$$

where

$$\tan \phi = \frac{f}{u/2} = \frac{2f}{u}. \quad (6)$$

The position of the intersection point of two refractive surfaces is calculated in Eq.(7).

$$Z = \tan \varphi \cdot X + d. \quad (7)$$

From Eq.(5)–(7), we can obtain the position of A (x_A, z_A).

$$x_A = \frac{du}{2f - u \tan \varphi}, \quad (8)$$

$$z_A = \frac{2df}{2f - u \tan \varphi}. \quad (9)$$

The ray from A to B is on the line expressed in Eq.(10).

$$Z = -\frac{1}{\tan(\varphi - \theta_2)}(X - x_A) + z_A. \quad (10)$$

From Eq.(3) and Eq.(11), we can obtain θ_2 .

$$\theta_1 = \varphi - \phi + \frac{\pi}{2}. \quad (11)$$

On the other hand, the refractive surface on the water side is expressed in Eq.(12).

$$Z = \tan \varphi \cdot X + d + \frac{t}{\cos \varphi}, \quad (12)$$

where t is the thickness of the glass (refractive surface).

From Eq.(10) and Eq.(12), we can obtain the coordinate of B (x_B, z_B).

$$x_B = \frac{x_A \cos \varphi + (z_A \cos \varphi - d \cos \varphi - t) \tan(\varphi - \theta_2)}{\sin \varphi \tan(\varphi - \theta_2) + \cos \varphi}, \quad (13)$$

$$z_B = x_B \tan \varphi + d + \frac{t}{\cos \varphi}. \quad (14)$$

The ray from B to C is expressed as follows:

$$Z = -\frac{1}{\tan(\varphi - \theta_3)}(X - x_B) + z_B, \quad (15)$$

where

$$\theta_3 = \sin^{-1} \left(\frac{n_2}{n_3} \sin \theta_2 \right). \quad (16)$$

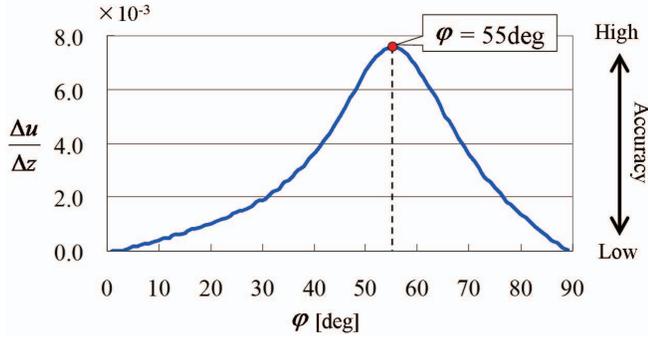
Finally, the position of the intersection point of two optical paths can be obtained.

$$\begin{aligned} z_C &= -\frac{1}{\tan(\varphi - \theta_3)}x_B + z_B \\ &= g(u, \varphi). \end{aligned} \quad (17)$$

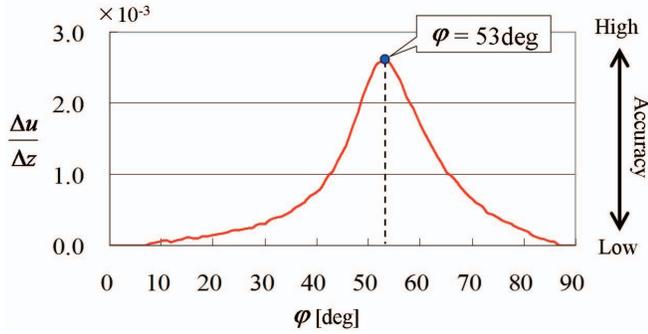
By differentiating Eq.(17) numerically, the following equation can be gained.

$$h \left(\frac{\Delta u}{\Delta z}, \varphi \right) = 0. \quad (18)$$

Equation (18) expresses the relationship between the disparity derivatives $\frac{\Delta u}{\Delta z}$ and the angle of the refractive surface φ .



(a) $z_C = 200$.



(b) $z_C = 500$.

Fig. 7. Optimal angle of the refractive surface. The accuracy of 3D measurement becomes high when the disparity derivatives $\frac{\Delta u}{\Delta z}$ is large. The optimal angle of the refractive surface changes when z_C changes. (a) $z_C = 200$. (b) $z_C = 500$.

The accuracy of 3D measurement becomes high when the disparity derivatives $\frac{\Delta u}{\Delta z}$ is large. However, the optimal angle of the refractive surface changes when z_C changes.

For example, the optimal angle φ is 55deg when $z_C = 200$ mm (Fig. 7(a)). The optimal angle φ is 53deg when $z_C = 500$ mm (Fig. 7(b)).

In this way, the optimal angle depends on the distance between the camera and the object.

The distance between the camera and the object is unknown before the measurement. Therefore, we consider that the angle should be optimized in a certain range $z_1 \leq z_C \leq z_2$. The worst accuracy in the range $z_1 \leq z_C \leq z_2$ is optimized as follows:

$$\varphi_{opt} = \arg_{\varphi} \left[\max_{\varphi} \left\{ \min_{z_1 \leq z_C \leq z_2} \left(\frac{\Delta u}{\Delta z} \right) \right\} \right]. \quad (19)$$

Figure 8 shows an example of the optimal angle. The accuracy in the case of φ_1 is higher than that of φ_2 when $z_C = z_1$. However, the accuracy of φ_1 is lower than that of φ_2 when $z_C = z_2$. Therefore, we regard φ_2 as the optimal angle of the refractive surface.

IV. EXPERIMENT

A. Optimal Design of Refractive Surface

In this paper, we used the camera whose focal length was 4.6mm and whose resolution was 2048×1536pixel. The

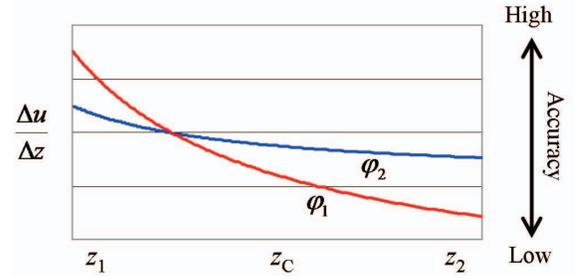


Fig. 8. Relationship between disparity derivatives and the angle of the refractive surface.

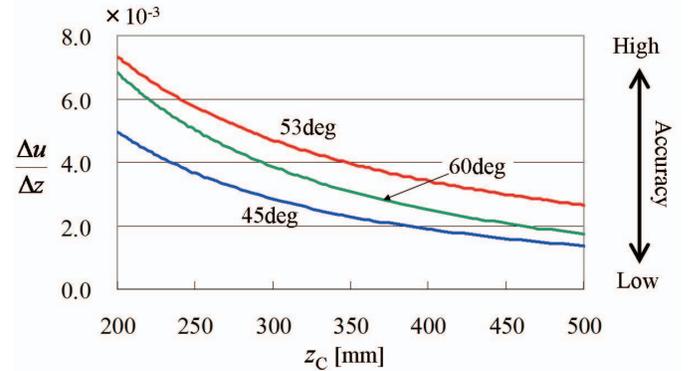


Fig. 9. Determination of the optimal angle of the refractive surface. Blue line: $\varphi = 45$ deg. Red line: $\varphi = 53$ deg. Green line: $\varphi = 60$ deg.

distance between the lens center and the intersection point of two refractive surfaces was 30mm ($d = 30$ mm) and the measurement range was set as $200\text{mm} \leq z_C \leq 500\text{mm}$.

Under the above conditions, the optimal angle of refractive surface is calculated. Figure 9 shows the accuracy of 3D measurement when φ equals to 45deg, 53deg, and 60deg. From the result of numerical calculation, the optimal angle was calculated as $\varphi = 53$ deg.

B. Prototype of Monocular Stereo System

We developed two prototypes of the monocular stereo vision whose angles of the refractive surface were 53deg and 45deg, respectively (Fig. 10). The prototype whose angle was 45deg was used for comparison. Refractive surfaces of a waterproof case are made with acrylic plates.

To calibrate the monocular stereo vision system, at first, the intrinsic parameters of the camera were calibrated without the waterproof case. Then the position of the intersection point of two refractive surfaces and the angle between the optical axis of the camera and each refractive surface were estimated precisely by observing the planar checkerboard pattern whose size was known.

C. 3D Measurement

To verify the accuracy of the 3D measurement quantitatively, a rectangular object whose upper surface size was 40mm×40mm was measured (Fig. 11).

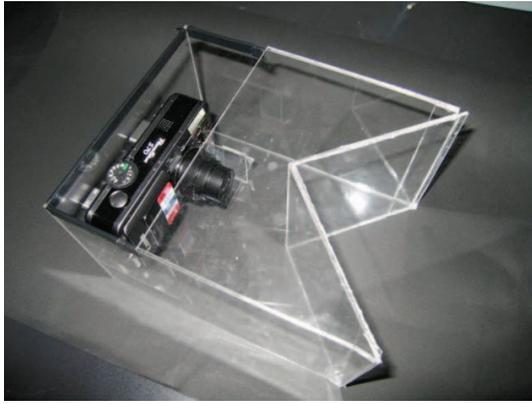


Fig. 10. Monocular stereo vision system. Refractive surfaces of a waterproof case are made with acrylic plates.

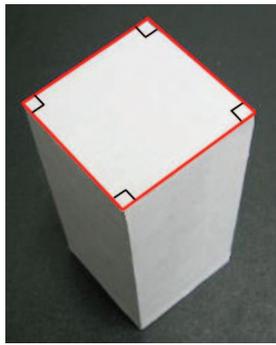


Fig. 11. Rectangular object whose upper surface size was 40mm×40mm.

Tables I and II show the accuracy of measurement results when the angle of the refractive surface $\varphi = 45\text{deg}$ and 53deg , respectively.

In the case of $\varphi = 45\text{deg}$, the average error of the upper surface edge was 0.93mm when $z_C = 200\text{mm}$. The error was 1.18mm when $z_C = 400\text{mm}$. The average angle error of 90deg corner was 1.15deg when $z_C = 200\text{mm}$, and the angle error was 1.18deg when $z_C = 400\text{mm}$, respectively.

On the other hand, in the case of $\varphi = 53\text{deg}$ (optimal angle), the average errors were 0.70mm when $z_C = 200\text{mm}$, and 1.13mm when $z_C = 400\text{mm}$, respectively. The average angle error of 90deg corner was 0.85deg when $z_C = 200\text{mm}$, and the angle error was 0.28deg when $z_C = 400\text{mm}$, respectively.

From these results, it is verified that the optimization can improve the accuracy of 3D measurement.

We also measured the 3D position of the fish-like object

TABLE I
MEASUREMENT RESULT ($\varphi = 45\text{DEG}$).

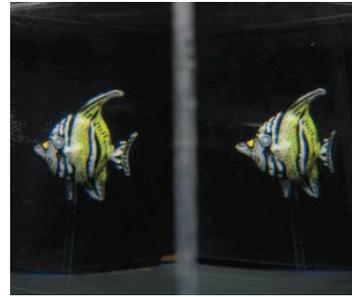
	length	angle
average error ($z = 200\text{mm}$)	0.93mm	1.15deg
average error ($z = 400\text{mm}$)	1.18mm	1.18deg
standard deviation ($z = 200\text{mm}$)	0.88mm	1.53deg
standard deviation ($z = 400\text{mm}$)	1.28mm	0.61deg

TABLE II
MEASUREMENT RESULT ($\varphi = 53\text{DEG}$).

	length	angle
average error ($z = 200\text{mm}$)	0.70mm	0.85deg
average error ($z = 400\text{mm}$)	1.13mm	0.28deg
standard deviation ($z = 200\text{mm}$)	0.73mm	0.96deg
standard deviation ($z = 400\text{mm}$)	1.13mm	0.34deg



(a) Overview.



(b) Acquired image.

Fig. 12. Fish-like object. (a) Overview. (b) Acquired image.

(Fig. 12). Figure 12(a) shows an overview of the fish-like object, and Fig. 12(b) shows an example of the acquired image.

Figure 13 shows the shape measurement result by using our monocular stereo vision system.

The fish-like object was moved in water by using a manipulator whose trajectory is known (Fig. 14). The distance between the camera system and the target object was about 350mm. The measured trajectory of the object is shown in Fig. 15. The target was measured near to the actual path.



Fig. 13. 3D measurement result.

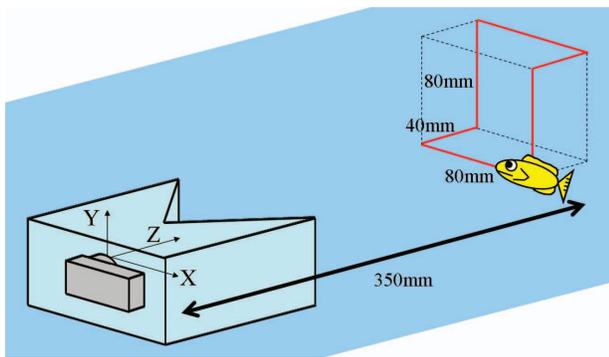


Fig. 14. Trajectory of target object. The target object was moved in water by using a manipulator whose trajectory is known. The distance between the camera system and the target object was about 350mm.

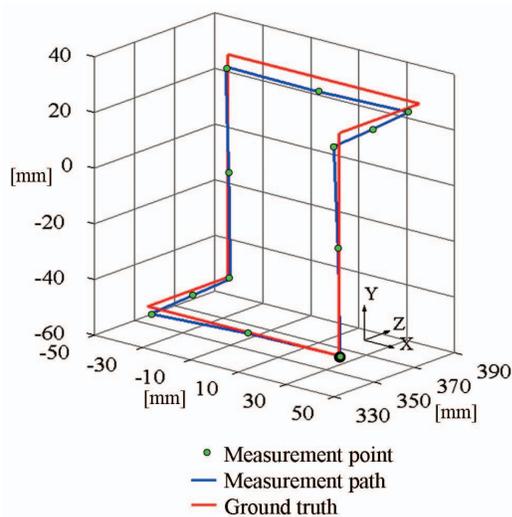


Fig. 15. Measurement result of trajectory. Green point: measured point. Blue line: measured trajectory. Red line: ground truth.

V. CONCLUSION

In this paper, we propose an underwater sensing method by using a monocular stereo vision system. The proposed method is based on a monocular stereo technique using a difference of appearance depending on optical paths. The angle of refracting surface is optimized by considering the measurement accuracy. Experimental results show the effectiveness of the proposed method.

As a future work, the accuracy of the monocular stereo vision system should be improved by improving the assembly accuracy of the camera system. Several phenomena typical of underwater environments such as absorption [10], forward scattering and backscattering should be also considered.

REFERENCES

[1] Rongxing Li, Haihao Li, Weihong Zou, Robert G. Smith and Terry A. Curran: "Quantitative Photogrammetric Analysis of Digital Underwater Video Imagery," *IEEE Journal of Oceanic Engineering*, Vol.22, No.2, pp.364–375, 1997.

[2] Mark R. Shortis and Euan S. Harvey: "Design and Calibration of an Underwater Stereo-Vision System for the Monitoring of Marine Fauna Populations," *International Archives of Photogrammetry and Remote Sensing*, Vo.32, No.5, pp.792–799, 1998.

[3] Atsushi Yamashita, Susumu Kato and Toru Kaneko: "Robust Sensing against Bubble Noises in Aquatic Environments with a Stereo Vision System," *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA2006)*, pp.928–933, 2006.

[4] Matteo Zoppi and Rezia Molfino: "ArmilleEye: Flexible Platform for Underwater Stereo Vision," *Journal of Mechanical Design*, Vol.129, No.8, pp.808–815, 2007.

[5] Atsushi Yamashita, Akira Fujii and Toru Kaneko: "Three Dimensional Measurement of Objects in Liquid and Estimation of Refractive Index of Liquid by Using Images of Water Surface with a Stereo Vision System," *Proceedings of the 2008 IEEE International Conference on Robotics and Automation (ICRA2008)*, pp.974–979, 2008.

[6] Visesh Chari and Peter Sturm: "Multi-View Geometry of the Refractive Plane," *Proceedings of British Machine Vision Conference 2009 (BMVC2009)*, pp.1–11, 2009.

[7] Hiroshi 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.

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

[9] Numo Gracias and Jose Santos-Victor: "Underwater Video Mosaics as Visual Navigation Maps," *Computer Vision and Image Understanding*, Vol.79, No.1, pp.66–91, 2000.

[10] Atsushi Yamashita, Megumi Fujii and Toru Kaneko: "Color Registration of Underwater Images for Underwater Sensing with Consideration of Light Attenuation," *Proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA2007)*, pp.4570–4575, 2007.

[11] Tali Treibitz, Yoav Y. Schechner and Hanuman Singh: "Flat Refractive Geometry," *Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2008)*, pp.1–8, 2008.

[12] Atsushi Yamashita, Etsukazu Hayashimoto, Toru Kaneko and Yoshimasa 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 (IROS2003)*, pp.1578–1583, 2003.

[13] Atsushi Yamashita, Hirokazu Higuchi, Toru Kaneko and Yoshimasa 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 (ICRA2004)*, pp.2736–2741, 2004.

[14] Atsushi Yamashita, Shinsuke Ikeda and Toru 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 (ICRA2005)*, pp.3923–3928, 2005.

[15] Hayato Kondo, Toshihiro Maki, Tamaki Ura, Yoshiaki Nose, Takashi Sakamaki and Masaaki 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 (IROS2004)*, pp.1103–1108, 2004.

[16] Ryohei Kawai, Atsushi Yamashita and Toru Kaneko: "Three-Dimensional Measurement of Objects in Water by Using Space Encoding Method," *Proceedings of the 2009 IEEE International Conference on Robotics and Automation (ICRA2009)*, pp.2830–2835, 2009.

[17] Edward H. Adelson and John Y. A. Wang: "Single Lens Stereo with a Plenoptic Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.2, pp.99–106, 1992.

[18] Ardeshir Goshtasby and William A. Gruver: "Design of a Single-Lens Stereo Camera System," *Pattern Recognition*, Vol.26, No.6, pp.923–937, 1993.

[19] Doo Hyun Lee, In So Kweon and Roberto Cipolla: "Single Lens Stereo with a Biprism," *Proceedings of the 1998 IAPR Workshop on Machine Vision Applications (MVA1998)*, pp.136–139, 1998.

[20] Doo Hyun Lee, In So Kweon and Roberto Cipolla: "A Biprism-Stereo Camera System," *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR1999)*, Vol.1, pp.82–87, 1999.