

# Scale Reconstructable Structure from Motion Using Refraction with Omnidirectional Camera

Hiromasa Kume\*, Hiromitsu Fujii\*, Atsushi Yamashita\*, and Hajime Asama\*

\*Department of Precision Engineering

The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Email: kume@robot.t.u-tokyo.ac.jp

**Abstract**—Structure from Motion (SfM) is a technique for 3D reconstruction using a single camera. The absolute scale of objects cannot be reconstructed by the conventional SfM. Recently, a method of SfM that can reconstruct the absolute scale has been proposed using refraction at the boundaries of a transparent plate in front of an ordinary camera. In this research, we develop a new method of scale reconstructable SfM with a wide field of view using an omnidirectional camera and a cylindrical transparent housing. After the essential equation is extended for the system, simulation experiments assuming air and underwater environments are conducted. The results show the theoretical validity of this method.

## I. INTRODUCTION

Structure from Motion (SfM) is a technique for 3D reconstruction using image information. In SfM, 3D measurements can be achieved by acquiring images of objects while a camera moves from one viewpoint to another. SfM needs only one camera, and therefore, its system is simpler than that of methods with stereo cameras or conventional active stereo vision methods because such methods require multiple cameras or a projector other than a camera. SfM has been actively researched in recent years [1], [2]. In particular, the recently developed SfM using an omnidirectional camera [3]-[5] with its wide field of view is suited to 3D reconstruction of surrounding environments. However, unlike methods that use stereo cameras or active stereo vision, SfM cannot reconstruct the absolute scale of objects.

Other than SfM, there are methods in which the absolute scale of objects can be obtained with one camera [6]- [8]. In these methods, the light paths are changed by using optical phenomena including refraction or reflection, and therefore parallax images of objects can be obtained. However, such methods have the problem that fields of view are small because the acquired images are divided into two parts that each contain a parallax image. In a similar study [9], the depth map is obtained from an image captured through a transparent plate between the scene and the camera and from an image captured without the plate. However, the fields of view are limited because an ordinary camera is used in the method.

In SfM, the camera motion and shapes of objects are estimated simultaneously by using the information of corresponding points in the acquired images. In conventional SfM, it is necessary to solve an equation based on a geometric relationship related to the directions from which the rays originate and the unknown translational movement of the

camera in order to estimate the camera motion. Although the rotation and the direction of the translational movement of the camera are estimated uniquely from the equation, the amount of the translational movement cannot be estimated by the equation. This is because the epipolar geometry allows for reconstructing the structure and motion up to an overall scale. Therefore, the 3D shape of an object expressed in only the relative size compared with the amount of translation of the camera can be reconstructed. Shibata *et al.* proposed a method of SfM that can reconstruct the absolute scale with a transparent plate in front of an ordinary camera that causes refraction [10]. However, since an ordinary camera is used in the method, the field of view is small.

In this paper, an omnidirectional camera is used to secure a wide field of view, and we solve the scale problem by means of the following approach: First, as an approach related to a measurement device, an omnidirectional camera is covered by a cylindrical transparent housing to generate refraction through a full 360 degrees. This refraction acts as a constraint in a particular geometric relation. The constraint is used to determine the amount of translational movement of the camera. Second, we derive an equation from the geometric relation. By solving it, the camera motion can be estimated including the amount of the translation and thus the 3D shape of an object can be reconstructed with the absolute scale.

The proposed method can be applied to measurements at both land sites and underwater sites by generating refraction. In particular, the proposed method is more effective at underwater sites than at land sites because the amount of refraction is larger in water than in air. Our simulation results demonstrate the validity of this method. They show that shapes of objects are reconstructed with their absolute scale by the proposed method.

## II. 3D MEASUREMENT METHOD

### A. Omnidirectional Camera

In this research, an omnidirectional camera [11] that has a hyperboloid mirror in front of a lens is used. Rays directed to one of the focal points of the hyperboloid mirror are reflected to the other, and therefore, the mirror is positioned so that the center of the lens and the focal point of the mirror are aligned.

In this research, the omnidirectional camera is covered with a cylindrical transparent housing (Fig. 1). This housing is also used for waterproofing in underwater environments.

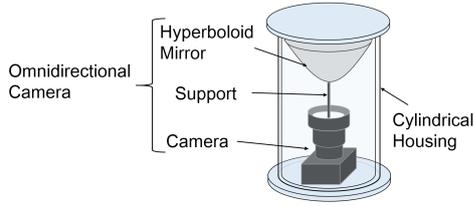


Fig. 1. Omnidirectional camera with a housing. The omnidirectional camera is covered by a cylindrical housing to generate refraction. The axes of the mirror and the housing are in alignment.

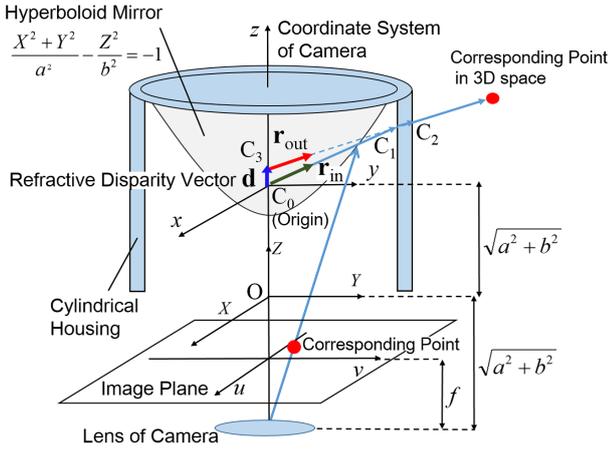


Fig. 2. Schematic of ray vector. The ray is refracted at  $C_2$  and  $C_1$ , and reflected at the mirror before arrival at the image plane. The inner ray vector  $\mathbf{r}_{in}$  is calculated from the image coordinate and known parameters. The vector  $\mathbf{r}_{out}$  represents the outer ray vector and  $\mathbf{d}$  the refractive disparity vector.

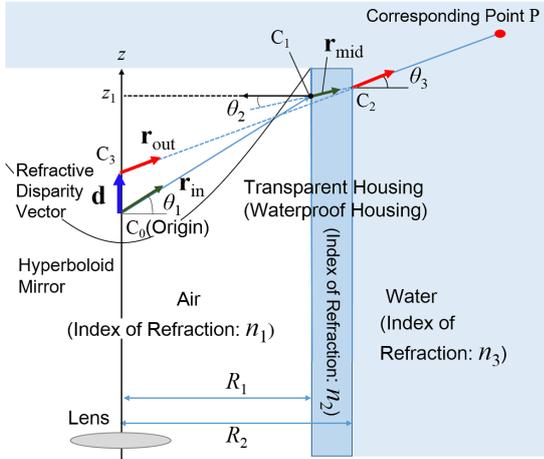


Fig. 3. Ray vectors and refractive disparity vector. The outer ray vector  $\mathbf{r}_{out}$  and refractive disparity vector  $\mathbf{d}$  are calculated from the inner ray vector  $\mathbf{r}_{in}$  by using ray tracing.

### B. Ray vectors

Rays are refracted at the boundary between the housing and its outer surface ( $C_2$ ), and the boundary between the housing and the inner surface ( $C_1$ ) as shown in Fig. 2.

First, an image coordinate  $(u, v)^T$  is transformed into the vector that is directed from the focus of the hyperboloid mirror to the refraction point  $C_1$  on the mirror surface as shown in Fig. 2. We define this unit direction vector as the “inner ray vector”  $\mathbf{r}_{in}$  (Fig. 2). This vector is calculated as given by [3]:

$$\mathbf{r}_{in} = \frac{1}{(su)^2 + (sv)^2 + (sf - 2c)^2} \begin{pmatrix} su \\ sv \\ sf - 2c \end{pmatrix}, \quad (1)$$

$$s = \frac{a^2(f\sqrt{a^2 + b^2} + \sqrt{b^2 f^2 + b^2(u^2 + v^2)})}{a^2 f^2 - b^2(u^2 + v^2)}, \quad (2)$$

where  $a, b$ , and  $c (= \sqrt{a^2 + b^2})$  denote the hyperboloid parameters, and  $f$  denotes the distance between the image plane and the center of the lens of the camera (Fig. 2).

Second, the inner ray vector  $\mathbf{r}_{in}$  is transformed into the vector that is directed from the refraction point  $C_2$  at the outside surface of the housing to the object point in 3D space. We define this unit direction vector as the “outer ray vector”  $\mathbf{r}_{out} = (x, y, z)^T$ . We define the start point of an outer ray vector as the intersection point between two lines: the line passing through the refraction point  $C_2$  and the object point in 3D space, and the  $z$  axis (Fig. 2, Fig. 3). When there is no refraction, the start points of outer ray vectors are always the origin of the coordinate system  $C_0$ . When there is refraction, each start point of them varies based on the amount of refraction. We define this position vector of the start point of an outer ray vector as the “refractive disparity vector”  $\mathbf{d} = (0, 0, d)^T$ , which contains information about the absolute scale of objects. The vectors  $\mathbf{r}_{out}$  and  $\mathbf{d}$  can be calculated by ray tracing [12] using the information of the shape of the housing and the refractive indices of each medium, which are known parameters in this research. The refraction occurring at each boundary is shown in Fig. 3. In this figure, the housing is shown immersed in water as a possible measurement environments in order to simplify our understanding of the figure. The inner ray vector is refracted at  $C_1$  and transformed into the vector  $\mathbf{r}_{mid}$  shown in Fig. 3. Subsequently, the vector  $\mathbf{r}_{mid}$  is refracted at  $C_2$  and transformed into the outer ray vector  $\mathbf{r}_{out}$ . The vectors  $\mathbf{r}_{out}$  and  $\mathbf{d}$  are calculated as follows:

$$\mathbf{r}_{out} = \frac{n_2}{n_3} \mathbf{r}_{mid} - \left\{ \frac{n_2}{n_3} \cos \theta_2 - \sqrt{1 - \left(\frac{n_2}{n_3}\right)^2 \sin^2 \theta_2} \right\} \mathbf{N}_1, \quad (3)$$

$$\mathbf{d} = (0, 0, R_1 \tan \theta_1 + (R_2 - R_1) \tan \theta_2 - R_2 \tan \theta_3)^T, \quad (4)$$

where

$$\mathbf{r}_{mid} = \frac{n_1}{n_2} \mathbf{r}_{in} - \left\{ \frac{n_1}{n_2} \cos \theta_1 - \sqrt{1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2 \theta_1} \right\} \mathbf{N}_1. \quad (5)$$

The parameters  $n_1, n_2, n_3$  denote the refractive indices of the inside of the housing, the housing itself, and the surrounding water respectively, and  $\theta_1, \theta_2, \theta_3$  denote the refractive angles shown in Fig. 3. The parameter  $\mathbf{N}_1$  denotes a unit normal vector at the refraction points  $C_1$  and  $C_2$ .

### C. Extended Essential Equation in Proposed SfM

In this section, first, the conventional method of SfM and the reason as to why the absolute scale of an object cannot be reconstructed are explained. Subsequently, the proposed method and the mechanism of reconstructing the shape of an object with the absolute scale are described. In physical terms, the key lies in the changes in the ray paths caused by refraction. We effectively use the information related to refraction, and we express it as the refractive disparity vector and introduce a new equation using the vector.

There are several SfM techniques including the eight-point algorithm that can estimate camera motion from at least eight pairs of corresponding points in the acquired images [13]-[15], and the five-point algorithm, which uses five pairs of corresponding points [16]-[18]. In this paper, the conventional method of SfM is explained based on the eight-point algorithm, which is one of the standard methods of SfM.

A corresponding point observed from two viewpoints of the camera before and after movement is considered. As shown in Fig. 4, the coordinate system  $C$  is introduced, which is fixed to the camera position before movement, and the coordinate system  $C'$  is fixed to that after movement. The origin of each coordinate system is the focus of the mirror. In this paper, variables in the coordinate system  $C'$  are expressed using the prime symbol “'.” For simplification, the world coordinate system is fixed to the coordinate system  $C$ .

In the conventional SfM, there is no refraction. Therefore, the outer ray vectors are identical to the inner ray vectors, and both are described as “ray vectors” in this paper. The ray vectors  $\mathbf{r} = (x, y, z)^T$  and  $\mathbf{r}' = (x', y', z')^T$  are obtained as described in Sec. II-B. The camera motion is expressed by the rotation matrix  $\mathbf{R}$  and the translation vector  $\mathbf{t}$ :

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}. \quad (6)$$

In conventional methods of SfM, the geometric relation indicating that the two ray vectors and the translation vector lie on the same plane (Fig. 4) is represented as follows:

$$\begin{pmatrix} xx' \\ yy' \\ zx' \\ xy' \\ yy' \\ zy' \\ xz' \\ yz' \\ zz' \end{pmatrix}^T \begin{pmatrix} r_{12}t_3 - r_{13}t_2 \\ r_{13}t_1 - r_{11}t_3 \\ r_{11}t_2 - r_{12}t_1 \\ r_{22}t_3 - r_{23}t_2 \\ r_{23}t_1 - r_{21}t_3 \\ r_{21}t_2 - r_{22}t_1 \\ r_{32}t_3 - r_{33}t_2 \\ r_{33}t_1 - r_{31}t_3 \\ r_{31}t_2 - r_{32}t_1 \end{pmatrix} = 0, \quad (7)$$

$$\iff \mathbf{b}^T \mathbf{e} = 0. \quad (8)$$

This equation is called the essential equation. By solving a system of equations about more than eight corresponding points, the rotation matrix  $\mathbf{R}$  and the vector of the direction of the translation  $\mathbf{t}$  can be obtained [14]. The relative scale of objects compared with the norm of the translation vector

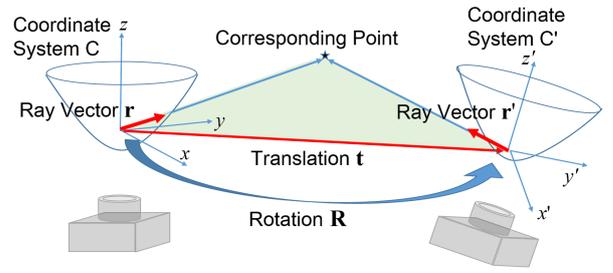


Fig. 4. Geometry in the conventional SfM. The two ray vectors ( $\mathbf{r}$  and  $\mathbf{r}'$ ) and the translation vector  $\mathbf{t}$  lie on the same plane.

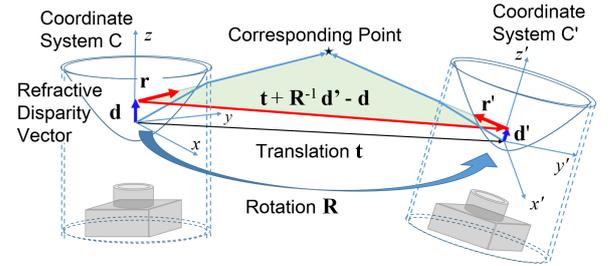


Fig. 5. Geometry in the proposed method. The two outer ray vectors ( $\mathbf{r}$  and  $\mathbf{r}'$ ) and the vector between their start points ( $\mathbf{t} + \mathbf{R}^{-1} \mathbf{d}' - \mathbf{d}$ ) lie on the same plane. Unlike the geometry in the conventional method, this includes refractive disparity vectors ( $\mathbf{d}$  and  $\mathbf{d}'$ ), which are related to the absolute scale of objects.

can be obtained. However, Eq. (7) is satisfied even if  $\mathbf{t}$  is multiplied by any constant, which means that the norm of the translation vector cannot be determined uniquely. Therefore the absolute scale of objects cannot be reconstructed.

In the proposed method, the essential equation (Eq. 8) is extended by also considering refraction. In our method,  $\mathbf{r} = (x, y, z)^T$  and  $\mathbf{r}' = (x', y', z')^T$  denote the outer ray vectors belonging to a corresponding point while  $\mathbf{d} = (0, 0, d)^T$  and  $\mathbf{d}' = (0, 0, d')^T$  denote refractive disparity vectors, which are the position vectors of the start points of the outer ray vectors. Unlike in the case of the conventional SfM, refractive disparity vectors are generally not zero vectors, and they contain information about the absolute scale of objects.

In the following explanation, the outer ray vectors and refractive disparity vectors are assumed to have already been calculated in the manner described in Sec. II-B.

In the proposed method, we use the following geometric relationship shown in Fig. 5: the two outer ray vectors ( $\mathbf{r}$  and  $\mathbf{r}'$ ) and the vector between their start points ( $\mathbf{t} + \mathbf{R}^{-1} \mathbf{d}' - \mathbf{d}$ ) lie on the same plane. This relationship is transformed into the following equation [10]:

$$((\mathbf{t} + \mathbf{R}^{-1} \mathbf{d}' - \mathbf{d}) \times \mathbf{R}^{-1} \mathbf{r}')^T \mathbf{r} = 0. \quad (9)$$

It is to be noted that  $\mathbf{r}'$  and  $\mathbf{d}'$  are expressed in the coordinate system  $C$  as  $\mathbf{R}^{-1} \mathbf{r}'$  and  $\mathbf{R}^{-1} \mathbf{d}'$ , respectively.

In order to simplify the equation, an alternative matrix  $\mathbf{Q}$  is introduced, and it satisfies the following equation about any vector  $\mathbf{x}$ :

$$(\mathbf{t} + \mathbf{R}^{-1} \mathbf{d}' - \mathbf{d}) \times \mathbf{x} = \mathbf{Q} \mathbf{x}. \quad (10)$$

$\mathbf{Q}$  is expressed as:

$$\mathbf{Q} = \begin{pmatrix} 0 & -t_3 - r_{33}d' + d & t_2 + r_{32}d' \\ t_3 + r_{33}d' - d & 0 & -t_1 - r_{31}d' \\ -t_2 - r_{32}d' & t_1 + r_{31}d' & 0 \end{pmatrix}. \quad (11)$$

Consequently, Eq. (9) is expressed as:

$$\mathbf{r}'^T \mathbf{R} \mathbf{Q} \mathbf{r} = 0. \quad (12)$$

This equation can be expressed as an inner product of a vector with unknown components and that with known ones. Using the orthogonality of the rotation matrix, we have

$$\begin{pmatrix} r_{12}r_{33} - r_{13}r_{32} \\ r_{13}r_{31} - r_{11}r_{33} \\ r_{11}r_{32} - r_{12}r_{31} \end{pmatrix} = \begin{pmatrix} -r_{12} \\ -r_{22} \\ -r_{32} \end{pmatrix}, \quad \begin{pmatrix} r_{22}r_{33} - r_{23}r_{32} \\ r_{23}r_{31} - r_{21}r_{33} \\ r_{21}r_{32} - r_{22}r_{31} \end{pmatrix} = \begin{pmatrix} r_{11} \\ r_{21} \\ r_{31} \end{pmatrix}. \quad (13)$$

Consequently, Eq. (12) the equation is simplified as:

$$\begin{pmatrix} xx' \\ yx' \\ zx' \\ xy' \\ yy' \\ zy' \\ xz' \\ yz' \\ zz' \\ dyx' + d'xy' \\ -dxx' + d'yy' \\ d'zy' \\ dyy' - d'xx' \\ -dxy' - d'yx' \\ -d'zx' \\ dyz' \\ -dxz' \end{pmatrix}^T \begin{pmatrix} r_{12}t_3 - r_{13}t_2 \\ r_{13}t_1 - r_{11}t_3 \\ r_{11}t_2 - r_{12}t_1 \\ r_{22}t_3 - r_{23}t_2 \\ r_{23}t_1 - r_{21}t_3 \\ r_{21}t_2 - r_{22}t_1 \\ r_{32}t_3 - r_{33}t_2 \\ r_{33}t_1 - r_{31}t_3 \\ r_{31}t_2 - r_{32}t_1 \\ r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{31} \\ r_{32} \end{pmatrix} = 0, \quad (14)$$

$$\iff \mathbf{u}^T \mathbf{g} = 0, \quad (15)$$

where  $r_{i,j}$  and  $t_i$  denote the  $(i,j)$ -th component of  $\mathbf{R}$  and the  $i$ -th component of  $\mathbf{t}$ , respectively.

By comparing  $\mathbf{e}$  of Eq. (8) with  $\mathbf{g}$  of Eq. (15), it is confirmed that the first to ninth components of  $\mathbf{u}$  and  $\mathbf{g}$  in Eq. (15) and the corresponding components of  $\mathbf{b}$  and  $\mathbf{e}$  in Eq. (8) are common. When  $d$  and  $d'$  are 0, from 10th to 17th components of  $\mathbf{u}$  are 0 and Eq. (15) is identical to Eq. (8). Therefore, it is confirmed that Eq. (8) is a special case of Eq. (15) with the condition of  $d = d' = 0$ . In other words, the condition indicates that there is no refraction, and the system of equations does not contain information regarding the absolute scale of objects. We define Eq. (15) as the ‘‘extended essential equation’’ because Eq. (15) is the equation which Eq. (8) is extended into; Eq. (15) subsumes Eq. (8).

Equation (15) is set up about each corresponding point, and consequently, the following simultaneous equation about all of the corresponding points is set up:

$$\mathbf{U} \mathbf{g} = \mathbf{0}, \quad (16)$$

where  $\mathbf{U}$  is a matrix in which row vectors  $\mathbf{u}^T$  about each corresponding point are arranged. One of the unknown parameters can be fixed because Eq. (16) is homogeneous, and therefore,  $N \geq 16$  is a necessary condition.

In actual measurements,  $\mathbf{g}$ , which satisfies Eq. (16) precisely, is not always obtained because errors can be caused by various reasons. Therefore, an approximate solution of  $\mathbf{g}$  that satisfies the following equation is obtained:

$$\|\mathbf{U} \mathbf{g}\|^2 \rightarrow \min. \quad (17)$$

Unlike  $\mathbf{e}$  in Eq. (8),  $\mathbf{g}$  has components consisting of only components of  $\mathbf{R}$ . Using relations among from the 10th to 17th components as constraints, Eq. (17) is solved with a least squares method. From the orthogonality of the rotation matrix, the constraints of Eq. (17) are written as:

$$g_{10}^2 + g_{11}^2 + g_{12}^2 = 1, \quad g_{13}^2 + g_{14}^2 + g_{15}^2 = 1, \quad (18)$$

where  $g_i$  denotes the  $i$ -th component of  $\mathbf{g}$ . The vector  $\mathbf{g}$  that minimizes  $\|\mathbf{U} \mathbf{g}\|^2$  is obtained by using the Lagrange multiplier method with the constraints and with an initial solution derived from an eigenvector belonging to the minimum eigenvalue of  $\mathbf{U}^T \mathbf{U}$ . Subsequently, the components of the rotation matrix and the translation vector are calculated.

By the definition of  $\mathbf{g}$  and the orthogonality, the estimated rotation matrix  $\mathbf{R}_{\text{est}}$  is calculated as follows:

$$\mathbf{R}_{\text{est}}(1) = [g_{10} \ g_{11} \ g_{12}], \quad (19)$$

$$\mathbf{R}_{\text{est}}(2) = [g_{13} \ g_{14} \ g_{15}], \quad (20)$$

$$\mathbf{R}_{\text{est}}(3) = \mathbf{R}_{\text{est}}(1) \times \mathbf{R}_{\text{est}}(2), \quad (21)$$

where the  $i$ -th row vector of  $\mathbf{R}_{\text{est}}$  is represented as  $\mathbf{R}_{\text{est}}(i)$ . The estimated translation vector  $\mathbf{t}_{\text{est}}$  can be obtained as:

$$t_{\text{est}3} = \frac{R_{\text{est}23} g_1 - R_{\text{est}13} g_4}{R_{\text{est}23} R_{\text{est}12} - R_{\text{est}13} R_{\text{est}22}}, \quad (22)$$

$$t_{\text{est}2} = \frac{R_{\text{est}22} g_1 - R_{\text{est}12} g_4}{R_{\text{est}12} R_{\text{est}23} - R_{\text{est}22} R_{\text{est}13}}, \quad (23)$$

$$t_{\text{est}1} = \frac{g_{10} t_{\text{est}2} - g_3}{g_{11}}, \quad (24)$$

where  $R_{\text{est} \ ij}$  denotes the  $(i,j)$ -th component of  $\mathbf{R}_{\text{est}}$ , and  $t_{\text{est} \ i}$  the  $i$ -th component of  $\mathbf{t}_{\text{est}}$ .

The parameters  $\mathbf{R}$  and  $\mathbf{t}$  are obtained as above, which means the relationship between the coordinate system  $C$  and  $C'$  has been obtained. Therefore, the 3D coordinates of the corresponding points can be obtained by using the outer ray vectors  $\mathbf{r} = (x, y, z)^T$ ,  $\mathbf{r}' = (x', y', z')^T$ , and refractive disparity vectors  $\mathbf{d} = (0, 0, d)^T$ ,  $\mathbf{d}' = (0, 0, d')^T$  belonging to each corresponding point. The intersection of two lines passing through the start points of the outer ray vectors in the directions of each outer ray vector is the 3D coordinate of a corresponding point. However, these lines do not always intersect exactly because of errors due to various reasons including numerical calculation. Therefore, the middle point of the line segment that links the two lines at the shortest

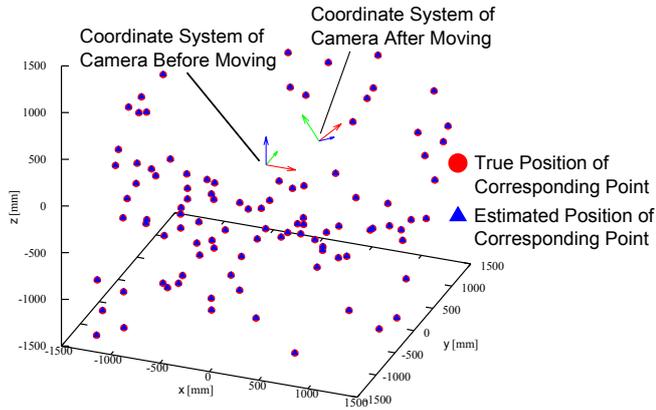


Fig. 6. Simulation result (in air,  $N = 100$ ). The red circles and the blue triangles indicate the true positions of the corresponding points and their estimated positions, respectively. The two data sets overlap because the errors between them are extremely small. The shape of the object is reconstructed with the absolute scale.

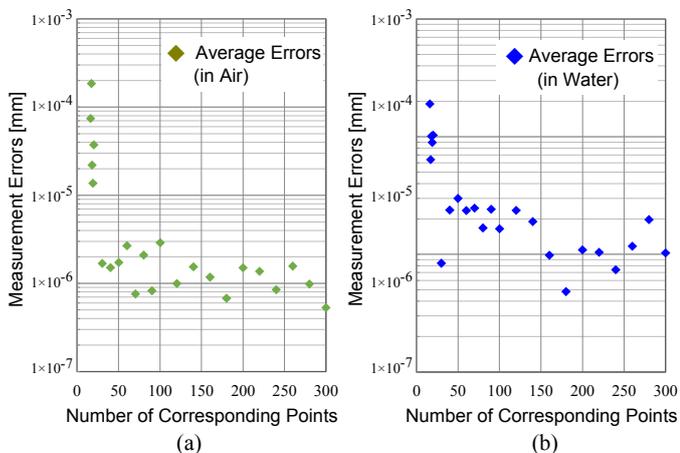


Fig. 7. The average errors in simulations in air (a) and water (b) for various values of  $N$  are plotted. The errors reduce as the number of corresponding points increases.

distance between them is regarded as the 3D coordinate of the corresponding point.

As described above, the 3D coordinate of each corresponding point surrounding the camera can be obtained, and therefore, the shape of an object is reconstructed with the absolute scale with a wide field of view.

### III. SIMULATION EXPERIMENTS

We conduct simulation experiments of measurements in air and water in order to verify that the 3D coordinates of the corresponding points can be reconstructed by using the proposed method. First,  $N$  corresponding points are arranged randomly within a distance of 1.5 m from the camera, and subsequently, the image coordinate of each corresponding point is calculated by simulating the refraction and reflection of rays. Second, based on the image coordinate obtained, the 3D coordinate of each corresponding point is estimated by using the proposed method. The errors between the estimated positions and true positions of the corresponding points are

calculated to evaluate the proposed method's performance. We apply the following simulation conditions: the thickness of the acrylic housing with refractive index of  $n_2 = 1.49$  is 3 mm, the inside of the housing is in air whose refractive index is  $n_1 = 1.00$ , and the surrounding environment is air or water whose refractive index is  $n_3 = 1.33$ .

The result of a simulation (in air,  $N = 100$ ) is shown in Fig. 6 as an example. The red circles and the blue triangles indicate the true positions of the corresponding points and their estimated positions, respectively. The two data sets exhibit an overlap, thereby indicating that the errors are extremely small. The positions of the corresponding points surrounding the camera are estimated precisely, and therefore omnidirectional 3D reconstruction with the absolute scale is achieved.

The average errors in simulations in air and water with for various numbers of corresponding points are shown in Fig. 7 (a) and (b), respectively. From the results, we observe that each 3D coordinate of the corresponding points is precisely obtained. The results also indicate that measurement errors reduce as the number of corresponding points increases. We speculate that this error is derived from numerical calculations, particularly in the stage of calculations related to refraction. In the simulations, refraction is calculated in two different ways. On the one hand, when calculating the image coordinates, a minimization problem about the optical path length is solved based on Fermat's principle. On the other hand, when the outer ray vectors are calculated from the image coordinates, Snell's law is used. These approaches differ in terms of numerical calculation although they are identical mathematically, and therefore, differences between the 3D coordinates of the true positions of the corresponding points and those of their estimated positions can be finally generated. The results show that this error is extremely small when the number of the corresponding points is large. We speculate that when the number of corresponding points is large the least squares fitting is performed more precisely, and thus, the result is significantly less affected by errors due to numerical calculations.

The result also indicates that errors in water are smaller than those in air. This is because the norms of the refractive disparity vectors in measurements in water are larger than those in air. In measurements in air, the directions of inner ray vectors and outer ray vectors are identical. In contrast, they are different in the case of measurements in water. Therefore, the norms of the refractive disparity vectors in measurements in water are larger than those in air. This fact can be confirmed by comparing the graphs shown in Fig. 8 and Fig. 9; the  $z$  and  $z'$  components of the refractive disparity vectors ( $d$  and  $d'$ ) in air and those in water are shown in Fig. 8 and Fig. 9, respectively. Figures 8 and 9 are examples of the simulation results. The maximum value of  $\|d\|$  or  $\|d'\|$  in water is 131 mm, which is considerably larger than that in air, which value is 3.12 mm. The parameters  $d$  and  $d'$  reflect the effect of refraction. In the extended essential equation (Eq. (15)), the small effect of refraction indicates that the values of the 10th to the 17th components in Eq. (15) are relatively small compared to the terms in the equation. This means that the contributions of

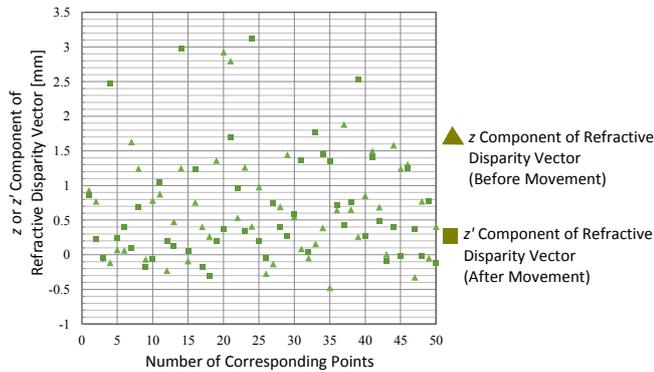


Fig. 8. Refractive disparity vectors (or to be more accurate, their  $z$  or  $z'$  components) belonging to each corresponding point in one of the simulations of measurements in air are plotted.

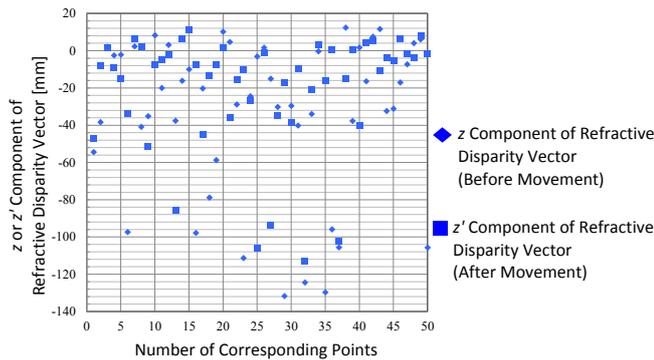


Fig. 9. Refractive disparity vectors (or to be more accurate, their  $z$  or  $z'$  components) belonging to each corresponding point in one of the simulations of measurements in water are plotted.

the 10th to the 17th components in Eq. (15) become small upon applying the least squares method. The limit of  $d$  and  $d'$  tending to 0 in the extended essential equation (Eq. (15)) yields the essential equation (Eq. (8)) in the conventional SfM in which the 3D coordinates of corresponding points cannot be obtained. Therefore, the proposed method for measurements in water is more robust against errors from numerical calculations compared to those in air. The proposed method is thereby more effective in underwater sites than in land sites.

#### IV. CONCLUSIONS

We proposed a new SfM capable of reconstructing the shapes of objects with the absolute scale using refraction with a wide field of view. First, we proposed a camera system using a cylindrical transparent housing covering an omnidirectional camera, in which refraction is caused in order to create image differences in perceived vision. The effect of refraction contains information related to the absolute scale. Second, we introduced the refractive disparity vector and extended essential equation in order to obtain the positions of corresponding points surrounding the camera and reconstruct the shape of an object with the absolute scale. Subsequently, simulation experiments assuming air and underwater environments were conducted, and the theoretical validity of our method was demonstrated.

In actual measurements, errors can be added to the image coordinates of corresponding points by axial misalignment or nonuniform thickness of a housing, and reading errors can occur. Conducting experiments in actual environments and improvement of the method's robustness against such errors form our future tasks.

#### ACKNOWLEDGMENT

This work was funded by Tough Robotics Challenge, IMPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).

#### REFERENCES

- [1] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. M. Seitz, and R. Szeliski: "Building Rome in a Day," *Communications of the ACM*, Vol. 54, No. 10, pp. 105–112, 2011.
- [2] I. Akhter, Y. Sheikh, S. Khan, and T. Kanade: "Trajectory Space: A Dual Representation for Nonrigid Structure from Motion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 7, pp. 1442–1456, 2011.
- [3] R. Kawanishi, A. Yamashita, and T. Kaneko: "Three-Dimensional Environment Model Construction from an Omnidirectional Image Sequence," *Journal of Robotics and Mechatronics*, Vol. 21, No. 5, pp. 574–582, 2009.
- [4] A. Pagani and D. Stricker: "Structure from Motion using full spherical panoramic cameras," *Proceedings of IEEE International Conference on Computer Vision Workshops*, pp. 375–382, 2011.
- [5] R. Kawanishi, A. Yamashita, T. Kaneko, and H. Asama: "Parallel Line-Based Structure from Motion by Using Omnidirectional Camera in Textureless Scene," *Advanced Robotics*, Vol. 27, No. 1, pp. 19–32, 2013.
- [6] D. H. Lee, I. S. Kweon, and R. Cipolla: "A Biprism-Stereo Camera System," *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 82–87, 1999.
- [7] A. Goshtasby and W. A. Gruver: "Design of a Single-Lens Stereo Camera System," *Pattern Recognition*, Vol. 26, No. 6, pp. 923–937, 1993.
- [8] A. Yamashita, Y. Shirane, and T. Kaneko: "Monocular Underwater Stereo -3D Measurement Using Difference of Appearance Depending on Optical Paths," *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3652–3657, 2010.
- [9] Z. Chen, K. Y. K. Wong, Y. Matsushita, and X. Zhu: "Depth from Refraction Using a Transparent Medium with Unknown Pose and Refractive Index," *International Journal of Computer Vision*, Vol. 102, No. 1–3, pp. 3–17, 2013.
- [10] A. Shibata, H. Fujii, A. Yamashita, and H. Asama: "Scale-Reconstructable Structure from Motion Using Refraction with a Single Camera," *Proceedings of the 2015 IEEE International Conference on Robotics and Automation*, pp. 5239–5244, 2015.
- [11] J. Gluckman and S. K. Nayar: "Ego-Motion and Omnidirectional Cameras," *Proceedings of the 6th IEEE International Conference on Computer Vision*, pp. 999–1005, 1998.
- [12] R. Li, H. Li, W. Zou, R. G. Smith, and T. A. Curran: "Quantitative Photogrammetric Analysis of Digital Underwater Video Imagery," *IEEE Journal of Oceanic Engineering*, Vol. 22, No. 2, pp. 364–375, 1997.
- [13] H. C. Longuet-Higgins: "A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature*, Vol. 293, pp. 133–135, 1981.
- [14] J. Weng, T. S. Huang, and N. Ahuja: "Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 5, pp. 451–476, 1989.
- [15] R. I. Hartley: "In Defense of the Eight-Point Algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 580–593, 1997.
- [16] D. Nister: "An Efficient Solution to the Five-Point Relative Pose Problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 756–770, 2004.
- [17] H. Li and R. Hartley: "Five-Point Motion Estimation Made Easy," *Proceedings of the 8th International Conference on Pattern Recognition*, pp. 630–633, 2006.
- [18] D. Batra, B. Nabbe, and M. Hebert: "An Alternative Formulation for Five Point Relative Pose Problem," *Proceedings of IEEE Workshop on Motion and Video Computing*, pp. 21–21, 2007.