# Camera Calibration for Underwater 3D Reconstruction Based on Ray Tracing using Snell's Law

Malte Pedersen, Stefan Hein Bengtson, Rikke Gade, Niels Madsen and Thomas B. Moeslund

Visual Analysis of People (VAP) Laboratory, Aalborg University

Rendsburggade 14, Aalborg, Denmark

mape@create.aau.dk, shbe@create.aau.dk, rg@create.aau.dk, nm@bio.aau.dk, tbm@create.aau.dk

## Abstract

*Accurately estimating the 3D position of underwater objects is of great interest when doing research on marine animals. An inherent problem of 3D reconstruction of underwater positions is the presence of refraction which invalidates the assumption of a single viewpoint. Three ways of performing 3D reconstruction on underwater objects are compared in this work: an approach relying solely on in-air camera calibration, an approach with the camera calibration performed under water and an approach based on ray tracing with Snell's law. As expected, the in-air camera calibration showed to be the most inaccurate as it does not take refraction into account. The precision of the estimated 3D positions based on the underwater camera calibration and the ray tracing based approach were, on the other hand, almost identical. However, the ray tracing based approach is found to be advantageous as it is far more flexible in terms of the calibration procedure due to the decoupling of the intrinsic and extrinsic camera parameters.*

## 1. Introduction

Research on marine animals is becoming increasingly popular in terms of studying pharmacology, genetics and marine ecosystems. Environmental studies being especially popular due to the increased focus on how emission of various pollutants, such as microplastics [2, 5], may affect the environment and us.

One approach to study the impact of marine pollution is to expose a model organism to a pollutant in a controlled environment and analyze the behavioral patterns before and after the exposure.

Mapping these behavioral patterns is a time-consuming process and has therefore inspired multiple vision based systems, which automate this process to some degree [16, 18]. However, a large part of these systems only supports tracking of animals in a single plane; i.e., shallow water

when working with marine animals. This is problematic as most marine animals naturally move in three dimensions and studying them in a single plane is insufficient in most cases [13].

Gathering 3D information about marine animals is not without difficulties in vision based systems as the light captured by the cameras will be exposed to refraction as it passes through different media. This is especially true when observing a fish tank with cameras placed outside the tank, which is a common setup as it imposes less restrictions on the placement of the cameras and their resistance to water.

The refraction occurring at the media interfaces, such as the aquarium boundaries, can be described using Snell's law [20] illustrated in Figure 1. It relates the angles of incidence and refraction of the light ray by

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2} = \frac{n_2}{n_1}, \tag{1}$$

where $\theta$ is the angle between the surface normal and light ray, $v$ is the velocity of light, and $n$ is the refractive index of the respective medium.
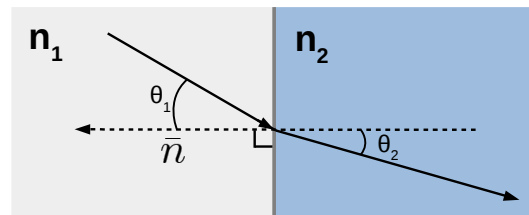


Figure 1: Illustration of the relationship between the angles of incidence and refraction described by Snell's law.

How refraction can affect a camera is illustrated in Figure 2, where multiple light rays travel from water to air until they are captured by the camera lens. The refraction essentially causes the camera to experience the scene as if it was observed from $n$ viewpoints, $vp_n$, and thereby invalidates the assumption of a single viewpoint which is prevalent in

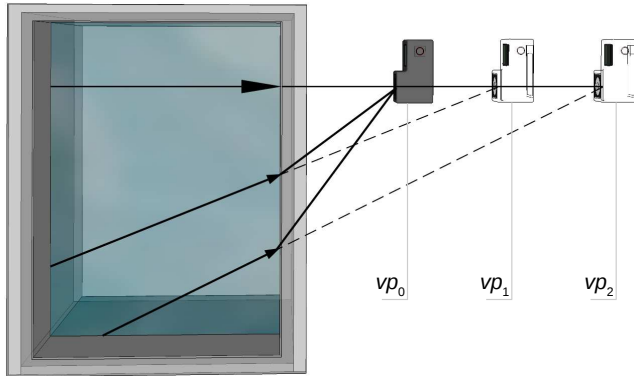most common camera models, such as the one described in [22].



Figure 2: Illustration of how light rays are bent due to refraction before entering the aperture of the camera. The transparent cameras represent virtual viewpoints caused by the refraction.

The focus of this paper is to outline and compare the precision of various ways of dealing with refraction when using a stereo vision based system to gather 3D information about marine animals in aquariums.

## 1.1. Related Work

The different ways of handling refraction when gathering 3D information can generally be divided into two categories: approaches that indirectly account for refraction by relying on the camera model to absorb the errors and approaches that directly try to account for the physics of refraction. These will be discussed in more details below.

Some approaches rely solely on the SVP (single viewpoint) camera model, which is advantageous due to its ease of use, as it is well documented and widely supported in many toolboxes. Its popularity is not without reason, as it is also simple to deal with mathematically, mainly because the camera is described using a linear projective transform.

Examples of using the SVP model for underwater purposes can be found in [3, 17, 19, 4]. The latter also tries to account for the problem of refraction by mounting a dome port instead of a flat port in front of the camera. Doing so essentially focuses the refracted rays into a single viewpoint. The drawback of such a solution is the precision required to manufacture the dome and align it properly with the given camera.

The main problem of these approaches is that the base assumption of having a single viewpoint is violated due to the refraction, as shown in Figure 2. However, it can be argued that the error caused by refraction can be absorbed, to some extend, by the focal length adjustment and radial distortion correction commonly found in SVP models.

This idea is the scope of [10], which explores the performance of the SVP camera model for 3D underwater reconstruction. The paper describes tests of different configurations of the SVP model with and without focal length adjustment and radial distortion correction. It is found that the adjustment of the focal length has the biggest impact on reducing the error caused by refraction.

A variation of this idea is to use multiple localized focal length adjustments instead of a single global focal length adjustment. An example of such an approach can be found in [11] which utilizes a pixel-wise varifocal camera model. Another example is the method presented in [12] which segments the scene into smaller partitions and calibrates localized SVP camera models for each partition.

The second group of approaches are the ones which actively seek to counteract the refraction error by modelling the underlying physics.

A straightforward example of such an approach can be found in [15], where two cameras are used to track a single fish in 3D. The cameras are modeled using the SVP camera model and their intrinsic parameters have been found through calibration in air. The extrinsic parameters of the camera are found in relation to the corners of the aquarium.

The intrinsic and extrinsic parameters are then used to project two rays, one for each camera, into the aquarium. Knowledge of the aquarium's corners are used to calculate the intersection between the rays and the aquarium. The refracted rays originating from the intersections are calculated and used for triangulation.

A somewhat similar approach is described in [9], where multiple cameras are used to track a single seahorse in 3D. This approach relies on the same combination of ray tracing, Snell's law, and the SVP camera model. However, it differs in the sense, that it includes an additional step where the intrinsic and extrinsic camera parameters are optimized. The estimated position and orientation of the refractive interface is also optimized during this step, based on reference points on a known calibration frame. The approach described in [6] performs a similar optimization step to refine the parameters used during ray tracing as well.

Other approaches discard the SVP camera model in favor of an axial camera model [1], as this model only assumes that the light ray will intersect along a common line and not in a common focal point. The use of axial camera models for marine research is however sparse, as they are deemed unpractical [19].

Two of the most frequently used approaches do hence appear to be the SVP camera model using a calibration frame placed under water [3, 17, 19] and ray tracing in combination with Snell's law [6, 9, 15].

Motivation in the respective works, as to why one approach is used instead of the other, is however lacking, as

no clear comparison has been made between the two. The SVP camera model has been popular for many years due to its simplicity and precision and it is therefore well documented and easy accessible. This is not the case with the ray tracing based approach which has not gained a lot of attention; possibly because the demand for precise 3D estimation of objects, in cases where light moves through different media, is not high. Due to this, the contributions of this paper will be

- A comparison between the approach relying on the SVP camera model with an underwater calibration frame and the approach based on ray tracing in combination with Snell's law.

  Furthermore, the impact of refraction is tested by comparing both of these approaches against the SVP camera model using a calibration frame placed in air.

- A description of each step needed to perform ray tracing in combination with Snell's law to account for refraction.

- A publicly available Python implementation of the ray tracing approach.

The rest of the paper is organized with first a thorough explanation of the ray tracing based approach, followed by a description of the evaluation process and lastly a comparison between the mentioned approaches.

## 2. Ray Tracing using Snell's Law

In this section, the ray tracing based method, to precisely estimate the 3D position of an underwater object placed in a tank of water, will be presented. All the steps are visualized in a simplified manner in Figure 4, where each number fits the number of the subsection describing the respective step.

The first part describes how to find the intrinsic and extrinsic camera parameters in a way that keeps them separated, which opens for a more flexible setup. This is followed by an outline of how ray tracing can be used in combination with Snell's law to expand the SVP model to account for refraction. By taking refraction into account using ray tracing, the system becomes unaffected by the size of the aquarium as no calibration frame is needed under water. In the end it is explained how the resulting rays can be used to estimate the 3D position of an underwater object using triangulation.

### 2.1. Camera Calibration

Calibration of the cameras is an essential part of 3D reconstruction as it relates 3D world coordinates to 2D image coordinates. This 3D to 2D relationship can be described by the extrinsic and intrinsic parameters as illustrated in Figure 3.
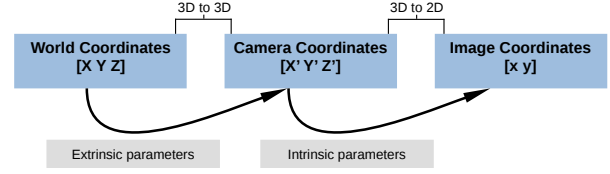


Figure 3: Illustration of the relationship between the extrinsic and intrinsic camera parameters.

The extrinsic parameters describe the transformation from world to camera coordinates and the intrinsic parameters describe the relationship between the camera coordinates and the 2D pixel coordinates in the image.

The intrinsic parameters are found for each camera using the method presented in [22], where a checkerboard in air is used for the calibration and 25 images are captured for each camera.

The extrinsic parameters are found using four 3D-2D point correspondences for each camera. The parameters are found by minimizing the reprojection error of the four points in the respective cameras, using an iterative approach based on Levenberg-Marquardt optimization [14]. This approach relies on the intrinsic parameters, as they are used in the calculation of the reprojection error, namely the 3D to 2D transformation between camera coordinates and image coordinates. The needed point correspondences are found through manual annotation of the intersections between the water surface and the corners of the aquarium.

### 2.2. Projecting a 2D Point Into a Ray

The second step is to project the 2D image coordinates into rays in the world space coordinate system. This step is repeated for each camera with each of the rays being characterized as

$$r(\lambda) = \lambda \bar{r} + r_0, \qquad (2)$$

where $\bar{r}$ is the direction of the ray, $r_0$ is a point on the ray and $\lambda$ defines all positions along the ray. The direction vector, $\bar{r}$, is found by

$$\bar{r} = R^{-1} K^{-1} \begin{bmatrix} x & y & 1 \end{bmatrix}^T, \qquad (3)$$

where $R^{-1}$ is the inverse rotation matrix of the camera, $K^{-1}$ is the inverse intrinsic camera matrix and $\begin{bmatrix} x & y \end{bmatrix}^T$ are the image coordinates of the 2D point to back project into a ray.

The point along the ray, $r_0$, is set to the camera center as the back projected ray must pass through this position, which can be found by

$$r_0 = -R^{-1} t, \qquad (4)$$

where $t$ is the translation vector of the extrinsic parameters.

## 2.3. Identifying the Plane-Ray Intersection

The point of intersection between the ray and the plane of the media interface, *e.g.*, the plane separating air and water, must be found in order to account for refraction. Determining this point is essentially a matter of identifying $\lambda_0$, such that $r(\lambda_0) = p$, where $p$ is a point on the plane. This point, $p$, must satisfy the plane equation:

$$(p - p_0) \cdot \bar{n} = 0, \tag{5}$$

where $\bar{n}$ is the plane normal and $p_0$ is an already known point on the plane. Combining Equation (2) and Equation (5) then yields

$$\lambda_0 = \frac{(p_0 - r_0) \cdot \bar{n}}{\bar{r} \cdot \bar{n}}. \tag{6}$$

The intersection point between the plane and ray, $I$, can then be found by inserting $\lambda_0$ into Equation (2)

$$I = \lambda_0 \bar{r} + r_0. \tag{7}$$

The above calculations require knowledge of the plane normal, $\bar{n}$, which can be found as $\bar{n} = \bar{v}_1 \times \bar{v}_2$. The vectors $\bar{v}_1$ and $\bar{v}_2$ being two vectors on the plane, which can be found from three non-collinear points on the plane. These three points are extracted from the set of points manually annotated during the camera calibration. The required known point on the plane, $p_0$, can be selected from these manually annotated points as well.

## 2.4. Calculating the Refracted Rays

The refraction of the ray, $r(\lambda)$, at the intersection between the media is calculated using Snell's law. The following describes the steps to calculate the refracted vector, $\bar{r}'$, of an incoming vector, $\bar{r}$, and is based on [7]:

1. Calculate the cosine of $\theta_1$ as

$$cos(\theta_1) = -\bar{n} \cdot \bar{r}, \tag{8}$$

   where $\theta_1$ is the angle between $\bar{r}$ and the surface normal of the interface between the media, $\bar{n}$.

2. Calculate the cosine of $\theta_2$ as

$$cos(\theta_2) = \sqrt{1 - \left(\frac{n_1}{n_2}\right)^2 (1 - cos(\theta_1)^2)}, \tag{9}$$

   where $\theta_2$ is the angle between $\bar{r}'$ and $\bar{n}$.

3. The refracted vector, $\bar{r}'$, can then be described as

$$\bar{r}' = \left(\frac{n_1}{n_2}\right)\bar{r} + \left(\frac{n_1}{n_2}cos(\theta_1) - cos(\theta_2)\right)\bar{n} \tag{10}$$

   where $n_1$ is the refractive index of the medium that the traced ray passes from and $n_2$ is the refractive index of the medium that the ray passes to.

For a camera placed in front of an aquarium, the indices $n_1 = 1.0$ and $n_2 = 1.33$ are used for air and water, respectively. The refraction caused by the plastic or glass sides of the aquarium is not taken into account, as the impact is minor except for large tanks where the sides are very thick.

## 2.5. Triangulation using Rays

The final step is to triangulate the 3D position of 2D image coordinates using the refracted rays, $r_1(\lambda)'$ and $r_2(\lambda)'$, from the respective cameras. The refracted rays still adheres to Equation (2) and are formed using the refracted direction vector, $\bar{r}'$, from Equation (10) along with the position of the plane-ray intersection, $I$, from Equation (7).

The triangulation method employed is commonly known as the midpoint algorithm. The idea is to identify the vector, $\bar{m}$, between the two rays, $r_1(\lambda)'$ and $r_2(\lambda)'$, such that the length, $||\bar{m}||$, is minimized. The final 3D position is found as the midpoint of the vector $\bar{m}$.

Other ways of triangulating 3D positions exist, such as the methods mentioned in [8]. However, the midpoint algorithm is chosen over other triangulation methods, as it operates on rays by default. Other approaches utilizes the camera matrix (formed by the intrinsic and extrinsic camera parameters) to triangulate a point, while minimizing the reprojection error. However, this assumes that the pinhole camera model holds, *i.e.*, a single viewpoint exists, which is not the case as discussed earlier.

The midpoint algorithm is based on the fact that the length, $||\bar{m}||$, must be at its minimum when $\bar{m}$ is perpendicular to both rays. The main idea is hence to identify a vector, $\bar{m}$, such that

$$\begin{aligned} \bar{m} \cdot \bar{r}_1' &= 0 \\ \bar{m} \cdot \bar{r}_2' &= 0, \end{aligned} \tag{11}$$

where $\bar{r}_1'$ and $\bar{r}_2'$ are the direction vectors of the two refracted rays and $\cdot$ is the dot product.

The vector, $\bar{m}$, is found by calculating the vector's start position, $M_1$, and end position, $M_2$, along the refracted rays. The two points can be calculated by

$$M_1 = I_1 + \bar{r}_1'\frac{-(\bar{r}_1' \cdot \bar{r}_2')(\bar{r}_2' \cdot \overline{I_1 I_2}) + (\bar{r}_1' \cdot \overline{I_1 I_2})(\bar{r}_2' \cdot \bar{r}_2')}{(\bar{r}_1' \cdot \bar{r}_1')(\bar{r}_2' \cdot \bar{r}_2') - (\bar{r}_1' \cdot \bar{r}_2')(\bar{r}_1' \cdot \bar{r}_2')} \tag{12}$$

and

$$M_2 = I_2 + \bar{r}_2'\frac{(\bar{r}_1' \cdot \bar{r}_2')(\bar{r}_1' \cdot \overline{I_1 I_2}) - (\bar{r}_2' \cdot \overline{I_1 I_2})(\bar{r}_1' \cdot \bar{r}_1')}{(\bar{r}_1' \cdot \bar{r}_1')(\bar{r}_2' \cdot \bar{r}_2') - (\bar{r}_1' \cdot \bar{r}_2')(\bar{r}_1' \cdot \bar{r}_2')}. \tag{13}$$

where $I_1$ and $I_2$ are the plane-ray intersections of the two views.

The final 3D position, $P$, of the triangulation process, is then calculated as
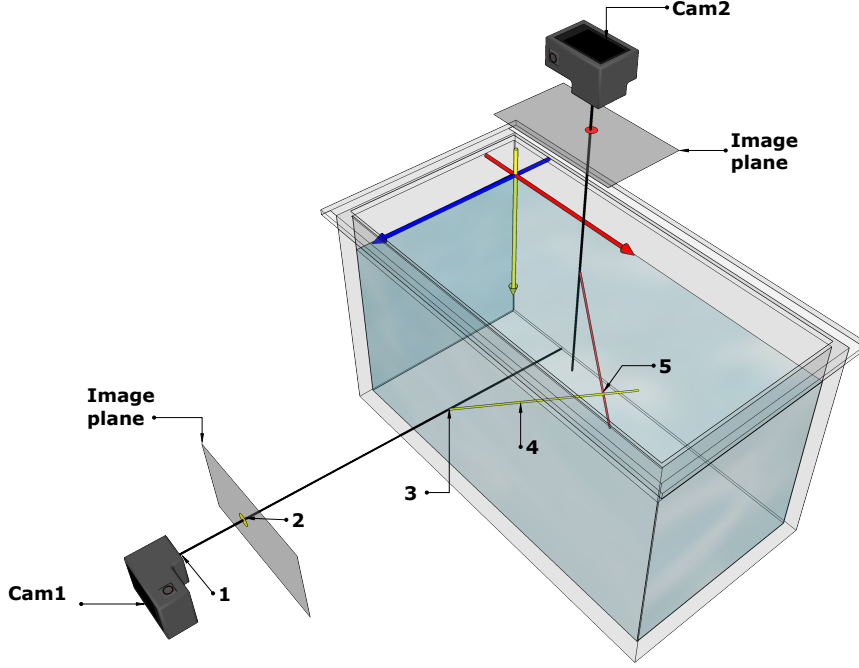
$$P = \frac{(M_1 + M_2)}{2}. \tag{14}$$

Figure 4: Illustration of the five 3D reconstruction steps. **1.** Camera calibration **2.** Projecting a 2D point into a ray **3.** Identifying the plane-ray intersection **4.** Calculating the refracted rays **5.** Triangulation using rays.

## 3. Evaluation

The described ray tracing approach is compared against two other approaches: calibration using a checkerboard placed in air and calibration using a checkerboard placed under water in the aquarium. Both methods are based on the calibration approach described in [22].

The test is performed by moving a calibration rod, with two brightly colored balls mounted on it, around in an aquarium with water. The rod has a bend near the location of the balls in order to make it easier to move it around without interfering with the cameras, *i.e.*, blocking their field of view or moving them. The cameras are placed as shown in Figure 5, which is a setup that has been used in several marine life behavioral analysis systems based on stereo vision [9, 15, 17].

The length of the aquarium is 40 cm and it has a width and height of 20 cm and 25 cm, respectively. A total of 15.6 liters of water, equalling a depth of 19.5 cm, has been used for the test.

The colored balls are used as markers and all three approaches have been evaluated by their euclidean inter-distance error, $e_i$, given by

$$e_i = d_w - |p'_1 - p'_2|, \qquad (15)$$

where $d_w$ is the known distance between the two markers, while $|p'_1 - p'_2|$ is the euclidean distance between the mark-

ers when their 3D positions have been estimated. This error will hence provide a measure of the ability of each approach to correctly judge distances while capturing underwater objects.

The actual triangulation steps remain unchanged throughout the test and are as described in Section 2.5. The main difference between the tested approaches is hence how the camera calibration is performed and how the rays for the triangulation process are found.

The positions of the two markers are automatically found in the recordings using color thresholding and Hough Circle Transform [21]. The recording used for the test consists of 2809 frames where a total of 2739 pair-wise marker detections were found. Care has been taken to move the calibration rod around the entire volume of the aquarium as it is expected that the inter-distance error can depend on the location of the objects for some of the approaches. Inaccuracies may be introduced due to the automatic detection of the markers, however, as the same detections are used for testing all three approaches it is assumed that it does not have any significance.

A refractive index of $n_1 = 1.0$ for air and $n_2 = 1.33$ for water is used when calculating the refracted rays in the test. Due to the thickness of the aquarium plastic being only 3mm the refraction caused by the plastic is deemed insignificant and therefore ignored.
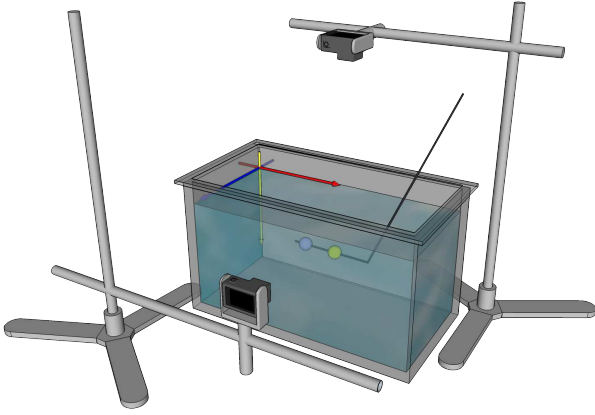
Figure 5: Illustration of the test setup and calibration rod with the two colored balls. The setup is used during evaluation of the different approaches. The red, blue, and yellow axis denotes the x-, y-, and z-axis, respectively.

## 3.1. Results

Histograms of the inter-distance error for the three approaches are shown in Figure 6. The errors appear to be normally distributed for all three, with the ray tracing based approach performing the best. The results have been gathered in Table 1.

| | In air | Under water | Ray tracing |
|---|---|---|---|
| Mean error ($\mu$) | 0.50 cm | 0.03 cm | 0.01 cm |
| Deviation ($\sigma$) | 0.18 cm | 0.14 cm | 0.09 cm |

Table 1: Results of the 3D reconstruction test.

The distribution of the inter-distance error with regard to the location of the markers is illustrated in Figure 7. Each of the 2739 dots represents the centerpoint between the reconstructed positions of the two markers in the aquarium and the color represents the magnitude of the error, *i.e.*, the absolute inter-distance error.

## 4. Discussion

The in-air camera calibration did not perform well in the conducted 3D reconstruction test, which is not surprising as it has no way of accounting for refraction. This is made apparent in the results as the inter-distance error varies to a large degree based on the location of the markers relative to the cameras, as seen in Figure 7.

The same is true for the underwater camera calibration approach, but to a lower extend. The degree of deviation of this approach rely on how well the calibration frame, in this case the checkerboard, covers the space of the container. In order to cover the entire container, the checkerboard must

be moved around in the aquarium and this may cause disturbances in the water surface that can lead to a further decrease in accuracy and thereby makes it prone to errors.

The only approach that seems independent of the location of the markers is the ray tracing based method, with the exception of a few outliers, which is most likely caused by noisy detections such as reflections.

The difference in mean error between the ray tracing based approach and the underwater camera calibration may be attributed to the precision of the measurement of the actual distance between the two markers on the calibration rod and the precision of the detections. These results are hence deemed insufficient to recommend the ray tracing approach over the underwater calibration approach, and vice versa.

The thing that really separates the two approaches is the degree of flexibility, where the ray tracing approach is superior to the under water calibration approach. This is mainly due to the decoupling of the extrinsic and intrinsic parameters, as it allows for a one-time only calibration of the intrinsic parameters. In other words, if the cameras are moved or the aquarium is replaced with another, it is only the extrinsic parameters that need to be estimated, which is easily done by manually annotating the four aquarium corners.

On the other hand, an entire re-calibration with a checkerboard is required for the underwater calibration method if a single feature is changed in the setup. This quickly becomes tedious and time consuming and may, furthermore, create issues with repeatability. This could to some extend be circumvented by replacing the checkerboard with a 3D frame with markers. However, such a frame would have to be manufactured to the specific tank making the setup even less flexible and impractical for larger setups.

To sum up the findings, the pros and cons of the tested approaches are presented in Table 2. Precision describes how well the respective approach is capable of reconstructing 3D positions independent of their location in the aquarium. Flexibility is an expression of its capabilities to adapt to variations in the setup, such as change in water levels, camera positions or aquarium size. Lastly, ease of use is an indication for how effortless the calibration of the cameras can be done.

| | In air | Under water | Ray tracing |
|---|---|---|---|
| Precision | ✓ | ✓✓✓ | ✓✓✓ |
| Flexibility | ✓ | ✓ | ✓✓✓ |
| Ease of use | ✓✓✓ | ✓ | ✓✓ |

Table 2: Simplified recap and comparison of the properties of the tested 3D reconstruction approaches.

## 5. Conclusion

The commonly used SVP (single viewpoint) camera model is not applicable when capturing underwater objects due to the refraction of light at the interface between air and water which invalidates the assumption of a single viewpoint. This observation has been confirmed during the test with the in-air camera calibration where the 3D reconstruction of underwater positions showed to be imprecise when refraction was not taken into account at all.

If the SVP camera model is calibrated with a checkerboard, placed under water, a higher precision is obtained as the intrinsic camera parameters can offset the error to some extend. However, this is problematic as even minor changes in the test setup will require the entire system to be re-calibrated, making this approach tedious to use.

Another approach is to use ray tracing in combination with Snell's law to model how the light is refracted and counteract the effect. Tests showed that this approach achieved the lowest mean error and deviation when estimating the position of underwater objects. Furthermore, it has a less restrictive calibration procedure, which makes it less prone to errors and more flexible than the other approaches, why it is recommended to use within the field of automated behavioral analysis of marine animals.

Future work could be to extend the ray tracing method to account for the refraction occurring due to light passing through the material of the aquarium. One solution could be to perform the refraction calculation twice if both the thickness and refractive index of the material are known.

A Python implementation of the ray tracing based method can be found at `https://bitbucket.org/aauvap/underwater-camera-calibration`.

## References

[1] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[2] A. L. Andrady. Microplastics in the marine environment. *Marine Pollution Bulletin*, 62(8):1596–1605, 2011.

[3] G. Bianco, M. T. Ekvall, J. Bckman, and L.-A. Hansson. Plankton 3d tracking: the importance of camera calibration in stereo computer vision systems. *Limnology and Oceanography: Methods*, 11(5):278–286, 2013.

[4] G. Bianco, A. Gallo, F. Bruno, and M. Muzzupappa. A comparative analysis between active and passive techniques for underwater 3d reconstruction of close-range objects. *Sensors*, 13(8):11007–11031, 2013.

[5] M. A. Browne, T. Galloway, and R. Thompson. Microplastic an emerging contaminant of potential concern? *Integrated Environmental Assessment and Management*, 2007.

[6] P. D. V. Buschinelli, G. Matos, T. Pinto, and A. Albertazzi. Underwater 3d shape measurement using inverse triangulation through two flat refractive surfaces. In *OCEANS 2016 MTS/IEEE Monterey*, pages 1–7, 2016.

[7] A. S. Glassner, editor. *An Introduction to Ray Tracing*. Academic Press Ltd., London, UK, 1989.

[8] R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997.

[9] S. Henrion, C. W. Spoor, R. P. M. Pieters, U. K. Müller, and J. L. van Leeuwen. Refraction corrected calibration for aquatic locomotion research: application of snell's law improves spatial accuracy. *Bioinspiration and Biomimetics*, 10(4), 2015.

[10] L. Kang, L. Wu, and Y.-H. Yang. Experimental study of the influence of refraction on underwater three-dimensional reconstruction using the svp camera model. *Appl. Opt.*, 51(31):7591–7603, 2012.

[11] R. Kawahara, S. Nobuhara, and T. Matsuyama. Dynamic 3d capture of swimming fish by underwater active stereo. *Methods in Oceanography*, 17:118 – 137, 2016.

[12] Y. Kwon and J. B. Casebolt. Effects of light refraction on the accuracy of camera calibration and reconstruction in underwater motion analysis. *Sports Biomechanics*, 2006.

[13] S. Macrí, D. Neri, T. Ruberto, V. Mwaffo, S. Butail, and M. Porfiri. Three-dimensional scoring of zebrafish behavior unveils biological phenomena hidden by two-dimensional analyses. *Nature*, 2017.

[14] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[15] K. Müller, J. Schlemper, L. Kuhnert, and K. D. Kuhnert. Calibration and 3d ground truth data generation with orthogonal camera-setup and refraction compensation for aquaria in real-time. In *International Conference on Computer Vision Theory and Applications*, volume 3, pages 626–634, 2014.

[16] A. Perez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. de Polavieja. idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, 11:743 – 748, 2014.

[17] Z.-M. Qian and Y. Q. Chen. Feature point based 3d tracking of multiple fish from multi-view images. *PLOS ONE*, 12(6):1–18, 2017.

[18] Z.-M. Qian, X. E. Cheng, and Y. Q. Chen. Automatically detect and track multiple fish swimming in shallow water with frequent occlusion. *PLOS ONE*, 9(9):1–12, 2014.

[19] E. Simetti, F. Wanderlingh, S. Torelli, M. Bibuli, A. Odetti, G. Bruzzone, D. L. Rizzini, J. Aleotti, G. Palli, L. Moriello, and U. Scarcia. Autonomous underwater intervention: Experimental results of the maris project. *IEEE Journal of Oceanic Engineering*, PP(99):1–20, 2017.

[20] 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. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1578–1583, 2003.

[21] H. K. Yuen, J. Princen, J. Dlingworth, and J. Kittler. A comparative study of hough transform methods for circle finding. In *Proceedings of the Alvey Vision Conference*, 1989.

[22] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
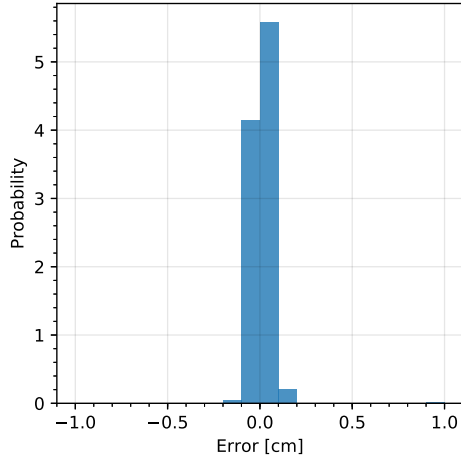
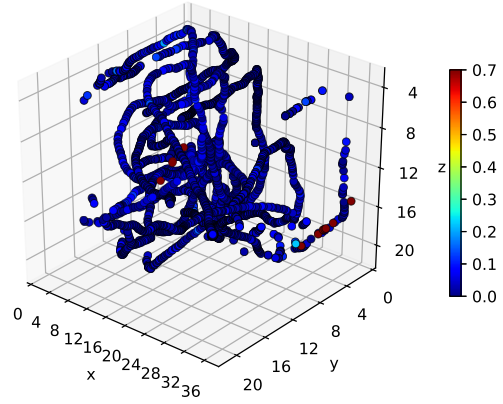(a) Ray tracing approach. ($\mu = 0.01$ cm, $\sigma = 0.09$ cm)
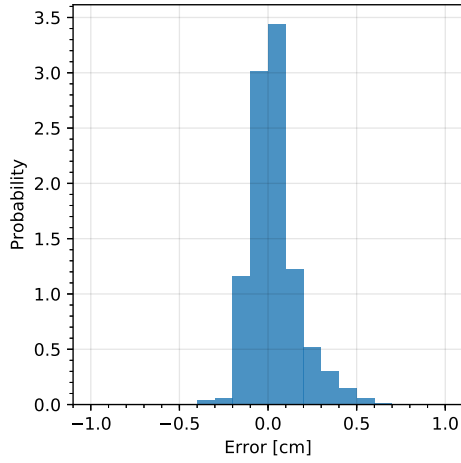


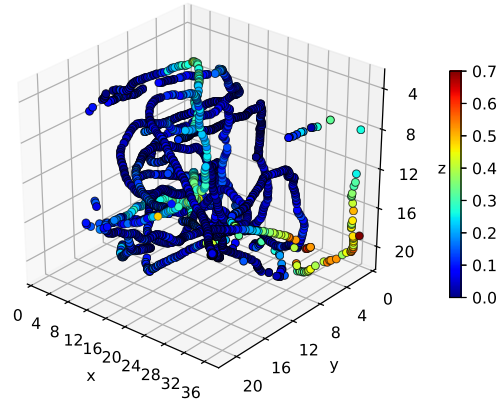(b) Underwater calibration. ($\mu = 0.03$ cm, $\sigma = 0.14$ cm)
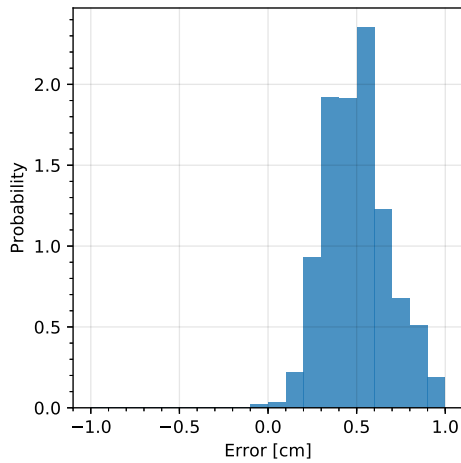


(c) In-air calibration. ($\mu = 0.50$ cm, $\sigma = 0.18$ cm)
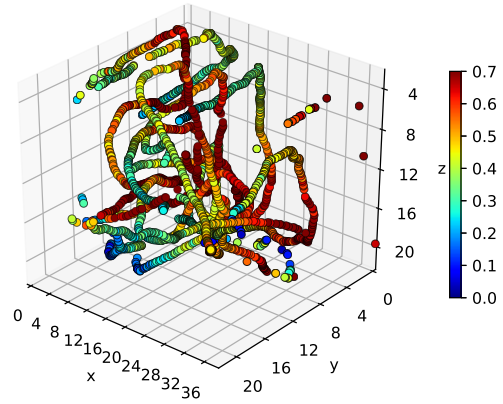
Figure 6: Histogram of the inter-distance error for each of the approaches.



(a) Ray tracing approach.



(b) Underwater calibration.



(c) In-air calibration.

Figure 7: Plot of the inter-distance error for each of the approaches. The colorbars depict the absolute inter-distance errors in centimeters.