

Article

Underwater Camera Calibration Based on Double Refraction

Yushan Sun, Tian Zhou *, Liwen Zhang  and Puxin Chai

Science and Technology on Underwater Vehicle Laboratory, Harbin Engineering University, Harbin 150001, China; sunyushan@hrbeu.edu.cn (Y.S.); zhangliwenhrbeu@163.com (L.Z.); chaipuxin@hrbeu.edu.cn (P.C.)

* Correspondence: zhout@hrbeu.edu.cn

Abstract: Underwater camera calibration plays a pivotal role in underwater positioning and underwater mapping reconstruction, making it crucial for achieving precise spatial measurements in underwater environments. To elevate measurement accuracy, we have refined the calibration methodology for underwater cameras. Firstly, we conducted an in-depth investigation into the intricate challenges posed by double refraction errors arising from light passing through water–glass–air interfaces. To address this issue, we established a double refraction model based on the actual underwater light paths, laying the groundwork for our calibration efforts. Furthermore, to tackle the problem of determining multiple camera parameters, we proposed a parameter optimization method based on genetic algorithms, capable of navigating the complex parameter space. Finally, compared with other algorithms, our method enables more precise determination of underwater camera parameters.

Keywords: camera calibration; underwater camera; double refraction

1. Introduction

Computer vision has been widely applied across various domains, including the medical field [1], industrial sector [2], military applications, and beyond. In recent times, with advancements in underwater equipment, such as autonomous underwater vehicles (AUVs) [3], remotely operated vehicles (ROVs) [4], etc., underwater vision has become an indispensable component of underwater exploration [5], aquaculture [6], underwater archaeology [7,8], and various other applications such as underwater target recognition [9,10], underwater target localization [11], underwater object reconstruction [12], and underwater simultaneous localization and mapping (SLAM) [13]. In these applications, the accuracy of camera calibration is paramount for tasks like underwater positioning and map reconstruction.

To establish the relationship between the three-dimensional geometric position of a point on the surface of an object and its corresponding point in the image, it is necessary to build a geometric model of camera imaging. The process of solving these parameters is called camera calibration. Due to the unique characteristics of the underwater environment, underwater cameras typically consist of camera components, a closed shell, and a glass cover, causing the underwater light path to pass through three media: water, glass, and air. This inevitably leads to refraction, which introduces errors in subsequent measurement processing.

In previous studies, some researchers [14,15] directly overlooked the influence of refraction and applied calibration algorithms designed for air environments [16,17] directly in underwater settings, inevitably leading to errors. Some scholars addressed the refraction effect by approximating it as changes in focal length and distortion: Lavest [18] suggests that if calibration methods designed for air are directly applied to underwater environments, the focal length needs to be multiplied by the water index; and Qin [19] utilized advanced camera distortion models to compensate for deviations caused by light refraction through water and air media. Li [20] equivalently treats the underwater camera model as



Citation: Sun, Y.; Zhou, T.; Zhang, L.; Chai, P. Underwater Camera Calibration Based on Double Refraction. *J. Mar. Sci. Eng.* **2024**, *12*, 842. <https://doi.org/10.3390/jmse12050842>

Academic Editor: Mohamed Benbouzid

Received: 17 April 2024

Revised: 15 May 2024

Accepted: 17 May 2024

Published: 19 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

a pinhole model with changes in focal length, and provides the corresponding process for calculating camera parameters. Sun [21] proposed a rapid underwater calibration method based on two sets of mutually orthogonal parallel lines, optimizing the vanishing point to determine its focal length, which is based on an air–water equivalent focal length model. Bashar [22] proposed the theory of refraction invariance, dividing the calculation of imaging parameters into refractive and non-refractive related sets, and utilizing an approximate linear model to compute camera parameters. In such algorithms, due to the complexity of simulating refraction, there are issues with low calibration accuracy. Fortunately, recently some scholars have conducted research on underwater cameras based on real light path models: Treibitz [23] derived forward projection based on Snell’s Law, and calibrated the distance between the camera optical centre and the nearest interface, and the intrinsic parameters and distortion of the camera need to be calculated through prior calibration in air. Agrawal [24] proposed an axial camera model based on flat refraction constraints and coplanarity constraints, and presented corresponding methods to solve the high-order forward projection equations. Gu [25] formulated the refractive projection function and calibrated the interface parameters with a quaternion-based optimal solution. However, due to the complexity of parameters in the real underwater camera model, there are the following issues: 1. parameter-solving conditions rely on strict imaging conditions; and 2. constraint adjustment cannot be accurately solved. With the advancement of research, recent studies have shown the potential of intelligent optimization algorithms in solving camera parameters. Tian [26] developed a monocular camera calibration algorithm for land-based applications using an improved particle swarm optimization algorithm. Du [27] utilized an enhanced slime mold algorithm for the calibration of underwater cameras, yet the imaging model remains based on the pinhole model.

To address these challenges, we propose a correction algorithm based on the real light path’s double refraction model and demonstrate its necessity in underwater camera usage. Furthermore, due to the need for properties such as high-pressure resistance in underwater cameras, many modern underwater camera structures are relatively integrated and difficult to disassemble and reassemble. As a result, measuring the thickness of the glass becomes relatively challenging. Our model algorithm addresses the issue of glass thickness. Subsequently, we employ a genetic algorithm to solve the parameters, incorporating various strategies during the algorithm’s iterations to overcome local optima. Experimental results demonstrate the accurate calculation of camera parameters using our algorithm, making the imaging calibration process straightforward, thus presenting a feasible method for underwater camera calibration.

For the remaining sections, we first introduce the double refraction model and derive the coordinate transformation of the real light path using our method in Section 2. In Section 3, we provide the mathematical proof of the necessity of considering double refraction in underwater camera calibration. This part discusses the impact of neglecting this factor and the resulting spatial measurement errors. In Section 4, we describe our underwater calibration algorithm in detail, including the calibration process using genetic algorithms and checkerboard pattern images. Experimental validation is presented in Section 5. Finally, in Section 6, we summarize the conclusions and propose future work.

2. Underwater Camera Model

Due to the existence of water medium in the underwater environment, to prevent the camera from entering the water and short-circuiting, the underwater camera is usually placed in a watertight and sealed pressure-resistant shell, and a transparent medium such as glass is used as a light-transmitting material. This causes the light reflected to pass through the three mediums of water–glass–air before the camera can capture it.

According to Fermat’s principle, light rays travel in a straight line in the same homogeneous medium, but they are refracted when passing through a different medium. Therefore, the camera will capture the light rays of underwater objects after two refractions. As

shown in Figure 1, we have to model the light path propagation model according to the real light path.

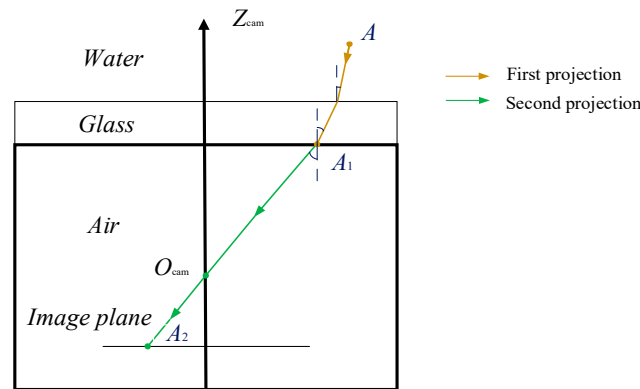


Figure 1. Schematic diagram of double refraction.

In the real light path, we propagate $A(x_G, y_G, z_G)$ in the world coordinate system to $A_2(u, v)$ in the pixel coordinate system. The method regards the process as two refractions: (1) coordinates derived from A to A_1 according to the refraction relation; and (2) coordinates derived from A_1 to A_2 according to the imaging principle.

First, we convert A from the world coordinate system (x_G, y_G, z_G) to the camera coordinate system (x_w, y_w, z_w) :

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \vec{\mathbf{T}} \\ \mathbf{0} & 1 \\ \mathbf{1} \times \mathbf{3} & \mathbf{4} \times \mathbf{4} \end{bmatrix} \begin{bmatrix} x_G \\ y_G \\ z_G \\ 1 \end{bmatrix} \tag{1}$$

where \mathbf{R} represents a 3×3 rotation transformation matrix, and its value is associated with the orientation of the camera coordinate system relative to the world coordinate system. Its specific definition is provided in Equation (2). $\vec{\mathbf{T}}$ represents a 3×1 translation vector, indicating the displacement of the camera coordinate system origin relative to the world coordinate system. Its specific definition is given in Equation (3).

$$\mathbf{R} = \begin{bmatrix} \cos \psi \cos \theta & \cos \psi \sin \theta \sin \phi - \sin \psi \cos \phi & \cos \psi \sin \theta \sin \phi + \sin \psi \sin \phi \\ \sin \psi \cos \theta & \sin \psi \sin \theta \sin \phi + \cos \psi \cos \phi & \sin \psi \sin \theta \sin \phi - \cos \psi \sin \phi \\ -\sin \theta & \cos \theta \sin \phi & \cos \theta \cos \phi \end{bmatrix} \tag{2}$$

$$\vec{\mathbf{T}} = \begin{bmatrix} x_O \\ y_O \\ z_O \end{bmatrix} \tag{3}$$

where ϕ, ψ, θ represents the rotation angle of the camera coordinate system relative to the world coordinate system; and (x_O, y_O, z_O) represents the position of the camera’s optical center in the world coordinate system.

$A \rightarrow A_1$: in the camera coordinate system, let the distance between a point A and the optical axis be r_w , and the distance between a point A_1 and the optical axis be r_a , and meet the following requirements:

$$\begin{cases} r_w^2 = x_w^2 + y_w^2 \\ r_a^2 = x_a^2 + y_a^2 \\ \frac{y_w}{x_w} = \frac{y_a}{x_a} \\ \sin \theta_a = \frac{r_a}{\sqrt{r_a^2 + h^2}} \end{cases} \tag{4}$$

The relation between point A_1 and point A can be obtained by geometric knowledge:

$$r_w = r_a + t \tan \theta_g + (z - t - h) \tan \theta_w \tag{5}$$

According to Snell’s law of refraction:

$$n_w \sin \theta_w = n_g \sin \theta_g = n_a \sin \theta_a \tag{6}$$

where h is the distance between the optical center and the lower surface of the glass; t is the thickness of glass; z is the depth of the target point in camera coordinates; n_w , n_g , and n_a represent the refractive index of water, glass, and air, respectively; θ_g is the refractive angle of light rays in the glass; θ_w is the incident angle of the ray in water; and θ_a is the refractive angle of the ray in air, and its value is $\arctan \frac{r}{f}$.

For a triangle, we can easily obtain the following equations:

$$\tan \theta = \frac{\sin \theta}{\sqrt{1 - \sin^2 \theta}} \tag{7}$$

At 20 °C and pressure 1 atmosphere, the refractive index of air is 1.00027, which we approximate to 1. Then, the following equation can be deduced from Equations (6) and (7):

$$\begin{cases} \tan \theta_g = \frac{\sin \theta_a}{\sqrt{n_g^2 - \sin^2 \theta_a}} \\ \tan \theta_w = \frac{\sin \theta_a}{\sqrt{n_w^2 - \sin^2 \theta_a}} \end{cases} \tag{8}$$

By incorporating Equations (4) and (5), we obtain:

$$\begin{aligned} r_w &= r_a + t \frac{\sin \theta_a}{\sqrt{n_g^2 - \sin^2 \theta_a}} + (z - t - h) \frac{\sin \theta_a}{\sqrt{n_w^2 - \sin^2 \theta_a}} \\ &= r_a + \frac{tr_a}{\sqrt{n_g^2 h^2 + (n_g^2 - 1)r_a^2}} + \frac{(z - t - h)r_a}{\sqrt{n_w^2 h^2 + (n_w^2 - 1)r_a^2}} \end{aligned} \tag{9}$$

To obtain r_a according to r_w , direct derivation is not feasible. Therefore, we opt to employ Newton’s method for iteration:

$$g(r_a) = r_a + \frac{tr_a}{\sqrt{n_g^2 h^2 + (n_g^2 - 1)r_a^2}} + \frac{(z - t - h)r_a}{\sqrt{n_w^2 h^2 + (n_w^2 - 1)r_a^2}} - r_w \tag{10}$$

The zero point of function, $g(r_a)$, corresponds to the solution of Equation (9). The iterative formula of Newton’s method is as follows:

$$r_{a(n+1)} = r_{a(n)} - \frac{g(r_{a(n)})}{g'(r_{a(n)})} \quad (n = 0, 1, 2, \dots) \tag{11}$$

We initiate the iterative process by substituting the projection position of the target point, under the condition of no refraction, as the initial point $r_{a(0)} = \frac{r_w h}{z}$ into Equation (11). This iterative calculation allows us to derive the solution for Equation (9).

The proof of the iterative convergence of Equation (10) is as follows:

$$g'(r_a) = 1 + \frac{th^2 n_g}{[n_g^2 h^2 + (n_g^2 - 1)r_a^2]^{\frac{3}{2}}} - \frac{(z - t - h)h^2 n_a}{[n_a^2 h^2 + (n_a^2 - 1)r_a^2]^{\frac{3}{2}}} \tag{12}$$

In the above Equation (12), both n_a and n_g are constants greater than 0. It can be easily obtained from Equation (11): within the interval of $r_a > 0$, $g'(r_a)$ is a monotonic decreasing function, and when $r_a \rightarrow \infty$, $g'(r_a)$ is the minimum value of 1. Therefore, when $r_a > 0$, $g'(r_a) \neq 0$ is constantly established. In the interval of $r_a > 0$, the second derivative of $g(r_a)$

to r_a exists and is continuous, according to the local convergence theorem of Newton iteration, for the initial iteration point, $r_a(0)$, close to the true value, r_a^* , the sequence $\{r_a(k)\}$ generated by Equation (11) always converges to r_a^* at a rate of no less than second order.

After obtaining r_a , the coordinate of A_1 in the camera coordinate system (x_a, y_a, h) can be obtained according to Equation (4).

$A_1 \rightarrow A_2$: the coordinate (x_{pic}, y_{pic}) in the imaging plane coordinate system can be obtained from the pinhole imaging principle:

$$h \begin{bmatrix} x_{pic} \\ y_{pic} \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_a \\ y_a \\ h \\ 1 \end{bmatrix} \tag{13}$$

In actual use, we have to consider the influence of distortion. We use Equation (14) to address the distortion effects.

$$\begin{cases} x_c = x_{pic}(1 + k_1r^2 + k_2r^4 + k_3r^6) + 2p_1x_{pic}y_{pic} + p_2(r^2 + 2x_{pic}^2) \\ y_c = y_{pic}(1 + k_1r^2 + k_2r^4 + k_3r^6) + p_1(r^2 + 2y_{pic}^2) + 2p_2x_{pic}y_{pic} \end{cases} \tag{14}$$

where x_c and y_c are real position coordinates, and k_1, k_2, k_3, p_1, p_2 are distortion parameters.

Due to the fact that computers typically utilize the pixel coordinate system when processing images, the corrected coordinates are converted into coordinates within the pixel coordinate system:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} \tag{15}$$

where dx and dy are the actual physical dimensions of a single photosensitive unit in the camera along the height and width directions, respectively; and (u_0, v_0) is the position coordinate of origin O_{pic} in the planar pixel coordinate system of the imaging plane.

3. Model Necessity Proof

To prove the necessity of the model, we set the “true point”, which ignores the first refraction effect, as point A' , and the mapping position of A and A' to the image is the same as shown in Figure 2. According to the geometric knowledge, we can find the distance r_w from point A to the central optical axis.

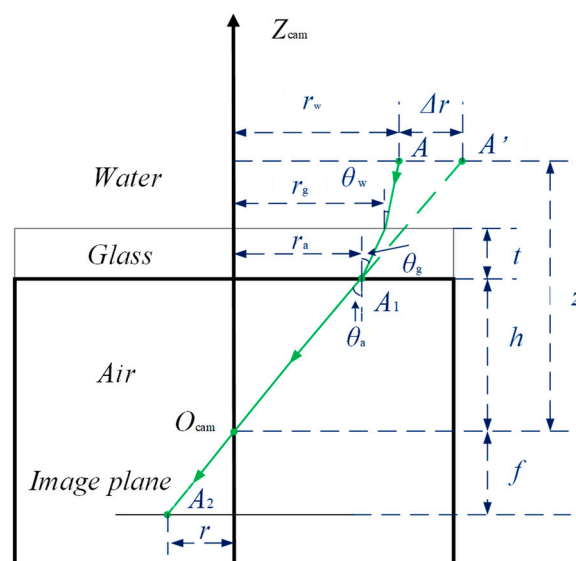


Figure 2. Underwater imaging light path.

$$r_w = \frac{rh}{f} + t \tan \theta_g + (z - t - h) \tan \theta_w \tag{16}$$

where r is the distance between the imaging point and the central optical axis in the imaging plane. Equation (17) can be obtained by substituting Equation (8) into Equation (16).

$$\begin{aligned} r_w &= \frac{rh}{f} + t \frac{\sin \theta_a}{\sqrt{n_g^2 - \sin^2 \theta_a}} + (z - t - h) \frac{\sin \theta_a}{\sqrt{n_w^2 - \sin^2 \theta_a}} \\ &= \frac{rh}{f} + \frac{tr}{\sqrt{n_g^2 f^2 + (n_g^2 - 1)r^2}} + \frac{(z - t - h)r}{\sqrt{n_w^2 f^2 + (n_w^2 - 1)r^2}} \end{aligned} \tag{17}$$

According to Snell’s law of refraction, glass in the air does not change the incident angle of the external light rays, and there is only a tiny steady-state translation. This paper ignores this small quantity, and glass does not affect the transmission of light rays when used in the air. As shown in Figure 2, we can calculate the deviation, Δr , between the target position in the air and the underwater environment based on the above derivation.

$$\begin{aligned} \Delta r &= \frac{z}{f}r - r_w \\ &= \frac{z-h}{f}r - \frac{tr}{\sqrt{n_g^2 f^2 + (n_g^2 - 1)r^2}} - \frac{(z-t-h)r}{\sqrt{n_w^2 f^2 + (n_w^2 - 1)r^2}} \end{aligned} \tag{18}$$

To see the influence of the distance of the target point from the central optical axis on Δr , we will obtain the partial derivative of Equation (18) to r :

$$\frac{\partial \Delta r}{\partial r} = \frac{z-h}{f} - \frac{tn_g^2 f^2}{[n_g^2, f^2, +, (n_g^2, -, 1), r^2]^{\frac{3}{2}}} - \frac{(z-t-h)n_w^2 f^2}{[n_w^2 f^2 + (n_w^2 - 1)r^2]^{\frac{3}{2}}} \tag{19}$$

In the above formula, t and f are constants greater than 0, n_g and n_w are constants greater than 1, and z is a variable greater than $t + h$. r is the variable greater than 0.

From Equation (19), it is easy to see that $\frac{\partial \Delta r}{\partial r}$ is the increment function of r , and the minimum value of Equation (19) is obtained when r is 0.

$$\left. \frac{\partial \Delta r}{\partial r} \right|_{\min} = \left(1 - \frac{1}{n_g}\right) \frac{t}{f} + \left(1 - \frac{1}{n_w}\right) \frac{z-h-t}{f} \tag{20}$$

In the above formula, n_g and n_w are constants greater than 1, and z is a value greater than $t + h$. It is easy to see that Equation (20) is greater than 0 and always holds.

From the nature of the increasing function, we can see that the value of Equation (20) is always greater than 0.

We know from the properties of derivatives that, for the z given value, Δr is an increment of r , and the bigger r is, the bigger Δr is. According to the principle of optical path reversible, it can be easily deduced that, when the z value is constant, the greater the distance from the target point to the central optical axis is, the greater is the distortion deviation.

To see the influence of the depth of the target point on Δr , we will obtain the partial derivative of Equation (18) with respect to z :

$$\begin{aligned} \frac{\partial \Delta r}{\partial z} &= \frac{r}{f} - \frac{r}{\sqrt{n_w^2 f^2 + (n_w^2 - 1)r^2}} \\ &= \frac{r}{f} \left(1 - \frac{1}{\sqrt{n_w^2 + \frac{(n_w^2 - 1)r^2}{f^2}}} \right) \end{aligned} \tag{21}$$

In the above Equation (21), t and f are constants greater than 0 and n_w are constants greater than 1. It is easy to see that Equation (21) is always greater than 0. We know from the property of the derivative that, at a given value of r , Δr is an increment of z , and the bigger z is, the bigger Δr is. According to the principle of optical path reversible, it can

be easily deduced that, when the r value is constant, the farther the target point is from the camera, the greater is the distortion deviation. It is necessary to study the camera imaging model in the underwater environment because the camera imaging model in the underwater environment is bound to cause the deviation of the target location.

4. Underwater Camera Calibration Based on Double Refraction

The accuracy of parameter calibration of the underwater camera-imaging model directly affects the accuracy of underwater light vision positioning. In this paper, we propose to use planar checkerboard pattern images and a genetic algorithm to calibrate the underwater camera model above and carry out calibration experiments. In the underwater camera model established above, there are parameters to be calibrated, as shown in Table 1.

Table 1. Summary table of parameters to be calibrated.

Parameter's Name	Describe
External Camera Parameters	ψ, θ, ϕ represents the rotation angle of the camera coordinate system relative to the world coordinate system; x_O, y_O, z_O represents the position of the camera's optical center in the world coordinate system.
Intrinsic Camera Parameters	(u_0, v_0) is the position coordinate of origin O_{pic} in the planar pixel coordinate system of the imaging plane; f_x, f_y are the horizontal and longitudinal digital focal lengths of the camera.
Distortion Parameters	k_1, k_2, k_3 is the radial distortion parameter; p_1, p_2 is the tangential distortion parameter.
Refraction Parameters	h is the distance from the optical centre to the lower surface of the glass; t is the thickness of the glass; n_g is the refractive index of light rays in the glass; n_w is the refractive index of light rays in the water.

Due to the speciality of the underwater environment, our calibration parameters are compared to air more out of refraction parameters, and this is our focus of calibration. The distortion parameters and intrinsic camera parameters can be calibrated in the air by Zhang's calibration [17]. For an image, the actual parameters that need to be calibrated in an underwater environment are only the six external parameters of the camera and four refraction parameters.

The calibration process of parameters is shown in Figure 3. Firstly, we use an underwater camera to take planar checkerboard pattern images in the underwater environment in the calibration process, and establish the world coordinate system on the checkerboard pattern. The distance between the corner points we can measure with a ruler, so the position of the corners in the world coordinate system is fixed, and we can obtain the coordinates of the corner points in the world coordinate system from the calibration plate. We use the external parameter matrix to transform the coordinate system to obtain the coordinates of the corners in the camera coordinate system, and use the double refraction model to predict the position of the corners on the image (expected position). At the same time, the position of the corresponding corners in the image (real position) is obtained through vision processing. In the initial state, the parameters in the camera model were inaccurate, so there was a deviation between the predicted position and the real position. We take the deviation as the optimization target, and the calibrated parameters are used as the variables to be optimized. The genetic algorithm is used to optimize the parameters continuously. The closer the parameter is to the accurate value, the smaller is the deviation between the predicted position and the true position, and accurate model parameters can be obtained after multiple iterations of calculations.

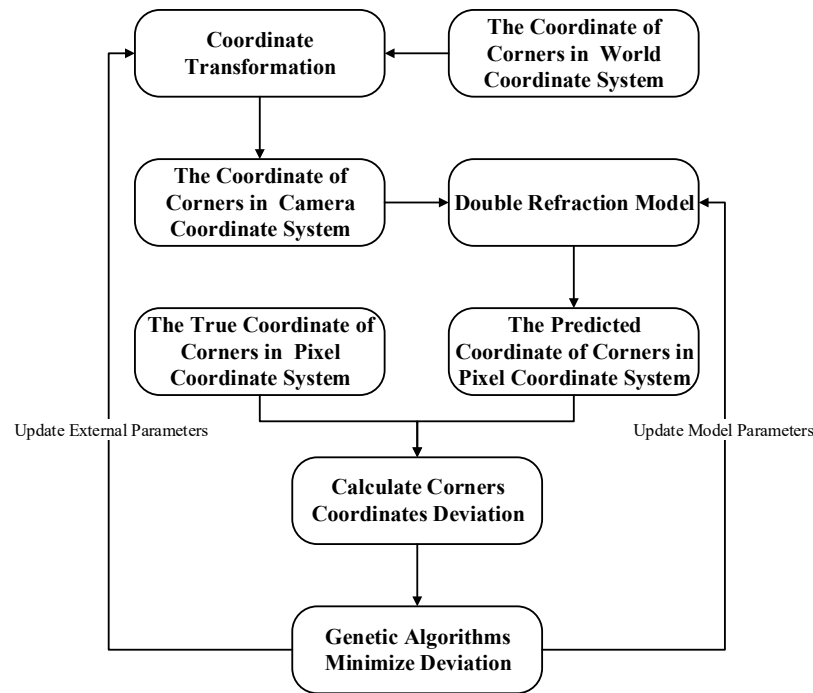


Figure 3. Parameter calibration process.

The position vector of the corners in the world coordinate system is $\mathbf{X}_G = (x_G, y_G, z_G)$; after the coordinate change in Section 2, we can obtain the position vector of the corners in the camera coordinate system: $\mathbf{X}_{cam} = (x_w, y_w, z)$; and, finally, we can obtain the position vector of the corners in the pixel coordinate system: $\mathbf{X}_{pix} = (u, v)$. Equation (22) can express this process:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{cases} F_1(\psi, \theta, \phi, x_O, y_O, z_O, h, t, n_g, n_w, \mathbf{X}_G) \\ F_2(\psi, \theta, \phi, x_O, y_O, z_O, h, t, n_g, n_w, \mathbf{X}_G) \end{cases} \quad (22)$$

We use vision processing to process images to obtain the true position of each corner: $\mathbf{X}_{pix}^* = (u^*, v^*)$. It can be seen from Equation (21) that two equations can be established for each corner, and a total of 10 parameters need to be calibrated. Therefore, in theory, only five corners' information is needed to determine the undetermined parameters. The actual calibration process, due to the existence of uncertain factors such as image noise using only five corners for calibration parameters, is inappropriate, and, therefore, the checkerboard pattern used in this article has dozens of corners, establishes an overdetermined equation set, calculates the sum of square errors of the corners according to the calculated value and the true value of the position of corners, and uses it as the evaluation function. As shown in Equation (23), we use the genetic algorithm to search for underwater camera parameters that minimize the sum of the squares of all corners' errors.

$$\varepsilon = \frac{1}{n} \sum_{i=0}^{n-1} (\|u_i - u_i^*\|^2 + \|v_i - v_i^*\|^2) \quad (23)$$

In the above formula, n represents the number of corners on the checkerboard pattern.

Optimization strategy. Since each picture has six external parameters, if m pictures are passed in, there are a total of $6m$ external parameters and 4 camera refraction parameters (camera refraction parameters are iteratively optimized in each optimization strategy), and the number of optimized parameters is $6m + 4$. There will be too many parameters, and it is easy to encounter local optimization. So here, we use three strategies to carry out, and analyze them.

1. The first strategy remains unchanged. We input ten calibration images into the aforementioned algorithm and optimize $6 \times 10 + 4$ parameters for each update.
2. The second strategy takes into account the number of parameters. With each image input, we update $6 + 4$ parameters and utilize the converged refraction parameters as the initial values for subsequent image updates.
3. The third strategy builds upon the second one. Here, we directly utilize Zhang’s [17] calibrated external parameters (without considering refraction, treating the image as if it were in the air for calibration) as the initial values for each iteration. Additionally, the external parameter $\times (1 \pm 0.5)$ serves as the upper and lower boundary values. This approach reduces the number of iterations and enhances local optimization capabilities.

5. Experiments

5.1. Experimental Data and Equipment

The experimental setup primarily comprises a custom-made underwater monocular camera (resolution: 1080 p, frame rate: 30 fps, interface: RJ45 Ethernet interface, field of view: 70°) and power supply, along with a checkerboard pattern (consisting of six rows and nine columns, totaling 54 target corners, with a 25 mm distance between corner points), and a water tank. To ensure the accuracy and reliability of the calibration results, we acquired calibration images from different distances and angles. Initially, without water in the tank, we captured images in the air to calibrate initial intrinsic parameters and distortion parameters. Subsequently, after filling the tank with water, we submerged the camera and checkerboard pattern, capturing underwater calibration images at various poses and distances.

The images captured in the air are shown in Figure 4, while the images captured in the water are shown in Figure 5. In the experiment, we utilized Python3.8.5 to handle the data and employed the OpenCV4.4.0 library for image data processing.



Figure 4. Images captured in the air (without water, used for obtaining initial parameters).

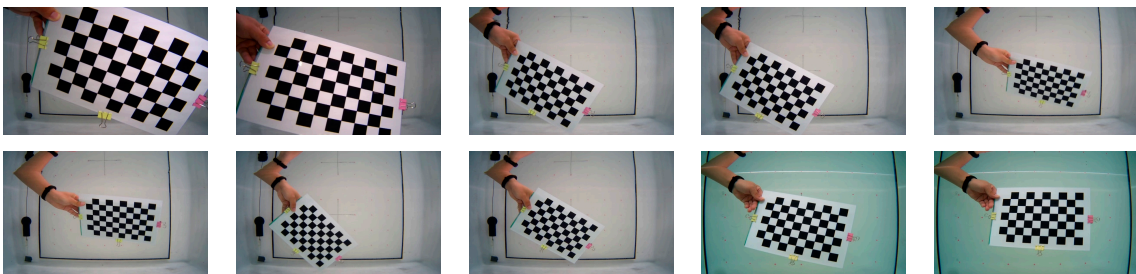


Figure 5. Images captured in the water.

5.2. Different Optimization Strategies

Firstly, regarding the choice of genetic algorithm, we adopt the elitist-reservation genetic algorithm [28], with the algorithm process shown in Table 2.

Table 2. Genetic algorithm flow chart.

1.	<p>Initialization:</p> <ul style="list-style-type: none"> • Define the objective function $f(x)$ as depicted in Equation (23) • Initialize a population of N individuals using various strategies. • Define the maximum number of iterations T_{max} and the target value F.
2.	<p>Iteration:</p> <ul style="list-style-type: none"> • If: the iteration count $< T_{max}$ or $f(x) \neq F$ Conduct statistical analysis on the current population, recording its best individual, average fitness. Independently select $N - 1$ parents from the current population. Independently perform crossover operations on these $N - 1$ parents. Independently perform mutation operations on these $N - 1$ individuals after crossover. Calculate the best individual of the current generation population and insert it at the first position among the $N - 1$ individuals after crossover to obtain the new generation population. • Else: Terminate.
3.	<p>Termination:</p> <ul style="list-style-type: none"> • End the algorithm.

In the experiment, we utilized ten underwater calibration images, as depicted in Figure 1.

The first strategy: We update the ten pictures together and set the upper and lower limits of the angle to $[-\pi, \pi]$, the upper and lower limits of the translation to $[-500, 500]$, set the population size to 100, update 2000 times, and obtain the errors. The objective function values of the population and the best individual for Strategy 1 are depicted in Figure 6, the images after reprojection are shown in Figure 7, the errors are illustrated in Figure 8, and the calibration result is shown in Table 3.

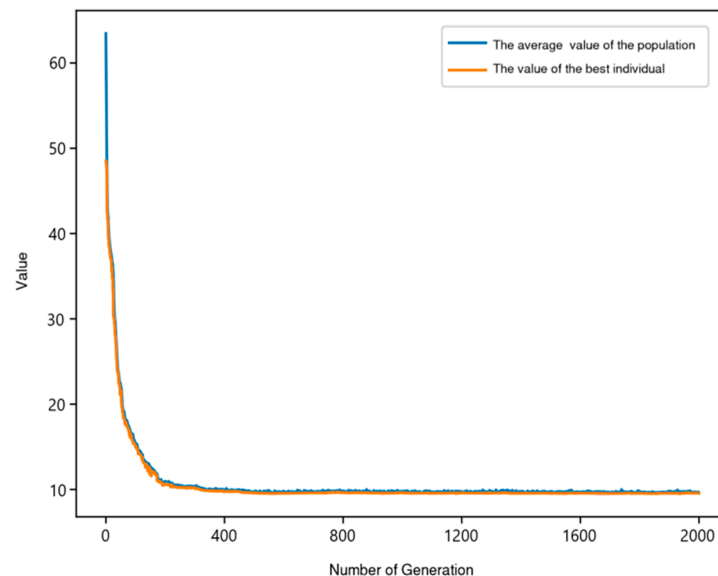


Figure 6. Population and individual information for Strategy 1.

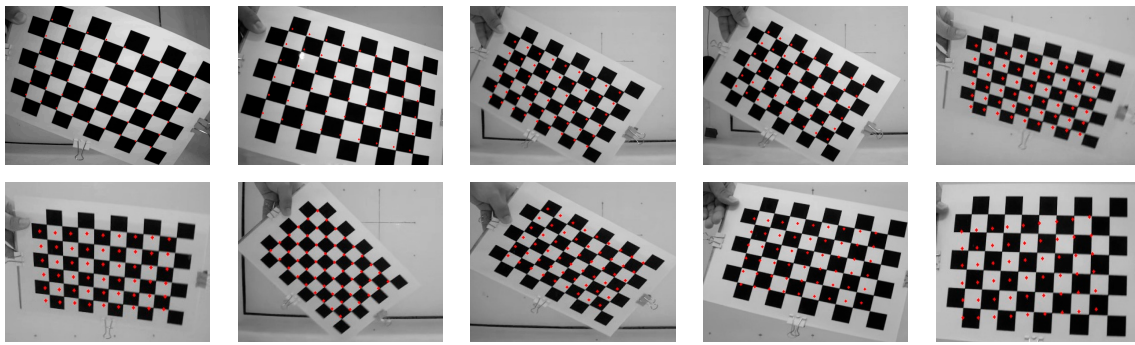


Figure 7. Re-projection images with the first strategy.

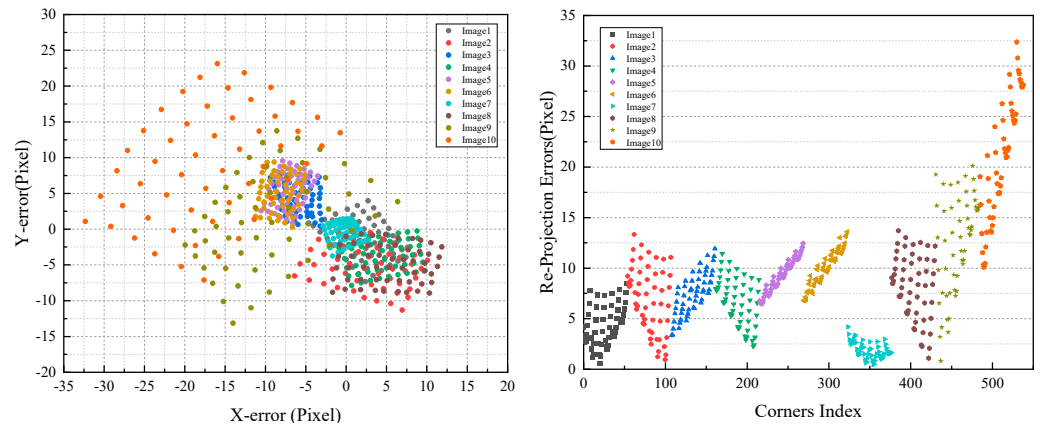


Figure 8. Error of each corner with the first strategy (on the left is the error in the X\Y direction, and on the right are the re-projection errors).

Table 3. Re-projection errors with the first strategy.

Images	Re-Projection Errors/Pixel
1	4.2749
2	6.8761
3	7.3536
4	6.5965
5	9.1818
6	9.9528
7	1.7927
8	8.0620
9	12.3986
10	20.9183
Average	8.7407

Based on our analysis, it can be inferred that the algorithm exhibits poor convergence, considerable re-projection errors, and encounters challenges associated with local optimization.

The second strategy: We pass each picture into the algorithm and pass the converged refraction parameters updated last time into the next image. We set the upper and lower limits of the angle to $[-\pi, \pi]$, the upper and lower limits of the translation to $[-500, 500]$, set the population size to 100, update 500 times, and obtain the errors. The objective function values of the population and the best individual for Strategy 2 are depicted in Figure 9, which is the result of the last image, the images after reprojection are shown in Figure 10, the errors are illustrated in Figure 11, and the calibration result is shown in Table 4.

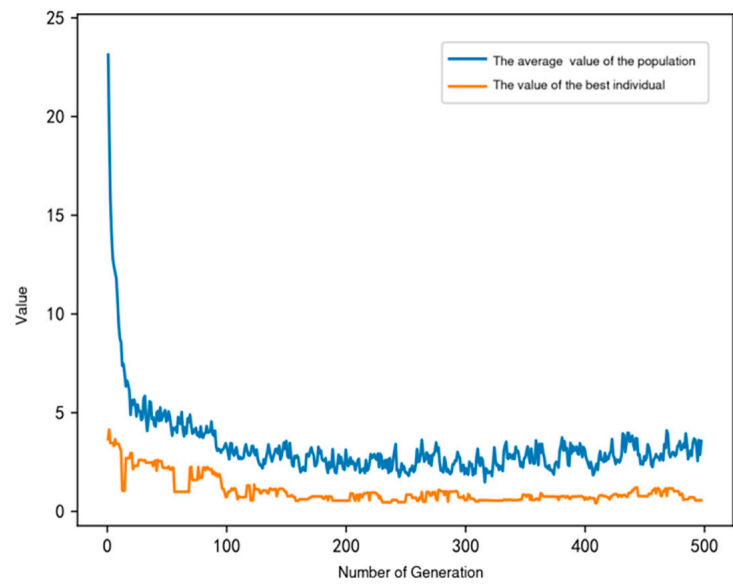


Figure 9. Population and individual information for Strategy 2.

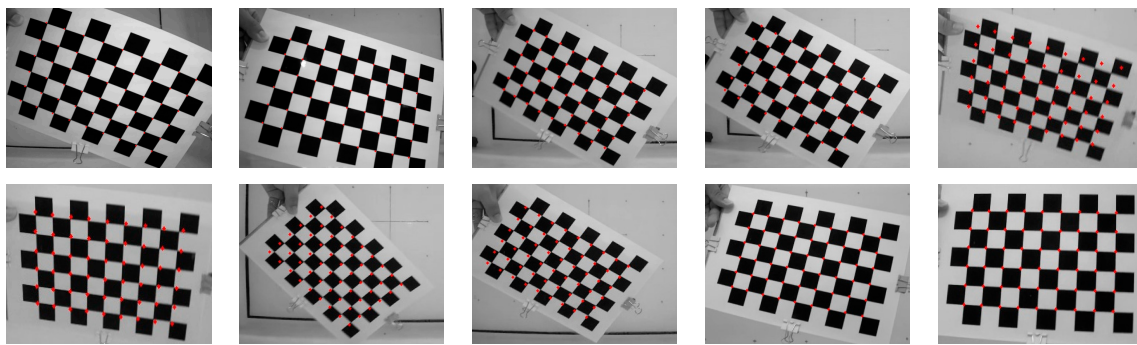


Figure 10. Re-projection images with the second strategy.

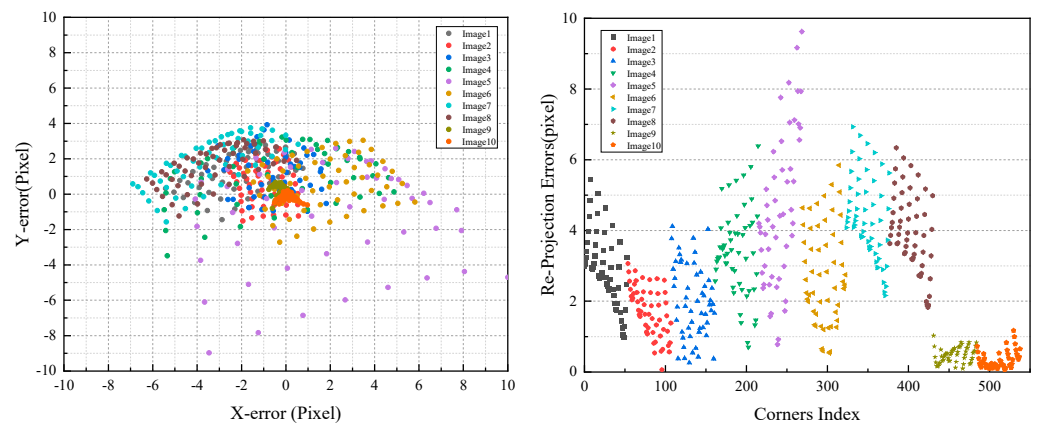


Figure 11. Error of each corner with the second strategy (on the left is the error in the X\Y direction, and on the right are the re-projection errors).

Table 4. Re-projection errors with the second strategy.

Images	Re-Projection Errors/Pixel
1	2.8960
2	1.6936
3	1.9811
4	3.3712
5	5.0023
6	2.7538
7	4.3243
8	3.9146
9	0.4683
10	0.3065
Average	2.6712

However, a notable issue arises in this context. For each picture, there are unique refraction parameters. Here we analyze that the upper and lower limits of the external parameters are too large to cause this. The calculated refraction parameters are not applicable to the previous pictures. It can be seen from Figure 11 that the last time the convergence is better, but the previous convergences are poor.

The third strategy: First, we directly apply Zhang’s calibration [17] to the underwater images to obtain the corresponding external parameters for use in the genetic algorithm. The external parameters are shown in Table 5.

Table 5. Approximate value of external parameters.

Rotating Vector	Translation Vector
−0.25208, −0.33484, −3.06365	113.5505, 62.3148, 540.9440
−0.20195, −0.34487, −2.89052	79.9068, 110.4759, 534.3428
0.11674, 0.32511, −2.53928	3.759, 147.6164, 750.6942
0.179883, 0.13025, −2.30014	−67.7814, 192.1184, 766.6448
−0.3604, 0.77093, −2.97292	155.2326, 75.8162, 714.0783
−0.32534, 0.76271, −2.80196	128.8083, 119.4154, 736.8143
−0.07531, −0.01894, −2.61955	−26.1963, 159.760, 567.3251
−0.25969, 0.25929, −2.61589	−17.6367, 152.3984, 579.6227
0.33259, −0.55940, −2.83098	79.1891, 111.7601, 340.2037
−0.21901, 0.53206, −2.70773	103.839, 87.2968, 327.9117

In order to avoid the “false” phenomenon (The parameters in Strategy 2 only apply to a subset of the images) of external parameters and refraction parameters, we multiply the approximate value of the external parameter obtained above by (1 ± 0.5) as the upper and lower boundary values. We pass each picture into the algorithm and pass the converged refraction parameters updated last time into the next image. We set the population size to 100, update 500 times, and obtain the errors. The objective function values of the population and the best individual for Strategy 3 are depicted in Figure 12, which is the result of the last image, the images after reprojection are shown in Figure 13, the errors are illustrated in Figure 14, and the calibration result is shown in Table 6.

After obtaining the calibration result, we pass the parameters into the model and calculate the error of each corner. Since we set the target value to 0, the iteration did not terminate, but it can be observed that the iteration has converged. The error is very small and performs well in all images. Therefore, we believe that we have obtained the optimal solution.

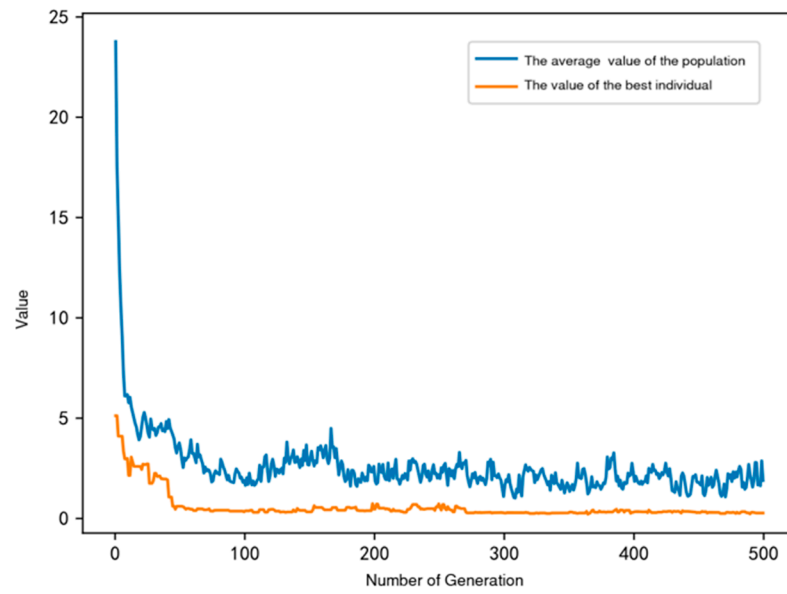


Figure 12. Population and individual information for Strategy 3.

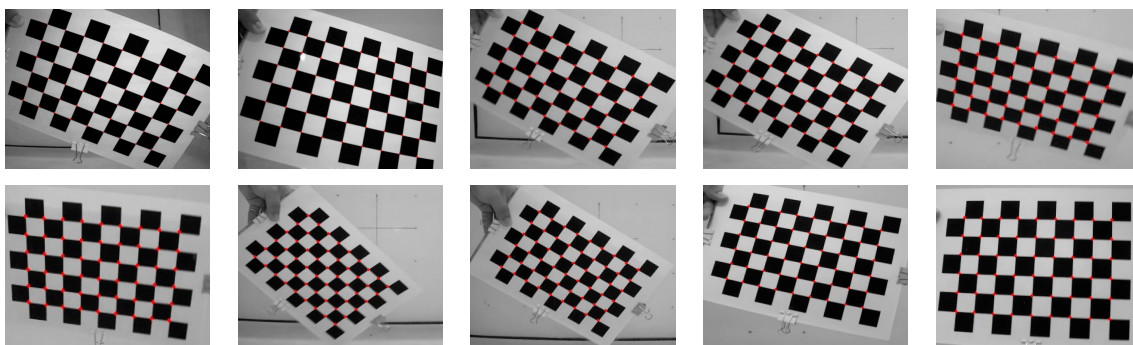


Figure 13. Re-projection images with the third strategy.

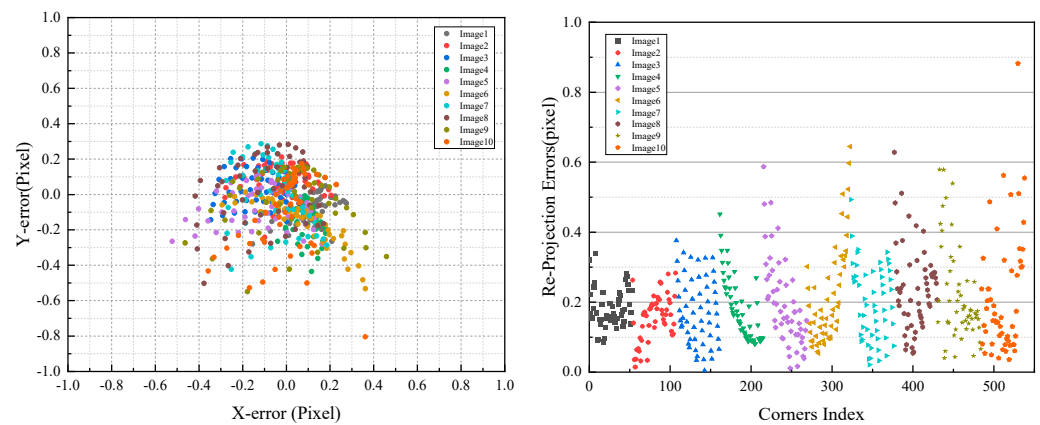


Figure 14. Error of each corner with the third strategy (on the left is the error in the X\Y direction, and on the right are the re-projection errors).

Table 6. Re-Projection errors with the third strategy.

Images	Re-Projection Errors/Pixel
1	0.1776
2	0.1513
3	0.1737
4	0.1772
5	0.1939
6	0.2168
7	0.1965
8	0.2381
9	0.2184
10	0.2146
Average	0.1958

5.3. Comparative Experiment

To validate the effectiveness of our model and algorithm, we conducted comparative experiments. We used ten checkerboard images for camera parameter calibration using Li’s [20] method based on equivalent focal length, Bashar’s [22] method based on an approximate linear model, and our proposed method. We then used the corresponding parameters for re-projection, and selected average error, root mean squared error (RMSE), and re-projection error as comparison metrics. The errors of the comparative methods are shown in Figure 15 and Table 7. It can be observed that our method performs better: compared with Li’s method, our method reduces re-projection error by 24.4%, and, compared with Bashar’s method, it reduces it by 40.2%.

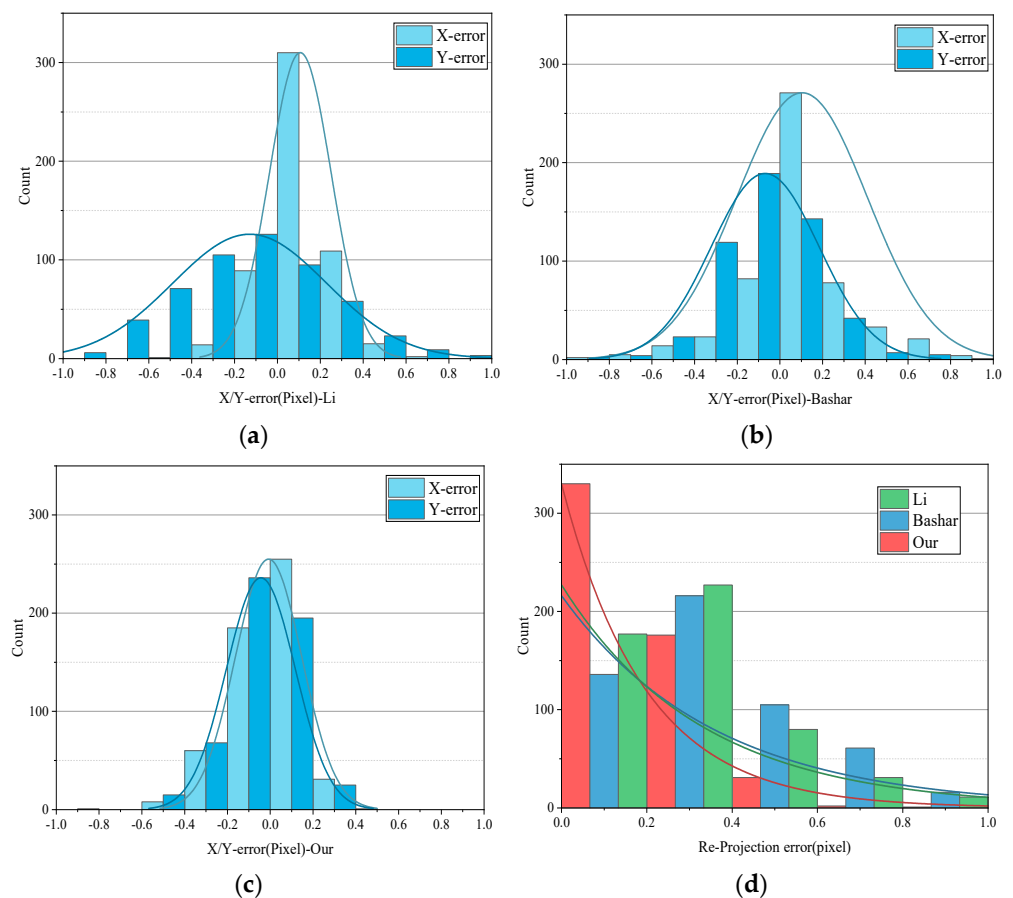


Figure 15. Error distribution charts of different methods. (a), (b), and (c), respectively, represent Li’s method, Bashar’s method, and our proposed method, showing the error distribution of corner points in the X/Y direction after re-projection. (d) displays the distribution of re-projection errors for each method.

Table 7. The various errors of different methods.

Method	Average Error	RMSE	Re-Projection Error
Li's method [20]	0.1544	0.4213	0.2590
Bashar's method [22]	0.2047	0.4166	0.3276
Ours	0.1264	0.2285	0.1958

6. Conclusions

This paper proposes a correction algorithm based on the real light path's double refraction model and demonstrates its necessity in underwater camera usage. During the iterations of the genetic algorithm, certain rules are applied to overcome local optima. Experimental results show that our algorithm accurately computes camera parameters, simplifying the imaging calibration process and providing a feasible method for underwater camera calibration. Comparative experiments reveal that our method reduces the re-projection error by 24.4% compared with Li's method [17] and by 40.2% compared with Bashar's method [19]. The model reduces errors in underwater measurements, improving measurement accuracy, and enhancing underwater positioning and SLAM. In future work, we will apply this model in real underwater environments and address some of the challenges it brings. Additionally, we will apply the model to underwater target positioning, and we also plan to develop target positioning software with this model and positioning technology. We intend to apply it in the field of underwater detection.

Author Contributions: Y.S.: funding acquisition, supervision, resources; T.Z.: methodology, writing—original draft, writing—review & editing; L.Z.: software; P.C.: validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Heilongjiang Province of China, grant number ZD2020E005, Financial support for Shaanxi Provincial Water Conservancy Science and technology program, grant number 2020slkj-5.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A Guide to Deep Learning in Healthcare. *Nat. Med.* **2019**, *25*, 24–29. [[CrossRef](#)] [[PubMed](#)]
2. Xu, S.; Wang, J.; Shou, W.; Ngo, T.; Sadick, A.-M.; Wang, X. Computer Vision Techniques in Construction: A Critical Review. *Arch. Comput. Methods Eng.* **2021**, *28*, 3383–3397. [[CrossRef](#)]
3. Yuh, J.; West, M. Underwater Robotics. *Adv. Robot.* **2012**, *15*, 609–639. [[CrossRef](#)]
4. Capocci, R.; Dooly, G.; Omerdić, E.; Coleman, J.; Newe, T.; Toal, D. Inspection-Class Remotely Operated Vehicles—A Review. *J. Mar. Sci. Eng.* **2017**, *5*, 13. [[CrossRef](#)]
5. Wynn, R.B.; Huvenne, V.A.I.; Le Bas, T.P.; Murton, B.J.; Connelly, D.P.; Bett, B.J.; Ruhl, H.A.; Morris, K.J.; Peakall, J.; Parsons, D.R.; et al. Autonomous Underwater Vehicles (AUVs): Their Past, Present and Future Contributions to the Advancement of Marine Geoscience. *Mar. Geol.* **2014**, *352*, 451–468. [[CrossRef](#)]
6. Zion, B. The Use of Computer Vision Technologies in Aquaculture—A Review. *Comput. Electron. Agric.* **2012**, *88*, 125–132. [[CrossRef](#)]
7. Meline, A.; Triboulet, J.; Jouvencel, B. Comparative Study of Two 3D Reconstruction Methods for Underwater Archaeology. In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura-Algarve, Portugal, 7–12 October 2012; pp. 740–745.
8. Janowski, Ł.; Skarlatos, D.; Agrafiotis, P.; Tysiąc, P.; Pydyn, A.; Popek, M.; Kotarba-Morley, A.M.; Mandlbürger, G.; Gajewski, Ł.; Kołakowski, M.; et al. High Resolution Optical and Acoustic Remote Sensing Datasets of the Puck Lagoon. *Sci. Data* **2024**, *11*, 360. [[CrossRef](#)] [[PubMed](#)]
9. Teng, B.; Zhao, H. Underwater Target Recognition Methods Based on the Framework of Deep Learning: A Survey. *Int. J. Adv. Robot. Syst.* **2020**, *17*, 1729881420976307. [[CrossRef](#)]

10. Mohamed, H.; Nadaoka, K.; Nakamura, T. Assessment of Machine Learning Algorithms for Automatic Benthic Cover Monitoring and Mapping Using Towed Underwater Video Camera and High-Resolution Satellite Images. *Remote Sens.* **2018**, *10*, 773. [[CrossRef](#)]
11. Jian, X.; Xiaoyuan, C.; Xiaoping, S.; Hang, L. Target Recognition and Location Based on Binocular Vision System of UUV. In Proceedings of the 2015 34th Chinese Control Conference (CCC), Hangzhou, China, 28–30 July 2015; pp. 3959–3963.
12. Guo, T.; Capra, A.; Troyer, M.; Grün, A.; Brooks, A.J.; Hench, J.L.; Schmitt, R.J.; Holbrook, S.J.; Dubbini, M. Accuracy Assessment of Underwater Photogrammetric Three Dimensional Modelling for Coral Reefs. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXIII ISPRS Congress, Prague, Czech Republic, 12–19 July 2016; Volume XLI, pp. 821–828.
13. Hidalgo, F.; Bräunl, T. Review of Underwater SLAM Techniques. In Proceedings of the 2015 6th International Conference on Automation, Robotics and Applications (ICARA), Queenstown, New Zealand, 17–19 February 2015; pp. 306–311.
14. Bodenmann, A.; Thornton, B.; Ura, T. Generation of High-Resolution Three-Dimensional Reconstructions of the Seafloor in Color Using a Single Camera and Structured Light. *J. Field Robot.* **2017**, *34*, 833–851. [[CrossRef](#)]
15. Iscar, E.; Skinner, K.A.; Johnson-Roberson, M. Multi-View 3D Reconstruction in Underwater Environments: Evaluation and Benchmark. In Proceedings of the OCEANS 2017—Anchorage, Anchorage, AK, USA, 18–21 September 2017; pp. 1–8.
16. Tsai, R.Y. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In Proceedings of the Computer Vision and Pattern Recognition, Miami Beach, FL, USA, 22–26 June 1986.
17. Zhang, Z. A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [[CrossRef](#)]
18. Lavest, J.M.; Rives, G.; Lapresté, J.T. Underwater Camera Calibration. In Proceedings of the Computer Vision—ECCV 2000, Dublin, Ireland, 26 June–1 July 2000; Vernon, D., Ed.; Springer: Berlin, Heidelberg, 2000; pp. 654–668.
19. Qin, D.; Mao, T.; Cheng, P.; Zhang, Z. The Underwater Camera Calibration Based on Virtual Camera Lens Distortion. In Proceedings of the SPIE—The International Society for Optical Engineering, Beijing, China, 24–26 May 2011; Volume 8193. [[CrossRef](#)]
20. Li, S.-Q.; Xie, X.-P.; Zhuang, Y.-J. Research on the Calibration Technology of an Underwater Camera Based on Equivalent Focal Length. *Measurement* **2018**, *122*, 275–283. [[CrossRef](#)]
21. Sun, J.; Wang, H.; Zhu, X. A Fast Underwater Calibration Method Based on Vanishing Point Optimization of Two Orthogonal Parallel Lines. *Measurement* **2021**, *178*, 109305. [[CrossRef](#)]
22. Bashar, E.; Sagi, F. Direct Linear and Refraction-Invariant Pose Estimation and Calibration Model for Underwater Imaging. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 259–271. [[CrossRef](#)]
23. Treibitz, T.; Schechner, Y.Y.; Singh, H. Flat Refractive Geometry. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8. [[CrossRef](#)]
24. Agrawal, A.; Ramalingam, S.; Taguchi, Y.; Chari, V. A Theory of Multi-Layer Flat Refractive Geometry. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3346–3353. [[CrossRef](#)]
25. Gu, C.; Cong, Y.; Sun, G.; Gao, Y.; Tang, X.; Zhang, T.; Fan, B. MedUCC: Medium-Driven Underwater Camera Calibration for Refractive 3-D Reconstruction. *IEEE Trans. Syst. Man. Cybern. Syst.* **2022**, *52*, 5937–5948. [[CrossRef](#)]
26. Qin, R.; Yang, Y.; Li, F.; Ji, T. Monocular Camera Calibration Based on Particle Swarm Algorithm with All Parameter Adaptive Mutation Mechanism. *Dongnan Daxue Xuebao (Ziran Kexue Ban)/J. Southeast Univ. (Nat. Sci. Ed.)* **2017**, *47*, 193–198. [[CrossRef](#)]
27. Du, S.; Zhu, Y.; Wang, J.; Yu, J.; Guo, J. Underwater Camera Calibration Method Based on Improved Slime Mold Algorithm. *Sustainability* **2022**, *14*, 5752. [[CrossRef](#)]
28. Jebari, K. Selection Methods for Genetic Algorithms. *Int. J. Emerg. Sci.* **2013**, *3*, 333–344.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.