Accuracy analysis of 3D points reconstructed from workspace of underwater robot

B A Skorohod¹, A V Statsenko¹, S I Fateev¹, P V Zhilyakov¹

¹Sevastopol State University, 33 Universitetskaya ave., Sevastopol, 299053, Russia

E-mail: boris.skorohod@mail.ru

Abstract. Video cameras are widely used in underwater robotics to construct 3D coordinates of the workspace. However, the accuracy evaluating that takes into account the properties of the underwater environment, and technical implementation in uncontrolled conditions is still a difficult task. This assessment is especially important for robots that are focused on performing operations with items. In this paper, we propose a novel technique for accuracy analysis and demonstrate its possibilities on real data. It is based on a statistical approach that allows estimating the influence of all sources of perturbations using only experimental data obtained in an underwater environment.

1. Introduction

It is known that the direct use of stereo vision algorithms employed in the air can cause severe difficulties in underwater conditions. The arising problems are related to the lack of robustness of these algorithms to changes in the underwater environment and the peculiarities of transmitting and receiving signals underwater [1, 2]. In addition to the physical features, there are many issues related to the technical implementation. This process means that the cameras must be placed in waterproof shells filled with air and a glass interface. At the same time, it is necessary to ensure sufficiently high accuracy of 3D coordinates of the robot's working space in an underwater environment. This assessment is especially important for robots that are focused on performing operations with items. Accuracy can be influenced by many factors (both in air and underwater environments) including in particular: measurement noise, non-linearity of optics, camera and object movement, estimation of key points and obtaining conjugates points from them, numerical and statistical properties of the triangulation. Numerous publications are devoted to the problem of accuracy of 3D images in the air. Note two approaches — analytical and statistical. Simplified approximations of image processing algorithms are constructed to get an analytical representation for the error. It is assumed that the data is subject to some distribution law, and the task is to find the distribution of the error [3]. The possible limitations of using such approaches are quite visible. Various assumptions can lead to the result that the algorithms have less accuracy than expected. Besides, building a workable model as well as its analysis for many applications can be a difficult task. In contrast to the analytical approach, the statistical approach does not involve the use of mathematical models. It allows estimating the influence of all sources of disturbances (both hardware and software) as a whole [4], relying only on underwater experimental data.

In this paper, we propose a novel technique for accurate analysis of an underwater stereo camera based on a statistical approach that allows estimating the influence of the perturbation using only experimental data obtained in an underwater environment. But unlike [4], we get the error distribution
using the measured values of the calibration sample and obtained by triangulation. This process allows evaluating the distribution of the 3D coordinate recovery error, in principle, for any point in the image.

2. Problem statement

It is assumed that the robot is equipped with a stereo camera that includes two cameras Basler acA1920-50gc rigidly fixed in parallel relative to each other with a stereo base of 10 cm. The cameras have a GigE interface with a Sony IMX174 CMOS matrix, a frequency of 50 frames per second at a resolution of 2.3 megapixels, a sensor size of 11.3 mm x 7.1 mm, Full HD, identical Lens TS0814-MP F1.4 f8mm 1" with a fixed focal length of 8 mm and an aperture F1.4-F16. The image of the cameras and lens are shown in Figure 1.

![Figure 1. The stereo camera and the aquarium.](image)

It is required to estimate the accuracy of restoring 3D objects coordinates within the robot working space (0.25 m – 1.25 m) using a perspective camera model based on the results of experiments in the aquarium shown in Figure 1.

3. Calibrating the stereo camera in underwater conditions

Figure 2 shows the geometry used for building a perspective camera model. Here C is the centre of the image, p is the main point of the image, f is the focal length, Z=f is the image plane perpendicular to the main optical axis, X is a point in 3D space, CZXY, pxy are coordinate systems of 3D space and the image, respectively.

![Figure 2. Geometry of the perspective camera model.](image)

With regard to figure 2, the point \( \mathbf{X} = (X, Y, Z)^T \) specified in the absolute coordinate system is mapped to the point \( \mathbf{x} = (x, y)^T \) of the image coordinate system with the help of the expressions

\[
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix}
= KR
\begin{pmatrix}
X/Z \\
Y/Z \\
1
\end{pmatrix}
+ KT,
\]

\[
K =
\begin{pmatrix}
f_x & 0 & c_x \\
s & f_y & c_y \\
0 & 0 & 1
\end{pmatrix},
\]

where \( K \) is the camera matrix, \( f_x, f_y \) are focal lengths in pixels, \( c_x, c_y \) are the coordinates of the main point relative to the camera origin, \( s \) is the skew coefficient between the x and y axes, \( R \) is the matrix describing the rotation of the camera coordinate system relative to the absolute coordinate system, and \( T \) is the offset vector of the absolute coordinate system origin relative to the camera coordinate system. Expressions describing radial and tangential distortions taking into account the influence of nonlinear optics have the form

\[
x_r = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6),
\]

\[
y_r = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6),
\]

where \( r \) is the radius of the point from the main point.
\[ x_t = x + [2p_1xy + p_2(r^2 + 2x^2)], y_t = y + [2p_2xy + p_1(r^2 + 2y^2)], \]

where \( x, y \) are undistorted pixel locations in normalized image coordinates, \( k_1, k_2, k_3, p_1, p_2 \) are radial and tangential distortion coefficients of the lens, \( r^2 = x^2 + y^2 \).

Comparison of calibration results in air and underwater environments is given in tables 1,2, where \((\alpha, \beta, \gamma)^T, (x_d, y_d, z_d)^T \) are vectors of rotation of camera two relatively to camera 1 (degrees), the vector of movement of camera two relatively to camera 1(mm). The results were obtained using Matlab. Let us make a few additional remarks about the calibration results. First, it follows from Table 1 that the ratio of focal distances in air and water is approximately equal 1.33. Second, it follows from Figure 3 that the camera optics in an underwater environment changes the type of distortion from negative ("pillow-like") to positive ("barrel-like"). You can also notice that the distortions of the left and right cameras are slightly different during experiments in an aqueous environment. The cause is the axes of the camera are not perfectly parallel to each other (Table 2), and therefore the lens of the right camera could not be brought close to the glass.

**Table 1.** Comparison of calibration results in different environments (internal parameters)

|          | Air Camera 1 | Air Camera 2 | Water Camera 1 | Water Camera 2 |
|----------|--------------|--------------|----------------|----------------|
| \( f_x \) | 1399.90      | 1402.70      | 1870.51        | 1877.17        |
| \( f_y \) | 1401.35      | 1401.60      | 1871.56        | 1878.81        |
| \( c_x \) | 973.61       | 965.13       | 966.54         | 959.44         |
| \( c_y \) | 598.41       | 581.10       | 602.66         | 591.67         |
| \( s \)   | 1.9511       | 1.9036       | 2.1567         | 2.4679         |
| \( k_1 \) | -0.1737      | -0.1780      | 0.0667         | 0.0308         |
| \( k_2 \) | 0.1174       | 0.1328       | 0.1348         | 0.3307         |
| \( k_3 \) | -0.0243      | -0.0391      | 0.5258         | 0.2075         |
| \( p_1 \) | 0.0008       | 0.0003       | 0.0042         | 0.0048         |
| \( p_2 \) | 0.0001       | 0.0009       | 0.0007         | 0.0009         |

**Table 2.** Comparison of calibration results in different environments (external parameters)

|          | Air | Water |
|----------|-----|-------|
| \( \alpha \) (°) | -0.075098 | 0.095958 |
| \( \beta \) (°)  | -0.36173  | -0.31579  |
| \( \gamma \) (°) | 0.085815  | 0.091879  |
| \( x_d \) (mm)   | -100.3152 | -100.3409 |
| \( y_d \) (mm)   | 0.0028    | -0.2089   |
| \( z_d \) (mm)   | -1.2879   | -0.7364   |
Analysis of the stereo camera accuracy in controlled underwater conditions

The essence of the proposed approach is as follows. The calibration board (CB) is located parallel to the image plane at different fixed distances from the outer plane of the glass interface within the workspace. The stereo camera complete coordinate system is connected with the left camera, and its origin is located at the point of the pole that is removed from the lens aperture in the negative direction by a distance equal to the focal length. The location of the CB during experiments and the numbering of its corners are shown in Figure 4.

Let the experiments be carried out for N different distances of the CB to the external plane of the glass interface \( Z_l, l = 1, 2, ..., N \) and each corner point of the CB, the estimates of the coordinates of \( X_{l,i,j}, Y_{l,i,j}, Z_{l,i,j} \) \( i = 1, 2, ..., M, j = 1, 2, ..., K, l = 1, 2, ..., N \) are known, where \( M, K \) are the number of angular points in horizontal and vertical, respectively, \( N \) is the number of repeated measurements in each corner point. The expression for the error in the determination of \( Z \) according to the available experimental data can be represented in the form \( e_{l,i,j} = Z_{l,i,j} - Z_l - d \), where \( d \) is the distance from the outer plane of the glass interface to the pole. The value \( d \) depends on the thickness of the glass \( d_g \), the width of the air gap in the protective cap \( d_a \), and the distance from the pole to the air gap \( d_p \). The parameter \( d_g \) is determined with an accuracy of 1 mm, the values of the parameters \( d_a, d_p \) are unknown and are determined by the characteristics of the lens (their value is about 2 cm). As the evaluation of \( d \), the median \( m \) of the known sampled values \( Z_1, 1, 2, ..., Z_N, M, K \) \( Z_N, M, K \), and the scatter about its by expression \( \delta Z_{l,i,j} = Z_{l,i,j} - Z_l - m \), \( i = 1, 2, ..., M, j = 1, 2, ..., K, l = 1, 2, ..., N \) are proposed to use. The sample values of the error for the coordinates \( X \) are defined by expressions \( \delta X_{l,i,j} = X_{l,i,j} - X_{l,i,j} \), \( \delta Y_{l,i,j} = Y_{l,i,j} - Y_{l,i,j}, i \leq M - 6 \), and the scatter \( \delta Z_{l,i,j}, \delta X_{l,i,j}, \delta Y_{l,i,j}, l = 1, 2, ..., M, j = 1, 2, ..., L \) and we characterize their using the histograms.

Six experiments were conducted in an aquarium filled with water. In each experiment, the calibration stand was positioned in parallel to the image plane at a certain distance \( Z_l(0.3; 0.5; 0.75; 1.0; 1.25; 1.5) \) meters (within the working area of the robot) from the inner surface of the aquarium glass to which the stereo camera was attached. Shooting was performed at a frequency of 1 frame per second for 1 hour. Thus, 3600 images were obtained for each experiment. The 3D coordinates of the CB points were determined using the algorithm for determining angular points on the CB and algorithms for moving from 2D image coordinates to 3D spatial coordinates.
coordinates. Let, for example, \( Z_l, l = 1, 2, ..., 6 \) are measured values of distances from the 28th corner point (Figure 4) to the outer plane of the glass interface, \( \hat{Z}_l, l = 1, 2, ..., 6 \) are estimates of distances from the same corner point to the center point of the image of the left camera. Figure 5 shows histograms of distributions of the error \( \delta Z \) of \( Z \), using the values \( Z_l - \hat{Z}_l, l = 1, 2, ..., 6 \) for the 28th corner point. When constructing histograms, the normalized probability \( p_i \) was used, determined by the expression \( p_i = c_i / (N \cdot w) \), where \( c_i \) is the number of elements in the interval, \( N \) is the total number of elements, and \( w \) is the width of the interval. Visual analysis of histograms shows that the error distribution laws for all ranges are close to normal; the average values for all ranges are approximately the same (differences up to 3mm); as the distance to the CB increases, the spread of error values increases; the observed error values are positive. Figure 6 shows a histogram of the distribution of values \( Z_l - \hat{Z}_l, l = 1, 2, ..., 6 \) for the combined sample, the median of which is approximately 27.7 mm. From figure 6, it follows that the distribution differs from the normal law and the error for all distances belonging to the working zone does not exceed 6 mm.

![Histograms](image)

**Figure 5.** Histograms of the \( \delta Z \) for different \( Z \) and the point №28: a – \( Z_1 = 0.3m \); b – \( Z_2 = 0.5m \); c – \( Z_3 = 0.75m \); d – \( Z_4 = 1m \); e – \( Z_5 = 1.25m \); f – \( Z_6 = 1.5m \).

![Histogram](image)

**Figure 6.** Histogram of the joint sample.

Study of the stereo camera accuracy was also performed independently for a specially prepared scene shown in Figure 7, including analogues of human-made underwater infrastructure objects. The numbers of the images correspond to the following experimental conditions: 1205 – the aquarium is closed, and one lamp is turned on; 1207 – the aquarium is opened, and two lamps are turned on (two lamps + the Sun); 1209 – an outside light is turned, no sun; 1210 – an outside light is turned, and the lid of the aquarium is opened, there is no sun. Two algorithms were used constructing stereo images based on dense 3D reconstruction (DS) and sparse 3D reconstruction (SS) [5].
Figure 7. Underwater scene: a – 1205; b – 1207; c – 1209; d – 1210.

The results of the processing are shown in Table 3. Let us make a few comments about the results shown in Table 3. First, the SS and DS algorithms allow getting 3D scene coordinates that are close in accuracy for all lighting options and the Z error does not exceed 15 mm. Second, the obtained error values are slightly larger than the error values obtained using the CB. We explain this by the fact that the accuracy of determining critical points as a result of triangulation is higher for the CB due to the particular requirements for its design.

| Object          | Z (mm) | 1207 | 1205 | 1210 | 1209 | 1207 | 1205 | 1210 | 1209 |
|-----------------|--------|------|------|------|------|------|------|------|------|
| Red ball        | 250    | 4    | 9    | 4    | 4    | 9    | 8    | 6    | 3    |
| Red cube        | 400    | 6    | 6    | 7    | 11   | 5    | 6    | 10   | –    |
| Parallelepiped 1| 600    | 3    | 4    | 3    | 4    | 4    | 5    | –    | –    |
| Hexahedron      | 750    | 2    | 3    | 4    | –!   | -1   | -1   | -1   | 0    |
| Parallelepiped 2| 1000   | 4    | 2    | 16   | 4    | 2    | 2    | 5    | 6    |
| Black cube      | 1150   | 1    | -2   | –    | 0    | 0    | 0    | –    | –    |
| Cone            | 1300   | 13   | 11   | 10   | –    | 13   | 12   | 14   | 15   |
| The aquarium wall| 1468  | -3   | 2    | 0    | -5   | -4   | 0    | 1    | 1    |

5. Conclusion
The main contributions of this paper are as follows. It was proposed a novel technique for accurate analysis of 3D points reconstructed from the underwater robot workspace. It was used the statistical approach that allows estimating the influence of all sources of perturbations using only on experimental data obtained in a marine environment. We demonstrate possibilities of our approach on real data.
6. Acknowledgments
Studies were supported by the Ministry of Science and Higher Education of the Russian Federation (agreement № 14.578.21.0264, project RFMEFI57818X0264).

References
[1] Bonin F, Burguera A and Oliver G 2011 Imaging systems for advanced underwater vehicles Journal of Maritime ResearchVIII (1) 65-86
[2] Dario L, Kallasi F, Aleotti J, Oleari F and Caselli S 2017 Underwater vehicle for pipe manipulation tasks Computers & Electrical Engineering58 560-571
[3] Blostein, D and Steven S 1987 Error analysis in stereo determination of 3-D point positions Pattern Analysis and Machine Intelligence IEEE Transactions6 752-765
[4] Cabrera E, Luis E, Bruno M, Esteban W, Luiz M and Gonçalves 2018 A versatile method for depth data error estimation in RGB-D Sensors Sensors (Basel)18(9) 3122
[5] Bradski, G, and Kaehler A 2008 Learning OpenCV: Computer Vision with the OpenCV Library (O'Reilly Media) p 555