

Structured Light System Calibration for Perception in Underwater Tanks

Flávio Lopes, Hugo Silva, José Miguel Almeida, Eduardo Silva

INESC TEC Technology and Science

Instituto Superior de Engenharia do Porto

flavio.lopes, hugo.m.silva, jose.m.almeida,eduardo.silva@inesctec.pt,
Rua Dr^o António Bernardino de Almeida 431, 4200-072 Porto, Portugal

Abstract. The process of visually exploring underwater environments is still a complex problem. Underwater vision systems require complementary means of sensor information to help overcome water disturbances. This work proposes the development of calibration methods for a structured light based system consisting on a camera and a laser with a line beam. Two different calibration procedures that require only two images from different viewpoints were developed and tested in dry and underwater environments. Results obtained show, an accurate calibration for the camera/projector pair with errors close to 1 mm even in the presence of a small stereo baseline.

1 Introduction

The process of visual exploring underwater environments is an important research challenge. The range of unmanned vehicles that require underwater perception capabilities is rapidly growing due to the more frequent use of this vehicles in inhospitable places such as oil and gas platforms and deep sea underwater mining.

In this paper, we propose to develop a structured light system (SLS) for conducting perception tasks in underwater infrastructures. Normally, pure visual methods struggle in underwater environments due to phenomena such as: poor scene illumination, lack of image texture and water disturbances like high turbidity, absorption and scattering [1]. Therefore, conventional computer vision approaches using stereo vision suffer from the difficulty of obtaining correspondences between image points. One way of overcoming such limitation is by using laser based techniques [2], or laser techniques combined with photometric stereo [3]. In [4] structured light systems are used to perform bathymetric operations using Remotely Operational Vehicles and discoverer diffuse flow on the seafloor.

To perform such operations in an accurate manner, an efficient calibration procedure that allows to calibrate the camera and the laser must be developed. Therefore, our main work was to develop, experiment and evaluate calibration methods that allow to perform all necessary SLS calibration.

The problem of calibrating a triangulation system between a laser (projector) and a camera has already been addressed by many authors, but the underlying

idea is still subject of ongoing research specially when trying to apply it for underwater operations. In [5] a calibration rig consisting of lines with known parameters is used. The intersection of the laser plane with each line gives an image point to world line correspondence. Having enough correspondences a 4×3 transformation matrix which maps a 2D image point to its 3D coordinate on the laser plane can be recovered. Reid et al [6] uses a similar solution, using a calibration rig consisting of planes with known parameters. Projecting the line laser onto the known planes gives an image point to world plane correspondence, and then the 8-point algorithm [7] is used to find the transformation matrix.

One of the methods used in our experiments was a modified version on the method developed by Huynh [8]. The method uses a calibration rig consisting of two planes with 12 points (control points) with known 3D coordinates. The points form 4 straight lines, and by using the cross-ratio invariance [7], the 3D coordinates of the point of intersection of the laser plane with these 4 lines can be determined, allowing to obtain 2D image point to 3D world point correspondences that are used to recover the transformation matrix.

Some of these methods are complex and so other simpler methods that require just a few measurements to a known planar target from different viewpoints can also be used to compute the camera and projector calibration [9], [10]. An image of the calibration board is captured, along with the line of its intersection with the laser plane. Then by correspondences between board points and their image coordinates, the pose of the calibration board with respect to the camera center can be determined, resulting in a known world plane. The image of the laser line is then back-projected to the world plane. Performing this operations with the calibration board at different orientations results in a 3D point cloud spanning the laser plane, from which the parameters that define the laser plane can be recovered.

2 Geometric Model

Our proposed structured light system is composed of a camera and a single projector equipped with a line beam. The main objective is to be able to obtain accurate 3D points information from objects that are illuminated by the projector, see Fig.1.

2.1 Camera Model

The camera model follows a classic pinhole model of perspective projection divided into intrinsic (1) and extrinsic parameters.

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where K is the intrinsic parameter matrix, that contains the focal length \mathbf{f} (f_x , f_y), the skew s , and the principal point \mathbf{c} (c_x , c_y) information. There is still one

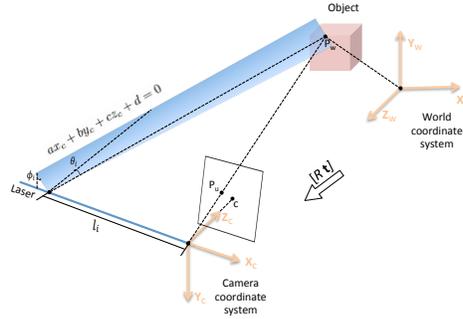


Fig. 1. Geometric model of the SLS system

parameter missing which is the lenses distortion that needs to be compensated when calibrating the camera [11].

For the extrinsic parameters, the projection of a 3D world point \mathbf{P}_w represented in the world coordinate system to a 2D image point \mathbf{p} (x,y), in the image plane is given by :

$$\tilde{\mathbf{p}} \simeq K[R|\mathbf{t}]\tilde{\mathbf{P}}_w \quad (2)$$

where R is the rotation matrix, K the intrinsic parameters and \mathbf{t} the translation vector that allows to translate a point between the world reference frame and the camera reference frame. While $\tilde{\mathbf{p}}$ is a 2D image point and $\tilde{\mathbf{P}}_w$ is the same point in the 3D world point reference frame, both in homogeneous coordinates representation.

2.2 Projector Model

For the projector calibration, the perspective projection does not hold since the light stripe is not directly illuminated from the optical center of the projector. Therefore, most applications do not use the projection of the light stripe, and instead prefer to represent the projector by applying the equation of a plane. The light stripe that appears in 3D space can be expressed in the camera coordinate frame by the plane equation, described by:

$$a_i x_c + b_i y_c + c_i z_c + d = 0 \quad (3)$$

where i is the light stripe number, a_i , b_i , c_i , and d_i are its coefficients. From here, it is possible to define the baseline l_i i.e the distance between the camera and the projector, and also the projection angle θ_i i.e the angle between the light stripe and the Z_c axis, and the tilt angle ϕ_i i.e the angle between the Y_c axis and the light stripe. The parameters can be obtained by the following expressions:

$$l_i = \left(\frac{d_i}{a_i} \right) \quad (4)$$

$$\theta_i = \arctan\left(\frac{-c_i}{a_i}\right) \quad (5) \quad \phi_i = \arctan\left(\frac{-b_i}{a_i}\right) \quad (6)$$

2.3 Triangulation

For obtaining the 3D world point information \mathbf{P}_w in the camera reference frame as $\mathbf{P}_c (x_c, y_c, z_c)$, a triangulation principle between the camera viewpoint and the light stripe is used. To do so, the camera and projector parameters are derived from (2) and (3), as given by:

$$\begin{bmatrix} f_x & s & 0 \\ 0 & f_y & 0 \\ a_i & b_i & d_i \end{bmatrix} \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \\ \frac{1}{z_c} \end{bmatrix} = \begin{bmatrix} x - c_x \\ y - c_y \\ -c_i \end{bmatrix} \quad (7)$$

So, we can estimate \mathbf{P}_c coordinates as:

$$x_c = \frac{x - c_x}{f_x} z_c \quad (8) \quad y_c = \frac{y - c_y}{f_y} z_c \quad (9)$$

$$z_c = \frac{l_i}{-\frac{c_i}{a_i} - \left(\frac{x - c_x}{f_x}\right) - \frac{b_i}{a_i} \left(\frac{y - c_y}{f_y}\right)} \quad (10)$$

The triangulation principle is used to compute the depth coordinate z_c (10). Afterwards, x_c, y_c coordinates are obtained based on the scaling relation with the camera (8)(9), we consider skew $s = 0$.

3 Calibration

Prior to be able to use the structured light system for obtaining accurate 3D underwater measurements, there is the need to perform a global system calibration. Not only we need to calibrate the camera (intrinsic and extrinsic parameters), but also to relate the projector position with respect to the image plane.

3.1 Camera Calibration

The first step to obtain a global system calibration is to recover all camera parameters. Therefore, we start by obtaining the camera intrinsic parameters by using an offline toolbox procedure.

For the camera extrinsic parameters, they are obtained by taking at least two images of a plain chessboard in different viewpoints. Afterwards, all chessboard corners present in each image are detected using the Harris corners detector [13]. The extrinsic parameters (rotation matrix R , translation vector \mathbf{t}) are recovered based on homography decomposition method [7].

We establish that 3D world points in the world reference frame have null component in Z_w axis, i.e, so for all the selected points z_w value is zero. These points can be represented in homogeneous coordinates by equation (11). The 2D points in image plane without distortion, can be represented in homogeneous coordinates by equation (12).

$$\tilde{\mathbf{P}}_{\mathbf{w}} = \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} \quad (11) \quad \tilde{\mathbf{P}}_{\mathbf{u}} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (12)$$

To define the set of points, $\tilde{\mathbf{P}}_{\mathbf{u}}$ and $\tilde{\mathbf{P}}_{\mathbf{w}}$, all corners of the chessboard are used. So, we can relate a world point $\tilde{\mathbf{P}}_{\mathbf{w}}$ and its image correspondent $\tilde{\mathbf{P}}_{\mathbf{u}}$ up to a unknown scale λ by means of homography H as given by:

$$\lambda \tilde{\mathbf{P}}_{\mathbf{u}} = H \tilde{\mathbf{P}}_{\mathbf{w}} \quad (13)$$

To calculate the rotation matrix and translation vector is necessary to perform homography H decomposition by:

$$H = K [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}] = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3] \quad (14)$$

With the intrinsic parameters known, the transformation from world coordinates to the camera coordinate frame can be obtained by:

$$\left\{ \begin{array}{l} \mathbf{r}_1 = \lambda_1 K^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 = \lambda_2 K^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} = \lambda_3 K^{-1} \mathbf{h}_3 \end{array} \right. \quad (15) \quad \left\{ \begin{array}{l} \lambda_1 = \frac{1}{\|A^{-1} \mathbf{h}_1\|} \\ \lambda_2 = \frac{1}{\|A^{-1} \mathbf{h}_2\|} \\ \lambda_3 = \frac{\lambda_1 + \lambda_2}{2} \end{array} \right. \quad (16)$$

Thus, the rotation matrix relating the world coordinate system to the camera coordinate system is described by:

$$R_{temp} = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3] \quad (17)$$

Since the matrix R_{temp} does not satisfy the orthonormality constraint of a standard rotation matrix. It is necessary to normalize and orthonormalize the matrix, using the SVD (Singular Value Decomposition) by:

$$R = UW'V^T \quad (18)$$

Where the matrix W' is:

$$W' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \|UV'\| \end{bmatrix} \quad (19)$$

With the calculation of the matrix R and the translation vector \mathbf{t} is possible to define the position of the optical center of the camera relative to the world's reference, thus completing the calibration process of extrinsic parameters.

3.2 Projector Calibration

For the projector calibration we implemented two different methods that are described in detail in the following sections. The first method denoted as **SLSC-CR** method is based on the cross-ratio invariance principle. While the second denoted as **SLSC-LP** is based on the robust fitting of the laser line projection.

The methods allow us to obtain control points belonging to the light stripe. Afterwards, by non-linear least squares fitting it is possible to define the equation of the plane in the camera coordinate frame using control points of two different images taken in different viewpoints. Thus, by equation (20) it is possible to determine each axis component (a, b, c, d) :

$$d_i = \frac{|ax_c + by_c + cz_c + d|}{(a^2 + b^2 + c^2)^{\frac{1}{2}}} \quad (20)$$

where x_c, y_c, z_c are the control points coordinates in the camera reference frame.

3.3 SLSC-CR Method

The objective for the projector calibration is to discover the equation of the plane that describes the projector light stripe in the camera reference frame.

Our first implementation is denoted as SLSC-CR (Structured Light System Calibration - Cross-Ratio). The cross-ratio allows to calculate the relationship between a set of collinear points belonging to the same line. This relationship is invariant under projective transformation, and can help us relate known laser points in the image reference frame with the laser points correspondences in the world reference frame.

The relationship used to establish the cross-ratio is displayed in Fig. 2. It contains the 3D camera coordinate frame with four collinear points (a, q, c, d) and 3D world coordinate frame defined on the calibration target with points (A, Q, C, D) . The plane π_C defines the camera plane, plane π_W is the calibration target plane and π_l is the projector plane.

We start by calculating the cross-ratio using a set of points in image coordinates for each horizontal line of the planar chessboard. The following procedure was defined:

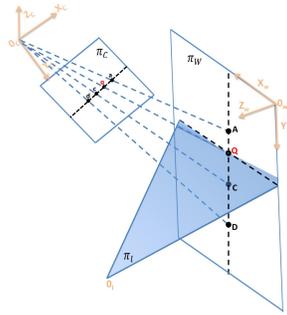


Fig. 2. Control points used to calculate the cross-ratio.

- a_i and d_i - corners of the right and left side of the chessboard.
- b_i or c_i - middle corner of the chessboard (the definition of this point depends of the q_i point position)
- q_i - intersection point between the projector plane and target horizontal line.

For computing the cross-ratio is necessary to preserve the correct order of selected points. Therefore, we considered two different situations for the projector line points, i.e, when the projector focus on the right side or left side of the calibration plane. In Fig. 3 is possible to see control points selection in both cases.

Afterwards, the value of the cross-ratios using 4 collinear points in image plane, can be defined as CR_i . As the value of cross-ratio is invariant, the value of CR_i in world plane is the same. So, we can calculate the equation (21) in order to control point Q , as given by (when the projector focus on the left side):

$$CR_i = \frac{(A_i - C_i)(Q_i - D_i)}{(A_i - D_i)(Q_i - C_i)} \iff Q_i = \frac{-A_i D_i + C_i D_i + A_i C_i CR_i - C_i D_i CR_i}{A_i CR_i - D_i CR_i - A_i + C_i} \quad (21)$$



Fig. 3. Point selection based on light stripe chessboard position

When the projector focus in the right side, a similar procedure can be used.

For this procedure was considered, for each image, at least six control points. The control points in world coordinates can be computed from their corresponding camera coordinates.

3.4 SLSC-LP Method

Our second implementation denoted as SLSC-LP (Structured Light System Calibration - Line Projection). With this method it is possible to increase the control points that defining the projector plane. With the 2D image points belonging to the light stripe and the camera calibration parameters is possible obtain these points in 3D coordinates. We used the projective equation and assuming that $z_w=0$ to calculate the points in world frame (x_w, y_w) :

$$\begin{bmatrix} P_u \\ P_v \\ 1 \end{bmatrix} = K [R \mid \mathbf{t}] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \iff \begin{bmatrix} P_u \\ P_v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_x \\ R_{21} & R_{22} & R_{23} & t_y \\ R_{31} & R_{32} & R_{33} & t_z \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (22)$$

$$\begin{cases} x_w \overbrace{(P_{x'} R_{31} - R_{11})}^{a1} + y_w \overbrace{(P_{x'} R_{32} - R_{12})}^{b1} = \overbrace{(P_{x'} t_z + t_x)}^{c1} \\ x_w \overbrace{(P_{y'} R_{31} - R_{21})}^{a2} + y_w \overbrace{(P_{y'} R_{32} - R_{22})}^{b2} = \overbrace{(P_{y'} t_z + t_y)}^{c2} \end{cases} \quad (23)$$

$$\begin{cases} (b_2 a_1 - b_1 a_2) x_w = (b_2 c_1 - b_1 c_2) \iff x_w = \frac{(b_2 c_1 - b_1 c_2)}{(b_2 a_1 - b_1 a_2)} \\ (a_2 b_1 - a_1 b_2) y_w = (a_2 c_1 - a_1 c_2) \iff y_w = \frac{(a_2 c_1 - a_1 c_2)}{(a_2 b_1 - a_1 b_2)} \end{cases} \quad (24)$$

$$P_{x'} = \frac{P_u - c_x}{f_x} \quad P_{y'} = \frac{P_v - c_y}{f_y} \quad (25)$$

4 Experimental Results

In this section results for both methods implemented with experiments in dry and underwater environments are presented.

4.1 Experimental Setup

Our experimental setup consists of a red 635 nm wavelength line laser beam, and a camera with SXGA resolution both contained in a water proof casing with a 10 cm baseline, see Fig. 4. The procedures were implemented in MATLAB as a proof concept. As for the processing time it is mainly dependent on the number of images that are used for calibrating the SLS system.

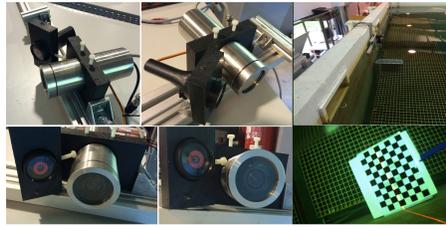


Fig. 4. Structured Light System Underwater Experimental Setup

In order to validate the proposed solutions, we compare the methods solution with the results obtained using known 6 points 3D information (world coordinates) obtained using the camera sensor model with 2D image information by applying the cross-ratio invariance. Then, both method results are compared against these obtained points. This can be considered a weak ground-truth system. In future work, we plan to use a LIDAR system that can map the real position of the planar chessboard and compare the results of both implementations with the 3D LIDAR information.

4.2 Results

The results presented in table 1, clearly show that the line projection method performs more accurately than the demonstrated by the cross-ratio method, with over 40% improvement. For obtaining this results we used 96 points (6 control points over 16 images). The same experiment was performed in underwater environment, see table 2. The line projection also outperforms the cross-ratio projection method. It is important to mention that both methods accumulate 2x more error in the underwater environment.

Table 1. RMS Error Comparison between **SLSC-CR** and **SLSC-LP** in dry environments

Methods	x (mm)	y (mm)	z (mm)	xyz (mm)
SLSC-CR	0.3207	0.3346	1.2340	1.8893
SLSC-LP	0.2134	0.3473	0.6491	1.0579

Table 2. RMS Error Comparison between **SLSC-CR** and **SLSC-LP** in underwater environments

Methods	x (mm)	y (mm)	z (mm)	xyz (mm)
SLSC-CR	0.1068	0.4215	1.8795	2.4078
SLSC-LP	0.1192	0.3642	1.7798	2.2632

5 Conclusions and Future Work

In this work we present two methods for performing the calibration of a structured light system for perception in underwater tanks. The first implementation was based on the cross-ratio invariance principle, while the second implementation was based on finding the line projection in the camera reference frame. The results comparison of both methods in dry and underwater experiments shows that the line projection method has lower root mean square error both in dry and underwater calibration scenario. The system will be in future work used for underwater inspection operations of ship hulls and harbor walls, and a new version for mapping in industrial indoor underwater tanks is also being currently develop.

References

1. Schechner, Y.Y.; Karpel, N., "Clear underwater vision," Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on , vol.1, no., pp.I-536,I-543 Vol.1, 27 June-2 July 2004
2. F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, A.V. Razonale, "Experimentation of structured light and stereo vision for underwater 3D reconstruction", ISPRS Journal of Photogrammetry and Remote Sensing, Volume 66, Issue 4, July 2011, Pages 508-518, ISSN 0924-2716
3. Srinivasa G., Narasimhan, S.K. Nayar, "Structured light methods for underwater imaging: light stripe scanning and photometric stereo," Proceedings of 2005 MTS/IEEE OCEANS, September, 2005, pp. 2610 - 2617.
4. Inglis, G., Smart, C., Roman, C., Carey, S. "Detection of diffuse sea floor venting using structured light imaging". Poster OS11B-1473 presented at 2011 Fall Meeting, AGU, San Francisco, Calif., 5-9 Dec.
5. C. Chen and A. Kak, "Modeling and calibration of a structured light scanner for 3-d robot vision," in IEEE International Conference on Robotics and Automation, vol. 4, 1987, pp. 807-815.
6. I. D. Reid, "Projective calibration of a laser-stripe range finder," Image and Vision Computing, vol. 14, no. 9, pp. 659-666, 1996.
7. Hartley, R., Zisserman, A., "Multiple View Geometry in Computer Vision" , Cambridge University Press , New York, NY, USA 2003.
8. D. Q. Huynh, R. A. Owens, and P. E. Hartmann, "Calibrating a structured light stripe system: A novel approach," International Journal of Computer Vision, vol. 33, pp. 73-86, 1999.
9. Yamauchi, K.; Saito, H.; Sato, Y., "Calibration of a structured light system by observing planar object from unknown viewpoints," 19th International Conference on Pattern Recognition ICPR 2008 , vol., no., pp.1,4, 8-11.
10. Fuqiang Zhou, Guangjun Zhang, "Complete calibration of a structured light stripe vision sensor through planar target of unknown orientations", Image and Vision Computing, Volume 23, Issue 1, 1 January 2005, Pages 59-67, ISSN 0262-8856,
11. Zhang, Z., "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations" ., in International Conference on Computer Vision ICCV 2009, pp. 666-673.
12. Harris, C., Stephens, M., "A Combined Corner and Edge Detector", in 'Proceedings of the 4th Alvey Vision Conference' , pp. 147-151, 1988.