Underwater Camera Calibration

J.M. Lavest, G. Rives, and J.T. Lapresté

LASMEA, UMR 6602 du CNRS, Blaise-Pascal University of Clermont-Ferrand,
F.63177 Aubière cedex. France
lavest@lasmea.univ-bpclermont.fr

Abstract. This article deals with optical laws that must be considered when using underwater cameras. Both theoretical and experimental point of views are described, and it is shown that relationships between air and water calibration can be found.

1 Introduction

Use of vision systems in media in which the wave speed propagation is not that one of air is a subject seldom treated in the Vision community. However, any trial to localize or reconstruct an object observed by an underwater camera (for instance) has to go through a calibration phase.

This article presents some optical considerations relating to underwater cameras.

We show the relationship between the current pin-hole model of the camera and the general optical model of the lens combination for the same camera in air and under water.

The relations found are verified in simulation and by experiments. We prove that the calibration of a camera working under water does not have to be carried out under water.

The intrinsic parameters of a camera immersed in any fluid can be computed from an air-calibration as soon as the optical surface between the two fluids presents some simple geometrical properties.

2 Optics

The classical model used in artificial vision for description of image formation is the perspective projection and thus the pine-hole (or *sténopé*) model of projection.

Some links between this model and classical optical laws were established in [5], but the object and its image were both in the same homogeneous medium, namely the air.

Underwater camera calibration must involve a slightly more general optical model, taking account of the different fluids in which the object and the image are situated .

D. Vernon (Ed.): ECCV 2000, LNCS 1843, pp. 654–668, 2000.

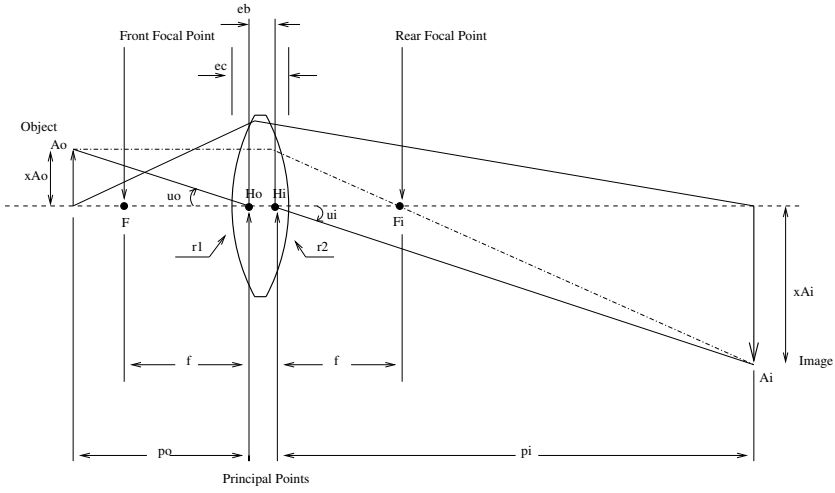


Fig. 1. Classical thick model, in an homogeneous fluid

2.1 Prerequisites

- **Conjugate planes:** if an optical system makes the rays from an object point A_o converge to a point A_i, then A_i is said to be the image or equivalently the conjugate of A_o.
- **Transversal magnification:** (Gt). If x_{Ao} and x_{Ai} are the respective distances of points A_o and A_i to the optical axis, the transversal magnification Gt is equal to the ratio of these distances:

$$Gt = \frac{x_{Ai}}{x_{Ao}}$$

- **Angular magnification:** (Ga). Angular magnification denotes the ratio of the incident and emergent angles (u_o, u_i) of an optical ray going through two conjugate points of the optical axis.

$$Ga = (\frac{u_i}{u_o})_{x=0,y=0}$$

- **Principal planes and points:** object and image principal planes are conjugate planes orthogonal to the optical axis such that the transversal magnification is **one**. This definition implies that the rays between these two planes are parallel to the optical axis.
  Principal points are the intersection of the principal planes with the optical axis.
- **Focal length:** fi. This is the distance (H_iF_i) where F_i is the image focal point. We recall that the image of an object located at infinity undergoes no blurring at the focal point.
- **Nodal points:** These are the pair of conjugate points on the optical axis N_o et N_i such that every ray through N_o emerges at N_i without change of direction (i.e. the angular magnification is **one**).

2.2 Thick Model, for Two Different Homogeneous Fluids

Most vision applications deal with a camera immersed in an homogeneous fluid, namely *air*. Under such an hypothesis some simplifications arise and it can be shown [5] that nodal points and principal points coincide. The use of a pin-hole model consists in merging the two principal planes in order to only retain rays through the equivalent optical center.

Paraxial formulas for a lens located between two distinct homogeneous fluids are found in most handbooks of geometrical optics [7]. We recall the expressions that will be used hereafter.

These formulas extend the properties of the lenses to the arbitrary refractive index of the object (n_1) and of the image (n_2) media, also involving mechanical specification of the lenses in which glass has an index equals to n:

For opticians, refractive index is the ratio of the speed of light in air and the speed of light in the considered medium. When the situation involves two different extremal indices, the focal length f has two distinct values f_o for the object medium and f_i for the image medium. Moreover nodal and principal points are now distinct.

The following relations hold between the different optical variables (C.f. Figure 2.2).

1. Lens constant

$$k = \frac{n - n_1}{r_1} + \frac{n_2 - n}{r_2} - \frac{tc(n - n_1)(n_2 - n)}{n.r_1.r_2} \tag{1}$$

2. Focal lengths:

$$f_o = \frac{n_1}{k}, f_i = \frac{n_2}{k} \tag{2}$$

3. Gauss relation

$$\frac{n_1}{p_1} + \frac{n_2}{p_2} = k \tag{3}$$

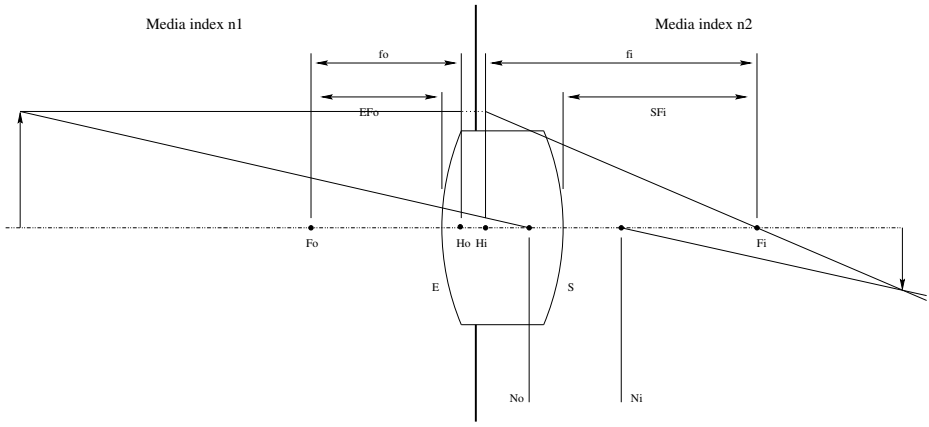


Fig. 2. Optical ways in fluids of different indices

4. Principal points locations

$$EH_o = \frac{n_1.tc}{k} \frac{(n_2 - n)}{n.r_2} \tag{4}$$

$$SH_i = \frac{-n_2.tc}{k} \frac{(n - n_1)}{n.r_1} \tag{5}$$

5. Nodal points locations

$$EN_o = EH_o + H_oN_o \tag{6}$$

$$SN_i = SH_i + H_iN_i \tag{7}$$

with

$$H_oN_o = H_iN_i = \frac{(n2 - n1)}{k} \tag{8}$$

2.3 Entry Surface Properties

For most applications involving an underwater camera, the lens system is set to be focused at infinity. This allows a focused image to be obtained from an infinite distance, up to a few centimeters of the entry surface (the minimum

focused distance decreases with short focal lengths).The above consideration implies that the photosensitive matrix is almost always located at the image focal point.

Underwater cameras often possess an entry surface where the external surface is a plane. This property can be explained by the necessity of obtaining focused images in water as well as in air conditions.

In fact, an image will remain focused independently of the object medium index if and only if its image focal point remains unchanged through index variation. The image focal point is determined by the distance SF_i (where S denotes the exit surface of the last lens of the optical system).

$$SF_i = SH_i + H_iF_i = SH_i + f_i = \frac{-n_2.tc}{k}\frac{(n-n_1)}{n.r_1} + \frac{n_2}{k} \tag{9}$$

When r_1 grows to infinity (which is equivalent to obtaining a plane surface at the air/water interface), the expression becomes independent of n_1:

$$k = \frac{(n_2 - n)}{r_2}$$

and SF_i can be written:

$$SF_i = \frac{n_2}{k} = \frac{n_2.r_2}{(n_2 - n)} \tag{10}$$

This location of the image focal point related to the out surface (the one nearest to the CCD matrix) is also independent of n_1: the image is focused in air as well as in water.

2.4 From Thick Model to Pin-Hole Model: The Nodal Points Influence

The *pin-hole* model merely consists of using only one optical ray through an equivalent point called the optical center. The extension of the optical model to different media indices, shows that the role of the optical center will be played by the fusion of the two nodal points that conserve the angular magnification.

In the vision community the focal length is defined by the distance between the CCD sensor and the optical center. It can be seen in figure (2.2) that this distance is equivalent to N_iF_i if the object is at infinity.

$$N_iF_i = N_iH_i + HiF_i = \frac{(n_1 - n_2)}{k} + \frac{n_2}{k} = \frac{n_1}{k} = f_o$$

$$f_o = N_iF_i = n_1\frac{r_2}{(n_2 - n)} = n_1 * Ct \tag{11}$$

This last relation is a major one for our purpose. It can be noticed that the (vision community) 'focal length' is directly proportional to the object medium index n_1, because n (glass index), n_2 (CCD medium index $n_2 = 1$) and r_2 are constant.

When the camera is underwater, the focal length is equivalent to the value measured in air multiplied by a factor 1.333***.***

2.5 Distortions and Changes of the View Cone

It is obvious that the variation of the focal length implies a decrease of the solid angle of view, when the canera is immersed. This variation is directly proportional to the index because the image size (and hence the CCD size) is constant. (C.f. Figure 3)

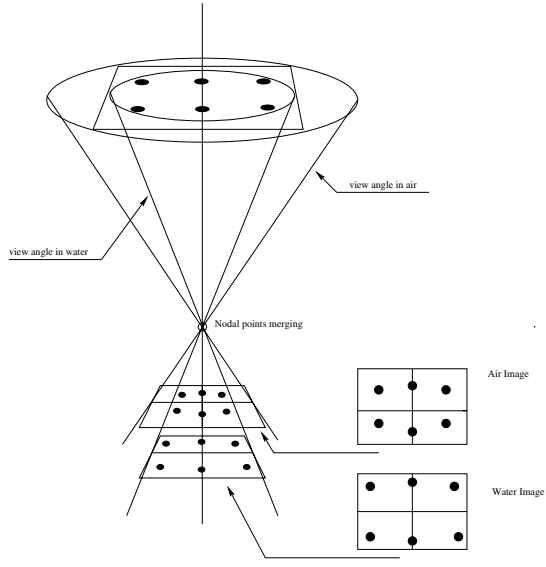


Fig. 3. *variation of the field of view, between air and water.*

What can be said about distortion ? How are air and water distortions related ?

Up to now we have not been able to determine in a theoretical way the mathematical relation between the two distortions, moreover it is doubtful if such a relation exists which does not involve the complete description of the lens systems. Even without any distortion, the image must be magnified with a factor 1.333.

Let u be the distorted image of a point in the air medium and du the distortion correction to obtain the perfect perspective projection. If now, in the same way u' is the distorted image of a point in the water medium and du' the new distortion correction, we must have:

$$1.333(u + du) = u' + du' \tag{12}$$

3 Camera Calibration

3.1 Protocol

This section briefly describes the camera calibration protocol we use in the experimental set-up in air as well as in water. The approach is mainly the photogrametric one [2] (i;e; *Bumble Adjustement*), and allows by the observation of a small number of views (approximately 10) the sensor calibration and the **calibration target reconstruction**. By this method, it is not necessary to take care in the calibration target measurement because, but the calibration sucess greatly depends on the accuracy of pattern detection in the calibration images set.The interested reader could refer to [6] for detailed issues.

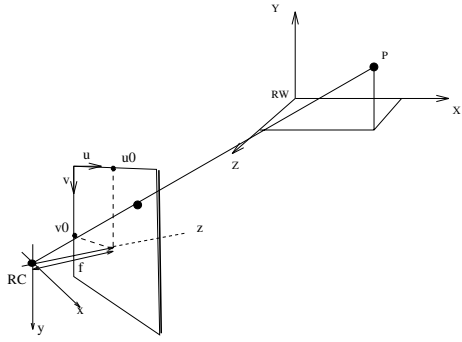


Fig. 4. *Coordinate systems.*

Under the perspective projection (pin-hole model), the relation between a point of an object and its image is given by the following expression:

$$\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \lambda_i \left[\boldsymbol{R} \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} + \boldsymbol{T} \right] \tag{13}$$

Where:

- (x_i, y_i, z_i) is a point defined in the camera frame (C.f. Figure 4 with $z(i) \equiv f$, i.e. the focal length of the camera,
- λ_i is a scale factor introduced when going from R^3 to R^2
- (X_i, Y_i, Z_i) are the coordinates of the target point in the world frame W–XYZ,

– (T_x, T_y, T_z) are the coordinates of the translation vector $\boldsymbol{T}$,
– $\boldsymbol{R}$ is the rotation matrix

Eliminating λ_i in (13) (and suppressing the i index for simplicity), we obtain the following expressions known as *colinearity equations* in photogrammetry:

$$\left.\begin{aligned} x &= f\frac{r_{11}X+r_{12}Y+r_{13}Z+T_x}{r_{31}X+r_{32}Y+r_{33}Z+T_z} \\ y &= f\frac{r_{21}X+r_{22}Y+r_{23}Z+T_y}{r_{31}X+r_{32}Y+r_{33}Z+T_z} \end{aligned}\right\}$$

If we express (x, y) in the image frame, we get:

$$\left.\begin{aligned} x &= (u + e_x - u_0)dx - do_x \\ y &= (v + e_y - v_0)dy - do_y \end{aligned}\right\} \tag{14}$$

In this expression e_x, e_y are the measure errors along coordinates x and y, (i.e., corrections to add to the measurements to fulfill the projection equations). do_x, do_y are the distortion components that can be split in two parts: *radial* and *tangential*, (i.e., $do_x = do_{xr} + do_{xt}$ and $do_y = do_{yr} + do_{yt}$).

The two following expressions are commonly used in photogrammetry [1], [3] and we will be adopt them:

$$\left.\begin{aligned} do_{xr} &= (u - u_0)dx(a_1r^2 + a_2r^4 + a_3r^6) \\ do_{yr} &= (v - v_0)dy(a_1r^2 + a_2r^4 + a_3r^6) \end{aligned}\right\} \tag{15}$$

$$\left.\begin{aligned} do_{xt} &= p_1[r^2 + 2(u - u_0)^2dx^2] \\ &\quad + 2p_2(u - u_0)dx(v - v_0)dy \\ do_{yt} &= p_2[r^2 + 2(v - v_0)^2dy^2] \\ &\quad + 2p_1(u - u_0)dx(v - v_0)dy \end{aligned}\right\} \tag{16}$$

In these expressions (14), (15), et (16),

– u, v are the image coordinates in the image frame,
– u_0, v_0 are the coordinates of the principal point in the image frame,
– a_1, a_2, a_3 are the radial distortion parameters,
– p_1, p_2 are the tangential distortion parameters,
– dx, dy are the sizes of the elementary pixel,
– $r = \sqrt{(u - u_0)^2dx^2 + (v - v_0)^2dy^2}$ is the distance of the image from the principal point.

Substituing (14), (15) and (16) in (3.1), we get the following system:

$$\left.\begin{aligned} u + e_x &= u_0 + (do_{xr} + do_{xt})/dx \\ &\quad + (\tfrac{f}{dx})\frac{r_{11}X+r_{12}Y+r_{13}Z+T_x}{r_{31}X+r_{32}Y+r_{33}Z+T_z} = P(\boldsymbol{\Phi}) \\ v + e_y &= v_0 + (do_{yr} + do_{yt})/dy \\ &\quad + (\tfrac{f}{dy})\frac{r_{21}X+r_{22}Y+r_{23}Z+T_y}{r_{31}X+r_{32}Y+r_{33}Z+T_z} = Q(\boldsymbol{\Phi}) \end{aligned}\right\}$$

so we have:

$$\left.\begin{aligned} e_x &= P(\boldsymbol{\Phi}) - u \\ e_y &= Q(\boldsymbol{\Phi}) - v \end{aligned}\right\} \boldsymbol{E}(\boldsymbol{\Phi}) \tag{17}$$

Perspective projection is always defined up to a scale factor. Conventionally, we put ($dx = 1$), then with $f_x = \frac{f}{d_x}$ and $f_y = \frac{f}{d_y}$, the parameter vector $\boldsymbol{\Phi}$ to estimate in the sensor/target joint calibration is:

$$\begin{aligned} \boldsymbol{\Phi}_{9+6m+3*n} = \Big[& u_0, v_0, a_1, a_2, a_3, p_1, p_2, f_x, f_y, \\ & X^1, Y^1, Z^1, \cdots, X^n, Y^n, Z^n, \\ & T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \cdots, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m \Big]^T \end{aligned}$$

Where n is the target number of points and m the number of images.

3.2 Initial Conditions

Optimisation of the non-linear system obtained is sensitive to the quality of the initial conditions. Generally the distortions coefficient are set to zero. The target is measured roughly (few millimeters); its planar structure eases the operation.

Initial locations of the camera in front of the target are estimated using Dementhon's algorithm [4] for planar objects. The principal point position is set at the image center, the focal length is set to the manufacturer estimate. Finally, the pixel size is set around 9 to 15 μm according to the camera manufacturer.

4 Experimentation

The experimental part consists of the self-calibration of an underwater camera in the two fluids air and water. We analyze results and try to express relations in regard with the previous theoretical developments.

4.1 Hardware

Two distinct cameras have been used for experiments. For the first one, the hardware system was an underwater camera made of a Sony CCD chip, a short focal length system and a special interface lens ensuring dry liaison between air and water. This lens allows an angular field of amplitude greater than 90 degrees. The whole video system is coupled with an automatic luminosity control device using two regulation loops, one for the mechanical regulation of the iris, the other being a gain controler for the video input signal.

The second one is also baseded on a sony CCD chip. However, all the optical system in this case, has been designed for experiments and the physical properties of each lens (index, size, position ...) is known. This permits a full optical simulation.

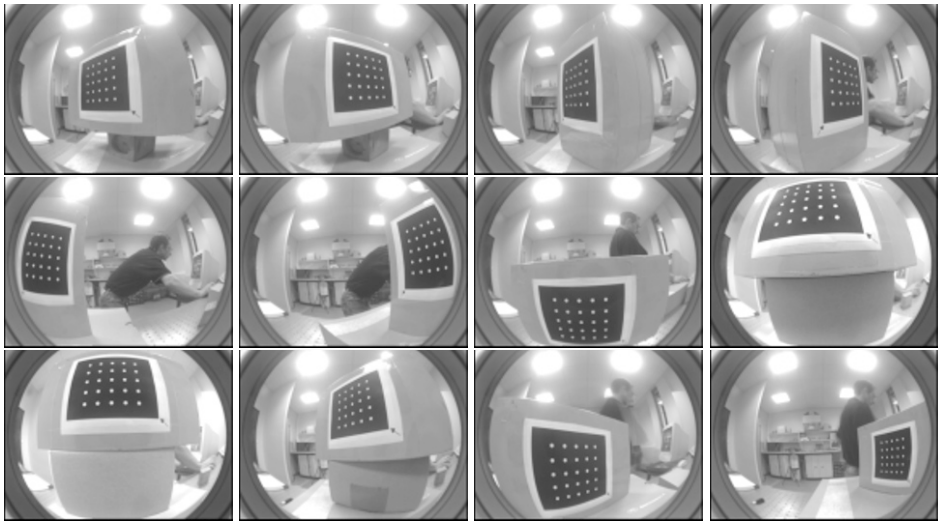


Fig. 5. Example of views for air calibration (768*576 pixels)

4.2 Air Calibration

For this first experiment we calibrate the system out of the water medium from twelve images of a plane target. Figure (5) shows some of the shots. The important radial distortion is to be noted. The dark circle around the image is due to the air medium and disappears when the camera is immersed in water.

Results : air calibration (media index = 1)
Under water camera 2

camera2	: Sony + full optical system
lens	: 4mm
digitizing card	: Silicon Graphics
algorithm	: Self-calibration

Number of images	: 12
Number of measures	: 283

Residuals mean e_x and σ (pixel)	2.69e-05	4.15e-02
Residuals mean e_y and σ (pixel)	1.49e-04	4.45e-02

		σ
fx(pix)	375.65	3.39e-01
fy(pix)	375.81	3.06e-01
u0(pix)	390.87	5.59e-02
v0(pix)	291.75	7.31e-02
a1	6.63e-01	7.52e-03
a2	-1.15e-00	5.02e-02
a3	6.83e+00	1.73e-01
a4	-1.20e+01	2.65e-01
a5	8.95e+00	1.56e-01
p1	-1.02e-03	1.98e-04
p2	1.22e-04	1.91e-04

Table 1. Self-Calibration in Air. Camera number 2

Notes:

- *The degree of the distortion polynomial has been increased in order to obtain a satisfactory fit with the image, due to this very distorted situation.*
- *The table (1) presents the computed values of the intrinsic parameters. It is to be noted that the residual at convergence is about* 0.04 *pixel along each*

coordinate: this is more than the usual results obtained with our method, which are turning at around 0.015 *pixels. They denote the difficulty of a very accurate detection of the target point in the image corner, and also the difficulty for the polynomial to fit such a large distortion.*
Nevertheless, the algorithm converges to a stable solution $f_x = 376$ *pixels and* $f_y = 376$ *pixels, corresponding to an observed angular field of* 110 *degrees when the distortion is compensated.*

– *The distortion polynomial can be seen in Figure 6. As we get further from the image center correction values become really high. As the target points measurements have been taken on a disk of radius of no more* 320 *pixels centered on the image (C.f. Figure 5), values of the distortion polynomial are mere extrapolation after this limit (Figure 6) and have no physical significance. The un-distorted view where the size has been increased by* 400 *pixels (line and columns) shows that inside the measurements field the distortion is properly corrected.*

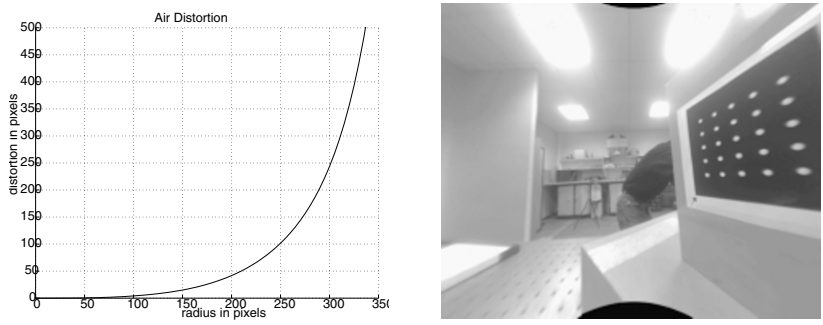


Fig. 6. radial distorsion and corrected image (1168*976 pixels)

4.3 Water Calibration

In a similar way we have calibrated the system underwater from 10 images of the target. Each point has been detected and matched along the sequence. As previously emphasised, the dark circle also desappears in images. It shows that the intrinsic parameters have been modified. We can also notice that the angular field has shrunk under water as the distortion.

Notes:

For this second experiment, the focal length has increased to 500 pixels and leads to an expected field of view in water close to 90 degrees.

We have also drawn the distortion curve. Also as expected, the point displacement is less than in air. Figure (7) presents views obtained with the underwater camera. In order to constrain triangulation angles involved in the self calibration algorithm, it can be noticed that the calibration target is observed from quite different points of view.

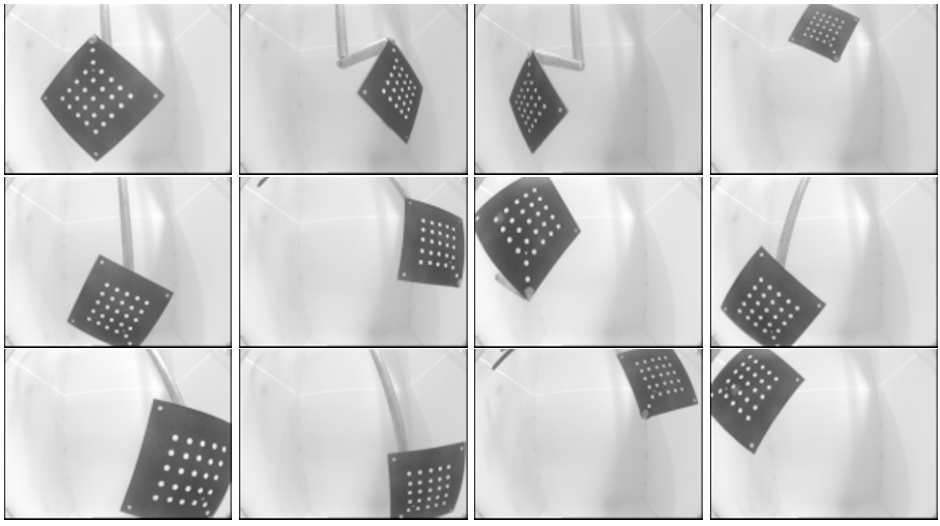
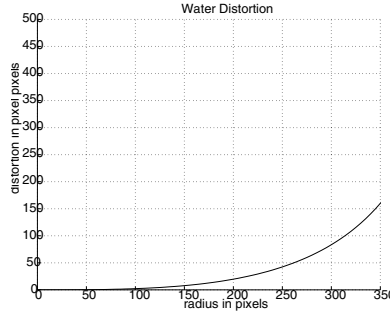


Fig. 7. Example of shots for underwater calibration, (768*576) pixels



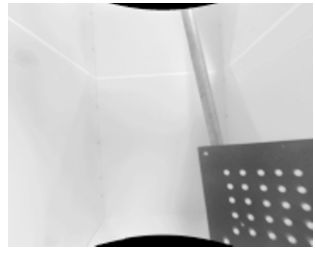


Fig. 8. radial distortion and corrected image in water, (968*776) pixels

Figure (8) shows one of the images after correction. The residuals remain at the same order of magnitude as in air. The black areas at the top and bottom correspond to image part that is not visible in the original view. (figure 7 number 10)

4.4 Relations between Water and Air Calibrations

Focal length: It appears that the theoretical relation is almost completely fulfilled:

The distance between the image nodal point and the CCD matrix is multiplied by the water index.

Results of water calibration (media index = 1.333)
underwater camera 2

camera 2	: Sony + full optical system
lens	: 4mm
digitizing card	: Silicon Graphics
algorithm	: Self-calibration

Number of images	: 13
Number of measures	: 325

Residuals mean e_x and σ (pixel)	1.36e-05	4.641e-02
Residuals mean e_y and σ (pixel)	-3.17e-05	5.164e-02

		σ
fx(pix)	499.12	5.51e-01
fy(pix)	501.97	5.22e-01
u0(pix)	391.81	9.27e-02
v0(pix)	292.30	1.25e-01
a1	7.52e-01	1.06e-02
a2	-1.84e-00	7.22e-02
a3	9.88e-00	2.54e-01
a4	-1.73e+01	4.11e-01
a5	1.30e+01	2.57e-01
p1	-1.66e-03	2.32e-04
p2	1.62e-03	1.76e-04

Table 2. Self-Calibration in Water. Camera number 2

	f-air	f-water	ratio (f-water/f-air)
fx(pix)	375.65	499.12	1.329
fy(pix)	375.81	501.97	1.336

Table 3. comparison of focal length in air and water media

Table (3) shows the ratio between the focal-length in water and air. If focal length uncertainties given by the calibration setup are taken into acccount, this ratio is almost $1,333$.

Self-calibration experiments have been carried out on many image sequences. The reproducibility of the results is ensured, but residuals at convergence are an order of magnitude larger than for calibration of a classical camera system.

(u_0, v_0) **location :**

	Air	Water
u0(pixels)	390.87	391.81
v0(pixels)	291.76	292.30

Table 4. comparison of principal point location in air and water media

The position of the image principal point (intersection of the optical axis and the CCD matrix) seems to remain unchanged between air and water calibrations (Table (4)).

However, it is well known that u_0 et v_0 are two parameters which are quite sensitive because they can (at a first order approximation) be compensated by object translation. The poor quality of the images delivered by the first underwater camera system has led to principal variation of up to 5 pixels.

Distortion:

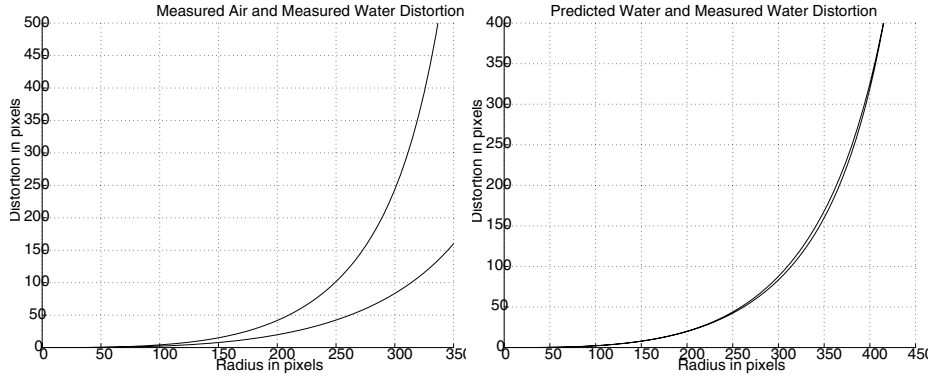


Fig. 9. Distortions: air and water curves (a), prediction of water distortion (b) from air data

Figure 9(a) shows a joint representation of the two distortion curves. As expected we can observe the large distortion in air medium compared with water measurements.

As seen in section 2.5, if we assume that the distortion is purely radial, the formula

$$1.333(u + d_a(u)) = u' + d_w(u')$$

would seem logical in order to give the natural relation between air and water distortion.

The right-hand part of Figure (10) shows the predicted water distortion compared to the measured one, according to the above expression. We can observe that the fit is quite good. Nevertheless, differences still remain if we try to overlap the two figures.

As the field of view in air is larger than in water, the predicted water distortion in the image border will correspond to very accurate data measured in the air sequence. This is not the case for the water sequence due to the difficulty in obtaining a full image in the image border.

5 Conclusion and Perspectives

This articles gives the basis of multi fluid sensor calibration and particularly air and water media. Different relations are shown between the index variation, the focal length, the field of view and the distortion function.

The fit between theoritical laws and measured data is almost completely fulfilled. For most applications that require an underwater camera, it is possible to carry out the sensor calibration in air and predict the intrinsic sensor parameters when the camera is immersed.

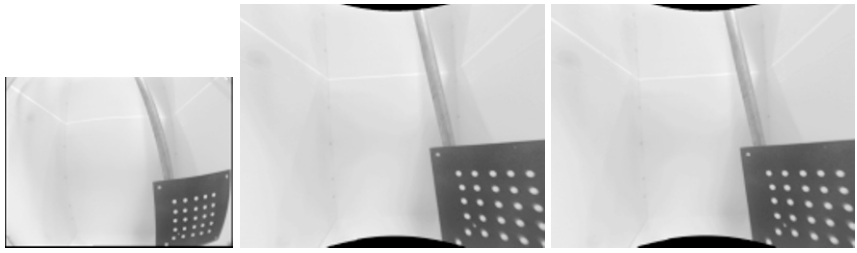


Fig. 10. Distortion: (a) Original view (768x576 pix), (b) Undistorted (968x776pix) from water data, (c) Undistorted from air data

To complete this study, we are working on a full optical simulation of the camera from the optical properties of each lens that compose the optical system. It will be possible to realize a more accurate comparison between simulated and real images.

Acknowledgements

This work has been done in collaboration with the french company POSEIDON [8]. The authors gratefully acknowledge all the staff for lending the underwater cameras.

References

1. American Society for Photogrammetry. *Manual of Photogrammetry*, 4th edition, 1984.
2. H.A. Beyer. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System.* PhD thesis, Institut fur Geodasie und Photogrammetrie, Nr 51, ETH, Zurich., May 1992.
3. D.C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 8(37):855–866, 1971.
4. L.S.Davis D.F.Dementhon. Model-based object pose in 25 lines of code. *International Journal of Computer Vison*, 15(2):123–141, 1995.
5. JM. Lavest, G. Rives, and M. Dhome. 3D Reconstruction by Zooming. *IEEE Trans. on Robotics and Automation*, Vol. 9(Nr 2):pp 196–207, April 1993.
6. JM. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration. In *Proc. of ECCV98*, pages 158–174, Freiburg, Germany, 1998.
7. J.Ph. Perez. *Optique géométrique et ondulatoire.* Masson,Paris, 1988.
8. Poseidon. *http://www.poseidon.fr.*