

Grouping and Establishing Correspondence of Calibration Holes for Underwater Machine Vision

E. Michaelsen, N. Scherer-Negenborn

FHG-IOSB, Gutleuthausstrasse 1, 76275 Ettlingen, Germany
eckart.michaelsen@iosb.fraunhofer.de

Underwater machine vision should start with careful calibration of the lens distortions. Inhomogeneous lighting and loss of contrast require special solutions. On the Fraunhofer/IOSB experimental platforms a calibration plate with a lattice of holes is used. Detection and grouping of these hole-objects from the images is presented. The discussion extends the investigation from the special application (calibration of distortions for underwater cameras) to more general problems of grouping and automatic gestalt perception.

Introduction

Machine vision under water should be aided by calibration of intrinsic camera parameters such as principle point, focal length and in particular distortion parameters. The measurements must be performed under water and should be repeated in advance of every campaign because the calibration is dependent on the optical density of the water, which may well vary with parameters such as temperature and salt content.



Fig. 1. Passive underwater inspection platform pasUWP of Fraunhofer-IOSB ready for launch on a slip ramp with calibration plate inserted

It is good advice to include removable calibration bodies or plates into the construction of underwater sensor platforms carrying visual sensors. Figure 1 shows an

example of a simple passive platform rolling under water on oblique infrastructure to be inspected. Two camera housings and two light sources are mounted pointing downward.

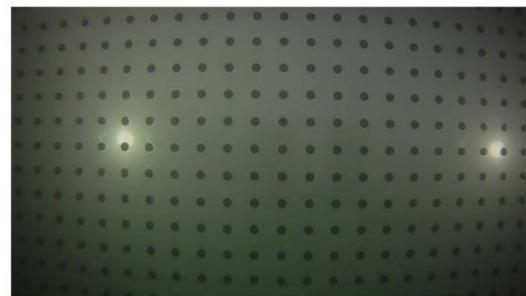
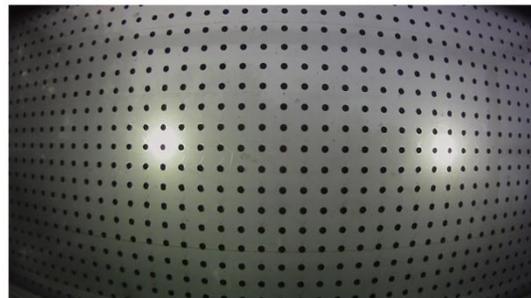


Fig. 2. Example pictures taken by pasUWP – above in air; lower in water

The calibration plate with a square lattice of holes slips in near the bottom of the vehicle and can easily be removed even while submerged. Figure 2 shows example images in air and water obtained from this platform. Notice, there is longer focal length and less distortion under water.

Related Work

Calibration: Most machine vision researchers active in the field emphasize the importance of lens distortion calibration – in particular for planar homography estimation. Standard state-of-the-art is the checker board method as presented by [7]. This is available as MatLab tool box or for free download from OpenCV. Such methods also estimate many other parameters such as image main point and focal length, etc. As point objects they use corner detection. For under water vision we are only interested in estimating a single distortion parameter and we have to cope with low contrast and inhomogeneous lighting. Therefore a specially tailored system is required. In [3] lens distortion models are discussed with particular attention to analytical un-distort functions.

Under water inspection: Vision is generally regarded as secondary sensor in the under-water community and thus people often have to get along with un-calibrated distortions. [1] discuss the possibility of calibrating lens distortions on such data afterwards. Some SLAM (simultaneous localization and mapping) work has been done for particular deep see biotopes [5] using the usual fundamental matrix and bundle adjustment methods.

Gestalt grouping: Lattice gestalts made of dot objects are among the favorite examples for the law of good continuation already presented in [6]. In these early publications the laws of perception are made plausible using the readers own perceptive apparatus on such dot patterns and illusory figures. Since the advent of machine vision methods are developed for the automation of gestalt perception. Latest advances are compared in a competition with benchmarks [2]. Emphasis is on lattice gestalts with perspective distortions. Lattice grouping with production systems has lately been described in [4]. There permanent scatterers in high-resolution space-borne SAR images and looking for facades was the application.

Correspondences for Calibration: A Grouping Task

The following four recognition steps are distinguished: Segmentation, clustering, lattice grouping, and establishing correspondence.

1) **Segmentation:** In images such as displayed in Fig. 2 the holes always appear as dark circular discs of roughly known diameter. Such objects can be extracted using the following standard image processing methods: 1) Convolution filter with a mask representing the expected target appearance (disk template in Matlab). 2) Pick local optima at the scale of the expected spacing of the images. Use a threshold for optimum acceptance and generate thus a set of *local-optimum objects*.

2) **Clustering:** Depending on the previous step an object of interest – here a hole – can be found uniquely or multiply. The method outlined above may give several local optimum objects clustered together in good separation from other such clusters. Figure 3 shows a section of our sample image with such clusters.



Fig. 3. Some detected local minima displayed as yellow crosses – hole-objects can be determined with sub-pixel accuracy

The clusters are found by a simple greedy clustering technique: A local-optimum object is picked and other such objects very close to it are added to its cluster and removed from the search list. The center of gravity of a cluster gives the new *hole-object's* location with sub-pixel accuracy.

3) **Lattice Grouping:** This can be treated by declarative definitions or by procedures. *Declaratively*, such grouping is handled e.g. by production systems [4]. In simple and special situations like this also a fixed *Procedure* can

be appropriate: First initialization - in particular of a generator vector – is required. Then stepwise prolongation – estimating the generator new from the observations – must be defined. Lens distortion and lighting problems are regarded low near the image center. Therefore, the grouping starts by picking the hole object closest to the image center (displayed in Figure 4 as red +) and its four closest neighbors in row and column direction. We call this central gestalt made of five hole-objects the *central cross*. Rather arbitrarily, in this work lattice grouping is performed row-wise in an inner loop and column-wise in the outer loop. The prolongation stops, when no hole-object is found where predicted (in the row) respectively no further row can be added to the column of rows. In Figure 4 such grouping can be seen as lattice of yellow crosses. Due to the high-light the grouping in the fourth row stops earlier than the other rows. One of the holes in the plate has not been found as hole-object.

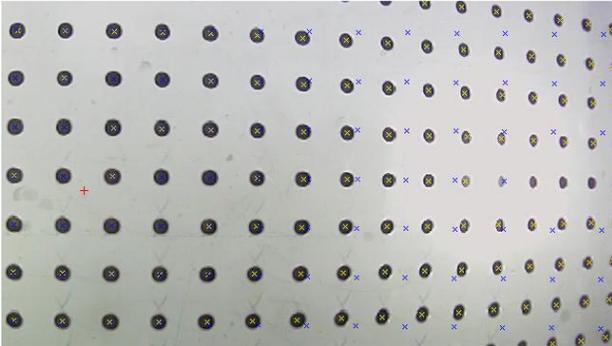


Fig. 4. Section of the example image (in air) with:
Lattice of distorted hole-objects (yellow x),
corresponding lattice of undistorted locations (blue x),
and image center (red +)

4) **Correspondence:** The *central cross* (five blue x in Figure 4 close to the red +) also sets the undistorted lattice. It is displayed as blue crosses in the Figure. With every hole-object appended to the distorted lattice also a corresponding node is appended to the virtual undistorted lattice. Thus correspondence is trivial. Such a pair of distorted and undistorted hole-objects is called a *correspondence-object*.

A Closed Form Distortion Model

It is standard to model the distortion function for the image height by a truncated polynomial series with odd powers, and here we hope to get along with only one coefficient giving already a positive curvature with $h > 0$:

$$h' = h + k_1 h^3 \dots \quad (1)$$

where h denotes the image height (distance from image center) in the distorted image and h' the image height in an ideal pin-hole projection. Fitting such a cubic least squares error sum to the empirical data is possible in closed form:

$$k_1 = \frac{\sum_i (h'_i - h_i) h_i^3}{\sum_i h_i^6} \quad (2)$$

where the (h'_i, h_i) are the correspondence-objects mentioned above. Breaking the sequence at k_1 leads to stable estimation even from low numbers of correspondences and an inversion that can also be given in closed form using Cardano roots. Provided

$$-4k_1 - 27k_1^2 h'^2 < 0 \quad (3)$$

two complex and one meaningful real root exists which can be obtained from

$$h = -\frac{1}{3k_1} \sqrt[3]{\frac{1}{2} (-r + \sqrt{r^2 + 108k_1^3})} - \frac{1}{3k_1} \sqrt[3]{\frac{1}{2} (-r - \sqrt{r^2 + 108k_1^3})} \quad (4)$$

where $r = 27k_1^2 h'$.

Such inversion is needed for production of re-sampled images such as presented in Fig. 4.

Discussion, Results, and Conclusion

Benign settings such as the one presented here hide the fact that grouping is a non-trivial task. The *cluster grouping* – according to the gestalt principle “proximity” can be regarded as benign if the size of the clusters is significantly smaller than their spacing (such as factor 10). Then a

threshold can be chosen in a not very sensitive way and a greedy cluster procedure can be used, as is done here.

The *lattice grouping* according to the principle “good continuation” can be regarded as benign if no clutter objects are present that may lead the prolongation astray. Here the calibration plate filling the picture has been made on purpose. Moreover the threshold for segmentation can be kept tight – losing some of objects near the margin in bad lighting does not matter much. So again a greedy strategy can be used. Completeness of the result is not important – i.e. no hypothesis and test procedure is needed to search for missing objects, no alternative grouping paths have to be added to the search. We should be aware that if grouping was column-wise first the loss of the hole-object in the (central) fourth row in Figure 4 would mean the loss of initialization of all following columns to the right – some less correspondence-objects less, but probably not a very different solution.

Correspondence is regarded as very benign when both gestalts are isomorphic (as graphs) and the mapping is known. This will be the case if both structures are built simultaneously in one greedy search sweep. If only one of the grouping steps was less benign more sophisticated search paradigms – e.g. such as outlined in [4] – would be appropriate.

The pictures presented in Figure 2 have been processed. Figure 5 shows the resulting residual error of the more distorted upper (in air) image.

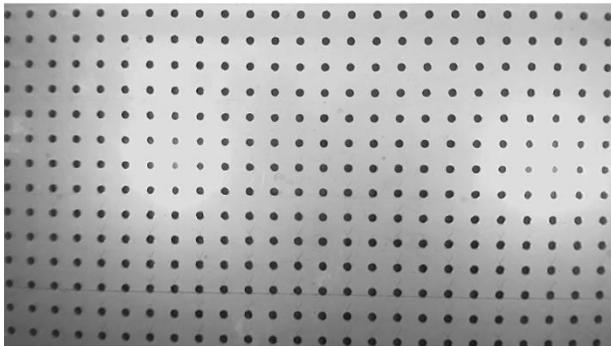


Fig. 5. Undistorted re-sampling of the upper image of Fig.2 using the estimated parameter k_l

The remaining projective distortions (slight rotations of the camera round the y- and z-axes

are not of interest here. Small non-projective distortions remain – as can be seen from the upmost row of holes with the image margin as reference. These can probably only be removed using higher order terms in (1). But the result as it suffices for planar homography estimation and panorama stitching.

References

1. A. Elibol, B. Moeller, R. Garcia. Perspectives of Auto-Correcting Lens Distortions in Mosaic-Based Underwater Navigation // Proc. of 23rd IEEE Int. Symposium on Computer and Information Sciences (ISCIS '08), Istanbul, Turkey - October 2008 - P. 1-6.
2. Y. Liu et al. The Symmetry Competition //, Pennsylvania State University - 2011 - <http://vision.cse.psu.edu/research/symmComp/index.shtml>
3. L. Ma, Y.Q. Chen, K.L. Moore. Rational Radial Distortion Models with Analytical Undistortion Formulae // arXiv:cs.CV/0307047 - Jul. 2003. – Vol. 1, No. 20.
4. E. Michaelsen, U. Soergel, A. Schunert, L. Doktorski, K. Jaeger. Perceptual grouping for building recognition from satellite SAR image stacks // PRRS 2010, IEEE Xplore - 2010. - http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5742807&tag=1
5. A. Sedlazeck, et al., 3D Modeling of Seafloor Structures from ROV-based Video Sequences // Univ. of Kiel - 2008 - <http://www.mip.informatik.uni-kiel.de/tiki-index.php?page=3DBlackSmoker>
6. M. Wertheimer: Untersuchungen zur Lehre von der Gestalt II // Psychologische Forschung. Vol 4, - 1923 <http://www.springerlink.com/content/u1w4k03566r51n68/fulltext.pdf>
7. Z. Zhang. A flexible new technique for camera calibration // IEEE PAMI, - 2000. - Vol. 22 , No. 11, P. 1330-1334.