

MULTI-CAMERA SYSTEM CALIBRATION OF A LOW-COST REMOTELY OPERATED VEHICLE FOR UNDERWATER CAVE EXPLORATION

E. Nocerino ^{a,b*}, M. M. Nawaf ^a, M. Saccone ^a, M. B. Ellefi ^a, J. Pasquet ^a, J.-P. Royer ^a, P. Drap ^a

^a Aix-Marseille Université, CNRS, ENSAM, Université De Toulon, LSIS UMR 7296,
 Domaine Universitaire de Saint-Jérôme, Bâtiment Polytech, Avenue Escadrille Normandie-Niemen, 13397, Marseille, France.
 erica.nocerino@univ-amu.fr

^b Institute of Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland – erican@phys.ethz.ch

Commission I, WG-7/WG-9

KEY WORDS: Photogrammetry, ROV, system calibration, relative orientation, synchronization, mobile mapping, underwater caves

ABSTRACT:

Exploration, documentation and mapping of underwater environment is one of the biggest open challenges for science and engineering. Humankind is not naturally designed to operate in water and, despite the enormous technological advancement that offers nowadays unprecedented opportunities, diving and working underwater is still very dangerous, especially in confined spaces such as underwater caves. Great research efforts are currently devoted to underwater autonomous navigation, but available solutions still mainly rely on complex and expensive systems, due to the difficulty of adapting localization and mapping sensors and algorithms suited for terrestrial or aerial applications. However, small and affordable underwater remotely operated vehicles (ROVs) are available, which offer good opportunities for underwater exploration and mapping. This paper focuses on the development of a small, low-cost ROV designed for 3D mapping of underwater environments, like caves. The system is based on a commercially available vehicle, the BlueRov2, and relies on the use of up to 12 action cameras (GoPro) mounted on it. A trifocal camera system for underwater real-time visual odometry can also be included. The work describes the photogrammetric procedure developed for the synchronization and calibration of the GoPro cameras and provides a thorough analysis on the achievable results.

1. INTRODUCTION

The ability of acquiring and recording precise, dense and geo-referenced 3D information is a constant request for a variety of applications, ranging from civil engineering and construction to cultural heritage, from environment to industry. When the area to be mapped is vast, mobile mapping systems (MMSs) are the optimal solution, allowing to derive 3D metric information of the environment while moving in it. Traditionally, MMSs are mounted on vehicles, i.e. cars, vans, but also planes and boats (Petrie, 2010). These systems are usually expensive and complex (Ellum and El-Sheimy, 2002), and are not suitable for mapping indoor and underground spaces. To overcome these limitations, portable MMSs have been developed both in the research and commercial domains and are becoming more and more popular for applications in scenarios characterised by harsh conditions (Lehtola et al., 2017; Nocerino et al., 2017; Tucci et al., 2018). While the last systems are designed to be carried around by an operator, ‘terrestrial’ robotic platforms (Nüchter et al., 2013; Ziparo et al., 2013) and autonomous aerial micro vehicles (Cieslewski et al., 2017) are under study with the aim of surveying complex scenes autonomously and, consequently, assuring the safety of human operators.

While it is undoubtful that there still is room for improvements towards more and more efficient 3D mapping and modelling approaches for ‘terrestrial’ scenarios, challenges are even more crucial when it comes to the underwater environment. It is estimated that about 95% of the ocean is unexplored (NOAA, 2018) and underwater caves represent probably the most fascinating and dangerous type of underwater exploration. Deep water applications commonly entail the use of expensive, huge and complex remotely operated (ROV) or autonomous underwater (AUV) vehicles. Recently, smaller and low-cost

systems have started to appear (e.g., BlueRov2, OpenROV, Sibiu Nano, etc.), revolutionizing the world of underwater investigation and opening unprecedented opportunities for researches in underwater caves, an environment which can be highly dangerous also for well-trained divers.

<i>Model</i>		BlueRov2
<i>Thrusters</i>		6 x Blue Robotics T200
<i>Lighting system</i>		4 x 1500 lumens each with dimming control Aquavolt Mini 5000 (2 x 5000 lumen)
<i>INS and additional sensors</i>		3-DOF Gyroscope 3-DOF Accelerometer 3-DOF Magnetometer Internal barometer Current and Voltage Sensing
<i>Camera</i>	<i>Integrated</i>	ELP-USBFHD06H-L36 Sensor: Sony IMX322LQJ-C Focal Length: 3.6 mm Underwater DOF (horizontal): 110° Light Sensitivity: 0.01lux Resolution: 1080 pixels
	<i>Trifocal system</i>	Sony IMX219PQ Focal Length 3.04 mm Sensor resolution 3280 x 2464 pixels
	<i>Additional</i>	Up to 12 x GoPro cameras (Hero5 Black and Hero6 Black)
<i>Control system</i>		Fathom-X Tether Interface Board M Robotics PixHawk

Table 1. Main characteristics of the modified version of the BlueRov2.

* Corresponding author

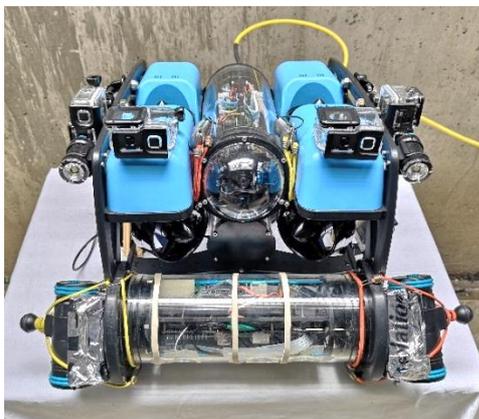


Figure 1. The BlueRov2 fully equipped for underwater cave exploration

1.1 Paper's aim and organization

Moving from the last considerations, this contribution presents the development of a small, low-cost ROV for 3D mapping of underwater cave based on photogrammetry. The system is a modified version of the commercially available BluRov2 (Table 1), equipped with a trifocal camera system (Nawaf et al., 2018) and with the possibility of mounting up to 12 GoPro action cameras (Figure 1).

The paper focuses on the methodology developed for the synchronization and calibration of the GoPro multi-camera system. In the current version, eight GoPro cameras, two Hero 5 Black and six Hero 6 Black, are mounted on the BluRov2.

The manuscript is organised as follows. First, an overview of related works is provided and, then, an in-depth analysis on the achieved results is presented. The paper concludes summarising the main outcomes, while further steps are will be required to combine the data from the GoPro cameras with the trifocal camera system and the inertial navigation unit embedded on-board the BluRov2.

2. RELATED WORKS

Action cameras have become very popular for photogrammetric applications thanks to their flexibility, lightweight, robustness and low cost.

Ballarin et al. (2015) used a GoPro Hero3 Black Edition on a small aerial drone to survey an archaeological site. The accuracy of orthophotos produced with the GoPro images was tested against check points measured with RTK-GNSS approach. Hastedt et al. (2016) investigated the reliability and accuracy of GoPro Hero4 for unmanned aerial vehicle (UAV) based photogrammetry, using different software packages, calibration setup, acquisition mode (i.e., single images vs video) and resolution.

A five-head GoPro system for indoor mapping purposes was proposed by Teo (2015). The five cameras were separately calibrated, their lever-arms and boresight angles estimated and video streams synchronised using an external timer. Keohl et al. (2016) presented a four GoPro Hero 4 multi-camera system for outdoor mobile mapping applications. A remote controller was used for the system synchronization, providing a not very accurate synchronization.

Schmidt & Rzhonov (2012) employed a stereo-camera systems based on two GoPro Hero2 for seafloor bathymetry measurements. The cameras set in video mode were synchronised using a synchronization cable. A similar stereo-camera system mounted on micro-ROV and combined with a

sonar scanner was proposed by Nelson et al. (2014) for surface reconstructions of an underwater archaeological site.

3. DEVELOPED APPROACH

The approach developed for the multi-camera system calibration, i.e. the computation of relative orientation (RORE) parameters between the cameras, is intended to be flexible. It follows the guiding principle of developing a procedure applicable in different environments, without the need of a dedicated facility. The main steps are summarised in Figure 2 and detailed in the following sections.

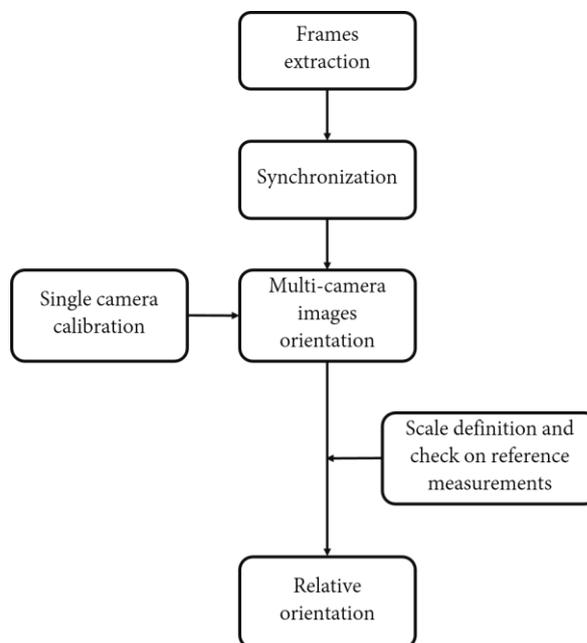


Figure 2. Flow-chart of the implemented procedure for the computation of RORE between the GoPro cameras.

The developed ROV system is designed to work in harsh underwater environments, characterised by narrow spaces and low light conditions. Consequently, the main idea is that the system is calibrated 'in-air' before the underwater mission and the computed RORE parameters are employed in the processing of the underwater dataset in two ways. The simplest consists in exploiting the RORE to infer a scale constrain between the camera centres. The second aims to speed up the processing, computing the approximate exterior orientation parameters of

slave cameras from the exterior orientation of the camera selected as *reference*.

For this specific application, where a high number of action cameras are used and the system is designed to operate underwater, a hardware-based synchronization approach is not feasible. Consequently, all the cameras are set in video mode and a post-processing software synchronization method is used, based on a well-recognisable common event visible in the single video streams.

4. PRELIMINARY STEPS

4.1 Scale definition and reference measurements

The multi-camera system calibration approach relies on the use of reference measurements for the definition of RORE parameters between the several cameras that compose the system.

Coded targets, arranged on planar sheets of known dimensions, are attached to a rigid structure; while the current experiment was performed in a basement (Figure 3), the procedure is easily exploitable in different closed spaces, like for example a small room or the back of a van.

In this study, a professional grade digital single lens reflex (DSLR) camera (Nikon D810) equipped with a 24 mm lens is used to measure the coordinates of 135 coded targets, with a final estimated accuracy of 0.5 mm. The self-calibrating image network, comprising 190 images, is shown in Figure 3, together with the calibration environment.

4.2 GoPro cameras set-up, video processing and frame extraction

To reproduce dark lighting conditions expected in real environments, the BlueRov2 lights are kept turned off and the GoPro cameras settings were configured to avoid motion blur in the video frames (*minimum exposure time* equal to $1/120$ seconds and 400 as minimum *ISO*). Following the outcomes of the experiments presented by Hastedt et al. (2016), the videos were recorded with the in-house GoPro camera fisheye distortion correction enabled (*FOV linear*), thus producing frames that can be assimilated to a central projective model. This choice is enforced by the evidence that underwater when flat ports are used, as for the GoPro cameras, the field of view (FOV) is limited by total internal reflection (Menna et al., 2016). The *video resolution* is set at 2.7K (2704 pixel x 1520 pixel) @ 30 frame per seconds (*fps*), corresponding to a nominal frame rate of 29.97 fps. The frames are extracted in the lossless png format from the original video stream using the FFmpeg free software (FFmpeg Development Team, 2010). The png frames are then converted in jpg at the highest possible quality, and the following exif tags are added using ExfiTool (Harvey, 2003): *Make*, *Model*, *FocalPlaneXResolution*, *FocalPlaneYResolution*, *FocalLength*, *FocalLengthIn35mmFormat*. Based on these parameters, photogrammetric software applications can automatically recognise images coming from different cameras and estimate the initial value for the principal distance.



Figure 3. Top (a) and later views (b, c) of the test area with the reference camera network

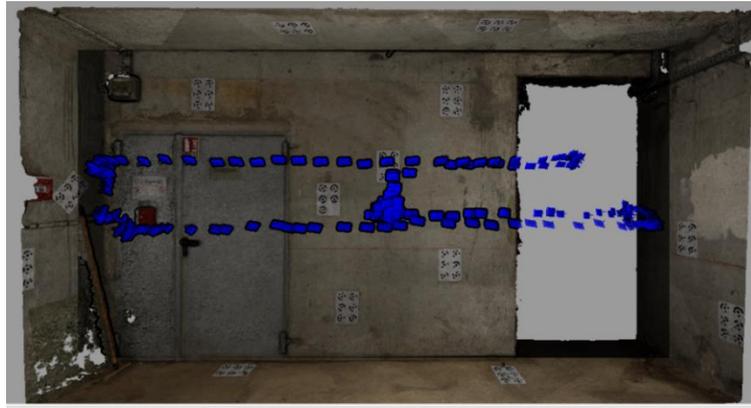


Figure 4. Typical self-calibrating network for the each GoPro

	1-GoPro6	2-GoPro6	3-GoPro6	4-GoPro6	5-GoPro5	6-GoPro5	7-GoPro6	8-GoPro6
Num. of images	262	313	258	258	309	306	268	196
Principal distance (mm)	2.2588	2.2633	2.2575	2.2591	2.2689	2.2527	2.2510	2.2594
σ_c (mm)	0.0001	0.0002	0.0001	0.0002	0.0002	0.0003	0.0001	0.0003
Principal point ppx (mm)	-0.0118	0.0005	-0.0057	-0.0434	0.0278	0.0099	-0.0421	-0.0065
σ_{ppx} (mm)	0.0001	0.0001	0.0001	0.0002	0.0001	0.0002	0.0001	0.0002
Principal point ppy (mm)	-0.0199	-0.0036	-0.0033	-0.0051	-0.0219	0.0101	-0.0131	-0.0144
σ_{ppy} (mm)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0001	0.0002
k1 (mm⁻²)	-5.75964e-004	-8.46643e-004	-5.91621e-004	-6.47151e-004	4.85256e-003	5.45661e-003	-3.47231e-004	-6.62475e-004
σ_{k1} (mm ⁻²)	4.97067e-006	7.18832e-006	3.59348e-006	1.04873e-005	1.42448e-005	2.40999e-005	4.89998e-006	1.20024e-005
k2 (mm⁻⁴)	-	-	-	2.77275e-005	-3.48628e-004	-3.43976e-004	2.16916e-005	-
σ_{k2} (mm ⁻⁴)	-	-	-	1.26451e-007	2.35563e-006	3.95520e-006	5.51442e-008	-
k3 (mm⁻⁶)	-	-	-	-2.96399e-006	-2.27144e-006	-2.33100e-006	-2.35193e-006	-
σ_{k3} (mm ⁻⁶)	-	-	-	1.33831e-008	3.59998e-008	6.96289e-008	5.96669e-009	-
P1 (mm⁻¹)	5.53300e-004	-1.78911e-004	1.97794e-004	2.19682e-003	-1.43115e-003	-6.12209e-004	2.09260e-003	2.46723e-004
σ_{P1} (mm ⁻¹)	5.17309e-006	7.89015e-006	3.90188e-006	1.01124e-005	6.06698e-006	1.16728e-005	4.41867e-006	1.33453e-005
P2 (mm⁻¹)	8.49805e-004	1.29264e-004	6.21069e-006	1.19171e-004	1.06867e-003	-5.32159e-004	6.74302e-004	6.24930e-004
σ_{P2} (mm ⁻¹)	5.18229e-006	6.01061e-006	2.71970e-006	7.65371e-006	8.26441e-006	9.29176e-006	3.61592e-006	1.01303e-005
Affinity B1	-1.39548e-003	-1.20837e-003	-1.29569e-003	-1.26928e-003	-1.78142e-002	-1.75087e-002	-1.49638e-003	-1.32110e-003
σ_{B1}	5.21193e-006	8.96003e-006	4.48993e-006	1.24481e-005	8.26441e-006	1.70022e-005	5.84069e-006	1.52401e-005
Shear B2	2.14500e-004	-3.84716e-006	-1.48764e-004	-1.00828e-004	-1.07170e-004	-1.50860e-004	-4.87838e-005	2.89534e-004
σ_{B2}	6.92324e-006	9.70481e-006	5.06160e-006	1.31367e-005	1.08494e-005	1.69015e-005	6.05329e-006	1.47391e-005
RMS re-projection error (pixel)	2.6	2.7	2.6	2.6	3.1	2.8	2.5	2.6
RMSE [on # check points] (mm)	0.9 [41]	0.6 [41]	0.7 [41]	0.8 [41]	1.2 [41]	0.7 [35]	0.9 [35]	0.6 [41]

Table 2. Results of self-calibrating bundle adjustment for the eight GoPro cameras. Interior orientation and additional parameters (correction coefficients) are reported along with internal assessment in image (root mean square - RMS of re-projection error) and object space (root mean square error - RMSE on check points).

4.3 Single camera calibration

The eight GoPro cameras are calibrated separately; a typical image network configuration is shown in Figure 4. The volumetric calibration testfield is composed of a vertical wall intersecting the floor and ceil.

From each video stream, the key frames are extracted, assuring a good image quality and including rolled imaging around the optical axis. Agisoft PhotoScan is used for the image processing.

Table 2 summarises the results of the calibration processing, where only the significant additional parameters are retained, and Figure 5 shows the distortion maps for the different cameras. Interestingly, the affinity term B1 is one order magnitude higher for the two GoPro 5 than the GoPro 6. The distortion maps feature very peculiar trends, with the two GoPro 5 showing a more pronounced vertical pattern.

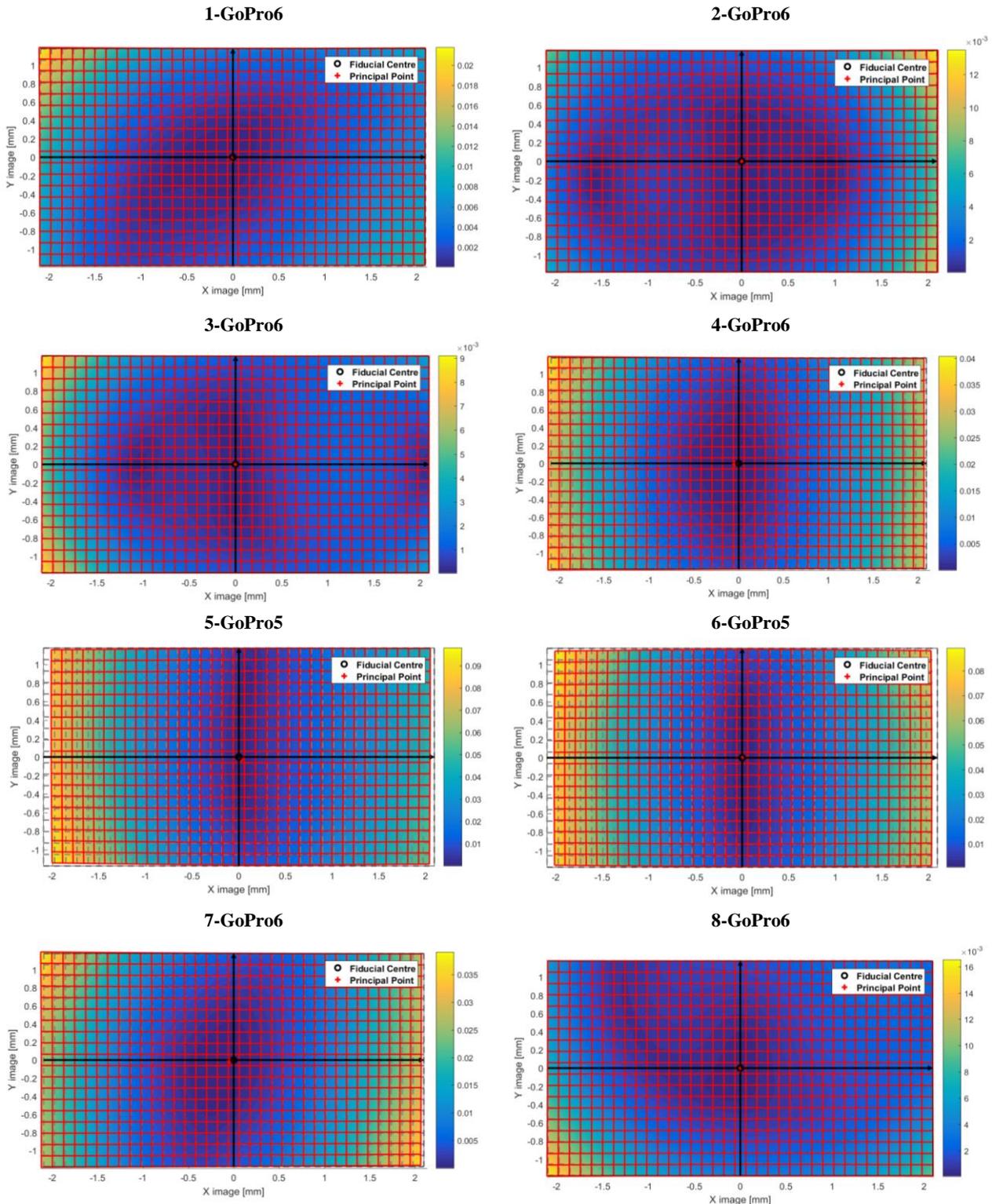


Figure 5. Distortion maps (difference in mm between ideal and actual distorted pixel position) for the eight GoPro cameras

5. EIGHT GOPRO CAMERA SYSTEM CALIBRATION

Once the single cameras are calibrated, they are mounted on the BlueRov2, which is positioned on a cart to avoid touching the cameras. After the system synchronization (section 5.1), video streams are acquired while the cart is moved in the calibration environment (Figure 6 a and b).

5.1 Synchronization

For the synchronization of the eight video streams, the system is put in complete darkness and a flash light is switched on and off (four times in this experiment). The frames are extracted at the original frame rate and, for each frame, the median intensity values is computed. The video streams are automatically synchronized by locating the peak (minimum and maximum) values in the times series of the median intensity difference (Figure 6c).

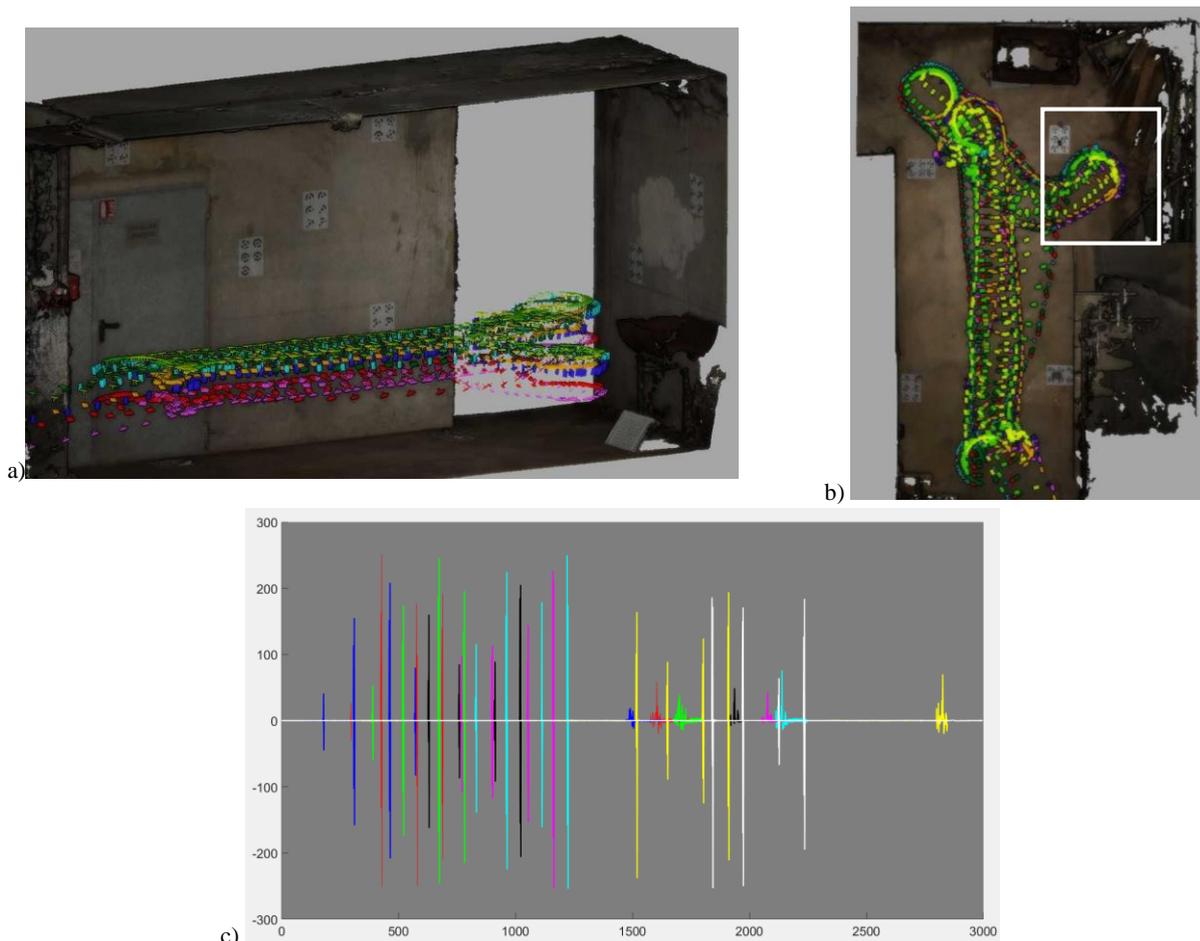


Figure 6. Camera network of the eight GoPro cameras system calibration (a, b). The white rectangle in b indicates the subset selected for the calibration approach 3 in section 5.2. (c) Automatic synchronization method based on the peak values of the median intensity values: four separate events are evident for the eight different cameras, represented by the different colours

	Self-calibration with approximate value from 4.3 [>2300 images]		Fixed pre-calibration from 4.3 [>2300 images]		Fixed pre-calibration from 4.3- subset [304 images]	
	FreeNet	Constrained	FreeNet	Constrained	FreeNet	Constrained
RMSE [on # check points] (mm)	1.6 [87]	1.0 [87]	1.6 [87]	1.1 [87]	7.1 [19]	1.2 [19]
Max baseline std [on camera pair] (mm)	6.1 [2-6]	5.2 [6-7]	2.7 [1-5]	2.6 [5-6]	4.6 [6-7]	4.2 [1-8]
Max baseline MAD [on camera pair] (mm)	2.4 [6-7]	2.2 [6-7]	2.2 [1-5]	2.1 [5-7]	4.0 [6-7]	3.6 [1-8]
Max omega std [on camera pair] (deg)	41.9 [2-6]	2.1 [4-7]	1.6 [5-7]	1.6 [5-7]	1.9 [5-7]	1.9 [5-7]
Max omega MAD [on camera pair] (deg)	9.8 [2-6]	0.9 [5-7]	1.2 [5-7]	1.2 [5-7]	1.5 [5-7]	1.5 [5-7]
Max phi std [on camera pair] (deg)	0.3 [1-7]	0.3 [1-7]	0.2 [1-7]	0.2 [1-7]	0.4 [4-8]	0.4 [4-8]
Max phi MAD [on camera pair] (deg)	0.05 [1-2]	0.04 [1-2]	0.02 [1-2]	0.02 [1-2]	0.08 [1-2]	0.08 [1-2]
Max kappa std [on camera pair] (deg)	3.2 [5-7]	2.2 [4-7]	1.6 [5-7]	1.6 [5-7]	1.9 [5-7]	1.8 [5-7]
Max kappa MAD [on camera pair] (deg)	0.6 [1-2]	0.4 [1-2]	0.3 [1-2]	0.2 [1-2]	0.3 [1-2]	0.3 [1-2]

Table 3. Results of multi-camera system calibration

The synchronization procedure with the flashlight is repeated at the end of the acquisition to verify that the synchronization is preserved. With the adopted approach, the maximum synchronization error can be equal to one frame, i.e. 1/29.97 (≈ 0.033) seconds that is the inverse of the actual frame rate.

5.2 Relative orientation

RORE parameters are computed for all possible camera pairs, at the each instant of time or epoch (t). According to equation 1, for an eight-camera system, 28 is the total number of RORE for each epoch t .

$$\forall t \in T: \# \text{ of } RORE_t = \frac{(\# \text{ of cameras})!}{(\# \text{ of cameras} - 2)!} / 2 \quad (1)$$

where T is the total test duration.

The camera showing with the smallest set of RORE std (or MAD) will be selected as *reference*.

For RORE computation, the definition and notation presented in Menna et al. (2013) are here adopted.

In the following equations, the letters G , R , S , indicate respectively the *global* (or *world*), *reference* camera and *slave* camera coordinate systems. The superscripts specify the frame where the quantity is defined, e.g. P^G is the position vector of a generic point P known in the *global* frame. The notation K_i^j expresses the transformation from the coordinate system $\{i\}$ to the coordinate system $\{j\}$. According to this rule, M_G^S is the rotation matrix from the *global* to the *slave* camera frame and O_S^G represents the origin of *slave* camera frame expressed in the *global* reference frame. Following this notation, the coordinates of the point P^G known in the *global* frame can be expressed in the *slave* camera reference frame as function of the exterior orientation parameters of the *slave* camera:

$$P^S = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^S = M_G^S \cdot \begin{bmatrix} X - X_{OS} \\ Y - Y_{OS} \\ Z - Z_{OS} \end{bmatrix}^G = M_G^S \cdot (P^G - O_S^G) \quad (2)$$

$$M_G^R = \begin{bmatrix} c \varphi c k & c \omega s k + s \omega s \varphi c k & s \omega s k - c \omega s \varphi c k \\ -c \varphi s k & c \omega c k - s \omega s \varphi s k & s \omega c k + c \omega s \varphi s k \\ s \varphi & -s \omega \cos \varphi & c \omega c \varphi \end{bmatrix} \quad (3)$$

$$O_S^G = \begin{bmatrix} X_{OS} \\ Y_{OS} \\ Z_{OS} \end{bmatrix}^G \quad (4)$$

where (3) and (4) are respectively the rotation matrix, containing the Euler angles, and the position vector of the *slave* camera perspective center in the *global* frame. Performing the RORE between the two cameras is equivalent to re-orienting the *global* reference system to be coincident with the *reference* camera frame. This transformation is expressed as:

$$P^S = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}^S = M_R^S \cdot \begin{bmatrix} X - X_{OS} \\ Y - Y_{OS} \\ Z - Z_{OS} \end{bmatrix}^R = M_R^S \cdot (P^R - O_S^R) \quad (5)$$

$$M_R^S = M_G^S \cdot (M_G^R)^T \quad (6)$$

$$O_S^R = M_G^R \cdot (O_S^G - O_R^G) \quad (7)$$

where (6) and (7) are respectively the rotation matrix (i.e., orientation or boresight angles) and coordinates of the *slave* camera in the *reference* camera coordinate frame. In other words, equation (7) represents the components of the baseline (or lever-arm) between the two cameras in the coordinate system centered on the *reference* camera. From the above transformation, the exterior orientation parameters of the chosen *reference* camera become null, and the exterior orientation of the *slave* camera with respect to the *reference* camera, i.e. the RORE, is obtained.

At each epoch t , the RORE of the eight GoPro cameras mounted on the BlueRov2 are estimated from the exterior orientation parameters. The mean and median values are then computed from the whole time series of duration T .

Starting from the last identified synchronization event (section 5.1), one frame per second is extracted from each camera video streams. All the extracted frames are processed together in Agisoft Photoscan, enforcing a camera-variant bundle adjustment, where a different set of calibration parameters is defined for each GoPro.

Three different approaches are tested:

1. the pre-computed camera calibration parameters (section 4.3) are used as initial values for a self-calibrating bundle adjustment, i.e. interior and exterior orientation parameters are estimated for the complete camera network shown in Figure 6 a and b;
2. the pre-computed camera calibration parameters (section 4.3) are kept fixed in the processing, i.e. only the exterior orientation parameters are estimated from the complete camera network shown in Figure 6 a and b;
3. as 2, i.e. only the exterior orientation parameters are estimated, but from a small subset of images (white rectangle in Figure 6 b), with the aim of simulating the camera calibration procedure in a small environment.

The coordinates of the coded targets previously measured with the reference photogrammetric system (section 4.1) are used to perform:

- a. a free network adjustment, where the coded targets are used to define the scale
- b. a constrained adjustment

for the three approaches described above.

The results of different processes are reported in terms of root mean square error (RMSE) on check points, maximum value of the standard deviation (std) and median absolute deviation (MAD) of the baseline (lever arm) and Euler (boresight) angles from the RORE between each camera pair (Table 3).

The RMSE on the check points improves for the constrained solution, especially in the case of the small image subset. The use of a robust statistic such as the MAD is highly significant for the omega angle when the self-calibration adjustment is performed. Figure 7 shows the mean and median baseline values with the associated std and MAD (error scale = 100). The robust statistic estimators show that the effect of residual outliers can be further mitigate, especially for approaches 1 and 3.

The absolute difference between the three different approaches, computed for the constrained adjustment solutions, is also reported, taking as reference the fixed calibration full dataset (Figure 8). It shows that the fixed-calibration full and subset approached provide more consist results, with difference below 2 mm.

¹ $c(\cdot) = \cos(\cdot)$, $s(\cdot) = \sin(\cdot)$

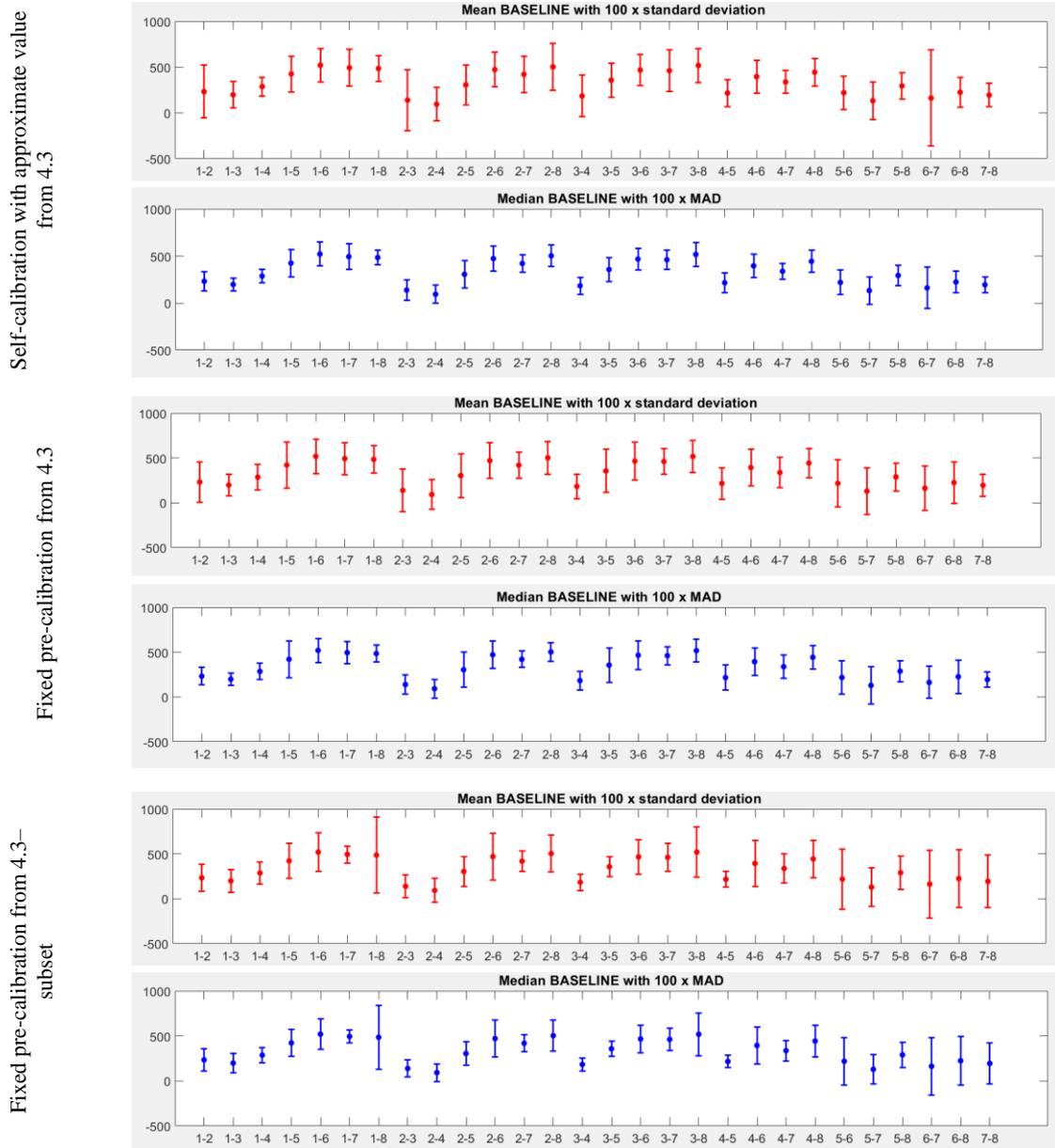


Figure 9. Mean and median baseline (lever arm) values (in mm) with the associated std and MAD (error scale = 100). The results are reported for the constrained solution.

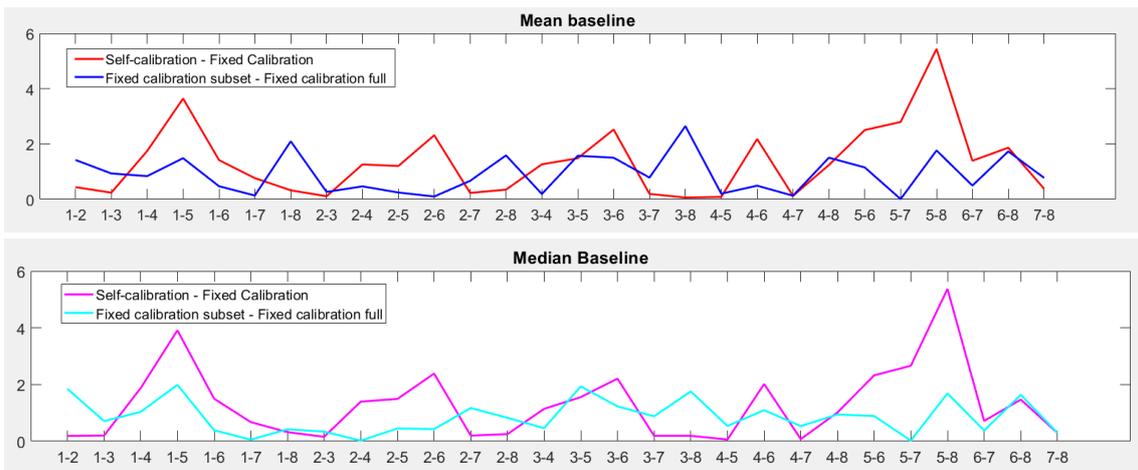


Figure 8. Absolute difference (in mm) between the three different approaches, taking as reference the fixed calibration full dataset. The results are reported for the constrained solution.

6. CONCLUSIONS AND OUTLOOK

The paper presented an in-depth analysis of the calibration of a multi-camera system, composed of eight action cameras (GoPro) mounted on a low-cost ROV. The system is designed as MMS for the photogrammetric survey of underwater caves.

The multi-camera system calibration approach entails a three-step procedure, consisting in the (i) measurement of a reference calibration environment, (ii) estimation of interior orientation parameters for each single camera, (iii) computation of the RORE (lever arm and boresight angles) between the multiple cameras.

The calibration results of the different GoPro cameras show that the interior orientation parameters are significantly different, while similar patterns in cameras of the same model can be identified.

The pre-calibration of the single camera provides results with smaller std (and MAD) in the estimation of the RORE. Normal and robust statistical parameters show that the effect of residual outliers can be further mitigate. Decreasing the number of images and reducing the space for the calibration leads to a difference in the RORE baseline less than 2 mm.

The next steps will involve the inclusion of the trifocal sensor and the inertial navigation unit in the system calibration, and tests in underwater environment.

ACKNOWLEDGEMENTS

The work described in this paper was partially supported by the French Single Inter-Ministry Fund (FUI), in the framework of the GREENEXPLORER project, and the EU H2020 project 'iMareCulture' (grant agreement No 727153, <https://imareculture.weebly.com>).

The authors thank Dr. Fabio Menna (3DOM, FBK) for the useful discussions and valuable comments.

REFERENCES

Agisoft Photoscan, Version 1.3.4 Professional Edition. <http://www.agisoft.com/> (last accessed, July 2018).

Ballarin, M., Balletti, C. and Guerra, F., 2015, June. Action cameras and low-cost aerial vehicles in archaeology. In *Videometrics, Range Imaging, and Applications XIII* (Vol. 9528, p. 952813). International Society for Optics and Photonics.

BlueRov2, <https://www.bluerobotics.com/>

Cieslewski, T., Kaufmann, E. and Scaramuzza, D., 2017. Rapid Exploration with Multi-Rotors: A Frontier Selection Method for High Speed Flight. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (No. EPFL-CONF-232505).

Ellum, C.M. and El-Sheimy, N., 2002. Land-based integrated systems for mapping and GIS applications. *Survey review*, 36(283), pp.323-339.

FFmpeg Development Team, 2010. Software version 3.4, <https://www.ffmpeg.org/> (last accessed, July 2018).

Harvey, F. 2003. ExifTool. Software version 10.71, <https://www.sno.phy.queensu.ca/~phil/exiftool/> (last accessed, July 2018).

Hastedt, H., Ekkel, T. and Luhmann, T., 2016. Evaluation of the quality of action cameras with wide-angle lenses in UAV photogrammetry. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41.

Koehl, M., Delacourt, T. and Boutry, C., 2016. Image capture with synchronized multiple-cameras for extraction of accurate geometries. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41.

Lehtola, V.V., Kaartinen, H., Nüchter, A., Kaijaluoto, R., Kukko, A., Litkey, P., Honkavaara, E., Rosnell, T., Vaaja, M.T., Virtanen, J.P. and Kurkela, M., 2017. Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods. *Remote Sensing*, 9(8), p.796.

Menna, F., Nocerino, E., Fassi, F. and Remondino, F., 2016. Geometric and optic characterization of a hemispherical dome port for underwater photogrammetry. *Sensors*, 16(1), p.48.

Nawaf, M.M., Merad, D., Royer, J.P., Boi, J.M., Saccone, M., Ben, M.E. and Drap, P., 2018. Fast Visual Odometry for a Low-Cost Underwater Embedded Stereo System†. *Sensors* (Basel, Switzerland), 18(7).

Nelson, E.A., Dunn, I.T., Forrester, J., Gambin, T., Clark, C.M. and Wood, Z.J., 2014, January. Surface reconstruction of ancient water storage systems an approach for sparse 3D sonar scans and fused stereo images. In *Computer Graphics Theory and Applications (GRAPP), 2014 International Conference on* (pp. 1-8). IEEE.

NOAA, 2018. How much of the ocean have we explored? National Ocean Service website, <https://oceanservice.noaa.gov/facts/exploration.html>, 16/04/2018

Nocerino, E., Menna, F., Remondino, F., Toschi, I. and Rodríguez-González, P., 2017, June. Investigation of indoor and outdoor performance of two portable mobile mapping systems. In *Videometrics, Range Imaging, and Applications XIV* (Vol. 10332, p. 103320I). International Society for Optics and Photonics.

Nüchter, A., Elseberg, J. and Borrmann, D., 2013. Irma3D—An intelligent robot for mapping applications. *IFAC Proceedings Volumes*, 46(29), pp.119-124.

Petrie, G., 2010. An introduction to the technology: mobile mapping systems. *Geoinformatics*, 13(1), p.32.

OpenROV, <https://www.openrov.com/>

Schmidt, V.E. and Rzhanov, Y., 2012. Measurement of micro-bathymetry with a GOPRO underwater stereo camera pair. In *Oceans*, 2012 (pp. 1-6). IEEE.

Sibiu Nano, Nido Robotics, <https://www.indiegogo.com/projects/sibiu-nano-affordable-portable-underwater-robot-drones#/>

Teo, T., 2015. Video-Based Point Cloud Generation Using Multiple Action Cameras. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(4), p.55.

Tucci, G., Visintini, D., Bonora, V. and Parisi, E.I., 2018. Examination of Indoor Mobile Mapping Systems in a Diversified Internal/External Test Field. *Applied Sciences*, 8(3), p.401.

Ziparo, V.A., Zaratti, M., Grisetti, G., Bonanni, T.M., Serafin, J., Di Cicco, M., Proesmans, M., Van Gool, L., Vysotska, O., Bogoslavskiy, I. and Stachniss, C., 2013, October. Exploration and mapping of catacombs with mobile robots. In *Safety, Security, and Rescue Robotics (SSRR)*, 2013 IEEE International Symposium on (pp. 1-2). IEEE.