# INTRODUCTION AND VALIDATION OF A NOVEL CALIBRATION FRAME

A. Mufti[1,2]*, P. Helmholz [1], I. Parnum [1], D. Belton [1]

[1] School of Earth and Planetary Sciences, Curtin University
GPO Box U1987, Perth WA 6845, Australia, alaa.mufti@postgrad.curtin.edu.au, (Petra.Helmholz, I.Parnum, D.Belton)@curtin.edu.au
[2] King Abdulaziz University, Jeddah 21442, Saudi Arabia

**Commission II WG II/7**

**KEY WORDS:** Calibration, Photogrammetry, Portability, camera calibration, self-calibration

**ABSTRACT:**

The use of Photogrammetry is increasingly used by several disciplines, including marine science. Among others, the accuracy of a 3D model generated from images, depends on the quality of the calibration and stability of the camera used to capture the images. For the calibration an optimum 3D geometry is essential to minimise correlations between the camera's interior orientation parameters. For the calibration, usually various different types of calibration frames are used. However, in practice, it can be challenging to use these frames when working in underwater environments. Calibration frames can be bulky, which makes them difficult to handle and transport, especially in boats where space is at a premium. This study aimed to develop a collapsible (and thereby portable) calibration frame, which is more practical for marine field data capture. The proposed collapsible calibration frame is validated in-air and underwater. Overall, three tests are performed. Firstly, the reliability of the frame is validated, i.e. if the collapsible frame can be put together in such a way that the Ground Control Points (GCPs) on the frame have unchanged positions relative to each other. The test showed a very small bias which could be removed by changing to a baseline assessment. Secondly, repeatability is validated, i.e. if the same results can be achieved for different software and camera combinations when using the same baselines. The test showed a clear downwards trend of the results for lower-grad cameras. However, all adjustments using the different software solutions and cameras show that the frame is suitable for application in-air. The final test is an underwater performance test which verified that the frame is usable achieving root-mean-squared error values of below 2 mm when using baselines.

## 1. INTRODUCTION

In recent years, there has been an increase in the use and development of marine resources, including the growth of aquaculture, deep-sea mining, shipping, and tourism. At the same time, observation, monitoring, and understanding of the marine environment are central topics of marine research that have received worldwide attention recently (Yuan et al., 2022).

The development and improvement of camera sensors and optical components have made it possible for underwater photogrammetry to become more accessible and affordable in recent years. As a result, non-professionals can now use off-the-shelf underwater photogrammetry action cameras with waterproof housings, such as the GoPro camera, to create 3D models of underwater areas. This has opened up new possibilities for a wide range of applications, including marine biology (Jalal et al., 2020), reef mapping (Guo et al., 2016) archaeology (Diamanti et al., 2017), oceanography, and more.

Before images acquired by a camera can be utilised to retrieve geometric information of a structure, distortions in the camera and lens must be addressed (Shortis, 2019). Hence, a good and reliable 3D model result that can be generated from these action cameras depends on the calibration quality and the system's stability. For instance, the refractive index of water is different to air. It is known that the refraction index of water varies with depth, and the complete light path must be coordinated, including the camera lens, housing port, and water medium (Shortis, 2019). Incorrect calibration parameters can create a doming effect known as the "bowing effect". This effect creates a scooping or bowing in the model's centre. This issue can be overcome or reduced by camera calibration (Samboko et al., 2022).

Hence, a properly calibrated underwater camera system is crucial for accurate and consistent 3D object measurement. One way to achieve this is through self-calibration (Fraser et al., 1995). The image quality, geometry, and redundancy of the calibration image network are critical factors that impact the reliability and precision of camera calibration in underwater photogrammetry. To ensure reliable calibration, several criteria have been proposed, including:

1. the use of three-dimensional camera and target arrays,
2. the acquisition of different convergent camera views of the targets,
3. the filling of the camera's field of view with the calibration fixture or range, and
4. the capture of different rotations of the camera(s) around the optical axis.

In practice, though, it can be challenging to meet all of these requirements, especially when working in underwater environments (Guo et al., 2016). Efforts should be made to meet as many of these criteria as possible to ensure the accuracy and reliability of the camera calibration. Therefore, usually a calibration frame is used and placed in the field of view of the camera. The diver then carries out "flying orbits" around this frame in a distance allowing to fulfil all the requirements listed

* Corresponding author

above (Helmholz et al. 2016). This procedure entails capturing multiple converging images of a calibration frame with known locations of target points such as circular dots. Alternatively, the checkerboard method is used where the target points are replaced with the corners of the checkerboard. Coded targets or image analysis techniques to automatically extract the locations of the target points such as through centroid fitting methods can improve efficiency (Shortis and Seager, 2014). The scale of the 3D measurement space is established by incorporating known distances between targets or by using the 3D coordinates relative to a target reference system.

A downside of calibration frames, is they can be bulky and difficult to handle especially underwater. This has been identified as an issue by marine scientists and the request was made to overcome this challenge. This study aims to introduce a new calibration frame, created to make it more practical for marine field data capture, and to test if the frame fulfils the photogrammetric (accuracy and repeatability).

The paper is structured as follows: in the next section, a literature review is presented to quantify the level of accuracy which is required for underwater marine photogrammetry after this, the new developed frame is introduced; this is followed by details of the methodology used to assess the frame; followed by the results of the assessment; and the paper ends with study's conclusions.

## 2. RELATED WORK

Calibration is a crucial component of underwater photogrammetry, because it is essential to correct for camera parameters, including distortions. In underwater photogrammetry, the significance of calibration cannot be overstated. Without precise calibration, the resulting 3D models may contain errors, which can have severe consequences for scientific research, marine conservation, and underwater engineering. For instance, if photogrammetry is used to assess the growth of coral, then distortions can skew distance observations and consequently derived growth information.

The calibration stability of underwater camera systems is highly dependent on the relationship between the camera lens and the housing port. It is crucial to rigidly mount the camera in the housing to maintain a consistent total optical path from the image sensor to the water medium. Studies have demonstrated that reliable calibration can only be achieved when the camera inside the housing is securely connected to the camera port (Shortis & Harvey, 1998; Shortis et al., 2000). Testing and validation have confirmed the importance of a rigid connection for accurate calibration.

Under optimal in-air conditions, the typical root-mean-squared error (RMS) of image observation error for control points have a range from 0.03 to 0.1 pixels (Shortis et al., 1995). However, when operating underwater, the RMS error degrades due to light attenuation, contrast loss, and minor non-uniformities in the medium, resulting in a range of 0.1 to 0.3 pixels. This degradation is a result of both increased statistical variations in image measurements and the influence of uncompensated systematic errors. In situations with poor lighting or visibility, the RMS error deteriorates rapidly (Wehkamp and Fischer, 2014).

The proportional error is another measure used to assess the calibration results of cameras. It is calculated by taking the ratio of the RMS error in the 3D coordinates of the targets to the largest dimension of the object. This indicator measures how well the camera's calibration performs, relative to the size of the object being captured. The average figure is approximately 1:5000 for underwater environments (Shortis, 2019).

Two main types of structures have been used to calibrate cameras for underwater photogrammetry: 2D checkerboards and 3D calibration frames. For instance, Bouguet (2017) used a small 2D checkerboard that had 36 squares: 18 black and 18 white. After the initial calibration, camera parameters were calculated in two subsequent steps: the initialization step and the optimization step. However, additional calibrations were required to obtain a pixel error of 0.19534. It is noteworthy, that this low error was achieved with a high-quality camera. Using a 2D checkerboard offers significant advantages, including the simplicity of the calibration fixture and the efficient measurement and processing of captured images. These benefits are achieved through the automated recognition of the checkerboard pattern (Zhang 2000), which enables swift and accurate analysis. The reliability and accuracy of measurements obtained through the checkerboard technique are constrained by the compact size and two-dimensional characteristics of the checkerboard. This method is more akin to a fixed test range calibration rather than a self-calibration, as the coordinates assigned to the checkerboard corners remain static. It is important to consider that inaccuracies in these coordinate values can significantly influence the calibration process, especially when the checkerboard deviates from a true two-dimensional plane. Such deviations have the potential to introduce systematic errors, thereby impacting the overall accuracy of the calibration results (Bouguet, 2017). In addition, the 2D checkerboard is a plane, which can lead to large correlation values between the camera calibration parameters (Shortis, 2019). For these reasons, a checkerboard was not used for this research.

Gourgoulis et al. (2008) employed a calibration frame which actually consisted of two aluminium frames: a large frame and a small frame. The dimensions of the large frame were 1m x 3m x 1m, and the dimensions of the small frame were 1m x 1m x 1m. Both frames had 32 designated points with known coordinates. In the case of the large frame, each of its eight vertical rods was marked with four points. Three of these served as control points for the calibration of the space, while the remaining fourth were used as Check Points to validate the calculations. For the small frame, six control points and two Check Points were inscribed on each of the four vertical rods. Hence, overall, both frames together carried 24 control points and 8 check points. They achieved an RMS in air for the small frame was 3.70 mm and for the large frame 4.66 mm. The RMS underwater for the small frame was 4.5 mm and the large frame was 5.92 mm.

Challis and Kerwin (1992) utilised a frame measuring 1.0 m x 0.6 m x 1.0 m, with the diminutive dimensions guaranteeing stability, to calibrate two single-lens reflex (SLR) 35 mm cameras: a Canon EOS 750 and a Canon EOS 620. The calibration structure was designed to include control locations throughout the calibration space. For the structure, twelve-millimetre-diameter steel tubing was used, with fifty 42-millimetre-diameter cylinders with central holes firmly attached. Spheres were utilised because they would be identifiable from any angle. Black matte paint was applied to the frame to reduce reflections. The frame provided a total of 51 control points, with the additional point located in the centre of the central cross. Using a laser-based surveying system, the locations of the control points were ascertained. The RMSE was 0.8 mm.

Helmholz et al. (2016) utilised a GoPro Hero 3 and an open cube calibration frame with a dimension of 60cm x 60cm x 50cm.

Along two planes, the calibration frame has 52 reference marks. 25 Ground Control Points (GCP) and 27 Check Points (CP) were randomly selected ensuring a god distribution for both sets of points. The aim of the paper was to assess the camera stability by performing different camera resolutions (7 MB and 12 MB) and different camera captured (e.g. shaking the camera between captures or removing the camera from the water tight housing between captures). The maximum RMSE from the tests carried out, was 0.45 mm for the 7 MB camera, and 2.5 mm for the 12 MB one.

Capra et al. (2015) utilised a frame made of PVC bars that form the approximate margins of a parallelepiped measuring 0.90m x 0.20 m x 0.15 m and weighing 3 kilogrammes. 34 Ground Control Points (GCP) and the targets are signalling with a 30 mm wide circles, alternately black and cross-printed rectangular target. All targets materialised on the frame have been numbered and measured, and their x, y, and z coordinates have been determined in an on-frame reference system. There were three cameras used: the Canon PowerShot G12 with an RMS of 0.524 mm, the GoPro Hero2 with an RMS of 43.037 mm, and the Intova Sport HD with an RMS of 11.33 mm. Both the GoPro Hero2 and Intova Sport HD have shown large RMS values for underwater environments. This is due to the strong distortion caused by the lenses with a
very small focal length.

Li et al. (1997) used a rectangular aluminium frame with a dimension of 1.4 m x 1.4 m x 0.7 m with 24 control targets marked with highly reflective circular discs with an 8-cm diameter and a well-defined centre with an accuracy of 0.8 cm along the X and Y direction and 1.2 cm along the Z with an overall RMS of 10.83mm.

| Manuscript | Frame design | GCPs/CPs | RMS |
|---|---|---|---|
| Gourgoulis et al. (2008) | large frame (1 m x 3 m x 1 m) in combination with a small frame (1 m x 1 m x 1 m) | 24 GCPs<br>8 CPs | In air:<br>3.7 mm – 4.66 mm<br>Underwater:<br>4.5 mm<br>5.92 mm |
| Challis and Kerwin (1992) | 1.0 m x 0.6 m x 1.0 m , | 51 GCPs | 0.8 mm |
| Helmholz et al. (2016) | 60 cm x 60 cm x 50 cm | 25 GCPs<br>27 GCPs | 0.45 – 2.5 mm |
| Li et al. (1997) | 1.4 m 1.4 m 0.7 m | 24 GCPs | 10.83 mm underwater |
| Capra et al. (2015) | 0.90 m x 0.20 m x 0.15 m | 45 GCPs | -Hero2: 43.03 mm<br>-Intova Sport HD: 11.330 mm<br>-Canon PowerShot G12: 0.524 mm |

**Table 1:** Summary of results achieved by selected calibration frames.

In conclusion, the design of calibration frame is usually a ridged construction, with dimensions of 1 m or more, which make it challenging to transport the frame. The number of GCPs and CPs is usually around 25. All points are equally distributed. RMS of 0.5 mm – 3 mm are considered acceptable for calibration outcomes in underwater environment (Shortis et al., 1995; Helmholz et al., 2016). The RMS is impacted by the camera used, the layout and geometry used to capture the images and finally the calibration frame. This covers a wide range of applications. The aim is to achieve similar RMS value when using the proposed calibration frame. Finally, while Shortis et al., (1995) define the range of acceptable point referencing error (PRE) with 0.1 to 0.3 pixels, our experiences show that 0.5 pixels are acceptable in air and 1 pixel is acceptable underwater (Shortis et al., 1995).

## 3. PROPOSED FRAME

The proposed solution is a triangular pyramid shape constructed with six metal, square-shaped bars (3 cm x 3 cm x 138 cm) and four 3D-printed apex joints connected with long screws and bolts (Figure 1). The joints ensure that the frame can be ensembled again in the same shape. The design was found to be light in weight approximately 3 kg, and it can be changed into two forms: a pyramid shape used for calibration, and two cylindrical shapes (made with three bars each) that easily fit into transportation tubes. White (paper), circular GCPs (diameter of 24 mm) were stuck on the bars (Figure 1). Each bar has a total of 20 dots (four on two of the sides and six each on the other two sides), which results in a total of 120 GCPs that could potentially be used. To distinguish and orientate the different bars of the pyramid, there are labels in the middle and the end of each bar (Figure 1). However, sometimes the labels may become blurred and difficult to read underwater, especially when the pyramid is in motion.

**Figure 1**. Proposed calibration frame assembled, with a scale bar (899.954 mm) inside it.

## 4. METHOD

### 4.1 Data collection

First the GCPs and CPs (in a local system) were determined in-air using a high-resolution SLR camera, namely a Nikon D750 SLR camera with a Nikon Zoom-Nikkor 24-70mm f/2.8 ED G AF-S Lens. Each image was captured with a fixed focus set and no alterations to the camera settings during data collection. The distance and height of the object were kept constant. For the data capture in air, a scale bar was placed together in the centre of the calibration frame. An example capture layout is provided in Figure 2 with the scale bar being highlighted in red. The calibration frame and image capture fulfilled all four criteria required for reliable calibration formerly mentioned.
For the underwater image capture, two cameras have been used: Canon G7X (Canon, 2014) and a GoPro HERO5 Black (GoPro, 2019) in Table 2. Both are low-cost cameras suitable for photogrammetric applications underwater. These are the cameras

which will be utilised for further research and therefore being the focus of this investigation. The GoPro 5 Black camera was operated in the linear mode.

| | Canon G7X | GoPro 5 Black |
|---|---|---|
| Dimension (pixel) | 5472x3648 | 4000x3000 |
| F-Stop | f/4 | f/2.8 |
| Exposure time | 1/500 sec | 1/330 sec |
| ISO speed | ISO-125 | ISO-100 |
| True focal length | 9mm | 3mm |
| Sensor size (mm) | 13.20 x 8.80 | 6.17 x 4.55 |

**Table 2.** Canon G7X and GoPro Hero 5 specifications

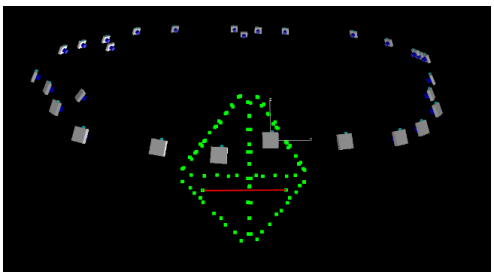

**Figure 2**. 3D view of camera locations during the data capture extracted from iWitnessPro. Green dots are the control points on the frame, the red bar is the scale bar used.

### 4.2 Data processing

Images were processed in commercial software, where a least-squares adjustment using the Brown camera model (Brown, 1971) was used to determine the values of the Interior Orientation Parameters (IOP) and Exterior Orientation Parameters (EOP), as well as the RMS of the GCPs after the adjustment and the coordinates of the CPs for an independent assessment of the accuracy. GCPs and CPs were observed in images using a centroid fitting method. The IOPs solved are: the principal distance (*c*) as the distance between the camera centre and the image plane, and the principal point offset (*xp, yp*) as the location of the principal point in the image plane. Furthermore, the following distortion parameters were used: radial lens distortion (*k1-k3* if not indicated otherwise), decentring distortion (*p1-p2*) and linear distortions (*b1-b2*). The adjustment was constraint by the dimension of the scale bar (899.954 mm).

### 4.3 Data analysis

The adjustments have been assessed using:

- Quality of Self-calibration (self-calibration score)
- Point referencing error.
- RMS of GCPs or control base lines
- RMS of CPs or check baselines
- Percentage error

The "Quality of Self-calibration" is calculated during the bundle adjustment procedure. "The Quality of Self Calibration has an optimal value of 1.0, and values to 1.5 are acceptable. Values higher than 1.5 indicate a weak network geometry and thus sub-optimal determination of camera parameters." (iWitness Manual, 2019). A good network geometry is critical for any Least Square Adjustment.

The RMS of the image point residuals (RMS Vxy), also called point referencing errors, is used as an indicator for the quality of the bundle adjustment. Assessing the RMS Vxy, allows to ascertain the likelihood of successfully orientating all images in the set to each other. A value less than 0.5 pixel in air and 1 pixel underwater is desired, and achievable considering the quality of the images, their overlap, and the high redundancy of the Least Squares Adjustment.

The Quality of Self-calibration (self-calibration score) and the Point referencing error, the RMS of GCPs or control base lines as well as the Percentage error, were used to assess the relative accuracy of the adjustment. The RMS values are outputted from the software for the GCPs and manually calculated for the CPs.

Where GCPs were calculated in an independent adjustment or if CPs were compared to a reference, the signed distances between the two datasets were calculated and tested for significance with a t-test. The same method was applied to assess using base lines. Where baselines were used, the baselines were split into two sets. One set was used to constrain the Least Squares Adjustment, while the other set was used for an independent accuracy assessment. This comparison was used to assess the absolute accuracy which could be achieved.

Three different tests were carried out:

1. *Reliability of the frame*: to check if the frame can be put together in such a way that the GCPs are unchanged. This was performed in-air only.
2. *Repeatability*: To see if the same results can be achieved for different photogrammetry software products, and using different cameras. This aspect is important for the practical use of the frame for further research. The test is only performed in-air, too.
3. Underwater performance test: The calibration frame is verified in an underwater setting as this is the intended application of the frame.

## 5. RESULT

The results of the different tests are presented below.

### 5.1 Reliability of the frame

To test the reliability of the frame, the capture of the in-air test was repeated on two different datasets captured during different days. Between the data captures the frame was dismantled and put away in the tubing simulating the field procedure. Both datasets were processed independently in a free adjustment. The only constraint introduced is the length of the scalebar. A total of 88 coordinate points were observed in each dataset and used in a free network Least Squares Adjustment utilising the iWitnessPro software.

The PRE and self-calibration score from the adjustments of the two datasets is presented in Table 3. The point referencing error for both datasets is below 0.5 pixels below the defined threshold. The self-calibration score is 1.1 for both datasets and below the threshold of 1.5. Hence, it is concluded that both adjustments have been successful.

| | Dataset 1 | Dataset 2 |
|---|---|---|
| PRE | 0.49 pixel | 0.37 pixel |
| Self-calibration score | 1.1 | 1.1 |

**Table 3.** Results of the LSA of dataset 1 and 2 using iWitness.

The derived 3D coordinates of the GCPs from both independent adjustments are used to calculate the residuals between the GCPs. Overall, 88 points could be utilised to calculate the residuals. The residuals are plotted in Figure 3, applying a uniform scale factor for on the residuals of 0.8335 mm to make any possible trends visible. The residual plot shows a bias towards the centre of the frame. This bias could be caused due to the frame not being able to be re-assembled tin exact the same manner.

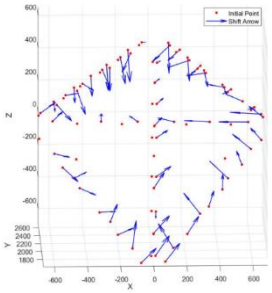

**Figure 3.** Residual (blue arrow) between the GCPs of dataset 1 (reference) and dataset 2. The residuals are scaled to make a possible trend visible.

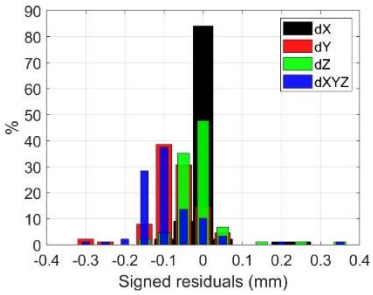


**Figure 4**. Residual plots of the X, Y, Z and overall residuals between the GCPs of dataset 1 and dataset 2.

The distribution of the signed residuals between the GCPs for the three axes and combined are presented in Figure 4. The detected bias is also visible here especially in the y coordinates of the GCPs.

The RMS of GCPs is presented in Table 4. Overall, RMS values were less than 2mm. Results of the t-test found that the residuals of the GCPs in X and Z are not significant. However, the Y value residuals show a significant difference which fits to the observations of Figure 3 and Figure 4.

| | |
|---|---|
| RMS X dimension | 0.4043mm |
| RMS Y dimension | 1.0142mm |
| RMS Z dimension | 1.1551mm |
| Overall RMS | 1.5894mm |

**Table 4.** GCP RMS of dataset 1 and 2 using iWitness.

Even though there are very small residuals, there is a systematic trend which could be caused by resembling the frame. Therefore, using the GCPs is assumed to be not sufficient. Instead, baselines will be used for any further processing.

### 5.2 Repeatability of results

Based on the results of the previous test, not GCPs but baselines were used to constraint the adjustment and to assess its processing results. On each of the six legs of the frame 2-3 baselines were defined (shown as red lines in Figure 5). As the baselines were all connected to a single leg of the pyramid, it is assumed that any bias from reassembling the frame can be removed. Overall, 14 baselines were introduced this way. Six of those baselines were kept fixed for the processing in the different software solutions. The remaining eight baselines were used for an independent accuracy assessment, to be validated if results achieved by different software products were comparable. The baselines extracted from dataset 1 was used as reference.

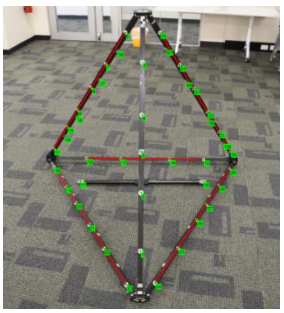

**Figure 5**. Baseline locations (red) on the frame derived y GCPS (green) on single parts of the frame.

Three software solutions were selected to process the Dataset 2 images. The software solutions were iWitness (version 4.105), Metashape (version 1.8.3) and ContextCapture (version 10.19.0.122). All adjustments using the Metashape software used the radial lens distortion parameters *k1-k4*. A summary showing all adjustment results is provided in Table 5. The adjustments were assessed to be successful, as the maximum point referencing error was 0.49 pixels, which was and under the defined threshold of 0.5 pixel. Furthermore, the maximum GCP RMS was 0.88 mm, which is within the defined range of 0.5 – 3 mm.

| | iWitness Pro | Context-Capture | Metashape |
|---|---|---|---|
| PRE | 0.49 pixel | 0.37 pixel | 0.49 pixel |
| RMS [mm] | 0.88 | 0.74 | 0.78 |

**Table 5**. Results of the LSA of dataset 2 using iWitness, Metashape and ContextCapture.

It was not possible to observed the points belonging to check baselines in Metashape as part of the bundle adjustment. Hence, using Metashape only, a 3D model of the frame had to be created based on a dense point cloud derived from the processed images. The baseline points were then observed as a 3D model using CloudCompare, serving as the foundation for all measurements resulting from the Metashape software. However, the accuracy of performing such observations is heavily reliant on the human eye and the ability to pick the points manually in the 3D model, as well as the quality of the model itself. Next to the adjustment errors, additional error caused by the dense reconstruction, as well as the creation of the model will impact the observations.

A histogram showing the residuals of the extracted check baselines from Dataset 2, compared to the reference baselines from Dataset 1 processed with iWitnessPro is shown in Figure 6. the distribution of values in Figure 6, suggested a negative bias for the iWitnessPro software and a positive bias for ContextCapture software visible.

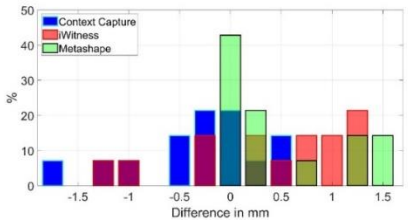


**Figure 6**. Residual plots of the different baselines distance depending on the software.

For a more detailed analysis, all baseline residuals were presented in Table 6. The RMS of the baseline residuals from all software solutions was within the acceptable magnitude of 3 mm, with values of 0.702 mm, 0.314 mm and 0.570 mm for the software iWintessPro, ContextCapture and Metashape, respectively. The average of the signed residuals shows a positive bias for Metashape (0.354 mm), a negative bias for ContextCapture (-0.132 mm) and no significant bias for iWitnessPro (0.019 mm). In contrast, the range of the residuals is the largest for the iWitness software.

To further investigate reliability, two additional cameras were added to the validation: the Canon G7X and the GoPro Hero5. By incorporating these different camera models, the aim was to assess the repeatability of the results across multiple devices, minimising any potential biases or limitations associated with a single camera. The additional cameras selected for the test were the same which will be used for Test 3 (underwater validation). For this test only the ContextCapture and Metashape software are used. The rationale was that these are the software solutions to be used in the next stage of the project.

| Differences | iWitness-Pro | Context Capture | Meta-shape |
|---|---|---|---|
| Baseline 1 | 0.809 | -0.038 | 0.123 |
| Baseline 2 | 0.462 | -0.344 | 0.111 |
| Baseline 3 | 0.291 | -0.325 | 0.132 |
| Baseline 4 | -0.945 | 0.344 | 0.169 |
| Baseline 5 | -1.204 | -0.495 | 0.237 |
| Baseline 6 | 0.702 | 0.066 | 1.350 |
| RMS | 0.795 | 0.314 | 0.570 |
| Average | 0.019 | -0.132 | 0.354 |
| Range | 2.013 | 0.839 | 1.461 |

**Table 6**. Baseline residual and RMS of the dataset 2 processing results using iWitnessPro, Metashape and ContextCapture. All values are in [mm].

A summary showing all adjustment results for the three cameras for the ContextCapture and Metashape is provided in Table 7. The adjustments were assessed to be successful for the Canon G7 camera, as the maximum point referencing error is 0.56 pixels, which only just above the defined threshold of 0.5 pixel for in-air applications. The point referencing error of the Canon camera was identical for the two software used. The RMS value was also identical for both software for this camera with 1.31 mm, which was nearly double of the value for the Nikon D750 camera, but still below our defined threshold of 3 mm. The results for the percentage were comparable to the results from the RMS analysis. The results of the proceeding of the GoPro camera were borderline. The RMS values for both software were 1.47 mm and 1.57 mm, respectively, and below the defined threshold. However, the point referencing error was above the threshold, and were 0.63 mm and 0.83 mm for the ContextCapture and Metashape software, respectively. It is possible that the GoPro cameras were borderline due to using only *k1-k4* parameters to model the radial lens distortion, which may not be sufficient for the fisheye distorted images.

| | ContextCapture | | | Metashape | | |
|---|---|---|---|---|---|---|
| | Nikon D750 | Canon G7X | Go-Pro | Nikon D750 | Canon G7X | Go-Pro |
| PRE pixel | 0.37 | 0.56 | 0.63 | 0.49 | 0.56 | 0.83 |
| RMS [mm] | 0.74 | 1.31 | 1.47 | 0.78 | 1.32 | 1.57 |

**Table 7:** Results of the LSA of dataset 2 using Metashape and ContextCapture in combination with three different cameras.

The RMS increased using the lower grad cameras, compared to the Nikon D750 SLR used for the previous tests, which reflects the results presented in Table 7. The negative bias of the ContextCapture software was visible for all cameras but significantly larger for the lower grade cameras (Table 8). In contrast, the positive bias of the Metashape software for the Nikon D750 SLR turns into a negative bias with by far the largest magnitude. The Metashape results could have been impacted by the method of how the check baseline observations were performed, as discussed previously. Overall, though, all RMS values were within the acceptable defined range of 0.5 – 3 mm.

| Base-line | ContextCapture | | | Metashape | | |
|---|---|---|---|---|---|---|
| | Nikon D750 | Canon G7X | GoPro | Nikon D750 | Canon G7X | GoPro |
| 1 | -0.04 | -1.77 | -0.91 | 0.12 | -1.52 | -1.28 |
| 2 | -0.34 | -1.08 | -2.42 | 0.11 | -1.41 | 1.46 |
| 3 | -0.33 | -1.10 | 0.67 | 0.13 | -1.34 | -1.49 |
| 4 | 0.34 | 0.89 | -1.17 | 0.17 | -1.32 | -1.52 |
| 5 | -0.50 | -1.49 | 1.53 | 0.24 | -1.24 | -1.99 |
| RMS | 0.34 | 1.31 | 1.47 | 0.16 | 1.32 | 1.57 |
| Avrg | -0.09 | -0.54 | -0.14 | 0.16 | -0.91 | -0.54 |
| Range | 0.84 | 3.08 | 3.95 | 0.13 | 2.89 | 3.56 |

**Table 8**. Baseline residuals [mm] and statistics of the dataset 2 processing results using Metashape and ContextCapture using three different cameras.

### 5.3 Underwater performance test

For the underwater performance test, the calibration frame was submerged in shallow and clear water, and images were captured using the previously tested Canon G7X (Figure 7) and the GoPro HERO5 Black (Figure 8). Images were captured on 29/5/2023 at Coogee Beach, Perth, Australia (S32.105569, E115.761822). The images were captured in the same manner as they were taken previously making them suitable for self-calibration.

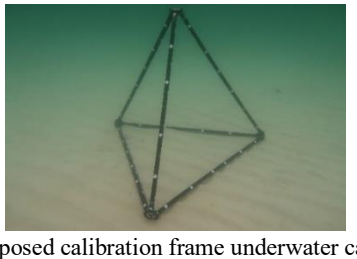

**Figure 7.** Proposed calibration frame underwater captured using the Canon G7X camera.

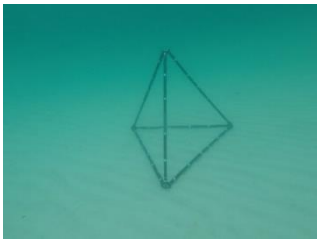

**Figure 8**. Proposed calibration frame underwater captured using the GoPro 5 Black camera.

Table 9, compares the performance of the two software programs: ContextCapture and Metashape, for the two different camera models: Canon and GoPro. In terms of point reference error in pixels, both ContextCapture and Metashape show higher values for the GoPro camera compared to the Canon camera. In ContextCapture, the point reference error was 0.91 pixels for the Canon camera and 1.39 pixels for the GoPro camera. In Metashape, the point reference error was higher for both, with values of 2.82 pixels for Canon and 3.6 pixels for GoPro. Considering the defined threshold of 1 pixel, only the Canon software processing the data with the ContexCapture software (0.91 pixel) passed the defined threshold requirements. The GoPro camera using the ContextCapture software was just above the threshold of 1 pixel with a value of 1.39 pixel. It can be concluded that the ContextCapture software was suitable to produce the required accuracy.

In terms of RMS error in mm, the Canon camera again performed better than the GoPro, but although ContextCapture again performed better than Metashape, the values were similar camera models. For the Canon, ContextCapture, the RMS error was 1.02 mm and in Metashape it was 1.32 mm; and for the GoPro, ContextCapture had an RMS of 1.90 mm and Metashape was 1.96 mm. Importantly, all values were within the threshold of 3 mm.

The RMS values for the check baseline residuals from the underwater test (Table 10), were generally slightly higher, compared to the in-air test (Table 8); however, still below the desired maximum threshold of 3 mm. In the ContextCapture software, the RMS value increased from 1.31 mm and 1.47 mm in air, to 1.63 mm and 1.90 mm underwater for the Canon and GoPro camera, respectively. The Metashape software again had larger values. For the Metashape software, the RMS value increased from 1.37 mm and 1.57 mm in air, to 1.80 mm and 1.96 mm underwater for the Canon and GoPro camera, respectively.

| | ContextCapture | | Metashape | |
|---|---|---|---|---|
| | Canon G7X | GoPro | Canon G7X | GoPro |
| PRE pixel | 0.91 | 1.39 | 2.82 | 3.63 |
| RMS [mm] | 1.02 | 1.90 | 1.32 | 1.96 |

**Table 9:** Results of the LSA of the underwater dataset using Metashape and ContextCapture in combination with two different cameras.

| Differences (mm) | ContextCapture | | Metashape | |
|---|---|---|---|---|
| | Canon G7X | GoPro | Canon G7X | GoPro |
| Baseline 1 | 0.92 | -2.28 | -1.38 | 2.00 |
| Baseline 2 | -1.01 | 1.46 | 1.35 | -2.75 |
| Baseline 3 | 1.03 | -1.49 | -1.36 | 1.25 |
| Baseline 4 | 1.13 | 2.12 | -1.35 | -2.14 |
| Baseline 5 | -1.02 | -1.99 | -1.24 | 1.25 |
| RMS | 1.02 | 1.90 | -1.32 | 1.96 |
| Average | 0.21 | -0.44 | -0.80 | -0.08 |
| Range | 2.15 | 4.40 | 2.73 | 4.75 |

**Table 10**. Underwater baseline residual and RMS of the dataset 2 processing results using ContextCapture and Metashape.

## 6. CONCLUSION

The accuracy achieved by the various calibrations carried out using the proposed calibration frame, with different cameras and software combinations, demonstrated that the proposed calibration frame is an effective tool for achieving reliable camera calibration both in air and in water. The results were comparable, or better, than similar studies, particularly when baselines were used (Table 11).

In air camera performance tests, this study found that the Nikon D750 consistently achieved the lowest RMS values with both ContextCapture (0.74 mm) and Metashape (0.78 mm), indicating superior accuracy and imaging quality compared to the other cameras. While the Canon G7X generally performed better than the GoPro in both software environments.

In water tests, the study found the Canon G7X outperformed the GoPro in both software environments. Capra et al. (2015) also achieved better results with a compact Canon camera over a GoPro (Table 11). Results were similar between software used. With ContextCapture, the Canon G7X achieved a lower RMS value of 1.02 mm, while the GoPro scored 1.90 mm. Similarly, using Metashape, the Canon G7X obtained an RMS value of 1.32 mm, and the GoPro scored 1.96 mm. Canon performance better than the GoPro, which is likely due to the better camera sensor and specifications (e.g. sensor and image size). This study achieved better results using Context Capture than MetaShape. Nevertheless, overall, it can be concluded that the results are acceptable for both cameras and software solutions.

| Study | RMS (mm) | |
|---|---|---|
| | In-Air | In-water |
| Gourgoulis et al. (2008) | 3.7 – 4.66 | 4.5-5.92 |
| Challis and Kerwin (1992) | 0.8 | N/A |
| Helmholz et al. (2016) | N/A | 0.45 – 2.5 |
| Li et al. (1997) | N/A | 10.83 |
| Capra et al. (2015) | N/A | -Hero2: 43.03 mm<br>-Intova Sport HD: 11.330 mm<br>-Canon PowerShot G12: 0.524 mm |
| This study | ContextCapture:<br>Nikon D750: 0.74<br>Canon G7X: 1.31<br>GoPro: 1.47<br>Metashape:<br>Nikon D750: 0.78<br>Canon G7X: 1.32<br>GoPro: 1.57 | ContextCapture:<br>Canon G7X: 1.02<br>GoPro: 1.90<br>Metashape:<br>Canon G7X: 1.32<br>GoPro: 1.96 |

**Table 11.** Comparison of calibration results from selected studies compared to the current study.

## ACKNOWLEDGEMENTS

The authors would like to thank King Abdulaziz University for funding Alaa Mufti's PhD study. We also express our thanks to Mr Malcolm Perry and Mr Luke Mickan from Curtin University for supporting to build the frame. Furthermore, we would like to acknowledge Curtin's 3D Hub and HIVE to utilise their facilities for the data processing and visualisation.

## REFERENCES

Bouguet., 2017. Camera calibration toolbox for MATLAB. *California Institute of Technology*. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

Brown., 1971. Close range camera calibration. *Photogramm Eng*, 37(8), 855–866.

Canon., 2014. Canon PowerShot G7X Mark II. Tokyo, Japan.

Capra, A., Dubbini, M., Bertacchini, E., Castagnetti, C., & Mancini, F,. 2015. 3D reconstruction of an underwater archaelogical site: Comparison between low cost cameras. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40, 67-72.

Challis, J. H., & Kerwin, D. G,. 1992. Accuracy assessment and control point configuration when using the DLT for photogrammetry. *Journal of Biomechanics*, 25(9), 1053-1058.

Diamanti, E., Spondylis, E., Vlachaki, F., & Kolyva, E,. 2017. Surveying the underwater arcaheological site of cape glaros at pagasetikos gulf. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 243-250.

Fraser, C. S., Shortis, M. R., & Ganci, G,. 1995. Multisensor system self-calibration. *Paper presented at the Videometrics IV*.

GoPro,. (2019). GoPro Hero 8 Black. GoPro, Inc, San Mateo, CA.

Gourgoulis, V., Aggeloussis, N., Kasimatis, P., Vezos, N., Boli, A., & Mavromatis, G,. 2008. Reconstruction accuracy in underwater three-dimensional kinematic analysis. *Journal of Science and Medicine in Sport*, 11(2), 90-95.

Guo, T., Capra, A., Troyer, M., Grün, A., Brooks, A. J., Hench, J. L., Schmitt, R. J., Holbrook, S. J., & Dubbini, M,. 2016. Accuracy assessment of underwater photogrammetric three dimensional modelling for coral reefs. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41(B5), 821-828.

Harvey, E., & Shortis, M,. 1995. A system for stereo-video measurement of sub-tidal organisms. *Marine Technology Society Journal*, 29(4), 10-22.

Helmholz, P., Long, J., Munsie, T., & Belton, D,. 2016. Accuracy Assessment Of Go Pro HERO 3 (Black) Camera In Underwater Environment. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41.

Jalal, A., Salman, A., Mian, A., Shortis, M., & Shafait, F,. 2020. Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecological Informatics*, 57, 101088.

Li, R., Li, H., Zou, W., Smith, R. G., & Curran, T. A,. 1997. Quantitative photogrammetric analysis of digital underwater video imagery. *IEEE Journal of Oceanic Engineering*, 22(2), 364-375.

Samboko, H. T., Schurer, S., Savenije, H. H., Makurira, H., Banda, K., & Winsemius, H,. 2022. Evaluating low-cost topographic surveys for computations of conveyance. *Geoscientific Instrumentation, Methods and Data Systems*, 11(1), 1-23.

Shortis, M. (2019). Camera calibration techniques for accurate measurement underwater. 3D recording and interpretation for maritime archaeology, 11-27.

Shortis, M., Abdo, E. H., & Dave,. 2016. A review of underwater stereo-image measurement for marine biology and ecology applications. Oceanography and marine biology, 269-304.

Shortis, M. R., Clarke, T. A., & Robson, S,. 1995. Practical testing of the precision and accuracy of target image centering algorithms. Paper presented at the Videometrics IV.

Shortis, M. R., & Harvey, E. S,. 1998. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. International Archives of Photogrammetry and Remote Sensing, 32, 792-799.

Shortis, M. R., Miller, S., Harvey, E., & Robson, S,. 2000. An analysis of the calibration stability and measurement accuracy of an underwater stereo-video system used for shellfish surveys. Geomatics Research Australasia, 1-24.

Shortis, M. R., & Seager, J. W,. 2014. A practical target recognition system for close range photogrammetry. The Photogrammetric Record, 29(147), 337-355.

Wehkamp, M., & Fischer, P,. 2014. A practical guide to the use of consumer-level digital still cameras for precise stereogrammetric in situ assessments in aquatic environments. *Underwater technology*, 32(2), 111-128.

Yuan, S., Li, Y., Bao, F., Xu, H., Yang, Y., Yan, Q., Zhong, S., Yin, H., Xu, J., & Huang, Z,. 2022. Marine environmental monitoring with unmanned vehicle platforms: Present applications and future prospects. *Science of The Total Environment,* 159741.

Zhang, Z,. 2000. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11), 1330-1334.