Underwater Videogrammetry with Adaptive Feature Detection at "See am Mondsee", Austria

MARCO BLOCK, Hochschule für Technik und Wirtschaft Dresden, Germany
CYRIL DWORSKY, Kuratorium Pfahlbauten, Austria
CARMEN LÖW, Kuratorium Pfahlbauten, Austria
HELENA SEIDL DA FONSECA, Kuratorium Pfahlbauten, Austria
BENJAMIN GEHMLICH, Hochschule für Technik und Wirtschaft Dresden, Germany
DENNIS WITTCHEN, Hochschule für Technik und Wirtschaft Dresden, Germany
NIKLAAS GÖRSCH, Freie Universität Berlin, Germany
PAULINA SUCHOWSKA, Uniwersytet im. Adama Mickiewicza, Poznań, Poland
BENJAMIN DUCKE, Deutsches Archäologisches Institut Berlin, Germany

We present a complete, video-based 3d documentation process for the submerged remains of Neolithic pile dwellings at the UNESCO World Heritage Site "See am Mondsee" in Austria. We discuss good practice routines and solutions, such as cable management, supporting the Unmanned Underwater Vehicle (UUV) when strong currents are prevalent, and documentation/record keeping. The recorded site is a Neolithic lake village dating to the 4th millenium BC. Based on initial reconstruction results, we improved the image matching process of our Structure from Motion (SfM) pipeline (built around the free end-user application VisualSFM), by replacing its default feature detector (SiftGPU) with our own implementation of adaptive feature detection. The campaign was accompanied by a German television film crew. Their documentary was shown on the German public television (ARD) broadcast "W wie Wissen".

1. MOTIVATION AND INTRODUCTION

The UNESCO World Heritage ensemble "Prehistoric Pile Dwellings around the Alps" consists of 111 archaeological sites in a region that includes adjacent parts of Germany, Slovenia, Italy, France, Austria and Switzerland. In order to manage the five sites on Austrian territory, the Kuratorium

Pfahlbauten[1] has been established. The Kuratorium's main objectives are to protect UNESCO-registered pile dwellings (Fig. 1) and to actively participate in research on this shared world heritage. In addition, the UNESCO World Heritage Convention[2] requires a World Heritage to be actively developed and used in a public function. Due to the sensitivity of the archaeological remains and their concealment under water, this proves to be a demanding task. Other World Heritage Sites are much more accessible, and can be traversed and experienced physically and in a more holistic way.

With regard to the pile dwellings, there is no physical accessibility; even for sport divers. Restricted diving areas covering the site have been designated and are indispensable to maintain and adhere to national protection laws. This reality requires new approaches for the general public to participate in UNESCO World Heritage Sites, i.e. conveying knowledge outside the actual site area, as well as the involvement of citizen scientists in research undertaken by the Kuratorium Pfahlbauten.

Figure 1. The UNESCO World Heritage Site "See" at lake Mondsee. The settlement remains dates to the 4th millennium BC and are exceptionally well preserved, because of their location in shallow water.

At "See" (am Mondsee/at lake Mondsee), diving operations are carried out by one of two site managers of the Kuratorium Pfahlbauten in cooperation with, in most cases, two scientific divers, working several days once a year [Pohl 2016]. During these monitorings, a grid of erosion markers is regularly read and documented. In this way, possible endangerment to the substance of the remains or their immediate surroundings, caused by the removal of the protective sediment layer, can be assessed.

An active involvement of citizen scientists would further facilitate denser surveillance of archaeological sites, and thus a more sustainable protection of dwellings belonging to the UNESCO World Heritage than our own resources could ever allow.

Due to strict national regulations, it is nearly impossible for citizen scientists to participate as divers. However, it is feasible to achieve something similar using videogrammetry supported by Unmanned Underwater Vehicles (UUVs), operating in accordance with the current judicial framework. It is

---

[1] http://www.pfahlbauten.at
[2] http://whc.unesco.org/archive/convention-en.pdf [01-07-2017]

essential for the implementation of this project that the technology used provides good results with regard to the most accurate documentation possible, is cost-effective and user-friendly. In addition, the recording process must not endanger the substance of archaeological sites.

The projects "Archaeocopter" and "Archaeonautic"[3], run by Hochschule für Technik und Wirtschaft (HTW) Dresden and Freie Universität (FU) Berlin, were initiated in cooperation with the Deutsches Archäologisches Institut (DAI) and the archaeological heritage office in the German federal state of Saxony. The philosophy of both projects is not to try to maximize the level of detail achievable using highest resolution sensors, but to focus on what is actually required for different practical purposes, and providing robust and cost-effective solutions.

2.   THE PILE DWELLINGS AT "SEE AM MONDSEE"

The pile dwelling, or lake village, "See am Mondsee" is perhaps the most famous Copper Age site of Austria. It dates back to the 4th millennium BC [Ruttkay et. al. 2004]. The site was discovered by Matthäus Much in 1872. He examined it during the 1880s and established the term "Mondsee culture". Elisabeth Ruttkay (†2009), one of the most significant researchers of the European Neolithic period in Austria, studied the Mondsee culture from the 1990s onward and came to be convinced that the Mondsee material is a sub-group of the Neolithic "Trichterbecherkultur" (Funnelbeaker culture), rather than an independent local culture. Therefore, the term "Mondsee culture" was replaced by the more appropriate name "Mondsee group" in the academic discourse.

Artefacts (Fig. 2) similar to those discovered by Much were also found at approximately twenty other underwater sites at lakes Attersee and Mondsee, as well as on dry land sites in the federal states of Lower and Upper Austria, Salzburg and – on the southern side of the Alps – in Styria [Maurer 2014]. The Mondsee group material is characteristic of the Neolithic pile dwellings (traditionally called "palafittes") settlements of the Austrian Salzkammergut region [Ruttkay 1981].

Figure 2. Pottery of the Mondsee group: black fired ceramics with incise-punctuation, limescale incrustation with typical decorations of wave-like lines, and circles often interpreted as symbols of the sun.

---

[3] www.archaeocopter.de

The rich inventory of finds, including well-preserved organic materials, at the settlement "See am Mondsee" constitutes the most comprehensive source for scientific exploration of Neolithic pile-dwelling cultures in Austria so far. After Much's collecting activities, L. Franz and R. Bernhart examined the site in 1938. In 1951, K. Willvonseder and K. Schaefer undertook the first diving examinations; and in 1961 salvages were carried out by the Mondseer Heimatbund under W. Kunze. In 1967/68 the state heritage protection office surveyed the site, and from 1982-86, had J. Offenberger undertake a surface documentation and a salvage of finds [Hirmann 1999].

Since 1989, the Austrian national research fund FWF has been supporting interdisciplinary projects regarding the prehistoric lake villages of Upper Austria. This has further facilitated research into trade contacts and comparison with contemporaneous pile-dwelling cultures. For example, objects made of the so-called "Mondsee copper", containing a high percentage of arsenic, have been found all across Europe [Dworsky 2016]. The site was granted UNESCO World Heritage status in 2011.

In 2013, the Kuratorium Pfahlbauten started to implement a monitoring system in each of Austria's World Heritage Sites. The results show very different topographical and hydrographical situations of the five sites and thus different states of preservation. The lake village of "See am Mondsee" is situated at the eastern end of lake Mondsee, in a small bay near the lake's outlet (Fig. 3).

Figure 3. See, Mondsee: location and extent of the World Heritage Site.

The outflow of the Mondsee, the river "Seeache", leads into the southern end of lake Attersee. The constant stream of the outflow/outlet causes erosion in the area of the prehistoric settlement, especially during periods of heavy rainfall and high water levels [Pohl 2014]. Therefore, and in contrast to better conserved lake village sites, "See am Mondsee" lies in a very dynamic and constantly changing environment of sedimentation processes, and thus challenging preservation conditions; prehistoric piles are visible and cultural layers with prehistoric material are exposed on

the surface (Fig. 4). Despite these difficult circumstances, prehistoric cultural layers are still preserved up to a thickness of about 50 centimeters in some areas [Pohl 2016].

Figure 4. See, Mondsee: pile-field and prehistoric material on the lake bottom.

In order to further and improve future protective measures, more emphasis has to be put on a detailed surface model of the lake bottom, accurate documentation of both erosion and sedimentation processes and the establishment of the exact locations of the most threatened areas of the site.

3.   VIDEOGRAMMETRY IN UNDERWATER ARCHAEOLOGY

Over the last two decades, archaeologists have profited from developments in UAV (Unmanned Aerial Vehicle) and rapid progress in digital camera technology. At the same time, expertise and methods from aerial archaeology and close-range photogrammetry intersect with each other [Luhmann et. al. 2007]. While it seems that most published research currently uses the proprietary software AgiSoft PhotoScan[4], there is also an increasing role for free and open source software, such as VisualSFM[5], MicMac [Deseilligny and Clery 2011], OpenMVG [Moulon et. al. 2013] and Bundler [Snavly et. al. 2006], that can reduce costs while still producing robust, high quality results [Rende et. al. 2015].

For our research, we have developed several tools in the Java programming language that each serve a well-defined purpose: JKeyFramer automatically extracts a set of key frames from a video stream that are well-suited for image-based reconstruction. Frame selection is based on a rapid assessment of quality via sharpness and other indicators. JEnhancer is a tool for automatic color correction of underwater imagery (for details see Block, Gehmlich and Hettmanczyk in this volume). JFeatureManager implements a novel, adaptive feature detector for use with blurry images. It is described in detail in section 5 of this paper. The programs JEnhancer and JFeatureManager were developed specifically for use in underwater 3d reconstruction.

---

[4] http://www.agisoft.ru/products/photoscan/
[5] http://ccwu.me/vsfm/

3.1   Videogrammetry versus photogrammetry

The videogrammetric approach [Greenwood 1999, Pappa et. al. 2003, Nistèr 2004. Pollefeys et. al. 2004] to 3d reconstruction (i. e. using frames extracted from video streams instead of single-shot image series as in the common photogrammetric approach [Hartley and Zissermann 2004]) was successfully introduced to aerial 3d reconstruction using Unmanned Aerial Vehicles (UAVs) in the project "Archaeocopter", among others. The results of several campaigns have shown that videogrammetry is a fully viable approach to reconstruct single objects as well as complete archaeological areas [e.g. Gehmlich and Block 2015]. The main challenge in videogrammetry is to compute a solution for a problem involving two contradicting constraints: On the one hand, we want to minimize the distances between corresponding images to maximize the intersection set and produce more 3d points. But on the other hand, we need to maximize the distances between corresponding images to reduce the locational uncertainty of the reconstructed 3d points, as only accurately reconstructed points will be kept for the 3d model.

In our experience with recording data while moving, videogrammetry is the more fault-tolerant, more cost-effective and easier-to-use approach, compared to single-shot photogrammetry. The software "JKeyFramer", an automatic keyframe selection tool and one of the most important outcomes of the project "Archaeocopter", has evolved to allow us to render fast preview models on site.

3.2   Aspects of underwater archaeology

The recording methods currently used in underwater archaeology are complex and expensive, even when sonar and laser scanning are deployed [Moisan et. al. 2015]. The importance of the photogrammetric approach is therefore analogous to its impact on aerial documentation, and the use of underwater photogrammetry is accordingly on the rise [Baletti et. al. 2015, Pruno et. al. 2015]. The results of land and water based photogrammetry are quite comparable. Depending on the setup used, only small details are lost in underwater scenes [Troisi et. al. 2015].

Currently, scientific divers use cameras with high resolution, plan sets of photos and need special training to work underwater [Papadimitriou 2015]. Their diving time is limited, depending on diving depth. Careful and systematic excavation under water is still a domain of manual human labor.

In comparison to aerial documentation with UAVs, underwater georeferencing is still a significant challenge. Natural or artificial markers on the ground (ground control points, GCP) need to be well detectable and recognizable. Determining the real-world coordinates of such points is a problem, because no GPS signal is directly available under water. To locate GCP markers with sufficient accuracy, indirect tracking solutions are often deployed. In areas with shallow water, the point to locate can be correlated with a GPS-provided point on the water surface via a perpendicular stick [Baletti et. al. 2015].

4.   DOCUMENTATION PROCESS FOR PILE DWELLINGS AT "SEE AM MONDSEE"

UUVs, can be used effectively for both, the documentation of registered archaeological sites, and the exploration of potential new sites. The small submarine "Eckbert-II", based on the OpenROV[6], was developed in the project Archaeonautic to fullfil both functions. At the time of writing, it is still under modification [Block et. al. 2016, Eckbert-II]. Two extensions, for cable management and for finding a balance setup for underwater taring in both salt and fresh water were designed.

4.1   Recording and computational setup

The main difference and the most critical technical challenge for the documentation of the pile dwellings at "See am Mondsee" in Austria, was to ensure that neither the UUV nor the attached cable touches the sea bottom or the salient lake dwellings. Stable balancing of the UUV with different attachments is an important aspect of successful and safe underwater recording. Promising 3d reconstruction results using GoPro cameras, Structure from Motion [Rende et. al. 2015] and stereo vision [Repola et. al. 2015] underwater were obtained. We added two diving torches on both sides of the OpenROV, to be able to get good light conditions if necessary, and also one GoPro Hero 4 BE on each side (Fig. 5, left).

Another practical aim of the Mondsee campaign was to evaluate whether the miniature UUV "Eckbert-II" would be able to document the pile dwellings so efficiently that it could be done periodically, making our approach suitable for long-term underwater site monitoring.

Figure 5. The images show "Eckbert-II" submerged in Lake Mondsee. Left: A flexible camera- and lighting setup can accommodate different recording needs and environmental conditions. To document the pile dwellings, three GoPro Hero 4 BE were attached to the UUV. Subsequently, weights and buoyancy bodies were combined to balance the UUV. Right: Recording can be done in complete darkness, thanks to the diving torch setup used. Under these extreme conditions, 3d models can still be produced successfully.

The Pro LED Scuba 860 with 860 lumens and a potential maximum depth of about 100 meters was chosen as diving torch. It outperforms the GoPro standard case that offers only about 40 meters of diving depth. One torch is placed 15 cm in front and the other 15 cm behind the centrally placed GoPro. Both torches produce a homogeneously lighted area around the field of recording (Fig. 5, right).
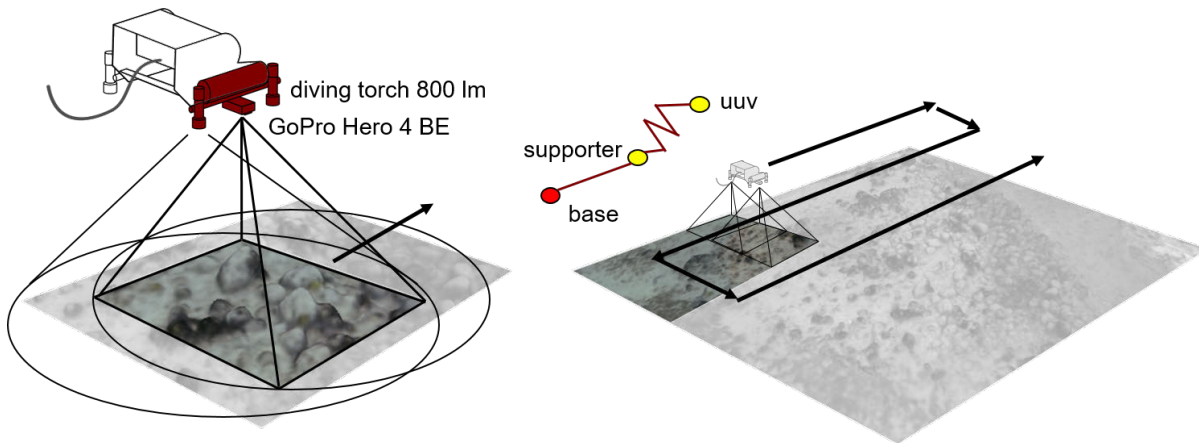
---

[6] http://www.openrov.com

Figure 6. Left: To each side of the OpenROV, two diving torches and one GoPro Hero 4 BE were extended. Middle: To improve the control of the UUV, a human diver supports the cable. In doing so, the diver decreases the pulling force from the base while increasing the level of movement control over the UUV. It is like «walking the dog». Right: The recording strategy with parallel and crossing stripes is similar the aerial documentation.
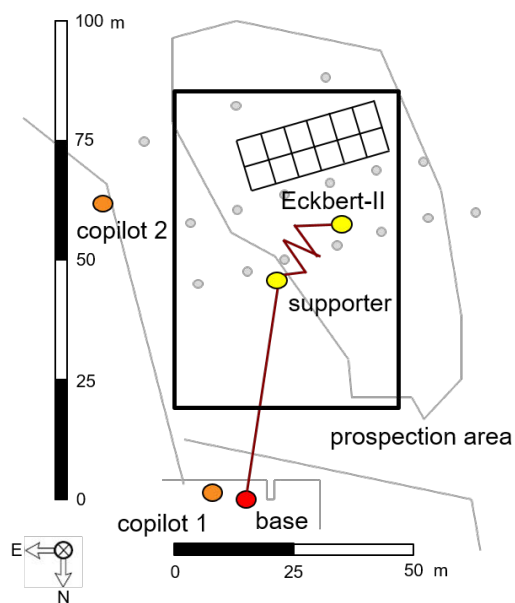

Figure 7. The base camp, placed to the north of the site, provides a workplace for the operator ("pilot"), who remotely controls the UUV via a laptop. Lower right: Equipment containers conveniently double as mobile control desk. "Copilot 1" stands close to the base to manage the cable and communicate via gestures with the UUV's supporter/diver. Similar to UAV-based documentation, when no live information about the UUV's current position is available, a second copilot is placed to the side of the surveyed area and provides an estimation of distance.

The strategy for recording video in underwater scenarios (Fig. 6) is quite similar to the grid-based version used in aerial recording [Gehmlich and Block 2015]. Compared to a UAV, "Eckbert-II" lacks one degree of freedom (it is not able to move sideways).

4.2   Documentation area

The base camp for the UUV's operation was set up on a meadow to the north of the site, since there was no public access to the (closer) areas south and east of site. Using a 100-meter-long cable, the area to survey could be covered in its entirety (Fig. 7).

As Fig. 10 (left) shows, a regular, georeferenced frame of about 10x30 meters was placed inside the archaeological site and relocated several times (archaeological field work supervised by Johann Offenberger[7], see also [Hirmann 1999]). The frame's position was recorded by the archaeologists and could therefore be used to georeference the resulting 3d model.

4.3   Team communication

A short briefing is part of the routine before every documentation job. Upcoming diving sessions are discussed among the team. During the diving session, communication between copilot 1 and copilot 2 is done via walkie-talkies with headsets (DeTeWe Outdoor 8500 PMR).

The pilot runs the base and controls the UUV via laptop. Control channel and video stream between UUV and pilot are realized via web services. Placed close to the pilot, the first copilot manages the cable to the UUV and communicates via gestures between the pilot and the UUV supporter/diver. Some gestures are equivalent to international hand signals used in diving, but additional communication codes are necessary when controlling the UUV (Tab. 1, see in Appendix). The base operator minds the diving time and keeps written protocols. To improve the workflow and identify bottlenecks, it is always advisable to document every step carefully.

4.4   Diving sessions and recorded data

In the end, there were about eight to ten diving sessions for the film crew and four diving sessions to record useful site data. The OpenROV is able to work three hours in a row, but the GoPro cameras proved to be the energy bottleneck. Even with the GoPros' energy saving video mode (1080p at 30 fps), we were only able to work 45 to 60 minutes. Therefore, we decided to limit the diving sessions to 45 minutes each.

As recording strategy, we used the orthographic double-grid-based approach [Gehmlich and Block 2015] with a stripe distance of about two meters. As a result, we had about five hours of useful video material out of four diving sessions, comprising more than a half million video frames.

4.5   Flexible taring setup

The basic OpenROV is well balanced for salt and fresh water. But if attachments like cameras and torches are added, a new taring setup is needed. To be more flexible, we designed a lightly oversized

---

[7] http://mondsee.salzkammergut.at/detail/article/pfahlbau-manuskript-mondsee-weltkulturerbe-see.html

buoyancy body for each side. To balance the UUV more sensitively depending on the current setup, additional plumbs were added. The positions of these plumbs are key to balancing the alignment while moving.

The postprocessing of recorded UUV sensor data, like inertial measurement unit or depth sensor readings, is also very important, as this can be used to analyze and understand the underwater behavior of the device.

In one recording scenario, we tested our flexible taring system. The UUV needed to hold its depth while moving (Fig. 8).
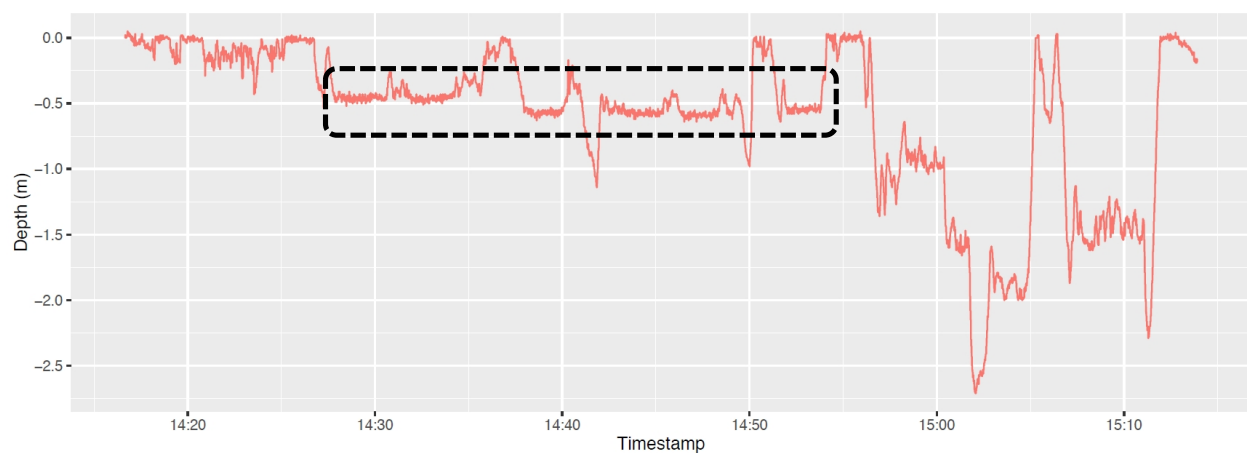
Figure 8. Most of the time, the UUV was close to the water surface, as mentioned above. While documenting the pile dwelling remains (section marked by frame) between 14:27 and 14:54, the UUV maintained its depth satisfactorily. When the supporter moved the UUV to the next position manually, small outliers occurred.

The results of the campaigns in Veruda [Scholz et. al. 2016] and Mondsee (Fig. 11) are promising and show that the flexible taring setup works equally well in both salt and fresh water.

5.   RECONSTRUCTION PROCESS AND INTERIM RESULTS

Before discussing some challenges and our ongoing work, we want to show some encouraging results. With a camera-to-bottom distance of about 1-3 meters, Archaeo3D is able to reconstruct gap-free 3d models from recorded video streams using the original frame data (Fig. 9).
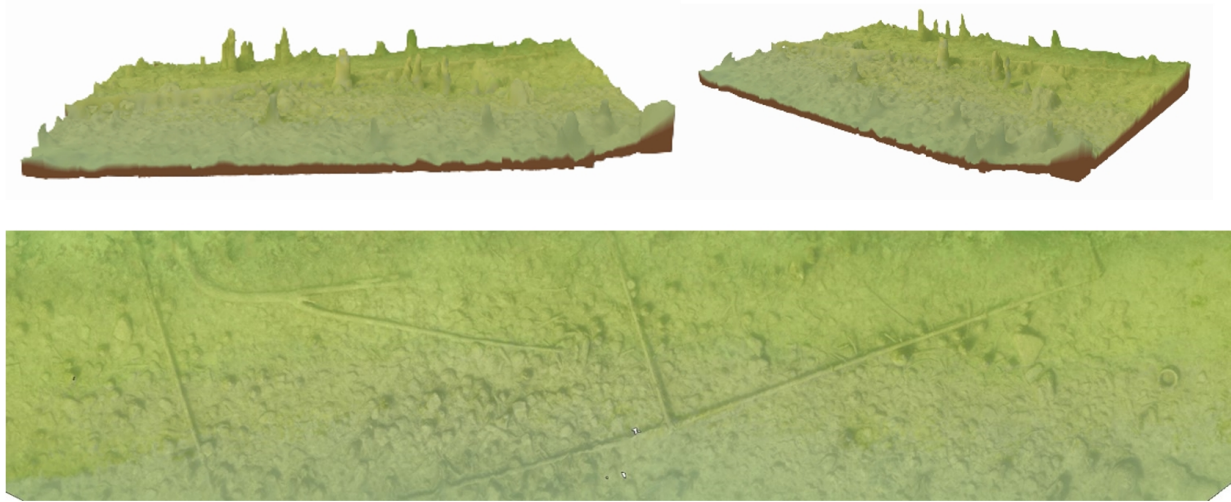
Figure 9. The 3d model above was reconstructed using the original image data with its greenish hue. Top: One patch of about five to six meters is shown from two different perspectives. The camera-to-bottom distance is about four meters. Although the muddy ground shows faded, blurry structures, the dwelling remains are well recognizable in the processed model. Bottom: One stripe of about four to twelve meters. Here, a frame with wood elements like trunks and branches can clearly be seen.

With a camera-to-bottom distance up to four meters, even images of the slightly muddy ground with its faded and blurry structures could be processed with quite promising results.

5.1   The Archaeo3D reconstruction pipeline

Within the scope of the Archaeocopter project, the semi-automatic tool chain "Archaeo3D" was developed to optimize and control the complete reconstruction process. Videos and photos are automatically imported and processed. The software is able to reorder or swap the processing pipeline's modules and adjust parameters, depending on the current hardware and the real recording situation and complexity. A combination of VisualSFM and CMPMVS [Jancosek et. al. 2011] provides the backbone of the tool chain in all Archaeocopter related projects.

5.2   Adaptive feature detection and pessimistic recording strategy

The main issue was the pessimistic recording strategy we used. We decided to record the data at a maximum depth of half a meter, due to the fact that we did not have enough knowledge about conditions inside the lake and we did not know the dimensions of the objects of interest, such as the lengths of the pile remains (as previously stated, the most important restriction was to never touch the latter).

At the beginning of Section 5 we presented some promising but interim results. We came to the conclusion, that a camera-to-bottom distance up to four meters is sufficient to produce the expected quality of the 3d model. Unfortunately, the continuously increasing depth of the prospection area reaches more than six meters. The quality of the recorded video data at a maximum distance of about

six meters (or more) was very dissatisfying, due to increasing haze. The more blurry a video frame, the less features are detectable. But a sufficient number of features are necessary to compute the pose estimation (camera reconstruction) in the SfM step, and thereby find the correct relationships between the images.

The most frequently used feature detection method is based on Lowe's SIFT with some pre-trained parameters [Lowe 1999]. VisualSFM uses SiftGPU by default, an accelerated, GPU-based implementation of SIFT. This approach is a good choice in general and works well in most cases. But when using small images with blurry content, feature detection often fails or detects insufficient numbers of features, so that image correlations cannot be found. To bring more images into the pose estimation step together, we improved our processing pipeline to detect more features by using an adaptive feature detection.

Common implementations of SIFT allow to adjust a number of parameters[8], like "octaves" or "sigma of gaussian". In most scenarios, a lot of features are detected and to reduce the time of matching, these parameters are used to filter out weak features. "Weak" is used here in the sense of "weaker than others", but most features are still good enough to produce reliable matches [May et. al. 2010]. In our case (where blurry images are processed and insufficient features are detected to match images), the SIFT parameter to adjust the contrast threshold t to filter out weak features in semi-uniform (low-contrast) regions is of particular interest. The larger this threshold, the fewer features are produced by the detector. Our adaptive feature detection starts with $t = T$ and reduces t stepwise by $t = \frac{t}{a}$ for all images with $t \in [T_{MIN}, T]$. The algorithms stops if the value of t reaches the minimum T_{MIN} or enough features are detected with $|F| \geq F_{Min}$. We empirically chose $T = 0.02$ (standard value uses by SiftGPU), $F_{MIN} = 0.0005$ and $a = 1.1$ to detect a minimum of 1800 features per frame. Our adaptive feature detection was implemented into the software modul JFeatureManager.

5.3   Comparison between standard and improved feature detection

To discuss the new feature detection approach, we selected one video sequence of about 750 seconds. The sequence starts with a camera-to-bottom distance of about four meters and ends with the deepest point of about six meters. The recorded color space has a greenish hue, nearly unaltered by the cameras' automatic white balancing.

As an experimental image set, we used 1560 video frames automatically extracted by JKeyframer. This set represents the original image data extracted from the GoPro videos with a resolution of about 1980×1080 pixels. To optimize the output of VisualSFM, we reduced the image size to 1600×900 pixels. We compared SiftGPU with pre-trained parameters (the standard feature extraction method used in VisualSFM) "M1" with JFeatureManager using the adaptive feature detection method "M2". The comparison between the number of detected features is shown in Fig. 10.

---

[8] http://docs.opencv.org/2.4/modules/nonfree/doc/feature_detection.html
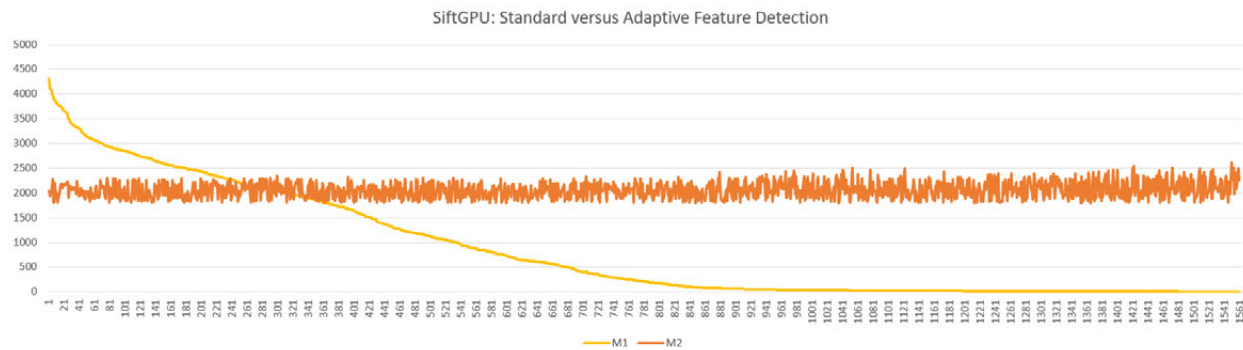
Figure 10. All 1560 frames are plotted on the horizontal axis. The yellow graph shows the number of features, detected by SiftGPU with pre-trained parameters in the set M1. The images are sorted by number of detected features. The orange graph shows the corresponding number of features detected by JFeatureManager. The sum of all detected features by SiftGPU in all frames is 1333287 with a mean of 855 features per video frame, while JFeatureManager detects 3209372 features with a mean of 2057 features per video frame.

The number of features detected using JFeatureManager was amplified by a factor of 2.4 in comparison with SiftGPU. Despite the property, that in good images (left margin of graph in Fig. 10) the amount of features will be reduced, we detected enough features in blurry images (right side of Fig. 10) to increase to chance to find matches. For the original image data set M1, the pose estimation using VisualSFM delivers one connected model (50659 3d points) with 350 images and for M2 one model (170502 3d points) with 1137 images. The adaptive value t was located between 0.001146 (minimum) and 0.007009 (maximum) with a mean of 0.003580. The average calculation time for one image was 23.65 sec.

Such an effective feature enrichment, is, however, a two-edged sword. On the one hand, the feature enrichment process offers the opportunity to integrate more images into one coherent 3d model. But on the other hand, the process slows down and parallel computations need more computing power, due to a quadratic increase of complexity in the matching process. In summary, however, a combination of JEnhancer for prior color enhancement (see Block, Gehmlich and Hettmanczyk in this volume) and JFeatureManager allows us to find more matches between clusters of images (Fig. 11).

This in turn increases the likelihood of obtaining a complete coherent site model. Images of insufficient quality remain a problem, but we believe that we have alleviated the worst constraints for underwater scenarios.
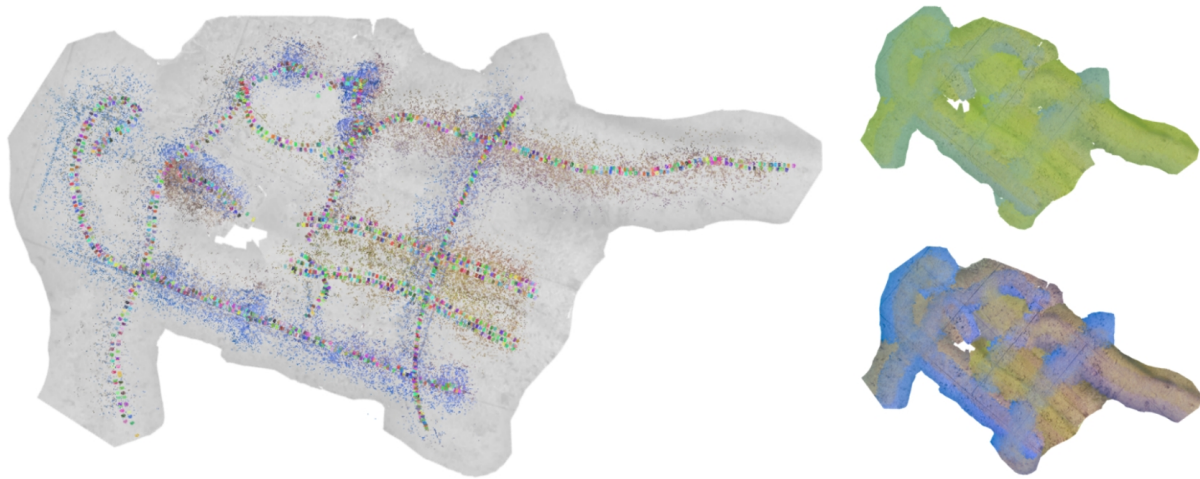
Figure 11. Left: Different 3d models based on different feature matching strategies. The largest model contains 1984 frames, but is still underachieved. Right: The image at the top gives an idea of the original color of the model. As mentioned before, the automatic white balance was active in the cameras, which strongly influences the outcome of color correction using our own JEnhancer software.

6.   CONCLUSIONS AND FUTURE WORK

The image-based 3d reconstruction of the pile dwelling field at "See am Mondsee" was challenging, due to problems with automatic white balancing and a pessimistic recording strategy. Most of the recorded data was blurry and far from suitable for a common SfM reconstruction process. Nevertheless, after replacing the feature extractor of VisualSFM with our own implementation, it was possible to integrate more frames into the 3d model. Experiments have shown that adaptive feature detection can significantly increase the number of detected features in an SfM processing chain, which is of crucial importance for constructing complete 3d models.

The software package JFeatureManager, which uses the adaptive feature detection, is not only designed to replace SIFT. It is in fact a basis for image analysis and a place where different feature extraction methods come together. There are many interesting feature extraction solutions, such as Harris Corner [Harris and Stephens 1988], SURF [Bay et. al. 2008], MSER [Matas et. al. 2004], FAST [Rosten and Drummond 2006], ORB [Rublee et. al. 2011], BRISK [Leutenegger et. al. 2011] or FERN [Ozuysal et. al. 2007], but all candidates for SfM need to be invariant to affine transformations (scaling, rotation, translation, shearing and any combinations of them). The combination of feature enrichment and image enhancement using JEnhancer is a promising approach to underwater videogrammetry.

Another technical challenge was the need for a more accurate georeferencing method for tracking the UUV while recording. We identified and designed a two-stage-technical solution: In the first stage, a GPS buoy is fixed at the cable and placed close to the UUV (Fig. 12). This initial solution is oriented towards improved UUV tracking.
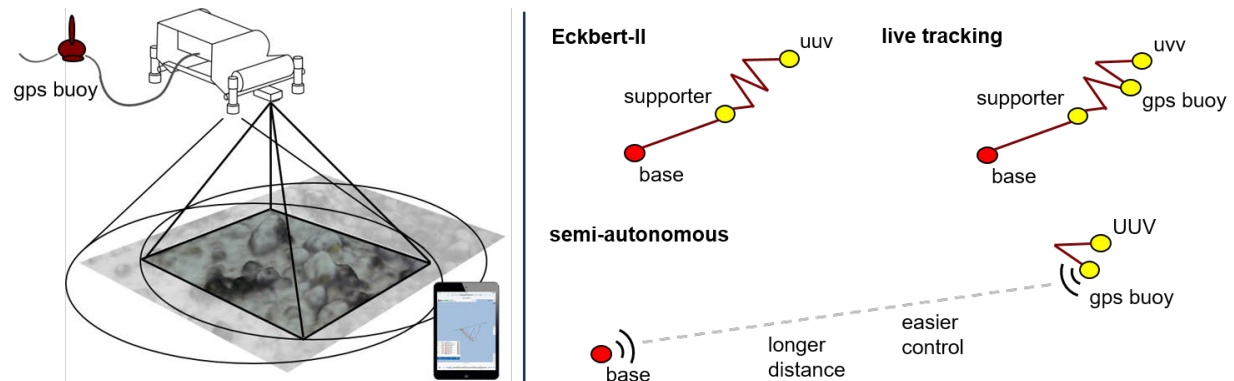
Figure 12. To locate and track the UUV automatically and in real-time, a GPS buoy is fixed to the cable between the supporter and the UUV. To extend the maximum achievable distance between the prospection area and the base, and to replace the cable altogether, we are also working on a fully wireless solution.

Later, in a second developmental stage, we will replace the cable-based setup with a completely wireless communication solution. This is fairly challenging, but the advantages are obvious: The maximum distance between the base and the prospection area would only be limited by the transmission power of the wireless module. The complex and constantly required cable management would be obsolete. The documentation team could be reduced to only two persons. With such a wireless solution, autonomously operating UUVs using fully automated tracking and mapping approaches (with the frameworks ROS[9] or MIRA[10]) would be conceivable.

7. ACKNOWLEDGEMENTS

8. REFERENCES

Caterina Balletti, Carlo Beltrame, Elisa Costa, Francesco Guerra, and Paolo Vernier. 2015. Underwater Photogrammetry and 3D Reconstruction of Marble Cargos Shipwreck. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc V. Gool. 2008. Speeded-up robust features (SURF). In *Computer Vision and Image Understanding (CVIU)* 110, 3 (2008), 346–359.

Marco Block, Benjamin Gehmlich, and Damian Hettmanczyk. 2016. Automatic Underwater Image

---

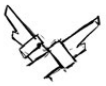[9] http://www.ros.org/

[10] http://www.mira-project.org/joomla-mira/

Enhancement using Improved Dark Channel Prior. In *Journal of Studies in Digital Heritage*. Indiana University. 21st Conference on Cultural Heritage and New Technologies (CHNT 21). Vienna, Austria.

Marco Block, Benjamin Gehmlich, and Dennis Wittchen. 2016. Eckbert-II: Mini-U-Boot zur Sondierung und Dokumentation von archäologischen Fundstellen. *Technical Report*. DD-2016-01, FIITS, HTW Dresden, Germany.

Marco Block, Benjamin Ducke, Estela Mora Martinez, Peter C. Kroefges, Raúl Rojas, and Paulina Suchowska-Ducke. 2015. Low-cost and efficient, UAV-based 3D videogrammetry in Tamtoc/Mexico. In *20th European Maya Conference, The Maya in a Digital World (EMC 2015)*. Bonn, Deutschland.

Paolo Cignomi, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. 2008. Meshlab: an open-source mesh processing tool. In *Sixth Eurographics Italian Chapter Conference*. 129–136. Eurographics.

Cyril Dworsky. 2016. Quer über die Alpen? Die Pfahlbauten im Norden und Süden Österreichs. 4.000 Jahre Pfahlbauten. In *Belegband zur Großen Landesausstellung Baden-Württemberg 2016*, 16.April-9.Oktober 2016, (Baden-Wüttemberg 2016), 119–121.

Marc Pierrot Deseilligny and Isabelle Cléry. 2011. Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images. In *Proc. of the ISPRS Symposium*. 3DARCH11.

Yasutaka Furukawa and Jean Ponce. 2010. Accurate, Dense, and Robust Multi-View Stereopsis. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 8 (2010), 1362–1376.

Benjamin Gehmlich and Marco Block. 2015. Diversity of Flight Strategies in UAV Recording. In *20th Conference on Cultural Heritage and New Technologies (CHNT 20)*. Vienna, Austria.

John A. Greenwood. 1999. Large component deformation studies using videogrammetry. In *The 6th International Workshop on Accelerator Alignment (IWAA 99)* 32, 50 (1999). Grenoble, France. 1-33.

Chris Harris and Mike Stephens. 1988. A combined corner and edge detector. In *Alvey Vision Conference*. 147–151.

Richard Hartley and Andrew Zisserman. 2004. *Multiple view geometry in computer vision. Second edition*. Cambridge University Press.

Holger Hirmann. 1999. *Unterwasserarchäologische Fundstellen in Österreich*. B.A. theses Univ. Vienna 1999.

Michal Jancosek and Tomas Pajdla. 2011. Multi-View Reconstruction Preserving Weakly-Supported Surfaces. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011)*. 3121–3128.

Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. 2006. Poisson surface reconstruction. In *Eurographics Symposium on Geometry Processing (2006)*. 61–70.

Stefan Leutenegger, Margarita Chli, and Roland Y. Siegwart. 2011. Brisk: Binary robust invariant scalable keypoints. In *IEEE International Conference on Computer Vision (ICCV 2011)*.

David G. Lowe. 1999. Object recognition from local scale-invariant keypoints. In *IEEE International Conference on Computer Vision (ICCV 99)* 2, 1150–1157.

Thomas Luhmann, Stuart Robson, Stephen Kyle, and Ian Harley. 2007. *Close Range Photogrammetry: Principles, Techniques and Applications*. Whittles: Dunbeath, UK.

Jiri Matas, Ondrej Chum, Martin Urban, and Tomas Pajdla. 2004. Robust wide-baseline stereo from

maximally stable extremal regions. In *Image Vision Computing* 22, 10 (2004), 761–767.

Jakob Maurer. 2014. Die Mondsee-Gruppe: Gibt es Neuigkeiten? Ein allgemeiner Überblick zum Stand der Forschung. In *Vorträge des 32. Niederbayerischen Archäologentages (2014)*. 145-190.

Michael May, Martin J. Turner, and Tim Morris. 2010. Scale Invariant Feature Transform: A Graphical Parameter Analysis. In *Proceedings of the BMVC 2010 UK postgraduate workshop*. 5.1-5.11. BMVA Press.

Emmanuel Moisan, Pierre Charbonnier, Philippe Foucher, Pierre Grussenmeyer, Samuel Guillemin, and Mathieu Koehl. 2015. Building a 3D Reference Model for Canal Tunnel Surveying Using Sonar and Laser Scanning. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

Pierre Moulon, Pascal Monasse, and Renaud Marlet. 2013. Global fusion of relative motions for robust, accurate and scalable structure from motion. In *IEEE International Conference on Computer Vision (ICCV 2013)*. 3248–3255.

Chris Musson, Rog Palmer, and Stefano Campana. 2013. Flights into the Past, Aerial photography, photo interpretation and mapping for archaeology. In *Aerial Archaeology Research Group, ArchaeolLandscapes (ArcLand)*. ISBN 978-3-00-044479.

David Nistér. 2004. Automatic Passive Recovery of 3D from Images and Video. In *Proc. The Second International Symposium on 3D Data Processing, Visualization & Transmission (3DPVT04)*. Thessaloniko, Greece.

Mustafa Özuysal, Pascal Fua, and Vincent Lepetit. 2007. Fast keypoint recognition in ten lines of code. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007)*. 1–8.

Kimon Papadimitriou. 2015. Course Outline for a Scuba Diving Speciality „Underwater Survey Diver". In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W5K*. 2015 Underwater 3D Recording and Modeling.

Richard S. Pappa, Jonathan T. Black, Joseph R. Blandino, Thomas W. Jones, Paul M. Daneby, and Adrian A. Dorrington. 2003. Dot-Projection Photogrammetry and Videogrammetry of Gossamer Space Structures. In *Proc. The 21st International Modal Analysis Conference (IMAC)*. Kissimmee (FL), USA.

Henrik Pohl. 2016. Drei Jahre unterwasserarchäologisches Monitoring an den österreichischen UNESCO-Welterbestätten. In *Archäologie Österreichs 27/1 (Wien 2016)*. 29–35.

Henrik Pohl. 2014. Erste Ergebnisse und Massnahmen zum Schutz der prähistorischen Seeufersiedlungen in Österreich. In *arcéologie & erosion – 3, Monitiring et Messer de protection pour la sauvegarde des Palastes préhistoriques Autor des Alpes, Altes de la troisiéme Rencontre Internatinale Arenenberg et Hemmenhofen,* 8-10 Oktober 2014 (Hemmenhofen 2014), 69–76.

Marc Pollefeys, Luc Van Gool, Maarten Vergauwen, Frank Verbiest, Kurt Cornelis, Jan Tops, and Reinhard Koch. 2004. Visual modeling with a hand-held camera. In *International Journal of Computer Vision* 59, 3 (2004), 207–232.

Elisa Pruno, Chiara Marcotulli, Guido Vannini, and Pierre Drap. 2015. Underwater Photogrammetry Methods for a Peculiar Case-Study: San Domenico (Prato-Italy). In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

Francesco S. Rende, Andrew D. Irving, Antonio Lagudi, Fabio Bruno, Sergio Scalise, Paolo Cappa,

Monica Montefalcone, Tiziano Bacci, Maria P. Penna, Benedetta Trabucco, Rossella Di Mento, and Anna M. Cicero. 2015. Pilot Application of 3D Underwater Imaging Techniques for Mapping Posidonia Oceanica (L.) Delile Meadows. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

Leopoldo Repola, Raffaele Memmolo, and Daniela Signoretti. 2015. Instruments and Methodologies for the Underwater Tridimensional Digitization and Data Musealization. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

Edward Rosten and Tom Drummond. 2006. Machine learning for high-speed corner detection. In *9th European Conference on Computer Vision (ECCV 2006)*. 430–443.

Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. 2011. Orb: An efficient alternative to sift or surf. In *IEEE International Conference on Computer Vision (ICCV 2011)*.

Elisabeth Ruttkay, Otto Cichocki, Ernst Pernicka, and Erich Pucher. 2004. Prehistoric lacustrine villages on the Austrian Lakes. Past and recent developments. In *F. Menotti (Hrsg.), Living on the lake in prehistoric Europe (London 2004)*. 50–68.

Elisabeth Ruttkay. 1981. Typologie und Chronologie der Mondsee-Gruppe. Das Mondsee-Land. Geschichte und Kultur. In *Ausstellungskatalog zur oberösterreichischen Landesausstellung*, 8.Mai-26.Oktober 1981, Heimatmuseum Mondsee (Mondsee 1981), 269–294.

Roman Scholz, Luka Bekic, and Marco Block. 2016. Shipwreck documentation in Veruda using Photo- and Videogrammetry. *21st Conference on Cultural Heritage and New Technologies (CHNT 21)*. Vienna, Austria.

Noah Snavely, Steven M. Seitz, and Richard Szeliski. 2006. Photo tourism: exploring photo collections in 3D. In *ACM transactions on graphics (TOG)* 25, 3 (2006), 835–846.

Salvatore Troisi, Silvio Del Pizzo, Salvatore Gaglione, Antonino Miccio, and R. L. Testa. 2015. 3D Models Comparison of Complex Shell in Underwater and Dry Environments. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5/W5*. 2015 Underwater 3D Recording and Modeling.

APPENDIX

Table I. A list of selected gesture codes, used for communication between the supporter and the two copilots. Due to the head gear, the snorkel and the distance of the supporter, no verbal communication is possible.

Base		Supporter	
Sign	**Description/meaning**	**Sign**	**Description/meaning**
	Is everything ok?		Yes, everything is ok.
	Check the UUV's status, the cable, the cameras, the lights and so on.		Something is going wrong, copilot 1 and 2 need to take note of the current position and orientation. Back to base.
	What is the UUV's orientation?		Showing current UUV orientation.
	The recording process is finished. Back to base.		Recording of strip is finished. Maybe repositioning.