

Photogrammetric models for marine archaeology

Martin Ludvigsen

Norwegian University of Science and
Technology

Department of Marine Technology

7491 Trondheim, Norway

martin.ludvigsen@ntnu.no

Ryan Eustice

University of Michigan

Department of Naval Architecture and
Marine Engineering

Ann Arbor, Michigan 48109-2145

eustice@umich.edu

Hanumant Singh

Woods Hole Oceanographic Institution

Woods Hole, MA, 02543, USA

hsingh@whoi.edu

Abstract- The paper presents a method to create small scale topographic models of a deepwater archaeological site using photogrammetric methods. An example data set from a Norwegian archaeological investigation is discussed.

I. INTRODUCTION

This paper presents a method to create high resolution spatial models of underwater scenes from optical images. Optical imagery contains more detailed information than its acoustical counterpart, but how can this information be exploited to produce a spatial model? High resolution spatial models could benefit many users' underwater imagery, e.g. forensic investigators, marine biologists, marine geologists and archaeologists. This text will pursue the needs of the latter group.

Constructing photogrammetric models of underwater scenes differs from corresponding terrestrial techniques in several aspects. The vertical extent of features on the seabed is often large compared to the seabed-camera distance due to the limited range of optics underwater. Underwater light suffers from attenuation and backscattering causing challenges like low contrast and non-uniform lighting patterns [1], and [2]. The vehicle carrying the camera also usually holds the light source, making the shadow pattern shift around the depicted objects from image to image. To construct a photogrammetric model of a large area there is a navigation problem to solve because there is a disparity between the accuracy present navigational aids and the resolution offered by imaging sensors.

Archaeologists need information about dimensions, shape, material, structure of surfaces, colors, orientations and positions of objects on an excavation scene to be able to identify artifacts, reconstruct events and understand the scene in an underwater investigation [3]. The properties of light scattering and attenuation in seawater [2] prevent larger sites to be depicted within one frame. Thus, to cover a larger site, a large number of images have to be combined in a model. In underwater archaeology time is an expensive resource, implying that the data acquisition to be efficient. The method must be easy to implement on ROV's or AUV's of opportunity,

requiring the hardware requirements should be modest. The integrity of the measurements is crucial to give credibility to the results, and has to be verifiable.

The methods most commonly used to measure geometrical shape of a site in underwater archaeology are mapping by tape measure [4], photogrammetry by manual point matching [5], and acoustical measurements by single beam or multi beam sonars [6]. These techniques are often combined with optical imagery and preferably put together in a mosaic and in some cases a combined model is constructed [6]. These mapping approaches have major limitations. Using sonars, basic acoustics in terms of frequency, transducer geometry and noise suppression will limit precision and data resolution. The manual methods are too labor intensive to produce a sufficient amount of data points.

Photogrammetry has been used in underwater archaeology for some time [7]. A substantial advantage of the presented method is the automatic interest point identification and processing. Automatic interest point identification and processing is adapted from the science of computer vision [8], [9] and [10] the topics SLAM (Simultaneous Localization and Mapping) and SFM (Structure From Motion). An alternative approach which automatically can produce larger 3-d models underwater is presented by Negahdaripour [11]. On land close range photogrammetry models is an established tool for surveying and industry dimension documentation [12]. In the computer vision community, technology based on feature matching have only been used experimentally for underwater navigation, [13] and [14].

In this paper we will present an automated procedure to construct larger photogrammetric model from sets of still images. The process is automated in the sense that there is no human interaction in the photogrammetric interpretation of the images. The method produces a sparse spatial model and the archaeologist will still need to lean on the optical imagery to complete the historical information content. The presented method will meet the above laid out requirements to

operational efficiency. The reliability of the model is estimated during the construction of the model.

In section II this paper presents the framework and methodology for producing large scale underwater photogrammetric models. The requirements to the data-acquisition process of the images for the photogrammetry are laid out. Local models are constructed from image pairs before they are assimilated into a global model while computing the internal error. Section III describes the data acquisition, the data processing, and the resulting model for a dataset collected at an archaeology excavation at 170 meters depth. Section IV discusses the method and the performance on the data set used in section III.

II. METHODS

This section will describe how 3-d models can be made from a set of optical images. Requirements and guidelines to the image acquisition are offered, together with a description on how the images are put together and how the 3-d information is extracted.

A. Imaging

In order to make spatial models the images need to overlap. When covering a wreck-site it is recommended that ~ 60% of the image foot-print overlap. Likewise across-track there should be ~25% sidelap to be able to close loops in the vehicle trajectory, and to ensure consistency between lines in e.g. a lawn-mower pattern.

The lighting patterns within the images should be as uniform as possible. This is preferably provided by a well configured light(s)-to-camera geometry and sufficient amount of light power, see [2]. If this is not possible an option is to remove lighting patterns by software in post processing e.g. [15]. A high dynamic range camera is useful in this context because a wider color span increases the potential for color correction.

The presented methodology requires a prior pose estimate for all images as position and orientation stamps for all images are used in the model, both in the interest point matching and in the creation of the photogrammetric model.

To provide the necessary relative pose estimate between two images dead reckoning from doppler velocity or inertial measurements is recommended. Most acoustical baseline system do not have sufficient accuracy, while dead reckoning from one image to the next usually allows for high precision due to the short time span between images.

B. Local models

This section describes the image registration pipeline used for producing local models as described by Eustice et al in [8][9][10][18], which is briefly recapped here. The photogrammetric process starts by identifying interest points. The Harris corner detector is used to detect interest points [16] together with the SIFT (Scale Invariant Feature Transform) operator [17]. The Harris operator will detect corners and

edges, and is based on the curvature of the autocorrelation function in the neighborhood of each point. The Harris operator is invariant to rotations and robust to small variations in scale [10].

To match interest points the routine need descriptions of the points. The Zernike moment is used to encode the Harris interest points with a descriptor while the SIFT points are encoded by a SIFT descriptor. In underwater applications changing lighting patterns and rotations are often experienced but the Zernike moment is rotationally invariant and to some level robust to changes in scale.

Comparing the Zernike moments and SIFT operators attached to the Harris- and SIFT-points in both images in an overlapping pair, the interest points present in both images are identified and localized. To ease the requirements for the uniqueness of the interest point descriptors, the pose prior is used to derive regions that are common in both images and where corresponding points are sought. The routine then searches for common points along an area defined by epipolar lines. Outliers are rejected applying a least median of squares (LMedS) registration methodology. The essential matrix E used in the LMedS routine is derived by a six point algorithm [18].

When the essential matrix is defined, the relative pose is estimated using Horns relative orientation algorithm. After a two-view bundle-adjustment, the relative pose is finally estimated. The two-view local model is established by triangulation using a two-view point transfer model [8], and the xyz positions of the interest points are calculated.

C. Global registration

The first step in the global registration is to identify common points in the global model and the local models. The local models are related to each other position-wise by the navigation data.

To determine the scale between two local models a set of distances between points present in both models are calculated for both point clouds. The scale is determined as the ratio between the distances.

By monitoring the standard deviation of the scale estimates, an indication of the quality of the scale match is provided. Outliers of the point clouds are detected by comparing each scale to the mean of scales. The translation necessary to add a local model to the global model is determined in a comparable way.

The distance between the corresponding points in the local model and the global model is computed, and there are as many distances as there are pairs of corresponding points. The standard deviation of the distances is a check for overall quality of the translation and is used to detect outliers. The process of scaling and translating the local model is repeated

and outliers peeled off until an acceptable level of standard deviation is achieved.

A local point cloud will typically overlap the global model at different points or with part of the global model that origin from different local models. Thus there will be two estimates of scale and translation. The deviation between the two estimates will give an indication of the integrity of the model.

When a global point cloud is established, the data are gridded to a mesh and plotted.

III. EXPERIMENTAL DATA

A. Background

In the pipeline route for the Ormen Lange gas field outside the west coast of Norway *Museum of Natural History and Archaeology*, NTNU discovered a historic shipwreck at 170 meters depth [19]. The wreck was believed to originate from the beginning of the 19th century. This triggered a large ROV-based underwater investigation [20]. During the investigation several image sets were collected. It was decided to pursue the possibility of extracting 3-d geometrical information from these data sets.

B. Setup for data acquisition

A 10 x 10 meters excavation frame was purpose built for the project and placed over the stern of the wreck, see Figure 1 and Figure 3. Connected to this frame, the ROV moved along rack rails over the site without touching the wreck or the seabed. The orientation angles of the vehicle were fixed while docked to the frame. With the camera placed on the frame-driven ROV, two image series were taken from the excavation area, one before and one after the excavation. The first series contained 270 images and the second 252 images, both with the camera mounted 1.5 to 2.2 meters above the seabed. In both image sets the imagery was taken in a pattern consisting of 15 and 14 lines containing 18 images each. A high dynamic range camera and continuous light was used. The camera was calibrated in a tank.

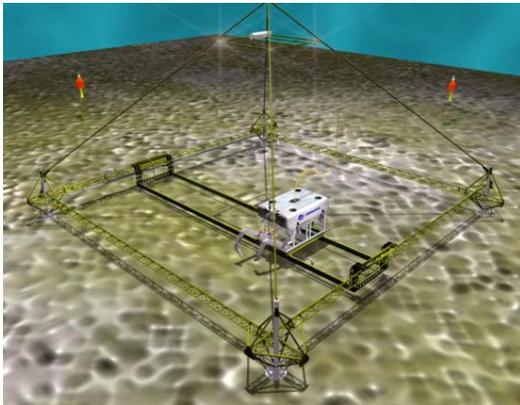


Figure 1 Artistic impression of the ROV in the support frame during excavation. LBL transponders are seen in the background. Image by Fredrik Søreide.

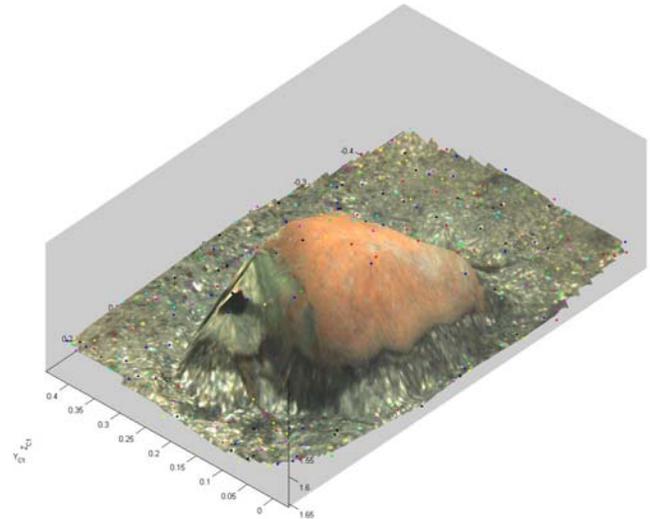


Figure 2 Local model of a large mug laying on the sandy seafloor. A surface was created from the xyz-points and one of the images draped over it. The data points are plotted on top.

C. Data processing

Overlapping imagery was identified from the position data associated with each picture. Applying the camera calibration parameters, the images were undistorted before they were used in the photogrammetrical model. The pose vector was used together with the images to make local models from all overlapping images. Each local model typically consisted of 300 to 1000 xyz points, see Figure 2.

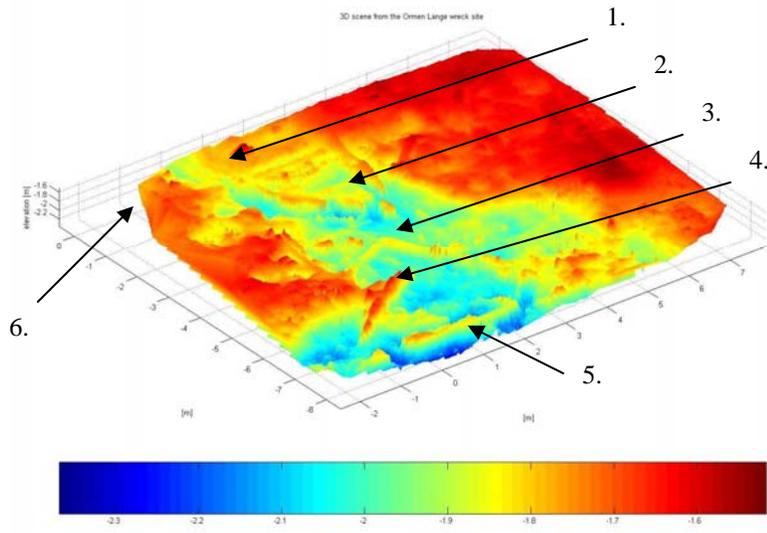
D. Resulting 3-d model

From the two data series more than 400 local models were produced and the local models were joined to two global models by adjusting the scale and translation as described in section II. The topographic model of the site is presented in Figure 4.A. The final model contained approximately 120,000 data points, but due to overlap most points had redundant point, so the actual information content was lower.



Figure 3 Picture of an ROV working on the investigation site. Image by Fredrik Søreide..

A:



1. Mug
2. Hull
3. Wooden plank 1
4. Wooden plank 2
5. Wooden plank 3
6. Sponge

B:

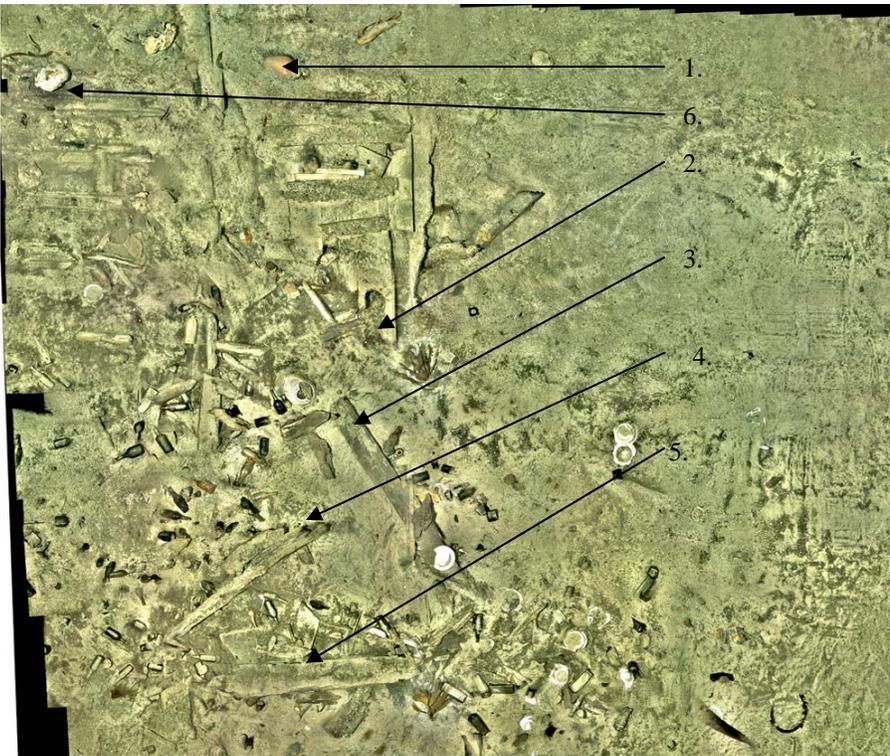


Figure 4.A: A 3-d model of the wreck site color-encoded for elevation. The elevation is local and derived from the camera centers. The grid size used in this mode is 25 mm.

Figure 4.B Shows a 2-d photomosaic made from the images used to construct the elevation model in 4.A. The mosaic is made from 252 images. Some characteristic features are pointed out in both figures.

The images used to construct the model in Figure 4 A are rendered to a photomosaic in Figure 4.B.

When Figure 4.A and Figure 4.B are compared some of the identifiable artifacts are a few wooden planks, an assembled part of wood appearing to be the bottom of the ship, a mug, and a sponge. A depression in the terrain in the middle of the scene is distinctive.

A model of the area before the excavation began was also constructed. When these two models are subtracted from each other, an estimate can be made of how much mass was physically removed during the excavation.

IV. DISCUSSION

At close range photogrammetrical methods have some advantages over corresponding acoustical methods. The method produces high density of data points due to the efficient interest point detector and matching routine. Photogrammetrical methods are not bound by the same limitations as acoustical methods, thus for close range the potential for accuracy are higher for photogrammetry underwater. The data density can be further by tuning the automatic point identification.

In the data set presented the camera-carrying vehicle was docked to a frame. The method presented should work well also with a free flying vehicle, as long as the attitude and position estimates are sufficiently accurate to form the pose prior vector data exist.

The global registration pattern applied here is basic, but during the registration of the example data sets it proved to perform satisfactorily. To add more robustness to the global registration, a more advanced approach will be considered in the future, e.g. a bundle adjustment procedure like the one described in [10].

To improve the scientific applicability, it should be attempted to proceed from the sparse photogrammetric model to a full photogrammetric model as Pollyfeys has discussed for terrestrial applications [21]. In the example data set it is hard to positively identify smaller objects like e.g. bottles merely from the shape of the model, but considering the 3-d model together with the images it is possible to identify dimensions and shape of the modeled objects down to bottle size.

The success of the local modeling depends on the image quality and the precision of the pose priors. In a lawn-mower pattern, some errors produced in the global registration are cancelled out and errors are not allowed to accumulate as local models are put together. The local models will in most cases be matched with models matched several steps back in the global registration process, limiting the aggregation of error.

V. CONCLUSION

We have presented a method to create small scale topographic models of a deepwater archaeological site, based on a SLAM algorithm previously developed by Eustice. We

have shown that the survey method is operational feasible, proved by processing an image set from a deepwater ship wreck.

ACKNOWLEDGMENT

The authors would like to thank Fredrik Søreide for providing the data set.

REFERENCES

1. Singh, H., Howland, J., Pizarro, O., Large Area Photomosaicking Underwater, *IEEE Journal of Oceanic Engineering*, pp. 872-886, vol 29, no 3, 2004.
2. Jaffe, J., *Computer modeling and the design of optimal underwater systems*. Journal Oceanic Engineering, 1990. **15**: p. 101-111.
3. Singh, H, Adams, J., Foley, B.P. Mindell, D., Imaging for Underwater Archaeology, *American Journal of Field Archaeology*, pp. 319-328, vol 27, no 3, Fall 2000.
4. Dean, M., *Archaeology underwater : the NAS guide to principles and practice*. 1992, [London]: Nautical Archaeology Society.
5. Green, J., S. Matthews, and T. Tufan, *Underwater archaeological surveying using PhotoModeler, Virtual Mapper: different applications for different problems*. The International Journal of Nautical Archaeology, 2002. **31**(2): p. 283-292.
6. Singh, H., et al., *Microbathymetric Mapping from Underwater Vehicles in the Deep Ocean*. Computer Vision and Image Understanding, 2000(79): p. 143-161.
7. Rosenscrantz, D.M., *Underwater photography and photogrammetry*. Photography, Archaeology research, 1975.
8. Eustice, R., *Large Area Visually Augmented Navigation for Autonomous Underwater Vehicles*. 2005, Massachusetts Institute of Technology/Woods Hole Oceanographic Institution joint program: PhD-thesis.
9. Eustice, R., H. Singh, and J.J. Leonard. *Exactly Sparse Delayed-State Filters*. in *International Conference on Robotics and Automation, IEEE*. 2005. Barcelona, Spain: IEEE.
10. Pizarro, O., *Large Scale Structure from Motion for Autonomous Underwater Vehicle Surveys*. 2004, Massachusetts Institute of Technology/Woods Hole Oceanographic Institution joint program: PhD-thesis.
11. Negahdaripour, S. and H. Madjidi, *Stereo Imaging on Submersible Platforms for 3-D Mapping of Benthic Habitats and Sea-Floor Structures*. Journal Oceanic Engineering, 2003. **28**(4): p. 625-650.

12. Mikhail, E.M., J.S. Bethel, and J.C. McGlone, *Introduction to modern photogrammetry*. 2001, New York: Wiley. IX, 479 s.
13. Gracias, N. and J. Santos-Victor. *Underwater mosaicing and trajectory reconstruction using global alignment*. in *OCEANS 2001*. 2001. Honolulu, Hawaii: MTS/IEEE.
14. Eustice, R., et al. *Visually Navigating the RMS Titanic with SLAM Information Filters*. in *Proceedings of Robotics Science and Systems*. 2005: MIT Press.
15. Zuiderveld, K., *Contrast limited adaptive histogram equalization*. *Graphic Gems*, 1994. **IV**: p. 474-485.
16. Chris Harris, M.S. *A combined corner and edge detector*. in *4th Alvey Vision Conference*. 1988. Manchester UK.
17. Lowe, D.G., *Distinctive image features from scale-invariant keypoints*. *International journal of computer vision*, 2004. **60**(2): p. 91-110.
18. Pizarro, O., R. Eustice, and H. Singh. *Relative Pose Estimation for Instrumented, Calibrated Imaging Platforms*. in *Digital Image Computing Techniques and Applications*. 2003. Sydney, Australia.
19. Jasinski, M. and F. Søreide. *The Ormen Lange marine archaeology project*. in *Oceans 2004*. 2004. Kobe, Japan: MTS/IEEE.
20. Søreide, F. *Ormen Lange: Investigation and excavation of a shipwreck in 170m depth*. in *Oceans 2005*. 2005: MTS/IEEE.
21. Pollefeys, M., et al., *Automated reconstruction of 3D scenes from sequences of images*. *Isprs Journal of Photogrammetry and Remote Sensing*, 2000. **55**(4): p. 251-267.