CrossMark

ORIGINAL PAPER

A Comparison Between Structure from Motion and Direct Survey Methodologies on the Warwick

Piotr Bojakowski[2,3] · Katie Custer Bojakowski[1,2,3] iD **·
Perry Naughton[4]**

Abstract Structure from Motion, the process of turning two-dimensional digital images into a three-dimensional digital model, is recognized as an emerging method in archaeological research. While some of the previous studies of Structure from Motion applied to underwater projects showed promise as an information rich and affordable survey method, the issue of accuracy remains. This study examined the efficacy of this new technology as a post-processing analytical tool on the early seventeenth-century shipwreck site, Warwick, from Bermuda. Using original digital images from the archaeological excavations, Structure from Motion was tested for suitability and accuracy, and the results compared against the Direct Survey Method. The outcome was an interdisciplinary effort that allowed for a better understanding of the process and the resulting limitations of Structure from Motion for underwater surveys and excavations.

Keywords 3D technology · 3D visualization · Structure from motion · Direct survey method · Warwick project

✉ Katie Custer Bojakowski
kmcuster@gmail.com

Piotr Bojakowski
bojakus@gmail.com

Perry Naughton
pnaughto@eng.ucsd.edu

1   Ashford University, San Diego, CA, USA

2   Atlantic World Marine Archaeology Research Institute, Everett, WA, USA

3   Institute of Nautical Archaeology, College Station, TX, USA

4   Center for Interdisciplinary Science of Art, Architecture and Archaeology (CISA3), Qualcomm Institute, Calit2, University of California, San Diego, CA, USA

Introduction

Maritime Archaeologists are tasked with an important obligation to discover, document and disseminate extraordinary caches of history. Unfortunately, the adverse environmental conditions posed by the ocean limit the types of technological advances that can be utilized. In many cases, the available land-based technology proves ineffective when adapted for use underwater. In other cases, underwater remote sensing or recording equipment is prohibitively expensive, hence only employed by a small number of individual researchers or institutions. Still, there have been some promising advancements in 3D technology and software that can be adapted to the marine environment. One such 3D technology, Structure from Motion (SfM), was selected as a post-processing step in the analysis and reconstruction of the shipwreck *Warwick*. In the most basic terms, the SfM process requires only an adequate coverage of the sites or artifacts with digital two-dimensional photographs, which in turn are converted into a three-dimensional digital model. The process is fast, relatively easy, and inexpensive as it utilizes an open source or commercially available software packages and standard digital cameras in watertight housings.

The English galleon *Warwick* wrecked in Castle Harbour, Bermuda when a hurricane hit the island in November 1619. As expressed in the historical records, *Warwick* belonged to Sir Robert Rich, the Second Earl of Warwick, but functioned during the 1619 voyage as a magazine ship under the mandate of the Somers Isles Company. According to the preliminary analysis and reconstruction, it was a relatively small and lightly-armed merchant vessel of about 180 tons. Almost 400 years after it wrecked, *Warwick* was excavated and recorded by the authors between 2010 and 2012 (Bojakowski and Custer-Bojakowski 2011). The wreck site represented the starboard side of the ship preserved from the turn of the bilge to just above the first deck providing a large intact structure to test new methodologies and refine various techniques of site recording. Although *Warwick's* hull was ideal for this task, any coherent hull remains would serve the same purpose. One of the major goals of the Warwick Project was to collect enough data to recreate the hull remains in 3D. The shipwreck was documented with digital photography, videography and the Direct Survey Method (DSM) (Rule 1989) in addition to traditional, or "by hand," archaeological recoding in 1:1 and 1:10 scale using rulers and measuring tapes. The data was used to facilitate the interpretation of the site and the reconstruction of the complex geometry of the *Warwick's* hull in 3D.

The digital media was not collected with SfM in mind. Due to the fast-paced advancements in technology, the authors used data that was collected for photomosaicing, and later realized that it was possible to also reconstruct the 3D geometry. After the completion of the excavation and reburial of the site, a collaborative project was initiated between Atlantic World Marine Archaeology Research Institute (AMARI) and the Center for Interdisciplinary Science of Art, Architecture and Archaeology (CISA3) at University of California, San Diego (UCSD). The aim was to use the original image dataset collected during the 2010–2012 excavations to build a 3D model of the preserved section of *Warwick's* starboard side. The broader context of this project, however, was to assess the applicability of SfM to the field of underwater archaeology. It should be noted that the 3D modeling was used exclusively as a post-excavation processing tool. Likewise, the research concentrated only on producing a 3D model of the articulated shipwreck structure and not the associated cultural material excavated from around the site. The working parameters of the project were that the 3D model had to be accurate to at least the level of traditional

"hand" recording or better. Given the average budgetary concerns of any archaeological project, it was also necessary to use free or inexpensive commercially available software and equipment that was user-friendly. To test for accuracy, the resulting SfM data was compared against the data from the Direct Survey Method (DSM). The project combined the perspectives of archaeology, engineering, and computer science to determine if SfM was a viable method for rendering underwater photographs into a subsequent 3D model.

Traditional Methodology of Recording Shipwreck Sites Underwater

The Warwick Project officially started in 2008 with a non-disturbance survey of the site. Using archival documents, photographs and original sketches of the site, and personal knowledge of a few local divers; the team quickly relocated the site. A small fragment of the *Warwick's* stern section was uncovered for documentation while a collection of artifacts associated with the site was studied at the National Museum of Bermuda (NMB). This test area was used to analyze the shipwreck and evaluate the extent of preservation of the hull timbers, but most of all to explore the viability of full-scale excavations of the site. The survey concluded with a high-resolution photomosaic and detailed archaeological drawings of the uncovered portion of the hull (Bojakowski and Custer-Bojakowski 2011).

The excavation started in 2010 and continued till 2012. The main goal was to scientifically and methodically uncover and record *Warwick's* hull remains and artifacts. Based on current archaeological standards, these objectives were accomplished by excavating the overburden of the shipwreck and then recording the hull using traditional methods as an intact structure in situ. Such approach was possible because *Warwick* represented a well-preserved articulated starboard portion of the hull, preserved from the turn of the bilge, where the ship broke off, to just below the first deck.

Preliminary examination of the site indicated that *Warwick* was a relatively small merchant vessel of not more than 200 tones burthen, and likely in the range of 180 tons. Its total length was estimated to about 30.5 m while the maximum breadth, at the midship, was about 7 m. Based on proportions of the hull, it was a fast and maneuverable ship with shallow draft making it well adapted to the uncharted waters of the New World. It also appeared that *Warwick* had two decks, and that it was armed for the voyage (Bojakowski and Custer-Bojakowski 2011).

The excavation of the *Warwick* proceeded in three stages: from the very stern section to just before the midship, the midship, and from the midship to the bow section of the ship.
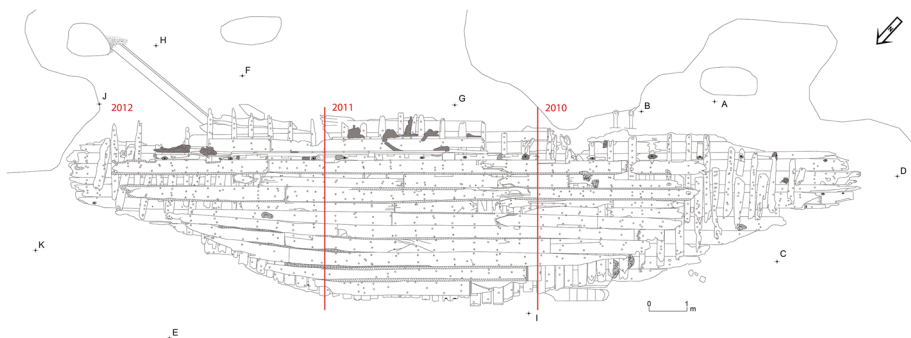

Fig. 1 *Warwick* site plan showing the progress of excavation from 2010 to 2012

Collectively, these three stages constituted roughly a 21 m by 6 m section of *Warwick's* starboard side (Fig. 1; Bojakowski et al. 2011). Using water dredges, the very top layer of the overburden consisting of silt and sand was removed, and the ballast was exposed. The ballast was thoroughly photographed for the purpose of creating a photomosaic and recorded. Below the ballast, there was a layer of thick grey clay that sealed the timbers creating a favorable anaerobic environment that promoted underwater preservation of the timbers.

Since Bermuda Wrecks Board (a government advisory body) did not grant permission to disassemble *Warwick's* hull underwater, each of the three sections of the shipwreck was recorded in 1:10 scale as an intact structure in situ. This was done by assigning small fragments of the shipwreck to different members of the team, who in turn were charged with sketching, manually recording, and transferring all visible measurements on the slates with Mylar (water resistant polyester film) (Fig. 2). In the lab, the sketches and measurements were then meticulously transcribed onto the metric graph paper producing a final hand drawing of a given section. The drawings included all surface details, marks, fasteners, wood grain pattern, but also the scale and other necessary spatial information to easily relate them to each other, to other major ship timbers, and to the fixed datum points. After the completion of all the individual scale drawings, these were transferred to a computer graphics program (e.g. Rhinoceros®, AutoCAD®, etc.), digitized, and combined into a large 2D site plan. The site plan was further tested for accuracy and georeferenced with the fixed datum points. The biggest issue with the site plan, in fact with any underwater archaeological site plan, is that it portrays and distorts a 3D structure of shipwreck as a flat 2D drawing.
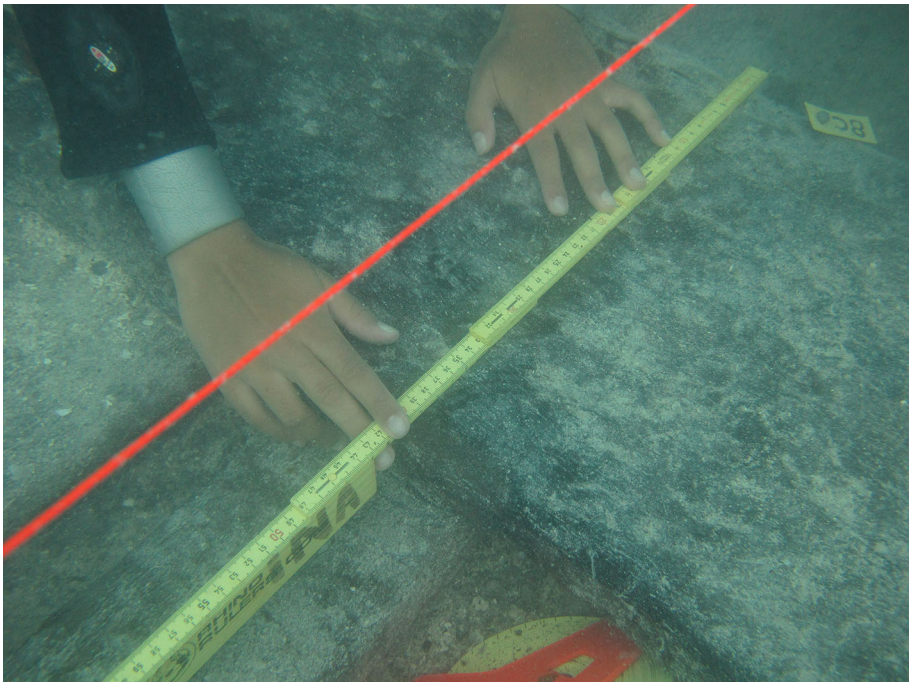

Fig. 2 Traditional or "by hand" recording of the shipwreck's structure

Major ship timbers such the ceiling planks, intact knees, cross beams, filler boards, and deck planks among others, were additionally traced in 1:1 scale on an acetate film underwater. To map framing timbers covered by the preserved ceiling planks, the team adopted a methodology that involved probing through the seams between the ceiling planks to find the edges of otherwise inaccessible framing timbers. This methodology was adapted from the technique developed by Batchvarov and Vallejos for the study of the framing system of the Swedish royal warship *Vasa*. The probing through seams provided information about the thickness of various ship components. Disarticulated or otherwise loose structural timbers were raised to the surface and traced on Mylar in 1:1 scale in two or three views, depending on the physical characteristics of the timbers. Later, these were re-deposited on the site. To capture the geometry of *Warwick's* hull, 18 sections spanning the entire uncovered structure of the shipwreck were taken using the Direct Survey Method (DSM). Although low-tech in nature, this recording methodology provided a great deal of spatial information about the structure of the ship. It also opened a window to more detailed analysis and research using other high-tech solutions.

Direct Survey Method

Underwater recording techniques began to change in the 1970s and 80s with advancements in technology and access to computers in universities and research institutions, and by the end of this period in the home. Maritime archaeologists around the world started realizing that conventional terrestrial techniques did not adequately record and map the complex nature of shipwrecks especially when coupled with issues caused by turbidity, visibility, and current. The impetus of this change was the underwater excavation of shipwrecks with large coherent sections of hull remains such as *Mary Rose*, Henry VIII's flagship which sunk in 1545 (Marsden 2009). The Direct Survey Method (DSM) was developed during the excavation of *Mary Rose* and has been an accepted methodology in maritime archaeology since Rule (1989) demonstrated its effectiveness. Instead of measuring horizontal (2D) distances with the use of a plumb line, DSM (or three-dimensional trilateration) allowed the archaeologist to measure direct distances between survey points in 3D space. This simple, yet capable method was one of the first applications of 3D recording by underwater archaeologists. The least squares estimate of the 3D position of the survey points could be calculated using these direct distances in a dedicated computer software (so-called WEB). This method provided two main benefits. First of all, measuring direct distances between datums became more accurate and efficient than trying to keep measurements planer. Second of all, solving the least squares solution in 3D space gave an estimate of the error in each measurement, hence the confidence intervals on the estimated locations of the survey points. Collectively, these benefits significantly increased the efficiency and efficacy of underwater surveys (Rule 1989, 1995).

Following up on this survey method, Holt (2003) published on the typical accuracy of this methodology. Through the experiments, a standard deviation of 25 mm was achieved. More interestingly, Holt (2003) concluded that most of the errors came from the actual transcription process by examining that the residual, or the amount of error in each measurement, was uncorrelated with the length of the measurement. Around 20 % of the measurements were outliers, a surprisingly large number. While this statistic can be reduced by experience (to about 10 %), the error rate demonstrates the difficulty of survey

work underwater, the diligence needed for effective data collection, and the importance of properly trained crew.

Datum Points

Using a standard DSM methodology with the aid of WEB computer program, the excavation and mapping of the *Warwick* site proceed in two steps. The first step involved the positioning of 11 primary datum points (designated with the letters A through K) around the site. The role of these main reference points was to define the perimeter or the maximum extent of the excavation area. Once secured, the direct distances (for which line of sight could be established) and relative depths were measured and entered into the WEB program. The second step involved the positioning of 92 secondary datum points within the physical structure of the shipwreck (Fig. 3). It should be noted that the mapping methodology of the internal structure of *Warwick* followed an earlier work by Rule and Adams who visited the site 1988 (Rule 1995). Initially, they mapped one section across the site with a combination of 14 distances, 10 slopes, and 17 offsets between 22 datums. During the 2011 and 2012 field seasons, the work was expanded into 18 sections (or profiles) along known frame stations spanning the entire preserved structure of the shipwreck. Each section had two to six secondary datum points permanently placed along its length. To provide a unique identification, the datums were designated with a distinctive number, which referred to a number of the section, and a letter A through G. For example, section number 4 had five, or in fact six, datums: 4A, 4B, 4C, 4D, 4E, plus an additional datum associated with a lodging knee K4. Similar to the primary datums, the direct distances among as many as possible secondary datum points were measured and transferred to the WEB program. To be statistically valid, the number of measurements had to be at least three (basic triangulation); however, more measurements were crucial to increase the statistical significance. To achieve the third dimension (along the *z* axis), the relative depth was measured at each datum point using a measuring tape tied to a floating buoy (Fig. 4).

Once computed inside the WEB program, a few of the measurements had to be rejected while others redone to correct for errors. One of the most common errors was simply a sagging of the measuring tape over longer distances. Second most common was a transcription error. Statistically, the maximum negative residual, or the amount of error, for 849 direct distances between the fixed datum points was −25 mm while the maximum
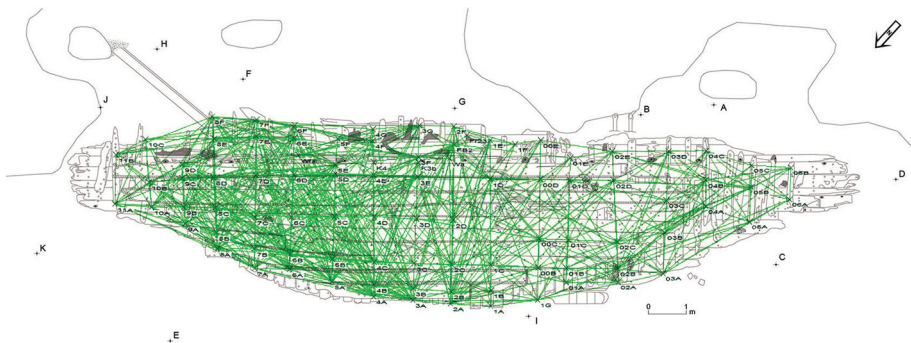

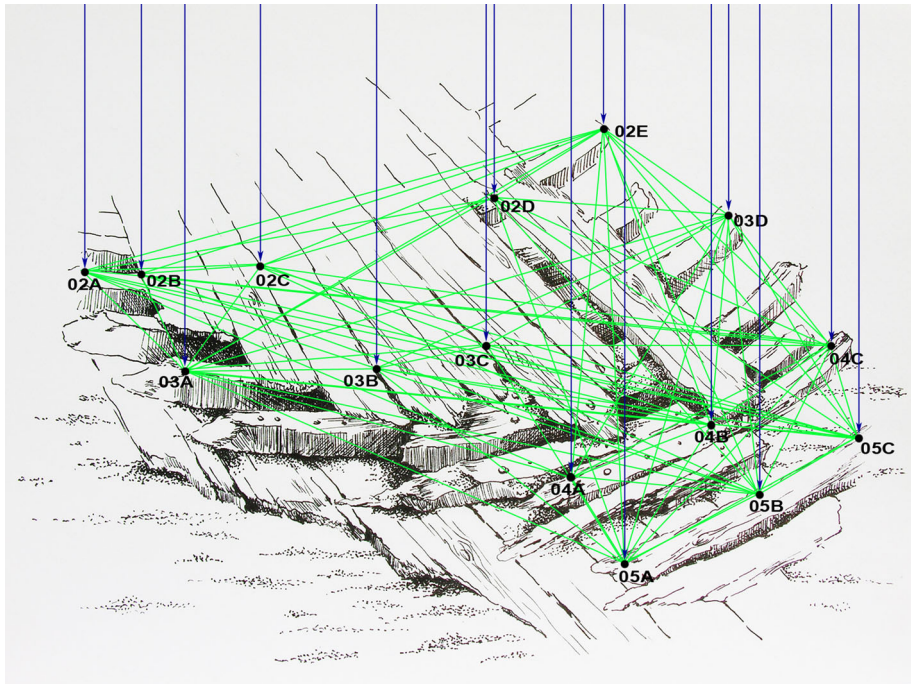**Fig. 3** DSM of 92 secondary datum points and 849 direct distances interpolated over the site plan

Fig. 4 Sketch of the site showing DSM (using WEB program) methodology between profile 02 and profile 05 at the stern; *green lines* show direct distances, *blue lines* show relative depths. Each profile consists between 3 and 5 datum points (Color figure online)
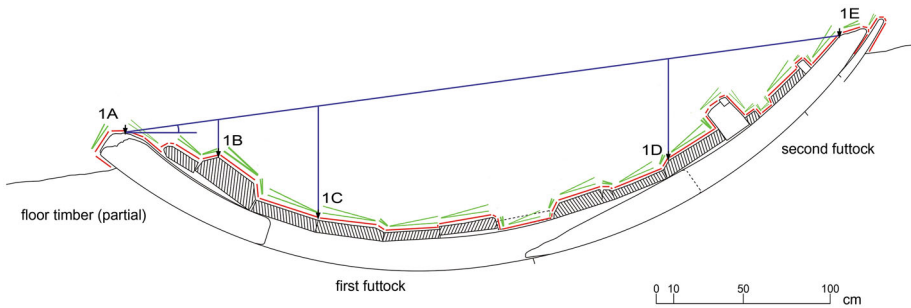

Fig. 5 Sketch showing the steps of traditional recording of the profiles, here profile 1, with offsets from a base line (*blue*), with angles along each surface (*green*), and with direct distance measurements along all the visible surfaces (*red*). The profile is also recorded with the exact position of the datum points (Color figure online)

positive residual was 32 mm, with the average absolute residual of 4 mm. The maximum negative residual for relative depths at the datum points was −21 mm, the maximum positive residual was 29 mm, and the average absolute residual was 8 mm. Finally, the maximum average residual associated with the physical location of each of the 92 datum points surveyed across the preserved structure of the shipwreck was 10 mm and the average absolute residual was only 4 mm. For the purpose of the Warwick Project, these

average residuals were satisfactory and provided for great accuracy of the 3D relationships among the datum points. The data was then exported from the WEB program to Rhinoceros® modeling software as a 3D cloud of points.

As one team was working on the DSM, another was recording the same 18 sections in 1:10 scale in a traditional manner with offsets from a base line, and with angles and direct measurements along the visible surfaces (Fig. 5). The offsets were measured with a plumb line from the base line to a given datum point. The angles were measured along all visible surfaces with a high precision digital goniometer. Likewise, the direct distances were measured along all visible surfaces, seams, and gaps with a ruler. Although time consuming, the recording provided a highly accurate representation of all features along the sections, including the datums themselves. The key, however, was the actual geometrical shape of the sections.

When the scale drawings of all 18 sections were completed, they were scanned and imported to Rhinoceros®. One by one, they were placed in the background of the screen and digitized as 2D line drawings. Using the datum points as guides, the section drawings were then aligned with the DSM cloud of points producing a remarkable 3D curvature of the entire *Warwick's* starboard side (Fig. 6). This process allowed not only for subsequent modeling in Rhinoceros®, but also facilitated the research and reconstruction of the missing bottom and port side of the hull. It culminated in the set of contour lines, in all three views, which after multiple fairing and readjustments lead to the development of a digital model of the ship (Fig. 7).

Photo and Video Recording

Traditionally, maritime surveys relied on hand drawings to convey the structure of an excavation. Photography, and more recently digital photography, became an informative way to document what structures looked like underwater. The images can record the individual features or the spatial context among multiple features of the shipwreck. They can be taken from different angles, distances, using different cameras, camera settings, and
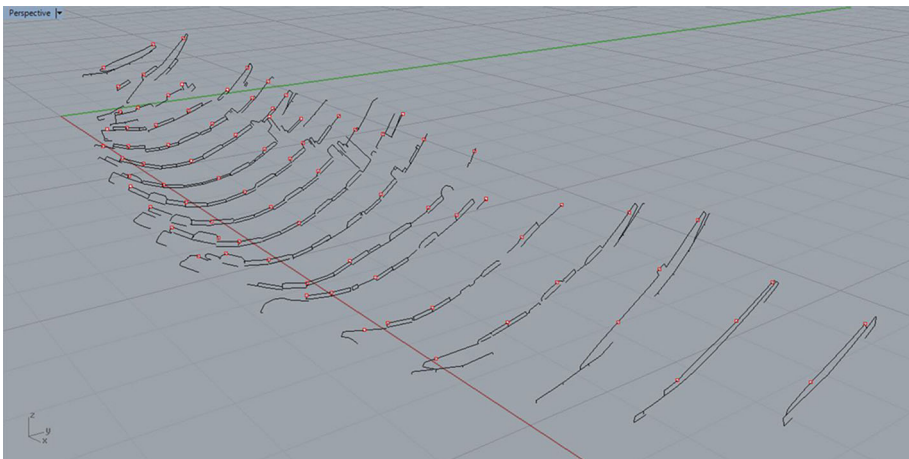

Fig. 6 DSM cloud of points (*red*) exported to Rhinoceros® and matched with corresponding 18 profiles producing a 3D representation of the preserved *Warwick's* starboard side (Color figure online)
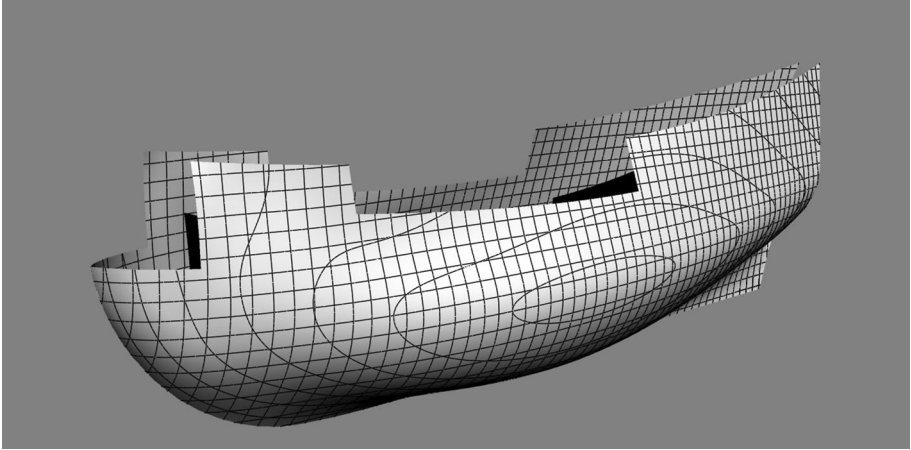
Fig. 7 A digital 3D model of *Warwick* based in part on the geometrical data produced through DSM
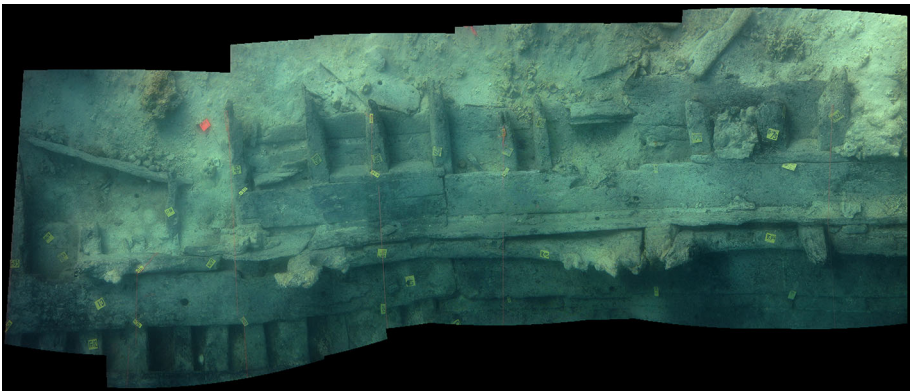

Fig. 8 Fragment of a photomosaic of the upper midship section of *Warwick*

light sources. They can also be processed in a variety of ways to provide information otherwise indistinguishable to a working diver. Digital photography is not just a recording tool; it is simply an integral part of any underwater survey or excavation. A fully-charged camera is a basic archaeological tool. The standard procedure on the Warwick Project was to use the first dive of the day, before any bottom sediment was disturbed, to photo and video record the entire site. It was time consuming but worth the effort as it provided an amazing visual catalog of the work progress, which in turn could be used to recreate the excavation as it proceeded day by day. In addition, hundreds of photographs were taken throughout the day to record the work inside individual sections of the shipwreck (Fig. 8).

Unfortunately, photography has always had a limited field a view. In many cases, a single photograph could not capture the whole site or an artifact. In other cases, the field of view hindered a proper understanding of the site scale or how everything fit together. To overcome this, archaeologists began constructing large multi-photo panoramas, generally known as photomosaics, from carefully overlapped images yielding a large photograph of

an entire site. Digital videos are particularly helpful for photomosaic purposes. They allow not only the extraction of individual video frames as images, but also provide for better control of the overlap between the frames. The most common underwater photomosaics are similar in nature to aerial photographs or "bird's eye" view and are produced with the camera facing vertically down. The process makes the site easier to understand visually and facilitates georeferencing the photos. If the correct scale is assigned, this two-dimensional (2D) representation of the site can also be used for measurements and mapping.

In the past, a lot of technical skills were required to make a good photomosaic. With advances in computer software, however, the process of blending individual images became simplified. There are many commercial software packages in which a user can upload a set of images and with a click of a button get a basic panoramic view of the site. Other software can be more complex and provide numerous adjustments, optimizations, or projection options. It is worth emphasizing that the physical properties of water, even under excellent visibility, can create problems that limit the detail and contrast. Error in photomosaics can also compound; a tiny error while blending two images can create much larger errors as more images are added to the panorama. Just as these problems are being remedied with new generations of computer graphics programs, SfM models are quickly displacing photomosaics as the next natural step. SfM has the potential to provide dense 3D structure in a way that is both visually informative and also geometrically accurate. Figure 9 shows a comparison of the different information forms now available to archaeologists.

Structure from Motion

Structure from Motion (SfM) compares multiple camera views to estimate the 3D geometry of an object. At a high level, it starts by picking out thousands of salient points in each image. It then finds matches of salient points between images, determining which images have overlap. With knowledge of point matches the software estimates the relative position and orientation of each frame. Finally, the 3D location of the points in the image can be deduced through triangulation. In general, SfM provides a geometrically accurate representation of a scene, with the caveat that there is no scale. When using a single camera, it is ambiguous whether a camera is moving a small distance around a small object or a large distance around a large object. For this reason, some scale must be defined by using known distances or dimensions of one or more objects. In other words, physical measurements of an object are still crucial to provide a correct scale. There are multiple implementations of SfM software, both free (Visual SfM,[1] Autodesks 123D Catch,[2] Bundler[3]) and for sale (Agisoft[4]). In this study the authors chose to use Agisoft because we have found that it is the most comprehensive and easiest to use.

There have been a few pilot studies of 3D vision technology being applied in maritime archaeology. Henderson et al. (2013) collected images from a custom stereo camera set up equipped with a strobe, an inertial measurement unit, and a GPS receiver. They swam transects about 2 m above the submerged city of Pavloperri off of the coast of Greece. The transects they used had about 50 % overlap between them. The data was post-processed

---

[1] http://ccwu.me/vsfm/.

[2] http://www.123dapp.com/catch.

[3] http://www.cs.cornell.edu/∼snavely/bundler/.
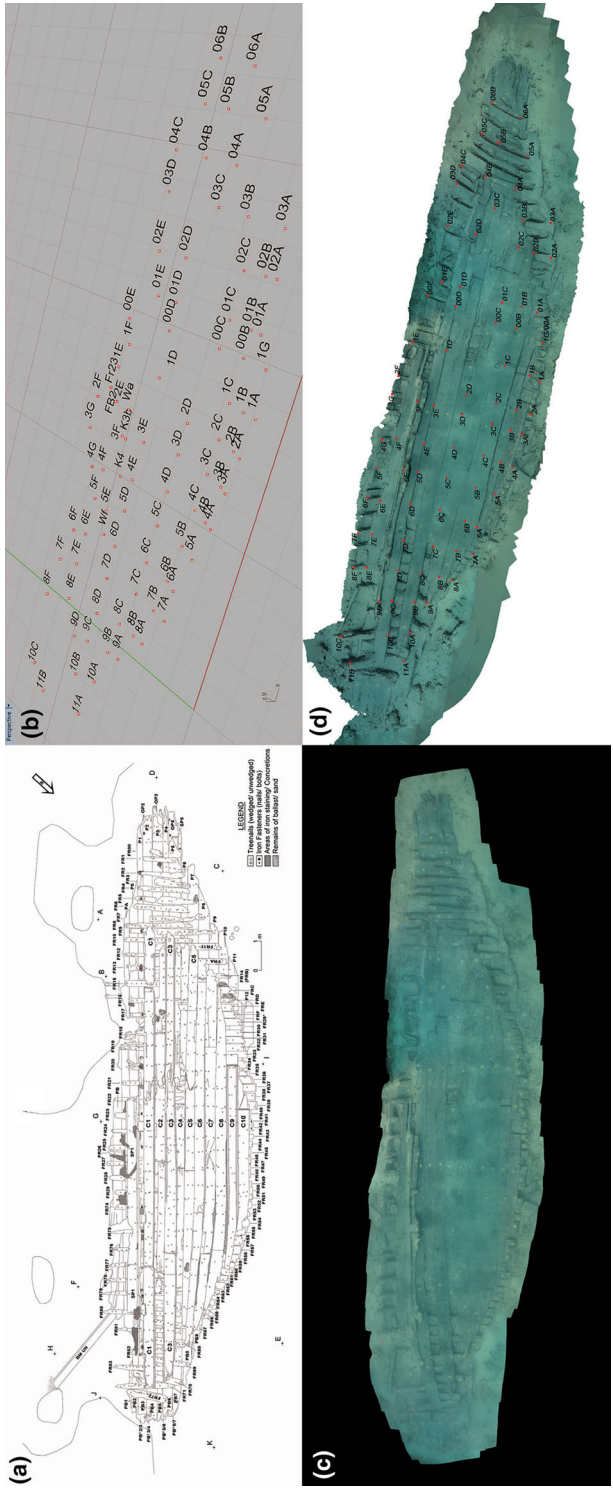
[4] http://www.agisoft.com/.

Fig. 9 The four different data representations available to maritime archaeologists: **a** Hand drawn result after years of measurements and excavation. **b** DSM model output, a sparse yet accurate point cloud. **c** Photomosaic, visually appealing but only representative in 2D. **d** Structure from Motion (SfM) model, both visually appealing and 3D

using a Simultaneous Localization and Mapping (SLAM) algorithm that they developed which estimates the position and orientation of the cameras at each time step and uses this information to build a geometry of the site (Mahon et al. 2008). The authors were able to map an area of about 450 m^2 and compared their SLAM survey with points collected from a total station with the largest error being 200 mm.

There have also been surveys of complex, intact shipwrecks. Mertes et al. (2014) used a diver-propelled rig with six 50-watt video lights, a dive computer for depth, a level, and a viewfinder that gave the diver feedback. They swam about 10, 25-min surveys with a constant height above the wrecks *Hetty Taylor* and a schooner *Home*. These wrecks lie at the bottom of Lake Michigan at 32 and 53 m of water, respectively. After constructing a point cloud using Agisoft they exported the model to Autodesk's 3DS Max to fit the model to a hand survey that they completed in previous years. After finding an optimal scaling, translation and rotation to fit the 3D model to the hand survey they compared the discrepancy between different parts of the model. They found that the percent difference in the error was inversely proportional to the length of the measurements. In other words, the longer the measurement, the higher the percentage of error between hand-measured lengths and lengths taken from the model. As such, the method worked best on smaller measurements, hence smaller objects or sites.

Not all SfM work has been done with expensive camera rigs or other technology that price out many archaeology budgets. McCarthy and Benjamin (2014) demonstrated a pragmatic method that relies only on consumer digital cameras and housings under ambient light conditions. They presented results on surveys with varying visibility off of the coasts of Denmark and Scotland. A 1-m scale bar was placed next to their artifacts to define a scale. Measurements were also taken using a tape measure for comparison. Their conclusions differed from the results of Mertes et al. (2014), they claimed that the error in measurements on the 3D model was mostly determined by where in the model the measurement was taken. They admitted that they found discrepancies as high as 133 %, but measurements taken close to the scale bar were acceptable and did not depend on the length of the measurement. The difference in the conclusions of Mertes et al. (2014) and McCarthy and Benjamin (2014) are probably the result of the different ways the two authors chose a scale for their model, using a 1 m bar versus scaling the model to best fit a hand survey. While these studies have shown promise in the use of 3D vision technology, they have not fully examined the use of a similar method to a complex archaeological excavation such as the Warwick Project.

Methodology

Collecting Images

The first and most important part of the SfM methodology is the collection of images. Without a "good" dataset of images, a complete and accurate model cannot be constructed. A couple of qualities make up a "good" dataset. The first is image clarity. The majority of image descriptors that are used to pick out salient points rely on having sharp contrast in texture. Imaging underwater presents many challenges in collecting sharp photos. Lighting, visibility and back/forward scatter can all negatively affect the clarity of the image. The photographer needs to be experienced with underwater imaging. Another important metric that determines the effectiveness of the dataset is coverage. There have been many "rule of thumb" descriptions of how much coverage is necessary for a complete model such as the 50 % between frames quoted in Henderson et al. (2013) or that each point needs to be seen

in at least three images.[5] While there may not be an agreed upon optimal amount of coverage, it is clear that too much coverage is better than too little. Notice that there are two parts of the coverage. There is the overlap between the images and the amount of coverage of the actual site both of which need to be adequate, albeit still not fully defined, to ensure a comprehensive model.

Preprocessing Images

Geometry SfM assumes that the camera capturing images follows a pinhole camera model. There are many aspects of consumer cameras that are different from the pinhole camera model. The lens creates a radial and tangential distortion and the camera's sensor is rarely mounted perfectly parallel to the camera's imaging plane. This problem has been studied extensively in computer vision and robotics communities. Subsequently, there are many examples of calibration software that correct for these distortions. Calibration software usually relies on taking pictures of a known calibration object, a popular one being a chessboard, from different angles and distances. The software then detects the corners of the chessboard in each image and calculates the calibration parameters (i.e. focal length, radial distortion, tangential distortion). Agisoft has the functionality to compute calibration parameters with pictures of a chessboard. Additionally, Agisoft has a feature called optimize photo alignment, which will try to compute the calibration parameters using the estimated structure of an object (such as the *Warwick* wreck) instead of a calibration target. While using the optimize photo alignment may seem more appealing to a user because it eliminates the extra step of collecting calibration data, it is preferable to collect calibration data because using a known object is more accurate. Once a camera is calibrated each pixel in an image defines a ray that extends from the focal center of the camera through that pixel and out into the world. The images must be representative of this capture model in order for the process of triangulation to produce accurate estimates of a point's 3D location.

Housing The choice of waterproof housing can also violate the pinhole camera model. The difference in the index of refraction (related to the speed of light in the medium) between the water, the material of the enclosure and air between housing and the camera causes the light to bend at these interfaces. There are two types of housings, dome ports and flat ports. Dome ports are usually more expensive but if used properly the images suffer less from the light bending at the enclosure. If using a flat port in front of your camera lens, the distortion will noticeably change your field of view and will strongly violate the pinhole camera model. In this case, the calibration procedures mentioned above may not adequately correct for all of the distortion caused by the housing interface. Correcting for these distortions is still an active area of research and theory and some software can be found in Treibitz et al. (2012) and Agrawal et al. (2012).

Color Light is absorbed by water at rates that are dependent on the depth and the color of the light. As a result, many images appear blue or greenish because the water absorbs most of the red light. There have been different techniques to post-process the images in order to get an image of true color. Additionally, lenses can cause vignetting where the outer edges of a frame appear darker because they receive less light from the camera lens. Software correcting for vignetting is also available. In our workflow, none of these calibrations where used.

---

[5] Based on the recommendations of the Agisoft software at http://www.agisoft.com/.

Annotating Images

Once an image dataset is finalized, the next thing is to annotate images with ground control points (GCPs). A GCP coordinate frame must be decided beforehand. In our case, we chose the coordinate frame that was defined by the DSM survey described earlier. Figure 10 shows screen shots of the annotation process, for which Agisoft provides a convenient interface. Once a datum is tagged, the coordinates for that datum can be entered, defining the scale of the final SfM model and helping the reconstruction accuracy.

Bundle Adjustment

Once the images are annotated, it is time to build the geometry of the model. While there are many implementations that do this, they all have the goal of minimizing the reprojection error. The reprojection error is the error, in pixels, of the observed pixel locations of a matched point in each image from the expected pixel location given an estimate of camera parameters and the location of the point in 3D space. It is difficult to find a solution that satisfies this minimum globally, and many results will give a suboptimal solution. We have found Agisoft's bundle adjustment to give consistent solutions compared to other software, a main reason why we chose to use it.

Model Representations

Once an initial geometry is built, there are several different processing routes depending on the intended use of the model. The point cloud that was generated by the bundle adjustment
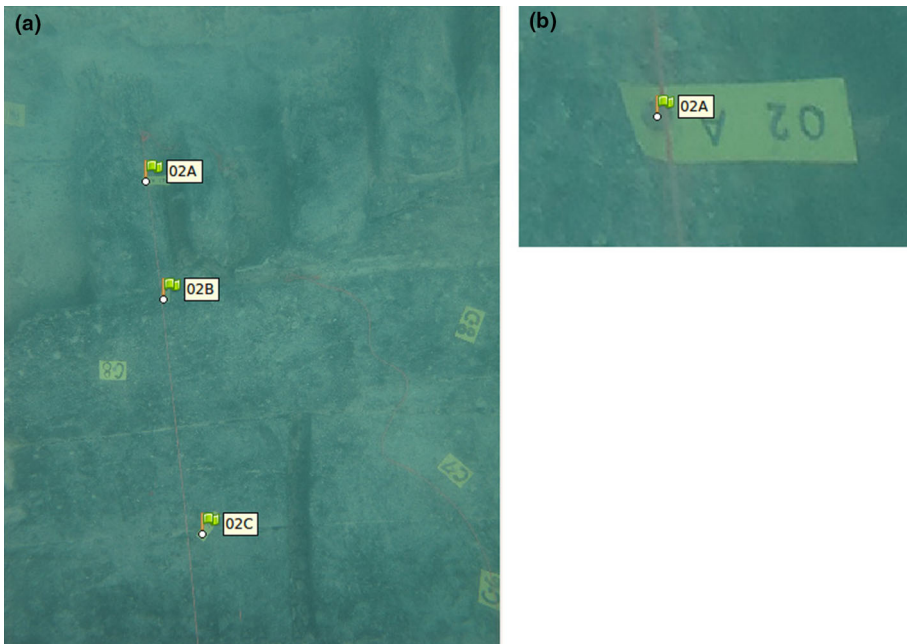

Fig. 10 Photographs presenting: **a** A screenshot showing the annotation of each image with the datums in Agisoft, **b** a close-Up of 02A datum

can go through a dense reconstruction process where parts of the original images that did not pass the original point detector (either from a cap on the number of point detectors computed or because of the texture level) can be reconsidered, matched and triangulated so that the point cloud can include more points. An alternative to this dense cloud is to compute a mesh of the original point cloud. A mesh triangulates a surface in between the points that were computed. After a surface is computed, texture from the original images can be applied to the surfaces. In either case, the result is a photo-realistic model. We can also mesh the dense point cloud, but this is often computationally taxing and in some cases computationally infeasible.

The choice of which processing path to go with depends on how one wants to represent the data. For example, with the proliferation of 3D printing, it is often desired to print a scaled down artifact. This requires a mesh model because it defines surfaces, not just free points in space. On the other hand, we typically opt for a dense point cloud for viewing in a virtual environment. We have found that this gives more detail in the 3D geometry of sites and artifacts when compared to meshing the sparse point cloud.

Structure from Motion Accuracy

An important consideration of the Warwick Project was that the images were not gathered with the intention of generating a 3D model. We believe that this has large implications for this technology. Sites such as *Warwick* are often reburied to protect them, making re-excavation politically difficult, expensive, and often times simply impossible. Fortunately, taking image or video record of a shipwreck is a common practice and all of the conclusions of this work can be extended to other sites that have adequate coverage of a site in digital photography. In our case, the images were gathered using a Canon Powershot G11 with an ISO of 80 with the intention of providing a visual record for the site and creating a photomosaic. Calibration data was not available because the photos were not taken with the intention of 3D vision. We did not implement any of the pre-processing steps to the images because we were less concerned with the true color of the shipwreck and we did not have calibration data for the camera.

To test the accuracy of the SfM method we first annotated all of the datums that were found in the images that were collected as per Fig. 11 so that when Agisoft generated a sparse point cloud of *Warwick* it would have each datum picked out. Next, we built SfM models giving Agisoft the locations of a subset of the datums that were computed from the DSM survey. We withheld another subset of ten datums that were never given to Agisoft to scale the model. These withheld points were then used to determine the accuracy of SfM. We compared the estimated 3D location given by Agisoft to the 3D position calculated in the DSM survey, which we took to be the ground truth of the locations of the datums on *Warwick*.

We compared this method for different subsets of points in order to get a feel for how our GCPs affected the accuracy of the model. We used this as an analogy to the question, "if this were to be used as a survey tool, what location information would we need to collect to make this model accurate?" These different cases are highlighted in Fig. 11 where the red points are the ten points that were withheld from the comparison and the points of other colors were input into the SfM model for that case. We chose points to input into the SfM model that varied in noticeable ways. For example, in Fig. 11b we only use the points along one vertical transect of *Warwick* and in Fig. 11d we used a horizontal
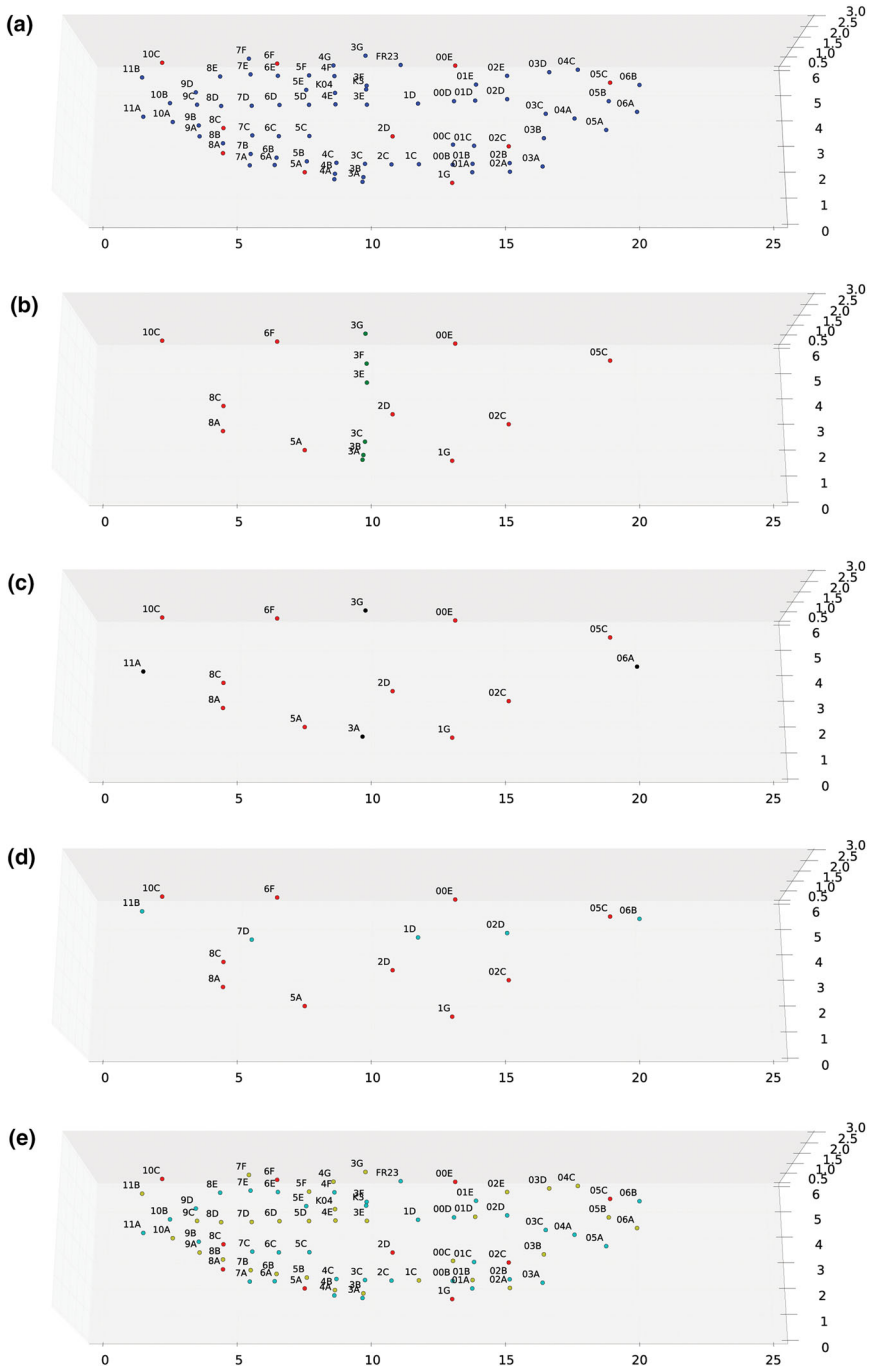
Fig. 11 Graphs presenting the experiments taken for SfM. In each graph the *red points* represent the datums that were used to compare against the DSM model. The other points represent datums that were input into the SfM method (Color figure online)

Table 1 Results from the SfM experiments

Name	Av. error	Max error	Max datum	Min error	Min datum	Av. X	Av. Y	Av. Z
1 A	0.096	0.4146	10C	0.014	5A	0.0123	0.0286	0.0893
1 A Opt	0.0682	0.2773	10C	0.0088	8C	0.0108	0.0231	0.0611
2 A	0.0961	0.4156	10C	0.0142	5A	0.0123	0.0287	0.0894
2 A Opt	0.0751	0.3314	10C	0.0082	8C	0.0118	0.0267	0.0664
1 B	0.1313	0.3808	10C	0.0173	2D	0.027	0.0337	0.122
1 B Opt	0.1757	0.5053	05C	0.0098	2D	0.0266	0.0377	0.1675
2 B	0.1308	0.3803	10C	0.017	2D	0.0267	0.0336	0.1215
2 B Opt	0.1741	0.4995	05C	0.0098	2D	0.027	0.0377	0.1658
1 C	0.0957	0.3811	10C	0.0292	8C	0.0159	0.0318	0.0851
1 C Opt	0.0921	0.3466	10C	0.0334	02C	0.0197	0.0304	0.0819
2 C	0.0959	0.382	10C	0.0292	8C	0.0159	0.0319	0.0853
2 C Opt	0.0921	0.3463	10C	0.0342	02C	0.0196	0.0304	0.0819
1 D	0.0963	0.3557	10C	0.018	1G	0.0125	0.0331	0.0872
1 D Opt	0.0881	0.353	10C	0.0121	02C	0.0137	0.0279	0.0811
2 D	0.0965	0.3563	10C	0.0182	1G	0.0125	0.0331	0.0873
2 D Opt	0.088	0.3529	10C	0.0124	02C	0.0136	0.0281	0.081
1 E B	0.1	0.419	10C	0.0161	5A	0.0125	0.0298	0.0931
1 E B Opt	0.0815	0.3421	10C	0.0107	8C	0.0119	0.0273	0.0732
1 E Y	0.1001	0.42	10C	0.0163	5A	0.0126	0.0299	0.0933
1 E Y Opt	0.0816	0.3421	10C	0.0108	8C	0.0119	0.0275	0.0731
2 E B	0.0937	0.4129	10C	0.0113	02C	0.0124	0.0289	0.0865
2 E B Opt	0.0779	0.3342	10C	0.0103	02C	0.0122	0.0281	0.0686
2 E Y	0.0939	0.4139	10C	0.0115	02C	0.0124	0.029	0.0867
2 E Y Opt	0.0778	0.3342	10C	0.0104	02C	0.0122	0.028	0.0686

The first column indicates the test number (either 1 or 2) followed by the Fig. 11 that the experiment corresponds to. Also, in the case of (e), which color it is corresponding to in the plot (either (B)lue or (Y)ellow)

transect. Notice that the last figure, Fig. 11e, has two different experiments. We first ran the model with half of the available datums and then with the other half to see if there was a noticeable difference.

Table 1 shows some statistics on the results of this experiment. It is important to realize that the SfM is not deterministic and can result in slight, or sometimes significant, deviations in the answer it provides. To show this deviation each experiment was run twice, indicated by either the 1 or 2 in the 'Name' portion of the table. The letter A–E in 'Name' corresponds to the A–E in Fig. 2 and defines which datums points where input as GCPs. The Opt, indicates the result after the optimize photo alignment feature of Agisoft was used. Lastly the 'Y' and 'B' correspond to the yellow and blue dots in E that correspond to different experiments.

The error associated with this experiment shows some promise, but is not to the level of the DSM survey. We can see that the error is robust to which GCPs we gave. When we gave all of the available datums besides the ones that were used for the comparison (A on the table) the average error was 9.6 cm with a maximum error of 41 cm and a minimum
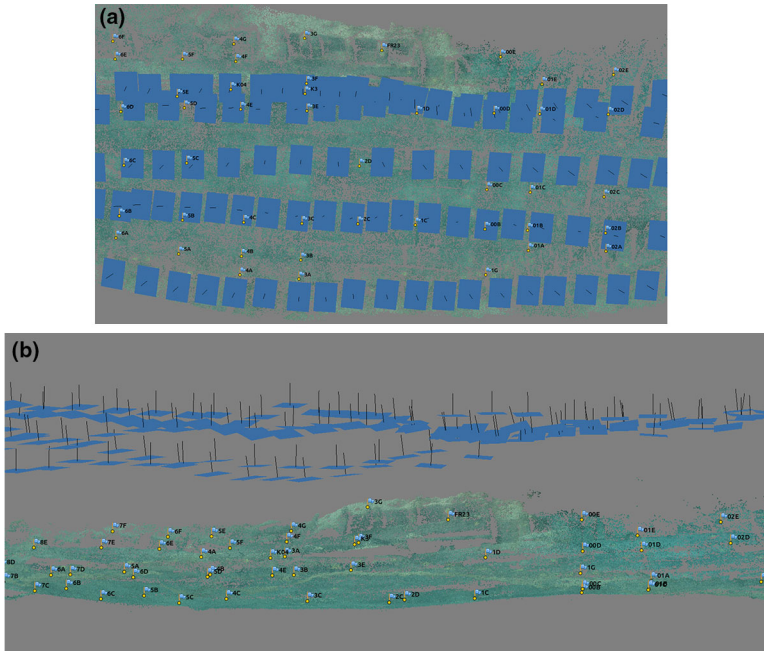
Fig. 12 Images showing the estimated camera positions from Agisoft with the *blue square* representing the imaging frame: **a** a *top view* of the coverage, **b** a *side view* of the coverage. Note that the *blue rectangles* are not representative of the entire image footprints. They represent only a scale version of the images (Color figure online)

error of 1 cm. While the worst average error was reported as 13 cm (coming from B configuration), which had a maximum error of 38 cm and a minimum error of 2 cm.

There are a couple of trends that are important. First, the average error holds the following relationship, x < y < z while this seems surprising at first. It can be explained by Fig. 12. Figure 12a shows the coverage in the x–y plane. It is clear that the overlap in the x direction is much greater than the y direction. We believe that this is the reason why the accuracy is better in the x direction. Figure 12b shows a side view of these camera orientations. We can tell from this figure that there are no camera angles that would capture distance in the z direction. While SfM is able to rectify the depth information in this case, we believe that the accuracy in the z direction would be improved with these camera views. It is important to note that these vertical camera views are much harder to come by because of the physical structure of *Warwick*.

Another interesting aspect of these results is that the 'optimize camera alignment' feature in Agisoft does not always improve the resulting model. We believe this is because the camera model that Agisoft assumes does not adequately describe the extra distortion cause by the enclosure interface. We can also see that the results generated from different tests produce similar results. This shows that although Agisoft cannot guarantee the exact solution every time, it does a good job at coming up with very similar solutions. This shows great promise in this software because it indicates that it can consistently find a good solution.

Another test that we ran is described in Fig. 13 where we plot the error of each datum in the set of ten versus the number of times that datum is projected into an image. This is a
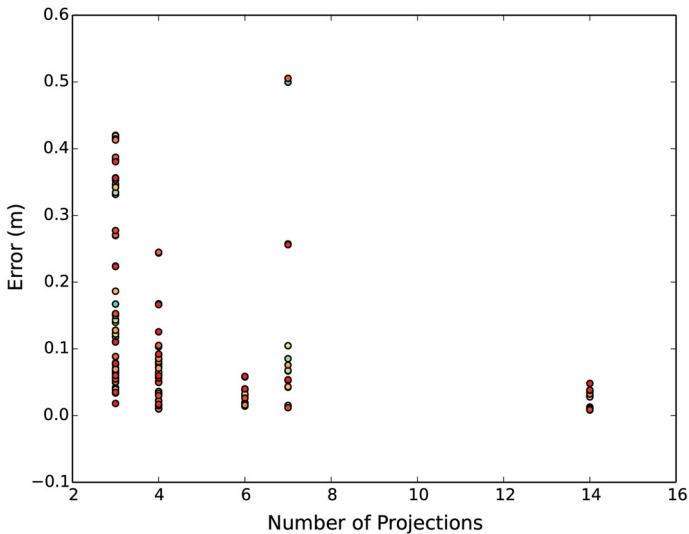
Fig. 13 Graph of the error for each point in all of the test cases versus the number of times that point is found in an image

rough measure of overlap. We can see that there is a slight trend in that the more times you see a point in an image the higher the accuracy will be. This is an important consideration and stresses the need for a detailed methodology that ensures adequate coverage of the site or artifact with digital photography. There is no easy metric to gauge the amount of coverage until after the processing is complete. Great care must be taken to take clear photographs and many views of the same points. This is a difficult consideration in the application of SfM as a survey tool. The most notable deviation from this is in datum with seven projections, if you cross-reference this with Table 1 you will notice that the outliers at seven correspond to the datum 5C and that these errors happen after the 'optimize photo alignment' step. This again demonstrates that this feature of Agisoft does not quiet work well with images taken underwater with a waterproof enclosure around it.

Methodological Considerations for Underwater Archaeology

Researchers found that within the context of terrestrial heritage management, the SfM methodology is fast and detailed without compromising the application of existing field-work techniques. The method can easily be applied by archaeologists and heritage professionals without a technical background and can be used to provide rapid and systematic recording of archaeological sites (Plets et al. 2012). The same holds true for underwater archaeological sites. Photo and video recording is not only a standard procedure within maritime archaeology, any project participant with an understanding of underwater photography can handle it. However, the issues of accuracy still persist when the method is applied underwater. Although SfM creates remarkable site maps and digital representation of artifacts and entire shipwrecks, it is not yet accurate enough to be used as the only recording technique, particularly for measurements to be performed off site or after the project.

It is clear from this study that SfM would not be an adequate replacement for DSM with the results given; SfM returns an average error of around 10 cm where DSM is an order of magnitude less (4 mm on average). However, there are many reasons why this is not a fair comparison and we should not write SfM off as not being a viable survey tool. We were not able to utilize many of the calibration procedures that we explained in the methodology section because calibration data was not collected (keep in mind that the authors did not consider SfM when these images were taken). Also, it is generally accepted to place scale bars and other known geometric objects in the scene while the collection of digital data is performed. This adds very precise geometry to the estimation of the model. Both of these have potential to increase the accuracy of the model.

Another important consideration in the comparison between DSM and SfM is the amount of time spent on each methodology and the amount of feedback that was utilized. In performing the DSM survey, we were able to input our measurements into the WEB program and examine the measurements that seemed off so that we could re-measure parts of the site. Without this feedback, DSM would not have been as effective. With SfM, we were only given one shot at the data collection; there was little feedback in the data collection process. Figures 14 and 12 demonstrate the importance of this limitation. With the information these figures provide it would be beneficial to go back and record more photographs of *Warwick* around the datum point 10C so that we could give the software more images to work with. As evident from the comparison of Fig. 12 and the relationship
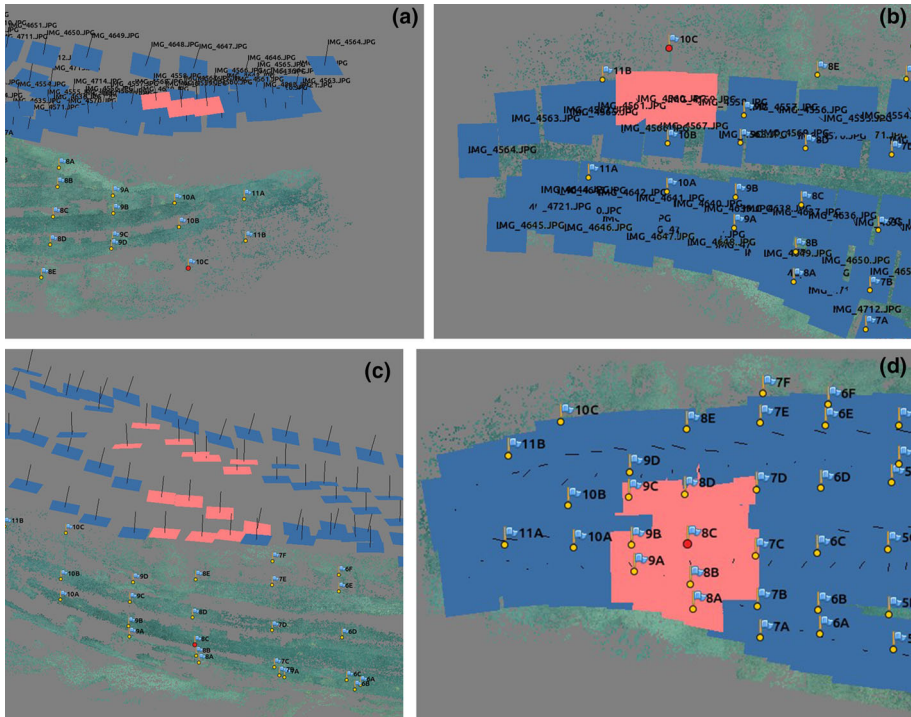

Fig. 14 Difference in coverage of the site with photographs: camera locations that contain datum 10C (**a**, **b**), and camera locations that contain datum 8C (**c**, **d**). Notice that there are more images that overlap with 8C than 10C and that there is more variance in the *z* depths of the cameras that contain 8C in the images

between the error in the x, y and z direction, we would receive better position estimates if we took images more densely in the y direction and more variation in the z direction, and possibly more oblique angles of the hull. Unfortunately, this data collection is next to impossible now. Admittedly, even if we were collecting imagery with SfM in mind it is still not easy to receive feedback. SfM is a computationally intensive method, and generating the full *Warwick* model (approximately 200 images) takes hours on a high-end personal computer, which may not be available at many field sites. As the number of images increases, the processing time goes up dramatically. However, as with most technology, we believe that this limitation will be quickly eliminated from advances in computing power.

Improving the accuracy of the method and software for use with underwater photography will have a significant impact on the field of maritime archaeology. Once a higher level of accuracy is reliably obtained and systematically tested across several different archaeological sites representing common underwater conditions (e.g. high versus low visibility, or various depth levels) SfM could revolutionize underwater heritage management. The implications for deep-water archaeological surveys, cultural resource management, academic surveys and excavations, and state and national institutions are significant. A SfM method optimized for underwater archaeological work would allow an institution to initially record and monitor changes to shipwrecks or other cultural features in a cost-effective, user-friendly, and accurate manner such as it is currently employed for terrestrial archaeology.

A combination of traditional underwater recording methods (for accuracy) and SfM (for visualization) is recommended until the SfM methodology becomes more mature. The main advantages of DSM is the simplicity of the recording method and the accessible software. Although DSM requires a complex web of measurements, the method is easily managed with a systematic approach to recording individual measurements. The WEB software is user-friendly and allows the researcher to input the daily measurements and to find and correct any resulting errors in a timely manner. As previously mentioned, the most common issues were human errors such as slack in the measuring tape or transcription mistakes, both of which are easily corrected. In comparison, it is often tough to process a full SfM model in the field, so feedback on data collection is limited.

The transfer of scaled drawings of individual shipwreck features (e.g. sections) and the DSM data into Rhinoceros® 3D modeling software allows a researcher to produce a remarkable three-dimensional rendering of the feature or site while maintaining accurate geometry. This part of the method is more complex and does require training with the Rhinoceros® software; however, the software is accessible to the average archaeology student or professional. While this study showed promise in the use of 3D vision technology as an information rich and relatively affordable survey method, the SfM technique was examined only as a post-processing analytical tool. Now, the question is if it can also be used as a stand-alone recording technique without being coupled with hand recording or DSM. If the cumulative (or average) error stays in the range of centimeters, than the technique might be applicable only to larger sites where such error would render statistically insignificant. However, for small sites with a lot of internal structural details such as *Warwick*, this level of precision is inadequate. In that case, SfM must be combined with other recording techniques that can capture these details with millimetric precision.

Regardless, the ability to quickly capture even very complex structures for an impressive visual representation makes SfM an effective 3D vision method. It can be particularly useful for 3D representations of previously excavated archaeological sites or heavily degraded objects. As on *Warwick*, where the images were not taken with the SfM in mind but only later applied to produce a 3D model, numerous archaeologists have similar data sets of pictures from the old excavations. Granted that there is enough

coverage, SfM enables archaeologists to represent these sites in 3D providing a new and exciting way to engage the public in the process of discovery and the interpretation of maritime archaeological sites. With more studies and refinements of the existing methodology, it is imminent that the 3D vision techniques will become standard archaeological methods for underwater surveys or excavations.

Compliance with Ethical Standards

Conflict of interest   The authors declare that they have no conflict of interest.

References

Agrawal A, Ramalingam S, Taguchi Y, Chari V (2012) A theory of multi-layer flat refractive geometry. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). S.l.: s,n, pp 3346–3353

Bojakowski P, Custer-Bojakowski K (2011) The *Warwick*: results of the survey of an early 17th-century Virginia Company ship. Postmediev Archaeol 45(1):41–53

Bojakowski P, Custer-Bojakowski K, Inglis D (2011) The Galleon *Warwick*: excavation 2010. INA Annu 2010:50–55

Henderson J, Pizarro O, Johnson-Roberson M, Mahon I (2013) Mapping submerged archaeological sites using stereo-vision photogrammetry. Int J Naut Archaeol 42(2):243–256

Holt P (2003) An assessment of quality in underwater archaeological surveys using tape measurements. Int J Naut Archaeol 32(2):246–251

Mahon I, Williams SB, Pizarro O, Johnson-Roberson M (2008) Efficient view-based slam using visual loop closures. IEEE Trans Robot 24(5):1002–1014

Marsden P (2009). Your noblest shippe: anatomy of a Tudor warship (Archaeology of the Mary Rose: vol 2). Mary Rose Trust, UK

McCarthy J, Benjamin J (2014) Multi-image photogrammetry for underwater archaeological site recording: an accessible, diver-based approach. J Marit Archaeol 9(1):95–114

Mertes J, Thomsen T, Gulley J (2014) Evaluation of structure from motion software to create 3d models of late nineteenth century great lakes shipwrecks using archived diver-acquired video surveys. J Marit Archaeol 9(2):173–189

Plets G, Gheyle W, Verhoeven G, De Reu J, Bourgeois J, Verhegge J, Stichelbaut B (2012) Three-dimensional recording of archaeological remains in the Altai Mountains. Antiquity 86:884–897

Rule N (1989) The direct survey method (DSM) of underwater survey, and its application underwater. Int J Naut Archaeol 18(2):157–162

Rule N (1995) Some techniques for cost-effective three-dimensional mapping of underwater sites. BAR Int Ser (Suppl) 598:51–56

Treibitz T, Schechner YY, Kunz C, Singh H (2012) Flat refractive geometry. IEEE Trans Pattern Anal Mach Intell 34(1):51–65