Applying Low Budget Equipment and Open Source Software for High Resolution Documentation of Archaeological Stratigraphy and Features

Undine Lieberwirth,
Free University Berlin, Germany

Bernhard Fritsch
Humboldt-University of Berlin, Germany

Markus Metz, Markus Neteler
Fondazione Edmund Mach, San Michele all'Adige, Italy

Kerstin Kühnle
Dresden University of Applied Sciences, Germany

Abstract:
Available technology and measurement methods currently exceed the requirements of archaeological documentation at excavation in terms of precision, accuracy and resolution. Hence, archaeologists face the challenge of deciding not only what and how to document in archaeological terms, but also what degree of precision, accuracy and resolution are necessary and which tools are the most suitable for their purposes. This paper discusses controversial opinions and methods to face this challenge. On one hand, if we are not concerned with price, there is a wide variety of software and hardware technology to meet requirements far beyond archaeological limits. On the other hand, a growing community supports low cost developments in this field, and advances have now reached a point where Open Source (OS) results can compete with those from proprietary devices. As stated above, results and purposes are strongly related to aims and means. Documenting archaeological stratigraphy requires different levels of precision and resolution than documenting architecture, objects or finds. The challenge here is to consider purpose, possibilities, costs and managing accruing amounts of data. This paper aims to act as a kind of guide for archaeologists and historians to find the most suitable documentation equipment out of a wide range of possibilities.

Keywords:
Structure-from-motion, laser scan, FOSS, computation, documentation

1. Introduction

Technological developments in the field of archaeological documentation and historical building research have increased in recent years at a rapid pace. Survey results will not only result in higher resolutions and more precise measurements, but these high-quality documentation procedures are normally also cost-intensive. Because of the latter fact, these methods are often beyond the budget of an archaeological project. Since these unfortunate constraints often set boundaries to scientific research, alternative methods are generally very welcome.

An option, which has already been widely accepted in archaeology, is the application of Open Source Software under the GNU General Public License GPL (Free Software Foundation 2014a). The ideology also meets the general aims of scientific work by applying the most transparent, independent research in terms of methodology and processing.

In this paper we aim to show and prove that these boundaries can be dissolved by alternative low-budget applications which can compete, in terms of quality, with established methods.

The low-cost equipment of this study consisted of a digital camera and two different software solutions for processing (one Open Source software, one proprietary software). The applied method for documentation was the so-called "structure-from-motion" (SfM), which was first successfully used in archaeological contexts two years ago (Ducke, Score and Reeves 2011).

The results of the SfM method used in this study were compared with those terrestrial 3D laser scanner (TLS) and its proprietary control and processing software, the use of which has already been established in archaeology. The first applications of a 3D laser scanner in archaeological research and monument management can be traced back one decade (Doneus, Neubauer and Studnicka 2003; Sachsen.de 2012). Since then, it has proven its quality beyond any doubt. For this reason, in this study the results of the TLS method were used as a reference for estimating the quality of the SfM outcome.

Both the photogrammetric SfM method and the TLS produce georeferenced 3D point clouds. Since these are in the same co-ordinate system they can be compared precisely in a 3D space.

The test took place in a variety of archaeological contexts in order to gain experience in practical use in different conditions and environments.

In this paper we give an overview of the results of this methodological comparison in terms of precision, quality, handling and costs from a user's perspective. We also discuss the advantages and disadvantages of the chosen low-budget methods BundlerTools (BT, Open Source) and VisualSFM (VSFM, proprietary), and in this context the newly developed software tools (GRASS GIS modules) (GRASS Development Team 2012) so that future users are able to choose the optimal way of documentation relative to their questions and needs.

1.1 The Archaeological Research Question

Archaeological digital documentation is now widely understood to result in a digital 3D ground plan including all features, stratigraphical surface information, architecture, finds, objects etc., expressed as vectors mainly in CAD-based systems.

Approximately 10 years ago colleagues in Europe (Doneus and Neubauer 2004; Doneus and Neubauer 2005; Doneus, Neubauer and Studnicka 2003) started to document archaeological surfaces using 3D laser scanners, thereby producing 3D point clouds which recreate archaeological 2.5 D surfaces. This precise documentation of continuous surfaces represents a great step in archaeology. It completes the former vector models with a precise and nearly gapless surface description.

Until then, point clouds generated by LiDAR (acronym of Light Detection And Ranging) or total stations were normally the basis for creating digital elevation models (DEMs) in Geographic Information Systems (GIS) in archaeology. With the introduction of 3D laser scanners, archaeologists got the chance to model much smaller areas precisely (i.e. the top and bottom surfaces of archaeological stratigraphy).

However, since 3D laser scanners are expensive and require proprietary analysis software, applications of this kind will be restricted to a few exceptional projects. Therefore, the aim of this study was to look for alternative methods to document archaeological, stratigraphic surfaces, architecture, features, etc., for further processing and/or analysis in a GIS or other software environment.

For testing the strength of the methodology in archaeological terms, we used three different archaeological environments.

Case Study 1

The first case study took place at the Western Porticus of the Main Forum in Ostia Antica, Italy. The stratigraphical surfaces we targeted contained material from the Late Roman Antiquity (100 – 200 CE), and consisted of ceramic objects, stones and marble slabs, bricks, floor pavement and floor filling, screed material and sediments. Since some of the stratigraphical layers were very thin and could hardly be recognised, precise measurement of the top surfaces covering the whole area was of great importance.

Corresponding author: undine.lieberwirth@fu-berlin.de

390

The study trench had an expanse of 2 x 3m and at its deepest spot approximately 1.3m. The documentation of one archaeological layer (SE201) was performed using by a terrestrial 3D laser scanner (Leica ScanStation 2) and a digital camera (Canon EOS D60). According to the on-site situation, the adjusted point cloud resolution for the laser scanner was 1 x 1cm. Due to the depth and accessibility of the trench the scanner's position had to be changed twice.

In order to execute the SfM methods BT and VSFM, overlapping pictures of about 20-25 images/ layer were taken with different digital cameras. The calibration of the cameras in the process of creating the 3D point cloud is included into the SfM software. The resulting point count and resolution is dependent on the quality of the pictures (see Paragraph 2 and Strecha et al. 2008). Thus, no extra step is necessary to prepare the data.

Case Study 2

For the second case study, a very different environment was chosen: a medieval cellar made of cobble stones in the Archaeological Park at Freyenstein, Germany. Again, the aim was to document the architecture of the preserved cellar walls in 3D quickly, economically and precisely. The documentation should be suitable for creating a construction drawing and generating several post-excavation analyses, e.g. measuring of details, creation of vertical and horizontal sections. We used the same equipment for the TLS, BT and VSFM methods as in case study 1. In order to execute the SfM methods BT and VSFM, overlapping pictures of about 20 images/wall were taken with the digital camera by a size of 3x5 m.

Case Study 3

The third study generated a 3D point cloud of an object. The aim was to find a suitable solution for 3D objects and features we encounter in archaeological environments. In this example, we chose a full-size portrait of an Antique Roman emperor with an uncertain identity (probably Antigonus II). The idea was to compare the digital 3D model of the sculpture with digital 3D models from other portraits in terms of physiognomic similarities in order to verify its identity.

For a third time, we used the same equipment for the TLS, BT and VSFM methods as in case study 1 and 2. In order to execute the SfM methods BT and VSFM, overlapping pictures of about 30 images of the complete portrait were taken with the digital camera.

1.2 The Methodological Research Question

In this contest we were looking not only for the best digital reconstruction result but also for other parameters which effect archaeologists' decisions by choosing the most suitable method for their special needs according to the aims of the documentation. The two alternative SfM methods we compared with the established 3D TLS method should therefore fulfil the following methodological requirements.

The methods should be able to compete in terms of:

- precision and quality with 3D laser scanner documentation (Precision & Accuracy).
- price should be affordable within a budget typical for an archaeological project (Costs).
- method should use Open Source Software because of its ideology to reveal its source code.
- algorithms used, which are vital for serious scientific research (Open Source & Open Access, Learning Curve Hardware & Software).
- easy to transport equipment since archaeological sites are often off-road (Equipment Size & Weight).
- few working steps for saving time since excavations, whether for rescue or research, are always time limited (Working Steps).
- processing time of the models which should be appropriate to execute them on-site for verification (Processing Time).
- being appropriate for a wide range of features, not only restricted to detailed architecture and not restricted by price to well-funded projects.

Given that applications in science should fulfil the requirements of being comprehensible, verifiable, and repeatable, access to the algorithms used in software tools as well as their free access for research meant we were biased in favour of the SfM method (Rocchini and Neteler 2012).

2. The Documentation Processes

During the tests on different archaeological sites, the laser scanning results acted as a reference model since we assumed this well established method is widely recognised and hence not controversial.

2.1 Method 1 – Photogrammetry: BundlerTools (BT)

A couple of years after the introduction of terrestrial laser scanners in archaeology, the method of image-based modelling, using a photogrammetric method to generate 3D point clouds of objects and architectural features and later called Structure-from-Motion (SfM), has started to become more common in the same subjects (e.g. El-Hakim et al. 2008). Research in this field (e.g. Debrevec, Taylor and Malik 1996; Pollyfeys et al. 2000; Snavely, Seitz and Szeliski 2006; 2007) over the past decade has led to the development of different software packages and web services. The choice of tools is still growing and new developments are announced almost monthly.

After the release of Bundler (Snavely, Seitz and Szeliski 2006; 2007) under the GNU General Public License and CMVS/PMVS2 (Furukawa and Ponce 2010; Furukawa et al. 2010) to create dense point clouds, usability and/or velocity are constantly being improved in such programs as VisualSFM (Wu 2007; Wu et al. 2011; Wu 2014), RunSFM (Ho 2014), SFMToolkit (Visual Experiment 2014), APERO-MICMAC (MICMAC 2014) or Phython Photogrammetry Toolbox (Arc-Team 2014) free available for academic use.

At the same time, such proprietary software solutions as Agisoft PhotoScan (Agisoft 2014) or aSPECT3D (ArcTron 3D 2012) have been developed. The unusual fact that software has been developed specifically for archaeological purposes shows the great attraction of this method in the field of archaeological data acquisition.

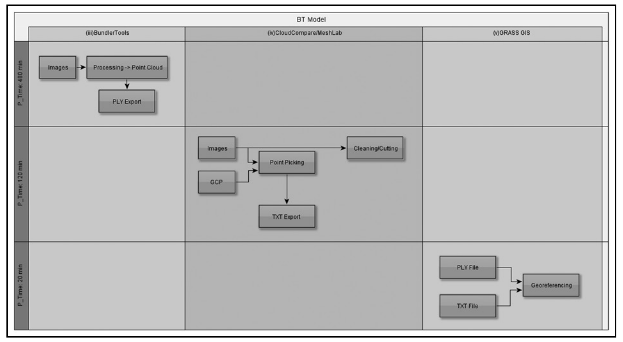


Figure 1. *Workflow Method 1, BundlerTools, with Processing Time (P_Time) and Working Steps within the different software packages.*

Additionally, there are several alternative web services like ARC3D (ARC3D 2013) and Autodesk ReCap (2014), which require an internet connection which is not available on every excavation.

Disregarding the chosen processing software, the method of Structure-from-Motion requires a set of unsorted digital images, marked and measured control points and software to calculate and create the 3D point cloud model. The process of work contains five steps (Fig. 1):

- First, taking pictures of the object or archaeological layer (surface). This requires nothing more than taking care to shoot a good image (sharpness, light conditions etc.) with a minimum overlap of 60 %. Control points also have to be placed in such a way that they are visible in as many pictures as possible.
- The second step is to measure the ground control points (GCP). They are vital for the georeferencing process and to align single layers and objects.
- The third step is the computation of the 3D point cloud. As mentioned above, there are a number of ways to do this. In this study, we used BundlerTools (BundlerTools 2014), a script which combines several FOSS modules like Bundler, CMVS and PMVS2 to create dense 3D point clouds in just one step on a UBUNTU/ FEDORA operating system.
- The post-processing in this study was done in MeshLab V1.3.0a (MeshLab 2011) and

CloudCompare 1.0 (Cloud Compare 2011). These programs offer various and manageable tools for cleaning the point cloud (MeshLab) and point picking (CloudCompare) of the control points.

- Finally, the georeferencing was done by a new, especially developed module of Grass GIS 7.

The Structure-from-Motion approach first produces first of all a 3D point cloud in an arbitrary coordinate system which needs to be georeferenced. The purpose of georeferencing is to convert the arbitrary coordinates into real-world coordinates, which in turn allows combining the 3D point cloud with other spatial data fur further analysis. Georeferencing is commonly performed by using Ground Control Points (GCPs) with known coordinates in the arbitrary coordinate system and the real-world coordinate system. In our case, GCPs were markers on the excavated structures which could easily be identified in the 3D point cloud and on the real structure. The arbitrary coordinates of the markers could be determined by querying digital point coordinates. The real-world coordinates of the markers on the real structure were measured using a total station. This set of GCPs could then be used to calculate 3D transformation equations.

The part of the georeferencing was done by the OS GIS software GRASS GIS 7. Unless you want to keep the point cloud in a GIS environment (Paragraph 3.2), GRASS GIS needs three modules for importing, transforming and exporting the georeferenced point cloud.

The application BT and GRASS GIS was tested on a Linux-based system (x64), CloudCompare and Meshlab on a MSWindows-based system (x64).

2.2 Method 2 – Photogrammetry: Visual SfM

Much has happened in the FOSS community working in this field since we started documenting the case studies in 2011. One and a half year can be enough time to develop the free software which combines all the steps we had to undertake in the BT method in just one software product.

Commonly, after creating overlapping pictures of the structure (we used the same images we created for the BT method), the free – but 'closed and proprietary' – software VisualSFM (VSFM) calculates the 3D point cloud and includes an option to georeference the result. The latter can be modified in the OS software MeshLab and Cloud Compare.

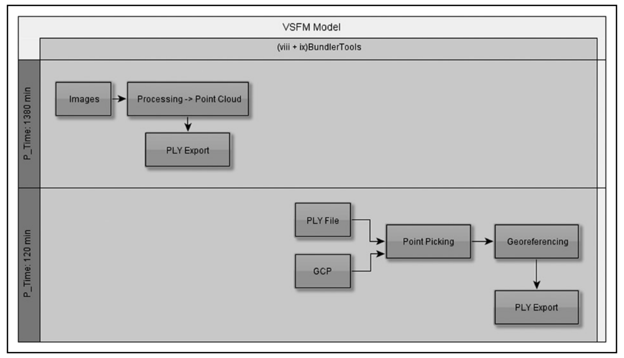


***Figure 2**. Workflow Method 2, VisualSFM, with Processing Time (P_Time) and Working Steps within the different software packages.*

Step 1 and Step 2 are the same like in the BundlerTools method.

The third step is the computation of the 3D point cloud. Although VSFM uses a slightly different algorithm for camera calibration (Wu 2007) and can even work without the EXIF-data (metadata) of the images, the computing steps are very similar to BT: Matching images -> reconstructing a sparse point cloud -> reconstructing a dense point cloud.

The georeferencing function is included into the same software. For the georeferencing, the GCP values can be attached to the images in a separate VSFM function. Their values are finally used for the transformation of the 3D point cloud into a real-world coordinate system (Fig. 2). The whole application was tested on an MSWindows-based system (x64) and runs furthermore on Linux and Mac OSX.

3. Results

3.1 Model comparison

Case study 1 – Ostia Antica

In summary, three studies of different layers in one archaeological trench were executed. As an example to present in this paper, we choose one part with a size of 50 x 100 cm of the archaeological layer

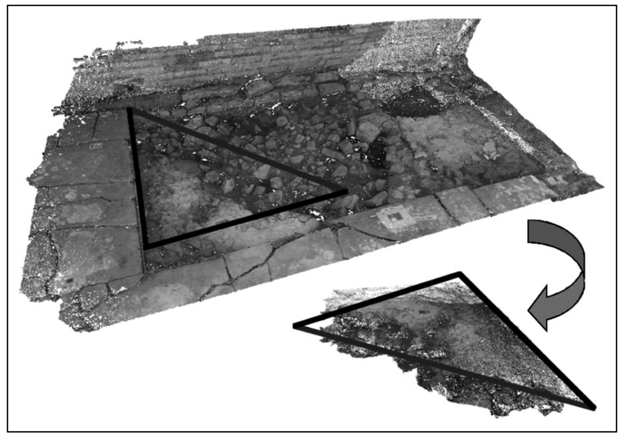


***Figure 3**. Excavation Trench 2, 3D point cloud generated by BT method, top surface of SE201 and example part, Ostia Antica/Italy.*

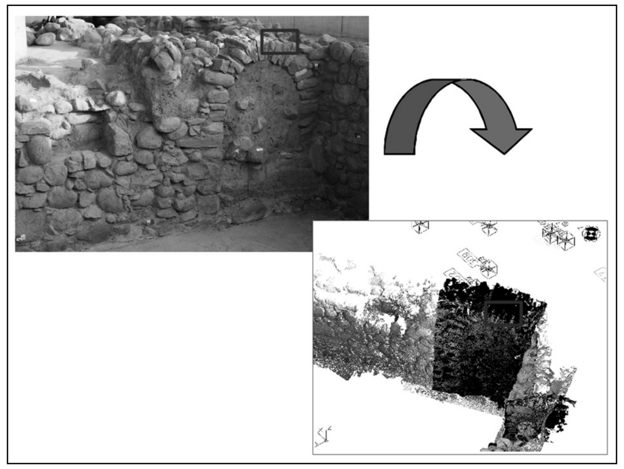


***Figure 4**. 3D point cloud generated by BT method, south wall of Middle Age cellar No. 2-1 Freyenstein/Germany and example part.*

SE201. It consisted mainly of antique floor filling with screed material and hence had a highly jointed topography (Fig. 3).

Case study 2 – Freyenstein

The example in case study 2 was a 50 x 50 cm part of the upper south wall of the medieval cellar No. 2-1 in Freyenstein/Germany. The surface we documented was made of cobble stones of different sizes with a sedimentological filling in between (Fig. 4). This resulted again in a highly jointed topography of different 3D shapes. The result is representative for four case studies we calculated in Freyenstein.

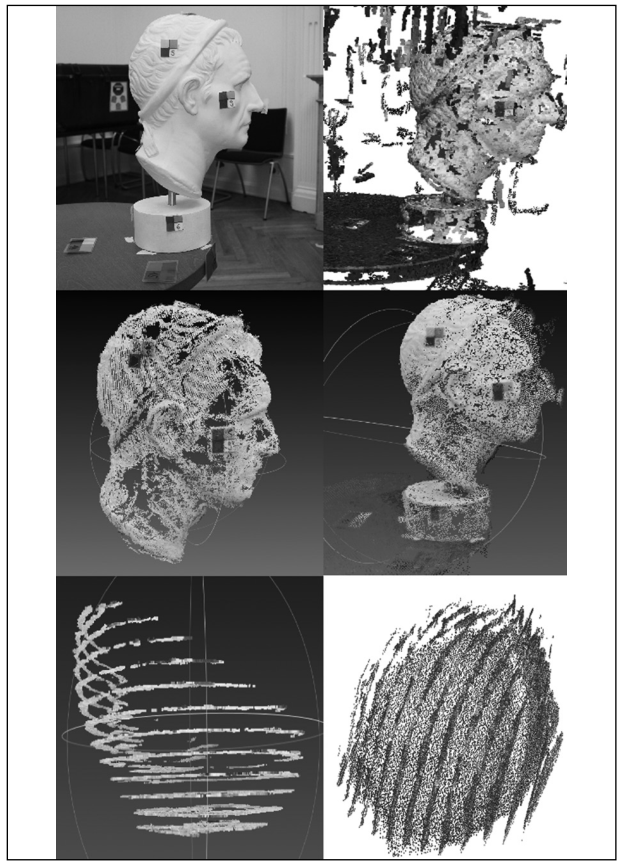


***Figure 5**. Portrait head of an Antique emperor, top right: 3D point cloud generated by VSFM, middle left: 3D point cloud generated by VSFM and cleaned in MeshLab, middle right: 3D point cloud generated by BT, bottom left: 3D point cloud generated by BT after 90° rotation, bottom right: 3D point cloud generated by BT after 90° rotation in blue in comparison with the TLS point cloud in red and the VSFM point cloud in green.*

Case study 3 – Sculpture

The copy of a full-size portrait head of an Antique emperor made of gypsum was chosen for the third case study (Fig. 5). The different material and the filigree surface were very different from case studies 1 and 2, but represent – in archaeological terms – a very common object and material.

Comparison

In order to compare the resulting 3D point clouds of the two SfM methods with the reference model of the 3D laser scanner in terms of precision, we investigated the behaviour of the 3D Point clouds

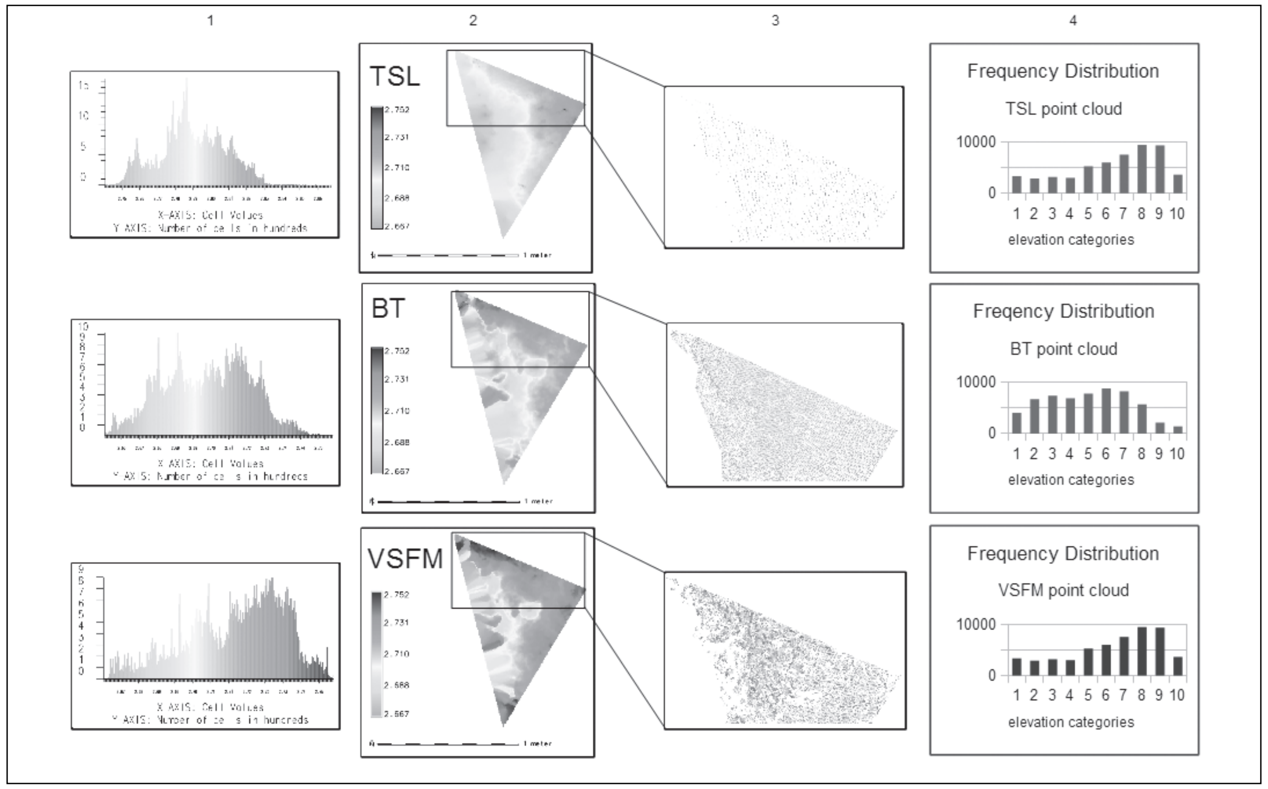
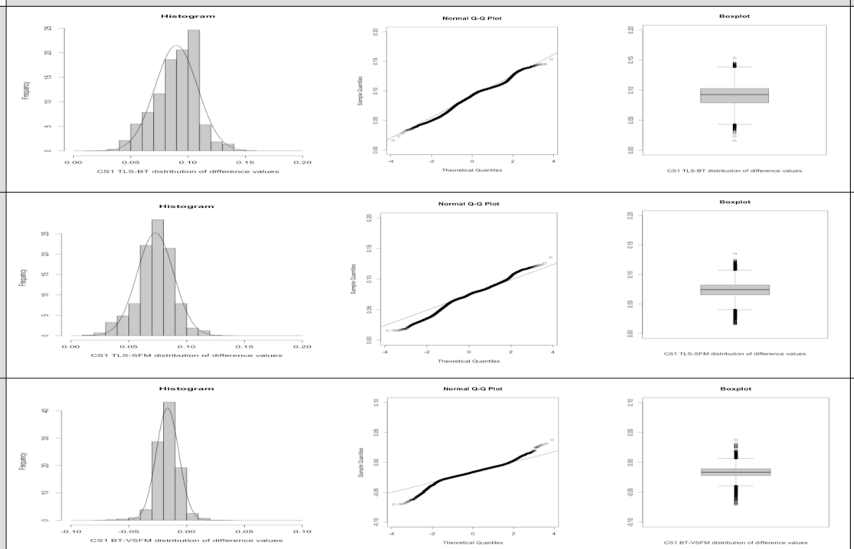


***Figure 6**. Case Study 1, sample part of SE201, column 1: Histograms of DEMs, column 2: DEMs, column 3: 3D point clouds (enlarged section of sample part), column 4: frequency distribution of 3D point clouds.*



1	3	4		5	6
	Histogram with normal distribution graph	QQ-plot	Boxplot	SW-Test (p-value) AD-Test (p-value) Sig. level 95%	Moran's I z-score p-value
TLS – BT				0.002122 0.0009808	0,960146 133,81 0.0
TLS – VSFM				0.0001718 2.444e-05	0,960308 134,12 0.0
BT - VSFM				5.739e-06 1.044e-08	0,913079 129,62 0.0

***Table 2**. (iii) Case Study 1- significance tests for difference maps; Legend: SW-Test = Shapiro-Wilk test, AD-Test = Anderson-Darling test.*

	Point Count	Point Resolution (inter-point distance)	Mean / STD
TLS	5856	0.01294	Ø 2.79 ± 0.015
BT	59162	0.00425	Ø 2.71 ± 0.018
VSFM	54709	0.004454	Ø 2.72 ± 0.012

***Table 1**. Case Study 1 - Statistics for 3D point clouds; The inter-point distance was obtained by GRASS GIS 7.0 module v.surf.bspline.*

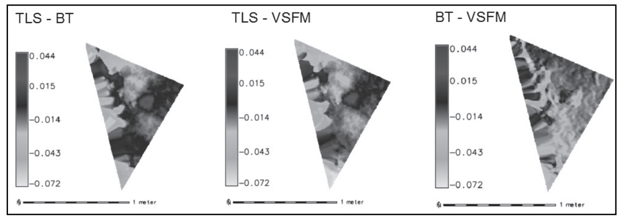


***Figure 7**. Case Study 1, difference maps displaying the deviation from each other below and above 0.*

	Overlapping area ±0.05m	Overlapping area ±0.01m
TLS - BT	3 %	0 %
TLS – VSFM	9 %	0 %
BT - VSFM	99 %	22 %

***Table 3**. Case Study 1 – difference maps and their congruent areas according to the given tolerance of ±0.05m and ±0.01m.*

	Cell Count	Cell Resolution	Mean / STD
TLS – BT	9850	0.007	0,0897 / 0,0185
TLS – VSFM	9891	0.007	0,0729 / 0,0157
BT - VSFM	10218	0.007	-0,0165 / 0,0097

***Table 4**. Case Study 1 - Statistics for difference maps; Cell Count = amount of the non-null cells, Cell Resolution = cell size, Mean = mean of absolute values, STD = standard deviation.*

(Fig. 6, column 3, 4, and Table 1). Using them as a basis, we interpolated raster surfaces (digital elevation models, DEMs (Conolly and Lake 2006, 90–111); Fig. 6, column 1, 2, via the GRASS GIS 7.0 module v.surf.bspline; Table 2.1).

As a third step, we calculated the overlapping parts between the DEMs (via GRASS GIS 7.0, module r.mapcalc) in order to measure differences between the SfM DEMs and the reference model at certain places.

The resulting raster surfaces (difference maps Fig. 7), their histograms and plots (Table 2, column 3, 4) show the difference from 0 between the reference, the laser scan DEM (TLS DEM), and the SfM DEMs as well as the difference between both SfM DEMs (BT-VSFM). This is expressed in the difference from 0 (the overlapping parts) in the negative and positive direction with a tolerance range of -0.08 and 0.05 m. If we accept a difference tolerance to the reference of ±0.01m, which might be sufficient in certain circumstances, we get the results of Table 3, column 3.

The results, even in a wider tolerance range of ±0.05m (Table 3, column 2), show clear similarities of the two SfM methods Bundlertools (BT) and VisualSfM (VSFM) with an overlap of 99% in the ±0.05m tolerance range and of still 22% with ±0.01m. The model created by the laser scanner does not share any overlapping areas in the acceptable range.

Considering this result, we looked at the quality of the DEMs by measuring the root mean square error (RMSE) of ground control points (GCPs, Paragraph 3.2) for getting an '…accurate assessment how well each cell in the DEM represents the true elevation.' (Wechsler 2007, 1483) and furthermore by studying the distribution and behaviour of the legacy data and their derivatives.

The point count in Table 1 shows a great difference in the total amount between the TLS and SfM methods of approx. 50 000. This difference can be explained by the TLS resolution adjustments which were chosen according to the topography of the archaeological surface by 0.01m x 0.01m with a devise distance of approx. 3m. The SfM methods, on the other hand, used approx. 20 photographs for the automatic generation of the 3D point clouds. As mentioned later in this paper (Paragraph. 4), the number of pictures is responsible for the point cloud resolution with a strong correlation to the inter-point distance (Table 1, column 2). A second indicator for differences is the mean value of the TLS point cloud which has a difference of max.

0.08m from the SfM mean values. This indicates, as a consistent appearance, that the mean value is located at higher elevations than for the SfM results. The low spatial resolution of the TLS point cloud (Table 1, column 2) could therefore not take into account the strong relief of the surface from case study 1. The low variance of the STD values supports this fact.

The post-processing TLS-result is a smoothed surface which explains the differences at certain places in the deviation calculations (Fig. 7 and Table 4). We hence decided that the TLS measurements could no longer act as reference in this study, but were still used as a comparison sample set.

Since a clustered pattern of elevation values of an artificial surface is evident, one cannot expect a random distribution of the measurement points. We therefore dismissed an investigation of statistical significance of the sample set.

However, the behaviour and distribution of the difference values in the difference maps was investigated by statistics to get information about similarities in the methodologies. For this, three different graphical plots (Table 2, column 3, 4) and two non-parametric one-sample tests, the Shapiro-Wilk-test (SW)

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

and the Anderson-Darling-test (AD),

$$A^2 = -n - S, \text{ with}$$

$$S = \sum_{k=1}^{n} \frac{2k-1}{n} \left[\ln F(Y_k) + \ln\left(1 - F(Y_{n+1-k})\right)\right]$$

which are powerful mathematical procedures for statistical hypothesis testing against the Null-hypothesis (H_0) for large sample sizes (Razali 2011; Shapiro and Wilk 1965) were applied.

The difference maps results in Table 2 show clear clustered results in the graphical plots with SW and AD p-values < 0.05 (Table 2, column 5). This result is supported by the high Global Moran's Indices very close to +1 for clustered patterns (Conolly and Lake 2006, 158–160; Moran 1950),

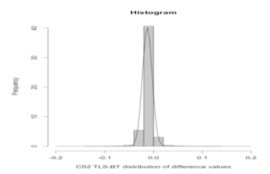
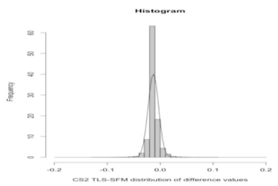
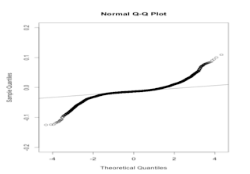
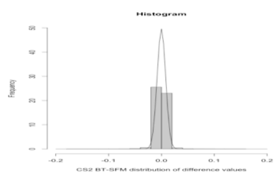
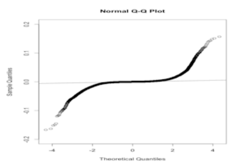
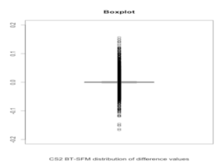
1	3	4		5	6
	Histogram with normal distribution graph	QQ-plot	Boxplot	SW-Test (p-value) AD-Test (p-value) Sig. level 95%	Moran's I: z-score p-value
TLS – BT				1.593e-08 3.154e-15	0,905362 324,43 0.0
TLS – VSFM				1.054e-08 1.028e-14	0,937849 336,07 0.0
BT - VSFM				1.682e-10 < 2.2e-16	0,870506 312,95 0.0

Table 5. *(iii) Case Study 2 - significance tests for difference maps; Legend: SW-Test = Shapiro-Wilk test, AD-Test = Anderson-Darling test.*

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2}$$

an indicator for spatial auto-correlation between the value and the topography of the modelled surface. The high z-scores in combination with very low p-values (Table 2, column 6) indicate high significance of the sample sets. Finally, the result of Table 3 supports the aforesaid assumption that the greatest differences occur between the laser scanner and SfM results.

Case study 2

The same procedures described above have been executed for case study 2 (Table 5) for obtaining statements about surface similarity of different measurement methods. Comparing the results of case study 1 and case study 2 the results are very similar from a proportional perspective (Table 3 and 6).

Table 6 shows nearly no difference between the two SfM DEMs in the tolerance range of ±0.01m and still ¼ conformity between the TLS and SfM maps. This is proportionally similar to the case study 1 outcome. However, in contrast to the wider tolerance range (±0.05m) there is nearly complete conformity between all maps which was not the case in case study 1.

Considering the behaviour and distribution of the point clouds and their corresponding DEMs, except for the point count and its corresponding spatial resolution, there is no great difference between all three point clouds at all. This initial similarity runs through all subsequent calculations (Table 7 and 8) and explains the results in Table 6. We suppose that the reason for this might lie in the smooth topography of the surface in case study 2. Such topographies have a greater chance for more realistic modelling for laser scan measurements with the same device adjustments as in case study 1. The amount of photographs for the SfM methods was as well the same.

Finally, the test of normality and auto-correlation in case study 2 (Table 5) revealed, like in case study 1, a significant clustered distribution of

	Overlapping area ±0.05	Overlapping area ±0.01
TLS - BT	99 %	25 %
TLS – VSFM	99 %	24 %
BT - VSFM	99 %	93 %

Table 6. *Case Study 2 – difference maps and their congruent areas according to the given tolerance.*

	Point Count	Point Resolution (inter-point distance)	Mean / STD
TLS	2289	0.004229	Ø 93.66 ± 0.0072
BT	19424	0.001431	Ø 93.67 ± 0.0069
VSFM	18391	0.001456	Ø 93.67 ± 0.0085

Table 7. *Case Study 2 - Statistics for 3D point clouds; The inter-point distance was obtained by GRASS GIS 7.0 module v.surf.bspline.*

	Cell Count	Cell Resolution	Mean / STD
TLS – BT	64460	0.0007	-0,0128 / 0,0099
TLS – VSFM	64460	0.0007	-0,0129 / 0,0099
BT - VSFM	64826	0.0007	-0,0007 / 0,0081

Table 8. *Case Study 2 - Statistics for difference maps; Cell Count = amount of the non-null cells, Cell Resolution = cell size, Mean = mean of absolute values, STD = standard deviation.*

the difference maps. This result suits to the mean and STD with very low STD values and the result for the BT-VSFM difference map with ±0.01m tolerance. Both difference maps in case study 1 and case study 2 of BT minus VSFM show a slight bias towards a normal distribution whereas the difference maps involving the TLS results are in both case studies very similar.

3.2 Georeferencing

The comparative calculation described above was only possible because we were able to georeference all models. Since 3D point clouds in archaeological contexts have to be located in a

GCP No.	TLS		BT		VSFM	
	Horz	Vert	Horz	Vert	Horz	Vert
1	0.010	0.015	0.0008	0.0010	0.0005	0.0005
2	0.011	0.012	0.0040	0.0045	0.0087	0.0083
3	0.002	0.001	0.0075	0.0083	0.0063	0.0059
4	0.001	0.001	0.0114	0.0126	0.0091	0.0009
5	0.002	0.001	0.0071	0.0077	0.0150	0.0132
6	0.001	0.001	0.0071	0.0079	0.0098	0.0088

Table 9*. Selection of RMS values occurring during registration (TLS) and georeferencing (BT and VSFM). The TSL RMSE measurement provided only three digits after comma.*

georeferenced environment – otherwise they do not fit to other documentation – we were interested to improve the BT software result in this way. This part in the working steps and function within the different software is hence very important and is therefore included in most proprietary and non-proprietary SfM software packages, like VSFM. However, BT still lacked this function at the time of the study. Therefore, for the Open Source GRASS GIS software package (GRASS Development Team 2012; Neteler et al. 2012) a new software module 'v.ply.rectify' has been developed, sponsored by Topoi, the Berlin Excellence Cluster at Humboldt-University of Berlin and Free University Berlin about 'The Formation and Transformation of Space and Knowledge in Ancient Civilisations' (Topoi 2014). The new GRASS GIS module performs 3D georectification based on GCPs and offers different transformation methods. The 3D transformation method most appropriate for 3D point clouds produced with the Structure-from-Motion approach is a Helmert transformation that preserves angles and proportions. This kind of transformation is also known as orthogonal transformation. Commonly used algorithms for Helmert transformation are found in the reprojection of spatial 3D data from one real-world coordinate system into another real-world coordinate system where the rotation angles are always very small, in the range of a few arc seconds. These algorithms are not suitable for georeferencing a 3D point cloud from an arbitrary coordinate system to a real-world coordinate system because the rotation angles can be as large as 180 degrees. Therefore a generalised version of the Helmert transformation was implemented, supporting large rotation angles. In addition, the polynomial 3D transformation equations of have been estimated. A first-order polynomial 3D transformation is also known as affine transformation with shift, scale, and rotation separately for each dimension. The new module tests whether the GCPs fulfil the requirements needed to perform the requested 3D transformation and provides diagnostics in terms of Root Mean Square Errors (RMSE) for the GCPs which help to identify outliers in the GCPs and quantify the accuracy of the coordinate transformation.

An extraction of the results generated in the case studies by the new GRASS GIS modules in comparison to the transformation results by the proprietary TLS software and VSFM are listed in the table below. Additional modules were developed for GRASS GIS to import and export 3D point clouds in the PLY format commonly used in archaeological SfM projects. The complete workflow in GRASS GIS thus consists of:

1. importing a 3D point cloud,
2. identifying and removing potential outliers in the GCPs based on the RMSE of the selected transformation method,
3. georeferencing the 3D point cloud with the selected GCPs and method, and

Method	Costs	Open Source & Open Access	Equipment Size & Weight	Steps	Time	Curve
BT	camera, notebook, external drive € 800 - 2500	OS & OA	hand luggage size, ca. 1.5 kg	5	620 min	low
VSFM	camera, notebook, external drive € 800 - 2500	Proprietary & OA, Free download	hand luggage size, ca. 1.5 kg	4	1500 min	low

Table 10*. Parameter comparison of SfM methods.*

4. optionally exporting the georeferenced 3D point cloud for further analysis in other software.

The precision of the results have been evaluated by both the deviation from measured real-world coordinates of the markers and visual inspection of the georeferenced objects (preserving angles and proportions). An additional module, to perform these parts in one step (v.ply.rectify), only runs in the Linux operating system at the time of this publication. The case studies were executed at FEDORA 18, a Unix-based operating system and GRASS GIS 7. The separate modules v.in.ply, v.out.ply, v.rectify are also available in the current GRASS GIS Windows version 7.

3.3 Processing time, effort and resources

According to the methodological research question, we compared not only the models but also the parameters (Table 10).

Model Precision & Quality

The precision of a measurement system (also called reproducibility or repeatability) is the degree to which repeated measurements under unchanged conditions show the same results (Taylor 1999, 128–129). In the fields of science, engineering, industry, and statistics, the accuracy (JCGM 200 2012, 21) of a measurement system is the degree of closeness of measurements of a quantity to that quantity's actual (true) value.

In this study, we dismissed the working hypothesis in order to refer to the 3D point cloud generated by the terrestrial laser scanner for accuracy. However, by comparing the transformation results of the RMS errors, we could still use TLS as a reference since the GCP's acquisition and their use for the transformation is decoupled from the generation of the 3D point cloud. Hence, in both case studies we got very accurate results.

Costs

Even if some archaeological projects are very well funded, costs are an important factor in all projects. Saving costs for technical equipment releases resources for other important needs (e.g. professional staff, laboratory examinations etc.). Costs in the case studies were very low or do not even occur since the SfM equipment which is needed belongs on every archaeological project as well as a total station for measuring the GCPs. However, since not every archaeologist is able to get low cost educational licences, we think it is important having an Open source alternative at hand.

Open Source and Open Access

Processing software in science can be seen as a breadboard assembly where legacy data are processed via algorithms in order to transform them to a new result. Since scientists need to understand how these processes transform their raw data, knowledge of the internal procedures (the breadboard assembly) is vital.

Open Source software which follows the Open Source Definition (Open Source Initiative 2014) and is developed under the GPL license (Free Software Foundation 2014b) and other related licenses fulfils these requirements. Since the BT application was tested on a Unix-based system, this application fulfils these requirements. The VSFM method was tested in an MSWindows-based system but its documentation also mentions a Unix version for Ubuntu. Open Access guarantees the free access to scientific literature which also means software documentation through the internet what we realised for both SfM methods.

Equipment Size and Weight

Transport and handling of equipment is a further important factor for archaeological excavations, which are often off-road. Hence, the size and weight of the additional equipment was also of great interest and therefore one study parameter. Finally, as was the case with the 'Cost' parameter, both SfM methods required no additional equipment.

Working Steps

The detailed working steps of the two methods are depicted in the two diagrams (Figs 1 and 2). The BT method required just one step more than VSFM due to the software change for applying the georeferencing.

Processing Time

We defined "processing time" in this study in terms of the time required from starting the processing of the images until the final georeferenced 3D point cloud since the documentation time – taking pictures – was the same for both SfM methods. In the example in Figure 2 we display an average time for both methods. The VSFM method in the case studies needed about 10 times more than the BT method using comparable hardware.

Learning Curve

This parameter investigated the effort to learn the handling of the hardware and working with the associated software (control software and/or processing software). Since both methods are very well documented (see the Project Website), even untrained users could follow their descriptions without further reading or workshop attendance.

4. Discussion and Conclusions

4.1. Analysis

The aim of this paper was to examine the results and the procedure of low budget equipment and Open Source or free available software for high resolution documentation of archaeological stratigraphy and features as alternative to high cost equipment. The three case studies provide just small glimpses of archaeological structures which were chosen in order to generate comparable sample sets. Although the study cannot give general advice, we hope that the analysis of the results of three very different archaeological examples can help support decisions regarding future applications in this field.

Case Study 1 and 2

During the first and second case study, it turned out that the reference we chose was not appropriate for the purposes of this study. Undoubtedly, the TLS method produces accurate and precise results. However, in our working hypothesis we assumed that both SfM methods cannot compete with its results. With the background knowledge of being also the excavators of the sample structures we finally had to conclude that both SfM methods created much more realistic models than the laser scanner. The reasons for this result have been already described in detail above. One could hence argue that the laser scanner was not used to its full extent. This is true at one hand. On the other hand, the comparison tried to act within a realistic environment where a higher resolution adjustment for the laser scanner would have been too meticulously for the structures we measured. The precision of the SfM models could be supported by the RMSE results described in Paragraph 3.2.

Case Study 3

As a contrast to the sample sets we used in case study 1 and 2, we applied the three methods to the challenging material gypsum. The TSL method produced reasonable results (Fig. 5, bottom right red point cloud). However, in order to generate a 3D point cloud, VSFM and BT referred more to the environment (Fig. 5, top right) than to the object itself. In Fig. 5 (left middle row) the cleaned point cloud has still many gaps on the object's surface. However, by changing the point of view one can recognise a strip-like order to the points (Fig. 5, left and bottom right) which do not represent the object's surface. In summary, the material seemed very problematic for the two SfM methods (Fig. 5, top left and bottom right).

Since SfM is based on feature extraction from the overlapping images, the colour, shape, material and roughness of an archaeological surface is vital for getting excellent results (Ducke, Score and Reeves 2011, 377–380). Objects without texture and a reflective surface are, according to our experiments, not suitable for the SfM method.

4.2. Advantages – Disadvantages

Since the keywords Open Access and Open Source policies are nearly already established in many scientific fields, archaeologists and cultural scientists are becoming more and more aware of their potential and sustainability. Hence, this paper gives an inside in the usefulness of alternative documentation methods at sufficient precision and accuracy.

However, answering the question of which method is more suitable for certain purposes strongly depends on the scientific objective, research aims, circumstances, environment, time, budget and available hardware. The ideal method can only be determined on-site and is defined by the site's specific conditions since the find spectrum at archaeological excavations cannot be predicted precisely. Hence, the challenge of archaeological documentation is to provide a flexible system. In consideration of effective data acquisition and storage possibilities, the methods should follow an objective data documentation system and take long-term file storage into account. With the BT method such a system is now at hand. It is a three-dimensional documentation method which is easy to handle, universal in scale and of low cost. Its advantages are:

- reliable results,
- fast post-processing,
- the 3D point cloud is already in a GIS system, if needed.
- The downsides of this application are
- manual GCP picking,
- many working steps/software packages,
- the BT software itself runs only on a Unix-based system.

The negative results in the third case study show the limits of the SfM method, which had nothing to do with the different processing algorithms we applied with both methods. The reason lies within the methodology of the generation of the 3D point cloud itself.

Returning to the 'Archaeological Research Question' we had in mind, neither SfM application is restricted to detailed architecture or highly funded projects. It is, on the contrary, open to a wide user community in archaeology and related disciplines.

Acknowledgements

The Topoi Summer School 2011 acted as part of the 'KBO Excavations' (Kent-Berlin Ostia Excavations 2011) which investigated inter alia the Main Forum of Ostia Antica under the direction of PD Dr. Axel Gering (Humboldt-University of Berlin) under the auspices of the Soprintendenza Archeologica di Ostia directed by Prof. A. Pellegrino).

References

Agisoft. 2014. "Image based 3D modelling." Accessed March 30. http://www.agisoft.ru/.

ARC3D. 2013. "Automatic Reconstruction Cloud. Convert your images into 3D." Accessed March 30. http://www.arc3d.be/.

Arc-Team. 2014. "Download/Python Photogrammetry Toolbox." Accessed March 30. http://184.106.205.13/arcteam/ppt.php.

ArcTron 3D. 2012. "Vereinen Sie Ihre Daten mit: aSPECT 3D. Innovative 3D-Information Management." Accessed March 30. http://www.arctron.de/fileadmin/Dateiliste/01_Home/aSPECT3D_Neu_2012.pdf.

Arc-Team. 2014. "ATOR."Accessed March 30. http://arc-team-open-research.blogspot.de/.

Autodesk ReCap. 2014. "Autodesk ReCap." Accessed March 30. http://www.autodesk.com/products/recap/overview.

BundlerTools. 2014. "BundlerTools." Accessed 30 March, https://github.com/qwesda/BundlerTools.

Cloud Compare. 2011. "3D point cloud and mesh processing software Open Source Project." Accessed March 30, 2014. http://www.danielgm.net/cc/.

Conolly, J., and M. Lake. 2006. *Geographical Information Systems in Archaeology*. Cambridge: Cambridge University Press

Debrevec, P. E., C. J. Taylor, and J. Malik. 1996. "Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach." In *SIGGRAPH '96: Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, 11–20. New York: ACM.

Doneus, M., and W. Neubauer. 2004. "Digital Recording of Stratigraphic Excavations." In *Enter the Past. The E-way into the four Dimensions of Cultural Heritage. CAA 2003, Computer Applications and Quantitative Methods in Archaeology*, edited by K. F Ausserer, W.

Börner, M. Goriany, and L. Karlhuber-Vöckl, 113–116. Oxford: Archaeopress.

Doneus, M., and W. Neubauer. 2005. "Laser scanners for 3D documentation of stratigraphic excavations." In *International Workshop on Recording, Modeling and Visualization of Cultural Heritage, Ascona, May 22–27, 2005*, edited by M. Baltsavias, A. Gruen, L. van Gool, and M. Pateraki. 193– 203. London: Taylor and Francis.

Doneus, M., W. Neubauer, and N. Studnicka. 2003. "Digital Recording of Stratigraphic Excavations." In *Proceedings of the XIXth International Symposium, CIPA 2003: New Perspectives to Save Cultural Heritage. The CIPA Int. Archives for Documentation of Cultural Heritage, Vol. XIX, 2003*, edited by M. O. Altan, 451–456. Istanbul: CIPA

Ducke, B., D. Score, and J. Reeves. 2011. "Multiview 3D reconstruction of the archaeological site at Weymouth from image series." In *Computers & Graphics 35 Issue 2, Virtual Reality in Brazil Visual Computing in Biology and Medicine Semantic 3D media and content Cultural Heritage*, edited by L. P. Soares, J. Kelner, B. Preim, C. P. Botha, B. Falcidieno, I. Herman, A. Chalmers, M. Mudge, and L. P. Santos, 375–382. Philadelphia: Elsevier.

El-Hakim, S., F. Remondino, F. Voltolini, and F. Gonzo. 2008. "Effective high resolution 3D geometric reconstruction of heritage and archaeological sites from images." In *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA), Berlin, 2–6 April 2007,* edited by A. Posluschny, K. Lambers, and I. Herzog, 43–50. Bonn: Dr. Rudof Habelt GmbH.

Free Software Foundation. 2014a. "GNU General Public License." Accessed June 24, 2012. http://www.gnu.org/licenses/gpl-3.0.html.

Free Software Foundation. 2014b. "What is GNU?" Accessed March 15. http://www.gnu.org/.

Furukawa, Y., B. Curless, S. M. Seitz, and R. Szeliski. 2010. "Towards Internet-scale Multi-view Stereo." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 13 – 18 June, San Francisco*, 1434-1441. Los Alamitos: IEEE Computer Society

Furukawa, Y., B. Curless, S.M. Seitz, and R. Szeliski. 2010b. "Towards Internet-scale Multi-view Stereo." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1434–1441. Los Alamitos: IEEE Computer Society.

GRASS Development Team. 2012. "Geographic Resources Analysis Support System (GRASS) Software, Version 7.0." Accessed March 30, 2014. http://grass.osgeo.org.

Ho, N. 2014. "RunSFM." Accessed March 30, http://nghiaho.com/?page_id=253.

JCGM 200. 2012. "International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)." Accessed on March 30, 2014. http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.

Kent-Berlin Ostia Excavations. 2014. Accessed March 30. "Ostia. Harbour city of ancient Rome." http://lateantiqueostia.wordpress.com/about/.

MeshLab. 2011. " MeshLab." Accessed March 30, 2014. http://meshlab.sourceforge.net/.

MICMAC. 2014. "APERO-MICMAC." Accessed March 30. http://www.tapenade.gamsau.archi.fr/TAPEnADe/Tools.html.

Moran, P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37(1-2):17–23.

Neteler, M., M. H. Bowman, M. Landa, and M. Metz. 2012. "GRASS GIS: a multi-purpose Open Source GIS." *Environmental Modelling & Software* 31:124–130.

Open Source Initiative. 2014. "The Open Source Initiative." Accessed March 15. http://opensource.org/.

Pollyfeys, M., R. Koch, M. Vergauwen, and L. Van Gool. 2000. "Automated reconstruction of 3D scenes from sequences of images." *ISPRS Journal of Photogrammetry and Remote Sensing* 55(4):251–267.

Razali, N., and Y. B. Wah. 2011. "Power comparison of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling test." *Journal of Statistical Modeling and Analytics* 2(1):21–33.

Rocchini, D., and M. Neteler. 2012. "Let the four freedoms paradigm apply to ecology." *Trends in Ecology & Evolution* 27 (6):310–311.

Shapiro, S. S., and M. B. Wilk. 1965. "An Analysis of Variance Test for Normality (Complete Samples)." *Biometrika* 52(3/4):591–611.

Snavely, N., S. M. Seitz, and R. Szeliski. 2006. "Photo tourism: exploring photo collections in 3D." In ACM *Transactions on Graphics (TOG) – Proceedings of ACM SIGGRAPH 2006* 25(3):835–846.

Snavely, N., S. M. Seitz, and R. Szeliski. 2007. "Modeling the World from Internet Photo Collections." *International Journal of Computer Vision* 80 (2):189–210.

Sachsen.de. 2012. '#d Laserscanning." Accessed March 30. http://www.sachsen.de/en/1842.htm.

Strecha, C., W. von Hansen, L. van Gool, P. Fua, and U. Thoennessen. 2008. "On benchmarking camera calibration and multi-view stereo for high resolution imagery." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 23-28 June, Great Boston*, 1-8. Los Alamitos: IEEE Computer Society.

Taylor, J. R. 1999. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Sausalito: University Science Books.

Topoi. 2014. " Welcome to the excellence cluster TOPOI!" Accessed March 15. http://www.topoi.org/.

Visual Experiment. 2014. "SFMToolkit." Accessed March 30. http://www.visual-experiments.com/demos/sfmtoolkit/.

Wechsler, S. P. 2007. "Uncertainties associated with digital elevation models for hydrologic applications: a review." *Hydrology and Earth System Sciences* 11:1481–1500.

Wu, C. 2007. "SiftGPU: A GPU implementation of Scale Invaraint Feature Transform (SIFT)." Accessed March 30. http://cs.unc.edu/~ccwu/siftgpu.

Wu, C., S. Agarwal, B. Curless, and S. M. Seitz. 2011. "Multicore Bundle Adjustment." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 20-25 June, Orlando,* 3057-3064. Los Alamitos: IEEE Computer Society.

Wu, C. 2014. "VisualSFM: A Visual Structure from Motion System." Accessed March 30. http://ccwu.me/vsfm/.