G Model
CULHER-2586; No. of Pages 10
ARTICLE IN PRESS

Journal of Cultural Heritage xxx (2012) xxx–xxx

Available online at
SciVerse ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Original article

Archeological excavation monitoring using dense stereo matching techniques

Matteo Dellepiane[a,*], Nicolò Dell'Unto[b], Marco Callieri[a], Stefan Lindgren[c], Roberto Scopigno[a]

[a] Visual Computing Lab, ISTI-CNR, Via G. Moruzzi 1, 56124 Pisa, Italy
[b] Department of Archaeology and Ancient History, Lund University, Sandgatan 1 Box 117, 221 00 Lund, Sweden
[c] Humanities Laboratory, Lund University, Helgonabacken 12 Box 201, 221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 26 October 2011
Accepted 23 January 2012
Available online xxx

Keywords:
Archeology
Archeology excavation
Documenting
Monitoring
Dense stereo reconstruction
Digital 3D models
Visualization
Assessment

ABSTRACT

Several new tools to obtain three-dimensional information from unorganized image sets are now available for the public use. The main advantage of this software, which is based on dense stereo matching, is the possibility to generate 3D content without the need of high-cost hardware (e.g. 3D scanning devices). Nevertheless, their use in real-world application domains (like cultural heritage) is still not very diffused, due to the non-straightforward usability of the raw data produced. In this paper, we investigate the use of automatic dense stereo reconstruction tools for the monitoring of an excavation site. A methodology for the effective acquisition and processing of data is presented. In addition, the results of the data assessment demonstrate the repeatability of the data acquisition process, which is a key factor when qualitative analysis is performed. The use of three-dimensional data is integrated in an open source mesh processing tool, thus showing that a spatio-temporal analysis can be performed in a very intuitive way using off-the-shelf or free/open digital tools. Moreover, the use of peculiar rendering and the creation of snapshots from arbitrary points of view increase the amount of documentation data, and suggest a perfect integration of data produced with dense stereo matching in the future standard documentation for excavation monitoring.

1. Research aims

The work presented in this paper investigates the use of dense stereo reconstruction tools for the diachronic documentation of the cultural layers during an ongoing archeological excavation.

Dense stereo reconstruction tools are able to produce accurate 3D models of an object starting from a set of uncalibrated images. The paper explores the possibility to have a three-dimensional documentation of the evolution of an excavation in a very simple manner, using low-cost acquisition hardware (a digital camera) and freeware software for data processing.

This could help in overcoming the main issues related to the applicability of 3D acquisition (hardware cost, amount of data, post-processing effort), and integrate 3D data to the classic documentation.

The availability of a "spatio-temporal" representation of an excavation evolution boosts the potentials for the analysis and interpretation work by archeologists.

The main goals of this work are:

- the definition of a workflow for the effective and easy acquisition and management of the 3D data acquired during the field work;
- a reliability test on the models generated with this technique;
- the proposal of new tools or features to support improved visualization and analysis of the acquired 3D archeological data.

2. Introduction

Digital technologies have produced an impact on our society that can be considered unique and unusual; archeology – as many other subjects of the Cultural Heritage domain – was strongly influenced by this revolution. In the last decade, the introduction of automatic instruments for data collection and for management of archaeological features connected with the World Wide Web has strongly influenced the way of how archaeologists today approach and perceive the material culture.

Ezra B. W. Zubrow provides an interesting overview of how digital technology is radically changing this discipline [1]: [...] "archeologists are becoming part of the widespread digital village [...] and [...] a new village is defined by archaeologists, including other specialists such as geologists, palynologists, geographers, economists, etc." [...].

The techniques for 3D digitization have been used to document individual archeological findings, but also the state of a whole excavation environment. Archaeological excavations are very dynamic

* Corresponding author.
E-mail addresses: surname@isti.cnr.it (M. Dellepiane), nicolo.dell_unto@ark.lu.se (N. Dell'Unto), stefan.lindgren@humlab.lu.se (S. Lindgren).

1296-2074/$ – see front matter © 2012 Published by Elsevier Masson SAS.
doi:10.1016/j.culher.2012.01.011

environments, and a documentation of their evolving state is needed almost every time a new historical feature is detected. This would imply doing a complete 3D acquisition of the working site every day, or even every few hours.

Despite the usefulness of such time-dense documentation, it is not always possible to employ 3D scanning devices on such tight schedule, because of time and resources constrains.

Photogrammetry (the process where, using geometrical and optical principles is possible to extract measures from a series of photos) has been used in archeology since a very long time. However, a survey done using the standard photogrammetry tools requires as much time and competence as a 3D scanning survey.

A possible alternative to the adoption of 3D scanning or photogrammetry may be the family of tools able to generate 3D geometry from photographic data, based on dense stereo matching.

The generation of 3D data from a video sequence, or from an unstructured photographic dataset, has sprouted, in the course of the years, a lot of research activity in the computer graphics (CG) and computer vision (CV) communities. This resulted in the creation of several tools, which started as research prototypes used in laboratories, and are now available to the general public. These tools seem particularly suited to the documentation of archeology, due to their low hardware requirement, wide range of application and easiness of use.

The dense matching tools apply most of the principles of photogrammetry, but most of the assumptions are relaxed to move towards a “completely automatic” approach, requiring less care when taking the photos, and not requiring user input or markers in the initial phase of photo calibration and orientation.

However, this choice impacts on the quality of the output data on their reliability and on the possibility to directly use their output in a metric way. These limitations are what have kept this technology from being effectively used to its full potential, relegating it to a role of a “toy”, usable only to generate approximate models used only for visual presentation.

It is our belief, however, that with a tolerable effort in the phase of data gathering and processing, it is possible to bring back these tools closer to a surveying method (able to provide more reliable and usable data), like photogrammetry or 3D scanning, in a way such that they remain cost-effective.

Unfortunately, the current available systems are still quite rough, and the generation of a 3D model is still not a completely automatic task. For these reasons, carefulness is needed while deploying these technologies in the archeological documentation, to avoid falling in the “toy” paradigm: produce something nice to see, but with limited usefulness.

2.1. From the single test case to the general use

The aim of this paper is not to present a stand-alone case study, where the technology is adapted to a test case, but to propose a methodology that could be generic enough to be applied to a wide range of cases.

In order to generalize a proposed methodology, there are some extremely important aspects to consider:

- technology assessment: for the general application, repeatability of the results is as important as the “precision” or “accuracy”. Are these methods suitable when the acquisition of the same object has to be repeated several times? Is it possible to use the produced 3D data for the purpose of qualitative and quantitative comparison between two stages of an excavation? What are the costs of the application of such technologies (in terms of money, time, resources and skill needed)?
- definition of a workflow: the theory of measurement and surveying relies, more than onto the technology used to take the measures, on the methodology by which it is applied. Our aim is to show how these technologies may be effectively employed using the same kind of rules and with the same rigorous approaches used in the standard documentation production process;
- presentation tools: the data acquired with the proposed methodology must add information and possibly integrate with what is produced using the “state-of-the-art” documentation workflow used in archeological excavations. Moreover, the tools to visualize and present the results should be easy to use and access (i.e. preference should be given to free or open software).

It is important to underline how this work does not try to propose something that should replace the current procedures, but to integrate this new kind of documentation inside the standard documentation workflow.

3. Material

The archeological site of Uppåkra (Figs. 1 and 2), Sweden, was chosen as the test case for the presented work. The acquisition was performed in 2 weeks of intensive excavation during the period June to July 2010.

Uppåkra was an urban settlement, in which intensive and complex activities were conducted during a period spanning more than 1000 years, from about the first century BC to about 1000 AD [19].

This archeological excavation, located five kilometers south of Lund and covering an area of nearly 100 acres, is considered one of the most important archeological contexts of south Sweden. The

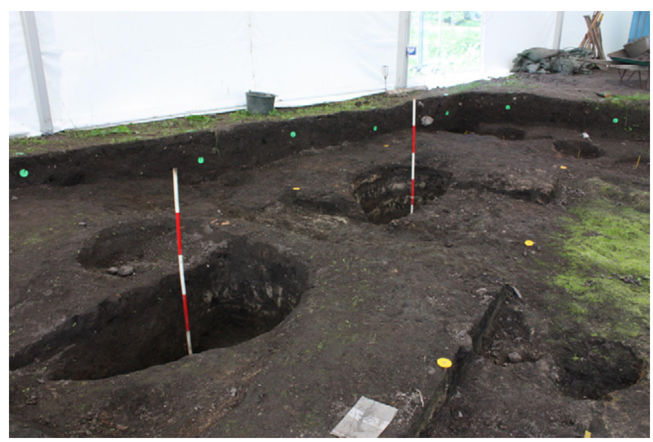

Fig. 1. A photo of the excavation used as a test case. Green and yellow markers, used for scaling the object to real measures, are well visible.

Fig. 2. A portion of the Uppakra excavation.

Please cite this article in press as: M. Dellepiane, et al., Archeological excavation monitoring using dense stereo matching techniques, Journal of Cultural Heritage (2012), doi:10.1016/j.culher.2012.01.011

long life of this site and its complex stratigraphy makes this area the perfect environment to introduce the above mentioned new documentation techniques.

In specific for this experiment, we decided to follow the investigation of a long house (approximately 500 AC) located in one of the main areas of the site. This building is still under investigation and presents a very complicated stratigraphy. Its interpretation is not an easy task, due to the size of the construction and the soil material. Therefore, before the beginning of the campaign we thought that probably the generation of 3D models could have been useful to better connect all the relations between the different cultural layers.

Today all the data that compose the excavation are recorded through the use of the total station, and afterward uploaded into a Geographic Information System (GIS) that allows reconstructing the relations between context and artifacts. This documentation method is very diffuse in Scandinavia because it allows having a high control on the information, keeping a good balance between acquisition time, data density and data quality. A disadvantage of this technique is represented by its limitation in connecting vector and raster data.

The application of a new documentation method, in support of the technique already used by the excavation team in Uppåkra, allowed to record important features otherwise impossible to describe with more traditional approaches. A fundamental role was played by the portability of the instrument (a digital camera) and the quick data acquisition to be elaborated in order to generate the 3D models.

In Uppåkra, we acquired seven sets of images for a total of nearly 400 pictures. Each acquisition process was very fast; in fact, no more than 20 minutes were required to collect each group of images. The use of the acquired data will be explained in detail in Section 5.

4. Methodology

4.1. Related work

4.1.1. Excavation interpretation and monitoring

The archeological excavation is a unique research context and its investigation is mainly based on analysis, connection and interpretation of data collected during the excavation process.

The archeological excavation is an unrepeatable experiment where all the information, once identified, are recorded and removed from their original context [2]. The main goal of this process is the interpretation and the reconstruction of the diachronic evolution of the site during the different ages.

The introduction of new visualization paradigms and tools for the spatial documentation represented an important opportunity for archeologists to record the data in a more complete and efficient way. In particular, the possibility to describe the spatial relations between all the elements that compose the cultural layers makes the visualization in real time possible with a high number of details. This operation provides not only a synthetic interpretation, but also a complete and measurable three-dimensional image of the real status of the excavation.

The process of recording features in an archeological context is usually managed by combining two different approaches: the first one – called direct documentation – concerns all the operations that imply a direct contact with the excavation; the second one – called indirect documentation – concerns all the operations that imply the use of optical, mechanical or informatics instruments for the description of the data [3].

During the normal evolution of the archeological campaign, both these methods are employed; however, with the diffusion and development of new technologies, the use of indirect methods has had an exponential growth inside the archeological excavation.

It is our opinion that the use of these methods to document specific features of the archeological site will not decrease the quality of the site interpretation, but instead it will make possible to keep track of the research activity developed into the site through a diachronic analysis of the entire work.

4.1.2. Dense stereo reconstruction tools

Dense stereo reconstruction tools produce three-dimensional data starting from a series of unordered images. The main advantage of having uncalibrated cameras, or other types of image acquisition devices, is the easiness of the acquisition process.

The process of 3D reconstruction is usually composed by different steps. Firstly, the input images are analyzed in order to find local features descriptors (for example, using SIFT). The descriptors are then matched from one image to another. Starting from the matched features, the perspective cameras associated to the images are calibrated and oriented. Once all the images have associated camera parameters, the system tries to match every pixel of every image (in this sense, the matching becomes dense), producing for each single matched pixel its corresponding coordinates in the 3D space. Several approaches to fulfill these steps have been proposed [4–7], and most of current research is focused on how to index and organize the images in order to reduce the image matching operations and possibly clustering similar images [8,9]. An overview of all the approaches in literature is well beyond the scope of this paper. In general, the amount of time needed for the processing of the image set is in the order of a few hours. The few methods which have been made available to the public are based on web services; an example of those is Arc3D [6], where the user uploads the data to a server connected to a cluster of computers, which process the data and send the results back. Recently, Bundler [10] was released as open-source software for structure-from-motion reconstruction of 3D scenes from unordered photographic dataset. One notable implementation of the Bundler tool is inside the PhotoTourism [11] application, which works quite similarly to Arc3D, since it is implemented as a web service. Bundler, coupled with a dense stereo matching tool (PMVS2 [12]), is able to generate a single point cloud from all the matched points in the input image dataset.

Cultural heritage may appear one of the most interesting fields of application for this kind of 3D reconstruction approach, being it a non-contact, low-cost alternative to current reference methods for geometry acquisition. However, while a massive amount of work has been done to compare 3D scanning and photogrammetry, comprehensive studies on the accuracy and repeatability of dense reconstruction tools are still missing. Some experimentation in this direction has started [13], but the main problem in its use for 3D acquisition is the lack of scale information as discussed in Section 5.1. Previous and current reconstruction processes [14,15] still rely on the integration with 3D scanning or photogrammetry. Hence, although someone claims that dense stereo reconstruction is a mature alternative to 3D scanning, no convincing comparison has been presented until now. Recently, some initial effort has been made in this direction [16], but an overall methodological definition and an accurate data assessment are still missing.

4.2. Theoretical and practical issues

Since our aim is to prove the effectiveness of dense stereo reconstruction as a fast but powerful way to gather 3D information in an archeological excavation, it is necessary to consider not only the results obtained, but also the applicability of this methodology to the specific domain.

Please cite this article in press as: M. Dellepiane, et al., Archeological excavation monitoring using dense stereo matching techniques, Journal of Cultural Heritage (2012), doi:10.1016/j.culher.2012.01.011

In particular:

- applicability to the object of interest: using dense stereo reconstruction, the completeness and quality of reconstruction changes drastically not only from object to object, but also among different areas of the same object. This happens with a higher variability with respect to 3D scanning and other surveying technologies. Additionally, there is a non-zero probability that the reconstruction may fail completely, without any clear cause of it. These problems seem to undermine the basic idea of this work; however, experimenting with different tools, we found that the archeology excavations are the ideal case for dense reconstruction. The excavations provide a mix of details at high frequency (the texture of the ground), medium frequency (small stones, masonry) and low frequency (the various levels and surfaces of the excavation) which perfectly adapts with the needs of dense stereo algorithms. The obtained 3D models are complete and the quality level of the data is uniform across the scene;
- cost of the acquisition equipment: with respect to 3D scanning devices, the hardware required for dense stereo reconstruction (apart from a PC for the data processing) is just a digital photo camera. We experimented with different cameras: from cheap compacts to high-end DSLR. While it was possible to process successfully all the dataset, the quality of the photo does affect the quality of the reconstruction. The quality of the lenses (i.e. sharpness of focus, low lens distortion) is more important than the number of megapixels. A higher resolution of the photos means a higher resolution in the extracted geometry. However, this requires more computation time and may cause problems during matching, if the pixels are much smaller than the distinguishable features. The general idea is that if it is possible to distinguish the details of interest in the photos, the resolution is adequate. A medium-range DSLR with a decent lens is usually perfect for the job;
- cost of training/expertise for acquisition and processing: the basic guidelines for the photographic acquisition are really simple and, with a short training, the archeologist himself can generate the photographic dataset. For the processing step, the processing tools proved to be effectively usable by Computer Graphics non-experts just after a few days (1–2 weeks) of experimenting;
- cost of data processing: the tools for 3D data generation from images adopt, at the moment, two different setups:
  - the ones based on a web service, where the user uploads the images and the computation happens on a remote server;
  - the ones based on stand-alone tools, where the computation is carried out on the user's computer. In both cases, there are free/open tools as well as commercial tools, and the computation may take from some minutes to some hours, depending on the dataset complexity. However, since no user intervention is necessary, this time does not impact too much on the real workload. Then, in most cases, the data produced from the dense reconstruction tools need some processing in order to produce a usable 3D model. All the required data processing may be carried out on a normal PC and, as we will show, using free and open source tools.

4.2.1. The scaling process

While dense stereo reconstruction tools are able to generate an accurate digital representation of an object, unfortunately, the generated 3D data/model is at an unknown scale. The system is able to determine the spatial relationship between the various reconstructed points (thus, producing a geometry with the correct proportions), but the general scale of the scene cannot be recovered.

Hence, the only unknown in the reconstruction is a scale factor, which may be easily and precisely obtained by a few measurements on the real-world scene. The same issue is found in most photogrammetric software tools, which require the manual specification of the distance between two (or more) points to recover the scale value. The basic idea is to find some points in the scene, measure them in the real world and in the 3D data, and use the ratio of these two measures as the scale factor.

Three approaches are possible to obtain these reference points:

- reference objects in the scene: if an object of known size is present in the scene, the scaling factor can be obtained using noteworthy points on its surface;
- points in the scene: the same reference can be obtained by selecting some points which are easy to recognize in both the real and the reconstructed scene and measure their real-world/reconstructed ratio;
- use of markers and/or total station: an alternative solution is to use some markers in the scene (Fig. 1), and measure their distances manually or with a total station.

The most straightforward process is to use as scaling factor the ratio between a single real-world distance and its reconstructed counterpart. However, to make things more robust, it is much better to measure multiple distances (possibly spanning across different directions, all over the scene) and use as the scaling factor the mean of these real-world/reconstructed ratios.

A step further would be using a series of marker or selected points to act as a reference "constellation", in order to determine the scaling of the scene and its relative orientation with respect to the other reconstructions.

A much better approach would be using the marker points, surveyed with a total station, during the camera calibration and orientation phase, to obtain a more precisely scaled reconstruction and, at the same time, its correct position in the chosen reference frame. Unfortunately, while some photogrammetry tools do provide this option, the class of automatic dense reconstruction tools we are considering in this work, does not yet offer this possibility.

4.2.2. Site preparation

No special preparation is required for the working site to be acquired. As stated, the standard excavation site is well suited for dense stereo reconstruction.

In order to be able to scale the reconstructed model correctly, a set of markers was put on the scene to be used as a reference for the scaling and alignment of the different models. In our case, we found that flat, round markers are the best solutions: they tend to show up in the reconstructed data as color blobs, making it easy to pick their center. The markers should be well distributed across the scene, if possible, avoiding too many coplanar positions. It is highly probable that some of the markers are removed during the excavation, when the material underneath is removed, for this reason, it is better to start with more markers than necessary, and placing some on the side of the trench (like the green ones visible in Fig. 1).

Also in this case, the nature of the archeological excavation turns to be useful, since the same grid used to document the archeological features with traditional methods, may also be exploited as reference to obtain the scaling factor, reducing the overhead.

4.2.3. On-the-field work

The image acquisition during the archeological campaign is a key step of the process. Every set of pictures has to be planned in advance, according to the kind of archeological context to document. Knowledge of basic photography does help in choosing the best camera parameters, and an inexperienced user may learn the tricks in a short time. In addition to the standard procedure for the acquisition for dense stereo matching (i.e. constant lighting conditions, sharp focusing on the photos, regular coverage of the site, no

Please cite this article in press as: M. Dellepiane, et al., Archeological excavation monitoring using dense stereo matching techniques, Journal of Cultural Heritage (2012), doi:10.1016/j.culher.2012.01.011

use of flash), some more specific guidelines are needed to obtain reliable data.

The basic idea is to take photos while moving around the site, trying to cover the entire area of interest (Fig. 9). It is important that each part of the site is covered from multiple angles and, possibly, keeping a uniform coverage.

Planning in advance the path inside the site is a good idea. Some reconstruction tools work better if there is strong continuity between subsequent photos. Close-ups of some more complex areas may be added at the end; again, it is better if the close-ups are not scattered shots, but a small sequence that, from afar, goes near the area of higher detail and covers the details.

It is much better to cover the site with a single set. There are, however, cases where multiple datasets are unavoidable. In these cases, it is advisable to have overlap between the datasets (to facilitate data integration) and to include in each dataset enough markers for the scaling process.

4.2.4. Data processing

We processed the test dataset with three different tools for the 3D reconstruction which are publicly available: the Arc3D web service, Photosynth/Bundler + PMVS2 and AutoDesk PhotoFly:

- Arc3D is organized as a web service. The 3D reconstruction via the web services is defined as a “black box”: the user uploads the data, and the server sends the results once ready. The resulting data consist in the matched images, a depth map for each image and all the matching information;
- the Photosynth/Bundler tools work in a two-step process: the photos are matched on a web service (Photosynth) or locally (Bundler) and then processed locally with another tool (PMVS2) for the dense matching. In both cases, the result is a single point-cloud;
- AutoDesk PhotoFly also works as a web service, but it is able to give back to the user a complete, triangulated 3D model, with a texture mapping built using the input photos.

Excluding the case of AutoDesk PhotoFly, the general output of this family of tools does require further processing, in order to obtain a complete and usable 3D model. A valid option for the data processing is MeshLab [17], an open source mesh processing tool.

Regardless of the technology used to acquire the geometry, the generation of a 3D model is a multi-step process and it is quite standard. For this reason, a tool which covers the various stages of the post-processing (mesh cleaning, data filtering, range scans alignment, surface merging..) is perfectly able to deal with data coming from dense stereo tools, as it is the case of MeshLab.

A complete description of this process is out of the scope of this paper, we will just outline here the peculiarity arising when processing this specific kind of data:

- data extraction: depending on the system used for 3D reconstruction, there is the possibility to “tune” the final results by combining different groups of images, or by changing the reconstruction parameters. The output of the reconstruction systems can be a triangulated surface or a point cloud;
- meshing and coloring: the triangulated surface is then meshed from the data using, in our case, the MeshLab Poisson reconstruction filter [18]. This filter removes part of the sampling noise and reconstructs an overall triangulated mesh from the sampled points by also closing all the small missing gaps (unsampled surface regions). Then, the color from the original data can be transferred to the model.

After the model generation, we need to perform the following post-processing actions:

- scaling: by measuring distance between markers on the 3D data and knowing their real-world distance, it is possible to recover the scale factor and bring the data in the correct scale;
- bringing the model in the correct reference frame: if some markers or known points have been acquired using a total station, the model can be easily brought into the same reference frame used for all the other analysis/surveys/measurements of the site. In this way, the produced documentation can be integrated in a GIS system or even geo-referenced.

Data processing and 3D model generation require practice, given the need to deal with geometry cleaning and meshing. Nevertheless, the proposed pipeline is consolidated enough to have non-CG personnel to manage the data. Archeologists have carried out most of the data processing needed in our experiment.

5. Experimental data and results

All the images used for data testing were acquired using a Nikon D70 with Nikkor lenses, with a resolution of 10 MPixel. The original dataset of 400 images was divided in several groups depicting the various stages of the excavation. Each subgroup was uploaded to the Arc3D server, and each level was generated with a number of images varying from 31 to 49. Unfortunately, part of the less interesting portions was not covered by the acquisitions. This suggested the need to acquire a bigger number of images, with an even more careful planning of the photographic campaigns.

After the processing pipeline described in previous section, seven models covering a period of 17 days (from 21st June to 7th July 2010) were created. The size of models goes from 1.2 to 1.8 million triangles. The complexity level of the geometry was chosen as the best trade-off between detail preservation and easiness of navigation.

The images presented in this paper are related to four time steps (Figs. 5 and 6), which have been selected to present easily visible changes in the excavation, like the enlargement of a posthole location and the creation of a new excavation.

5.1. Data assessment

The lack of a reference model of the excavation (e.g. a model produced with active 3D scanning) prevented from a point-to-point evaluation of the precision of the obtained data. Some preliminary work in this direction has been proposed [20,21], but a comprehensive testing activity would be useful.

For this evaluation, we used some of the markers that were present in the scene (but not used in the scaling process on purpose), but also measuring distances between reference points available in the original excavation documentation. In this way, it was possible to evaluate the precision of the 3D model with respect to the “manual”, on-site measurements.

The results were encouraging, with an average of 1.8 mm discrepancy in measures ranging from 1 to 2.5 m. Considering that the average pixel size on the ground is 3 mm, even with the reconstruction noise, the level of error is more than acceptable for the proposed activity. We left an extensive comparison with 3D scanning devices on an excavation to future tests, but we analyzed a possibly more critic feature of the reconstruction methods: the repeatability.

5.1.1. Repeatability

Repeatability is an essential property for any acquisition tool used to produce 3D models at different times and to support any

Please cite this article in press as: M. Dellepiane, et al., Archeological excavation monitoring using dense stereo matching techniques, Journal of Cultural Heritage (2012), doi:10.1016/j.culher.2012.01.011

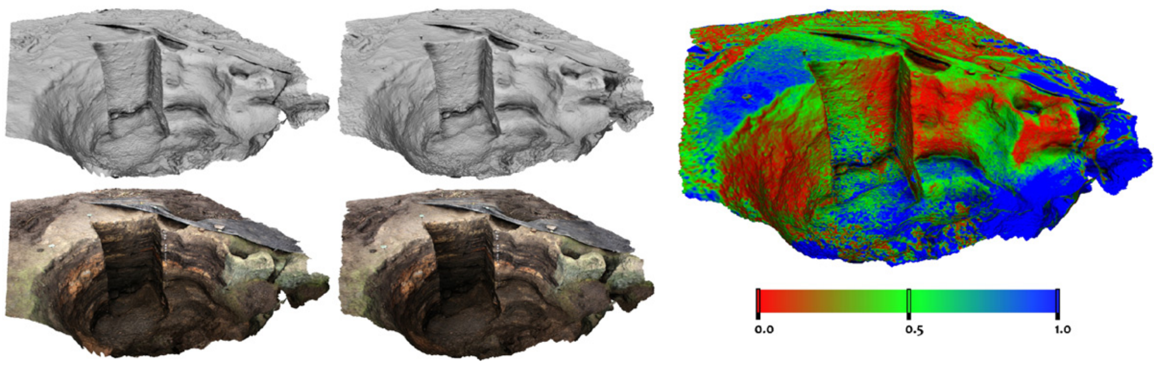

Fig. 3. Repeatability Test. Left: the two datasets of the same object (pure geometry above and with mapped color below). It is possible to appreciate their really similar appearance. Right: the color-coded deviation between the two models; 90% of the model is below 1 cm deviation. Reference color scale is shown below the model (unit is in cm).

further metric comparison between those models. Starting from different photographic dataset of the same object, we need the system to be able to produce reconstructions with a difference below the minimal amount of modification we want to measure between different stages of the excavation. Our concern is that, when acquiring two subsequent steps in the excavation, the unchanged area should produce two (almost) identical geometries, to avoid to report false changes in the excavation shape.

To evaluate repeatability of dense stereo matching reconstruction, we acquired two sets of photos of the same object and processed them independently using Arc3D, finally comparing the obtained models.

We show here a comparison between two models of an area of the Uppakra site (one of the postholes in the excavation, around 2 m wide). In order to be able to compare the data, the models were scaled and aligned using some markers in the scene, as previously described.

Looking closely at the geometries obtained from the two dataset (as visible in Fig. 3), it is possible to observe how the reconstructed features visible in the first model are clearly recognizable also in the second model, making possible an effective qualitative comparison.

The measurement of the geometrical deviation between the two models showed that more than 90% of the surface has a deviation of less than 1 cm, and 50% of less than 0.5 cm: it is possible to see a graphical representation of this data in Fig. 3, where the deviation error is color coded. The error distribution does not present any particular pattern or distinct accumulation (which would indicate some kind of non-rigid deformation or bias).

5.1.2. *Considerations on the reconstruction tools*

The data assessment has been carried out on 3D models generated by three tools (Arc3D web service, Photosynth/Bundler + PMVS2 and AutoDesk PhotoFly), obtaining quite similar geometrical results. A simple geometrical comparison between the geometries obtained from the three tools showed a level of discordance similar to the one found in the repeatability evaluation (less than 1 cm for geometries around 2 m in size).

We can safely say that the three systems may be used in this kind of activity.

There are, however, differences in how the tools behave in terms of output data density, resilience to non-optimal photo dataset (insufficient photo coverage, photos too far apart, blurred photos, lighting problems), visual/geometrical quality of output data and tool flexibility.

Arc3D proved to be extremely flexible. The possibility to tweak extraction parameters for each processed photo make sometimes possible to obtain usable results also from difficult dataset; also the density of reconstruction and the data cleaning are fully in the hand of the user. On the downside, Arc3D seems to more often fail to reconstruct the 3D data if the photo dataset is non-optimal (or even if the photos are not in the correct order).

Photosynth/Bundler + PMVS2 is a bit less flexible and easy to use; however, the user can operate on data density and cleaning operations. Moreover, the system is very robust even when dealing with non-optimal photographic sets.

PhotoFly tool is completely automatic, and only the final resolution of the model may be selected (using a simplistic low-medium-high option). The visual quality of results is usually very good, but the surface creation algorithm tends to automatically fill unsampled areas: while this may not be a problem for small holes, it is often difficult to distinguish between a correctly reconstructed area and a filled one, making somehow unreliable the data comparison between time steps.

Since we already had some background experience in using Arc3D, we chose to use this tool for the generation of the bulk of the 3D models used in the next sections (the different time stages of the excavation).

5.2. *Visualization*

The test case presented in Section 5 produced several layers of the same excavation. The geometric data were all scaled to real measures and aligned using the measurements made with the total station on the markers in the scene.

The aim of the test was to show that this type of data can be a very valuable instrument for monitoring and documenting an archeological excavation. Hence, a series of ad-hoc features have been implemented in the context of MeshLab [17] by following the initial requests of the test users. These initial features were chosen to show that the work of the archeologists could be improved by an easier way, obtaining some of the typical data used for documentation and the possibility to extract additional information in a straightforward way. A snapshot of the interface of this specific MeshLab feature is shown in Fig. 4.

5.2.1. *Spatio-temporal visualization*

Once we have put all the acquired 3D models in the same reference frame, we are able to visualize spatial and temporal information at the same time.

If a label with the acquisition date is attached to each model, a very simple functionality can order the 3D models by date, and show them one by one using a slider. Hence, the evolution of the excavation can be easily visualized by changing the position of the slider.

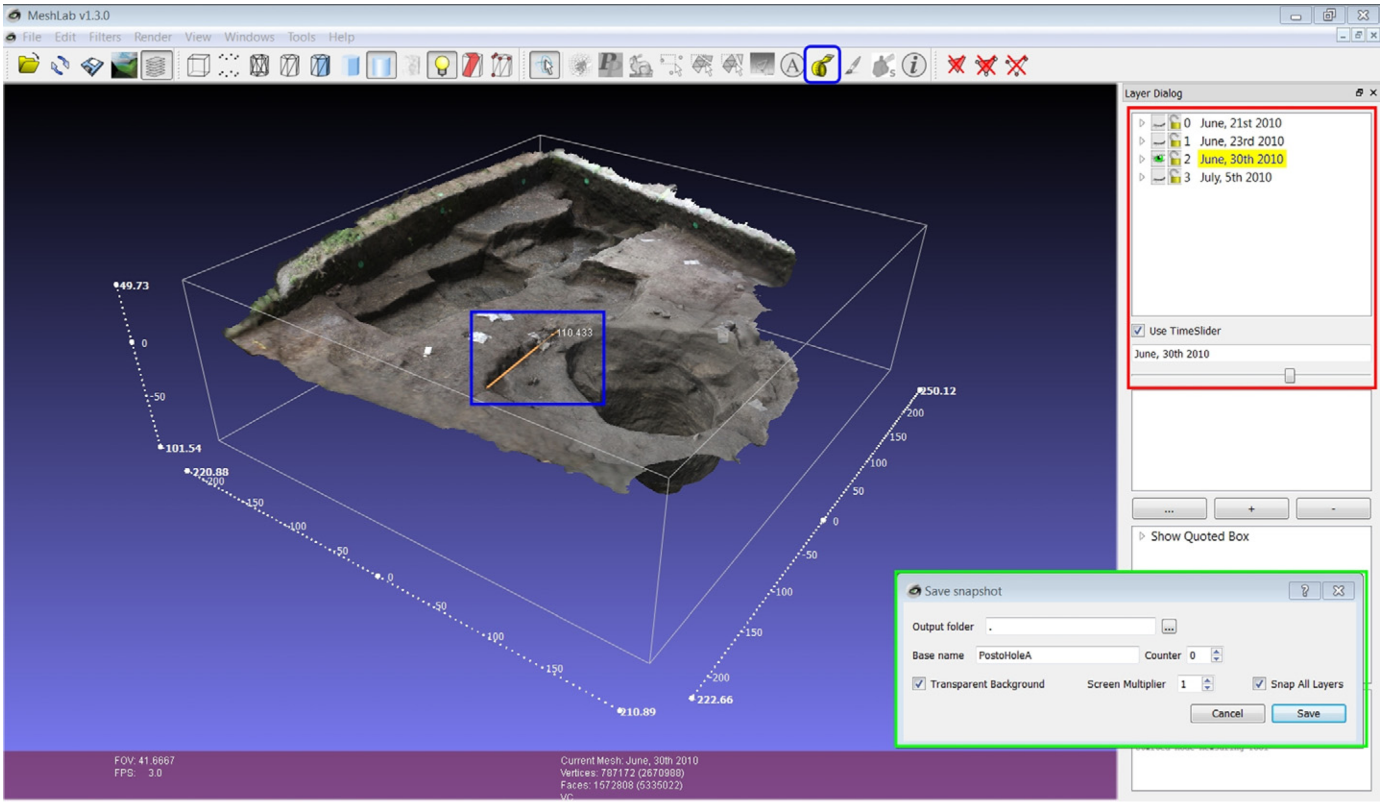

Fig. 4. A snapshot of MeshLab interface: the red box shows the Time Slider, the blue boxes shows a single measurement tool result, the green box shows the multiple snapshots tool.

This provides a very intuitive way to find the areas that were modified, and the "story" of the excavation is always available. Generated data (measures, snapshots) may be automatically marked with the correct time step. In addition, visualizing the levels from different points of view can provide additional information, and possibly suggest future actions. The red box in Fig. 4 shows the intuitive interface of Time Slider, where moving the slider can change the acquisition day.

5.2.2. *Data extraction tools*

The visualization of the set of 3D models in the context of MeshLab gives direct access to some simple features which can be extremely important to extract significant data. One of them is the measurement tool, that allows the user to select two points and computes the distance between them (see the blue box in Fig. 4, inserted in the image to highlight a selected distance and the numeric value computed by MeshLab).

The *snapshot* feature is another simple but effective tool. The resolution of the screen is usually not sufficient to produce good quality illustrations (e.g. to be printed at the large scale). Regardless the result produced on the screen, MeshLab gives the possibility to re-generate it at a user-defined high resolution, using the same view, illumination or shader, making it a useful instrument for documenting on paper or with illustrations a geometric 3D model.

To make the documentation of this kind of datasets easier, the possibility to take, with a single button, a snapshot for each layer was added (green box in Fig. 4).

The snapshot functionality is important for two main reasons:

- it easily creates the type of data which is used by archeologist for current monitoring analysis. Snapshots can be used as "photos" of the excavation, so that most of the typical analysis can be done on them;
- it generates an image of the various excavation phases from a perfectly overlapping point of view, which is hard to obtain using photos. In addition, canonical views (i.e. orthographic view from the top, front, left, right) are created with a click. For example, Fig. 5 shows the evolution of the excavation from the top of the azimuthal direction, a point of view that is almost impossible to obtain with a photograph. Fig. 6 describes the same evolution from the left side, showing only the geometry information.

Moreover, by saving the camera parameters (position, orientation and FOV), it is possible to memorize an interesting view and being able to replicate it later, for example, as new data become available.

The interactive slicing rendering is another visualization mode that enables the local inspection of the 3D models. The idea is to clip the model with a slicing plane, which can be moved along one of the main axis. In this way, the profile of the excavation in a particular area becomes more apparent.

The user may select the axis along which to move the plane and control the slicing offset. It is possible to see one half of the sliced model (the one behind the slicing plane), or to render only the cross section. This visualization paradigm is useful both on a single mode (to better appreciate the cross section of the excavation) and on multiple visible layers (to compare the evolution of such cross sections, like in Fig. 7). By mixing and matching the various options, it is possible to obtain interesting rendering, like the one in Fig. 8. This is just a visual representation but, using the measuring tool, it is possible to easily measure the displacement between the time steps in any spot.

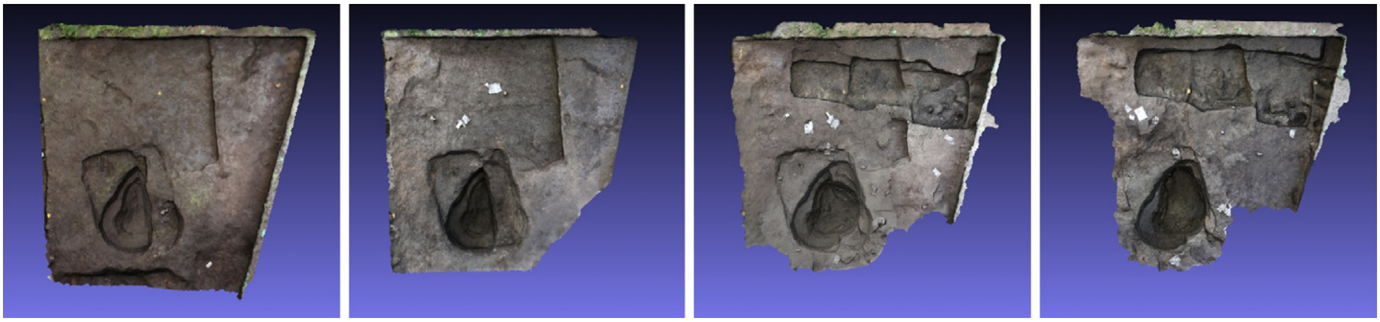

Fig. 5. Snapshots of the various time steps from the canonical "top view", showing the evolution of the excavation.

MeshLab will very soon include also a filter that will allow saving cross sections in a vector format (SVG) that is now under development. The possibility to generate and export cross sections in a format usable inside desktop publishing software (but still containing the correct measures) is an incredibly useful feature, making it easy to generate scale drawings and schemes, integrable in an existing documentation protocol.

5.2.3. *Photographic reference*

Since the photos used in the 3D reconstruction have been, in the matching process, geo-referenced in the 3D model space, it is possible to retrieve their position w.r.t. the reconstructed 3D model, and possibly take advantage of it.

Fig. 9 shows a snapshot of the 3D model obtained with one photographic acquisition, together with the corresponding positions in space of all the images used for the reconstruction. The acquisition path is clear, with the coverage going from the whole scene to the most important features (the postholes).

Moreover, it is also possible to go beyond this, and reference, with a minimal manual intervention, new images to the 3D dataset. This may be useful to locate/project in their specific spatial context photos with special illumination (grazing light, fluorescence...), photos with reference objects, but also sketches done by the archeologist or historical photos.

In this way, the typical documentation material could be geo-referenced, creating a spatial documentation, easier to access and navigate.

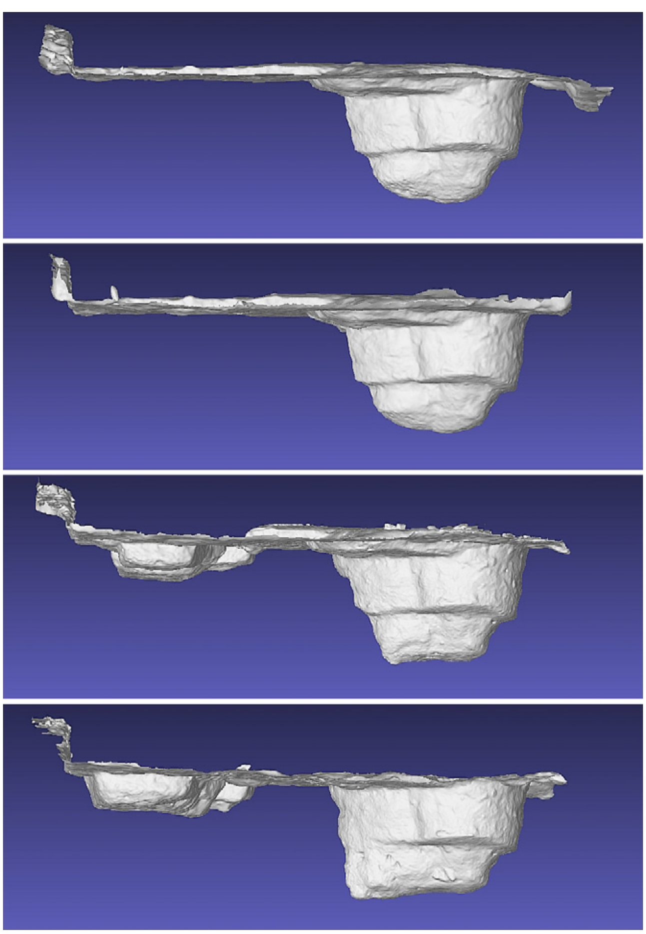

Fig. 6. Snapshots of the various time steps from the canonical "left side" view, showing the evolution of the geometry of the excavation.

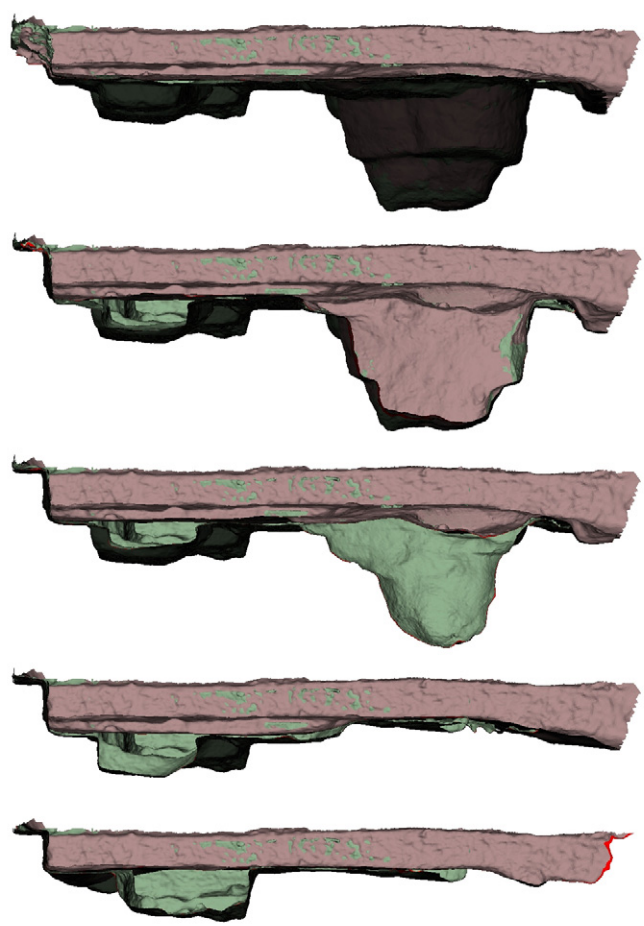

Fig. 7. Interactive slicing visualization. The dataset is viewed from left, and the slicing plane is moving farther along the view direction, showing the evolution between two time steps.

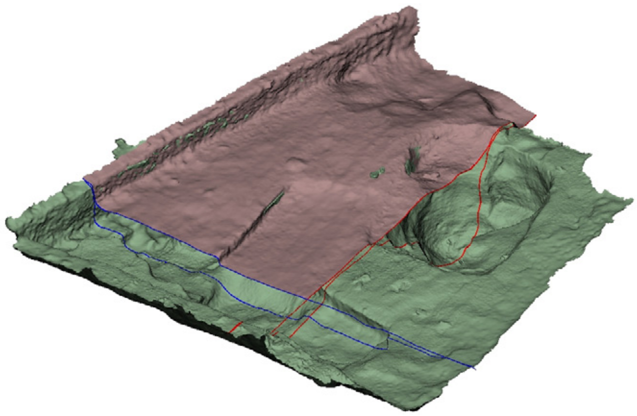

Fig. 8. Interactive slicing visualization. From a three-quarter view, using two slicing plane and the cross section visualization, it is possible to appreciate the evolution of the excavation in specific areas of the site.

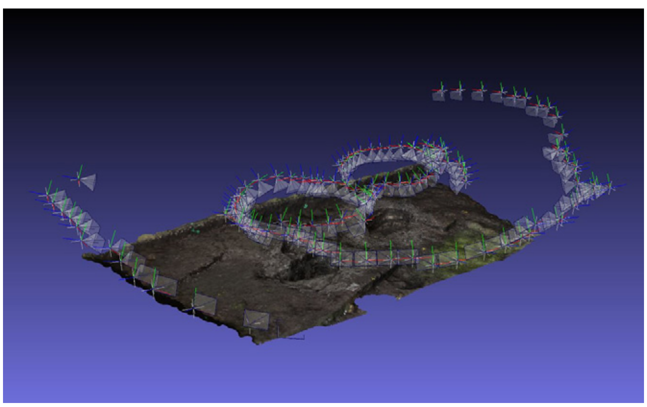

Fig. 9. Snapshot of a reconstructed excavation, with the respective positions of the images used for dense stereo matching.

6. Conclusions and future improvements

In this paper, we presented a methodology for the use of dense stereo matching systems for the monitoring of an archeological excavation and an assessment of the results obtained in an on-site campaign.

The results of this experiment demonstrate that, by following a careful acquisition and a processing pipeline, it is possible to produce quite accurate and reproducible data. This work proved how the use of this methodology for the archeological investigation exponentially increases the documentation quality of the site, reducing in most cases the time spent to collect the data. This method, if properly combined with other technologies such as Total Station or GPS (GNSS), can generate very powerful spatio-temporal information.

The material acquired and elaborated during this experiment was used primarily for data assessment, and for showing a new perspective in visualization and analysis: a spatio-temporal context where the evolution of the excavation is exposed in a very intuitive way. In addition, simple tools for the creation of snapshots and enhanced visualization showed that 3D models can improve the quality of documentation strategies, with a minimal overhead of time for data acquisition.

Another advantage of the proposed system is the integration with the current documentation pipeline workflow: three-dimensional models can be added to the usual data, and it is also very simple to provide new information in the state-of-the-art format (i.e. creating snapshots or measuring distances, or even including 3D models in text documents, for example using the Acrobat3D data format to add interactive 3D models to pdf documents).

Our belief is that this family of 3D-from-images tools may be of great help also in the framework of archeological studies: being able to visualize the complete geometry of the excavation, with its evolution in time, may be an invaluable instrument when teaching the excavation techniques and the management of a site. Like chess players may learn by studying important matches of the past, it would be possible for the students to explore (spatially and temporally) important excavations carried out all around the world.

Possible future improvements are:

- improvement of output data: during the presented test, the bulk of the dataset has been processed using Arc3D (while the other tools have been used for the data assessment), but a web-service-based system prevents from having any kind of feedback during acquisition. The next tests will be performed directly on the field using also Bundler and PM2VS [22] in all the time steps. If a robust processing pipeline will be found, it could be possible to refine the acquisition on-site.

  Another possible extension could be the improvement of color detail, using the input images to re-project the color information on the 3D model [23]; a new version of MeshLab is planned on fall 2011, including this enhanced management of color and re-projected textures.
- definition of a formal guideline for archeologists: based on the results of the tests and on the feedback by the users, a more detailed guideline for the monitoring using dense stereo matching could be defined in order to standardize the acquisition process. Some of the important issues to take into account could be: markers placement, photographic campaign strategy, processing and integration of data.
- annotation: in order to perfectly integrate the 3D models in the documentation process, simple annotation tools should be made available, so that text, images and other contents can be spatially and temporally referenced directly on the 3D models.

Acknowledgements

We want to thank the all excavation staff of Uppåkra for their help and assistance during the entire duration of the project. In particular, we want to thank the Professor Lars Larsson and the excavation leaders Birgitta Pitz Williams and Karl Magnus Lentorp.

The research leading to these results has also received funding from the European Community Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 231809 (IST IP "3DCOFORM").

References

[1] E. Zubrow, Digital Archaeology A historical context, in Digital Archaeology bridging method and theory, edited by Thomas Evans and Patrick Daly, 2006, pp. 10–31.
[2] P. Barker, Techniques of archaeological excavation, B.T. Batsford, London, 2005.
[3] M. Bianchini, Manuale di rilievo e di documentazione digitale in archeologia, edited by Aracne, 2008.
[4] M. Brown, D.G. Lowe, Unsupervised 3d object recognition and reconstruction in unordered datasets. In: Proc. of. Int. Conf. on 3-D digital imaging and modeling (Washington, DC, USA, 2005), IEEE Computer Society, pp. 56–63.
[5] D. Martinec, T. Pajdla, Robust rotation and translation estimation in multiview reconstruction, Comput.Vision Patt. Recogn. IEEE Computer Society Conference on 0 (2007) 1–8.
[6] M. Vergawen, L. Van Gool, Web-based 3d reconstruction service, Mach. Vis. Appl. 17 (2006) 411–426.
[7] M. Goesele, N. Snavely, B. Curless, H. Hoppe, S.M. Seitz, Multi-view stereo for community photo collections, Proceedings of ICCV 2007, Rio de Janeiro, Brasil, October 14–20, 2007.
[8] M. Havlena, A. Torii, T. Pajdla, Efficient structure from motion by graph optimization, ECCV 2 (2010) 100–113.

[9] M. Farenzena, A. Fusiello, R. Gherardi, Structure-and-motion Pipeline on a Hierarchical Cluster Tree, Proceedings of “IEEE International Workshop on 3-{D} Digital Imaging and Modeling”, Kyoto, Japan, October 3–4, 2009, pp. 1489–1496.

[10] M. Snavely, S. Seitz, R. Szeliski, Photo tourism: exploring photo collections in 3d. In: ACM SIGGRAPH 2006 Papers (New York, NY, USA, 2006), SIGGRAPH 06, ACM, pp. 835–846.

[11] Phototourism: exploring photo collections in 3D. More info on: http://phototour.cs.washington.edu/.

[12] Y. Furukawa, J. Ponce, Accurate, dense, and robust multiview stereopsis, pattern analysis and machine intelligence, IEEE Transactions on 32 (8) (2010) 1362–1376.

[13] K. Rodrigueza-Echavarria, D. Morris, D. Arnold, Web based presentation of semantically tagged 3D content for public sculptures and monuments in the UK. In: Proceedings of the 14th International Conference on 3D Web Technology (New York, NY, USA, 2009), Web3D 09, ACM, pp. 119–126.

[14] A. Gruen, F. Remondino, L. Zhang, Image-based automated reconstruction of the Great Buddha of Bamiyan, Afghanistan, Comp. Vision and Pat. Rec. W. 1 (2003) 13.

[15] S. El-Hakim, J.A. Beraldin, M. Picard, L. Cournoyer, Surface reconstruction of large complex structures from mixed range data - the Erichtheion Experience. IAPRS&SIS, 37(B5), pp. 1077–1082, XXIth ISPRS Congress, Beijing, China, 2008.

[16] S. Hermon, D. Pilides, N. Amico, A. D’Andrea., G. Iannone, M. Chamberlain, Arc3D and 3D laser scanning a comparison of two alternate technologies for 3d data acquisition. In: Proceedings CAA, Granada, April 2010, pp 55–58.

[17] Meshlab: an open source, portable, and extensible system for the processing and editing of unstructured 3D triangular meshes. More info on: http://meshlab.sourceforge.net/.

[18] M. Kahzdan, M. Bolitho, H. Hoppe, Poisson surface Reconstruction. In: Proc. of Eurographics symp. on Geometry processing (Aire-la-Ville, Switzerland, Switzerland, 2006), Eurographics Association, pp. 61–70.

[19] L. Larsson, The uppakra project. Preconditions, performance and prospects, Ac. Arch. Lund. 80 (2003) 3–26.

[20] E. Fratus de Balestrini, F. Guerra, New instruments for survey: on-line softwares for 3D recontruction from images, 3DARCH 2011, Trento.

[21] S. Hermon, D. Pilides, N. Amico, M. Chamberlain, A. D’Andrea, G. Iannone, P. Ronzino, The Hill of Agios Georgios, Nicosia: 3D Analysis of an on-going Archaeological Excavation, CAA 2010, Granada.

[22] Bundler: structure from motion for unordered image collections. More info on: http://phototour.cs.washington.edu/bundler/; 2010.

[23] M. Callieri, P. Cignoni, M. Corsini, R. Scopigno, Masked photo blending: mapping dense photographic dataset on high-resolution 3D models, Comput. Graphics 32 (2008) 464–473.

Please cite this article in press as: M. Dellepiane, et al., Archeological excavation monitoring using dense stereo matching techniques, Journal of Cultural Heritage (2012), doi:10.1016/j.culher.2012.01.011