IMAGE-BASED VIRTUAL RECONSTRUCTION OF COMPLEX ARCHITECTURES USING SOPHISTICATED AUTOMATED MATCHING PROCEDURES – THE ARTEMIS TEMPLE IN JERASH, JORDAN

E. Baltsavias[1], F. Remondino[1], R. Weckerle[1], J. Ohlmann[1,*], N. Al-Hanbali[2], O. Al Bayari[2], B. Saleh[2], H. Almasri[2], J. Al Azizi[2]
[1]Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland
[2]Surveying & Geomatics Eng. Dept., Al-Balqa Applied University, Al Salt, Jordan
[1](manos, fabio)@geod.baug.ethz.ch
[2]nedalalhanbali@yahoo.ca, obayar@yahoo.it, Bsaleh@hotmail.com, husam_masri@yahoo.com
http://www.photogrammetry.ethz.ch

COMMISSION V – WG 4

KEYWORDS: 3D Modeling, Preprocessing, Orientation, Matching, Texturing, Automation, Cultural Heritage

ABSTRACT:
Compared to other recording and modeling methods, images can be acquired with almost inexpensive systems and contain all the information for the generation of detailed and photo-realistic 3D models. But recovering a complete, accurate and realistic 3D model from images is still a difficult task, in particular if uncalibrated or widely separated images are used or in case of complex architectures like ruins of temples with lots of details and often not well defined edges. In this article, the 3D image-based modeling of the Artemis temple in Jerash, Jordan, is reported. Jerash is the best preserved Roman provincial city in the world and includes monuments of various civilizations from the Neolithic epoch to the Muslim one.

1. INTRODUCTION

3D modeling of an object can be seen as the complete process that starts from the data acquisition and ends with a virtual model in three dimensions that can be visualised interactively on a computer. Nowadays, the generation of computer 3D models is mainly achieved with range sensors or image data while in some cases other information like CAD, surveying or GPS data are also integrated in the project. Different applications and fields require 3D models, from the traditional industrial inspection and robotics to the recent interest for visualization, animation and documentation of Cultural Heritage objects, sites and landscapes.

Compared to other recording and modeling methods, images can be acquired with inexpensive systems and contain all the information for the generation of detailed and photo-realistic 3D models. But to recover a complete, accurate and realistic 3D model from images is still a difficult task, in particular if uncalibrated or widely separated images are used or in case of complex architectures like ruins of temples with lots of details and often not well defined edges. In addition, automation of the whole process and in particular of tie point transfer in triangulation and 3D surface generation via matching is in most cases not possible due to the complex objects and the close-range imaging but also due to the poor performance of commercial software systems.

In cooperation with the Dept. of Surveying and Geomatics Eng., Al Balqa Applied University (BAU) in Jordan, our aim is to establish a 3D GIS and a virtual reconstruction of the whole Jerash (Jordan) archaeological site in order to enable documentation, digital preservation, reconstruction and visualization and thus support archaeologists and the general public administration. For the main monuments of the site, we aim at a detailed generation of textured 3D models. We started the work with one of the most important monument, the temple of Artemis. 3D digital models enhance the understanding of heritage sites that need to be preserved and shown to more people, increasing the awareness and understanding of the past local culture. In this article, the whole image-based 3D modeling methodology of the temple is presented, reporting also critical factors and encountered problems. This work is parallel and partly complementary to similar work performed at BAU (see Al-Hanbali et al., 2006) using different approaches and software for preprocessing, orientation, matching and 3D surface generation. The main difference to previous BAU work was on image preprocessing and automated matching and 3D surface generation. The work was performed within a diploma thesis and thus had time constraints, not allowing a complete modeling of the temple.

2. JERASH AND THE ARTEMIS TEMPLE

Jerash city, about 70 km N-NW of Amman, has a significant importance in the ancient history of Jordan(see Figures 1 & 3). The ancient town of Jerash (Gerasa or Jarash), Jordan, has a remarkable record of human settlement since the Neolithic times. Few Roman ancient towns are as complete and well-preserved as Jerash, a city complex that once was a thriving commercial zone and part of the Decapolis. Built in the 2nd century BC, the city was conquered in 63 BC by the Roman general Pompey. The town reached its peak in the 2nd century and declined after a series of Christian and Muslim invasions and earthquakes in the mid 8th century. Jerash is a high-profile archaeological site with 4 forums, theatres, public squares, baths and 41 temples. Although well-preserved, the great majority of the monuments are unrecorded and unprotected and are constantly endangered by development projects. Due to the population growth and the expansion of modern Jerash, many important monuments in the site begun to disappear (cemeteries, monuments from Bronze and Iron Age) and therefore the monitoring and the digital documentation of the site, including both the built and the natural environment are

* Exchange student for the summer semester 2006 from Technical University Munich, Germany.

really essential and necessary for the conservation and protection of the cultural heritage area.

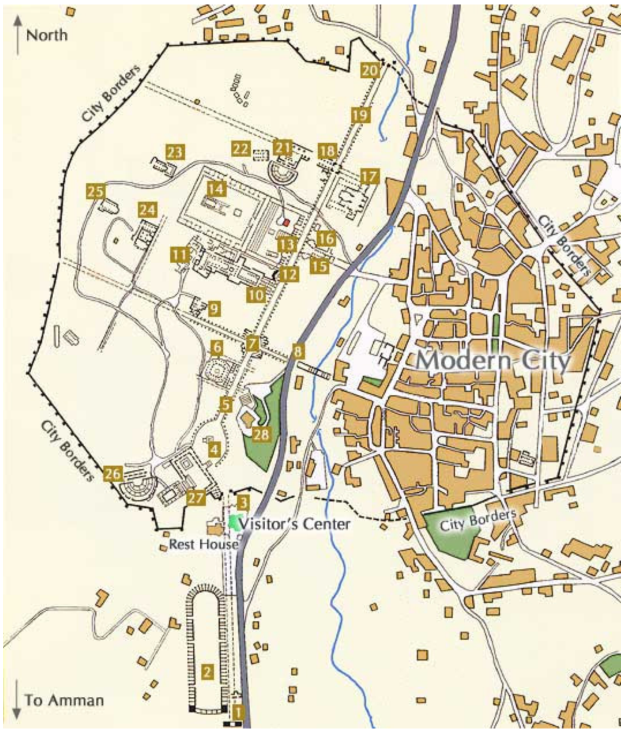

Figure 1: The ancient and modern Jerash City. The temple of Artemis is marked with 14.

BAU started to use Geomatics technologies, such as GPS, satellite remote sensing, photogrammetry and classical geodetic instrumentation, to build a 3D GIS and precise base maps for documentation purposes of this archaeological site in 2005. The objective is to start a national project that can be organized in cooperation with the Department of Antiquities, Jordan for comprehensive documentation of the Jerash city, and possibly other sites in Jordan. This work is planned to be carried out from macro level, i.e. the modeling of the modern and ancient city via 3D GIS, to micro level, i.e. building 3D virtual models as well as 3D GIS databases for each monument (Al Bayari, 2005; Al-Hanbali et al., 2006).
In the following, we report the first modeling results of the south side of the Artemis Temple (Figure 2). The entire Artemis temple is quite large: It is about 53 m long with a 13 m part for the entrance and stairs, 22 m high and 23 m wide. The interior chamber is 26 m long and 13 m wide (Figure 2). The final and complete 3D virtual reality model will serve as an important and accurate documentation of the temple. It was never documented as such before. The complete Artemis model will be then used in a 3D GIS model.

Figure 2: An overview of the south side of the Artemis temple.

3. IMAGE-BASED MODELING OF THE ARTEMIS TEMPLE

Often 3D modeling is meant only as the process of converting 3D geometric measurements into a geometric mesh and maybe also textured surfaces, while some applications require additional steps. The overall image-based 3D modeling process (Remondino and El-Hakim, 2006) consists of the following steps:

1. Design (sensor and network geometry)
2. Measurements (point clouds, lines, etc.)
3. Structuring/Modeling (geometry, texture)
4. Visualization/Analysis

Currently, the recovery of the sensor and network geometry and the measurement phase are often separated from the modeling and visualization parts. But in many applications, this gap has to be bridged in order to acquire suitable images, perform correct measurements and recover realistic 3D models.
In the next sections, the entire modeling process of the Artemis temple is reported.

Figure 3: The Forum in Jerash, with the Artemis temple on the far left and the modern city on the far right.

3.1 Data Acquisition

For our modeling project, we employed a Hasselblad camera with Kodak DCS Pro Back Plus (16MPixel chip), 9 microns pixel spacing, 38 mm lens, an average image scale of 1:330, and a ground pixel size of about 3 mm (Figure 3). The camera was calibrated in the lab using the 10-parameter Brown model, colour coded targets and the program iWitness^{TM}. Over 800 images (inside and outside) were acquired for the modeling of the object. The images in most parts were taken at 3 different heights (at 1.6 and 2 m with a tripod, and about 4.5-5 m with a ladder), as 3 strips with camera stations in each strip at about 4 m distance, resulting in a forward overlap of about 75%, a vertical overlap between 1st and 3rd strip of about 60%. In addition, at most camera stations except the frontal images, images at about 45 deg to the left and right were taken, in order to increase the number of images and reduce occluded surfaces.
The digital camera was heavy and cumbersome to use. For unknown reasons, in some cases it stopped working for hours. The LCD viewer had poor contrast, in some cases the images were shown with artifacts, although the saved images were correct, and in other cases the LCD was showing nothing, thus images were taken blindly and several times to ensure full coverage of the object. Due to different image acquisition dates and illumination conditions, the images had quite varying radiometry, while due to the wide FOV, images, especially at borders were not sharp.

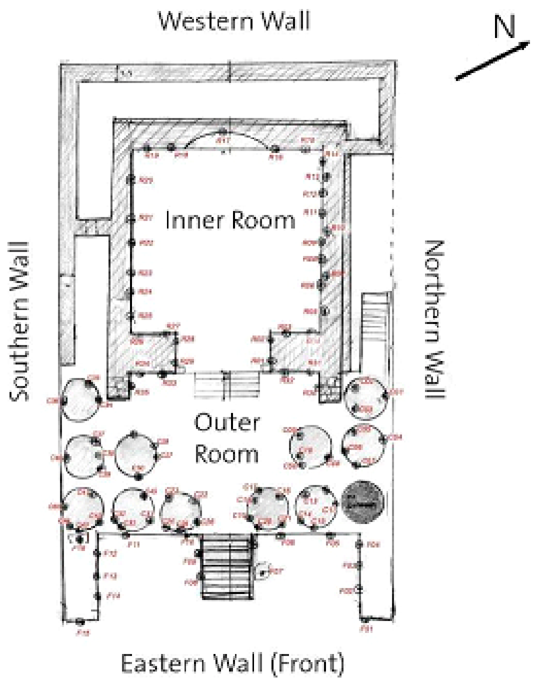

Figure 2: An overall view of the Artemis temple in Jerash, with some of the control points in the inner and outer room and front part.

Moreover, a geodetic network of 16 stations was constructed around and in the temple using GPS and laser total stations, for the datum definition and to determine the coordinates of over 160 points at the temple which were used as control and check points. These points were mostly corners of dark green tape stuck on the monument and, in some cases, black tennis balls stuck with nail and tape on some temple positions, offering less perspective changes than the tape points in the images (Figure 3). Their average accuracy was 1 cm in height and 1.5-2 cm in planimetry. The tennis balls were scarce and due to the ball weight instable and were not further used. The tape control points could not be placed at high positions on the walls and the entrance columns, leading to a suboptimal control point distribution.

Figure 3: The used Kodak DSC Pro Back Plus camera (left) and the signalized tape and tennis ball control points (right).

3.2 Image Pre-processing

The original DCR images in a proprietary Kodak format were transformed using the Kodak PhotoDesk software to TIFF. There are various versions of PhotoDesk and plug-ins and various program parameters that can influence significantly the resulting TIFF images, especially their brightness, contrast, color balance, noise, sharpness etc., which should be taken into account, especially if no further preprocessing is applied. Although the images could be saved as 12-bit and then maybe reduced to 8-bit, this was not done to save storage space and mainly because we estimated that the difference between 8 and 12 bit would be minimal for the task at hand.
Different image pre-processing techniques are applied to (1) replace permanent defects in the sensor elements, (2) reduce the noise and enhance edges, (3) enhance the contrast while simultaneously equalizing the image radiometry and (4) estimate the misregistration of the colour channels. Procedures 1-3 were applied to the green channel which was selected due to its better contrast and less noise for the further automated DSM generation. Procedure 4 was performed to estimate lateral color lens aberration, but no correction of the color images was performed, due to time constraints and relatively small effect of the corrections.
The pixel defects were detected from various images taken with a cover on the lens and were corrected using different filter mask sizes (3^2 or 5^2 pixels) and interpolation procedures (e.g. inverse distance weighting). The defects include single pixels, groups of pixels (e.g. crosses) and a part of a column (see Figure 6 top).
For the noise reduction and edge enhancement, a nonlinear, local adaptive edge preserving smoothing filter was employed. Since the noise level, estimated using homogeneous areas, was already low (0.5-1 grey value), the 20%-40% noise reduction after employing this filter did not have a significant effect, except maybe from the homogeneous areas (e.g. sky), which were however masked out in automated processes like matching.
Contrast enhancement and simultaneous equalization of radiometry is performed using the Wallis filter (Wallis, 1976). The filter enables a strong enhancement of the local contrast, especially at dark areas, while retaining edge details and removing low-frequency information in an image (Figure 4, bottom). Careful choice of the Wallis parameters is needed to optimise the results for subsequent matching and reduce artefacts that can be introduced by Wallis filter like noise in homogeneous areas, overshoots at high-contrast edges, and visible radiometric jumps at the borders of the filter blocks.

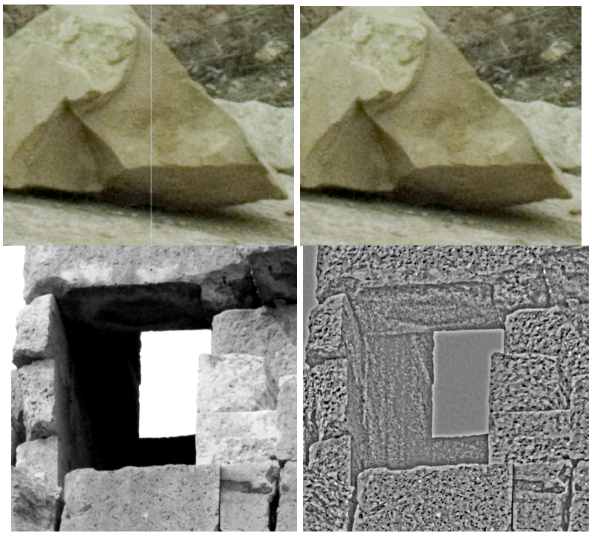

Figure 4: Results of the column pixel defect correction (top middle) and image enhancement, especially in dark regions, by means of Wallis filter (bottom, before and after Wallis).

Finally, the misregistration of the 3 colour channels, due to lens lateral chromatic aberrations, was estimated (see also Kaufmann and Ladstädte, 2005). This is a common problem, especially in commercial of-the-shelf digital cameras, and increases with increasing chip size (number of pixels) leading to clearly visible colour misregistration of several pixels at edges, especially at the image borders and against the sky. This effect is not always symmetric around the image centre and is made worse by use of color filters on the sensor elements (in this case Bayer filter). To evaluate the channel misregistration, a very dense matching of a relatively planar surface with good texture for matching was performed between the Red and Blue with the Green channel. The dense matching results (ca. 650,000 points) were binned in larger image blocks (200x200 pixels) to smooth the results and account for matching errors. The shifts between the channels were at max a bit less than one pixel, supporting the view that much of the color misregistration is due to interpolation and the use of color filters on the sensor elements.

3.3 Image Network Orientation

The image orientation is a very important step within the 3D modeling pipeline. To achieve the best results together with accuracy estimates, the image orientation is generally performed by means of a photogrammetric bundle block adjustment. In order to speed up the entire modeling pipeline, automation in the orientation step was highly investigated (Beardsley et al., 1996; Fitzgibbon and Zisserman, 1998; Pollefeys et al., 1999; Roth, 2004; Roncella et al., 2005; Läbe and Förstner, 2006; Remondino and Ressl, 2006). This has to operate on two issues: (i) the automatic measurement of tie points (without requiring sticking markers or targets on the object) and (ii) the automatic provision of initial orientation parameters for the bundle block adjustment. Whereas (ii) is nowadays easily solved – once the image correspondences are given – using perspective- (Cronk et al., 2006) or projective-geometry (Hartley and Zisserman, 2001) based formulations of the relative orientation, the automatic measurement of markerless tie points is still a challenging topic, especially in close-range images where no commercial solutions are available.

For the south side of the temple, between the over 100 images available, 30 are selected and used for further processing. For the orientation, we did not employ any automated tie point extraction procedure, due to the complexity of the object, the quite different imaging viewpoints, and the lack of appropriate automated software. Tie points, and also control points, were therefore measured manually and monoscopically using PhotoModeler, with an average of 20-30 tie points measured in at least 3 images (about 450 tie points in total), a time-consuming process. Among the signalised points, 16 were used as control and 16 as check points. Major problems were the lack of clear edges, occlusions, holes and plants on the temple walls. After a first bundle block adjustment using the partly "blackbox" Photomodeler to provide approximations, the IGP program SGAP was used for the final bundle adjustment. Its results include: RMSE of 16 checkpoints: X =12 mm, Y = 15 mm, Z = 36 mm ; and a sigma 0 of 9 microns (1 pixel).

3.4 Surface Measurement and 3D Model Generation

The pre-processed and oriented images are afterwards used for the subsequent automated image matching and DSM generation For the object reconstruction, our in-house multi-image, multi-primitive geometrically-constrained matching strategy is applied, to recover dense 3D point clouds. The multi-image matching approach was originally developed for the processing of satellite and aerial images (Zhang, 2005) and it has been extended to a certain extent to process close-range images (Remondino and Zhang, 2006).

The approach uses a coarse-to-fine hierarchical solution with an effective combination of several image matching algorithms and automatic quality control (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Starting from the known calibration and orientation parameters, the approach performs a matching of multiple primitives (feature points, grid points and edges) and afterwards optionally refines the results with a multi-photo geometrically constrained least squares matching (Gruen and Baltsavias, 1986). As multi-image concept, contrary to stereo-pairs approaches (Scharstein and Szeliski, 2002; Strecha et al., 2003; Pollefeys et al., 2004), highly redundant matching results are obtained. The high redundancy also allows automatic blunder detection. Mismatches can be detected and deleted through the analysis and consistency checking within a small neighbourhood.

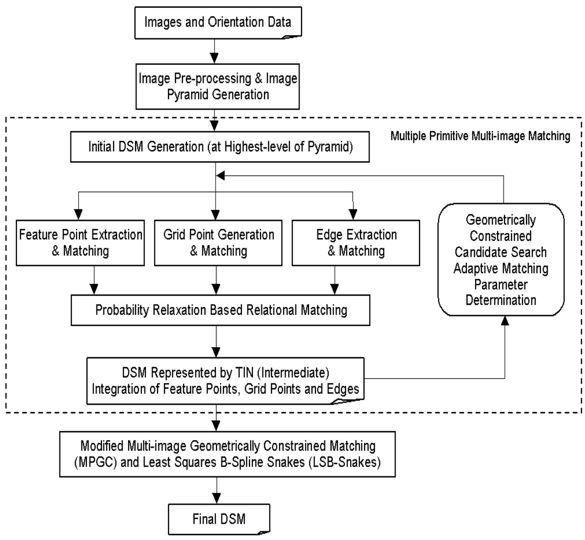

Figure 5: Workflow of the automated DSM generation approach. The approach consists of three mutually connected components: the image

pre-processing, the multiple primitive multi-image matching and the refined matching procedure.

For the south side of the temple, different groups of images are processed. The final DSM has a very dense regular grid (1.5 cm grid spacing) and the fact that the process uses an unrestricted number of images simultaneously (more than 2) led to reduction of problems like occlusions, multiple solutions etc.
Due to the complexity of the object, the generation of a surface model from the measured point cloud required some manual editing and additional measurement of some features, especially edges, using a commercial digital photogrammetric software. The automatic and manual measurements are then combined to generate a 3D triangular mesh. Finally, a textured 3D model is produced (
Figure 6), as well as virtual flights for visualization purposes.
The final 3D model presents some gaps in the geometry due to its complexity as well occlusions and untextured areas.

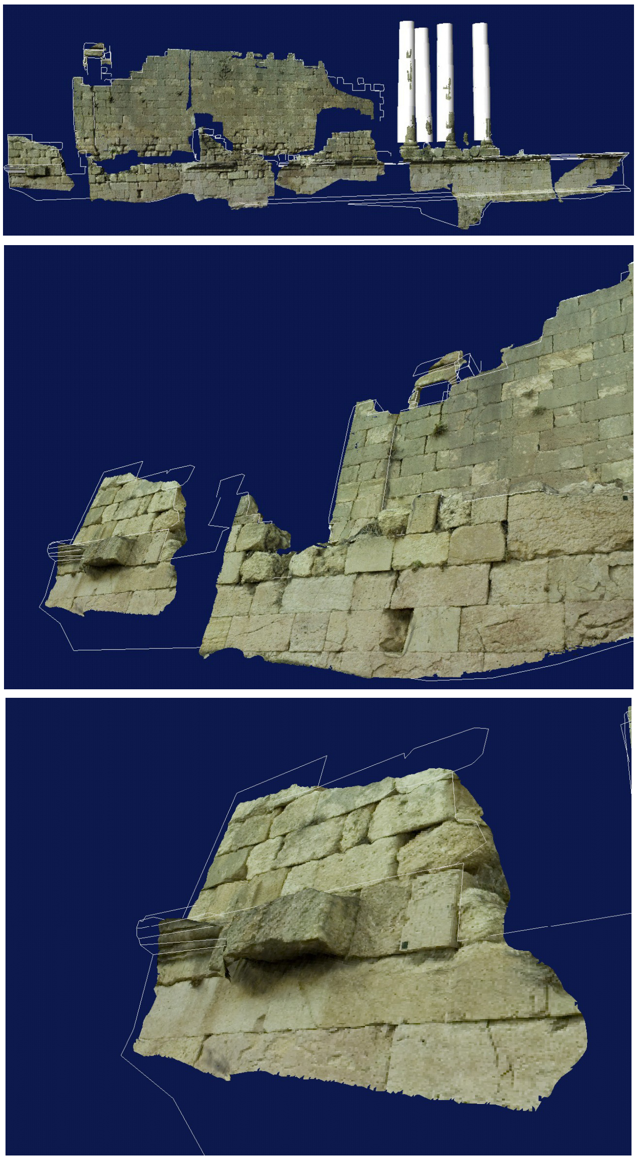

Figure 6: Some views of the textured 3D model of the south side of the Artemis temple. Most part of the surfaces are automatically reconstructed with a multi-photo surface matcher, expect the columns and some object edges which are manually measured with PhotoModeler.

3.4.1 **DSM Comparison.** A comparison between the results obtained with our in-house surface measurement program and a commercial photogrammetric station (Leica Photogrammetric Suite) was performed to asses a quantitative analysis of the 3D results. Using a stereo-pair, both algorithms are employed and afterwards the generated DSMs are compared within Geomagic^{TM} reverse engineering software (Figure 7). The average Euclidean distance between the two DSMs ranges between -10 and +12 cm.

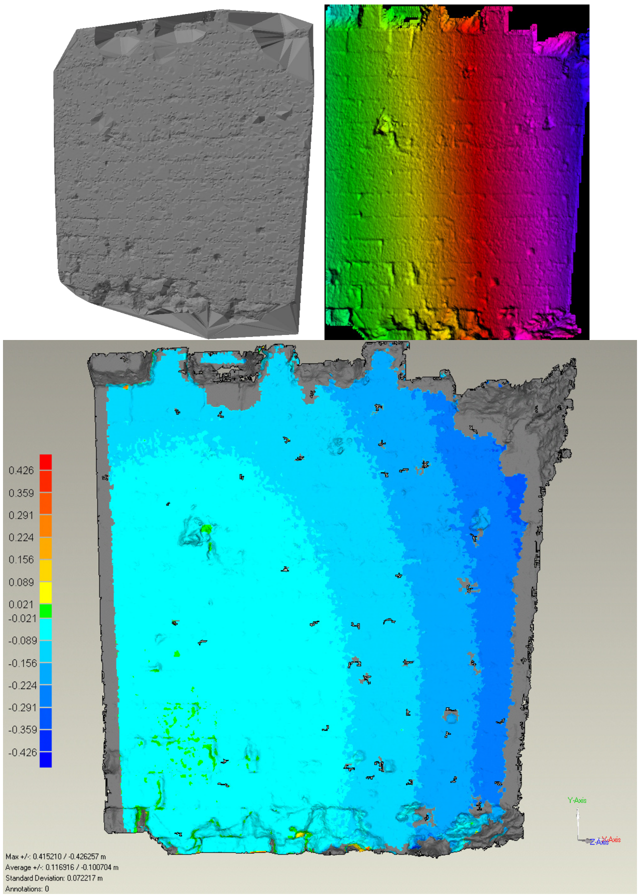

Figure 7: DSM generated using LPS and our in-house matcher (above, left and right) and 3D comparison between the two results (below).

4. CONCLUSIONS

The Artemis temple of Jerash is a complex and large structure archaeological object, which needs a good planning and data capture for a complete and detailed digital documentation. By applying digital technologies to archaeological sites and monuments, archaeologists can greatly improve the quality and usability of the surveyed and recorded objects.
In this contribution, we have shown an image-based approach, starting with the image acquisition and pre-processing, orientation, surface measurement and final 3D model generation.
For the surface generation, we employed our multi-image approach, as more reliable and accurate than simple stereo-pairs approaches. It requires quite precise image orientation parameters to fully exploit the collinearity constraint within the least squares matching estimation. The maximum orientation errors in image space should be less than 2-3 pixels. Providing few seed points in the regions where big discontinuity changes are present, the algorithms is able to derive a dense and detailed point cloud of the measured surface.

The two major contributions that separate this work from previous one on 3D reconstruction of cultural sites is (i) the sophisticated image pre-processing and (ii) the high-accuracy automated generation of a dense surface model. Although with such complex objects full automation still has a long way to go, it is shown that a certain degree of automation in the surface measurement can reduce production time and improve the results.
Alternative modeling technique, like active sensors (in particular laser scanner) would also have problems in the recording of this kind of objects. Probably a combination of image- and range-based techniques would help in providing a complete and detailed 3D model.
In the future we will test automated tie point extraction and try to complete the other sides of the temple with automated surface reconstruction methods.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to BAU that supplied hardware, software and logistics support for executing this work, the Swiss Government who provided funding to BAU to purchase hardware and software used in this work and to the Dept. of Antiquities, Jordan for kind permission to perform and support during the field work.

REFERENCES

Al Bayari, O., 2005: New Survey Technologies for Production of GIS Model of the Ancient Roman Jerash City in Jordan. Proc. CIPA XX International Symposium, Torino, Italy, Sept. 26-Oct. 1, 6 p. (on CD-ROM).

Al-Hanbali, N., Al Bayari, O., Saleh, B., Almasri, H., Baltsavias, E., 2006. Macro to Micro Archaeological Documentation: Building a 3D GIS Model for Jerash City and the Artemis Temple. Proc. International Workshop on 3D Geoinformation 2006, 7-8 August, Kuala Lumpur, Malaysia (in press).

Beardsley, P., Torr, P. and Zisserman, A., 1996: 3D model acquisition from extended image sequences. Proc. ECCV'96, Lecture Notes in Computer Sciences, Vol. 1065, pp. 683-695.

Cronk, S., Fraser, C.S. and Hanley, H., 2006: Automatic calibration of colour digital cameras. The Photogrammetric Record (in press).

Eisenbeiss, H., Lambers, K., Sauerbier, M., Zhang, L., 2005. Photogrammetric documentation of an archaeological site (Palpa, Peru) using an autonomous model helicopter. Proc. CIPA XX International Symposium, Torino, Italy, Sept. 26-Oct. 1, 7 p. (on CD-ROM).

Fitzgibbon, A. and Zisserman, A., 1998: Automatic 3D model acquisition and generation of new images from video sequences. Proc. European Signal Processing Conference, pp. 1261-1269.

Gruen, A. and Baltsavias, E., 1986: Adaptive least squares correlation with geometrical constraints. Proc. of SPIE, Vol. 595, pp. 72-82.

Hartley, R. and Zisserman, A., 2001: Multiple View Geometry in Computer Vision. Cambridge University Press, UK.

Kaufmann, V., Ladstädter, R., 2005. Elimination of color fringes in digital photographs caused by lateral chromatic aberration. Proc. CIPA XX International Symposium, Torino, Italy, Sept. 26-Oct. 1, 6 p. (on CD-ROM).

Läbe, T. and Förstner, W., 2006: Automatic relative orientation of images. Proc. 5th Turkish-German Joint Geodetic Days, 29-31 March, Berlin , ISBN 3-9809030-4-4.

Pollefeys, M., Koch, R. and Van Gool, L., 1999: Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. International Journal of Computer Vision, 32(1): 7-25.

Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. and Koch, R., 2004: Visual modeling with a hand-held camera. International Journal of Computer Vision, 59(3): 207-232.

Remondino, F. and El-Hakim, S., 2006: Image-based 3D modeling: a review. The Photogrammetric Record (in press).

Remondino, F. and Ressl, C., 2006: Overview and experiences in automated markerless image orientation. IAPRSSIS, Vol. 36, Part 3 (in press).

Remondino, F. and L. Zhang, 2006: Surface reconstruction algorithms for detailed close-range object modeling. IAPRSSIS, Vol. 36, Part 3 (in press).

Roncella, R., Remondino, F. and Forlani, G., 2005: Photogrammetric bridging of GPS outages in mobile mapping. Videometrics VIII, Beraldin/El-Hakim/Grün/Walton (Eds.), SPIE Electronic Imaging, Vol.5665, pp. 308-319.

Roth, G., 2004: Automatic correspondences for photogrammetric model building. IAPRSSIS, Vol. 35, Part B5, pp. 713-718.

Scharstein, D. and Szeliski, R., 2002: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47(1/2/3): 7-42.

Strecha, C., Tuytelaars, T. and Van Gool, L., 2003: Dense Matching of Multiple Wide-baseline Views. IEEE Proc. ICCV'03, Vol. 2, pp. 1194-1201.

Wallis, R., 1976: An approach to the space variant restoration and enhancement of images. Proc. of Symposium on Current Mathematical Problems in Image Science, Naval Postgraduate School, Monterey, CA.

Zhang, L., 2005: Automatic Digital Surface Model (DSM) Generation from Linear Array Images. PhD Dissertation, Report No. 88, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland. http://e-collection.ethbib.ethz.ch/ecol-pool/diss/fulltext/eth16078.pdf.