

THE ROMAN HEIDENTOR AS STUDY OBJECT TO COMPARE MOBILE LASER SCANNING DATA AND MULTI-VIEW IMAGE RECONSTRUCTION

N. Studnicka, C. Briese, G. Verhoeven, M. Kucera, G. Zach, C. Ressel

1. INTRODUCTION

Laser scanning (LS), which can be utilized in both a static mode (terrestrial LS or TLS) or kinematic mode (airborne LS (ALS) or mobile LS (MLS)), has established itself in the last decade as an appropriate approach to densely sample a whole variety of objects and scenes in 3D (Vosselman and Maas, 2010). Besides differences in their applications (ALS and MLS being more suited for large area mapping), the georeferencing of different LS approaches differs. TLS data are typically georeferenced by indirect techniques using control and tie points, while ALS and MLS require the direct georeferencing of the moving sensor coordinate system based on a global navigation satellite system (GNSS) receiver and an inertial navigation unit (INS). Obviously, these position and orientation (POS) data of the platform must be perfectly synchronised with the simultaneous LS observations and possible image data. Besides LS-based methods, research in computer vision and photogrammetry lead to advanced automated procedures in image orientation and image matching. An important impact for 3D reconstruction computer vision algorithms was gained by so-called structure-from-motion (SfM) algorithms, as they allow reconstructing 3D scene geometry and camera motion from a sequence of 2D imagery captured by a camera moving around the scene (Szeliski, 2010). To do this, the SfM algorithms use image matching to detect image feature points and subsequently monitor their movement throughout whole image collection. Using this information as input, the locations of those feature points can be estimated and rendered as a sparse 3D point cloud. As the SfM heavily depends on accurate knowledge of camera positions, estimating the latter is one of the core components in SfM (Hartley and Zisserman, 2001; Szeliski, 2010). Using the output of these SfM algorithms (i.e. the sparse point cloud and the camera positions, orientation, and calibration parameters) as input, multi-view stereo (MVS) reconstruction algorithms can generate very dense 3D geometry that present the majority of geometric scene details (Scharstein, 2002).

This paper focuses on the analysis of 3D geometry acquired by LS data and simultaneously acquired imagery from the same moving MLS platform. The Roman monument Heidentor, part of the archaeological site of Carnuntum (Austria), was selected as test object. This monument measures approximately 15 by 15 m and has a height of circa 14 m (Jobst, 2001).

2. THE MLS SYSTEM

For this comparison, the Heidentor was scanned with the RIEGL MLS system VMX-450 (see Figure 1). The hardware of the VMX-450 consists of two synchronously operated RIEGL VQ-450 laser scanners, a portable control unit box and a GNSS/INS-unit that comprises the electronics for real-time kinematic (RTK) measurements and three sensors. The modular VMX-450-CS6 camera system complements the acquisition of LS data with the



Figure 1: *Figure 1 . RIEGL MLS system VMX-450 at the Heidentor; Roman city of Carnuntum.*

recording of high-resolution (5 MP) colour images. Up to six individually selectable, fully calibrated industrial digital colour cameras with electronic shutters and 5 mm optics can be integrated. In conjunction with the known mounting parameters, precisely time-stamped position and orientation of the cameras are defined on an image-by-image basis. These still images can be used to colour the scan data, but are also the basis for the photogrammetric processing presented here (further key data of the complete MLS system can be found in Briese *et al.*, 2012). Data acquisition and processing

The MLS and optical still imagery of the Heidentor were acquired on the 29th of March 2012. The complete monument surface was sampled by the MLS with approximately 1 point/cm² (see Figure 2), while four cameras were triggered to acquire an image every 3 m. Overall, approximately 64.5 million laser points were recorded and subsequently processed with the RIEGL software RiProcess. The 1,156 images acquired images were available in the raw *.pgm (portable graymap) format. After a necessary debayering step (executed with RiProcess) the jpg-compressed imagery were subjected to additional post-processing (shadow brightening and sharpening), necessary to tackle the unfavourable illumination conditions during the data acquisition (cloudy sky with locally penetrating sunlight).

For the subsequent image orientation and surface model generation 165 images were selected and processed with PhotoScan (Agisoft). Using PhotoScan's SfM algorithm (cf. Doneus *et al.*, 2011; Verhoeven, 2011; Verhoeven *et al.*, 2012), the camera positions, orientation and calibration parameters were computed. Since highly accurate D-GNSS positional data were embedded as image metadata, the same coordinate frame could be defined for the SfM result as for the original MLS data. In the end, only 144 photographs were used for the next processing steps, since some images could not be matched while others had overly large positional errors.

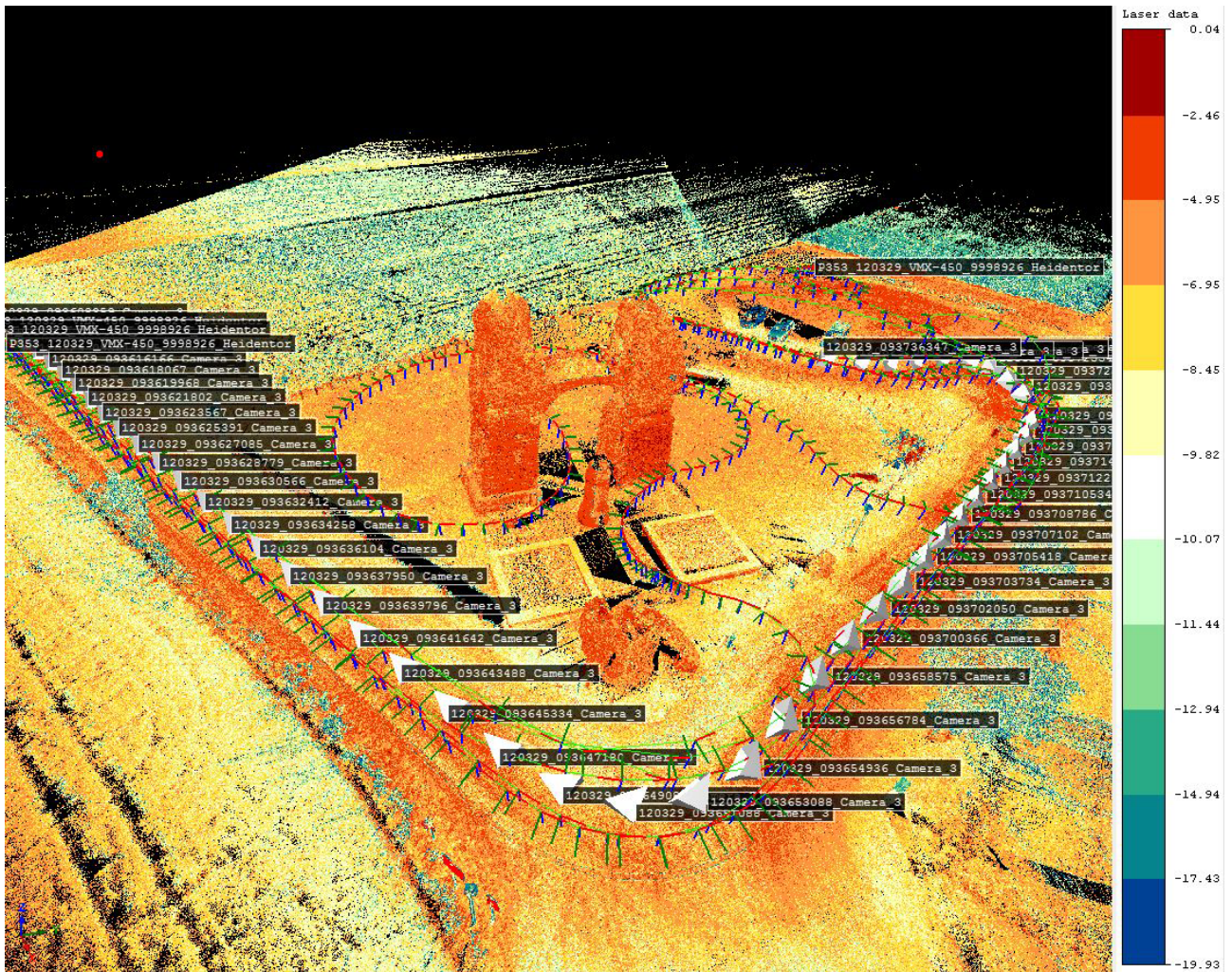


Figure 2: MLS-sampled point cloud of the complete monument surface.

Afterwards, PhotoScan complemented the SfM approach with a dense MVS algorithm to compute the surface geometry of the observed scene. Although the image stations are accurately known in the MLS coordinate reference system, the output of the 3D PhotoScan mesh in the MLS reference frames failed. Hence, the final 3D model needed to be re-georeferenced in a local reference frame using four ground control points extracted from the MLS point cloud. After deleting some noise, the vertices extracted from the mesh represented the final point cloud resulting from the image matching step.

3. EVALUATION

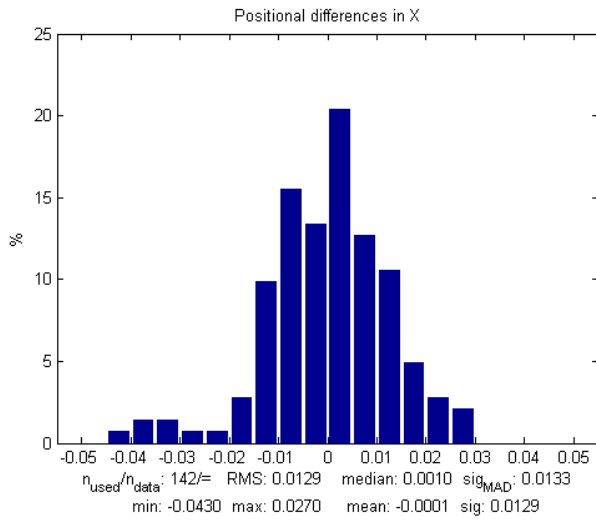
On the one hand, the images were georeferenced using the GNSS/INS system on board the RIEGL MLS system. On the other hand 144 images were successfully oriented and georeferenced using an SfM approach in PhotoScan. However, further analysis revealed that two images had significantly big positional and orientation errors compared to the direct georeferencing result. For the subsequent steps, these two images were removed and the analysis is just based on the remaining 142 images.

Figure 3 presents histogram visualisations of the differences between direct and indirect georeferencing of the images. While the positional differences do not show significant systematic

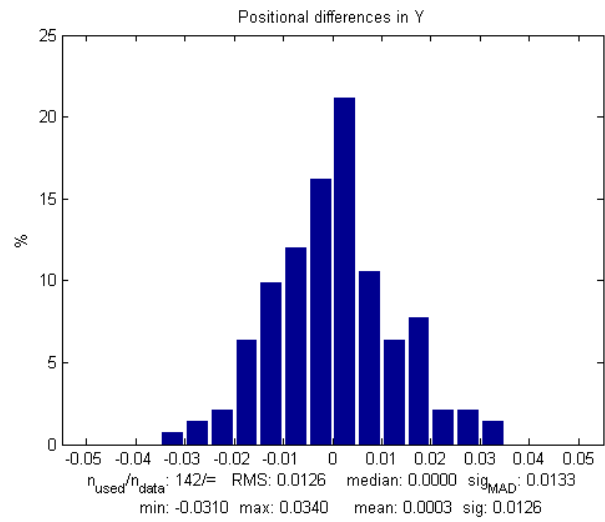
errors and the standard deviations of the differences (approx. 0.01 m) correspond well to the accuracies from the bundle block adjustment, the differences for the orientation angles show a small systematic error for roll and nick, but a significant median difference for the yaw component of more than 1° . The corresponding standard deviations are all higher than 1° . Analysis of the histograms showed that these high values for the standard deviation are introduced by a few gross errors. By calculating a robust measure for the standard deviation (sigma mad (median absolute deviation)) these values could be significantly reduced (more specifically from 1.467° , 1.002° , and 1.132° to 0.271° , 0.175° , and 0.203° respectively).

For the evaluation of the two surface models, one of the main facades of the Roman monument was selected for the evaluation. All LS scan data of this facade were transformed into a 2.5D facade coordinate system, yielding 6.9 million points. From those points, a 2.5D surface model was generated by a moving planes interpolation with the software OPALS. The result of the dense surface matching of the selected facade resulted in 6.4 million points. Although this result is quite similar to the LS data set, the density on the left part of the facade is significantly less than on the right part and some structural details seem to be a little bit more smoothed when compared to the LS model.

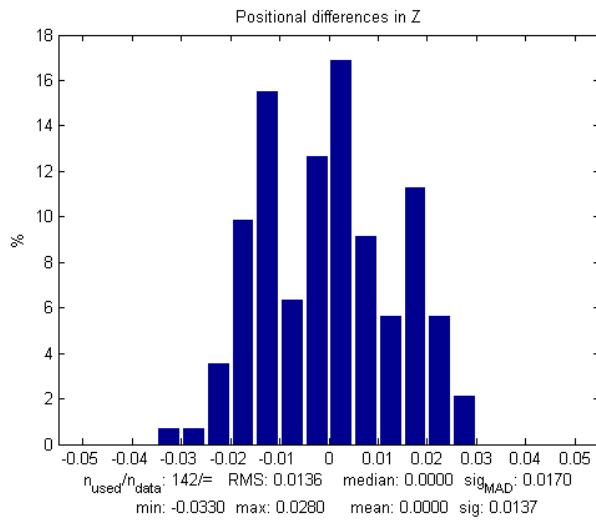
In order to compare both results analytically a difference



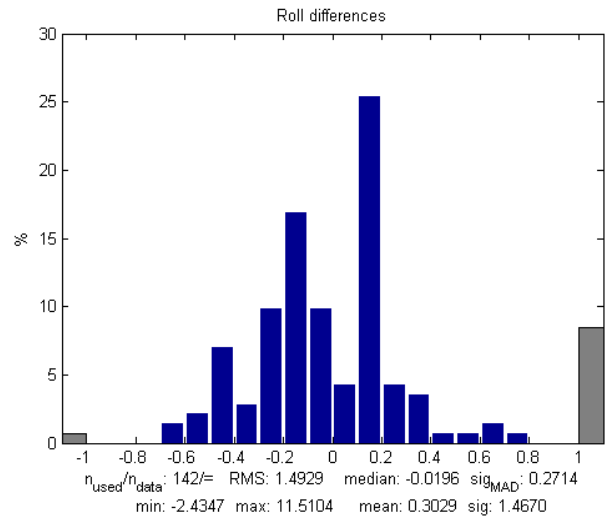
(A) X coordinate



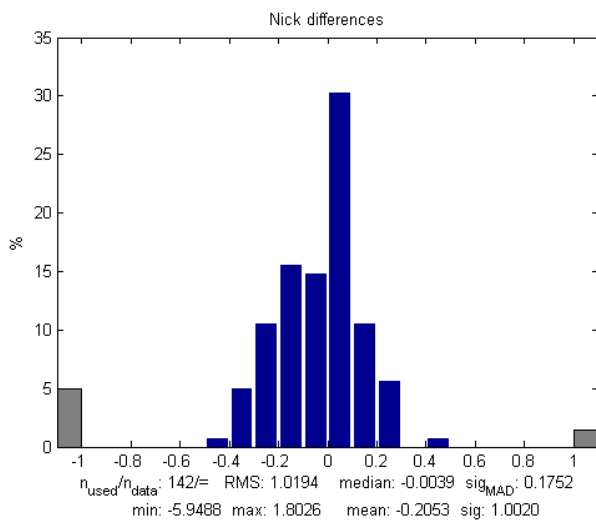
(B) Y coordinate



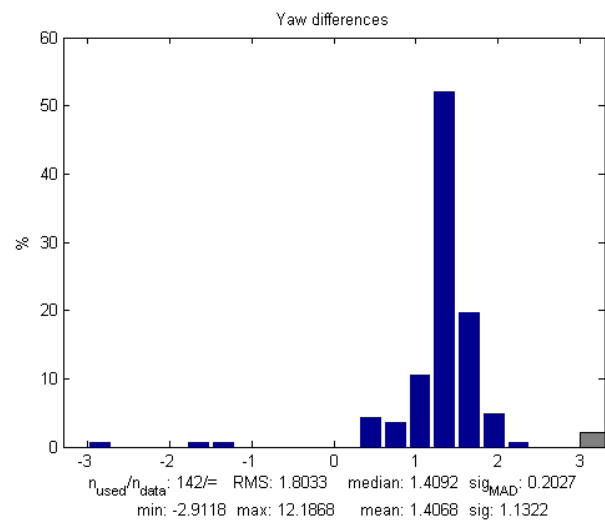
(C) Z coordinate



(D) Roll angle



(E) Nick angle



(F) Yaw angle

Figure 3: Histogram visualisations of the differences between direct and indirect georeferencing of LS images.

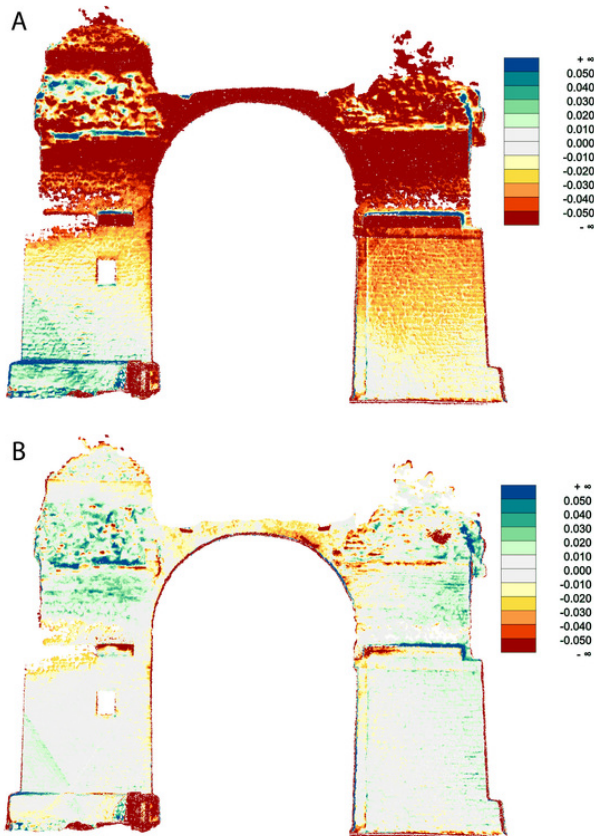


Figure 4: Facade of the Heidentor created from LS data.

model was calculated with OPALS. The sigma mad of the differences is about 0.04 m. In general, the differences are significantly lower for the lower part of the facade (Figure 4A). The bigger differences in the upper part seem to be mainly caused by a vertical tilt between the two models, which might be introduced by the necessary realignment step in PhotoScan. After a least squares adjustment (LSM) of the two models, the resulting difference model (displayed in Figure 4B) indicated an improved value for sigma mad (from 0.04m to 0.01m) and most of the systematic differences in the upper part of the monument could be reduced significantly. The remaining differences are mainly caused by the different surface smoothing (stronger in the image-based model), by the different view geometry and on sharp edges.

4. CONCLUSION

This paper presents a study on the comparison of simultaneously acquired LS data and image data from an MLS platform. The results indicate individual strength and weaknesses. While the indirect georeferencing worked quite well for most of the images that represented sufficient image content, some imagery definitely needs external POS data to achieve an accurate georeferencing (especially for images with homogeneous background

with similar features). However, in the case of POS errors or the absence of sufficient GNSS signals, automated image orientation based on SfM algorithms might allow the determination of the position and orientation of the moving platform. However, image arrangement is of the utmost importance to fulfil this aim. While the image orientation and surface matching works typically quite well for images that are specifically acquired for a good geometric reconstruction, the surface matching based on constantly triggered cameras delivered sub-optimal results. They do not consider nor adapt to the actual shape of the observed object of interest, hence causing much more problems and reducing the degree of automatization.

ACKNOWLEDGEMENTS

The Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (archpro.lbg.ac.at) is based on an international cooperation of the Ludwig Boltzmann Gesellschaft (A), the University of Vienna (A), the Vienna University of Technology (A), the Austrian Central Institute for Meteorology and Geodynamic (A), the office of the provincial government of Lower Austria (A), Airborne Technologies GmbH (A), RGZM-Roman-Germanic Central Museum Mainz (D), RAÄ-Swedish National Heritage Board (S), IBM VISTA-University of Birmingham (GB) and NIKU-Norwegian Institute for Cultural Heritage Research (N).

REFERENCES

- BRIESE C, ZACH G, VERHOEVEN GJJ, *et al.* 2012. Analysis of mobile laser scanning data and multi-view image reconstruction. In *Proceedings of the XXII ISPRS Congress*, Shortis MR, Mills J (eds.). Technical Commission V. ISPRS: 163-168.
- DONEUS M, VERHOEVEN G, FERA M, BRIESE C, KUCERA M, NEUBAUER W. 2011. From deposit to point cloud a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. *Geoinformatics (XXIIIrd International CIPA Symposium)* 6: 81-88.
- HARTLEY R, ZISSERMAN A. 2001. *Multiple View Geometry in Computer Vision*. Cambridge University Press: Cambridge.
- JOBST W. 2001. Das Heidentor von Carnuntum. Ein spätantikes Triumphalmonument am Donaulimes. Verlag AW, Wien.
- SCHARSTEIN D, SZELISKI, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47: 7-42.
- SZELISKI R. 2010. Computer vision: Algorithms and applications. Springer: London.
- Verhoeven G. 2011. Taking computer vision aloft archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeological Prospection* 18: 67-73.
- VERHOEVEN G, DONEUS M, BRIESE C, VERMEULEN F. 2012. Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science* 39: 2060-2070.
- VOSSELMAN G, MAAS H-G (eds.). 2010. *Airborne and terrestrial laser scanning*. Whittles Publishing: Dunbeath.