

3D Reconstruction of Archaeological Trenches from Photographs

Robert Wulff, Anne Sedlazeck, Reinhard Koch

Abstract This paper presents a method for 3D reconstructions of archaeological excavation sites. The method extends a 3D reconstruction algorithm for general rigid scenes to better fit the special archaeological needs and to integrate easily into the documentation process. As input, an ordered image sequence captured with a calibrated standard digital camera is required, along with a small set of 3D points from the trench with well-known coordinates. The 3D points are used to transform the model into the world coordinate system used at the excavation site, so measuring in the model and fusing it with other models becomes possible. Furthermore, a new algorithm called *LoopClosing* is introduced to minimize drift and increase accuracy. The resulting models provide lasting 3D representations of the trenches and allow the user to explore the scene interactively, not being restricted to a photographer's point of view. True orthographic views can be generated from the 3D models that can be correlated with other archaeological data.

1 Introduction

When working in archaeological excavations, the configuration of finds and features needs to be well-documented. A lot of techniques are used in the documentation procedure including drawings, measuring, photogrammetry, photographs, and CAD drawings—most of them being very time-consuming. This extensive documentation mainly serves the purpose of retaining representations of the configuration for later research because the configuration is usually destroyed when the next layer in a trench is unveiled.

Robert Wulff, Anne Sedlazeck, Reinhard Koch
Multimedia Information Processing Group, Department of Computer Science,
Christian Albrechts University of Kiel, Germany,
e-mail: {rwulff, sedlazeck, rk}@mip.informatik.uni-kiel.de

We therefore propose the computation of digital 3D models of a trench by extracting the implicitly contained geometric properties of the scene from a sequence of images. This is achieved by adapting an existing algorithm for 3D reconstruction of general rigid scenes (Pollefeys et al. [12]) to meet the special needs of archaeologists. The resulting 3D models offer an intuitive way for visualizing the configuration interactively and hence can help in retrospective interpretations of the finds and features. Another advantage is the ability to create real orthographic views from any direction. So far, such views are approximated by rectifying a perspective image. However, this rectification has its limitations due to occlusions and protruding objects. In addition, the models allow for measuring and offer new possibilities for public presentations in museums, lectures, talks, and multimedia applications.

The idea of reconstructing 3D models and even their usage in archaeology is not new. Several different methods exist for general scenes and provide a basis for further research. Among them are the works of Hartley and Zisserman [6] and Pollefeys et al. [12]. Both are solely based on images sequences.

Some existing systems focusing on archaeology or architecture are 3D Murale [2], 3D-Arch [14], ARC3D [17], and the works [13], [16], [19], and [9]. The focus of 3D Murale and ARCH3D is broader than the one in this work. Image-based 3D reconstruction is only one method for acquiring data. In addition, laser scanners are used and data integration is of interest. ARC3D is a web-based service computing 3D models for users who upload their data. However, the resulting models do not offer measuring capabilities, so it is difficult to combine different excavation layers. The designated use of the methods introduced in [16] and [19] is the reconstruction of small finds, so they are not applicable to reconstructing trenches. A method for large-scale reconstruction was suggested in [9]. This approach is based on laser scanners, and detailed models can be achieved. The two main drawbacks of this method are the required expensive equipment and the time-consuming data acquisition.

In contrast to the existing systems, the goal of this project is to adapt an existing method to integrate well into the documentation procedure and to meet the special needs in archaeology, e. g. measuring. The algorithm is based on the work of Pollefeys et al. [13] and does not require any extra equipment besides a standard digital camera. However, the intrinsic parameters of the camera need to be known. They can be acquired by a camera calibration method [1]. In contrast to [13], SIFT keypoints [10] are used because they provide invariance against changes in lighting, rotation, and scale. During the documentation procedure, usually a set of 3D points is surveyed by photogrammetric methods to compute the above mentioned orthographic views. We call these points *photogrammetry points* and reuse them here to transform the model into the world coordinate system used at the excavation site. Besides yielding absolute scale and hence allowing measuring, it also enables the correlation of models from different layers or other models within the same coordinate frame. For improved accuracy and robustness, a loop closing procedure is applied in case of an orbital camera path.

This work is structured as follows. In the next section the data acquisition process is described. Afterwards, the description of the reconstruction algorithm is given,

with special emphasis on the loop closing procedure and the transformation based on the photogrammetry points. The results are presented in the experiments section followed by the conclusion.

2 Data Acquisition

The proposed method requires an ordered sequence of photographs, taken with a standard digital camera with known intrinsic parameters. These parameters include focal length, principal point, aspect ratio, and radial distortion coefficients. They need to be kept constant for the whole image sequence. The calibration can be performed with the free implementation in [1], which is based on [7] and [18]. As long as the same intrinsic parameters are used, the calibration photographs can be taken before or even after the excavation.

The most important requirement for the input sequence is that consecutive images overlap by a large proportion (about 80%). This requirement is necessary to ensure stable keypoint matching. In order to compute the absolute scale of the model, the markers of the photogrammetry points need to be visible in the images. If the camera is moved in an orbit around the trench so that the first and last image overlap as well, an optional loop closing procedure can be applied. This procedure exploits the known camera path to reduce the reprojection errors. Other assumptions for the algorithm include a rigid scene, no reflections, and a relatively constant brightness.

3 Reconstruction Process

We extend the structure-from-motion approach for general rigid scenes described in [12] to better fit the archaeological needs. The single steps are visualized in fig. 1.

First, the input images need to be prepared for further processing by converting them to gray scale and compensating lens distortion.

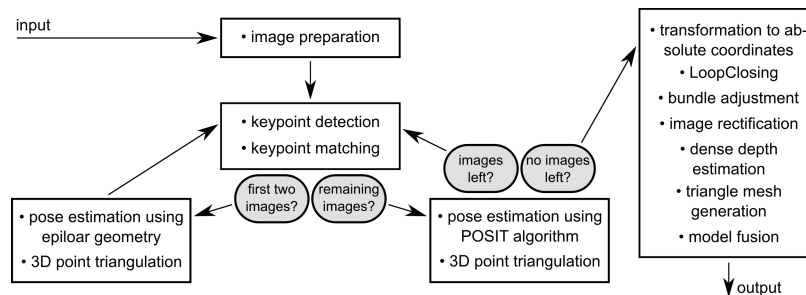


Fig. 1 Flowchart of the reconstruction process.

The next step is to detect keypoints in each image automatically. Since the camera’s orientation changes relative to the ground throughout the image sequence, rotation invariant keypoints are needed—we use the SIFT keypoints described in [10]. The keypoints then have to be matched to establish 2D–2D correspondences between each successive image pair. If the viewports overlap enough, more stable results can be achieved by matching each image with its *two* predecessors. To improve the performance of the keypoint similarity evaluation, it is sufficient to restrict the actual computation on a neighborhood around the current keypoint, e. g. 20 % of the image’s width and height, respectively. Note that this technique also reduces the number of outliers.

The configuration of the scene’s geometry is initialized using epipolar geometry on the first two cameras of the sequence. Since the camera’s intrinsics are known, the essential matrix can be used, so the reconstruction is performed in a metric frame (see [6] for more details). To solve for the essential matrix, we use [11] combined with a RanSaC approach to deal with outliers [4]. At this point the reconstruction is initialized such that the first camera is aligned with the coordinate system. The poses of the remaining cameras are determined using the *POSIT* algorithm [3], which needs 2D–3D correspondences. The required 3D points are triangulated (see [5]).

3.1 The LoopClosing Algorithm

In this scenario it is likely that the camera was moved in an orbit around the trench. This implies that the first and the last camera share a large proportion of their viewports. If this is the case, attaching the first image again at the end of the image sequence enables us to perform the keypoint matching and pose estimation between these cameras as well.

Let n denote the number of input images, $n + 1$ the index of the attached camera, c_i the position of camera i and q_i the orientation of camera i in quaternion representation, where $0 \leq i \leq (n + 1)$. Since the first camera is aligned with the coordinate system, its position is $c_1 = (0, 0, 0)^T$ and its orientation is $q_1 = (0, 0, 0, 1)$. Ideally, $c_1 = c_{n+1}$ and $q_1 = q_{n+1}$ hold, but in practice errors in camera calibration, measuring, and rounding will lead to a discrepancy between these values. The *LoopClosing* algorithm distributes these discrepancies between all cameras according to a weighting function so that the poses of the first and the attached camera will match perfectly. Furthermore, the reprojection error is minimized—see sect. 4 for results.

The first step is to compute the discrepancies. For the position, the difference vector is given by $\Delta_c := c_1 - c_{n+1}$. Since $c_1 = (0, 0, 0)^T$ this simplifies to $\Delta_c = -c_{n+1}$.

Using the quaternion representation, the discrepancy in the orientation is given as the conjugate of the quaternion of the orientation of camera $n + 1$, i. e. $\Delta_q := \bar{q}_{n+1} = (q_1, -q_2, -q_3, -q_4)$, where $(q_1, q_2, q_3, q_4) = q_{n+1}$.

In the second step, a weighting function $w : \{1, \dots, n + 1\} \rightarrow [0, 1]$ is computed. Under the reasonable assumption that the discrepancies were accumulated over the

sequence and grow with an increasing number of images, we choose this function so that the following conditions are met:

- $w(1) = 0$, i. e. the first camera shall not be transformed at all
- $w(n+1) = 1$, i. e. the attached camera shall be fully transformed so that the poses of the first and the attached camera are equal
- $w(i) < w(j)$ for all $i < j$, i. e. the front cameras are transformed less than the latter ones

The weights are computed recursively over the distance of each camera to the first camera along the camera path: let $L_1 := 0$ and $L_i := L_{i-1} + d(c_i, c_{i-1})$, for $1 < i \leq n+1$, where $d(.,.)$ is the Euclidean distance. The total length of the camera path is given by $L := \sum_{i=1}^{n+1} d(c_i, c_{i-1}) = L_{n+1}$. Now we can define the weighting function $w(i) := L_i/L$, for all $1 \leq i \leq n+1$, which meets the above mentioned conditions.

The third step is to transform the cameras. The location of camera i is replaced by its adjusted position $c_i + w(i)\Delta_c$. For the orientation, the computation is slightly more complex. The angle of the rotation represented by the quaternion $\Delta_q = (\Delta_{q,1}, \Delta_{q,2}, \Delta_{q,3}, \Delta_{q,4})$ is given by $\alpha := 2\cos^{-1}(\Delta_{q,1})$ and the axis can be computed as $a := (1/\sin(\alpha/2))(\Delta_{q,2}, \Delta_{q,3}, \Delta_{q,4})^T$. Using the weight $w(i)$, we compute the adjustment angle $\kappa_i := w(i)\alpha$ for camera i and obtain the quaternion $r_i := (\cos(\kappa_i/2), \sin(\kappa_i/2)a)$. The orientation of camera i is then replaced by its adjusted orientation $r_i q_i$.

Finally, the camera path is closed by correlating the correspondences of the keypoints of camera $n+1$ with the keypoints of camera 1. Afterwards, camera $n+1$ can be removed from the sequence.

The loop closing procedure is concluded by a global bundle adjustment (refer to [15]). Note that bundle adjustment is a mandatory step for the LoopClosing algorithm to perform well because the cameras were moved with respect to the 3D points.

3.2 Absolute Position, Orientation and Scale

After the loop closing procedure, the scene is transformed into the global coordinate system that was used at the excavation site. This coordinate system is clearly determined by the photogrammetry points. For transforming the scene, we use algorithm [8]. This algorithm requires a set of 3D–3D correspondences between the source and the destination coordinate system. The 3D points in the source coordinate system can be computed by triangulating the 2D projections of the 3D world points, which are in turn identified manually.

3.3 Model Generation

Up to this point, our algorithm estimated the poses of the cameras and a sparse cloud of 3D points. In order to compute a detailed 3D model from this data, an image rectification and dense depth estimation as described in [12] is applied. The dense depth estimation yields depth maps, containing the distance between camera center and a 3D point for each pixel. This information is utilized to generate triangle meshes. These triangle meshes have to be combined to compensate for occlusions and to fill areas where the dense depth estimation in single images did not succeed.

4 Experiments

The algorithm was tested on synthetic and real data. A synthetic scene was used to evaluate the keypoint matching and the LoopClosing algorithm under the influence of camera calibrations with a growing systematic error. In contrast to [13], where only the directly preceding image was used for keypoint matching, the performance is much better when matching keypoints across two preceding images instead of one. In fig. 2 (left) the x-axis shows runs with increasing systematic errors on the camera calibration or more precisely on focal length, principal point, and radial distortion, while the y-axis shows the Euclidean distance between the first and the appended camera (which equals zero if no drift occurs—see sect. 3.1). Clearly, the usage of two preceding images instead of one helps to minimize the accumulation of drift. The reason is to be found in the more robust triangulation due to longer baselines.

The LoopClosing algorithm spreads the accumulated error due to drift across all cameras, and the global bundle adjustment restores the consistency between 3D

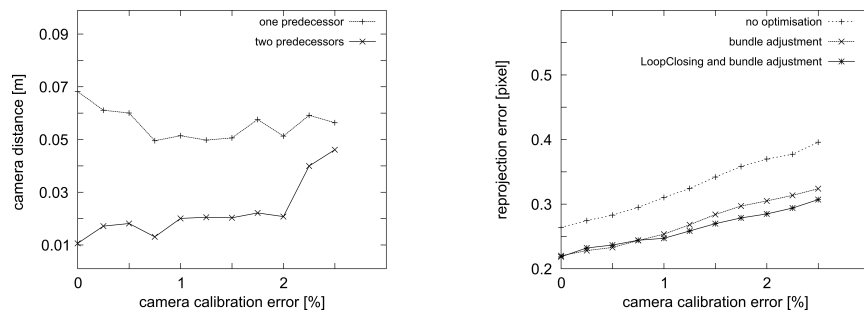


Fig. 2 Left: drift depending on matching with one or two predecessors. The x-axis denotes an increasing error in the camera calibration: compared to the true intrinsic camera parameters, a systematic error was added to focal length, principal point, and radial distortion. Right: reprojection error with and without the LoopClosing procedure with the same systematic error on camera calibration as on the left.

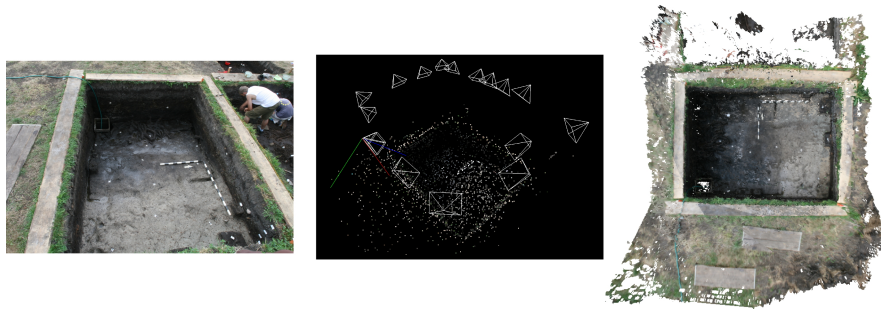


Fig. 3 Bruszczewo scene—left: input image, middle: camera path and 3D points, right: resulting model

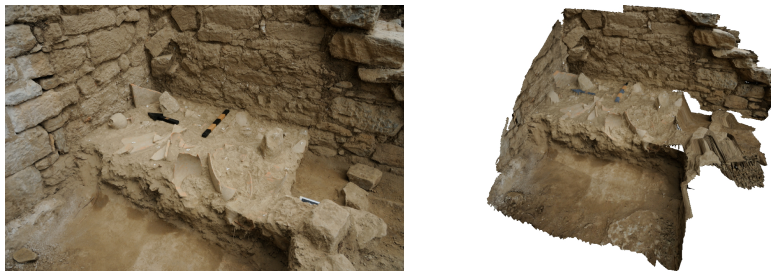


Fig. 4 Priene scene—left: input image (provided by the Department of Classics, Kiel University), right: resulting model

points and cameras. The success of this method was measured by comparing the reprojection errors of the whole sequence. The results are shown in fig. 2 (right). Especially with increasing systematic error on the camera calibration, the Loop-Closing algorithm reduces the reprojection error.

The algorithm performs well on real data and was tested with two different scenes. The first one (fig. 3) was recorded in Bruszczewo, Poland, while the second one (fig. 4) was captured in Priene, Turkey. For both scenes, the algorithm delivered detailed 3D models, which have absolute scale and can be used for measuring.

5 Conclusions

This paper presented a method for the 3D reconstruction of archaeological excavation sites. Since the models have absolute position, orientation, and scale, measuring in the model and fusing different models becomes possible. In the future, we hope to use exposure bracketing to benefit in high contrast lighting situations, e. g. in deep trenches. Other goals are a more sophisticated method for fusing several triangle meshes into one model and the automatic detection of the photogrammetry points, if special markers are used.

Acknowledgements The authors would like to thank Dr. Jutta Kneisel of the Institute of Prehistoric and Protohistoric Archaeology at the University of Kiel for the possibility to gain insight into the archaeological work and to take images at excavation sites. In addition, the authors would like to thank Prof. Rumscheid and his staff of the Department of Classics at the Kiel University for providing images of the excavation in Priene.

References

1. *OpenCV (Open Computer Vision Library)*. Intel Corporation, 2008. <http://www.opencv.org/>.
2. J. Cosmas, T. Itagaki, D. Green, E. Grabczewski, F. Weimer, L. J. Van Gool, A. Zalesny, D. Vanrintel, F. Leberl, M. Grabner, K. Schindler, K. F. Karner, M. Gervautz, S. Hynst, M. Waelkens, M. Pollefeys, R. DeGeest, R. Sablatnig, and M. Kampel. 3d murale: A multimedia system for archaeology. In *Virtual Reality, Archeology and Cultural Heritage*, pages 297–306, 2001.
3. D. F. DeMenthon and L. S. Davis. Model-based object pose in 25 lines of code. In *ECCV*, pages 335–343, 1992.
4. M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. of the ACM*, 24(6):381–395, June 1981.
5. R. Hartley and P. Sturm. Triangulation. *CVIU*, 68(2):146–157, November 1997.
6. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2004.
7. J. Heikkilä and O. Silvén. A four-step camera calibration procedure with implicit image correction. In *CVPR*, pages 1106–1112, 1997.
8. B. K. Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA*, 4(4):629 et seq., April 1987.
9. M. Ioannides and A. Wehr. 3d-reconstruction and re-production in archaeology. *Proc. of the Int. Workshop on Scanning for Cultural Herit. Rec.*, 2002.
10. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
11. D. Nistér. An efficient solution to the five-point relative pose problem. *PAMI*, 26(6):756–777, 2004.
12. M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual modeling with a hand-held camera. *IJCV*, 59(3):207–232, 2004.
13. M. Pollefeys, M. Vergauwen, K. Cornelis, F. Verbiest, J. Schouteden, J. Tops, and L. Van Gool. 3d acquisition of archaeological heritage from images. In *Proc. of the CIPA conference, Int. Arch. of Photogramm. and Remote Sens.*, 2001.
14. F. Remondino, S. El Hakim, S. Girardi, A. Rizzi, S. Benedetti, and L. Gonzo. 3d virtual reconstruction and visualization of complex architectures: The 3d-arch project. In *3DARCH09*, 2009.
15. B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment—a modern synthesis. *LNCS*, 2000.
16. V. Tsioukas, P. Patias, and P. F. Jacobs. A novel system for the 3d reconstruction of small archaeological objects. *ISPRS (Comm. V)*, 2004.
17. M. Vergauwen and L. J. Van Gool. Web-based 3d reconstruction service. *MVA*, 17(6):411–426, December 2006.
18. Z Zhang. A flexible new technique for camera calibration. *PAMI*, 22(11):1330–1334, 2000.
19. J. Zheng, W. Yuan, and S. QingHong. Automatic reconstruction for small archeology based on close-range photogrammetry. *ISPRS (Comm. V)*, 2008.