FAST ON-SITE RECONSTRUCTION AND VISUALIZATION OF ARCHAEOLOGICAL FINDS

Konrad Schindler, Markus Grabner, Franz Leberl
Computer Graphics and Vision
Graz University of Technology
Inffeldgasse 16, A-8010 Graz, Austria
{schindl,grabner,leberl}@icg.tu-graz.ac.at

KEY WORDS: Archaeology, Excavation, Close-Range Photogrammetry, Automatic Reconstruction, Interactive Visualization

ABSTRACT

We report on a study carried out within the 3D MURALE project, in which the complete workflow necessary for automatic recording, reconstruction and interactive visualization of artefacts on archaeological excavation sites has been tested. With state-of-the-art technologies from digital photogrammetry, computer vision and computer graphics a prototypical pipeline for recording, 3D modeling and visualization of archaeological artefacts has been implemented and tried out on site on the excavation of Sagalassos (Turkey). The purpose of the presented work was to test the complete pipe-line in a real archaeologic environment and assess its usability.

1 INTRODUCTION

To meet the practical requirements of archaeological heritage recording and documentation, technology must be usable under field conditions on archaeological excavation sites. Besides being fast and easy to use, systems for photogrammetric recording and modeling as well as visualization should interfere as little as possible with the excavation work and use only equipment that can be found on archaeological excavations.

We present a system which has been designed for the purpose of fast on-site recording, reconstruction and local as well as remote visualization of archaeological finds. The system combines methods from photogrammetry and computer vision with state-of-the-art computer graphics algorithms to accomplish the task. Objects are recorded directly on site with an off-the-shelf digital camera and are reconstructed in an almost fully automated photogrammetric modeling process, in which the only user interaction is to segment the artefact to be modeled from the background. Robust algorithms for orientation, adjustment and dense object reconstruction are applied, which deliver accurate, textured 3D object models.

New methods for compressed adaptive multi-resolution encoding of large models are used to enable visualization of the resulting 3D models at interactive frame-rates even on computers with inexpensive graphics hardware. An extension of this technology allows smooth remote interaction with the models via low-bandwidth connection using a standard Internet browser. This gives archaeologists in the field the possibility to instantly communicate about new finds with people who are not present on site. In contrast to existing simplification techniques object identities and attributes are preserved through the dynamic simplification process to provide a practically useful way of interaction. Again the processing is fully automatic.

The proposed system has been successfully tested within the 3D MURALE project (Cosmas et al., 2001) on the Greco-Raman excavation site of Sagalassos (Turkey) during the 2002 excavation campaign. Several artefacts have been recorded without any impact on the excavation routine, reconstructed on the same day using a standard laptop computer, and visualized on a different computer via a local network connection. We will present the utilized workflow and technology with real examples from Sagalassos.

2 RECORDING

A classical photogrammetric approach is used for on-site recording. Multiple overlapping images of the object to be modeled are recorded with a calibrated digital consumer camera. In our experiments we have used a high-quality SLR camera as well as a cheap consumer camera (a Canon D30 with a resolution of 2160×1440 Pixels and a Nikon Coolpix 995 with 2048×1536 Pixels). The standard setup with fixed focus close to infinity has proven to yield sharp images in most practical situations. We therefore use it, because it provides the best accuracy and robustness. It should be mentioned, however, that within the 3D MURALE project it has been shown that recording and reconstruction on archaeological sites is also possible with a completely uncalibrated digital video camera (Pollefeys et al., 2001).

2.1 Camera Calibration

Camera calibration is fully automatic. The user takes 4–6 pictures of a planar calibration target from arbitrary, different positions. The calibration target consists of a grid of circles with known positions and has a unique, easily detectable marker (a black ring) in the center, which allows to aoutomatically establish the correspondence between image points and target points. The calibration target is depicted in figure 1. With a calibration algorithm optimized for this configuration (Heikkilä, 2000) a full set of camera parameters is recovered, i.e. the principal point, the focal length, two coefficients for a symmetric 4^{th}-order radial distortion polynome and two coefficients for a similar tangential distortion polynome. Calibration is an offline process, which needs to be carried out only once for a given, stable camera geometry, i.e. it is sufficient to calibrate a camera (with each lens that should be used) once before the campaign, unless it is subject to unexpected physical impact.

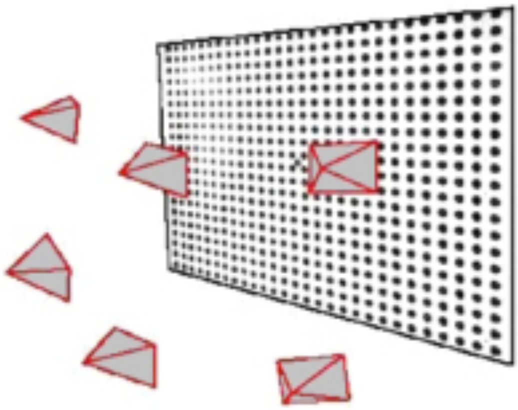


Figure 1: Calibration target for camera calibration. The red pyramids are the camera positions from an exemplary calibration. In the center of the target one can see the marker for automatic registration of the circles.

2.2 Data Capture

The actual recording consists of taking hand-held pictures of the object, as it is nowadays done for any close-range application. In order to provide sufficient redundancy for reliable automatic reconstruction it is however necessary to make sure that all relevant parts of the object are visible in at least 3, better 4 images (rather than the theoretical minimum of 2 for reconstruction from calibrated views). Furthermore we recommend an overlap of 80% (rather than 60%), which helps to ensure good image matching results at no extra cost (except for longer computation times). Figure 2 shows an example of a recording sequence captured in Sagalassos.

3 RECONSTRUCTION

3.1 Preprocessing

Before starting the actual reconstruction process lens distortion is corrected, i.e. the images are warped to the ideal pinhole camera geometry with the distortion parameters determined through camera calibration. Furthermore the user interactively has to segment the object to be modeled from the rest of the image by digitizing its outline. The segmentation is the only user interaction required for 3D reconstruction apart from defining the sequence (see next section) – it cannot be automated, because it is a decision of the user, which cannot be derived from the scene geometry alone. Note however that digitizing only needs a few mouse-clicks, if it is assisted by the life-wire method, as it is implemented in any image-processing toolkit (see Figure 3.

3.2 Matching and Orientation

Before homologous points are extracted for image orientation, the images are partitioned into regular tiles to ensure a good coverage of the object for orientation. We

(a)

(b)

Figure 2: Recording for automatic reconstruction. (a) Four images of a piece of the 'dancing girls frieze' from the Augustan 'Northwest Heroon' in Sagalassos. (b) 3-dimensional view of the recording setup.

use a partitioning of 8×6 tiles (which leads to roughly squared tiles for standard cameras). A number of homologous points is extracted for each tile with a hierarchical area-based matcher (Klaus et al., 2002). The same points are used for a set of 5 consecutive frames of the sequence to achieve a stable image block. In our experiments we have manually ordered the images (see Figure 4). In principle it is possible to match all possible pairs of images and determine the ordering automatically using the homologous points, if one is willing to accept the considerably higher computational cost.

With the homologous points an initial estimate for the relative orientations is sequentially recovered. Since the automatically matched correspondences can contain a large percentage of false matches (in some experiments up to 30%), it is necessary to use a robust estimation method. The standard algorithm for the task is RANSAC (Torr and Murray, 1997). Then the approximate 3D points are computed by forward intersection and the final orientations are estimated via bundle adjustment with iterative outlier elim-

Figure 3: Interactive object segmentation for modeling.

ination. Then dense image matching is performed to yield a dense grid of homologous point with a user-defined grid spacing and a dense cloud of 3D points is recovered through robust forward intersection.

3.3 Model Generation

As we need a surface model for visualization the point cloud is triangulated to a triangular mesh. In our experiments we have reprojected the points to one of the images and used $2\frac{1}{2}$-dimensional triangulation algorithm, which was readily available. For correct modeling it is however necessary to implement a more qualified method, which is able to produce a true 3D model, such as for example the volumetric integration method (Curless and Levoy, 1996). This is left for future work.

Many artefacts contain planar regions, where a dense point mesh is not only redundant, but also inaccurate because of the inevitable noise of the reconstruction process. At the triangulation stage it is also possible to reduce the amount of data and filter the noise by detecting the planar regions in the point cloud with a robust linear regression algorithm (an example is depicted in Figure 5). The regression algorithm should make use of the points' uncertainties, which are known from the bundle block adjustment to yield correct planar regions (Schindler, 2003). The points inside of planar regions can be removed and only the polygonal outline triangulated. Finally the triangulated surface models are textured with photo-texture from the original images. The textured model of the dancing girls' frieze (Waelkens et al., 2000) is shown in Figure 6. Another example of a textured model can be seen in Figure 7.

Figure 4: User interface for defining the image ordering.

Figure 5: Detection of planar regions for data reduction.

Figure 6: All 12 pieces of the 'dancing girls' frieze reconstructed and arranged in their original positions.

4 VISUALIZATION

4.1 Managing Large Models

Automatic modeling techniques, such as for example the one described in Section 2, deliver highly detailed 3D object models. However the resulting meshes often by far exceed the rendering capabilities even of high-end graphics workstations. Even worse, if it comes to deliver those models over the Internet, the transmission times become prohibitive if standardized file formats, e.g. VRML97 (The Web 3D Consortium, 1997), are used. It has therefore been an active field of research in the last years to provide tools to manage large and complex 3D data sets.

There are at least three concepts that are useful in dealing with huge 3D objects:

First, the raw data (e.g., the output of some acquisition and reconstruction process) contains much more detail than required for visualization. Thus the original object can often be replaced by a simplified one. This becomes particularly important in large scenes, where only a small fraction of

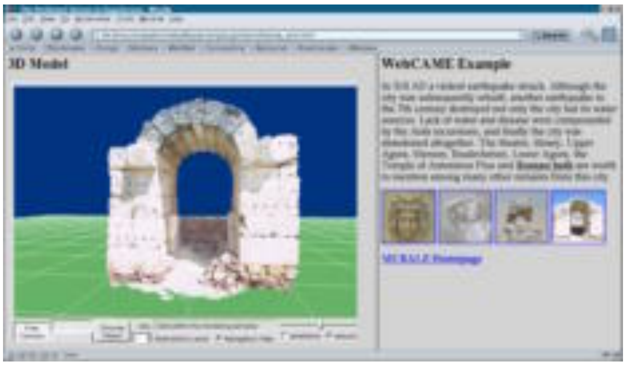


Figure 7: Screenshot of WebCAME embedded into a web page displayed in the Mozilla web browser showing an arch of the Roman baths at Sagalassos.

the entire data is visible and close to the virtual camera: objects which are close to the viewer have to be rendered at high detail, while more distant objects, where the detail is no longer visible, should be simplified. Since the viewpoint constantly changes in interactive applications, the mesh has to be adapted in real-time. The *progressive meshes* method (Hoppe, 1996) and some more involved techniques, such as (El-Sana and Chiang, 2000), belong to this class of algorithms.

Second, geometry information stored in a straightforward way (i.e. coordinates for points and indices for connecting lines and surfaces) contains a large amount of redundancy. By exploiting coherence within the data set the storage size can be reduced. This principle has first been described by (Deering, 1995), more efficient techniques have been developed later, e.g. (Taubin and Rossignac, 1998, Alliez and Desbrun, 2001b).

Third, if the simplified and compressed model is still too large to be displayed immediately after the user requested it, progressive transmission can be used to give the user a fast first impression and refine it progressively, as new data arrives. The user can start to operate on the (possibly low-quality) object with almost no delay. This feature is often integrated into simplification and compression techniques, e.g. (Khodakovsky et al., 2000, Alliez and Desbrun, 2001a).

In our visualization system we combine all three concepts. As stated above, we use triangle meshes as the representation of 3D objects. However, even under this restriction there are some properties that affect the choice of a particular method: generic meshes cannot be guaranteed to be two-manifold, and even simplifying a two-manifold mesh can create a non-manifold one. Therefore we will not rely on the two-manifold property (which greatly simplifies mesh simplification), but rather use a technique that treats non-manifold meshes and topology simplifications in a natural way (see Section 4.2).

4.2 Compressed Adaptive Multiresolution Encoding

The *CAME* data structure (Grabner, 2002), which we have developed, is an extension and enhancement of the *meta-node* approach (El-Sana and Chiang, 2000). It achieves a compression ratio of approximately 1:25 compared to the uncompressed meta-node data, while still providing view-dependent access to multiresolution triangle meshes based on an estimation of the screen-space error.

The key idea in CAME is to identify mesh vertices by *bit strings* indicating the path to be taken in the simplification hierarchy to reach the leaf node corresponding to the vertex. Topology compression is achieved by omitting redundant prefixes of bit strings. Mesh connectivity is implicitly maintained by independently traversing the simplification hierarchy according to the bit string stored with each triangle. For details we refer to the original publication.

After reconstruction the original meshes are converted to the CAME format by a fully automatic encoder and stored into a database. The rendering system used to view the 3D models progressively reads the multiresolution mesh from the database and provides a view-dependent rendering.

4.3 Internet-based Application

The CAME technology has been integrated into a lightweight web browser plugin called *WebCAME* (Grabner, 2003b). It allows to progressively transmit compressed textured 3D models with view-dependent simplification over the Internet. Although being a general-purpose tool, the navigation facilities have been designed with virtual archaeology applications in mind. The requirements of archaeologists and of visitors an archaeological sites are reflected in the navigation: both viewpoint-centered navigation to move around in the virtual world and object-centered navigation for careful exploration of single objects are provided. Figure 7 shows a screenshot of the WebCAME plugin with an arch from the Roman baths of Sagalassos.

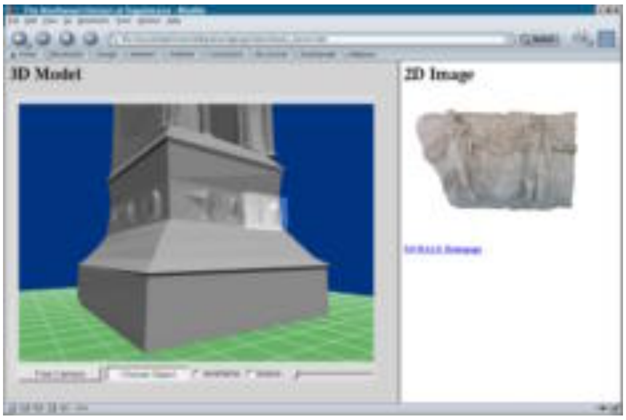
4.4 Smart Web Objects

To increase the semantics of the models, the traditional multiresolution concept, which only handles pure geometry, has been extended to preserve object identities and enable access to attribute data associated with the objects. The object identities are included in the CAME structure to support their compressed adaptive transmission.

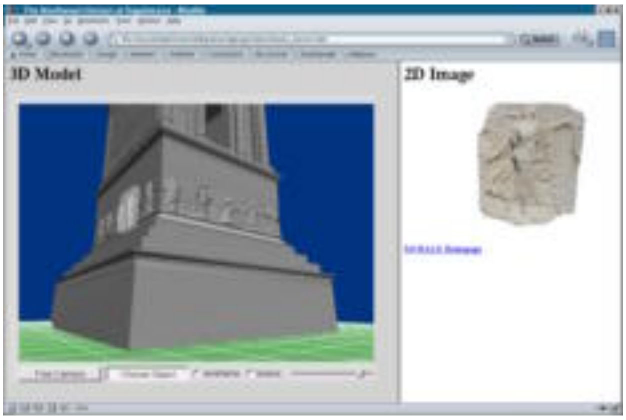
Application-specific properties of objects in the scene (metadata) are typically stored at a separate place, such as in our example the 3D MURALE multimedia database. In order to access this data, it must be possible to associate each piece of geometry in the scene with the appropriate metadata record. A simple way to solve this problem is to add a unique object identifier to every triangle in addition to color and texture information. This allows to perform the following important operations on the client side:

- Querying additional data about any part of the multiresolution mesh
- Manipulating (e.g. translating or rotating) parts of the mesh

Note that these operations are common for 3D modeling packages, but were previously not supported by multiresolution techniques. Since we address transmission over low bandwidth networks, our method is designed to efficiently represent the desired object information. By exploiting spatial and hierarchical coherence within the multiresolution data structure, the object assignment can be encoded with as little as 0.1 bits per triangle, which is negligible compared to the object geometry.



(a) high resolution



(b) low resolution

Figure 8: Querying meta-data about objects in the scene. Objects can be selected and their photographs displayed independent of the resolution of the rendered mesh resolution, as illustrated by the two screenshots with higher resolution (top) and lower lower resolution (bottom).

The object identification subsystem communicates with the Web browser (embedding the WebCAME plugin) by requesting URLs to be loaded into new browser windows or frames. This is a simple, but very general communication method. It enables the access to arbitrary database content independent of the momentary object resolution, because the object ID is propagated with every triangle through the multiresolution hierarchy. Figure 8 demonstrates a possible application of this technique: an object is selected by clicking into the multiresolution mesh and the corresponding photograph is loaded and displayed. The communication framework also directly allows the execution of JavaScript programs, which gives it great flexibility and facilitates the design of more complex interaction modes.

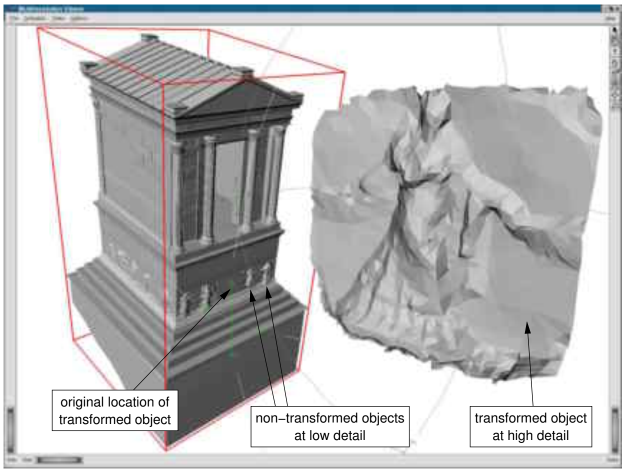


Figure 9: Northwest Heroon of Sagalassos with frieze of dancing girls, one of the objects is extracted from its original location and viewed at a close distance (at high detail), while the non-transformed objects in the background are rendered at low detail.

We will give an example how the ability to distinguish objects in the multiresolution framework also supports interaction: parts of the mesh can be extracted from their original location for close inspection by the user. The level-of-detail selection procedure takes the transformation into account, so that the examined object, which is closer to the viewer than in its original position, is presented at a higher resolution than it would appear in its original, more distant location. In Figure 9, one of the pieces of the 'dancing girls frieze' has been brought to the foreground and rotated by the user to study it in detail.

We have also examined the possibility to include attribute data (such as normal vectors) into the CAME framework. While an efficient and conceptually simple encoding could be found (Grabner, 2003a), it turned out that the implementation of the encoder/decoder stages is a sophisticated task due to the different orders of he involved hierarchical data structures. Decoding normal vectors for the rendering system is thus left for future work.

5 CONCLUSION

We have presented a work-flow for automatic 3D modeling and fast visualization of artefacts recorded on archaeological excavation sites with a digital camera. A set of robust multi-view-modeling techniques developed in the past decade in the field of digital photogrammetry and computer vision are used to reconstruct a 3D model with minimal user interaction, and the CAME data structure, a newly developed technique for efficient transmission and visualization of large models, is used for real-time viewing and interaction.

The purpose of the presented work was not to develop new technology, but to integrate the different state-of-the-art techniques into a complete work-flow and demonstrate its practical application in a veritable archaeological environment. Therefore only standard hardware has been used, namely cheap digital consumer cameras and average PCs, and the recordings have been made on site without specifically preparing the artefacts for recording. It has been shown that the described technologies are mature and can be applied under realistic conditions.

ACKNOWLEDGMENTS

This work has been supported by the European Commission under contract IST-1999-20273. The authors would like to thank Mark Waelkens from KU Leuven for his archaeological comments and for the logistical support in Sagalassos, and Joachim Bauer from the *VRVis research center* for the camera calibration software.

REFERENCES

Akeley, K. (ed.), 2000. SIGGRAPH 2000 Conference Proceedings. Annual Conference Series, ACM SIGGRAPH, ACM Press / ACM SIGGRAPH / Addison Wesley Longman, New York.

Alliez, P. and Desbrun, M., 2001a. Progressive compression for lossless transmission of triangle meshes. In: E. Fiume (ed.), SIGGRAPH 2001 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, ACM Press / ACM SIGGRAPH.

Alliez, P. and Desbrun, M., 2001b. Valence-driven connectivity encoding for 3D meshes. In: A. Chalmers and T.-M. Rhyne (eds), Proceedings Eurographics, Computer Graphics Forum, Vol. 20number 3, Blackwell Publishers. ISSN 1067-7055.

Cosmas, J., Itegaki, T., Green, D., Grabczewski, E., VanGool, F. W. L., Zalesny, A., Vanrintel, D., Grabner, M., Karner, K., Leberl, F. and Schindler, K., 2001. 3d murale: A multimedia system for archaeology. In: Proc. ACM SIGGRAPH Conference on Virtual Reality, Archaeology, and Cultural heritage (VAST01).

Curless, B. and Levoy, M., 1996. A volumetric method for building complex models from range images. In: H. Rushmeier (ed.), Computer Graphics, Proc. SIGGRAPH 96, Addison Wesley.

Deering, M. F., 1995. Geometry compression. In: R. Cook (ed.), SIGGRAPH 95 Conference Proceedings, Annual Conference Series, Vol. 29, ACM SIGGRAPH, Addison Wesley, pp. 13–20.

El-Sana, J. and Chiang, Y.-J., 2000. External memory view-dependent simplification. In: Proceedings Eurographics, Computer Graphics Forum, Vol. 19number 3, Blackwell Publishers, pp. 139–150. ISSN 1067-7055.

Grabner, M., 2002. Compressed adaptive multiresolution encoding. Journal of WSCG 10(1), pp. 195–202. ISSN 1213-6972.

Grabner, M., 2003a. Compression of arbitrary triangle meshes with attributes for selective refinement. Journal of WSCG 11(1), pp. 181–188. ISSN 1213-6972.

Grabner, M., 2003b. WebCAME: A light-weight modular client/server multiresolution rendering system. In: S. E. Spencer (ed.), Web3D 2003 Symposium Proceedings, ACM SIGGRAPH, Saint Malo, France, pp. 63–72. ISBN 1-58113-644-7.

Heikkilä, J., 2000. Geometric camera calibration using circular control points. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(10), pp. 1066–1077.

Hoppe, H., 1996. Progressive meshes. In: H. Rushmeier (ed.), SIGGRAPH 96 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison Wesley, pp. 99–108. ISSN 0097-8930.

Khodakovsky, A., Schröder, P. and Sweldens, W., 2000. Progressive geometry compression. In: (Akeley, 2000).

Klaus, A., Bauer, J., Karner, K. and Schindler, K., 2002. MetropoGIS: A semi-automatic city documentation system. In: Proc. Photogrammetric Computer Vision 2002 (PCV'02). ISPRS – Commission III Symposium, Graz.

Pollefeys, M., Vergauwen, M., Cornelis, K. and Gool, L. V., 2001. 3d acquisition of archaeological heritage from images. In: Proc. International CIPA Symposium, Potsdam, Germany.

Schindler, K., 2003. On robust regression in photogrammetric point clouds. In: Proc. 25th DAGM Pattern Recognition Symposium, Magdeburg, Germany.

Taubin, G. and Rossignac, J., 1998. Geometric compression through topological surgery. ACM Transactions on Graphics 17(2), pp. 84–115.

The Web 3D Consortium, 1997. The Virtual Reality Modeling Language: International standard ISO/IEC 14772-1:1997. http://www.vrml.org/technicalinfo/specifications/vrml97.

Torr, P. and Murray, D., 1997. The development and comparison of robust methods for estimating the fundamental matrix. International Journal of Computer Vision 24(3), pp. 271–300.

Waelkens, M., Vandeput, L., Arikan, B., Berns, C., Poblome, J. and Torun, E., 2000. The northwest heroon at sagalassos. Sagalassos V pp. 553–594.