

3D MODEL GENERATION AND VISUALIZATION OF WARTBURG CASTLE

Helmut Mayer^a, Matthias Mosch^b, Juergen Peipe^c

^a Bundeswehr University Munich, D-85577 Neubiberg, Germany - Helmut.Mayer@unibw-muenchen.de

^b Albert-Ludwigs-Universität Freiburg, D-79106 Freiburg, Germany - matthias.mosch@felis.uni-freiburg.de

^c Bundeswehr University Munich, D-85577 Neubiberg, Germany - j-k.peipe@unibw-muenchen.de

KEY WORDS: Cultural Heritage, Multimedia, Internet, Calibration, Orientation, Modelling

ABSTRACT:

In a joint project between Bundeswehr University Munich and Albert-Ludwigs-University Freiburg different methods to generate a 3D model of Wartburg Castle have been investigated. Several data acquisition devices were used, i.e. high resolution digital cameras and a low resolution camcorder. The virtual reconstruction of the historical building was accomplished by means of the low-cost (photogrammetric) software packages Canoma and PhotoModeler as well as by fully automatic computer vision methods for camera calibration, image orientation, and model generation. In this paper, the survey of Wartburg Castle is described and results of the modelling procedures are presented.

1. WARTBURG CASTLE

Founded in 1067, the *Wartburg* is a famous cultural heritage monument with regard to German history and an outstanding landmark of touristic interest. Wartburg Castle is located near *Eisenach/Germany* on top of a rocky hill and surrounded by the *Thuringian Forest* (Fig. 1 and 2). The terrain slopes steeply away on all sides of the castle complex. The inner courtyards are only accessible via a small drawbridge.

The photogrammetric recording and 3D model generation is part of the EU funded research project "Multimedia Geo-Information for E-Communities in Rural Areas with Eco-Tourism" (Frech & Koch, 2003). Within the project an on-line tourist information system of the Thuringian Forest has been developed providing also visualization by means of, e.g., 2D tourist maps, aerial images, 3D models and 3D sceneries. First

results of the photogrammetric work concerning Wartburg Castle were presented in Mayer et al. (2003).

Medieval castles are, in most cases, large and complex heritage sites showing a mixture of walls, gates, courtyards, towers, and buildings formed, extended, destroyed and reconstructed over centuries. Their survey is a relatively complicated and expensive task (e.g. Hanke & Oberschneider, 2002; Gonzo et al., 2004). In this paper, two image-based approaches for the documentation, 3D modelling and visualization of Wartburg Castle are presented. Subsequent to the acquisition of digital images (chapter 2), the photogrammetric object reconstruction was performed by the low-cost softwares Canoma and PhotoModeler (chapter 3) as well as by fully automatic procedures using computer vision techniques (chapter 4 and 5).

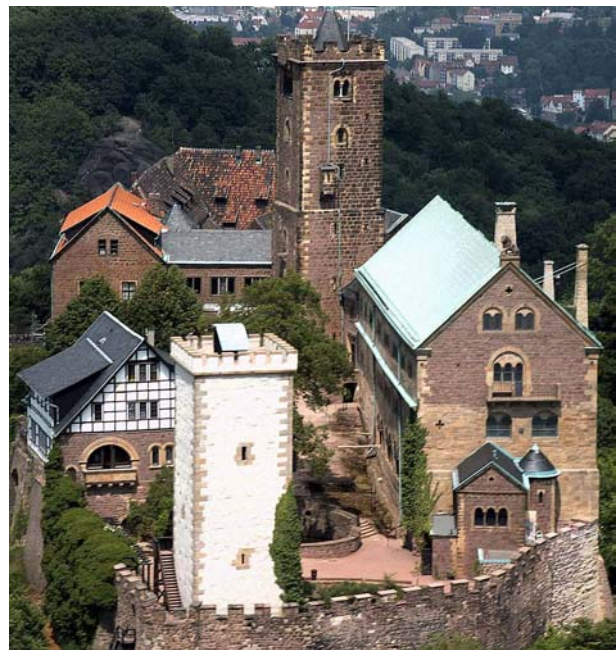


Figure 1 and 2. Wartburg Castle

	Rollei d30 metric ⁵	Canon EOS D60	Sony DCR-TRV738E
Number of pixels	2552 x 1920	3072 x 2048	720 x 576
Sensor format	9 mm x 7 mm	22 mm x 15 mm	
Lens (focal length)	10 mm - 30 mm (= 40 mm - 120 mm for 35 mm camera)	20 mm (SLR) (= 32 mm for 35 mm camera)	3.6 mm - 54 mm (= 48 mm - 720 mm für 35 mm camera)
Image data	6.4 MB uncompressed raw data per image	7.4 MB uncompressed raw data per image	0.8 MB

Table 1. Camera specifications

2. DATA ACQUISITION

Digital images were acquired by means of three imaging systems, a 5 MP Rollei d30 metric⁵, a 6 MP Canon EOS D60 SLR camera, and a Sony camcorder (Tab. 1). The Rollei provides an optical zoom with two pre-calibrated focal lengths (15 mm and 30 mm). The camera was used to take some hundred photographs from the towers and windows of the castle, and from the ground. The Canon D60 with wide angle lens was suited to record several parts of the courtyards from a shorter distance. Additionally, the camcorder served to collect some overviews and short image sequences (movies) to be presented in the tourist information system, e.g., Luther's workroom and the famous banquet hall. Due to the location of the castle, images of the outer facades had to be taken from an ultra-light airplane.

3. OBJECT RECONSTRUCTION AND 3D MODELLING BY MEANS OF LOW-COST SOFTWARE

The digital images were measured and object coordinates were determined within the low-cost software packages Canoma and PhotoModeler. The visualization tools of these programs were used to create a photorealistic model of the castle.

Canoma (2004) was originally developed and sold by MetaCreations Corporation, meanwhile acquired by Adobe. Unfortunately the future status and availability of Canoma is not known. The software aims at modelling man-made structures, but not irregular shapes like trees or people. Canoma produces geometrically simple models for the Web and is not intended to replace detailed CAD models. Yet, it is easy to learn and to use, and one is able to get quick results, i.e. a photorealistic 3D model. Results can be posted on websites or exported into common 3D modelling and animation software via standard formats (e.g. DXF, VRML2). Also interactive walk-through animations and movies (recordable to Quicktime movie format) can be produced. Fig. 3 and 4 show images cut out of a virtual flight around Wartburg Castle.

The well-known PhotoModeler software (Version 5.0; PhotoModeler, 2003; for camera calibration see also Wiggengagen, 2002) was additionally applied to some regions of the castle. PhotoModeler offers advanced photogrammetric

features such as camera calibration and bundle adjustment to achieve more accurate modelling results necessary in the courtyards and interior rooms of Wartburg Castle.

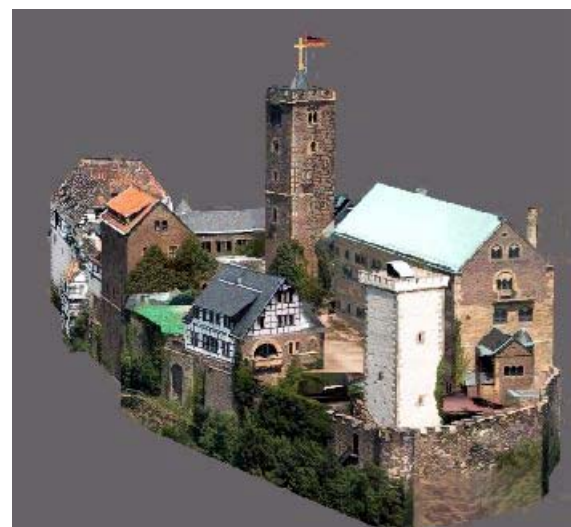


Figure 3 and 4. Images extracted from a virtual flight around Wartburg Castle

4. FULLY AUTOMATIC ORIENTATION AND CALIBRATION

The approach presented here is an extension of (Mayer et al., 2003). It rests on the scheme outlined in (Hao and Mayer, 2003). I.e., we match images in a pyramid and compute the uncalibrated orientation based on RANSAC (Fischler and Bolles, 1981) for pairs based on the fundamental matrix and for triplets based on the trifocal tensor (Hartley and Zisserman, 2003). The triplets are then linked via the direct linear transform (DLT; McGlone et al., 2004). Recent improvements comprise the substitution of RANSAC by an improved version termed geometric information criterion (GRIC; Torr, 1997) and the implementation of a multi image matching approach replacing the original pair-wise matching. The biggest change is the incorporation of the linear self-calibration approach of (Pollefeys et al., 2002, 2004) based on the dual absolute quadric making use of constraints on the parameters of a standard camera: It is assumed that the shear is small, that the principal point is close to the center of the image and that the ratio of the focal length and the image size is in a range between one third and three. Everything together with a robust bundle adjustment taking into account radial distortion made it possible to automatically obtain for the Wartburg castle scene a calibration corresponding well to the known for the calibrated camera Rollei d30 metric⁵ and a back projection error in the range of 0.2 pixels. The latter is an improvement by a factor of two compared to the results reported in (Mayer et al., 2003).

We wish to emphasize that the complete orientation and calibration is done totally automatically only based on the images and the assumption, that the images were taken by a perspective camera, possibly distorted by image deformations. There were no markers, ground control information, or approximate values used. The model needs to be rotated about three axes, shifted in three directions and scaled to comply with a real world coordinate system, but in spite of this, it is Euclidean, i.e., angles and distance ratios are correct. The bundle adjustment for nine images used for the results on the Wartburg castle scene given below was done with more than twenty five thousand observations and eight thousand unknowns from more than 2800 points, close to 1100 of them being 6-fold points and above.



Figure 5. First courtyard of Wartburg Castle, partly used for the automatic measurement and modelling approach

5. DETERMINATION OF PLANES AND THEIR EXTENT FOR THE INTERPRETATION OF BUILDING FACADES

We have recently obtained a research grant on building façade interpretation. Therefore, we have started to describe the Wartburg castle scene in terms of a façade model. This model basically consists of the description of façade parts by means of planes. The planes are detected from the relatively dense 3D points obtained by the orientation and calibration procedure described above by a GRIC based approach and then least squares matching is used to determine the actual extent of the planes.

Before detecting planes, we orient the whole 3D scene in the vertical direction, facilitating a later detection of architectural elements such as windows and doors, which are usually vertically aligned. As we assume that we deal with architectural scenes there is always a larger number of vertical lines available, i.e., lines which intersect in the vertical vanishing point in the image. The user just has to tell the system if the vanishing point is in the x- or in the y-direction, and then the vanishing point for an image is determined by means of a RANSAC based approach for line intersections. We compute vertical vanishing points for one or more images and together with the known relative orientations of the images we determine an average vertical direction used for vertical alignment.

To generate hypotheses for planes, we compute a large number of planes from randomly chosen triplets of points. Then, all points in the 3D scene are checked via the GRIC criterion, which takes into account also the precision of the 3D points obtained by the bundle adjustment, if they lie on a hypothesized plane. If the GRIC value lies below a threshold it is summed up, otherwise the threshold is summed. Finally, the hypotheses with the lowest sum is taken. As there can be more than one meaningful plane in our scene, we iterate the above procedure, generating additional planes which have a limited overlap in terms of common points (we use 30%) with already accepted plane hypotheses.

The points defining the above planes give only a sparse representation of the façade. To determine their actual extent including the possibility of several topologically separated parts, but also to improve the precision of their parameters, we match the plane hypotheses in all images. We use the given orientation and plane parameters to project the image parts corresponding to the bounding box of the points defining the plane onto the plane. If all parameters were correct, the resulting images were perfectly aligned.

Neglecting errors in the orientation of the cameras, differences between the projected images result from the unprecise plane parameters, but also from areas which do not lie on the plane. For façades this could be doors and windows lying behind the plane or drain pipes and columns lying in front of the plane. For more complex examples two parts of a façade residing on the same mathematical plane can also be separated by a part of the building. We treat both cases by means of robust least squares matching. The unknowns are the parameters of the plane as well as a bias and drift parameter for the brightness of every image. To separate areas actually residing on the plane from other areas, we use as threshold for the brightness difference to the average value of all images the double median

value. To avoid a too strong bias from areas not on the plane on the parameters of the plane and particularly on the bias and drift parameter for every image, we start the adjustment with dilated areas around the points known to lie on the plane.

Results of this approach on nine images of Wartburg castle with a size of 5 MP acquired with a Rollei d30 metric⁵ camera, the first image of which is given in Fig. 6, are shown in Figs. 7 and 8. As can be seen, we have extracted the seven major planes describing the two parts of the two façades, the roofs, and also the wall in front of the building. For the latter we also detected the holes. Please note, that due to limitations of computing resources the least squares matching was carried out in images of reduced resolution eliminating some details.

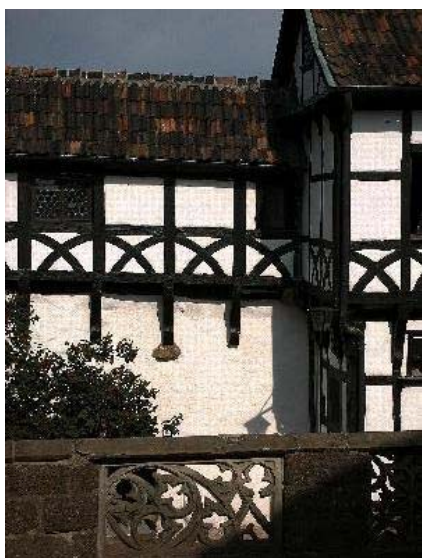


Figure 6. First image of nine of Wartburg castle acquired with a Rollei d30 metric⁵ camera used for orientation, calibration, and façade plane estimation.

Our ultimate goal is to describe the façades in terms of its elements such as doors, windows, etc. To cope with the statistical nature of the problem, we want to use a statistical sampling approach based on reversible jump Markov chain Monte Carlo (RJMCMC) as proposed by (Dick et al., 2004) for façade interpretation. One interesting feature of this approach is, that one can simulate the model by sampling into the distribution without using image data, making it possible to check a model for its plausibility. One can also substitute parts not seen by a sensor, e.g., the back of a building, in a way consistent with the rest of the building.



Figure 8. Frontal view, seen slightly from above of the same scene as in Fig. 7. On the lower side the green pyramids show the estimated positions of the cameras. For the wall in front of the building, the parts which have a pattern of holes can be distinguished from the solid parts.

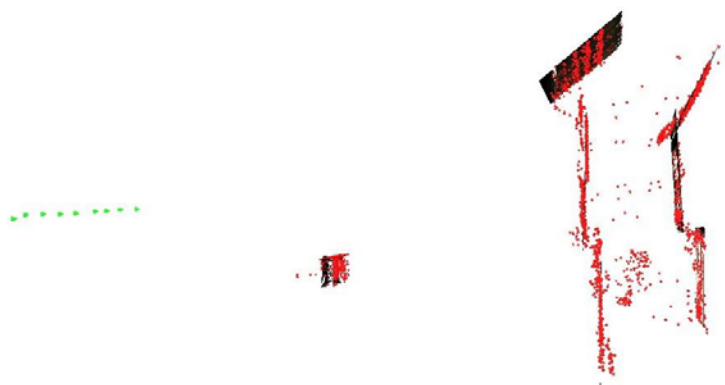


Figure 7. Side view on the fully automatically generated 3D points as well as the planes computed by RANSAC and least squares matching. The green pyramids on the left side show the estimated positions of the cameras.

6. CONCLUDING REMARKS

The use of 3D models of historical buildings has significantly increased over the last years. In this contribution the generation of a 3D model of Wartburg Castle as part of a tourist information system is described. Different image-based methods were applied, i.e. photogrammetric low-cost software and an automatic measurement and modelling approach. The latter allow fully automatic feature extraction, orientation, and 3D object reconstruction, even if the knowledge about the geometry of the image sequence is weak. Auto-calibration is possible in case that the interior orientation data have changed during the image recording (focusing, zoom) or are not available at all. It is advisable to combine several techniques to achieve virtual models from images as precise and reliable as required, especially in the case of a complex building like a medieval castle.

7. REFERENCES

- Canoma, 2004. Resources for Canoma Users. www.canoma.com (10 Sept 2004).
- Dick, A., Torr, P. and Cipolla, R., 2004. Modelling and Interpretation of Architecture from Several Images. *International Journal of Computer Vision* 60(2), pp. 111-134.
- Fischler, M. and Bolles, R., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24(6), pp. 381-395.
- Frech, I., Koch, B., 2003. Multimedia Geoinformation in Rural Areas with Eco-Tourism: The ReGeo-System. In: *Information and Communication Technologies in Tourism 2003*, Springer, Vienna/New York, pp. 421-429.
- Gonzo, L., El-Hakim, S., Picard, M., Girardi, S., Whiting, E., 2004. Photo-Realistic 3-D Reconstruction of Castles with Multiple-Sources Image-Based Techniques. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. (35) B5, pp.120-125.
- Hanke, K., Oberschneider, M., 2002. The Medieval Fortress Kufstein, Austria - An Example for the Restitution and Visualization of Cultural Heritage. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. (34) 5, pp. 530-533.
- Hao, X. and Mayer, H., 2003. Orientation and Auto-Calibration of Image Triplets and Sequences. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. (34) 3/W8, pp. 73-78.
- Hartley, R. and Zisserman, A., 2003. *Multiple View Geometry in Computer Vision - Second Edition*. Cambridge University Press, Cambridge, UK.
- Mayer, H., Mosch, M. and Peipe, J., 2003. Comparison of Photogrammetric and Computer Vision Techniques - 3D Reconstruction and Visualization of Wartburg Castle. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. (34) 5/C15, pp. 103-108.
- McGlone, C., Mikhail, E., Bethel, J. (2004): *Manual of Photogrammetry – Fifth Edition*. American Society for Photogrammetry and Remote Sensing (ASPRS), Bethesda, USA.
- PhotoModeler, 2003. *User Manual PhotoModeler Pro 5.0*. Eos Systems Inc., 488 p.
- Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K. and Tops, J., 2004. Visual Modeling with a Hand-Held Camera. *International Journal of Computer Vision* 59(3), pp. 207-232.
- Pollefeys, M., Verbiest, F. and Van Gool, L., 2002. Surviving Dominant Planes in Uncalibrated Structure and Motion Recovery. In: *Seventh European Conference on Computer Vision*, Vol. II, pp. 837-851.
- Wiggenhagen, M., 2002. Calibration of Digital Consumer Cameras for Photogrammetric Applications. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. (34) 3b, pp. 301-304.
- Torr, P., 1997. An Assessment of Information Criteria for Motion Model Selection. In: *Computer Vision and Pattern Recognition*, pp. 47-53.

ACKNOWLEDGMENTS

We want to thank Marc Pollefeys for making us available his MATLAB implementation of the linear self-calibration code based on the dual absolute quadric.