Generating a 3D model of a Bayon tower using non-metric imagery

Author(s):
Niederöst, Jana; Zhang, Li; Grün, Armin

Publication date:
2001

Permanent link:
https://doi.org/10.3929/ethz-a-004331583

GENERATING A 3D MODEL OF A BAYON TOWER USING NON-METRIC IMAGERY

Jana Visnovcova, Li Zhang, Armin Gruen
Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology Zurich
ETH – Hönggerberg
CH-8093 Zurich, Switzerland
Tel.: +41-1-633 31 57, Fax: +41-1-633 11 01
E-mail: <jana> <zhangl> <agruen>@geod.baug.ethz.ch

KEY WORDS: cultural heritage, non-metric camera, 3D reconstruction, texture mapping

ABSTRACT

The Hindu and Buddhist monuments of Angkor in Cambodia are the legacy of highly developed Khmer empires. Well-known for their structural and surface complexity they constitute a great challenge to any attempt towards precise and detailed 3D measurements and modeling. This paper reports about a pilot project using modern techniques of analytical and digital photogrammetry to derive a photorealistic 3D model of one of the very complex towers of the famous Bayon temple of the ancient city of Angkor Thom. The project is based on the use of small format analogue tourist-type photography. Described processing steps include phototriangulation on an analytical plotter, automated image matching for surface model generation and the use of novel point cloud editing and view-dependent texture mapping techniques.

1. INTRODUCTION

The Khmer kingdom of Angkor in Cambodia is one of the most remarkable empires in human history. On 75 square miles of fertile plains, north of the modern town of Siem Reap, a succession of 42 kings built during the Angkor Period 802-1432 A.D. about 72 major monuments, temples, palaces, canals with dikes, moats and reservoirs. This extended complex ranks among the most spectacular sites currently listed in the UNESCO World Heritage List. Many of the treasures are now lying in ruins or have been partially restored in the past by various international expert teams. One of the many monuments has been subject to particular attention throughout the years (Preston, 2000, White, 1982): The Temple of Angkor Wat.

Today, after a decade of forced neglect during the rule of the Khmer Rouge, the precious Angkor monuments receive much help again from various countries. Among those is the JSA project (Japanese government team for Safeguarding Angkor), which aims at the conservation and restoration of the Northern Library of the Bayon temple, the Royal Plaza (Prasat Suor Prat and its Terraces) of Angkor Thom and the Northern Library of Angkor Wat (JSA News, 1999). While on sabbatical leave at the Asian Institute of Technology (AIT) in Bangkok in spring 1999 the third author was invited to join the Japanese team of geographers and surveyors from the Keio University, Tokyo as photogrammetrist for a balloon photogrammetry mission over the Bayon temple. The main purpose of this mission was to generate a 3D computer model of the very complex structure of Bayon from small format balloon images. The results of this work will be published elsewhere. During this one week mission the author took a number of amateur

photographs of one of the many Bodhisattva-faced towers. This paper reports about the processing of these images and shows the resulting texture mapped 3D model of the tower.

2. THE TEMPLE OF BAYON

Bayon was the last great temple built in Angkor and remains one of the most enigmatic. Jayavarman VII became king of the Khmer in 1181. After the Cham (ancient Vietnamese) sacked Angkor in 1177 he conquered their homeland Champa and began a frenzy of buildings in Angkor, among which are Preah Khan, Ta Prohm, and Banteai Kdei, all adjacent to the city Angkor Thom, which he also created. Angkor Thom was then the Khmer capital, a four-sided enclosure surrounded by a wall 8 miles long and a vast 100 m broad moat.

Since the old Hindu gods failed to support the Khmers in 1177 he dedicated his largest temple, the Bayon, to Buddha. Today Bayon is considered an interesting mixture of both Hindu and Buddhist elements of style. Bayon is situated right in the center of Angkor Thom. It is an extremely complex structure, which is best revealed by aerial photographs (Figure 1). It is said to represent the Mount Meru of an old Hindu legend. It consists of three levels of platforms and two galleries with spectacular and well-preserved basreliefs, showing the live of the ordinary people and stories from the Hindu mythology.

Figure 1: Aerial image of Bayon, taken from a Helium balloon

Figure 2: The smiling faces of Bayon (one of 13 non-metric images used for the photogrammetric processing)

A special feature of Bayon are its 54 towers with four large faces on each, pointing in all four geographical directions (such adding up to 216 faces in total, not counting the many more on other objects). These faces on the towers and gateways of Angkor Thom represent Lokesvara, a Bodhisattva from the Mahayana Buddhism, a holy one who stayed at earth to do good work and help people. The origine and purpose of these faces are still under scientific dispute. Their symbolism is not yet untangled and they continue to amaze visitors from all over the world (Figure 2).

3. DATA ACQUISITION

The main goal of the Bayon field campaign was the taking of small format camera balloon images over Bayon for the 3D reconstruction of the complete and utterly complex temple. Balloon images where also taken over the Northern Library of Angkor Wat and over the several Prasat Suor Prat buildings. During the mission a sequence of small format terrestrial images was taken with a Minolta Dynax 500si camera (c = 35 mm) of one of the Bayon towers on the third platform in the north-eastern corner of the temple. The 13 images covering the full horizon were meant as a test of photogrammetric procedures rather than a serious project aiming at a complete recording of the object. Figure 3 shows the arrangement of images, as they were used for bundle triangulation around the tower. Since a 360 degree azimuth coverage was necessary and the light conditions were fairly extreme the production of good, evenly illuminated pictures was practically impossible without artificial lighting, which was not available at the site. Therefore, the images suffer under strong variations of the illuminated and shadow areas. Also, the shadow and light parts will vary from image to image, depending on the time of the day the images were taken. This will cause problems with texture mapping from multiple images, if an even distribution of light is aimed at allover the 3D model. Thus a great deal of our work went into a modification of the standard procedure of texture mapping. We developed a specific technique of view-dependent texture mapping, which is object face oriented and picks for each object face a combination of the best possible corresponding image patches for texture mapping.

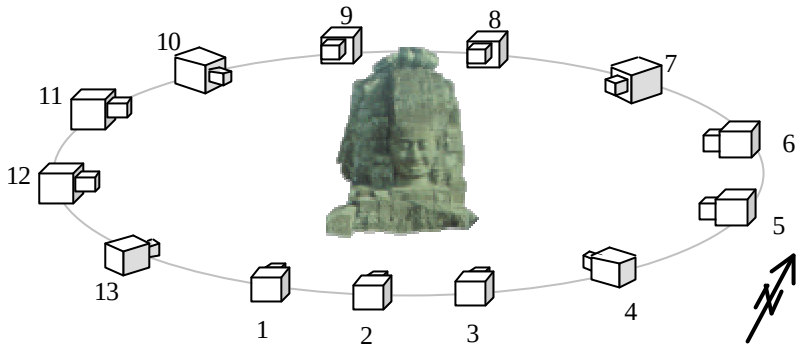


Figure 3: Images used for bundle triangulation

4. TRIANGULATION

In order to obtain a realistic and geometrically correct 3D model of the Bayon tower, 13 images covering the whole horizon were triangulated as follows:

- Determination of coordinates of fictitious fiducial marks
- Image measurement for triangulation
- Bundle adjustment with self-calibration

4.1 Determination of coordinates of fictitious fiducial marks

Unknown and unstable parameters of interior orientation, film unflatness and the lack of fiducial marks are essential features of non-metric small format cameras. Several methods have been developed to eliminate or reduce these shortcomings. To overcome the lack of fiducials, corners of the frame are commonly calculated as intersections of frame edges.

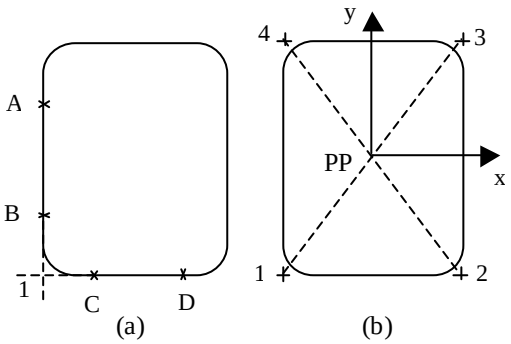


Figure 4: (a) Indirect measurement of fiducial marks on an analytical plotter (points A – D for the fiducial mark 1), (b) Image coordinate system with the principal point located by intersection of diagonals

In this paper, the coordinates of fictitious fiducial marks were determined by indirect measurement in several analog images. On an analytical plotter two points on adjacent edges of the image frame were measured and intersecting points were calculated (Figure 4a). As a result individual sets of coordinates in the machine coordinate system were obtained for all images. An affine transformation was performed in order to relate the different coordinate systems to a unique image coordinate system (Figure 4b).

4.2 Image measurement for triangulation

The image measurements for triangulation were performed on an analytical plotter using the approximate interior orientation (camera constant 35 mm, no lens distortion, zero coordinates of the principal point and estimated coordinates of fiducial marks). The optimal spatial distribution of orientation points was not possible due to incomplete frame fill. Altogether about 170 points were measured in 13 images around the tower. Each object point is imaged in two to six images.

For the absolute orientation of the whole block two full control points and one depth control point were used. Control points 1 and 2 are defined as marks on a vertical scale-bar. Depth control point 3 is situated on a small decorated stone to the left of the southern tower side. These three control points define the right-handed Cartesian object coordinate system with z-axis going towards the southern projection centers (Figure 4). Since control points are only visible from the southern tower side, the absolute orientation on the analytical plotter was performed in two models only.

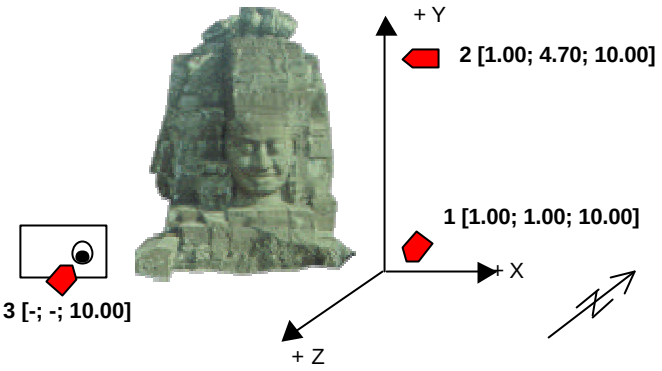


Figure 5: Cartesian object coordinate system and three control points with coordinates given in meters

4.3 Bundle adjustment with self-calibration

The bundle adjustment of the whole image sequence was performed in a stepwise mode. For that purpose the procedure was started with the first absolutely oriented images number 1, 2 and 3. Consequently additional images were added with manually obtained approximations for exterior orientation elements and new object points. This stepwise mode was necessary because, with the available coarse first approximations, the complete block did not converge simultaneously.

After the establishment of a stable adjusted image block, ten additional parameters (Brown, 1971) were used to model systematic errors: focal length correction, principal point coordinate offsets, five parameters modelling radial and decentering lens distortion and two parameters for a differential scale factor and shear (Beyer, 1992). The radial distortion amounted to 810 microns at an image radius of 20 mm.

5. IMAGE MATCHING

For the extraction of the 3D surface geometry the images were digitized with a resolution of 1200 dpi (pixelsize 21 microns). Image matching was performed with the commercial software package MATCH-T. However, MATCH-T can only generate the 2.5D model from one stereopair. This is necessarily incomplete for close-range applications because occluded parts of the scene cannot be processed. Therefore four image pairs taken from south, east, north and west (images number 2/3, 5/6, 8/9 and 11/12 in Figure 3) were selected and matching was performed in each model separately. This procedure required a transformation of orientation parameters from the Cartesian object coordinate system to local systems for each geographic direction (so that – like in an aerial case – the z-axis in each model is directed towards the projection centers). As a result of fully automatic matching, four separate surface models with a grid width of 3 cm were constructed. The visualization of matched points has shown that the image matching procedure in MATCH-T works reasonably well in this case. Within the individual models only about 0.2% of the matched points had to be edited manually (Figure 6). In the next step, the four separate point clouds were transformed back to a joint object coordinate system and merged together. In the complete 3D point cloud containing 46 850 points overlapping areas occurred and gaps and outliers between the adjacent models showed up (Figure 7). To achieve a good visualization result, these errors were eliminated automatically in a second editing step (Chapter 6).

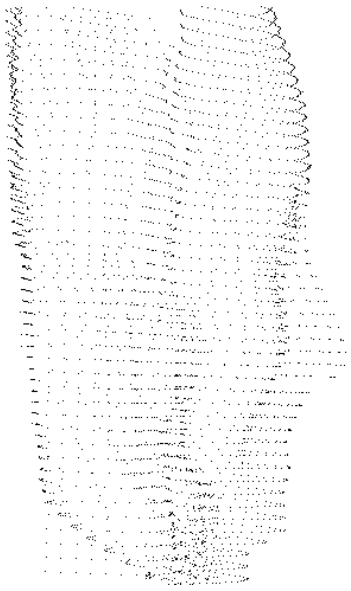


Figure 6: Matching result within one model (a detail of the southern face profile)

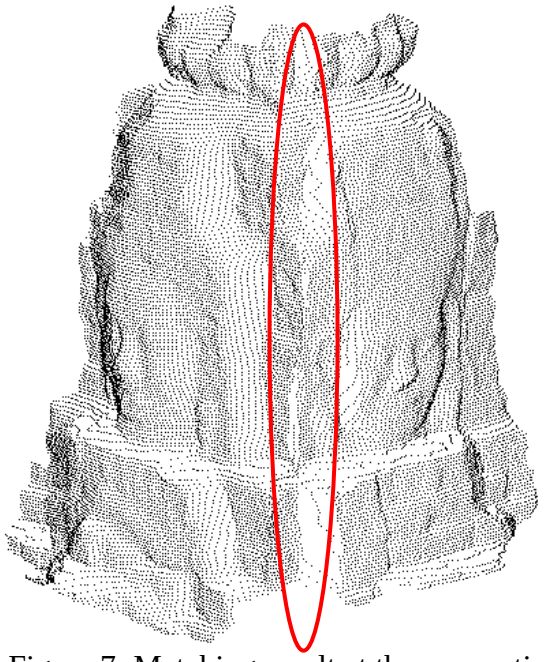
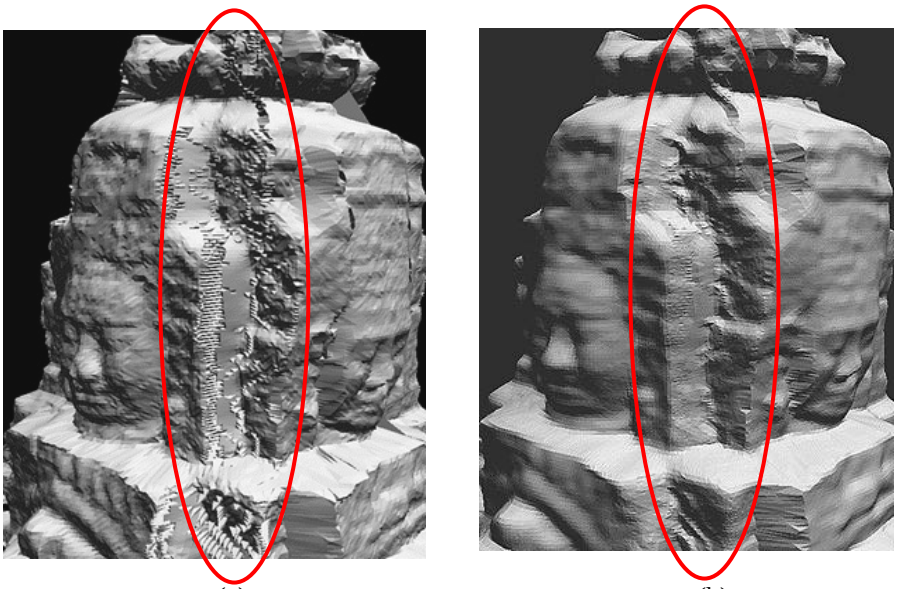


Figure 7: Matching result at the connection of two separately processed models with marked area needing additional editing

6. EDITING AND TRIANGULATION OF THE POINT CLOUD

Determining surfaces from a set of 3D points containing outliers is a complex task. Many approaches have been designed to treat this problem, such as 3D deformable surfaces (Cohen et al., 1991) and iterative local surface fitting (Fua and Sander, 1992). The last algorithm fits second order patches around each 3D point and groups points into surfaces. In this iterative process errors are eliminated without smoothing out relevant features. Our procedure is an adapted version of this algorithm without performing resampling and clustering. Because all the points should belong to one unique surface, only those outliers are deleted whose derivation from the fitted surface exceeds after 3-5 iterations the predefined threshold. Finally the errors are eliminated while preserving essential surface features (Figure 8).

For the conversion of the point cloud to a triangular surface mesh the 2.5D Delaunay triangulation was applied. Without losing its topology, the 3D surface model of the Bayon Tower was expanded to a plane by transforming the Cartesian coordinate system to a cylinder coordinate frame. In the defined rqz cylinder frame z is the vertical cylinder axis crossing the model center and parallel to the original Y-axis of the Cartesian object coordinate system. r is the Euclidean distance from the surface point to the z-axis and q is the angle around the z-axis. The 2.5D Delaunay triangulation was done in the qz plane. The final shaded model of the triangulated mesh is shown in Figure 8b.



(a) (b)

Figure 8: The shaded triangulated point cloud before (a) and after editing (b). Most outliers and gaps appeared at the connection of two separately matched models (marked area)

7. VIEW-DEPENDENT TEXTURE MAPPING AND VISUALIZATION

With the technique of texture mapping, gray-scale or true color imagery is mapped onto the 3D geometric surface in order to achieve photorealistic virtual models. Knowing the parameters of interior and exterior orientation, to each triangular face of the 3D surface the corresponding image coordinates are calculated. The gray-scale or color RGB values within the projected triangle are then attached to the face.

A common approach of texture mapping is to use one frontal image for a related part of the object. In close-range applications this is often not satisfactory, because not enough image information is available for fully or partially occluded object parts. In the 3D model the texture appears as "stretched" (Figure 9). Moreover, often varying light conditions during image acquisition do not allow regular light distribution allover the object. This causes sharp transitions between neighbouring object parts, which are texture-mapped from different frontal images. To overcome these problems, a new method of texture mapping was developed – a view-dependent texture mapping.

Our method of view-dependent texture mapping is based on the selection of a combination of optimal image patches for each triangle of a 3D model. According to the best possible geometric and radiometric conditions a combination of image content is calculated from all images where a particular triangle appears. The locations of the image triangles are computed from object faces via collinearity equations. The procedure consists of three steps: pre-processing, selection of a geometrically optimal image and texture weighted averaging.

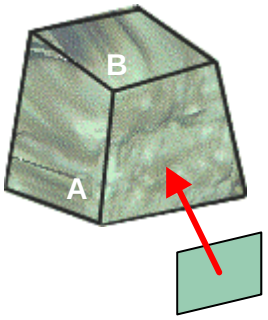


Figure 9: Texture mapping with one frontal image – sloped faces A and B appear with "stretched" texture

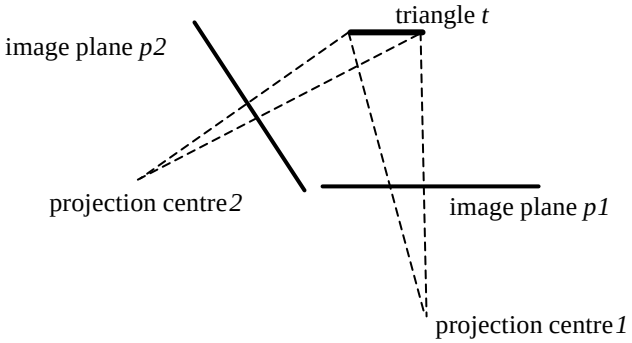


Figure 10: View-dependent texture mapping: for each object triangle the optimal texture image is selected (image *1* with the bigger area of projected triangle *t*)

(1) Pre-processing. In order to achieve a bigger local contrast and nearly the same color balance in all images high pass filtering and histogram equalization are used for each RGB channel separately. Similarly to Wallis filtering, the high pass filter enables a strong enhancement of the local contrast by removing low-frequency information in an image. It retains edge details in the specified radius where sharp color transitions occur. We apply high pass filtering with a radius of 50 pixels. Additionally, the equalization is performed for all three channels in all images. With these procedures we avoid color shifts when merging the separate channels together.

(2) Selection of a geometrically optimal image (Figure 10). In all image planes where a particular triangle *t* is projected, the area of the triangle is calculated. The image where the triangle *t* appears largest contains the most texture information. Thus, it is considered as a geometrically optimal image candidate for texture mapping. However, if we use only one image patch, in case of large differences in image radiometry and in case of strong local variability of the surface patch normals, the result will be a checkerboard-type 3D texture map. Therefore in our procedure the radiometry of adjacent images related to the selected "optimal" image is also considered.

(3) Texture weighted averaging. We consider all images where a particular triangle *t* appears simultaneously. The gray values of these images are averaged according to equation (1)

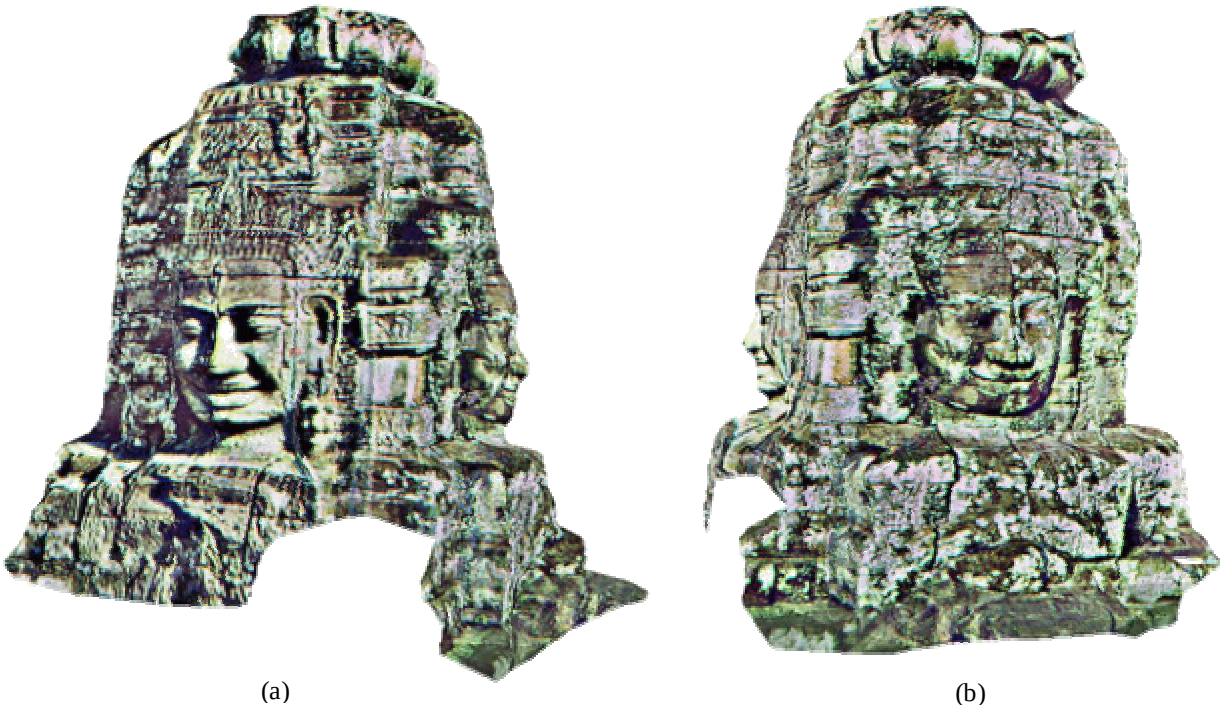
$$\overline{g}_i = \frac{\sum_{j=1}^{n} g_{ij} \cdot w_{ij}}{\sum_{j=1}^{n} w_{ij}}, \tag{1}$$

where $\overline{g}_i$ .... gray values of the "new" image
g_{ij} .... gray values of "old" image j
n .... number of image patches used for texture weighted averaging
w_{ij} .... weight factors calculated from the areas of projected triangles (equation 2)

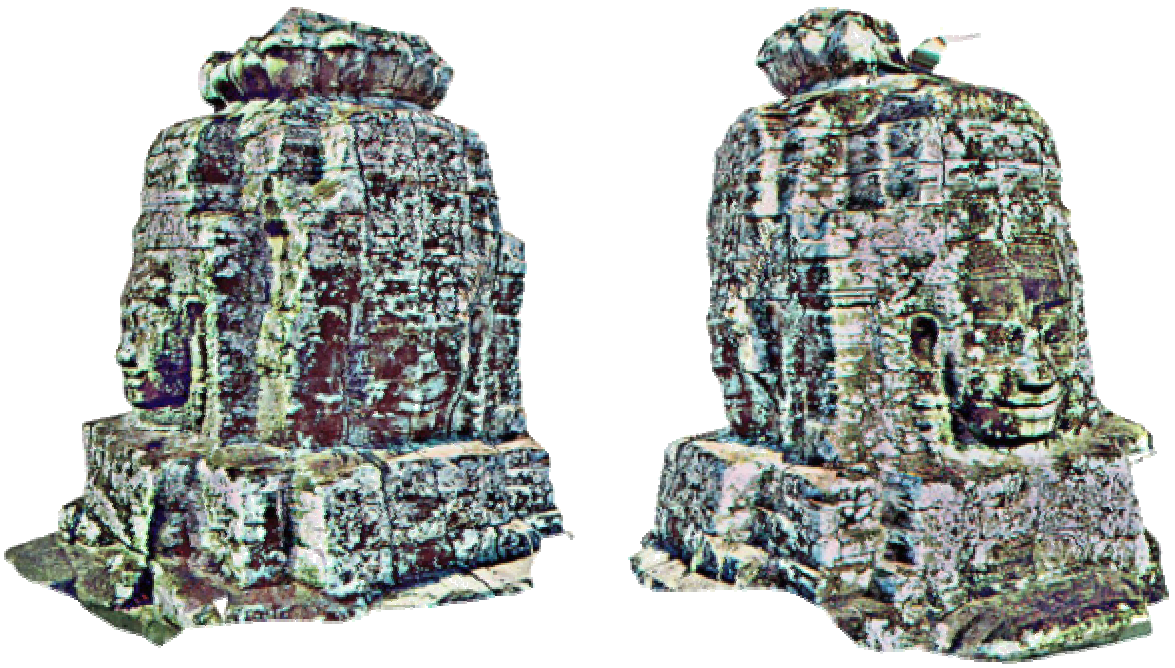
$$w_{ij} = \frac{area_j}{area_{\max}}. \tag{2}$$

Obviously, the weights are chosen proportional to the area of the image patches, which gives priority to the image patch with the better geometry. Before this weighted averaging can take place the different image patches are transformed by an affine transformation to the geometry of the "optimal" master image. The weighted averaging reduces the effects of radiometric differences in adjacent images.

The final textured model can be viewed with our own graphics program disp3D or it can be converted to VRML2 for viewing it with standard visualization packages. The results of view-dependent texture mapping are depicted in Figure11.



(a) (b)



(c) (d)

Figure 11: Texture mapped 3D model of the Bayon Tower: (a) southern view, (b) eastern view, (c) northern view, (d) western view. Display produced by disp3D.

8. CONCLUSIONS

We have shown that precise and very detailed complex photorealistic 3D models can be derived using tourist-type amateur photographs and automated processing techniques for image matching, point cloud editing and view-dependent texture mapping. The only manual steps were the establishment of fictitious fidutial marks for fixing the interior orientation and the phototriangulation. While the automation of the first procedure constitutes no problem, the automation of triangulation would require a denser sequence of images in order to simplify tie point measurement by image matching.

The geometric and texture quality of the models could be improved and the automated procedures would work more successfully if greater attention would be paid to a balanced illumination of all parts of the object. In the case presented here however, this would require the use of artificial lighting. Also, such cases of very explicit and complex 3D object models need certain modifications to be integrated into the matching algorithm (e.g. multi-images, geometrical constraints, edge matching, handling of occlusions), which were not available in this project.

Acknowledgement. This is to acknowledge the support given by the JSA (Japanese government team for Safegarding Angkor) project. Above all we would like to thank Prof. Sachio Kubo, Keio University, Tokyo for giving the third author the opportunity to participate in this campaign. Also appreciated is the great help which we received in the field from Prof. Kubo's young and ambitious team of surveying, climatology and economy students. Finally, our thanks go to Yoshito Miyatsuka for navigating and controlling the balloon even under the greatest difficulties, such that no harm

was done to the precious instrument. We also thank Nicola D'Apuzzo from our Institute for providing us with his software for visualization and editing of a 3D point cloud.

REFERENCES

Beyer, H. A., 1992. Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System. IGP ETH Zürich, Mitteilungen Nr. 51, 186 p.

Brown, D.C., 1971. Close-range camera calibration. Photogrammetric Engineering, 37(8), pp. 855 – 866.

Cohen I, Cohen L., Ayache, N., 1991. Introducing Deformable Surfaces to Segement 3D Images and infer differential structure. Technical Report of INRIA.

Fua, P., Sander, 1992. Reconstructing Surface from Unstructured 3D Point. Proceedings of the Image Understanding Workshop, San Diego, California.

JSA News, 1999. JSA Project No 4, Commemorating the Completion of the Conservation and Renovation of the Northern Library of Bayon. JICE, Tokyo, Japan.

Preston, D., 2000. The Temples of Angkor – Still Under Attack. National Geographic Magazine, August 2000, pp. 84 – 103.

White, P.T., 1982. The Temples of Angkor – Ancient Glory in Stone. National Geographic Magazine, Vol. 161, No. 5, pp. 552 – 589.