

Image-based Automated Reconstruction of the Great Buddha of Bamiyan, Afghanistan

Armin Gruen, Fabio Remondino, Li Zhang
Institute of Geodesy and Photogrammetry - ETH Zurich, Switzerland
agruen, fabio, zhangl@geod.baug.ethz.ch

Abstract

In the great valley of Bamiyan, Afghanistan, two big standing Buddha statues were carved out of the sedimentary rock of the region around the second to fourth centuries AD. The larger statue was 53 meters high while the smaller Buddha measured 35 m. The two colossal statues were demolished on March 2001 by the Taleban, using mortars, dynamite and rockets. After the destruction, a consortium was founded to rebuild the Great Buddha at original shape, size and place. Our group performed the required computer reconstruction, which can serve as the basis for the physical reconstruction. The work has been done using three different types of imagery in parallel and in this paper we present our results of the 3D computer reconstruction of the statue using automated approaches of image analysis.

1. Introduction

The town of Bamiyan, 200 km north-west of Kabul, was one of the major Buddhist centers from the 2nd century AD up to the 9th century. The town was situated in the middle of the Silk Route and it was a common meeting place for many ancient cultures and later tourists. In the north valley of Bamiyan, two big standing Buddha statues were carved out of the sedimentary rock of the region, at ca 2500 meters of altitude, between the 2nd and 4th century. The larger statue was 53 meters high while the smaller Buddha measured 35 m. They were cut from the sandstone cliffs and they were covered with a mixture of mud and straw to model the expression of the face, the hands and the folds of the robe. The two statues were demolished on 8 March 2001 by the Taleban, using mortars, dynamite, anti-aircraft weapons and rockets. The Buddhists, the world community, UNO and UNESCO failed to convince the Taleban to leave such works of cultural heritage and now they are nothing but piles of sandstone rubble and rocks. After the destruction, a consortium was established with the goal to rebuild the Great Buddha of Bamiyan at original shape, size and place. This initiative is lead by the global heritage Internet society New7Wonders [8], with its founder Bernard Weber and the Afghanistan Institute & Museum, Bubendorf (Switzerland), with its director Paul Bucherer. Our group has volunteered to perform the required computer reconstruc-

tion, which can serve as the basis for the physical reconstruction. No other technique (like laser scanner) can be applied for the reconstruction and only an image-based approach can be used. In the literature many procedural techniques have been described to model historical sites or monuments from images: often an handheld camera is used, the calibration is reduced to a minimum and the visual quality of the results is realistic. But when very low quality images are used or big baseline is present, these automatic approaches can easily fail.

In this paper we present our photogrammetric approaches and results of the automated computer reconstruction of the Great Buddha in Bamiyan.



Figure 1: The Great Buddha of Bamiyan before the destruction (left); the explosion of March 2001 (center) and the hole left in the cliff after the destruction (right).

2. Images of the Great Buddha

Our work is based on the use of three different types of imagery in parallel:

1. A set of images acquired from the Internet ('Internet images');
2. A set of tourist-type images acquired by Harald Baumgartner, who visited the valley of Bamiyan between 1964 and 1970 ('Tourist images');
3. Three metric images acquired in 1970 by Prof. Kostka, Technical University of Graz [7].

The computer reconstruction is performed with automatic and manual procedures and the three data sets are analyzed separately. The first two sets of images are used mainly to test our newly developed reconstruction algorithm while the third one is used to collaborate in the reconstruction project. Only the computer model generated with manual measurements on the metric images [6] will be used for the physical reconstruction of the statue, but in this paper we

report in particular the results obtained with automated approaches.

3. Photogrammetric reconstruction

The reconstruction process consists of phototriangulation (calibration and orientation of the images), image coordinate measurement (automatic matching or manual procedure), point cloud and surface generation, texture mapping and visualization. The phototriangulation is based on a photogrammetric bundle adjustment; therefore, as it is a non-linear problem, it needs initial approximation of the unknowns. Moreover, from a contour plot of the statue (scale 1:100) [7], we derived some control points used in the adjustment for the correct definition of the scale and the orientation of the coordinate system.

After the establishment of the adjusted image blocks, the 3D reconstruction of the statue was performed with automatic procedures on Internet, tourist and metric data sets, and with manual measurements only on the metric images. In all cases, Wallis filtering was applied to remove the low frequencies and enhance the texture information in the images.

3.1 The metric image data set

The metric images were acquired in August 1970 with a TAF camera [2] by Prof. Kostka (Technical University of Graz). The TAF (Terrestrische Ausrüstung Finsterwalder) is an old-fashioned photo-theodolite camera that acquires photos on 13×18 cm glass plates. The original photos were scanned by Vexcel Imaging Inc with the ULTRA SCAN 5000 at a resolution of 10 micron. The resulting digitized images resulted in a format of $16\,930 \times 12\,700$ pixels each (Figure 2).



Figure 2: The three metric images acquired by Prof. Kostka (the right sides are cut off here for display).

3.1.1 Calibration and Orientation

Concerning the interior orientation parameters, the focal length and the principal point of the TAF camera were known, as well as the acquisition procedure [2, 7]. The images were acquired in normal case, with a double baseline and a distance of ca 130-150 m from the statue. Using this information and some control points measured on the contour plot, we obtained the first approximations of the exterior orientation. The final orientation of the three images was achieved using a joint bundle adjustment with 35 tie points and six control points. The final standard deviation a posteriori of unit weight was 0.019 mm while

the average standard deviations of the object point coordinates located on the Buddha itself and on its immediate vicinity were $\sigma_{x,y} = 0.07$ m and $\sigma_z = 0.14$ m. These values are not representative for the reconstruction accuracy, which should be based on relative precision values, rather than on absolute ones.

3.1.2 Surface reconstruction with automatic matching

A multi-photo geometrically constrained (MPGC) least squares matching software package, developed at our Institute, was applied to the metric images [5]. The automatic surface reconstruction works in fully automated mode according to the following procedure (Figure 3):

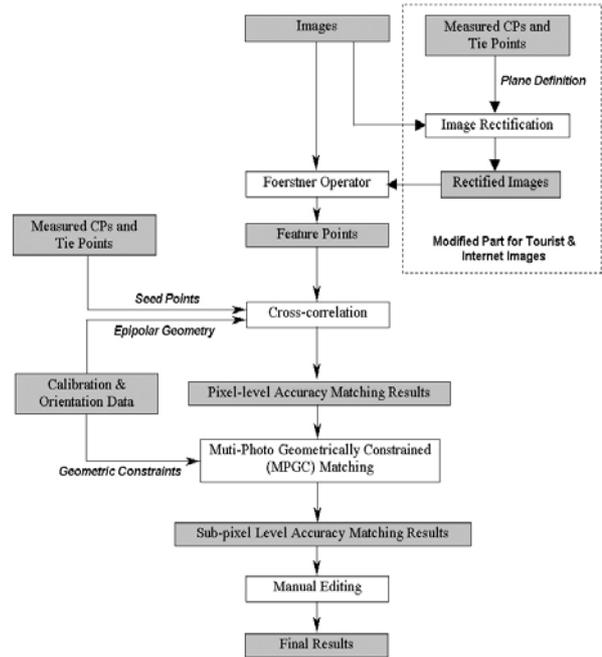


Figure 3: Data flow diagram of the matching procedure.

1. Selection of one image as the master image. In this case (metric images), the center image was selected.
2. Extraction of a very dense pattern of feature points in the master image using the Foerstner operator. At this stage, the master image was subdivided into 7×7 pixel image patches and within each patch only one feature point which gains the highest interest value was selected. In our implementation, the threshold for the Foerstner parameter roundness was set to 0.65 and the gray value variance of the image window was not allowed to drop below 5.0.
3. For each feature point, using the epipolar geometry determined by phototriangulation, we obtained the approximate matches for the following MPGC (Multi-Photo Geometrically Constrained) matching procedure by standard cross-correlation. The threshold of the normalized correlation coefficient was set to 0.70. The position and size of the search areas were determined by using the already known approximate surface model. In order to get this approximate surface, image pyramids and a matching strategy based on region growing,

which takes the already measured control points as seed points, were used.

4. In the final step MPGC matching was applied for fine matching, including patch reshaping parameters. MPGC exploits a priori known geometric information on orientation to constrain the solution and allows for the simultaneous use more than two images [1,3,4]. This algorithm can use blunder and occlusion detection, quality control and self-diagnosis.

In our implementation, for each feature point in the master image, all 3 metric images were employed for matching. Also, those, which failed with smaller patch size, will be re-matched with larger patch size in order to improve the matching success rate. Four different patch sizes (17×17 , 25×25 , 31×31 and 41×41 pixels) were used for the MPGC matching approach.

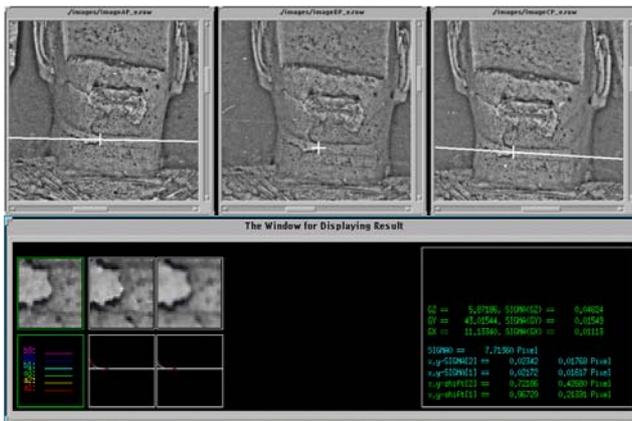


Figure 4: Epipolar geometry between the metric images to get the approximate values for the matching (upper row); MPGC matching results and computed 3D object space coordinates via forward intersection (lower row).

MPGC is a non-linear algorithm. Its linearization requires the procedure to be iterated. Considering the need to speed up the convergence rate, the geometric constraints were applied in a fully constrained mode in the first and second iterations. After that the weights of the geometric constraints were decreased in order to consider the errors in orientation data. To reduce the oscillations and divergence problems the reshaping parameters were constrained according to the image signal content in the image patches based on the analysis of the covariance matrix of least squares matching. With the MPGC approach, we can get sub-pixel accuracy matching results and 3D object coordinates simultaneously (Figure 4) and also, through covariance matrix computations, a good basis for quality control.

The procedure results in fairly reliable and precise matching results. From the three metric images, 49 333 points (without the surrounding rocks) and 73 640 points (with part of the surrounding rocks) were obtained. The point cloud data is shown in Figure 5, center. Although we use an automatic blunder and occlusion detection, some blunders are present in the final 3D point cloud and we removed them with a manual editing. As shown in Figure 5, there are some gaps in the point cloud, mainly due to the shading ef-

fects caused by the variation of the illumination conditions during the image acquisition (typical areas are marked in Figure 5).

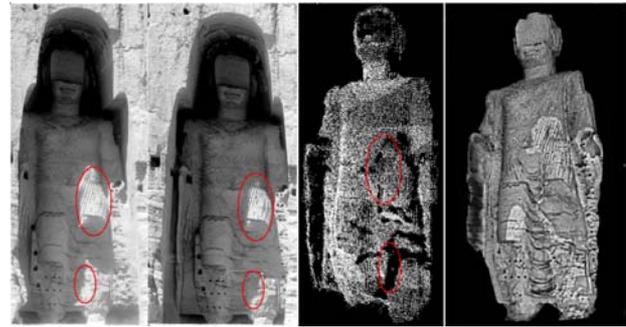


Figure 5: Left: two original metric images (part). Center: 3D point cloud generated by automatic matching on the metric images. Right: textured 3D model.

A textured model of the statue was also created and it is shown in Figure 5, right.



Figure 6: The folds on the dress of the Buddha had to be reconstructed with manual measurements.

The dress of the Buddha was rich in folds, which were between 10 and 15 cm in width (Figure 6). Many folds of the dress of the Buddha could not be reconstructed automatically. This is caused by the following reasons. Firstly, the image patches of least squares matching are assumed to correspond to planar object surface patches. Along the folds this assumption is not valid any more and these small features are smoothed out (Figure 7). Secondly, taking a very small patch size could avoid or reduce the smoothing effect theoretically, but may result not suitability for the determination of the reshaping parameters. Figure 7 compares the cases of small and large patch sizes over the folds of the dress. While a large patch tends to smooth out the local 3D information, a small patch may not include sufficient signal content for the matching. In our future work, edge-based matching will also be explored to reconstruct those very small features of the Buddha statue.

Figure 8 shows some of the problems that the automated matcher had to deal with: high noise values in the shadow areas, large differences between illuminated areas (left) and shadow areas (right), and artifacts (straight lines) from the photographic development process.

Therefore we used precise manual measurements to reconstruct the exact shape of the dress and other very small features of the statue and to provide accurate data for the future physical reconstruction [6].

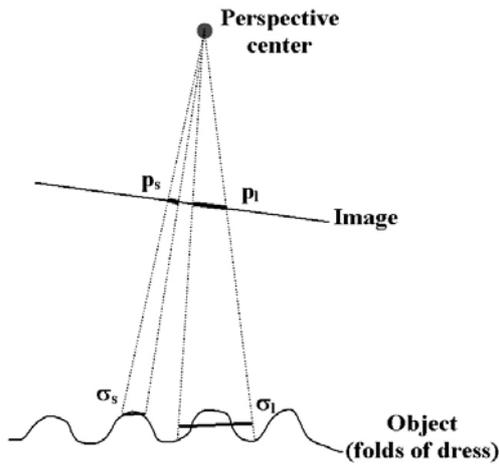


Figure 7: Patch definition in least squares matching.
 $p_s, p_l \dots$ image patches (small, large)
 $\sigma_s, \sigma_l \dots$ object planar patches (small, large)

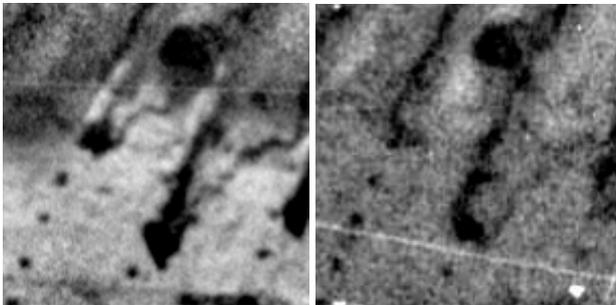


Figure 8: Corresponding left and right image patches, showing large differences in image content.

3.2 The Internet and tourist data sets

Out of five available tourist digital images (ca. $1\,840 \times 1\,232$ pixels), four were selected for processing (Figure 9). Instead, between the 15 images found on the Internet, four were selected for processing (Figure 10): two in front of the Buddha, one from the left side and one from the right side. All others were not suitable for photogrammetric processing because of very low image quality, occlusions or small image scale. The main problems with these two datasets are their differences in scale, their big baseline, the unknown original format, pixel size and interior orientation and most of all the different times of acquisition (Figure 11). Therefore some parts visible in one image are sometimes missing in others. Also, the illumination conditions are very different and this can create problems with the automatic matching procedures.



Figure 9: The tourist images used for 3D reconstruction.

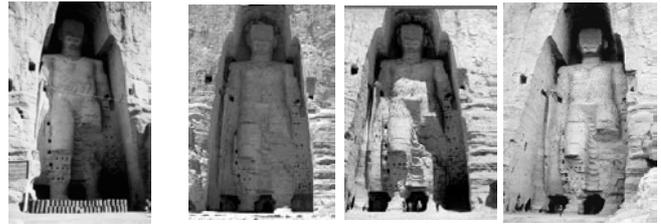


Figure 10: The Internet images used for the reconstruction.

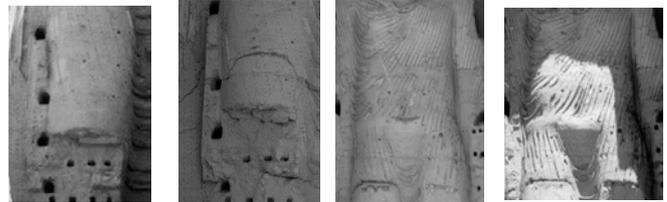


Figure 11: Changed of object details between images (left two) and changes of shadows (right two) between images.

3.2.1 Calibration and orientation

For the interior orientation, the pixel size and the focal length of each image were assumed. The principal point was fixed in the center of the images. With this last assumption, we considered the format of the found images as the original dimension of the photo, while they could be just a part of an originally larger image. The assumed pixel sizes are between 0.03 mm and 0.05 mm. Concerning the exterior orientation parameters, as no information was available, we first performed an interactive determination of the camera positions, varying also the value of the focal length and using some control points measured on the contour plot of Prof. Kostka. Then we refined these approximations with single photo spatial resection solutions.

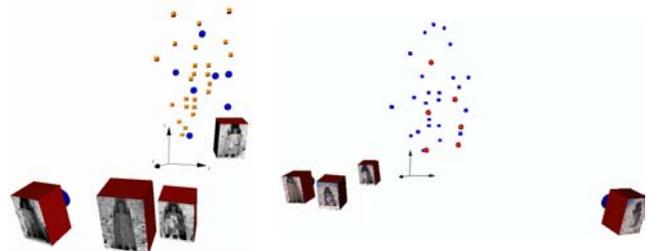


Figure 12: Restored camera poses of the tourist images (right) and the Internet images (left) after the bundle adjustment.

The final orientation of the two data sets of images was achieved using bundle adjustments while the required tie points between the images were measured semi-automatically with an Adaptive Least Squares matching algorithm [4]. The final average standard deviations of the object point coordinates located on the Buddha itself and on its immediate vicinity were $\sigma_{x,y} = 0.19$ m and $\sigma_z = 0.30$ m for the Internet images and $\sigma_{x,y} = 0.12$ m and $\sigma_z = 0.17$ m for the tourist one. Again these values are not representative for the reconstruction accuracy. The restored configuration of the cameras is shown in Figure 12.

3.2.2 Surface Reconstruction with automatic matching

The main problems of these images, which disturb the automatic matching procedure, are their differences in image size and scale, the different texture information due to the different times of acquisition and the different shadows, caused by different illumination conditions. Also the big baseline between the images can create some problems to common reconstruction algorithms. For this kind of images, feature-based matching may be a better solution, because matching of linear and point features can handle images taken from very different viewpoints or even include depth discontinuities. However, the results of feature-based matching are usually too coarse for complex surface reconstruction. The matching procedure for the tourist and Internet images is a modified version of that mentioned under section 3.1.2. The following paragraphs give a detailed description (see also Figure 3).

In the first step, the cross-correlation technique was used for the extraction of approximations for the MPGC matching. Normally the cross-correlation methods encounter difficulties when the image scale varies and if there are rotations between the images. In order to solve these restrictions, the so-called “rectified” images were used.

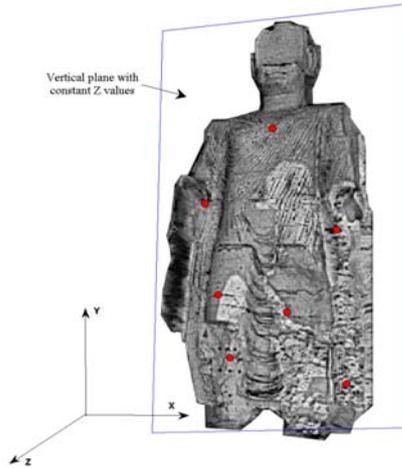


Figure 13: Definition of the vertical plane with constant Z_c , used for the image rectification.

The rectified images were generated according to the following procedure. (1) A vertical plane was defined to fit the control and tie points (Figure 13). Then each pixel of the original image is projected onto the vertical plane by using the known image orientation parameters. The mathematics relationship between the original and rectified images can be calculated by using the collinearity equations:

$$X_p = (Z_c - Z_0) \frac{r_{11}x_p + r_{21}y_p - r_{31}f}{r_{13}x_p + r_{23}y_p - r_{33}f} + X_0 \quad (1)$$

$$Y_p = (Z_c - Z_0) \frac{r_{12}x_p + r_{22}y_p - r_{32}f}{r_{13}x_p + r_{23}y_p - r_{33}f} + Y_0$$

$$x_p = -f \frac{r_{11}(X_p - X_0) + r_{12}(Y_p - Y_0) + r_{13}(Z_c - Z_0)}{r_{31}(X_p - X_0) + r_{32}(Y_p - Y_0) + r_{33}(Z_c - Z_0)} \quad (2)$$

$$y_p = -f \frac{r_{21}(X_p - X_0) + r_{22}(Y_p - Y_0) + r_{23}(Z_c - Z_0)}{r_{31}(X_p - X_0) + r_{32}(Y_p - Y_0) + r_{33}(Z_c - Z_0)}$$

where, Z_c is the constant value of the vertical plane; (X_0, Y_0, Z_0) is the position of the perspective center; $(r_{11}, r_{12}, \dots, r_{33})$ is the rotation matrix calculated from the attitude of images; f is the camera constant; (X_p, Y_p) and (x_p, y_p) are the pixel coordinates of the rectified and original images respectively.

(2) After generation of the rectified images, feature points are extracted in one rectified image and their conjugate points in other rectified images are computed by standard cross-correlation technique and matching strategy based on region growing. The already measured control and tie points are used as seed points. The threshold of the normalized correlation coefficient is set to a relatively low value (0.60).

This procedure results in ca. 8 700 match points for the tourist image data set. All of these points were converted to the original images by equation (2) and their object coordinates were also computed.

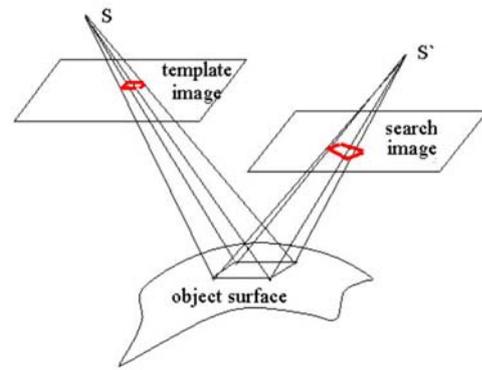


Figure 14: Calculation the initial reshaping parameters of MPGC. S, S' are the perspective centers of the template and search image respectively.

(3) The approximate matching results were refined by MPGC. This procedure was performed on the original images. For solving the problems caused by different image scale and rotation, the initial values of the reshaping parameters in MPGC matching can be predetermined by using the collinearity conditions for the four corner points of the image patches. The corresponding rays of these four corners in the template image should intersect in object space and their object coordinates can be determined (Figure 14). Through the collinearity equations, the corresponding image coordinates in the search image can be determined. The initial values of the four reshaping parameters were determined from these four points and their correspondences.

The matching procedure on the *tourist images* resulted in 5585 points. Some blunders were deleted by manual editing and the final point cloud and related textured model are shown in Figure 15. The relatively low image resolution of the tourist data set results in a coarse but quite complete 3D model.

Concerning the *Internet images*, four were used for the matching and a point cloud of ca 6 000 points was obtained (Figure 16, left). Some holes are present in the results because of surface changes (due to the different times of image acquisition) and little texture in some areas. For the conversion of the point cloud to a triangular surface mesh, a 2.5D

Delauney triangulation was applied. The final textured 3D model is shown in Figure 16, center and right.



Figure 15: Point cloud and textured 3D model, obtained from the tourist images by automated matching.

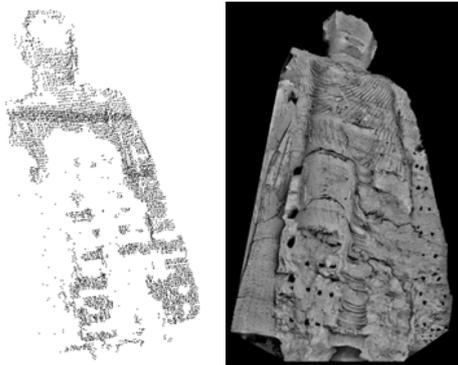


Figure 16: Point cloud and textured 3D model, obtained from the Internet images by automated matching.

4. Conclusions

The photogrammetric reconstruction of the Great Buddha of Bamiyan has been achieved successfully. We have produced various 3D computer models of different quality, all based on automated image measurements and using different types of images. The automated point cloud generation methods (with the Internet, tourist and the metric images) provide for dense point clouds (in presence of good texture), but they can fail to extract the very fine details of the statue, like the folds of the robe. Therefore only manual measurements allowed the generation of 3D models which are precise, reliable and complete enough to serve as a basis for a physical reconstruction. The problems encountered here with the orientation of amateur and tourist images and with the automated matching could be solved in an acceptable manner. The main difficulties of this project consisted in the transition from the point cloud to a surface model, which can satisfy high modeling and visualization demands. Some approaches usually cover the incorrect measurements with textures, but we preferred to do some man-

ual post-processing in order to obtain more accurate 3D data.

Acknowledgements

The authors would like to thank all the people who took part in this project and in particular: Prof. Kostka, Technical University of Graz, for the 3 metric images; Vexcel, for scanning the metric images; B. Weber, Founder of the New7Wonders society; P. Bucherer, Director of the Afghanistan Museum in Bubendorf, Switzerland; Robert Zanini and Joachim Wirth, Institute of Virtual Manufacturing, ETH Zurich, for the physical reconstruction of the statue at scale 1:200; Harald Baumgartner, Enzersdorf, Austria, for the "tourist images"; Tom Bilson, Courtauld Institute of Art, London, for some Internet images; all the web sites where we found images of the statues; Yuguang Li for the manual measurement on the metric images.

References

1. Baltasvias, E., "Multiphoto Geometrically Constrained Matching. Dissertation", IGP, ETH Zürich, Mitteilungen No. 49, 221 pages, 1991
2. Finsterwalder, S., Hofmann, W., "Photogrammetrie", De Gruyter Lehrbuch, Berlin, pp. 119-120, 1968
3. Grün, A., Baltasvias, E., "Geometrically Constrained Multi-photo Matching", *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 5, pp. 633-641, 1988
4. Gruen, A., "Adaptive Least Squares Correlation: A powerful Image Matching Technique", *South Africa Journal of Photogrammetry, Remote Sensing and Cartography*, 14 (3), pp. 175-187, 1985
5. Grün, A., Zhang, L., Visnovcova, J., "Automatic Reconstruction and Visualization of a Complex Buddha Tower of Bayon, Angkor, Cambodia", *Proceedings of 21th Wissenschaftlich Technische Jahrestagung der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung (DGPF)*, 4-7 September, Konstanz, Germany, pp. 289-301, 2001
6. Grün, A., Remondino, F., Zhang, L., "Reconstruction of the Great Buddha of Bamiyan, Afghanistan", *International Archives of Photogrammetry and Remote Sensing*, 34(5), pp. 363-368, Corfu (Greece), 2002
7. Kostka, R., "Die stereophotogrammetrische Aufnahme des Grossen Buddha in Bamiyan", *Afghanistan Journal*, Vol.3, Nr.1, pp. 65-74, 1974
8. New7Wonders Foundation, <http://www.new7wonders.com>