AUTOMATIC MULTI-IMAGE PHOTO-TEXTURING OF COMPLEX 3D SCENES

Y. Alshawabkeh, N. Haala
Institute for Photogrammetry (ifp), University of Stuttgart, Germany
Geschwister-Scholl-Strasse 24D, D-70174 Stuttgart
yahya.alshawabkeh@ifp.uni-stuttgart.de

KEY WORDS: Cultural Heritage, Photogrammetry, Laser Scanner, Visibility, Texture Mapping

ABSTRACT

The paper presents an approach for projective texture mapping from photographs onto triangulated surfaces from 3D laser scanning. By these means, the effort to generate photo-realistic models of complex shaped objects can be reduced considerably. The images are collected from multiple viewpoints, which do not necessarily correspond to the viewpoints of LIDAR data collection. In order to handle the resulting problem of occlusions, the visibility of the model areas in the respective images has to be established. For this purpose, the algorithm works in both image and object space and efficiently detects ambient, back-face and view frustum occlusions. Occluding polygons are labelled and separated with their connectivity to texture them recursively using the optimal of the available images until the final textured model is produced. After this visibility processing, colour values will be correctly assigned from the photograph to the visible polygons. In order to gain a high quality texture, lens distortion and colour corrections are applied during processing. The approach is demonstrated by generating high realistic 3D textured models for the Alkasneh monument in Petra city and a Romanian Theatre in the ancient city of Jerash.

1. INTRODUCTION

Cultural heritage applications frequently require data collection by terrestrial laser scanning in very complex structural environments. Thus, compared to similar applications in industrial environments, this requires more complex processing to generate geometric models of sufficient realism. In addition to geometric data collection, texture mapping is particular important in the area of cultural heritage to have more complete documentation. Photo-realistic texturing can for example add information about the structure condition, which is not present in the 3D model such as decay of the material. Additionally, color image information is also indispensable for features like frescos and mosaics. In addition to that, texture mapping considered as a requirement application for visualization and animation purposes.

For this reason, some commercial 3D systems already provide model-registered color texture by capturing the RGB values of each LIDAR point using a camera already integrated in the system. However, these images frequently are not sufficient for high quality texturing, which is desired for documentation, since the ideal conditions for taking the images may not coincide with those for laser scanning (El-Hakim et al, 2002). In addition, laser scanning from many viewpoints, as it is required to capture complex structures, is still relatively time consuming. For outdoor applications these large time differences will result in varying light conditions and changing shadows, thus the recorded images will have different radiometric properties. Such problems may disturb the appearance of the resulting textured model. So it is therefore more useful to acquire geometry and texture by two independent processes and allow for an image collection at optimal position and time for texturing. This is especially true for the high requirements of realistic documentation of heritage sites.

In order to warp independent images onto the collected object surfaces using perspective or affine transformation projection, different approaches are available (Kada et al, 2003; Remondino et al, 2004; Visnovcova et al, 2001). However, such approaches can only be applied for simple objects with restricted number of surfaces with no occlusions. Otherwise, if direct projection transformation is applied without considering occlusions, the mapping between object geometry and image will be incorrect. Warping the image over the geometry associated each 3D point to a pixel in the color image, the texture pierce through the geometry and gets mapped onto all occluded polygons on the path of the projected ray. So the geometry that is occluded in the image will receive incorrect texture coordinates instead of remaining in shadow. Figure 1 shows this problem and defines the three types of occlusion problems as they are known in computer graphics applications; ambient, self and view frustum occlusions. The occlusion problem can be avoided by manual texture extraction and mapping, however, such tedious task can take up to several days for good results.

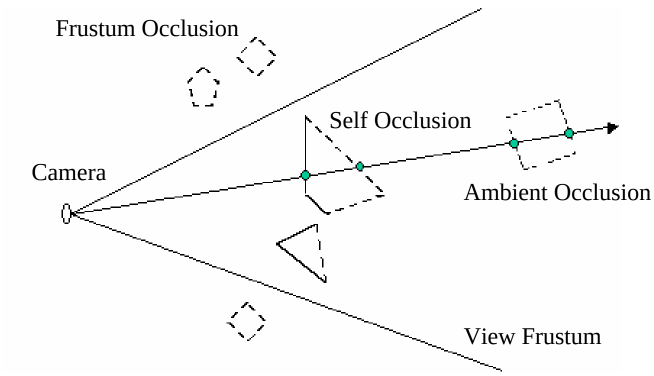

Figure 1. Visibility occlusions and double projection problem

Fully and semi-automatic texture extraction and placement of the real scene have been presented in different works (Grammatikopulos et al, 2004; Sequeira et al, 2000; Neugebauer et al, 2000) but with more complex processing. Those approaches use automatic techniques for occlusion detection. (Catmull, 1975) presents an image space Z buffer occlusion algorithm, the visibility is computed by comparing the surface depths at each pixel position on the projection plane. If the depth value is greater than the current value for the 3D point, the point associated to the previous depth value will be labeled as occluded. Compared to other occlusion algorithms, it is easy to implement but requires large memory and long computation time. (Grammatikopulos et al, 2004) speed up the searching process of the Z buffer algorithm by tessellating the textured image area into a rectangular grid with cells larger than those of the original image. In general, the Z_buffer algorithm shows quantization errors and bad performance as the complexity of scene is increasing (Sutherland, 1974). On the other hand, object space visibility algorithms, which can be used alternatively, show high precision and good results in high complex scenes but they are computationally expensive and difficult to implement. One example, which is very widely used, is the painter's algorithm. First, the algorithm sorts polygons by z-coordinates in object space. If the polygon's extents overlap, the overlapped triangles need to be spitted, then the polygons are sorted and rendered in back to front order (Martens, 2003).

If multiple images are used for texture mapping, calibration of the camera and image radiometric correction is required to guarantee the quality of the mapping. Otherwise artefacts such as line discontinuity will be visible at common edges of adjacent triangles mapped from different images. Geometric distortion resulting from incorrect camera calibration and orientation can be minimized by lens distortion correction and accurate image registration (El_Hakim et al, 2003). Additionally, radiometric distortions occur from different sensed brightness between different camera positions. To create seamless texture maps blending methods have been developed (Baumberg, 2002). Those methods use the weighted average of pixel values of triangles in the overlapping sections between two adjacent textures.

In this paper we will present an effective automatic approach to the problem of high-resolution photo-realistic texture mapping onto 3D complex model generated from range images. Our approach allows taking the images at different time from laser scanning and at whatever locations that will be the best for texturing. Through the approach we will introduce our visibility algorithm for occlusion detection and image fusing to have a realistic final model. For this purpose, the algorithm works in both image and object spaces. The main characteristics of this algorithm are the ability of detecting all types of occlusion problems; ambient, self and view frustum occlusions. In addition to that, it is flexible to allow the visibility detection of different resolution of geometric and 2D images.

Within the paper the presented approaches are demonstrated in the framework of a project aiming at the generation of a 3D virtual model of the Al-Khasneh, a well-known monument in Petra, and the Romanian Theatre in ancient Jerash city, Jordan. In section 2 the collection of the relevant image and LIDAR data is discussed. Data pre-processing and our approach for texture mapping will be presented in section 3, the pre-processing is mainly required in order to have perfect coregister laser and image data in order to gain high quality mapping. Section 4 contains the result and conclusion.

2. DATA COLLECTION AND PREPROCESSING

The collection of the data, which has been used for our investigations, was performed in cooperation with the Hashemite University of Jordan. One of the project goals is the generation of a 3D documentation of the Al-Khasneh monument in Petra city, and the north Theatre of ancient Jerash city.

2.1 Al-Khasneh Monument

The ancient Nabataean city of Petra has often been called the eighth wonder of the ancient world. Petra city in southwestern Jordan prospered as the capital of the Nabataean empire from 400 B.C. to A.D. 106. Petra's temples, tombs, theaters and other buildings are scattered over 400 square miles, these architectures are carved into rose-colored sandstone cliffs. After a visitor enters Petra via Al-Siq, a two-kilometer impressive crack in the mountain, the first facade to be seen is Al-Khasneh, which is considered as the best-known monuments in Petra city. The Al-Khasneh facade is 40m high and remarkably well preserved, probably because the confined space in which it was built has protected it from the effects of erosion. The name Al-Khasneh, as the Arabs call it, means treasury or tax house for passing camel caravans, while others have proposed that the Al-Khasneh Monument was a tomb. Behind the impressive facade of Al-Khasneh, large square rooms have been carved out of the rock (Sedlaczek, 2000).

2.2 North Theater of Jerash City

Jerash, formerly known as Gerasa, is considered to be one of the best-preserved and most complete Roman cities in the world. It was a member of the Decapolis group of cities; one of the ten biggest Roman cities. There are a large number of striking monuments: Gates, Arches, temples, colonnaded street, forum, theatres, baths, Byzantine churches and an almost complete circuit of city walls. The North theatre, which we used for exemplarily data collection was inaugurated in 165 A.D. and was extended during the 3rd century to bring the number of seats up to about 1600. It was used as an artisan workshop for the production of pottery. (Belloni, 2000).

2.3 Sensors Applied And Measurement Configuration

Depending on the application and the complexity of the environment, achieving high geometric and realism 3D model may require a large number of 3D and 2D images. In our project, the 3D laser scanning system GS100, manufactured by Mensi S.A., France was applied. The scanner features a field of view of 360° in the horizontal and 60° in the vertical direction, enabling the collection of full panoramic views. The distance measurement is realized by the time of flight measurement principle based on a green laser at 532 nm. The scanning range of the system allows distance measurements between 2 and 100 meters. The scanner's spot size is 3 mm at a distance of 50 meters; the standard deviation of the distance measurement is 6 mm for a single shot. The system is able to measure 5000 points per second. During data collection a calibrated video snapshot of 768x576 pixel resolution is additionally captured, which is automatically mapped to the corresponding point measurements. Because it is not possible to have a complete 3D coverage for outdoor complex structured sites based on data collected from a single station, different viewpoints have to be done to resolve the occlusion. As it can be seen in the top row of figure 2, the different acquisition time of the images results in considerable differences in brightness and colour. This will definitely disturbe the appearance of the textured 3D model. For this reason, additional images were collected by a Fuji S1 Pro camera, which provides a resolution of 1536x2034 pixel with a focal length of 20 mm. These images depicted in the bottom row of figure 2 were collected at almost the same time.

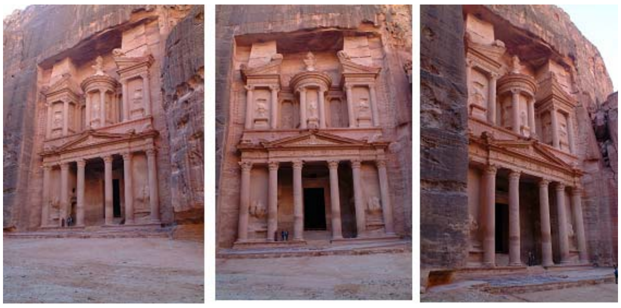

Figure 2. In the first row Mensi laser scanner images (768x576), second row shows images collected using Fuji camera (2304x1536), the mid image is the master image used for texture mapping.

For Alkasneh three different viewpoints with five laser scans were collected to resolve the occlusions. The problem to choose the viewpoint positions represents an important phase of the survey for such a monument since the mountainous environment surrounding Al-Khasneh restricts potential sensor stations. For covering the North Theater of Jerash and the surrounding environment seven scans were done.
All the acquired 3D models have been processed using Innovmetric Software, PolyWorks. Registration of the scans for both models was done using corresponding points. The produced models have an average resolution of 5 cm with more than 2 million triangles for Alkasneh model given in Figure 4, and 4.6 million triangles for modeling the Theater, given in figure 5.

Figure 3. North Theatre, Jerash

3. TEXTURE MAPPING

The purpose of texture processing is to integrate the 3D measurements from the laser scanner with 2D information taken with an external or internal camera. The reconstruction of texture maps from multiple views needs the information concerning the positions of a pixel in texture space in all images. This information is present in the transformation from texture to model and further on to image space. In our approach the warping is defined by standard co-linearity equations. Because the triangles are relatively small, no perspective correction was applied, and the original high-resolution 2D images are used for processing. Before performing projective texture mapping, we need to compute visibility information to map the image only onto the portions of the scene that are visible from its camera viewpoint. In order to guarantee high quality texture mapping, preprocessing steps for generating geometric and radiometric distortion free images is necessary. Otherwise artefacts such as line discontinuity will be visible at common edges of adjacent triangles mapped from different images. In this section an overview of the procedure with its different steps will be presented exemplarily for the Alkasneh model.

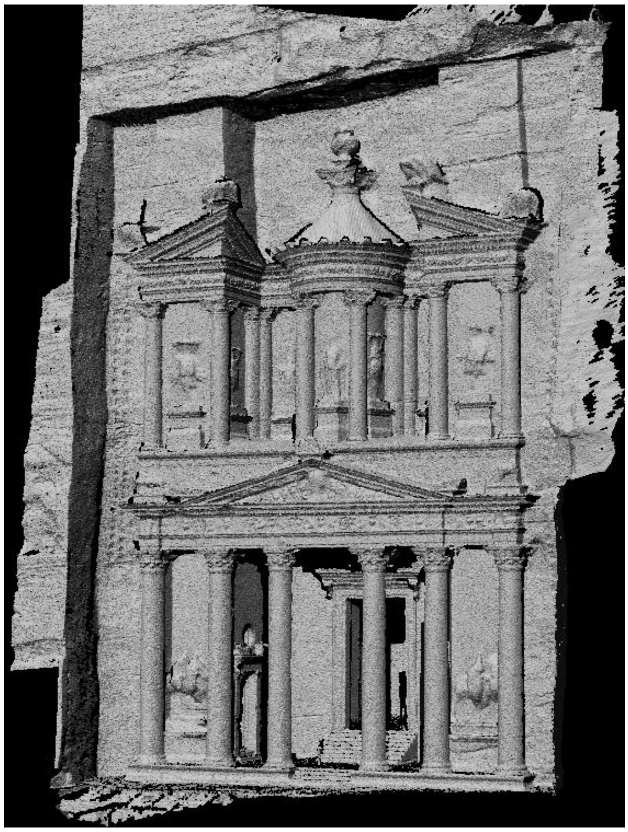

Figure 4. 3D model of Al-Khasneh

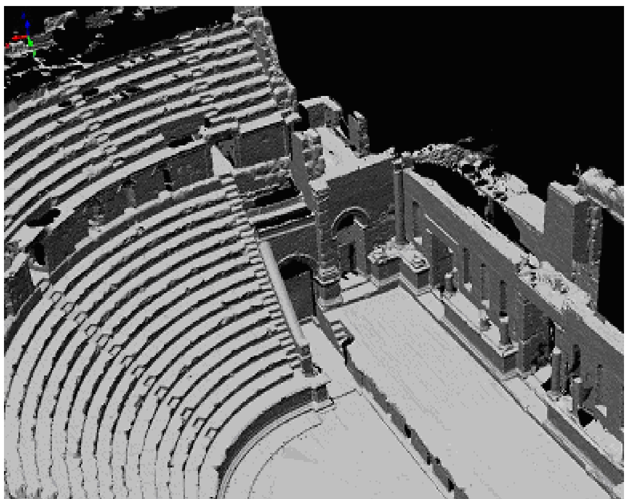

Figure 5. 3D model of the North Theatre of Jerash city

3.1 Camera Calibration and Data Co registration

Distortion free images and high quality of the registration process are crucial factors for getting a realistic looking nature of the model. This requires an accurate determination of the cameras interior and exterior orientation parameters. For our investigations the interior orientation parameters were computed using a calibration computed by the Australis software. These parameters were used to creating zero distortion images as intermediate step. Corresponding coordinates between image and laser scanner were measured using the photomodeler software, which was also applied to calculate the camera position and orientation in the coordinate system of the geometric model.

3.2 Colour Correction

Different illumination conditions prevent color continuity at the borders of each image, leading to observable discontinuity in the color and the brightness. For high realistic 3D texture mapping the artifacts due to illumination changes between photographs have to be removed. We use ENVI 4.1 software to perform the histogram equalization for all three channels in all textured images in order to diffuse the texture.

3.3 Summary of the Overall Procedure

During texture mapping every point on the surface of the object (in x, y, z coordinates) must be mapped to some point (in u, v coordinates) on the texture image. If no special care is taken, then erroneous pixel values are extracted for the occluded parts of the scene as it can be seen in figure 6. To avoid such artifacts, occlusion detection parts related to the novel image is realized. In our approach invalid pixels that belong to the shadow polygons will be identified and marked. The occluding parts will be masked out and be used as input model for the second selected textured image, the process for checking visibility and separating will be repeated automatically until the model is textured from all the available images. For the procedure, the following input data are required:

a) Triangulated 3D mesh in the form of successive XYZ, describing the object surface.
b) Calibrated colored images along with their exterior orientation parameters.
c) Two thresholds values: object space threshold (T1) represents maximum length of the triangle edge in the scene space (sampling interval), and image space threshold (T2) represents the maximum number of pixels of triangle edge projected onto the image.

One advantage of using thresholds values is the simple visibility detection for different resolutions of 3D geometric model or 2D images. Although the algorithm shows low sensitivity, choosing smaller threshold values may wrongly classify some visible triangles as occluded. Whereas choosing larger thresholds may classify the occluded triangle as a visible one. Our overall procedure for texture mapping summarized in Figure 7.
The different steps are:

1- Select most appropriate image that depicts most parts of the object (master image). Such selecting will speed up the further processing steps.
2- Given the exterior orientation parameters, the texture coordinates associated to each polygon vertex can be located using the co-linearity equations.

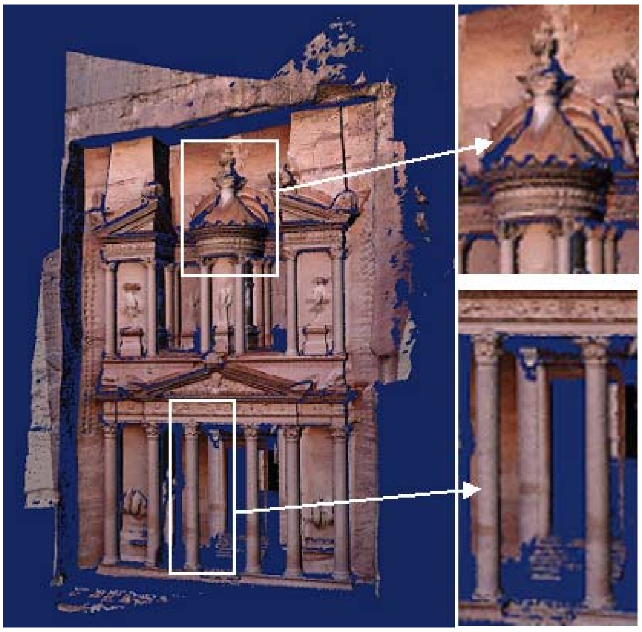

Figure 6. Double projection problem

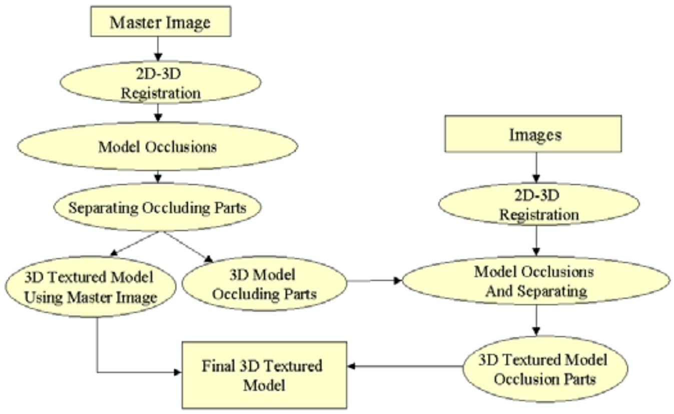

Figure 7. Texture mapping procedure

3- Check for frustum occlusion to exclude texturing the coordinates that are not within the range of the image. Store the other texture coordinates in a matrix with the Z value of the corresponding vertex.
4- In image space define searching area using threshold (T1), search and sort all the triangles within the specified area, then select the nearest triangle dependent on the Z value and the (T2), give an ID for such polygons as an occluder one.
5- The occluded polygons vertices will take a zero textured values and masked out from the original model, see figure 8.
6- The separated parts will be used as input model for the second selected textured images. The process for visibility and separating of occluding parts will be repeated automatically until the model is textured from all the available images as it can be seen in Figure 9. Then separated parts merge with the master image part to have the final model, see figure 10.

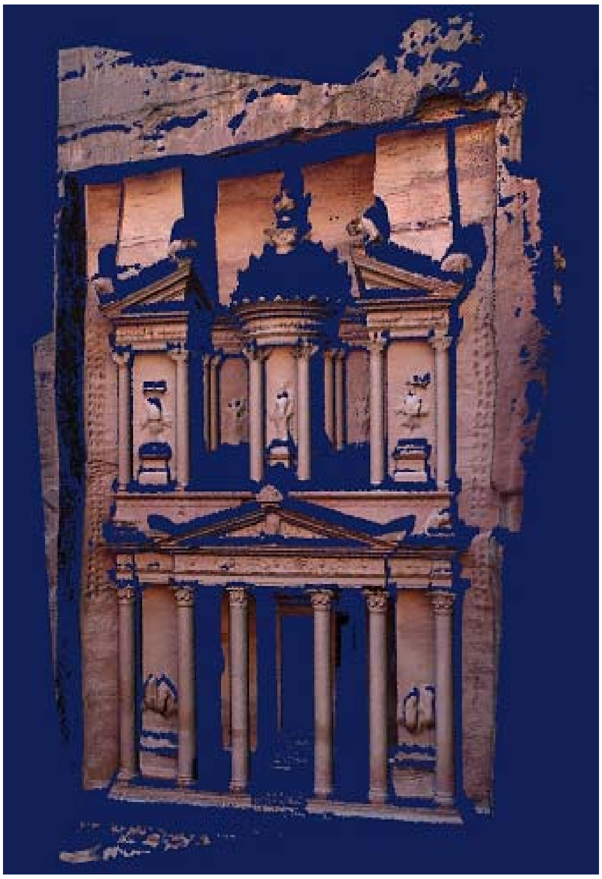

Figure 8. 3D Textured model using master image after separating of the occluding parts

One crucial point during the separating process is to maintain connectivity of the polygons; the coordIndex of the model. This is realised by preserving the ID of each polygon during all processing steps. For Alkasneh texture mapping, we have used five coloured images 2304 x 1536 pixels for texturing 2 millions triangles. The performance analysis has been conducted on a standard PC with an Intel 4, 3GHz Processor, and 1GB RAM. In our approach VRML modelling language is used. The approach has been also demonstrated for texture mapping the North Theatre of Jerash city. Figure 11 depicts a part of the model where four images were used for texture mapping.

4. SUMMARY AND CONCLUSION

The work described in this paper was developed as a part of an ongoing project, which aims to underline the necessity to integrate image-based measurements and laser scanner techniques in order to optimize the geometric accuracy and the visual quality of 3D data capture for historical scenes. In our paper, we presented a new automatic technique to enhance the 3D geometric model of real world objects with texture reconstructed from separate sets of photographs. The purpose is to generate photo-realistic models of complex shaped objects with minimal effort. Before performing projective texture mapping, we compute visibility information to map the image only onto the portions of the scene that are visible from its camera viewpoint. For this purpose an efficient algorithm addressing the image fusion and visibility is used. One advantage of our algorithm that it is easily allows the visibility detection of different resolution of 3D geometric models and 2D images. The approach has been demonstrated for texture mapping different heritage sites in Jordan, the results are extremely realistic 3D models.

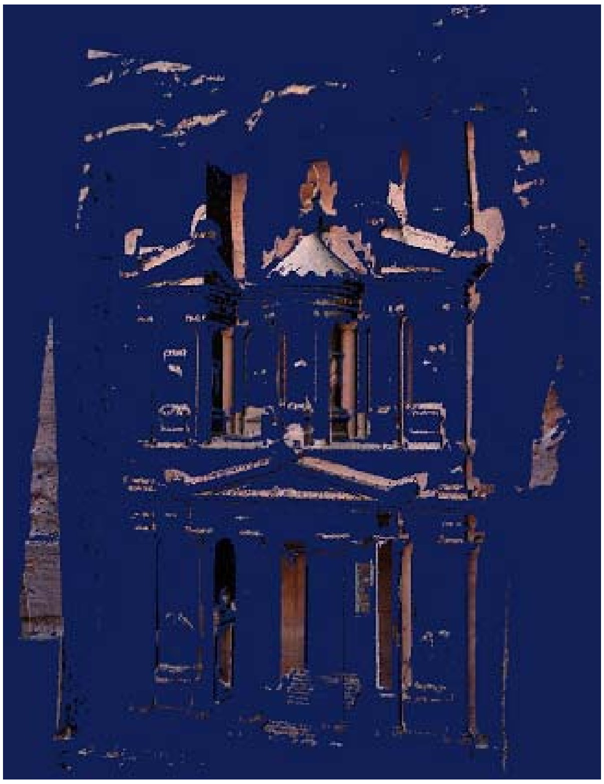

Figure 9. Texturing 3D model of occluding parts using 4 images

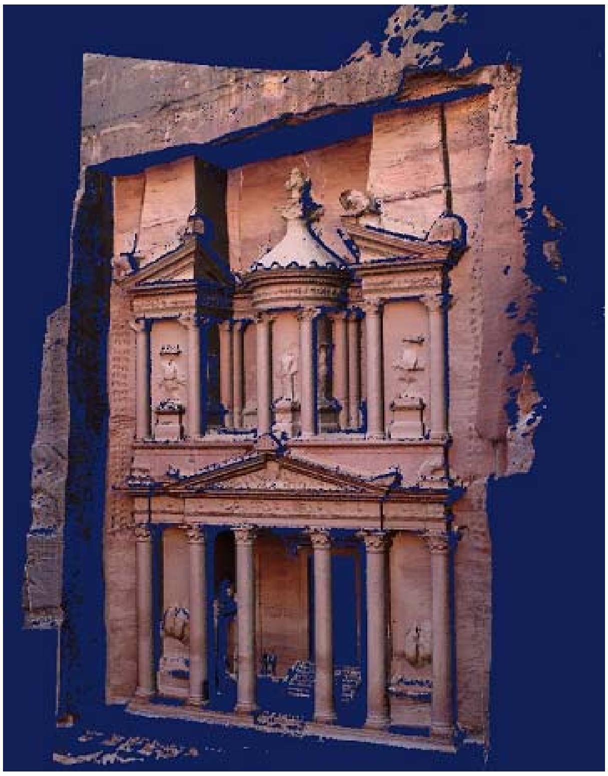

Figure 10. Final textured model of Alkasneh using 5 images

Figure 11. Part of the textured model of the North Theatre using 4 images

ACKNOWLEDGEMENTS

Special thanks to the Hashemite University of Jordan, Petra Region Authority and Jerash Municipality for support during the data collection.

REFERENCES

Baumberg, A., 2002. Blending images for texturing 3D models. Proc. Conf. on British Machine Vision Association, Cardiff, pp 404-413.

Belloni,S.,, B., 2000. Jerash, The Heritage of Past Cultures. Plurigraf, Terni, Italy.

Catmull, E.E. 1975. Computer display of curved surfaces. In

Proceedings of the IEEE Conference on Computer Graphics,

Pattern Recognition, and Data Structure, 11–17.

El-Hakim, S., Beraldin, A., and Picard, M., 2002. Detailed 3D Reconstruction of Monuments Using Multiple Techniques. Proceedings of the CIPA WG 6 International workshop, Corfu, Greece, pp.58-64.

El-Hakim, S., Gonzo, L., Picard, M., Girardi, S., Simoni, A. 2003. Visualization of Frescoed Surfaces: Buonconsiglio Castle-Aquila Tower, "Cycle Of The Months"Proceedings of International Workshop on Visualization and Animation of Reality-based 3D Models, Tarasp-Vulpera, Switzerland. February 24-28, 2003. NRC 44995.

Grammatikopoulos, L., Kalisperakis, I., Karras, G., Kokkinos, T., Petsa,E., 2004. Automatic Multi-Image Photo-texturing of 3D Surface Models Obtained With laser Scanning. CIPA International Workshop on "Vision Techniques Applied to the Rehabilitation of City Centres", Lisbon, 25-27 October.

Kada, M., Roettger, S., Weiss, K., Ertl, T. and Fritsch, D., 2003.

Real-time visualisation of urban landscapes using open-source

software. Proceedings of ACRS 2003 ISRS, Busan, Korea.

Sutherland, I., Sproull,R., and Schumacker,R.,1974. A Characterization of Ten Hidden-Surface Algorithms", Computing Surveys, Vol 6, No 1.

Neugebauer, P., Klein, K., 1999. Texturing 3D Models of Real. World Objects from Multiple Unregistered Photographic Views Proc. Eurographics'99.

Remondino, F., Niederoest, J., 2004. Generation of High-Resolution Mosaic for Photo-Realistic Texture-Mapping of Cultural Heritage 3D Models. proceedings of the 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST). K. Cain, Y. Chrysanthou, F. Niccolucci, N. Silberman (Eds), pp. 85-92 .

Sedlaczek, B., 2000. Petra, Art and Legend. MP Graphic Formula, Roma, Italy, pp.4-20.

Sequeira, V., Gonçalves, J.,2002. Reality Modelling: Photo-Realistic 3D Models of Real World Scenes, Proc 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT 2002), Padova, Italy, pp 776-783, June 19-21.

Visnovcova, J., Gruen, A., Zhang, L., 2001. Generating a 3D Model of a Bayon Tower Using Non-metric Imagery, Proc. Intern. Workshop Recreating the Past– Visualization and Animation of Cultural Heritage, pp. 30-39, Ayutthaya, Thailand.