OCCLUSION-FREE IMAGE GENERATION FOR REALISTIC TEXTURE MAPPING

Diego Ortin[1], Fabio Remondino[2]

[1]Department of Computer Science and System Engineering
University of Zaragoza, Spain
e-mail: dortin@unizar.es

[2]Institute of Geodesy and Photogrammetry
ETH Zurich, Switzerland
e-mail: fabio@geod.baug.ethz.ch

ABSTRACT:

Photo-realistic 3D models are nowadays required in many applications. The 3D modelling pipeline can be image- or range-based and often ends up with a visualization of a virtual textured model. One of the main problems encountered in texture mapping is the disturbance in the images by occlusions, which do not allow the generation of photo-realistic textured 3D models. Occlusions can be moving (e.g. pedestrians) or static (e.g. monuments) objects, which occlude the full and free visibility of the surface to be textured. In this paper we present the generation of occlusion-free artificial view for the realistic texturing of 3D models. Images of planar objects as well as complex 3D objects are considered.

KEYWORDS: Surface modelling, Visualization, Photo-realism, Texture, Rectification

1. INTRODUCTION

3D modelling from images is still a great topic of investigation in the research community, even if range sensors are becoming more and more a common source and a good alternative for 3D information. 3D image-based modelling should be meant as the complete process that starts from the data acquisition and ends with a virtual model in three dimensions visible interactively on a computer. Different applications require 3D models, like traditional inspections and robotics applications to the recent interest for visualization, animation and multimedia representation. The involved fields vary from cultural heritage to movies production, from industry to education. Of course the requirements for 3D modelling change according to the applications: most of them require only good visual quality of the digital product, city modelling has higher demands while medicine and industrial applications need very accurate measurements. 3D models are present everywhere in the actual society, their use and diffusion are becoming very popular, also through the Internet and even low-cost computers can display them. But, although it is very easy to create a simple 3D model of an object, the generation of precise and photo-realistic results of complex scenes still requires great modelling efforts.
After the reconstruction of the geometric model image data are often projected onto the geometry to achieve realism. The image information (texture) can be the same used for the geometric modelling or it can come from different acquisition sessions (e.g. external camera in case of laser scanner, terrestrial images in case of aerial city modelling).

2. TEXTURE MAPPING

Once a 3D digital model of a scene is created, with the technique of texture mapping, grey-scale or true colour images are mapped onto the 3D geometric surface in order to achieve photo-realistic virtual models. Knowing the parameters of interior and exterior orientation of the cameras, to each triangular face of the 3D surface the corresponding image coordinates are calculated. Then the grey-scale or colour RGB values within the back-projected triangle are attached to the face. Generally texture mapping is performed using a frontal image for a related part of the object. Unfortunately in close-range applications this is often not satisfactory, due to varying light conditions during image acquisition and not enough image information for fully or partially occluded object parts. Furthermore, different factors affect the photo-realism of a textured 3D virtual model:

- Image radiometric distortion: this effect comes from the use of different images acquired in different positions, with different cameras or in different daily moments (e.g. varying light conditions). Therefore in the 3D textured model discontinuities and artefacts are present along the edges of adjacent triangles textured with different images. To avoid these, blending methods based on weighted functions can be used.
- Image dynamic range: digital images have always a lower dynamic range than the scene. Therefore bright areas are often saturated while dark parts contain low signal to noise (S/N) ratio. Radiometric adjustment can be performed with common image processing tools.
- Scene geometric distortion: this kind of error is generated from an incorrect camera calibration and orientation, an imprecise image registration, too large triangles mesh or errors in the surface reconstruction. All these sources do not preserve detailed contents like straight edges or big discontinuity changes of the surface. Accurate photogrammetric bundle adjustment, precise and complete camera calibration and polygons refinement must be employed to reduce or minimize possible geometric errors.
- Object occlusions: static or moving objects, like pedestrians, cars, monuments or trees, imaged in front of the modeled objects might cause lack of quality in the texturing. Occlusion-free images can be generated as described afterwards.

To overcome all these problems and achieve photo-realistic 3D models, different pre- and post-processing techniques as well as methods of realistic texture mapping were presented [Haeberli and Segal, 1993; Niem and Broszio, 1995; Havaldar et al., 1995; Dallas, 1996; Weinhaus and Devarjan, 1997; Wang et al., 2001; El-Hakim et al., 2003; Grammatikopoulos et al., 2004]. The main ideas are to use high dynamic-range images [Cohen et al., 2001], weighted functions related to the camera's viewing angle [Debevec et al., 1996; Pulli et al., 1998], accurate piecewise local registration of the original images [Rocchini et al., 2002], compute the response function, exposure and white balance from a sequence of images [Kim and Pollefeys, 2004], rectify the images with a mesh-wise affine transformation [Remondino and Niederost, 2004], combine optimal image patches for each back-projected triangle of the 3D model [Neugebauer and Klein, 1999; Bernardini et al., 2001; Visnovcova et al., 2001], performing pixel-wise ray-tracing based on 3D model and camera parameter [Mayr and Heipke, 1988; Butler, 1989; Foley, 1990] or correct overlapping texture areas computing a common lighting of the textures [Beauchesne and Roy, 2003].

In particular, for the texture mapping of building facades, a large number of approaches have been presented [e.g. Brenner and Haala, 1998; Coorg and Teller, 1999; Karner et al., 2001; Kada et al., 2003; Varshosaz, 2003; Boehm, 2004; Zhang and Kang, 2004; Ulm, 2005] and the results (geometry and texture) are becoming more and more realistic and impressive.
One great problem in texturing buildings façades is the presence of obstacles in front of the façade. Often is difficult to select viewpoints where the complete façade is visible, e.g. in case of narrow streets. Moreover, for stereo processing, at least two images are required, leading to more difficult situations.

The removal of image disturbances by partial occlusions of building facades is the major motivation of this paper.
In fact, unwanted objects, occlusions or obstacles are often present in the texture data, reducing the realism of the generated 3D models. Looking at the literature [Boehm, 2004; Zhang and Kang, 2004; Eisert et al., 2005; Ulm, 2005], we can summarize the successful methods for occlusion removal like this:

- background learning and subtraction [Toyama, 1999]: it needs a very long sequence of images, acquired from the same standpoint. These methods are also applied to people detection and identification in forensic CCTV.
- rectification of multiple images in image space and average estimation of the unoccluded pixel values [Boehm, 2004]. Because of the involved projectivity, it can be applied only to planar patches.
- manual retouch the artefacts/occlusions in a single image [Ulm, 2005].

We could also split the occlusions removal algorithms in methods requiring the 3D object geometry and camera parameters (ray tracing) [Debevec et al., 1998; Gan et al., 2004] and methods that do not require any 3D information [Levoy and Hanrahan, 1996; Boehm, 2004]. The latter requires a large number of images, the former is very computationally expensive and has problems near the objects boundaries.

Compared to other solutions, our approach for occlusion removal from planar surfaces (section 3) employs a simple method that relies only on the homography between adjacent images (projective transformation), generating a new virtual texture free of occlusions. The requiring number of images is strictly related to the quantity of occlusions present in the image and if they are static (e.g. monument) or in movement (e.g. pedestrians). Furthermore, the process can be performed with an automatic computation of the image correspondences required for the 8 parameters projective transformation.

In section 4, we also illustrate how a similar approach can be used to deal with general scenes, and determine the radiometric appearance of the three-dimensional structure via the combination of multiple affine transformations.

3. OCCLUSION REMOVAL FROM PLANAR SURFACES

When determining the textured appearance of a model from real images, it is often the case that there are partial occlusions. Although these occlusions could theoretically be detected and corrected once the three-dimensional structure and the position of the cameras are known, it is not always practical or feasible to recover a complete model of the scene.

For instance, there could be some occlusions due to complex, non-interesting objects (e.g. the leaves of a three) or dynamic ones (e.g. moving people or vehicles). Correctly modelling these effects is often too complex or time-consuming for texture mapping applications.

Below, we present a simple method to determine the real appearance of facades when there are partial, non-modelled occlusions. By exploiting the fact that most facades can be locally well approximated by planar facets, we can transfer the texture among different views using a projective transformation. This produces alternative, independent estimates for the appearance of each facade. Rather than using a direct average of these values, we exploit the redundancy in the data to robustly determine the most likely texture of the wall.

Although the proposed method is limited to planar features, it has many advantages when compared with ray-tracing algorithms: it requires no explicit reconstruction of the occluding elements nor of the plane itself, it is not computationally demanding, and it can be directly applied to non-calibrated or even unoriented images (e.g. those acquired with an amateur camera or not used in the photogrammetric restitution).

This section presents the methods used here to determine the appearance of a planar facade given a set of views. First, we review the projectivities generated by a plane and show how new, synthetic views of that plane can be generated from novel viewpoints. Then, we show how these estimations can be combined in a robust way to determine the most likely appearance of the facade.

3.1 Transfer of Planar Texture

It is well known that the coordinates of a three-dimensional point [X,Y,Z] are related to its imaged position [x,y] through the collinearity model (often extended by means of additional parameters to model possible systematic errors). In terms of projective geometry, the collinearity equations can be described by the Direct Linear Transform (DLT):

$$x = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \qquad y = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}. \tag{1}$$

generally with p_{34}=1.
When additionally the point [X,Y,Z] is contained in a particular plane [a,b,c,d], then the following equation holds:

$$aX + bY + cZ + d = 0. \tag{2}$$

Therefore, given an image point [x, y], equations (1) and (2) can be used to uniquely determine the three unknown object coordinates [X,Y,Z]. If s is a scale factor then,

$$\begin{bmatrix} sx \\ sy \\ s \\ 0 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ a & b & c & d \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \qquad (3)$$

and the object coordinates of a point can be computed as:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = H^{-1} \begin{bmatrix} sx \\ sy \\ s \\ 0 \end{bmatrix}. \qquad (4)$$

The combination of equations (3) and (4) allows to determine the image coordinates of one point [x_w,y_w] given the projection of the same object point in a different image [x,y]:

$$\begin{bmatrix} s_w x_w \\ s_w y_w \\ s_w \\ 0 \end{bmatrix} = H_w H^{-1} \begin{bmatrix} sx \\ sy \\ s \\ 0 \end{bmatrix} = \begin{bmatrix} M & n_{3\times 1} \\ p_{1\times 3} & q_{1\times 1} \end{bmatrix} \begin{bmatrix} sx \\ sy \\ s \\ 0 \end{bmatrix}. \qquad (5)$$

Neglecting the last row and the last column of matrix $H_w H^{-1}$, equation (5) results in:

$$\begin{bmatrix} s_w x_w \\ s_w y_w \\ s_w \end{bmatrix} = M \begin{bmatrix} sx \\ sy \\ s \end{bmatrix}. \qquad (6)$$

M is generally called 'transfer matrix' and its elements can be computed when the interior and the exterior orientation of the cameras are known [Hartley and Zisserman, 2000], or directly from equation (6), when at least four corresponding image points are identified. Therefore *M* represents a projective transformation between two images that can be derived making use just of image information. This allows forecasting the projection of a point in the plane without the need of an explicit 3D reconstruction of the camera position.

Using *M*, we can also predict the appearance of an image point according to the radiometric information of the picture. In the following, we will make use of these properties to determine multiple independent estimations for the appearance of a plane. We will also illustrate how these synthetic views can be robustly combined in order to deal with occlusions.

3.2 Robust Texture Detection

The projective transformation *M* of equation (6) codifies the geometrical transfer from points in the original image *I* to the synthetic one I_w.

When we want to determine the appearance of the synthetic image, the inverse mapping should be applied: given a grid of points in I_w, their grey (or RGB) values could be determined by searching in the corresponding positions of image *I*. Because of the linear properties of equation (6), it is simply the inverse of the *M* matrix that codifies this projective transformation.

In practice, the transferred points will not exactly correspond to the centre of the pixels in image *I*, so their estimated colour should be interpolated from those of their neighbouring pixels. We are using cubic interpolation to preserve edge details [McGlone et al. 2004], but alternative methods (such as the nearest neighbour, bilinear interpolation or splines) could also be applied.

When the texture coming from different views I_i (i=1…, n) is transformed according to its corresponding projective transformations M_1, M_2 ... M_n, there are multiple estimations for the appearance of I_{w1}, I_{w2} ... I_{wn}. Least Squares approaches would compute an average of these estimates, as theoretically it would improve the signal to noise ratio of the resulting image. However, the presence of blunders due to non-modelled errors (in many cases occlusions) could seriously corrupt the quality of the estimation.

In order to determine a robust estimation for the appearance of the transformed image plane, we use a median of the different estimates rather than a mean. So, if $I_{w1}(i,j)$, $I_{w2}(i,j)$ ... $I_{wn}(i,j)$ are the possible colour values for the pixel (i,j), the final value assigned to the image $I_w(i,j)$, corresponds to:

$$I_w(i,j) = \underset{(i,j)}{med}\left(\left[I_{w1}(i,j), \quad I_{w2}(i,j), \quad \ldots \quad I_{wn}(i,j)\right]\right). \qquad (7)$$

In contrast to the average, the median is hardly influenced by errors not modelled in the distribution, whatever their magnitude. In this way, the contribution of the occluding elements can be easily reduced or neglected.

Other approaches determine the final texture of the model doing a robust regression of the grey values, by means of an exhaustive search in the space of possible solutions [Boehm 2004]. However, given that all non-occluded estimations should have similar colour values, one valid solution can be directly found seeking for the median of all the estimates. This allows for a simpler implementation. In our experiments, we observed no significant differences between the final results achieved using any of the mentioned approaches.

From a theoretical point of view, three views would suffice to determine the occluded appearance of a plane and remove any occlusions, as the median will work with a rupture point of the 50%. In practice, the radiometric properties of the image may change within the exposures, or the occluding elements could be close to the plane, producing little parallax, or the may move. In these cases, more than three views would be required to observe the real appearance of the plane in at least half of the pictures.

Note that the proposed method relies on the fact that the projecting rays are different for every view. The method cannot deal with static occlusions when the projective transformation M_i effectively corresponds to a rotating camera: in this case all the projective rays are the same and static occlusions consistently appear in all the views.

4. OCCLUSION REMOVAL FROM COMPLEX 3D OBJECTS

The generation of occlusion-free images for complex three-dimensional objects (e.g. Figure 6) is a much more difficult problem. Knowing the 3D model of the imaged scene and the camera parameters, the occlusion can be removed with a pixel by pixel ray-tracing approach.

However, it is necessary an accurate restitution of the model, and often, a complete reconstruction of the occluding elements (even when the occluding parts are not always of interest). Moreover, computationally demanding techniques must be applied to determine the visibility of each pixel from the different viewpoints. In contrast, our approach works directly in image-space with a warping of texture, and does not require the parameters of the camera or three-dimensional model.

Similarly to the previous case, we consider a general scene as composed of multiple small patches, roughly planar. Although a projective transformation can fully describe the perspective deformation of any planar surface under general viewing conditions, an affine one is also valid, if it is applied only to a small patch of the image (the template least squares matching [Gruen, 1985] is based on the same principle). Therefore, with this slightly modification, the method presented before can also be used and applied to small image patches to produce an occlusion-free view of complex models.

In particular, one picture (real or synthetic) must be considered as the reference one and all the remaining ones as 'slave' images. Homologous areas (corresponding to planar or quasi-planar patches) are (manually) identified in the different pictures using only three corresponding points (the minimum to define a planar patch). The patches of the entire slave image are then transformed to the reference view and combined using a median of the different colour channels.

In the next section, the performance of the presented methods on planar and complex scenes is presented on real data.

5. EXPERIMENTAL EVALUATION

For the methods evaluation, we focus on cases where the nature of the occlusions makes it difficult to geometrically model them, or when the manual segmentation of the occluding parts in the original images is rather complex or impractical. After the generation of the occlusion-free images, a three-dimensional model can be textured with any of the methods mentioned earlier.

5.1 Planar surfaces

Three possible cases are presented: recovering the appearance of a façade when the occlusions cover an important part of the visible area and there is small parallax between the occlusions (Figure 1), the automatic removal of multiple and complex occlusions when only a few views are available (Figure 3) and the occlusions removal when the reference planar surface is not fully visible in all the available images (Figure 4).

Figure 1-(a, b, c) depicts three sample views of a façade (ca 30 m long and 20 m height) partially occluded by some moving flags and their corresponding poles. As the poles are close to the facade, there is little parallax among the views. Moreover, as the flags occlude a significant part of the façade, several views with a large baseline had to be considered. In particular we used eighteen images to derive the clear appearance of the façade.

After the selection of a reference image (Figure 1-(d)), corresponding points were manually marked in the views. Although four points would be enough to identify the parameters of the projectivity *M*, in many views some of the selected points would be occluded. By using eight points in the reference view, we were always capable of identifying at least four of them in the other images, and therefore able to compute the parameters of the projectivity.

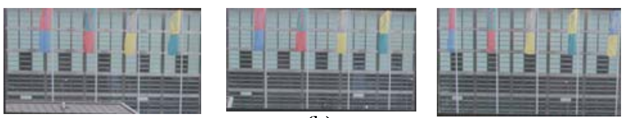
(a) (b) (c)

(d)

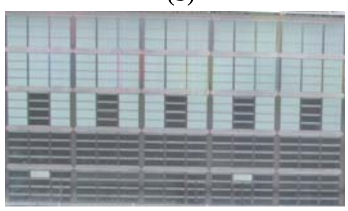
(e)

Figure 1: (a),(b),(c) sample views of the sequence; (d) reference view; (e) synthetic view in which the poles and the flags have been automatically removed.

The final occlusion-free image of the façade is shown in Figure 1-(e) and a closer view in Figure 2-(d). Figure 2-(a, b, c) shows a closer view of the rectified façades according to the images in Figure 1-(a, b, c). Note how the poles and the flags have been correctly removed, revealing the texture behind them. The base of the poles and part of the flag on the right (all of them outside the considered plane) could not be not removed.



(a) (b) (c)

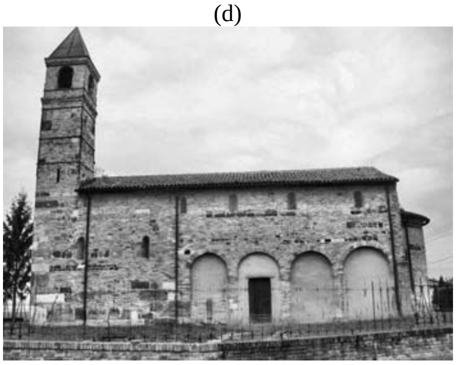


(d)

Figure 2: Appearance of the facade according to the views in Figure 1-(a), (b) and (c), and detail of the final texture of Figure 1-(e).

(a) (b) (c)

(d)



(e)

Figure 3: (a), (b), and (c) are the views used for the correction of texture; (d) corresponds to the reference view and (e) to the synthetic image in which most of the fence and the occluding person in (b) (bottom-left) have been removed.

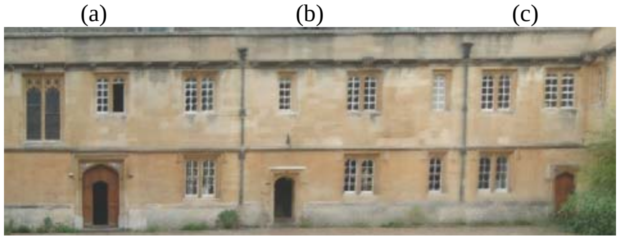
Figure 3 illustrates another example of occlusion removal. In this case the upper part of the facade of the building is free of occlusions, while in the lower part a person (in the central view) and a small fence occlude the church.

The geometrical modelling of the whole fence just to remove the occlusions using ray tracing algorithms is clearly impractical, whereas the manual editing of the images would be too time consuming.

As before, Figure 3 (d) and (e) show correspondingly the reference image and the synthetically generated one. Most of the occlusions were automatically removed, but there are still some traces of the fence in the final result. This could be due to the fact that the fence has a regular, repetitive pattern, and that only four images (a, b, c and d) where available to generate the occlusions-free image.

Finally, Figure 4 illustrates a third example in which a new view of a plane is rendered from a novel, frontal viewpoint without the occluding column. In none of the original views the plane is completely visible or free of occlusions. The closer views of Figure 5 show how not only large occlusions far of the plane (such as the column) but also some other small ones near the plane (such as the lamp over the door) were successfully removed.

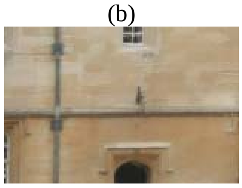
(a) (b) (c)



(d)

Figure 4: Sample views (out of five) of a building façade (upper row). Frontal view of the main façade in which the column has been removed (d).

(a) (b) (c)



(d)

Figure 5: Closer views onto small objects out of the façade plane removed in the new synthetic image (d)

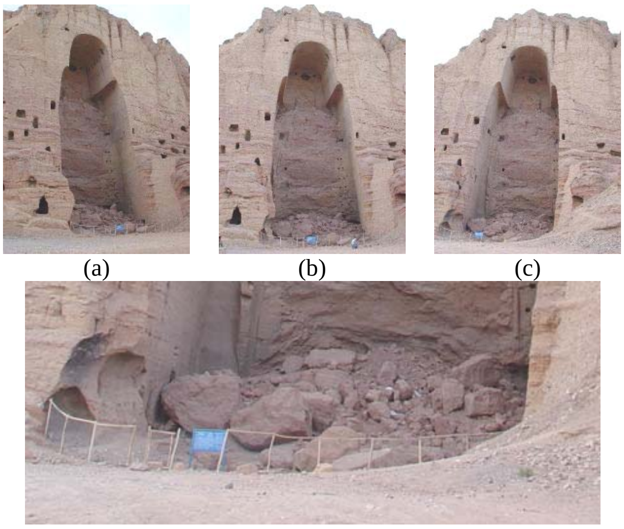
5.2 Complex 3D Surfaces

As shown in the previous results, the applicability of the method is restricted to large planar surfaces.

In practice, it is often the case that the modelled areas have non-planar elements that must be preserved in the final, textured model. An example is presented in Figure 6. The lower part of the object is occluded by a wooden fence together with a sign (Figure 6-d), which prevents the automatic mapping of the texture onto the reconstructed 3D geometry.
For the modeling of the cultural heritage object [Gruen et al., 2004] and the realistic mapping of the texture, the fence was eliminated with manual retouching in image-editing software (Figure 7).

As an alternative approach, we manually selected some corresponding points on the rocks behind the fence (when visible). The structure was then regarded as composed of multiple small planar facets, and the transformations relating these facets in the multiple views were codified by means of affine transformations.

The texture was then automatically transferred from the four ‘slave’ images to the reference one, without any explicit reconstruction, generating four synthetic images from the same viewpoint of Figure 5. These four images were then combined with the reference image as in section 3.2, generating the view of Figure 8.

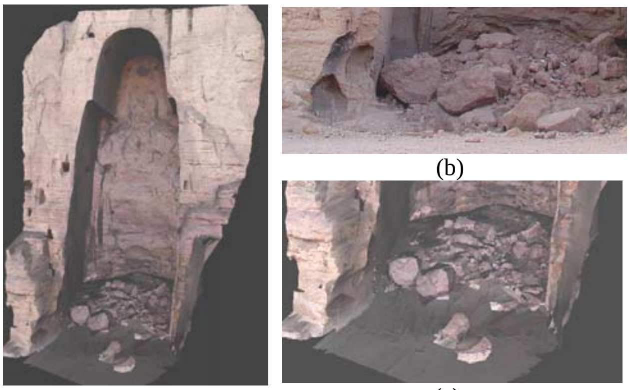


(a) (b) (c)

(d)

Figure 6: Three images (out of five) of the empty niche of the Great Buddha of Bamiyan, Afghanistan (a, b, c). A closer view on the lower part of the niche in the reference image, where the stones are occluded by a wooden fence and a big blue sign (d).

Figure 8 shows how most of the elements of the fence and the sign were successfully removed. Although small parts of the fence and the sign are still visible, the image can be directly used for texturing or visualization purposes. Probably, the availability of additional views (observing some 'dead-areas' behind the fence) could easily improve the quality of our results.



(b)

(a) (c)

Figure 7: The textured 3D model of the empty niche (a) and a closer view to the lower part without occlusions (b=image, c=3D model)

In this last example, the image points corresponding to the different facets where manually selected in the different views. However, they could also have been determined from the geometrical model (by back-projection) allowing for a fully automatic generation of the occlusion-free texture.

6. CONCLUSIONS

In this paper we present a method for the automatic recovery of the visual appearance of planar façades as well as complex 3D objects. By exploiting some assumptions on planarity, the texture of a building or of a small image patch can be transferred among different views without the need of calibrating or even orienting the images.

Figure 8: Occlusion-free image generated with a semi-automated method.

This provides us with alternative estimates for the appearance of the modelled surfaces. By exploiting the redundancy of these synthetic views via robust estimation techniques, occlusions can be easily identified and corrected.

We experimentally illustrated how the method performs well even in those cases with little parallax or when the occluding surface covers large areas of the facade. Moreover, our approach suggests a high potential for dealing with partial occlusions in complex, unstructured scenes.

However, we experimentally observed a decrease of its performances when the number of available views is too small (four or five images). More work is needed in order to refine the method and extend it to deal with more complex scenes, to automatically determine the minimal number of views needed for effectively removing a particular occlusion, or to mathematically quantify the probability that the computed texture corresponds indeed to the real one.

7. REFERENCES

Beauchesne, E. and Roy, S., 2003: Automatic Relighting of Overlapping Textures of a 3D Model. Proceedings of CVPR, pp. 166-173

Bernardini, F., Martin, I., Rushmeier, H., 2001: High-quality texture reconstruction from multiple scans. IEEE Trans. On Visualization & Computer Graphics, Vol. 7(4), pp. 318-332

Boehm, J., 2004 Multi-image Fusion for Occlusion-free facade texturing. IAPRS Vol. XXXV, part 5, pp.867-872.

Brenner, K. and Haala, N., 1998: Rapid acquisition of Virtual reality city models from multiple data sources. IAPRS, Vol., XXXII, part 5, pp. 323-330

Butler, T., 1989: Three approaches to terrain rendering. Proc. of SPIE, Vol. 1075, pp. 217-225

Cohen, J., Tchou, C., Hawkins, T., Debevec, P., 2001: Real-time High Dynamic Range Texture Mapping. Eurographics Rendering Workshop

Coorg, S. and Teller, S., 1999: Extracting textured vertical facades from controlled close-range imagery. IEEE Proc. of CVPR, pp. 625-632

Dallas, R., 1996: Architectural and archeological Photogrammetry. In Atkinson (Ed), 'Close Range Photogrammetry and Machine Vision', Whittles Publishing

Debevec, P., Taylor, C., Malik, J., 1996: Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach. SIGGRAPH 96 Conference Proceedings, pp. 11–20,

Debevec, P., Borshukov, G., Yu, Y., 1998: Efficient view-dependent image-based rendering with projective texture-mapping. Proc. of 9th Eurographics Rendering Workshop

El-Hakim, S., Gonzo, L., Picard, M., Girardi, S., Simoni, A., 2003: Visualization of frescoed surfaces: Buonconsiglio Castle – Aquila Tower, "Cycle of the Months". IAPRS, Vol. XXXIV, part 5/W10 (on CD-Rom)

Eisert, P., Rurainsky, J., Guo, Y., Höfker, U., 2005: Creation and Scene Composition for High-Resolution Panoramas. IAPRS, Vol. XXXVI, part 5/W8

Foley, J. D., 1990: Computer Graphics: Principles and Practice. Reading, Mass.: Addison-Wesley

Gan, Z.F., Chan, S., Ng, K.T., Chan, K.L., Shum, H.Y., 2004: On the rendering and post-processing of simplified dynamic light fields with depth. Proc. of IEEE Int. Conf. Acoust., Speech and Signal

Grammatikopoulos, L., Kalisperakis, I., Karras, G., Kokkinos, T., Petsa, E., 2004: On automatic orthoprojection and texture-mapping of 3D surface models. IAPRS, Vol. XXXV, part 5, pp. 360-366

Gruen, A 1985: Adaptative least squares correlation: A powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography, Vol. 14, No. 3, 1985, pp. 175-187.

Gruen, A., Remondino, F., Zhang, L., 2004: 3D modeling and visualization of large cultural heritage sites at very high resolution: the Bamiyan valley and its standing Buddhas. IAPRS, Vol. XXXV-B5

Haeberli, P. and Segal, M., 1993: Texture mapping as a fundamental drawing primitive. Proc. Of the 4th Eurographics Workshop on Rendering, pp. 259-266

Hartley, R and Zisserman, A., 2000: Multiple View Geometry in Computer Vision. Cambridge University Press

Havaldar, P., Lee, M.-S., Medioni, G., 1996: View synthesis from unregistered 2-D images. Proceedings of Graphics Interface '96, pp. 61–69

Kada, M., Roettger, S., Weiss, K., Ertl, T., Fritsch, D., 2003: Real-time visualization of urban landscape using open-source software. Proc. of ACRS 2003

Karner, K., Bauer, J., Klaus, A., Leberl, F. and Grabner, M., 2001: Virtual Habitat: models of the urban outdoors. Int. workshop on 'Automatic extraction of man-made objects from aerial and space images (III)', Baltsavias, Gruen, Van Gool (Edtors), pp. 393-402, Ascona, Switzerland

Kim S.J. and Pollefeys, M., 2004: Radiometric Self-Alignment of Image Sequences. Proceedings of CVPR, pp. 645-651

Mayr, W. and Heipke, C., 1988: A contribution to digital orthophoto generation. IAPRS, Vol. XXVII, part B/11, pp. 430-439

McGlone, C., Mikhail, E. and Bethel, J, 2004: Manual of Photogrammetry, ASPRS Publications.

Neugebauer, P. and Klein, K., 1999: Texturing 3D models of real world objects from multiple unregistred photographics views. Proc. of Eurographics '99, Computer Graphics Forum, Vol. 18(3)

Niem, W. and Broszio, H., 1995: Mapping texture from multiple camera views onto 3D-object models for computer animation. Proceedings of the International Workshop on Stereoscopic and Three Dimensional Imaging

Pulli, K., Abi-Rached, H., Duchamp, T., Shapiro, L., Stuetzle, W., 1998: Acquisition and visualization of colored 3-D objects. Proc. ICPR, pp. 99-108

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

Rocchini, C., Cignoni, P. Montani, C., Scopigno, R., 2002: Acquiring, stiching and blending diffuse appearance attributes on 3D models. The Visual Computer, 18, pp. 186-204

Ulm, K., 2005: 3D city models from aerial imagery – Integrating images and the landscape. GEOInformatics, January/February, Vol. 8, pp. 18-21

Varshosaz, M., 2003: Occlusion-free 3D realistic modelling of buildings in urban areas. IAPRS, Vol. XXXIV, part 5/W10 (on CD-Rom)

Visnovcova (Niederoest), J., Zhang, L., Gruen, A., 2001: Generating a 3D model of a Bayon tower using non-metric imagery. IAPRS, Vol. XXXIV, Part 5/W1, pp. 30-39

Wang, L., Kang, S.B., Szeliski, R., Shum, H.-Y., 2001: Optimal texture map reconstruction from multiple views. IEEE Proc. of CVPR, Vol.1, pp. 347-354

Weinhaus, M. and Devarjan, V., 1997: Texture mapping 3D models of real-world scenes. ACM Computing Survey, 29(4), pp. 325-365

Zhang, Z. and Kang, Z., 2004: The rendering of building texture from land-based video. IAPRS, Vol. XXXV, part B3, pp. 732-737