

EXPERIENCES AND CONSIDERATIONS IN IMAGE-BASED MODELING OF COMPLEX ARCHITECTURES

F. Voltolini ^a, F. Remondino ^b, M. Pontin ^a, L. Gonzo ^a

^aCentre for Scientific and Technological Research, ITC-IRST, Trento, Italy - (fvoltolini, pontin, lgonzo)@itc.it

^bInstitute of Geodesy and Photogrammetry, ETH Zurich, Switzerland - fabio@geod.baug.ethz.ch

COMMISSION V – WG 4

KEYWORDS: Calibration, Orientation, Modeling, Precision

ABSTRACT:

Nowadays image-based modeling is receiving much attention and many applications require precise and photo-realistic 3D models. The camera calibration and orientation phases are key steps in the 3D modeling process. If these phases are not accurately performed, there will be some errors in the final model and for some applications low accuracy results are not accepted. The goal of this work is to investigate the influence of wrong camera parameters or bad image configuration in object reconstruction. The analysis is performed with a bundle adjustment solution perturbing the interior camera parameters and using different network configurations. We analyze the effects of wrong focal length and principal point as well as absent distortion parameters with images acquired under typical project configurations. Finally we report some examples of 3D modeling of complex architectures where the theoretical considerations cannot always be fulfilled.

1. INTRODUCTION

One of the main goals of photogrammetry is to achieve very high accuracy in the measurements and object reconstruction. To achieve these objectives, assuming a good image resolution, it is very important to have both a precise calibration of the camera and a correct external camera configuration. In this paper, both aspects are investigated, in particular for applications like 3D modeling of complex architectural objects. While there is a great amount of publications on digital cameras [Kunii and Chikatsu, 2001; Läbe and Förstner, 2004], camera calibration [Fraser and Shortis, 1995; Wiley and Wong, 1995; Clarke and Fryer, 1998; Jantos et al., 2002; Cronk et al., 2006] and optimal network design [e.g. Fraser, 1996], there are not many deep studies dealing with the design and problems related to an image-based 3D modeling project. Kahl et al., [2001] and Sturm [2002] reported critical configuration in the projective reconstruction and self-calibration problems. El-Hakim et al. [2003] showed the errors obtained by taking images of an object with different focal lengths and assuming one focal length for all images in the modeling process. On the other hand, the effect of image resolution on the computed 3D object coordinates was investigated in [Chikatsu, 2001; El-Hakim et al., 2003].

Traditionally we can distinguish between 3 types of digital cameras and their expected accuracy on well-defined targets in ideal or perfect conditions:

- (i) *Amateur / consumer cameras*, employed for VR reconstructions, quick documentation and forensic applications; the relative accuracy potential is around 1:25 000.
- (ii) *Professional cameras*, used in industrial, archaeological, architectural or medical applications, which allow up to 1:100 000 accuracy.
- (iii) *Photogrammetric cameras*, used in industrial applications and measurements with accuracy potential of 1:200 000.

Nowadays different consumer digital cameras come with 8 Mega pixels or more, therefore highly precise measurements can be achieved also with these sensors, even if the major problems in these types of cameras are given by the objectives.

Usually artefacts and object details define the acceptable accuracy and quality of a project. The accuracy of the computed object coordinates (σ_{XYZ}) depends on the image measurement precision, image scale and geometry as well as on the number of exposures [Fraser, 1996]:

$$\sigma_{XYZ} = q S \sigma_{xy} / k^{1/2}$$

with :

σ_{xy} = standard error of image measurements;

q = empirical factor;

S = scale number (mean object distance / camera focal length);

k = number of images per station.

The standard error of the image coordinates plays a very important role: well defined targets can be measured with an accuracy of 1/25th of a pixel while with natural features the accuracy is around 1/3th of a pixel. Generally in 3D modeling of complex architectural objects, performed with amateur cameras and manual measurements (which gives ca one pixel measurement accuracy), a relative accuracy in the range 1:10 000 – 15 000 is generally expected.

In this article, using an already calibrated camera as reference, we study the influence and relevance of some camera parameters in typical practical network configurations. The internal parameters are perturbed by introducing some errors and the bundle adjustment is executed to simulate the behaviour of a wrong camera calibration. Moreover different camera configurations are considered: the effects of a bad B/D ratio and the weak image redundancy are evaluated to analyze which configuration is the most appropriate for better accuracies. All the tie points are measured manually using natural features, without targets.

The goal is to quantify the entity of wrong camera parameters and bad image configuration on the modeled object. The output parameters considered are the residuals in image space, the a posteriori standard deviation of the adjustment (sigma naught), the theoretical precision of the computed object coordinates and the comparison between the computed object coordinates and check point coordinates (RMSE).

2. INVESTIGATION ON CAMERA PARAMETERS

The goal of this analysis is to check the influence of wrong camera interior parameters on the imaged scene. Three images of a building's façade (Figure 1) are taken to analyze the effects of wrong focal length, wrong or absent distortion parameters and wrong displacement of the principal point.

The used camera (Sony DSC F828, 5 Megapixel, pixel size = 3.4 μm) is pre-calibrated in our laboratory, by means of a self-calibrating bundle adjustment. Afterwards, the recovered interior parameters are used to orient the three reference images and compute the 3D object coordinates. The bundle adjustment is performed in Australis™ using 22 tie points manually measured. The results ($\sigma_0 = 3.105 \mu\text{m}$, $\sigma_x = 2.0 \text{ mm}$, $\sigma_y = 4.0 \text{ mm}$, $\sigma_z = 1.9 \text{ mm}$) are considered as reference and compared with those obtained using wrong interior parameters.

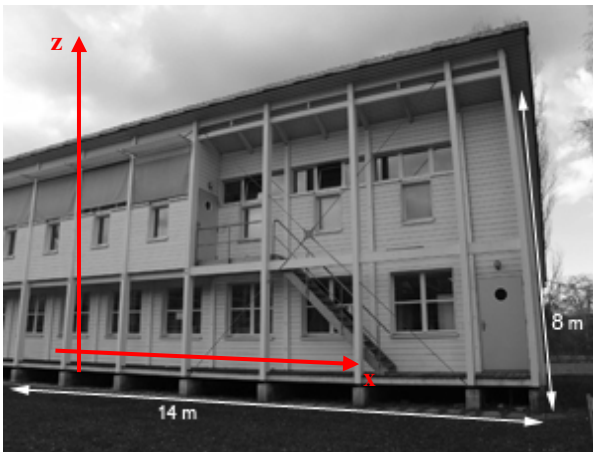


Figure 1: The object selected for the investigation: a façade spanning approximately (14 x 8 x 2) m. The distance camera-object is ca 10 m while the recovered focal length is 7.32 mm.

In the next sections different errors on the camera interior parameters are applied and the bundle adjustment results analyzed and commented.

2.1 Error in the focal length

The influence of the focal length in the adjustment is analyzed introducing an error of 1%, 3%, 5%, 10%, 15% and 25% of the correct calibrated value. The 3% error corresponds to the value readable in the EXIF header of a file. In Figure 2 and Figure 3 the behaviour of sigma naught, image residuals and standard deviations of object coordinates are reported. The linear behaviour is clear and similar for all the analyzed quantities.

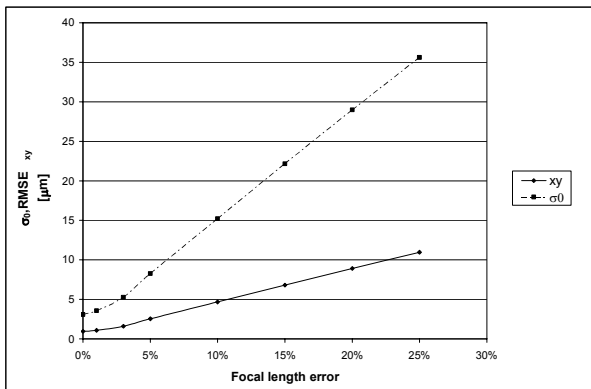


Figure 2: Error due to incorrect focal length value expressed with the image residuals and σ_0 .

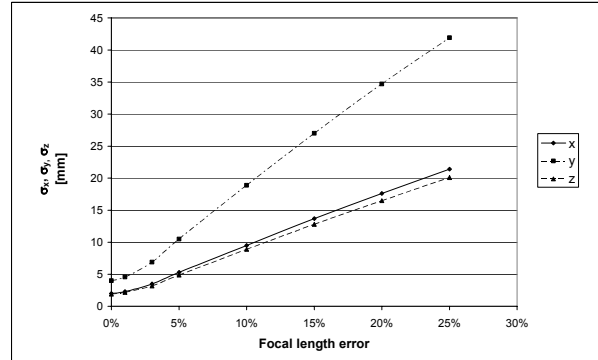


Figure 3: Error in object space represented with the standard deviation of the computed object coordinates. The y-axis is parallel to the depth.

In Figure 4 the RMSEs with respect to the check points are shown. The trend is the same as with the obtained theoretical precision of the computed object coordinates.

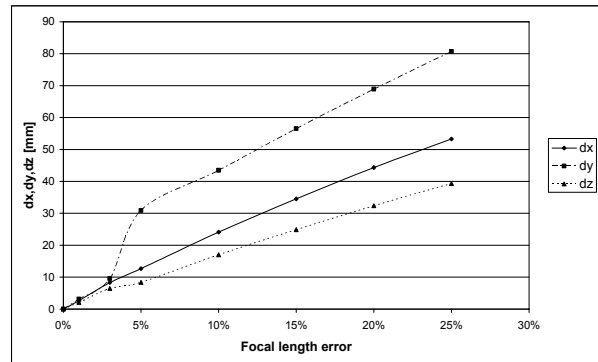


Figure 4: RMSEs of the computed object coordinates with respect to the check points.

2.2 Wrong or absent distortion parameters

The most common set of Additional Parameters (APs) employed to compensate for systematic errors in digital cameras is the 8-term 'physical' model originally formulated by Brown (1971). The three APs used to model radial distortion Δr are generally expressed via the odd-order polynomial $\Delta r = K_1 r^3 + K_2 r^5 + K_3 r^7$, where r is the radial distance. Radial distortion varies with focus and its coefficients K_i are usually highly correlated, with most of the error generally being accounted for by the cubic term $K_1 r^3$. The K_2 and K_3 terms are typically included for photogrammetric low distortion and wide-angle lenses, and in higher-accuracy vision metrology applications. On the other hand *decentering distortion* is due to a lack of centering of lens elements along the optical axis. The decentering distortion parameters P_1 and P_2 (Brown 1971) are invariably strongly projectively coupled with the principal point position. Decentering distortion is usually an order of magnitude, or more, less than radial distortion and it also varies with focus, but to a much less extent compared to radial distortion.

We performed different adjustments, without the three K_i parameters, without the two P_i parameters and finally without any distortion parameter. In Figure 5 and Figure 6 the results are graphically represented. The series marked as "OK" is the reference one, with all the correct camera internal parameters. In the "NO K " series the K_i parameters are removed, in the "No P " series the P_i parameters are removed and in the "No Par " series no distortion parameters are used.

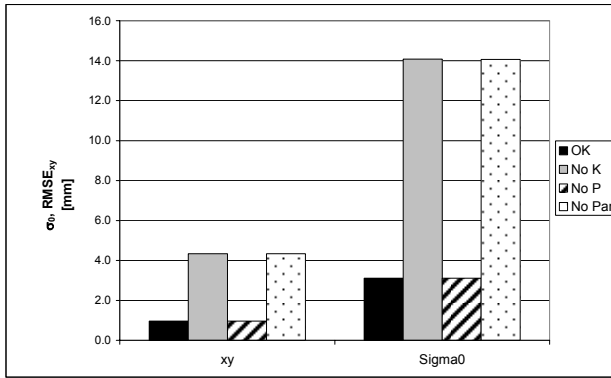


Figure 5: Error due to wrong/absent distortion parameters expressed with the image residuals and σ_0 .

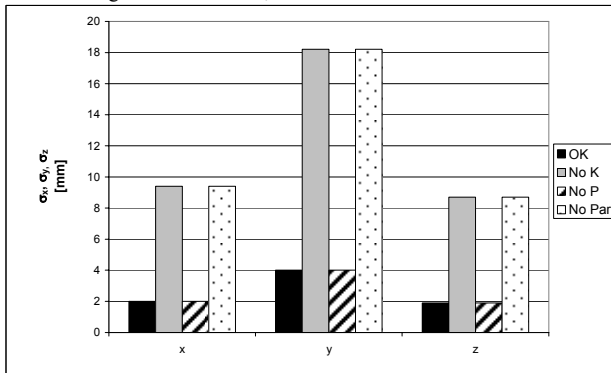


Figure 6: Error in object space due to wrong/absent distortion parameters represented with the standard deviation of the computed object coordinates.

In image space as well as in object space it's clearly visible that the parameters that mainly affect the adjustment are the radial distortions K_i . In fact there is no meaningful difference between the results without K_i parameters and without all distortion parameters. This was expected as the commonly encountered third-order barrel distortion seen in consumer-grade lenses is accounted for by K_1 . The influence of P_i parameters is very small. They have a much important role for applications like industrial metrology and deformation analysis, where very high accuracy is required.

As in our specific case the K_i parameters showed to be able to absorb all the systematic errors of the lens, we performed further tests introducing an error on the radial distortion parameters (1%, 5%, 10%, 15%, 25% of the calibrated values). Figure 7 shows the behaviour in image and object space of the analyzed quantities and Figure 8 shows the RMSEs of the computed 3D points.

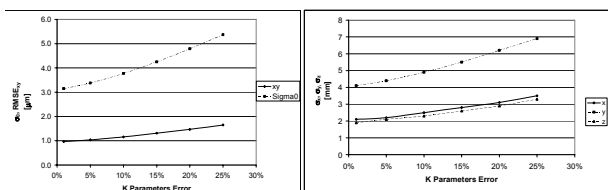


Figure 7: Error due to wrong K parameters expressed with the image residuals and σ_0 (above) as well as standard deviation of the computed object coordinates (below).

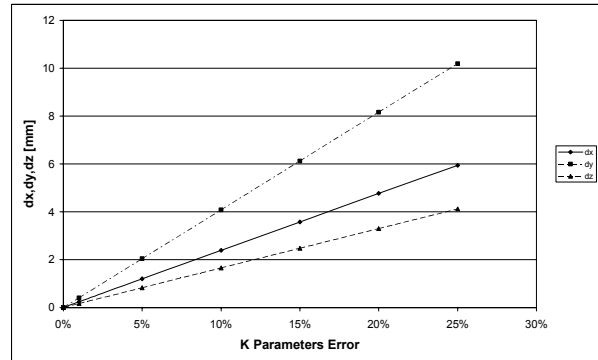


Figure 8: RMSEs of the computed object coordinates with respect to the check points in the case of wrong radial distortion parameters.

2.3 Error in the principal point

We consider the case where the principal point is located in the center of the image (0, 0). The adjustment output reported no significant effects in either standard deviations or residuals and very similar results with or without the correction of the principal point location. The value of σ_0 is 3.105 μm if the principal point is calibrated and 3.120 μm otherwise. This can be explained by the fact that a small principal point displacement is generally compensated by the exterior orientation parameters.

3. INVESTIGATION ON IMAGE CONFIGURATION

As previously practically demonstrated, the interior parameters cover a key role in the 3D reconstruction process. The positions where the photos are taken, the number of employed images and the number of points per image are also very important factors to obtain accurate models. In this section the effects of bad image configuration and image redundancy are analyzed. In particular, widely separated images are compared with closely spaced images. We consider a set of images tied with 20 points manually measured.

3.1 Effect of bad image configuration

A reasonable B/D (base-to-depth) ratio between the images should ensure a strong geometric configuration and reconstruction that is less sensitive to noise and measurement errors. A typical value of the B/D ratio given in the photogrammetric literature should be around than 0.5 – 0.75, even if in practical situations it is often very difficult to fulfil the requirement. In Figure 9 two different configurations of our tests are shown. The first one is the case with the smallest B/D ratio (about 0.1) and the second one is the case with the biggest B/D ratio (about 1.4).

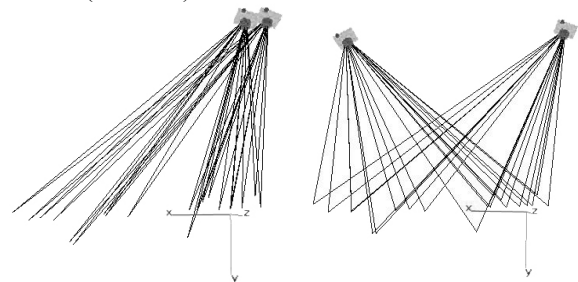


Figure 9: Configurations with a small B/D ratio (about 0.1) and a large B/D ratio (about 1.4) between the images.

Firstly we tested the influence of increasing the baseline between two images. It resulted that the computed theoretical precision of the object coordinates decreases until a stable value (Figure 10). Similarly we used more cameras (3 and 4), achieving the same behaviour in the results (Figure 11). In the graphs it can be noticed how, at the same B/D value (e.g. ~ 0.45), the standard deviations decreased (e.g. in y direction) almost of a factor 4, because of the higher image redundancy (and therefore images per point).

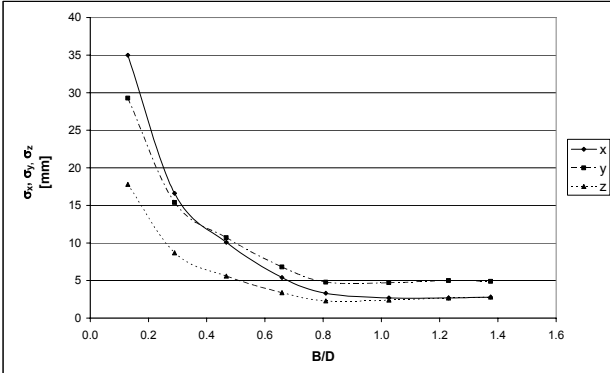


Figure 10: Error in object space with 2 cameras and different B/D ratio represented with the standard deviations of the object coordinates.

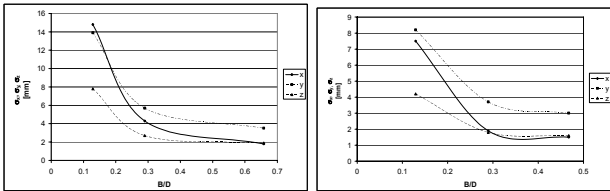


Figure 11: Error in object space with 3 (left) and 4 (right) cameras and different B/D ratio represented with the standard deviations of the computed object coordinates.

The 3D points obtained with these configurations are afterwards compared with those obtained by using 8 cameras. The results reported similar values and trend to the computed theoretical precision.

3.2 Effect of image redundancy

We consider the number of images where a point appears keeping fixed the number of points per image. Three configurations with 3, 5 and 9 cameras are analyzed. In Figure 12 the standard deviations of the computed object coordinates are shown. The theoretical precision decreases with the increase of the number of used images (i.e. higher redundancy). The higher difference is between the configuration with two and the one with three cameras.

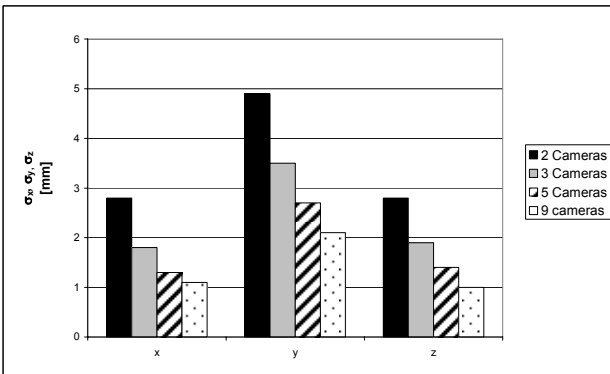


Figure 12: Effect of image redundancy represented with the theoretical precisions of the computed object coordinates.

Figure 13 shows the 3D points mean error with 2, 3 and 5 cameras.

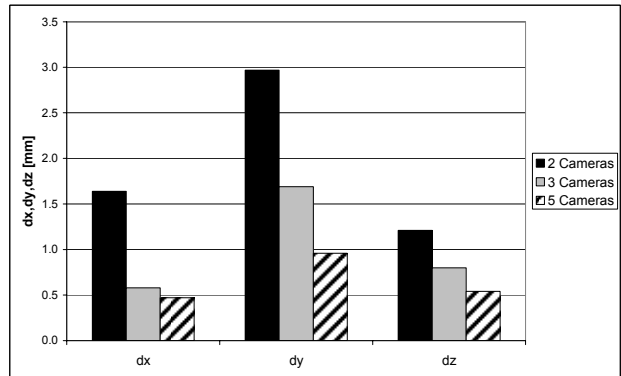


Figure 13: RMSEs of computed 3D coordinates using 2, 3 and 5 cameras.

4. PRACTICAL 3D MODELING PROJECTS

We have so far quantitatively described how camera interior parameters and image network can influence the 3D reconstruction results. Generally, in practical 3D modeling projects, the only parameter which cannot always satisfy the theoretical experiments is the B/D ratio, i.e. a good image network. Many difficulties in this way can be found during the image acquisition, due to occlusions or narrow areas. This will be reflected in the final accuracy and looking of the reconstructed 3D model.

In this section we report some 3D modeling projects of internal and complex courtyards or narrow alleys, located within fortresses or castles (Figure 16). In these situations, as a good baseline between the images and a high image redundancy are not always assured, the image registration becomes very complex and robust procedures should be used for the orientation phase. Moreover, for the modeling, powerful tools and algorithms should be employed. As man-made objects generally contains planes, right angles and parallel lines, these cues and geometric constraints (perpendicularity or orthogonality) together with image invariants should be used to retrieve 3D information of occluded objects.



Figure 14: An aerial view of the Valer Castle, located near Trento, Italy. This is a typical example of architecture containing streets and courtyard very narrow and quite complex for successful and precise 3D modeling.

Figure 15 reports the example of an open courtyard (approximately 10 m high with a base of 4 x 9 m). The narrow area limited the location of the image acquisition. Moreover the

vegetation on the walls could not be removed, creating a lack of quality in the produced 3D model. Different methods were published to overcome this problem [Boehm, 2004; Ortin and Remondino, 2005]. These methods are based on statistical image analysis techniques and therefore require a large numbers of views to remove the unwanted occlusions, redundancy which is not always assured in this kind of modeling project.



Figure 15: A narrow courtyard with vegetation occluding corners and walls (above). Two views of the generated 3D model (below).

Figure 16 shows a typical narrow path (ca 2 meters wide) of a medieval city with an overhanging balcony (ca 7 meters high) that should be virtually reconstructed.



Figure 16: Narrow medieval street (left) and 3D model of the balcony in the background (right).

Finally Figure 19 shows a 3-levels courtyard (approximately 12 m high with a base of 6 x 7 m) full of arches and columns. Some snapshots of the virtually reconstructed courtyard are reported in Figure 18.



Figure 17: Two views of the narrow and high courtyard.

5. CONCLUSIONS

In this article we have reported practical and quantitative analysis of image-based modeling projects. The initial investigations were performed on an architectural object (and not with a testfield) as it is quite similar to practical modeling objects. Moreover the measurements were performed manually as generally architectures 3D models are generated interactively, by means of manual measurements. The used networks consisted of images acquired along horizontal shifts in front of the object, a typical acquisition procedure of a virtual reconstruction project.

We have analyzed some of the interior camera parameters to understand which one have more influence on the final 3D reconstruction. As expected, the focal length and symmetrical radial lens distortion are the most significant parameters and an error on these values affects in a remarkable way the image residuals and the computed object coordinates accuracy. On the other hand, the assumption of the principal point located in the centre of the image did not show significant effects on the 3D reconstruction, as this error is usually absorbed by the exterior orientation parameters.

Considering the different camera configurations, the tests performed with closely separated images and points visible only in few images, produced very poor results. While combining images with good network geometry and increasing the number of images where a point appears, lead to reduce the standard deviation by a factor 4 in depth direction. Therefore the accuracy of a network increases by increasing convergence angles of the imagery (and therefore increasing the base-to-depth (B/D) ratio). The global network accuracy is enhanced by increasing the number of rays to a given object point, even if the rate of improvement is proportional to the square root of the number of images 'seeing' the point. Finally the accuracy increases with the number of measured points per image, but the incremental improvement is small beyond a few tens of points. More important is that extra points within an image offer better prospects for modelling departures from collinearity throughout the full image format.

The statistical evaluations in object space have been done using the covariance matrix, which is better to recall that might give too optimistic results if low redundancy, weak network and unmodelled systematic errors are present in the least squares adjustment. Nevertheless, in our case, a comparison with some check point gave similar results to the achieved theoretical precisions.

In practical 3D modeling cases, it is difficult to achieve an optimal network design, mainly due to constraints in the image

acquisition or because the images are acquired by non-expert. This was shown in the final examples. Therefore it is very difficult to obey to the theoretical rules as each modeling project is unique and different from others. Theory and basic principles are only one element: the creation of a 3D model, from the acquisition to the visualization is a different matter.

Moreover the image network which is optimal for camera calibration (well distributed images of a 3D object with 2-3 rotated acquisitions) is generally not the best for the object modeling phase. For this reason, it is always better to separate the calibration and the reconstruction steps to achieve as accurate results as possible.



Figure 18: Some renderings of the generated virtual courtyard shown in Figure 17.

ACKNOWLEDGMENTS

The authors would like to thank the helpful contributions of Alessandro Rizzi, ITC-IRST Trento, Italy, for the rendering of the generated 3D models.

6. REFERENCES

Boehm, J., 2004: Multi-image fusion for occlusion-free façade texturing. IAPRS, 35(5), pp. 867-872, Istanbul, Turkey

Brown, D.C., 1971: Close-range camera calibration. PE&RS, Vol. 37(8), pp.855-866

Clarke, T.A. and Fryer, J.F., 1998: The development of camera calibration methods and models. Photogrammetric Record, 16(91), pp 51-66

Chikatsu, H., 2001: Rapid acceleration of resolution of amateur camera and applications. 3rd Int. Image Sensing Seminar on New Development in Digital Photogrammetry, Gifu, Japan

Cronk, S., Fraser, C.S. and Hanley, H.B., 2006: Automatic Calibration of Colour Digital Cameras. The Photogrammetric Record (in press).

El-Hakim, S. F., Beraldin, J. A., Blais, F., 2003: Critical factors and configurations for practical image-based 3D modeling. Proceedings of 6th Conference Optical 3D Measurements Techniques. Zurich, Switzerland. Vol. II, pp 159-167

Fraser, C. S. and Shortis, M., 1995: Metric exploitation of still video imagery. The Photogrammetric Record, 15(85), pp. 107-122

Fraser, C.S., 1996: Network design. In 'Close-range Photogrammetry and Machine Vision', Atkinson (Ed.), Whittles Publishing, UK, pp. 256-282

Jantos, R., Luhmann, T., Peipe, J., Schneider, C.-T., 2002: Photogrammetric Performance Evaluation of the Kodak DCS Pro Back. IAPRS, Vol. 33(5), Corfu, Greece

Kahl, F., Hartley, R., Astrom, K., 2001: Critical configurations for N-views projective reconstruction. IEEE Proc. of CVPR01

Kunii, Y. and Chikatsu, H., 2001: On the application of 3 million consumer digital camera to digital photogrammetry. Proceedings of SPIE Videometrics VII, Vol. 4309, pp. 278-287

Läbe, T. And Förstner, W., 2004: Geometric stability of low-cost digital consumer cameras. IAPRS, Vol. 35(5), pp. 528-535, Istanbul, Turkey

Sturm, P., 2002: Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image and Vision Computing, 20, pp. 415-426

Ortin, D., Remondino, F., 2005: Generation of occlusion-free images for texture mapping purposes. IAPRS, 36(5/W17), on CD-Rom, Venice, Italy

Wiley, A.G. and Wong, K.W., 1995: Geometric calibration of zoom lenses for computer vision metrology. PE&RS, Vol. 61(1), pp. 69-74