Presented paper on the 3rd Conference on Optical 3-D Measurement Techniques, October 2-4, 1995, Vienna, Austria. Published in "Optical 3-D Measurement Techniques III", Gruen/Kahmen (Eds), Wichmann Verlag, Heidelberg, 1995. pp. 305-314.

VIDEOGRAMMETRY AND CAAD FOR ARCHITECTURAL RESTITUTION OF THE OTTO-WAGNER-PAVILLON IN VIENNA

André Streilein

Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
CH-8093 Zurich, Switzerland
email: andre@p.igp.ethz.ch

Abstract

Although there is an unchanged demand for the recording and documentation of the cultural heritage, modern acquisition and analysis techniques are rarely used. The paper shows results of the application of videogrammetry and CAAD for the architectural restitution, contributing thus to the CIPA test "Wagner-Pavillon". The images were acquired with a S-VHS camcorder, which today are considered to be "house-hold items". The results are compared to the results obtained by other institutes, which have used conventional photogrammetric techniques. The accuracies of about 1 cm on the object which have been obtained are compatible to those derived by classical photogrammetry, while the unique characteristics of digital photogrammetry (automation, computer compatibility, relaxation of expertise) give videogrammetry distinct advantages.

1. Introduction

Nowadays surveying methods based on imaging sensors are used in many applications in Computer Vision, Robotics, Machine Vision and Industrial Metrology. Although there is an unchanged demand for the surveying and documentation of the cultural heritage, modern image acquisition and analysis techniques are rarely used (*Waldhäusl, 1992*). Improvements and new developments in the fields of sensor and computer technology have had a major impact in many fields including photogrammetry. On the other hand such technological advances have led to the production of low-cost consumer products (e.g. video-cameras), which today are considered to be "house-hold items". World-wide, only a small percentage of the buildings of cultural interest are documented, but many images are taken by architects, historians and others, who are interested in the world heritage or just in souvenir photography. There is a great potential for such imagery taken by accident, and it would be advantageous to use this resource in addition to professional photogrammetric imagery for the restitution of the world's heritage (*Waldhäusl, Brunner, 1989*).

The purpose of this work is to investigate the use of video-camcorders in architectural recordings, contributing thus to the CIPA test "Wagner-Pavillon". One of the Otto Wagner's Stadtbahn station buildings on the Karlsplatz in Vienna was chosen as a test object for a CIPA (International Committee for Architectural Photogrammetry) project, which aims on an investigation on the state-of-the-art in Architectural Photogrammetry. Results from the restitution of the Otto-Wagner-Pavillon derived by digital architectural photogrammetry, using image data acquired with an inexpensive S-VHS camcorder, are presented. The results are passed to a data structure useful for CAAD (Computer Aided Architectural Design). The CAAD system is able to store the data in structures adapted to the architectural purposes and to find efficiently special data in a large data volume. It can also offer the capabilities of Virtual Reality in terms of visualization and animation, and is suitable for complex simulations, manipulations and analysis of the object.

2. Project definition

The Otto-Wagner-Pavillon is one of Otto Wagner's Stadtbahn Station buildings (see Fig. 1) on the Karlsplatz in Vienna, a masterpiece of Art Nouveau, built in 1898/1899. The dimensions of the building are 15 x 8 x 10 m^3. For the determination of object control points a 6-station surveying network has been established around the building and the polar coordinates of 44 non-signalized (but well defined in the majority) control points have been measured. After the adjustment of the surveying measurements, the local cartesian coordinates of the control points, which cover all four exterior facades, have been determined with an accuracy of 2 mm. The work described in this paper was done as part of the contribution to the CIPA project "Wagner-Pavillon". The idea and the initiative of this project belongs to P. Waldhäusl (*Waldhäusl, 1991*) and the aim was to check the current state-of-the-art in architectural photogrammetry. Therefore a small test object has been selected, photographed, measured and documented, in order to have well-checked materials to train students and photogrammetrists as well as to evaluate internationally the results of the analytic photogrammetric process with various cameras, with different software and with different kinds and amount of control information.

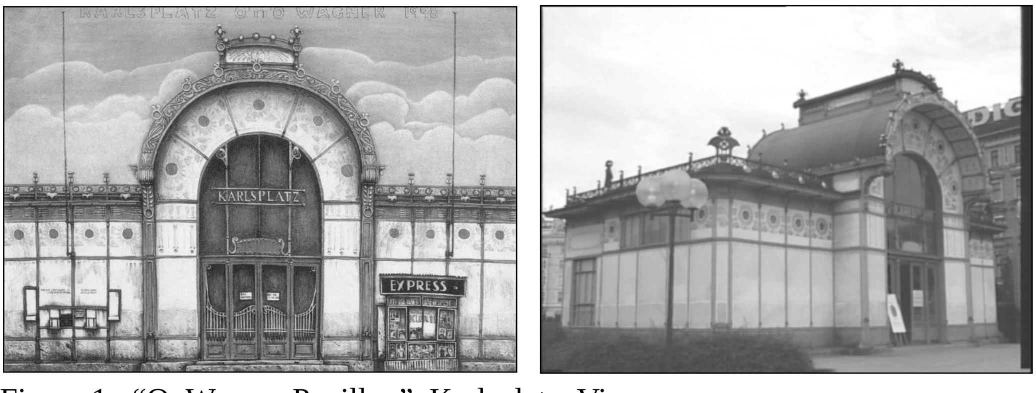

Figure 1: "O. Wagner Pavillon", Karlsplatz, Vienna.

3. Image acquisition

Images for digital photogrammetry are captured either with film-based cameras and subsequent digitization, or with cameras using solid-state sensors. Conventional film-based cameras still provide for an unsurpassed resolution, but the film must be developed and digitized before its data becomes available for digital photogrammetric techniques. Solid-state cameras on the other hand provide immediate access to the digital image data. Today a variety of solid-state imaging systems, ranging from inexpensive camcorders to high resolution systems, are available on the market (*Luhmann, 1992; Lenz, Lenz, 1993; Seitz et al., 1995*). For the application in architectural photogrammetry different requirements are of importance, e.g. high-resolution to image small object features, portability, high dynamic range to image large variations in scene brightness common to outdoor scenes, the ability to store a huge amount of image data and the metric accuracy performance. In this test the image acquisition was performed with an inexpensive S-VHS camcorder. Although camcorders are not intended for photogrammetric use, they exhibit many useful characteristics. They are inexpensive and widely used for other purposes as well, they are portable and free-hand, they need no special equipment, they offer the ability of on-site quality control. Furthermore they provide very inexpensively means for storage of huge amount of video data on video tapes.

Figure 2: Camcorder JVC GR-S77E.

The JVC GR-S77E camcorder (see Fig. 2) is a free-hand portable camera, which allows on-site control for the acquired imagery via an internal monitor. The camera incorporates a 1/2” color sensor (6.4 x 4.8 mm^2). The analog images are stored on a S-VHS video tape and have to be digitized by a framegrabber. The digitized images have a size of 728 x 568 pixel, which results in a pixel spacing of 8.8 µm in the horizontal and 8.5 µm in the vertical direction. The focal length (zoom lens) can vary between 8.5 mm and 65 mm.

Using such an imaging system in architectural photogrammetry requires that most functions which are implemented for the convenience of a common user (e.g. autofocus, zooming functions or image stabilizer) are disabled. During the image acquisition the zoom lens was fixed at its shortest focal length. Accuracies exceeding $1/20^{th}$ of the pixel spacing and a relative accuracy of 1 part in 15000 in object space can be obtained under laboratory conditions with this sensor (*Beyer et al., 1992*). Dealing with low-cost camcorders instead of metric cameras, the problem of low sensor resolution due to small imaging areas must be faced. To obtain sufficient accuracy in object space with such an imaging device, it is advisable to take as many images as reasonable and to use multiple camera arrangements with convergent rays instead of being restricted to special camera arrangements (e.g. stereopairs). Thus the number of measurements increases rapidly with the number of images, but one of the advantages of digital photogrammetry is the establishment of semi-automatic or automatic measurement routines.

4. Data processing

The digital photogrammetric system DIPAD (*Streilein, 1994*) was used as measuring and data processing device. The system combines digital photogrammetric methods with the capabilities of CAAD, while performing semi-automatic measurement routines in a CAAD environment. The overruling principle of DIPAD is that a human operator assumes responsibility of the image understanding part (assignment of feature attributes/semantics), while the measurement is automatically handled by the computer. The user indicates relevant parts of the object by approximating a geometric topology to it. This description is matched with the image data of multiple images and refines the coarse given model of the object iteratively until finally a detailed object description is generated.

By customizing the functionality of an existing software package with a true programming interface (AutoCAD) it is possible to perform the creation of the coarse model in the CAAD environment while the objects are projected run-time into the images. This means that DIPAD can make full use of the modelling capabilities of the CAAD environment, getting instant feed-back about how closely the model matches the images. The camera-icons in the CAAD system (see Fig. 3) are linked to the corresponding interior and exterior orientation parameters of an image viewer. Changes of the parameters of a camera-icon or the object model in the CAAD system can be entirely monitored in the image viewers. This serves as a general feedback generators during modelling and facilitates the initial positioning of the camera stations with respect to the model. These functions are not only relevant in creating a precise model of an existing building, they also serve well when creating new buildings in existing contexts.

A main software component of DIPAD is a three-dimensional feature extraction routine (3D-FEX). It is a semi-automatic measurement routine, where a series of model- and data-driven processes interact to extract geometric information from the digital imagery. The coarse object model is used to detect the features described by this model. Therefore, only relevant features (as defined by the user) are extracted and redundant or useless informations are reduced to a minimum. In addition, the use of a priori knowledge makes explicit assumptions, that allows the checking of whether or not these assumptions are fulfilled in the images. This approach acts rather globally, small deformations (e.g. image noise, blurred image structures,

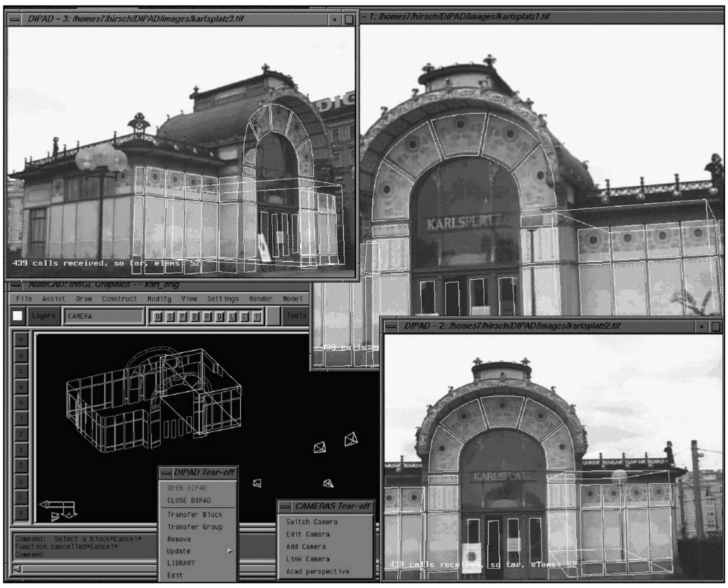

Figure 3: Modelling in the CAAD environment can be entirely monitored in the viewers linked to the model. Each camera icon corresponds to one viewer.

occlusions, etc.) in general do not effect the estimated result. The three-dimensional position of the object is derived by a simultaneous multi-frame feature extraction, where the object model is reconstructed and used to triangulate the object points from corresponding image points.

In architectural photogrammetry it is evident that linear boundaries (edges) of an architectural feature contain more information than the vertices (corners) of this feature. Although edges are only a small percentage of the whole image content, they have major importance for the description of object discontinuities. The 3D-FEX routine takes advantage of this knowledge, by first locating the edges of the features to be measured and then deriving the vertices as intersections of the appropriate lines. The object coordinates of a point detected in two or more images are calculated either by a spatial intersection, by a bundle adjustment or by a bundle adjustment with self-calibration, depending on which model parameters are treated as a priori known or unknown. The derived object coordinates are then projected into each image and used to restart the image-based feature extraction. Figure 4 gives an example of the perfor-

Figure 4: Performance of 3D-FEX routine of a single feature in three images simultaneously.
(a.) images from a sequence
(b.) initial 3D position of feature from approximate model, projected into image planes
(c.) final 3D position of feature after 5 iterations, projected into image planes

mance of this iterative procedure with a spatial intersection of a single feature in three images simultaneously. A more detailed description of the 3D-FEX routine is given in (*Streilein, 1994*).

5. System calibration

It is typical for low-cost imaging sensors to suffer from systematic errors. This especially so for camcorders employing off-the-shelf lenses with large distortions from an ideal perspective transformation. Previous investigations have shown that just the radial symmetric distortion of the lens used with the JVC camcorder can reach 70 µm (*Grün, Kersten, 1995*). The systematic errors are accounted for by extending the collinearity equations with functions of additional parameters. Many additional parameter sets have been developed to meet various requirements. In close-range CCD-sensor based systems the following set of additional parameters has proven to be effective (*Beyer, 1992*):

$$\Delta x = \Delta x_p - \frac{\bar{x}}{c}\Delta c - \bar{x}\, s_x + \bar{y}\, a + \bar{x}r^2k_1 + \bar{x}r^4k_2 + \bar{x}r^6k_3 + \left(r^2 + 2\bar{x}^2\right)p_1 + 2\bar{x}\bar{y}p_2$$
$$\Delta y = \Delta y_p - \frac{\bar{y}}{c}\Delta c + \bar{x}\, a + \bar{y}r^2k_1 + \bar{y}r^4k_2 + \bar{y}r^6k_3 + 2\bar{x}\bar{y}p_1 + \left(r^2 + 2\bar{y}^2\right)p_2 \quad (1)$$

where:

$\Delta x_p, \Delta y_p, \Delta c$ represents changes of the interior orientation elements,
s_x represents the scale factor in x direction,
a represents a shear factor,
k_1, k_2, k_3 are the first three parameters of radial symmetric lens distortion, and
p_1, p_2 are the first two parameters of lens decentering distortion,
$\bar{x} = x - x_p$, $\bar{y} = y - y_p$, $r^2 = \bar{x}^2 + \bar{y}^2$.

Parameters of interior orientation		**Additional parameters (value, $\hat{\sigma}$)**			
xp [mm]	+0.170685	Δx	-4.167*10^{-3}	1.033*10^{-3}	mm
yp [mm]	-0.172591	Δy	+1.744*10^{-3}	1.029*10^{-3}	mm
c [mm]	8.957499	Δc	+2.465*10^{-4}	1.001*10^{-3}	mm
		s_x	-8.512*10^{-5}	5.488*10^{-5}	-
		a	-3.240*10^{-4}	2.739*10^{-5}	-
		k_1	-1.284*10^{-3}	9.385*10^{-5}	mm^{-2}
		k_2	+2.678*10^{-5}	1.509*10^{-5}	mm^{-4}
		k_3	+1.366*10^{-6}	7.276*10^{-7}	mm^{-6}
		p_1	+1.366*10^{-4}	1.215*10^{-5}	mm^{-1}
		p_2	+6.258*10^{-5}	1.179*10^{-5}	mm^{-1}

Table 1: Result of sensor calibration.

To compare the results of the object reconstruction derived by a bundle adjustment with self-calibration with the results derived by a calibrated video camera a testfield calibration was performed. The testfield of ETH Zurich spans 2.6 x 2.0 x 1.1 m^3 and contains 162 targets.

Two images at each of four camera stations were taken, one image being acquired with the camera in upright position and the other with the camera rotated by 90 degrees around its optical axis. The above described set of additional parameters (1) was used for the calibration. The pixel coordinates of the targets were measured with Least Squares Matching (*Grün, 1985*). The data was then introduced into a bundle adjustment with self-calibration. The resulting parameters of interior orientation and the additional parameters are listed in Table 1. The values of the additional parameters entering the above calibration model have been estimated with a precision of 10^{-3} to 10^{-5}. Results of more detailed investigations on the accuracy performance of this type of video camera can be found in (*Höflinger, Beyer, 1993*) and in comparison with other solid-state imaging systems in (*Beyer et al., 1992*).

6. Object reconstruction

The image acquisition in-situ took place in a way that a human operator “filmed” an image sequence by walking around the object with an object distance of approximately 15-20 m, and another sequence by meandering across the front facade with an object distance of approximately 6 m. The idea of taking additional close-up images was to compensate the reduced number of measurable details due to the low sensor resolution, at least at the front facade. For the generation of the object model 38 single frames out of the video-sequences were digitized with a framegrabber. Depending on the original sequence two different types of images are collected in the data set (see Fig. 5). Type A with 25 images and an average image scale of 1 : 2400 and 13 images of type B with an average image scale of 1 : 700. This results in an image scale of 1 : 1800 for the entire data set of 38 images.

Figure 5: Examples for the two different types of images.

The measurement of image coordinates was performed with the 3D-FEX routine as described above. All image coordinates were introduced into a bundle adjustment. Several versions have been calculated. Concerning the type of object points it was distinguished between all object points which have been used to reconstruct the object (versions 1-5, 21&22) and a subset of these points, which has been used for the contribution to the CIPA test (versions 6-10, 23&24). Furthermore it was distinguished between the two types of images used and the used set of additional parameters (no additional parameters used, set of pre-calibrated additional parameters used, or values for additional parameters estimated by self-calibration). The results of these calculations are compiled in Table 2 and Table 3.

For the subset with the number of 145 “CIPA points” the results indicate a precision of 1.0 cm in object space, whereas the results derived for all object points a precision of 1.5 cm indicate. This is at a first look surprising, but explicable due to the higher number of image rays per object point (6.2 vs. 4.8) and due to the fact that for the subset of “CIPA-points” contains without exception well defined points, whereas the reconstruction of the entire object has to deal with any type of points, also poor defined ones. Comparing the versions 1, 2, 4 and 6, 7, 9 respectively shows that an increasing number of images and object point connections increases also the precision of the object point coordinates asymptotically towards a certain treshold. The comparison of the versions 3, 4, 5 and 8, 9, 10 respectively shows that the estimation of additional parameters is essential for that type of imaging system. On the other hand the versions calculated with a pre-calibrated camera and with self-calibration show no significant difference. This means that the additional amount of work for a sensor calibration can be avoided, if one is just interested in the values of interior orientation and additional parameters as intermediate results.

Version	Type of		St	Ap	Co	Ob	Im	Pc	r	$\hat{\sigma}_0$	Object Space [cm]			Image [μm]	
	obpts	images									$\hat{\sigma}_X$	$\hat{\sigma}_Y$	$\hat{\sigma}_Z$	μ_x	μ_y
1	CIPA-points	A	17	pre	40	141	540	3.83	675	6.09	1.227	1.540	0.744	4.71	4.63
2			25		41	145	811	5.59	1160	5.52	1.053	1.265	0.608	4.70	4.35
3		A+B	38	none	41	145	899	6.20	1258	8.75	1.388	1.607	0.899	6.27	7.12
4			38	pre					1258	5.70	0.887	1.022	0.578	4.89	4.41
5			38	self					1248	5.54	0.889	1.021	0.584	4.77	4.28
6	all object points	A	17	pre	40	517	1977	3.82	2421	6.71	1.564	1.890	0.873	5.05	5.35
7			25		41	521	2248	4.31	2906	6.39	1.469	1.761	0.808	5.00	5.18
8		A+B	38	none	41	537	2589	4.82	3462	8.65	1.783	2.087	1.001	6.31	7.20
9			38	pre					3462	6.90	1.389	1.622	0.780	5.35	5.84
10			38	self					3452	6.76	1.383	1.619	0.786	5.24	5.71

Table 2: Results of bundle adjustment for the Otto-Wagner-Pavillon.

Version	Type of object points	Ap	Ob	Im	r	$\hat{\sigma}_0$	Check points [cm]			Object points [cm]			Image [μm]	
							μ_X	μ_Y	μ_Z	$\hat{\sigma}_X$	$\hat{\sigma}_Y$	$\hat{\sigma}_Z$	μ_x	μ_y
21	CIPA-points	pre	145	899	1180	5.43	1.573	1.525	1.649	1.059	1.268	0.685	4.54	4.20
22		self			1170	5.32	1.823	1.573	1.418	1.106	1.294	0.731	4.46	4.08
23	all points	pre	537	2589	3384	6.83	1.769	1.883	1.413	1.488	1.760	0.855	5.23	5.78
24		self			3374	6.71	1.592	1.567	1.347	1.499	1.768	0.881	5.15	5.64

Table 3: Results of bundle adjustment for the Otto-Wagner-Pavillon (check point analysis).

Index for Table 2 and Table 3:

St	Number of stations/images
Ap	Additional parameters (none = no set used; pre = pre-calibrated set; self = estimated by self-calibration)
Co, Ob, Im	Number of control points, object points, image points
Pc	Point connection (mean number of image rays per object point)
r	Redundancy
$\hat{\sigma}_0$	Standard deviation of unit weight *a posteriori*
μ_X, μ_Y, μ_Z	Root Mean Square Error from comparison to check point coordinates in object space [cm]
$\hat{\sigma}_X$, $\hat{\sigma}_Y$, $\hat{\sigma}_Z$	Theoretical precision of object point coordinates [cm]
μ_x, μ_y	Root Mean Square Error from comparison to object point coordinates in image space [μm]

In an additional step the versions 4, 5 and 9, 10 were calculated with minimum instead of maximum control information (see Table 3). The estimated object coordinates of 26 check points were compared with the geodetic reference coordinates. For all four versions an accuracy of about 1.5 cm for the object coordinates of the 26 check points was achieved. This matches quite good with the theoretical precision of the 145, resp. 537 object points.

These results are comparable to the results derived with classical photogrammetric equipment in the CIPA test (*Patias et al., 1993*). There a total of 51 solutions was processed with different hard- and software. The average precision of these solutions is in a range of 0.3 cm to 2.0 cm. Among these 29 solutions with small format cameras were calculated. The precision of these solutions is in a range of 0.6 cm to 1.1 cm. This is comparable to the results of versions 4 and 5 in Table 1.

The final product of the photogrammetric analysis with DIPAD is a three-dimensional geometric and semantic object description of the Otto-Wagner-Pavillon in the CAAD environment (see Fig. 8). The representation of the architectural object can be given either by its object points, lines (wireframe), surfaces or any combinations of these. The representation in a CAAD environment offer also the capabilities of Virtual Reality in terms of visualization and animation. The original image data can be draped onto the derived object model and animations (walk-through, fly-through) can be performed.

7. Conclusions

Results from the restitution of the Otto-Wagner-Pavillon derived by digital architectural photogrammetry, using image data acquired with an inexpensive S-VHS camcorder, have been presented. Although camcorders are not intended for photogrammetric use, they exhibit many useful characteristics. They are inexpensive and widely used for other purposes as well, they

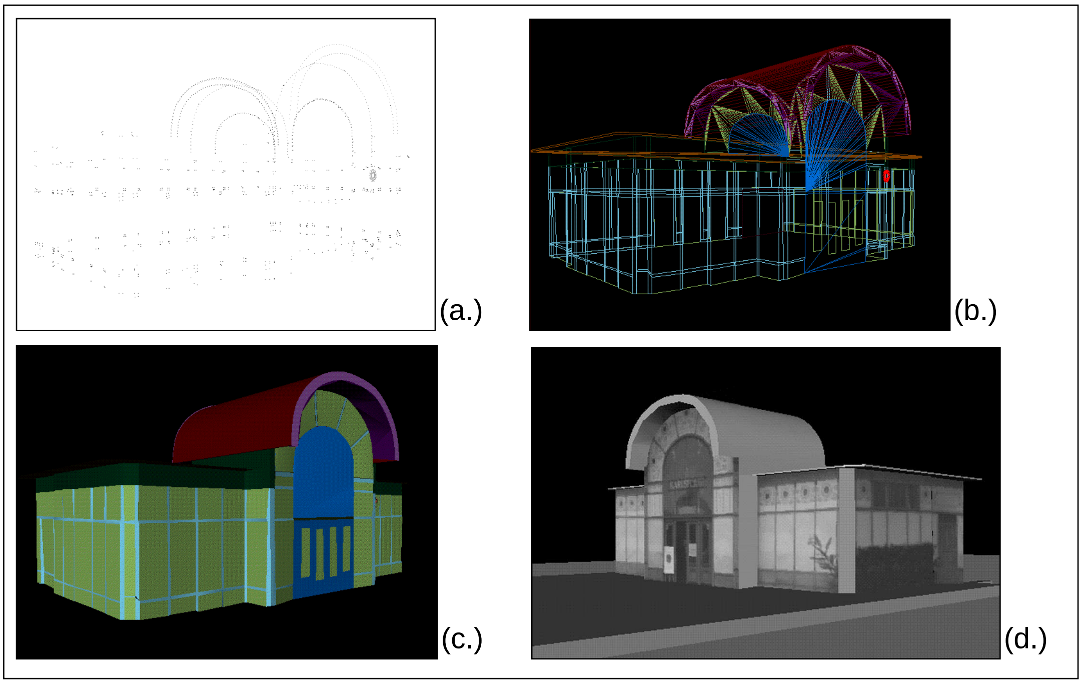

Figure 9: Photogrammetrically generated CAD-model of O-Wagner-Pavillon.
(a.) point model (b.) wireframe model
(c.) surface model (d.) texture model draped with original image data

are portable and free-hand, they need no special equipment, they offer the ability of on-site quality control, they provide very inexpensively means for storage of huge amount of video data on video tapes, and there is a potential for such imagery showing buildings of cultural interest taken by accident.

The accuracy of the results derived with the video-camcorder are comparable to the results derived with small format film-based cameras and classical photogrammetric techniques. The use video-data in addition to professional photogrammetric imagery for the restitution of the world's heritage is practical. The problem of low sensor resolution due to small imaging areas can be faced by taking as many images as reasonable and using multiple camera arrangements with convergent rays instead of being restricted to special camera arrangements (e.g. stereopairs). Thus the number of measurements increases rapidly, but with the semi-automatic and automatic measurement routines of digital photogrammetry this is not any longer a disadvantage.

8. References

Beyer, H., 1992. "Geometric and Radiometric Analysis of a CCD-camera-based photogrammetric close-range system", Dissertation No. 9701, ETH-Zurich, 1992.

Beyer, H., Kersten, T., Streilein, A., 1992. "Metric Accuracy Performance of Solid-State Camera Systems". Proc. SPIE Vol. 1820, Videometrics, pp. 103-110.

Gruen, A., 1985. "Adaptive Least Squares Correlation: A Powerful Image Matching Technique". South African Journal of Photogrammetry, Remote Sensing and Cartography. Vol. 14, No. 3, pp. 175-187.

Grün, A., Kersten, T., 1995: "Sequential Estimation in Robot Vision". Photogrammetric Engineering & Remote Sensing, Vol. 61, No. 1, January 1995, pp.75-82.

Höflinger, W., Beyer, H., 1993. "Characterization and calibration of an S-VHS camcorder for digital photogrammetry". Optical 3D Measurement Techniques II, A. Grün, H. Kahmen, Editors, Proc. SPIE Vol. 2252, pp. 133-140.

Lenz, R., Lenz, U., 1993. "New developments in high resolution image acquisition with CCD area sensors". In Optical 3D Measurement Techniques II, A. Grün, H. Kahmen, Editors, Proc. SPIE Vol. 2252, pp. 53-62.

Luhmann, T., 1992. "Bilderfassung in der Nahbereichsphotogrammetrie - Aktuelle Tendenzen", Vermessungswesen und Raumordnung (VR 54/8), Wichmann-Verlag Karlsruhe, 1992, pp. 400-410.

Patias, P., Rossikopoulos, D., Georgoula, O., 1993. "CIPA Test O. Wagner Pavillion". Preliminary Report. Presented at the CIPA XV International Symposium, Bucharest, Romania, September 22-26, 1993. (Executive summary under publication in the proceedings).

Seitz, P., Vietze, O., Spirig, T., 1995. "From pixels to answers - recent deveoplments and trends in electronic imaging". International Archives of Photogrammetry and Remote Sensing. Vol. 30, Part 5W1, pp. 2-13.

Streilein, A., 1994. "Towards Automation in Architectural Photogrammetry: CAD-Based 3D-Feature Extraction". ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 49 No. 5, October 1994, pp. 4-15.

Waldhäusl, P., 1991. "A test object for architectural photogrammetry: Otto Wagners underground station Karlsplatz in Vienna", Proceedings of the XIV. International Symposium of CIPA, October 2-5, 1991, Delphi, Greece, pp. 247-251.

Waldhäusl. P., 1992. "Defining the Future of Architectural Photogrammetry". International Archives of Photogrammetry and Remote Sensing, Vol. 29, Part B5, pp. 767-770.

Waldhäusl, P., Brunner, M., 1989. "Architectural Photogrammetry World-wide and by Anybody with Non-metric Cameras?". Proceedings of XI. International Symposium of CIPA, October 1988, Sofia, Bulgaria, pp. 35-49.