Reconstruction of the Disentis monastery from high resolution still video imagery with object oriented measurement routines

André Streilein, Markus Niederöst

Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
CH-8093 Zurich, Switzerland
e-mail: {andre, markus}@geod.ethz.ch

Commission V, Working Group V/2

KEY WORDS: CAD, CAAD, cultural heritage, digital photogrammetry, high resolution CCD camera, model based analysis, object oriented measurement

ABSTRACT

The performance and results of the three-dimensional reconstruction of the Disentis monastery from high resolution still video imagery with the software environment DIPAD are described. DIPAD combines digital photogrammetric methods with the capabilities of a CAD system. CAD models are used in an a priori and a posteriori mode. The semi-automatic processing follows the HICOM principle, where a human operator assumes responsibility for the image understanding part and the computer for the object oriented measurement and the data handling. The imagery was acquired with a Kodak DCS 460 still video camera from a helicopter and from terrestrial viewpoints. The resulting CAD-model of the monastery has a geometric accuracy of about 1.5 cm in all three coordinate axes.

KURZFASSUNG

Die Vorgehensweise und Resultate der drei-dimensionalen Rekonstruktion des Klosters Disentis aus digitalen Bilddaten einer hochauflösenden Still-Video Kamera mit dem Softwarepaket DIPAD werden dargestellt. DIPAD vereinigt digitale photogrammetrische Methoden mit den Möglichkeiten eines CAD-Systems. Dabei werden CAD Modelle sowohl im a priori als auch im a posteriori Modus benutzt. Die halbautomatische Messmethode folgt dem HICON-Prinzip, wobei der Operateur für die Bildinterpretation verantwortlich zeichnet, während der Computer für die Messungen und die Verwaltung der Daten zuständig ist. Die Bilddaten wurden mit einer Kodak DCS 460 Still Video Kamera von einem Helikopter und von terrestrischen Standpunkten aus aufgenommen. Die Genauigkeit des resultierenden CAD Modells beträgt 1.5 cm in allen drei Koordinatenachsen.

1. INTRODUCTION

Many tasks require the generation of precise as-built CAD models of an object, such as art historian studies, monument preservation, archaeology, town and regional planning, renovations, redevelopments, data acquisition for building information systems and others.

Especially for the tasks in art historian studies and monument preservation it is essential to have a very precise and reliable foundation for the research about the monument. And there is an unchanged demand for the surveying and documentation of the cultural heritage. For example the world heritage list (UNESCO, 1997) has at the moment about 552 properties in 112 states parties inscribed, but only a minority of them is sufficiently documented.

In the field of regional and town planning, there is a constant need for studies and simulations for the effects of planned changes, which require up-to-date data of the environment. Renovations and redevelopments of buildings take over an continuous growing part of the overall building activities. Therefore existing plans and models have to be updated or to be generated at all.

For the maintenance of existing buildings the use of building information systems increases. This requires the acquisition of up-to-date three-dimensional data of the existing building, which requires approximately 80-90% of the work for such an information system (Runne, 1993).

Most of the photogrammetric work for the reconstruction and documentation of buildings and monuments for cultural heritage purposes is still performed with analytical plotters (Dallas, 1996). However, there are systems available that establish an on-line dataflow based on digital monoscopic or stereoscopic image measurements (e.g. Elcovision 10 (PMS, 1998), CDW (RolleiMetric, 1998), PhotoModeller (Eos, 1997), PHIDIAS (Benning und Effkemann, 1992), Hazmap (Chapman et al., 1994), DVP (Nolette et al., 1992)). Moreover there are systems that make use of some automatic image measurement techniques like image matching or feature extraction (e.g. StereoView (Pitschke und Gorny, 1993), DPA (Peipe et al., 1993), VCM (El-Hakim and Pizzi, 1993)). All of these systems are using CAD models, but most of them only for the representation of the photogrammetric generated results.

On the other side currently more and more systems come up, that accept CAD information prior the measurement process. Such a system is the modelling-and-rendering system developed at the University at Berkeley (Debevec et al., 1996). It uses a rough object description to guide a stereo matching technique for the reconstruction of object details. Another one is developed at the University of Helsinki (Haggrén and Mattila, 1997), where a functional 3D model of indoor scenes is built first and the measurements of the geometry based on video images are performed thereafter. And a system under development at the University of Delft (van den Heuvel and Vosselman, 1997) makes use of a priori geometric object information in the form of parameterized object models with image lines as the main type of observations.

This paper describes the performance and results of the three-dimensional reconstruction of the Disentis monastery from high resolution still video imagery with the software environment DIPAD (digital system for photogrammetry and architectural design). The object oriented measurements in DIPAD are guided by a topologic model of the object. In Sec. 2 an overview of the project is given. Sec. 3 describes the use of object models for the object oriented measurement approach within DIPAD, and in Sec. 4 the results of the processing for the monastery are presented.

2. PROJECT DESCRIPTION

2.1. Object

The monastery of Disentis (see Fig. 1), located in the region of Surselva in the Swiss Alps, is one of the oldest monasteries of Switzerland.

Figure 1: Disentis Monastery
(viewed from south and east)

Its history reaches back into the 7th century, where the bishop Ursizin initiated a monastery in Disentis. After the reformation the monastery became a center for humanistic and natural sciences. During acts of war at the end of the 18th century the monastery was destroyed and it took a century before it was rebuild.

The actual building is a uniform complex which matches quite well with a rectangular ground plan. However, the structure of the building has changed many times during its long history. Parts of the building were added, destroyed, or after destructions rebuilt. In the 11th century a cloister and two centuries later the eastern tower and the western living rooms were added to the building. The south facade and the middle wing were finished in the baroque period and finally, at the beginning of the 18th century, the church with two towers was added.

The approximate dimensions of the building are 120 m by 60 m in plane and each of the two towers has a height of about 45 m. More detailed information about the monastery and its history can be found in (Schönbächler, 1992) and (Bürke, 1984).

2.2. Camera

The imagery for this project was acquired with a high resolution still video camera. The main advantage of still video cameras is the fast and easy image acquisition with the feel of an SLR camera. These systems combine image acquisition, analog-digital conversion, storage device and power supply in one system, and offer the possibility to control the quality of the acquired images immediately on the spot and use them directly for further processing.

The Kodak DCS 460 still video camera (Kodak, 1998) (see Fig. 2) with its high resolution CCD sensor (2036 x 3060 pixel) provides an internal slot for PCMCIA drives as image storage device with a capacity up to 54 digital images. The camera features virtually all standard functions, including autofocus, automatic exposure, metering modes, flash and self-timing.

The sensor unit is mounted on a slightly modified Nikon camera body. All standard Nikon lenses can be applied with the restriction of a slightly reduced viewing angle due to the smaller photosensitive area of the CCD chip (18.4 mm x 27.6 mm) compared to a standard film.

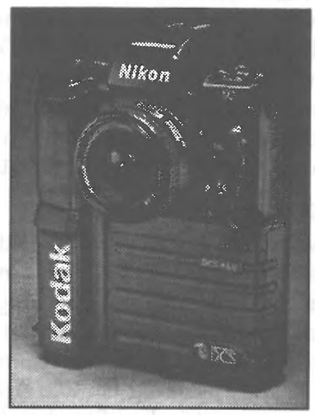


Figure 2: Kodak DCS 460

2.3. Image Acquisition

During the project basically two types of images were acquired. One set of images was taken from a helicopter (see Fig. 3a) and another set of images was taken from terrestrial viewpoints (see Fig. 3b)

Depending on the set of images different facts restricted the image acquisition. The restrictions for the terrestrial images were mainly given on the north and west side of the building due to occlusions and surrounding objects. In addition the situation allows only relatively short object distances. The image acquisition from the helicopter was restricted due to the minimum flying height permitted. In this case it was 50 m above ground. Thus the imagery was taken using two different lenses (28 mm and 18 mm), depending on the average object distances of the two image sets.

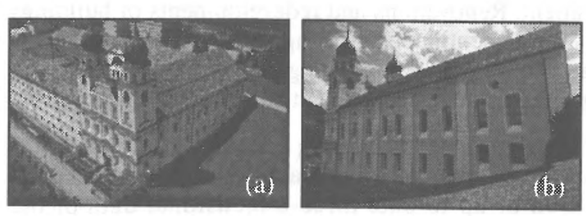


Figure 3: Examples for acquired images
(a) image from an aerial viewpoint with the 28 mm lens
(b) from a terrestrial viewpoint with the 18 mm lens

A total number of 49 images was acquired during the project. 19 images were taken from the helicopter with the 28 mm lens. These images have an image scale of about 1 : 2850 and the average object distance for theses images is about 80 m. A total of 30 images was acquired from terrestrial viewpoints using the 18 mm lens. These images have a mean object distance of about 25 m, which results in an image scale of about 1 : 1400. An overview of the camera configuration is given in Figure 4.

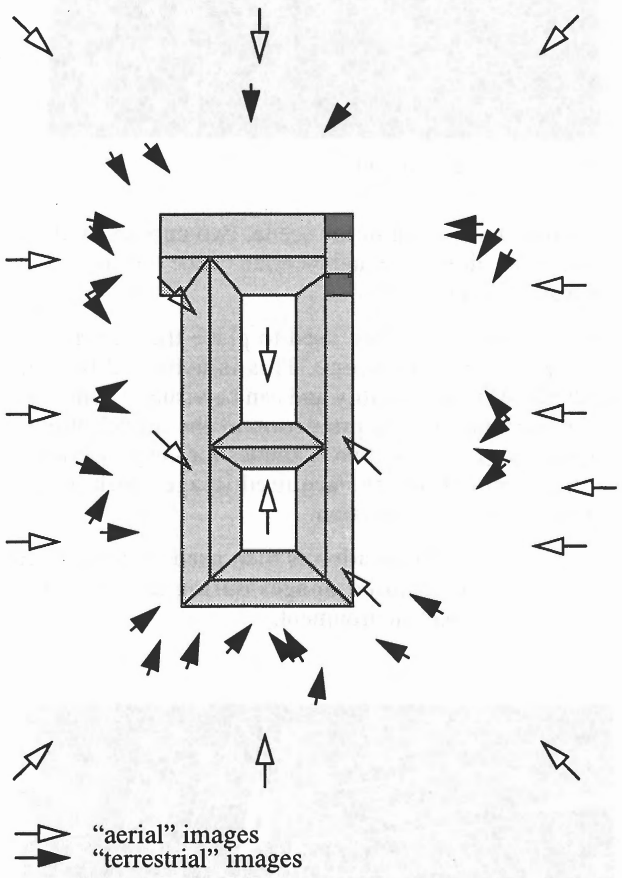


Figure 4: Camera configuration

3. OBJECT ORIENTED MEASUREMENT WITH DIPAD

The reconstruction of man-made objects is a non-trivial task. They are often complex, irregular, appear different according to their function or context, etc. In a typical non-controlled environment like outdoor scenes, their extraction from imagery is difficult due to occlusions (from other objects or due to perspective projection), illumination effects (shadows or weak contrast), radiometric interferences or varying background.

DIPAD aims on the automated and object oriented generation of as-built CAD models. It combines a CCD-sensor based image acquisition with a semi-automatic processing of the image data in a CAAD controlled environment. The main features of the system are:

- possibility of self-diagnosis (quality control),
- potential for high accuracy and reliability (redundant sensor data),
- flexibility with respect to the three-dimensional reconstruction of buildings or parts of buildings, and
- performance of object oriented measurements.

The problem of object recognition and measurement is solved in a way that the image interpretation task is done by the user (architect, art historian, etc.) and the reconstruction and measurement of the precise geometry from multiple images is performed automatically by the computer (HICOM principle). A human operator makes easily use of his knowledge about the real world (mostly subconsciously) while looking at an image and can easily filter the necessary information from the images for his task and/or complete the missing information in his idea. Following a combined top-down and bottom-up strategy a coarse given CAD model will be iteratively refined until the desired degree of detail is achieved.

3.1. Object Models

Object models can be treated as abstractions of real world objects. These are necessary in order to process objects of the complex and extensive reality in a computer environment. Each attempt to represent reality is already an abstraction. The most important role played in the definition of models is the proper balance between correctness and tractability, i.e., the results given by the model must be adequate both in terms of the solution attained and the cost to attain the solution.

There are several ways to describe an object in a CAD environment. In general, 3D models can be divided into three different classes of models: wireframe models, basically defined through vertices and their connecting edges, surface models, describing objects as an ordered set of surfaces, and volumetric models, describing objects by volumes. The class of volumetric models as the most interesting one comprises more sub-classes, such as parametric models, sweep representation schemes, cell decomposition schemes, boundary representations, constructive solid geometry, hybrid models and others. Each of these classes has its specific advantages and disadvantages for different tasks. But there is no class which is optimal for all tasks.

The formal data structure in DIPAD (see Fig. 5) consists of two data sets, the photogrammetric data, which contains all the information about cameras, images and stations, and the object data. The object data consists of three related data structures: the geometric data, the topologic data and the thematic data of the object. The topologic part of the object model consist of six classes, which represent the hierarchical structure of the object. The elements of a hierarchical class consists of the elements of the next lower hierarchical class. The six classes for the topologic data are vertices, edges, areas, volumes, constructive units and objects. The three classes of geometric primitives contain the geometric description of the corresponding topologic element. These classes are points, lines and surfaces.

Beside the topologic and the geometric classes there are also five classes of thematic attributes, which correspond to the topologic primitives. These attributes contain informations which are not of topologic or geometric nature but are necessary information for other tasks than the measurement process. Such information for example can be the history of a construction unit, the material or the condition.

The relations between these classes are basically of three different types:

- complex objects are defined by elementary objects or primitives (e.g. vertices define edges),
- relations, that can be determined from coordinates (e.g. intersections, neighbourhoods, “part of”),
- and coded relations, that can not be determined from coordinates (e.g. relations between attributes, relations between topologic and thematic data).

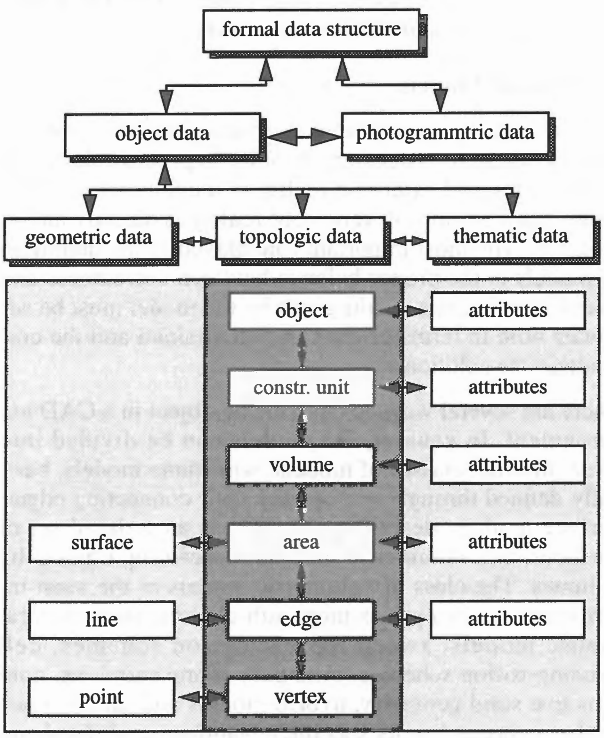


Figure 5: Formal data structure in DIPAD

3.2. Object Modeller

The object modeller figures as a user interface in order to generate the topology of the object which is then used for matching with the actual image data in the measurement routine. Practically the object modeller makes use of a standard CAD system (AutoCAD) which is extended by photogrammetric functionality. This is achieved by making use of the command language AutoLISP and AutoCAD’s Advanced Modelling Extension (AME) (Hirschberg and Streilein, 1995), (Hirschberg, 1996).

As the topology in AutoCAD can not be given without a geometry two tasks are fulfilled within the object modeller, the modelling of the object topology with an approximate geometry and the approximate positioning of the cameras in the scene.

Figure 6 shows the initial model used for the reconstruction of the monastery. It is basically a cuboid with the approximate dimensions of the building and for the sake of

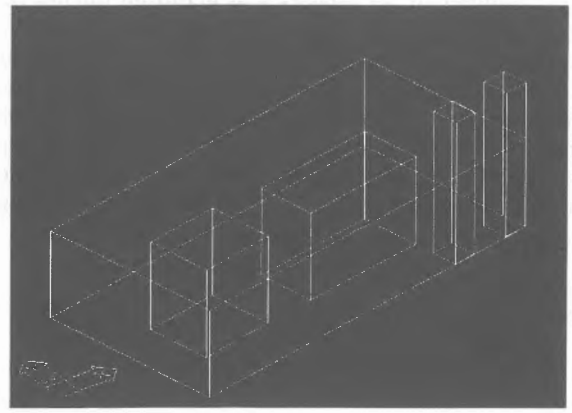


Figure 6: Initial model

convenient navigation in the scene, two cuboids with the approximate dimension and location of the towers and the courtyards respectively.

This initial model is then used to place the cameras approximately into the scene. This is achieved through standard CAD functionality and can be visually controlled by the user through the projection of the model into the images. Figure 7 shows two examples for the projection of the initial model into the acquired images with the approximate camera orientation.

The approximate orientation is then used to get a visual feedback in the acquired images during the modelling process in the CAD environment.

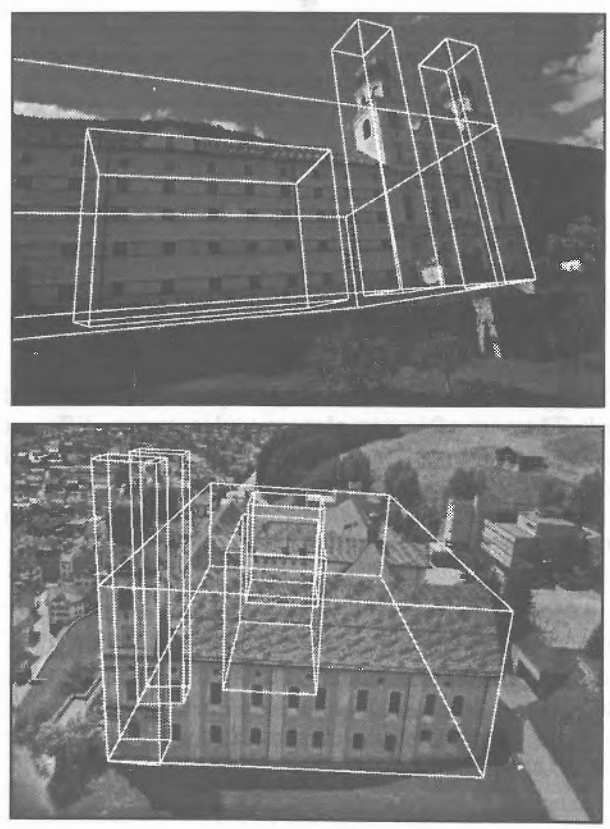


Figure 7: Two examples for the projection of the initial model into acquired images

3.3. Automatic Measurement

The automatic feature measurement employed in DIPAD is guided by the topologic object description created by the object modeller or by known a priori CAD information about the object and calculates simultaneously the best match of the elements described by the object model with the image data of multiple images.

Due to the guidance of the measurement by the topologic object model, only relevant features (as defined by the user) are extracted and redundant or useless information is reduced to a minimum. This strategy follows the native theory of human perception (Gibson, 1950), where information is defined by regularity in contrast to coincidence and not by contents and meaning. This means that a human being can identify a signal much easier the more redundant the signal is. The same counts for computer algorithms. 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. The three-dimensional position of the object is derived by a simultaneous multi-frame feature measurement, where the object model is reconstructed and used to triangulate the object points from corresponding image points.

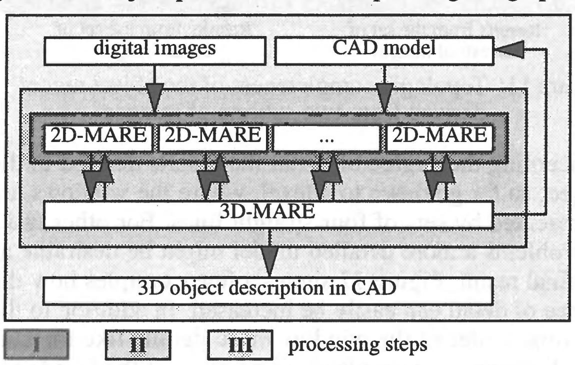


Figure 8: Processing steps in DIPAD (MARE = automatic measurement of architectural elements)

The whole process of object reconstruction can be divided into three processing steps, which run consecutively (see Fig. 8). In the first processing step the simultaneous determination of the object geometry in image space based on the radiometric information in the digital images and the topologic information about the object feature is performed. The routine uses straight lines as the basic entities, by first locating the edges of the features to be measured and then deriving the vertices as intersections of appropriate lines. In the second processing step the precise geometry of the architectural features in object space is determined. Therefore the image coordinates of the first processing step are used to estimate the object coordinates of the architectural feature. In the case of unknown or approximate camera parameters these data are estimated as additional unknowns during the bundle adjustment. The geometric improved features are reprojected into the images in order to restart the first processing step. This iterative procedure continues until the position of the feature in object space after each loop is stable. The third processing step enables the user to increase the degree of detail for the topologic object model. Here more object details can be added to the measurement routine. A more detailed description of the processing steps is given in (Streilein, 1994) and (Streilein, 1996).

An example for the performance of the first processing step on a window feature of the monastery is given in Figure 9. The feature with its approximate geometry is projected with the approximate camera parameters into the image and used as starting value for the automatic measurement. The linear feature boundaries are extracted and straight lines are fitted to the linear feature boundaries. Finally the image coordinates of the vertices are calculated by line intersections.

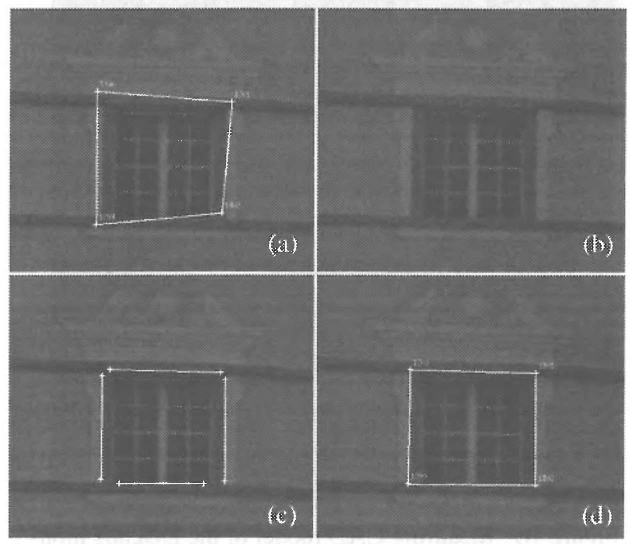


Figure 9: Example for the performance of the first processing step on a window feature
(a) approximate feature position
(b) extracted linear feature boundaries
(c) straight line fitting to linear feature boundaries
(d) vertex computation by straight line intersection

4. RESULTING MODEL

The final result of the processing of the digital images with DIPAD is a topologic and geometric object description of the monastery in the CAD environment. Figure 10 shows the resulting model from different view points and in different representations. The model exists of about 1'800 object points and 1'200 geometric entities.

The 3D geometry of the object was derived by a free network bundle adjustment with self-calibration. The systematic errors of still video cameras employing off-the-shelf lenses with large distortions from the ideal perspective transformation are accounted by extending the colinearity equations with functions of additional parameters. Many additional parameter sets have been developed to meet various requirements, in close-range CCD-sensor based systems a set of ten additional parameters has proven to be effective (Beyer, 1992). These parameters are three changes for the elements of the interior orientation, a scale factor in x direction. a shear factor, the first three parameters of radial symmetric lens distortion and the first two parameters of lens decentering distortion.

The processing of the 3D data was performed in two steps. In a first step the parameters of exterior and interior

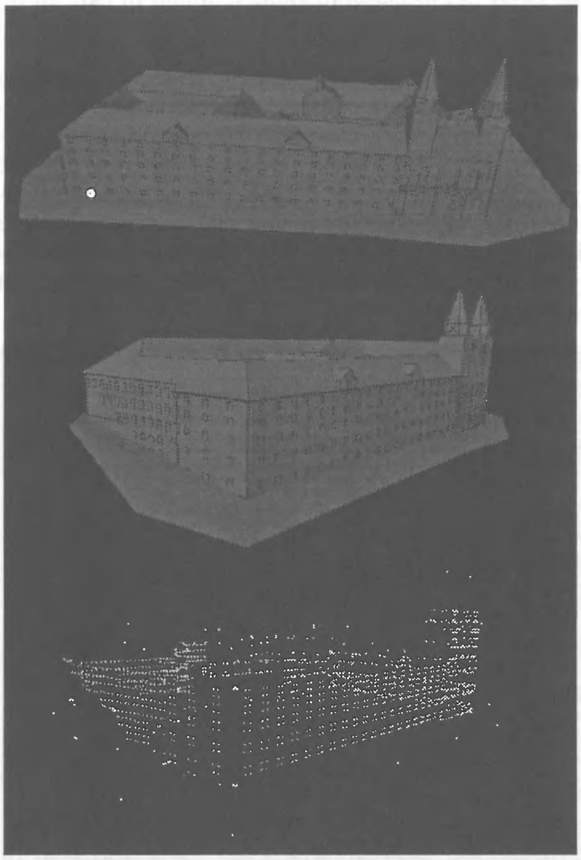
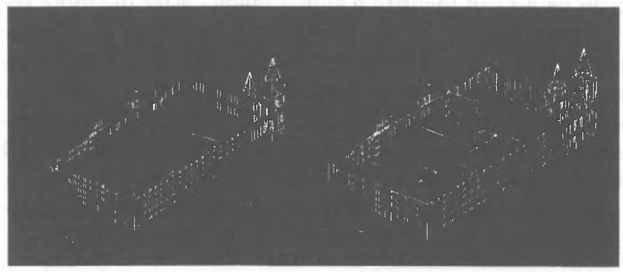


Figure 10: Resulting CAD model of the Disentis monastery

orientation of the cameras were determined together with a subset of the object point coordinates. Due to the two different lenses used, two different sets of additional parameters were determined accordingly. In the bundle adjustment routine 147 parameters of exterior orientation, 20 parameters of interior orientation and 771 objectpoint coordinates were determined from 2620 image coordinate observations. The standard deviation of the unit weight a posteriori ($\hat{\sigma}_0$) of the adjustment was 5.5 μm, which is approximately half the size of a pixel. The theoretical precision of the object point coordinates is 1.3 cm ($\hat{\sigma}_X$), 1.7 cm ($\hat{\sigma}_Y$) and 1.4 cm ($\hat{\sigma}_Z$).

In a second step all object coordinates of the objet model are determined by a bundle adjustment routine using the previously estimated parameters of interior and exterior orientation. This is practically equivalent with a multiple ray forward intersection. The 3D-coordinates of 1836 object points were determined from the image coordinate observation of 9573 image points in 49 images. In average each object point is present in 5.2 images. The accuracy of the object can then be estimated by the theoretical precision of the object point coordinates, which averages to 1.4 cm ($\hat{\sigma}_X$), 1.8 cm ($\hat{\sigma}_Y$) and 1.5 cm ($\hat{\sigma}_Z$). According to the dimensions of the building this is a relative accuracy of about 1 : 8'000.

Beside the geometric accuracy the topologic correctness and completeness of the resulting model is of interest. Whereas the topologic correctness of the model is already supervised by the user during the modelling process, the completeness depends on the degree of detail and the used imagery during the reconstruction process. To demonstrate the latter the topologic object description is matched with the set of aerial images and with the set of terrestrial images separately. The results with respect to the topology are given in Figure 11. It is obvious that the aerial images allow the reconstruction of most parts of the object, whereas the terrestrial images have a significant drawback in the reconstruction of the roof parts and some parts on the facade, which are occluded. Moreover the geometric accuracy of the two subsets suffers from the inferior number of image rays and the worse intersection angles. Thus the accuracy of the set of aerial images is worse by a factor of 2 and the set of terrestrial images by a factor of 3 than the combined set with all images.



Results from the set of terrestrial images | Results from the set of aerial images

Figure 11: Topologic completeness of the object model derived from two subsets of images

Concerning the degree of detail the results derived in the project so far go down to a level, where the windows are represented by sets of four straight lines. For other tasks or problems a more detailed model might be desirable as the final result. Figure 12 shows a few examples how the degree of detail can easily be increased. In addition to the existing model of the window more details like the surrounding parts and details in the painting of the facade are added.

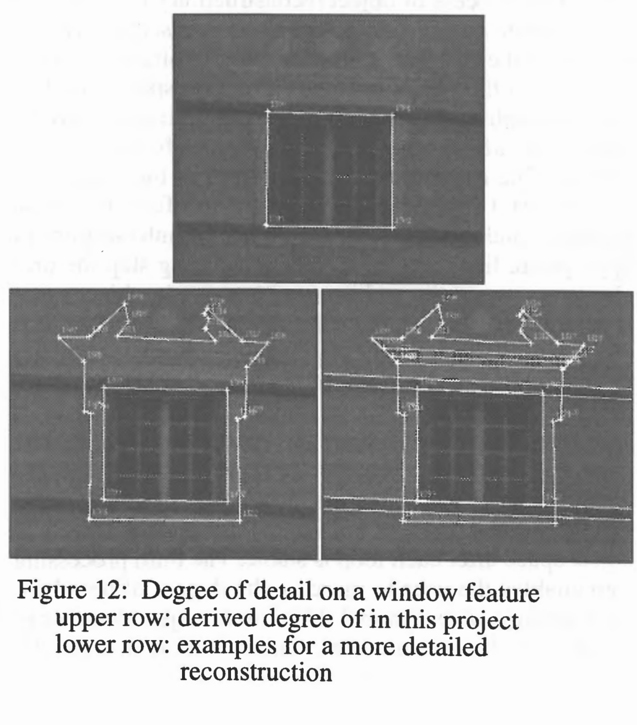


Figure 12: Degree of detail on a window feature
upper row: derived degree of in this project
lower row: examples for a more detailed reconstruction

5. CONCLUSIONS

This paper described the performance and results of the three-dimensional reconstruction of the Disentis monastery from high resolution still video imagery with the software environment DIPAD. The semi-automatic processing follows the HICOM principle, where a human operator assumes the responsibility for the image understanding part and the computer for the object oriented measurement and the data handling. The measurement approach makes use of CAD models for the initialization of the automatic measurement process and for the verification of the measurement results. Accurate image measurement using conventional technology demands expertise, is time-consuming, tiring and not without errors. The semi-automated approach delivers results in less time and more reliable than a user with conventional (manual) tools. The final solution can easily be interpreted and judged by an operator in the CAD environment, which still requires expertise, but is feasible for an untrained person. Thus even larger objects, like the monastery of Disentis, can be reconstructed by untrained persons in a short time period.

6. ACKNOWLEDGEMENTS

The data set used in this example was acquired during the work for a diploma thesis at ETH (Busslinger and Koller, 1998). We would like to thank Matthias Busslinger and Markus Koller for their fine work during the four weeks of their thesis.

7. REFERENCES

Benning, W., Effkemann, C., 1992. Von der digitalen Phototechnik zur digitalen Nahbereichsphotogrammetrie, Entwicklungstendenzen und Anwendungen mit dem System PHIDIAS. Vermessungswesen und Raumordnung, VR 54/8, pp. 420-427.

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

Bürke, B., 1984. Die Marienkirche im Benediktinerkloster Disentis. Festschrift zur Neuweihe 1984, Kloster Disentis, 1984.

Busslinger, M., Koller, M., 1998. Photogrammetrische Generierung eines digitalen 3D-Modells des Klosters Disentis. Diploma thesis, Institute of Geodesy and Photogrammetry, ETH Zürich, 1998. unpublished.

Chapman, D.P., Deacon, A.T.D., Hamid, A., 1994. Hazmap: a remote digital measurement system for work in hazardous environements. Photogrammetric Record, 14(83). pp. 747-758.

Dallas, R.W.A., 1996. Architectural and archaeological photogrammetry. Kapitel in "Close Range Photogrammetry and Machine Vision", Edited by K.B. Atkinson, Whittles Publishing, Caithness, U.K., 1996. pp. 283-302.

Debevec, P.E., Taylor, C.J., Malik, J., 1996. Modelling and Rendering Architecture from Photographs: A hybrid geometry- and image-based approach. Technical Report UCB//CSD-96-893, Computer Science Division, University of California at Berkeley. 1996. 33 pages.

El-Hakim, S., Pizzi, N., 1993. Multicamera vision-based approach to flexible feature measurement for inspection and reverse engineering. Optical Engineering, Vol. 32, No. 9, pp. 2201-2215.

Eos, 1997. PhotoModeller. User Manual, Eos Systems Inc., Vancouver, Canada, 1997.

Gibson, J.J., 1950. The perception of the visual world. Houghton-Mifflin, Boston, 1950.

Haggrén, H., Mattila, S., 1997. 3-D indoor modeling from videography. Videometrics V, Proceedings SPIE 3174. San Diego, 1997. pp. 14-20.

van den Heuvel, F., Vosselman, G., 1997. Efficient 3D-modeling of buildings using a priori geometric object information. Videometrics V, Proceedings SPIE 3174. San Diego, 1997. pp. 38-49.

Hirschberg, U., 1996. Object-Oriented Data-Integration between Digital Architectural Photogrammetry and CAAD. International Archives of Photogrammetry and Remote Sensing, Vol. 31, Part B5, pp. 237-242.

Hirschberg, U., Streilein, A., 1995. CAAD meets digital photogrammetry - modelling "weak forms" for computer measurement. Proceedings of ACADIA '95 "Computing in Design - Enabling, Caputring and Sharing Ideas", 19-22 October 1995, Seattle, Washington, USA. pp. 299-313.

Kodak, 1998. DCS 460 Produkt Information. URL: www.kodak.com/daiHome/DCS/dcs460.shtml.

Nolette, C., Gagnon, P.A., Agnard, J.P., 1992. The DVP: Design, Operation and Performance. Photogrammetric Engineering and Remote Sensing, Vol. 58, 1/92, pp. 65-69.

Peipe, J., Schneider, C.-T., Sinnreich, K., 1993. DPA - Entwurf und Realisierung einer PC-basierten digitalen Arbeitsstation für die Nahbereichsphotogrammetrie. Zeitschrift für Photogrammetrie und Fernerkundung, 2/1993. pp. 75-82.

Pitschke, I., Gorny, P., 1993. StereoView - Ein System zur interaktiven 3D-Geometrie-Rekonstruktion. Zeitschrift für Photogrammetrie und Fernerkundung 5/1993. pp. 185-192.

PMS, 1998b. Elcovision 10 - Produktbeschreibung. URL: www.pms.co.at.

RolleiMetric, 1998b. CDW - Produktbeschreibung. URL: www.rolleimetric.de/deutsch/the_products/software/cdw.html.

Runne, H., 1993. Geodätische Datengewinnung für Gebäudeinformationssysteme unter Anwendung reflektorloser tachymetrischer Verfahren. Geodätische Schriftenreihe der Technischen Universität Braunschweig, Nr. 11, 1993.

Schönbächler, D., 1992. Die Benediktinerabtei Disentis. Gesellschaft fuer Schweizerische Kunstgeschichte GSK, Bern, 1992. 38 pages.

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.

Streilein, A., 1996. Utilization of CAD models for the object oriented measurement of industrial and architectural objects. International Archives of Photogrammetry and Remote Sensing, Vol. XXI, Part B5, Vienna 1996, pp. 548-553.

UNESCO, 1997. World Heritage List. URL: www.unesco.org/whc/heritage.htm