Towards Automation in Architectural Photogrammetry: CAD-Based 3D-Feature Extraction

André Streilein

Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
8093 Zurich, Switzerland

ABSTRACT

Improvements and new developments in the fields of sensor and computer technology have had a major influence on photogrammetry and today a large variety of cameras and photogrammetric systems is available The application of a CAD-based 3D-Feature Extraction routine within a system for Digital Photogrammetry and Architectural Design (DIPAD) is described. This system enables the operator to perform the whole reconstruction of the three-dimensional object without any manual measurement. His task is only to interpret the scene qualitatively while selecting the features to be measured, whereas the quantitative measurement and calculations are performed by the computer. An automatic data transfer of the photogrammetrically generated three-dimensional object description to a CAAD (Computer Aided Architectural Design) system provides new capabilities for architects and conservationists.

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 (*Waldhäusl, 1992*), modern image acquisition and analysis techniques are rarely used. Towards automated extraction of architectural features from terrestrial images the general aim is the development of automated measurement routines, that are adapted to the special needs of architecture and related to CAD databases. As photogrammetry generally records all data as a series of three-dimensional coordinates, it makes a natural partner for today's three-dimensional CAD systems.

Research on Digital Architectural Photogrammetry at the Institute of Geodesy and Photogrammetry of ETH Zurich is focused on the development of a system for Digital Photogrammetry and Architectural Design (DIPAD), which allows the efficient use of image analysis algorithms for the measurement of architectural features and guidance by additional topological information. The development aims at the improvement of photogrammetric data acquisition and processing, the creation of a three-dimensional geometric object description, as well as creating a photo-realistic visualization and guided architectural processing. The results are passed to a data structure useful for CAAD (Computer Aided Architectural Design), including the automatic generation of a Digital Surface Model. The operator is able to perform the whole reconstruction of the three-dimensional objects without manual measurements. His task is to interpret the scene qualitatively while selecting the features to be measured, whereas the extraction of metric information is performed by the computer. As well as the advantage of an on-line flow of data, the operator is supported by the system through the display of all measurements and other known information in the original image.

Published in: ISPRS Journal of Photogrammetry & Remote Sensing,
Vol. 49 No. 5, October 1994, pp. 4-15.

2. PRINCIPLES OF DIPAD

The research environment DIPAD primarily consists of the Digital Photogrammetric Station (DIPS, *Grün and Beyer, 1990*) for photogrammetric processing and a CAAD-system for architectural processing (see Fig. 1). Aims are:

- the accurate and reliable creation of a three-dimensional geometric and semantic object description by photogrammetric means,
- to allow fast and comfortable photogrammetric data acquisition and processing,
- and to provide for visualization and further architectural processing (e.g. simulation, manipulation, animation) in a CAAD environment.

To reach these goals it is necessary to establish an on-line data flow for the entire process from image acquisition to the three-dimensional geometric object description in CAAD. As the hardware platform of DIPS and the CAAD-system are identical, one objective of DIPAD is the considerable integration and interaction of the software of both systems. This results in common data structures for the photogrammetric and CAAD systems.

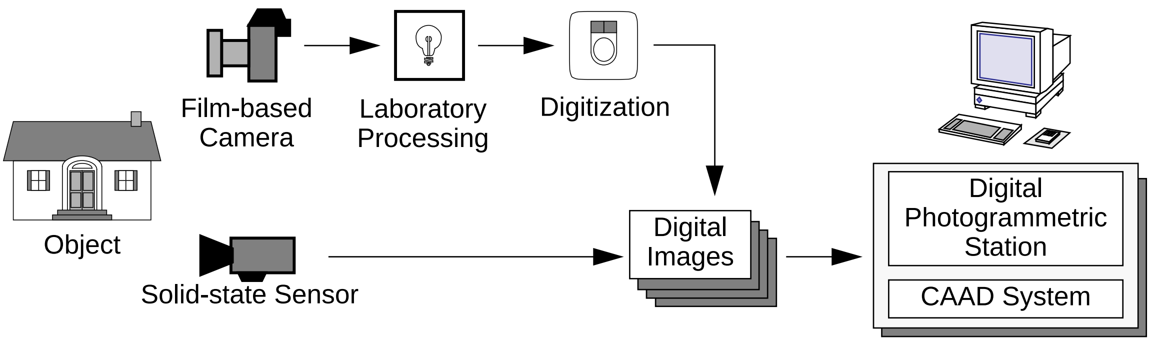


Figure 1: Digital Architectural Photogrammetry with DIPAD.

Digital image data can be acquired with film-based cameras and subsequent digitization, and with cameras using solid-state sensors. Devices of the latter type can be standard solid-state cameras with a video recorder or a computer with framegrabber, still video cameras, video cameras, and several types of high resolution CCD cameras. Conventional film-based cameras still provide an unsurpassed photographic resolution, but the film must be developed and digitized before its data is available for photogrammetric processing. Solid-state cameras, on the other hand, provide immediate access to the image data, which offers the advantage of on-site quality control for the acquired imagery. Today a variety of solid-state imaging systems, from inexpensive camcorders to specialized systems with high-resolution sensors, is available on the market. An overview of solid-state cameras and film-based cameras used in close-range photogrammetry is given in (*Luhmann, 1992*).

It should be mentioned that one major objective of DIPAD is the ability to process image data acquired by any type of solid-state sensor, even inexpensive camcorders or amateur video cameras. The reason for this is that, world-wide, only a small percentage of the buildings of cultural interest are documented, but many images are taken by architects, historians or others, who are interested in architectural photography 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*). If one assumes that the bulk of these images are taken with low-cost equipment instead of metric cameras, the problem of low sensor resolution must be faced. To obtain sufficient accuracy in object space with such low-cost equipment, it is advisable to take as many images as is reasonable and to use multiple camera arrangements with convergent rays, instead of being restricted to special camera arrangements (e.g. stereo pairs). On the other hand, the number of measurements increases rapidly with the number of images. However using digital imagery in a digital system makes it is possible to establish semi-automatic or automatic measurement routines.

To automate architectural photogrammetry several questions have to be posed. For example: what is the type of building; what is the semantic and geometrical resolution required in data acquisition and in the representation; what is the type and quality of the data source; and what should be the degree of human intervention? Complex image data, typical in architectural photogrammetry, requires complex object models. Therefore, we favour semi-automatic measurement routines which follow the principle of using humans for interpretation (high level grouping) and the computer for measurement (fast and accurate). Thus the task of the operator is to interpret the scene qualitatively, whereas the quantitative statement and the data management are performed by computer. This greatly relieves the user from laborious manual work. DIPAD currently supports two different semi-automatic measurement routines for use in architectural photogrammetry. These are Least Squares Template Matching (*Gruen, 1985*) to measure the precise position of signalised or predefined points and the CAD-based 3D Feature Extraction routine for the measurement of architectural features. The first routine is assumed to be well known, whereas the latter is described in the following chapter.

Just as photogrammetry has taken advantage of the new technologies, the methodology of architects also changes. For instance, architects now work with digital models instead of using plans (*Schmitt, 1993*). In many cases, line drawings are not sufficient for the representation of an architectural object. Photogrammetric analysis with DIPAD delivers a three-dimensional geometric and semantic object description, which includes the coordinates of the object vertices with high accuracy as well as the semantic information concerning the features. This allows the automatic creation of a Digital Surface Model in a CAAD environment which is independent of scale and viewing direction (see Fig. 10). 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.

3. CAD-BASED 3D FEATURE EXTRACTION

The basic software component of DIPAD is the CAD-based three-dimensional feature extraction routine. It is a semi-automatic routine, where a series of model- and data-driven processes interact to extract geometric information from the digital imagery. The main principle for this routine is “human for interpretation and computer for measurement”. In order to establish feature extraction that provides for high precision as well as for reliability, a top-down strategy is chosen. A semantic object model is used to detect the features described by this model. Therefore only relevant features are extracted and redundant information and data complexity are reduced to a minimum. In addition, the use of a priori knowledge makes explicit assumptions, which allows the checking of whether or not these assumptions are fulfilled in the images. Another advantage is that this approach acts globally. Small deformations (e.g. image noise, blurred image structures, 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. The geometric information of the object in the digital images is represented by locations and gray values of pixels.

Looking at images taken for architectural photogrammetry it is evident that in the most cases 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 CAD-based 3D feature extraction routine takes advantage of this knowledge. It first locates the edges of the features to be measured and then derives the vertices as intersections of appropriate lines. The idea is to reverse the designing process of an architect, who usually designs in terms of lines or more sophisticated geometric features and constructs the object points as intersections of these lines. In the current version, five different feature types are supported by the algorithm. These are points, lines, open polygons, closed polygons and planes. They are based on straight lines as the basic entity.

Figure 2 gives the flowchart of the current implementation of the CAD-based 3D feature extraction routine. This routine consists of three loops, which are contained respectively in each other. First the inner loop performs the feature extraction, which is based on the radiometric information given by the digital imagery. The middle loop allows the robust determination of object coordinates from multi-frames. Finally the outer loop allows the estimation of camera parameters (interior/exterior orientation) by a bundle adjustment as well as an increase in the degree of detail for the object model. The main processing steps of this routine are addressed in detail below.

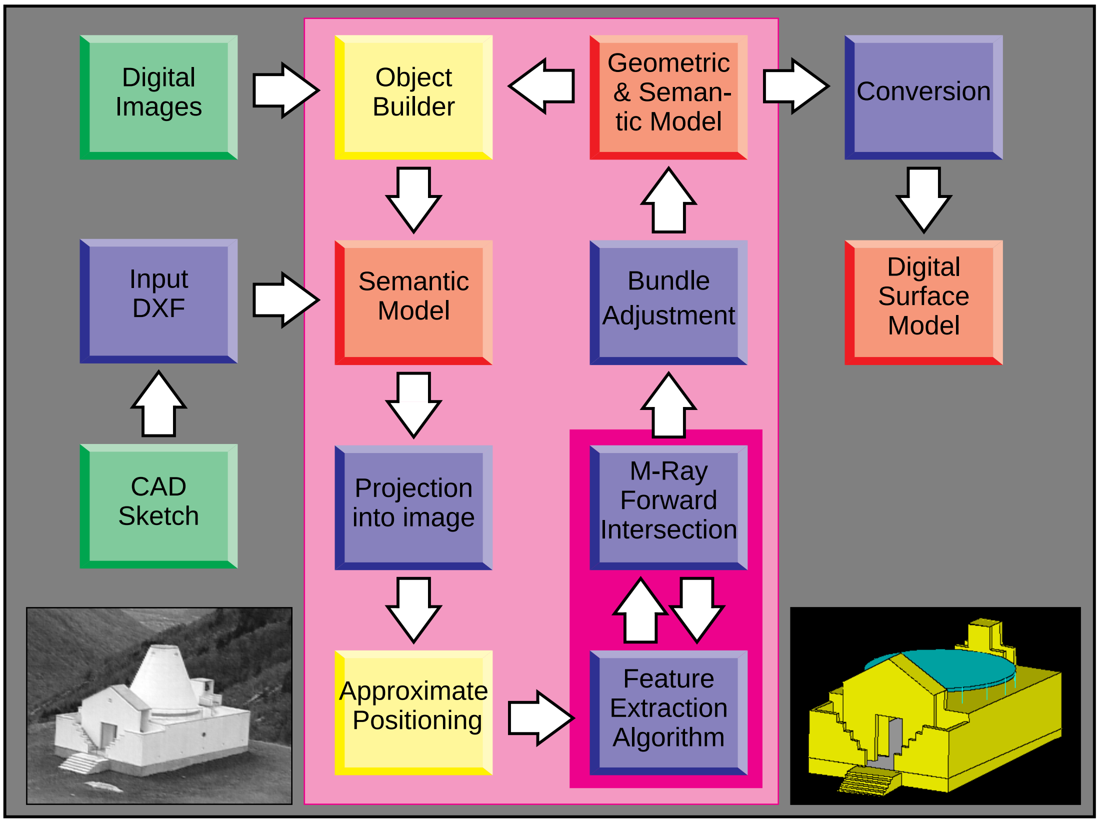


Figure 2: Flowchart of CAD-based 3D Feature Extraction Routine.

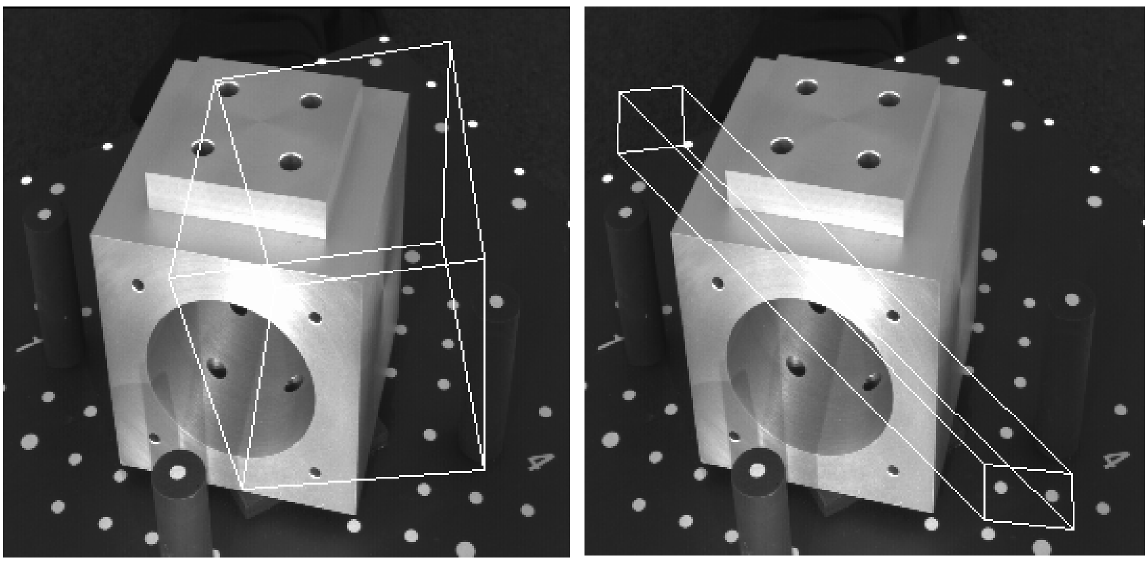
3.1. Input Sources and Semantic Model

The current system allows two different sources for input, the digital image data and a priori architectural knowledge of the object. As mentioned before, the digital imagery can be acquired with solid-state sensors as well as with film-based camera systems and subsequent digitization. The available architectural knowledge about the object (e.g. existing plans or models, construction rules, etc.) can be represented in form of a coarse CAD-model (CAD-sketch). This coarse CAD-model contains the information about the relationship between geometric primitives like points, lines or planes.

The basic idea of the CAD-based 3D feature extraction routine is the refinement of a coarse initial CAD model. Therefore, in a first step, such an initial model (semantic model) has to be created. This problem is currently solved in two different ways, depending on the available a priori architectural knowledge. Existing (and available) object information can be added via DXF (drawing interchange file), an informal standard for CAD systems and, additionally, an object builder enables the operator to add interactively approximations of image space features. This tool provides the display of two images of corresponding object parts on the screen and the user defines image points with the cursor. The system calculates automatically the approximate position of these features in object space by a two-ray forward intersection. This procedure for the definition of object space features is not very accurate, but sufficient to give approximations for the semantic object model.

3.2. Approximate Positioning

In order to guide the feature extraction algorithm properly, the three-dimensional approximation of the features and of the exterior orientation of the cameras can be changed interactively in an optional processing step. Window buttons, as well as the keyboard, allow changes to be made to the six degrees of freedom for the exterior orientation of the camera separately. The visual feedback (see Fig. 3) provides for the easy and fast performance of this tool. In a similar way, the position and form of the topological model can be changed by dragging the model points via the mouse buttons.



(a.) object model correct, orientation of image wrong.

(b.) object model wrong, orientation of image correct.

Figure 3: Interactive improvement of exterior orientation and topological model.

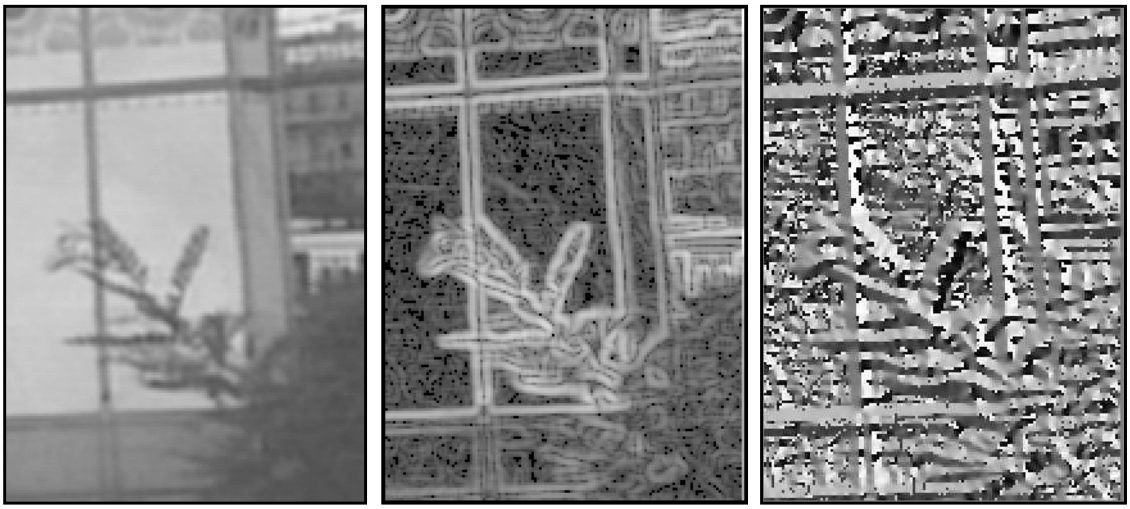
3.3. Feature Extraction Algorithm

The approximation of object space features transformed into image space are used as starting values for the automatic feature extraction algorithm. This algorithm is based on local gradient operators to determine local intensity variations in the image signal. It is based on the assumption that discontinuities or rapid changes in the intensity of the image signal often occur at the physical extent of objects within the image. The localization of an edge in image domain is equivalent to finding the maximum of the first partial derivatives (gradients) of the image intensity function. Gradients for the discrete image function are approximated by filter operators in orthogonal directions. In this algorithm, a row gradient $g_r(j,k)$ and a column gradient $g_c(j,k)$ are determined by the convolution of the image signal with the sobel operators ∇_r and ∇_c. The following convolution kernels are used:

$$\nabla r = \frac{1}{4} \times \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad \nabla c = \frac{1}{4} \times \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Every pixel is assigned a gradient value (see Fig. 4), which is a vector containing a magnitude $G(j,k)$ and a direction $\Theta(j,k)$:

$$G(j,k) = (g_r(j,k)^2 + g_c(j,k)^2)^{1/2}, \qquad \Theta(j,k) = \arctan(g_r(j,k)/g_c(j,k)).$$



(a.) part of original image data. (b.) visualization of magnitude. (c.) visualization of direction.

Figure 4: Gradient values derived by sobel operator.

The position of the edge is then determined with subpixel precision by fitting a second-order polynomial in the direction of the gradient. The maximum point of the fitting curve corresponds to the subpixel position of the edge (see Fig. 5). The covariance matrix of the estimated polynomial parameters represents the accuracy of the edge point.

At the end of this process, a list of single edge points with subpixel precision is obtained, but which are still linked to the image raster. In order to reduce the number of edge points, which contain redundant information and also disturbances/artifacts for the entire edge, this raster data is converted into vector format. All edge points belonging to the same edge are used to determine the edge parameters. The underlying edge model is currently a straight line. The observation values are given by the subpixel position of each edge point and its weight is proportional to the precision of this position. The edge parameters are determined by a Least Squares Estimation and the covariance matrix of the estimated edge parameters represents the accuracy of the edge.

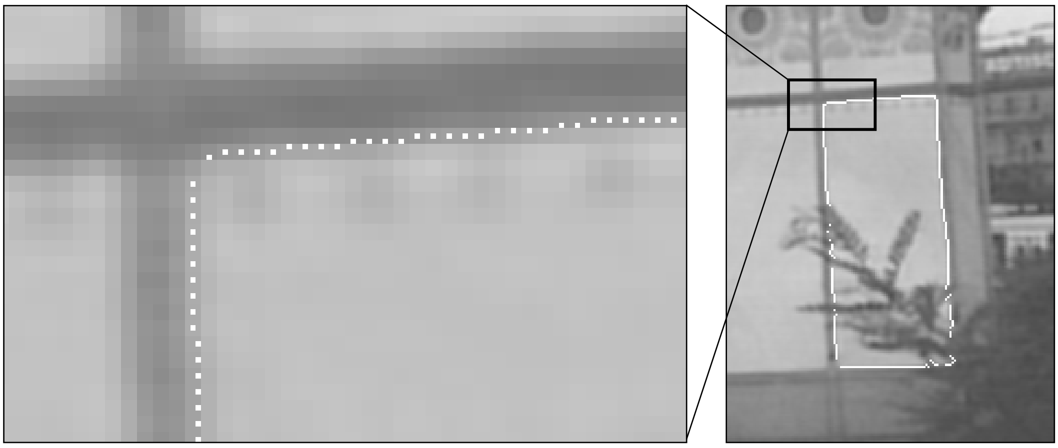


Figure 5: Zoom of extracted image edge points (sub-pixel precision).

If a vertex is defined by the intersection of two or more lines in the image (e.g. L-, T- or Y-junctions), the image coordinates of the vertices of the features are determined automatically as intersections of the appropriate lines. The precision of the vertex coordinates is determined by error propagation. Figure 6 shows the main processing steps of the feature extraction algorithm.

These are:

- Input of approximate feature position (Fig. 6a)
- Following the linear feature boundaries (Fig. 6b)
- Straight line fitting to linear feature boundaries (Fig. 6c)
- Vertex computation by straight line intersection (Fig. 6d)

Beside the precision of the detected edge, another important aspect for the extraction of features is the significance of the detected structures. In architectural photogrammetry problems typically occur due to occlusion, illumination effects (e.g. shadow edges), features fading into background or varying background. In order to handle such cases and to exclude insignificant structures, the algorithm rejects all edge points which do not fulfil two criterions from the list of points contributing to the entire edge. The first criterion is that the orientation of the gradient does not deviate too much from the orientation of the entire edge; the second is that the magnitude of the gradient has to be higher than a user defined treshold. An example of the performance of these criterions can be seen in Figure 6. The occluding structure of the bush in the foreground does not disturb the estimation of points 11 to 14. But there are typically additional circumstances that may create corners in image space, which have no representation as a corner in object space (e.g. a T-junction is created if one object edge is occluded by another). The algorithm handles such cases by consulting the underlying semantic object model, which contains the necessary information about existing (“real”) object points.

This image based feature extraction algorithm is performed sequentially in all images.

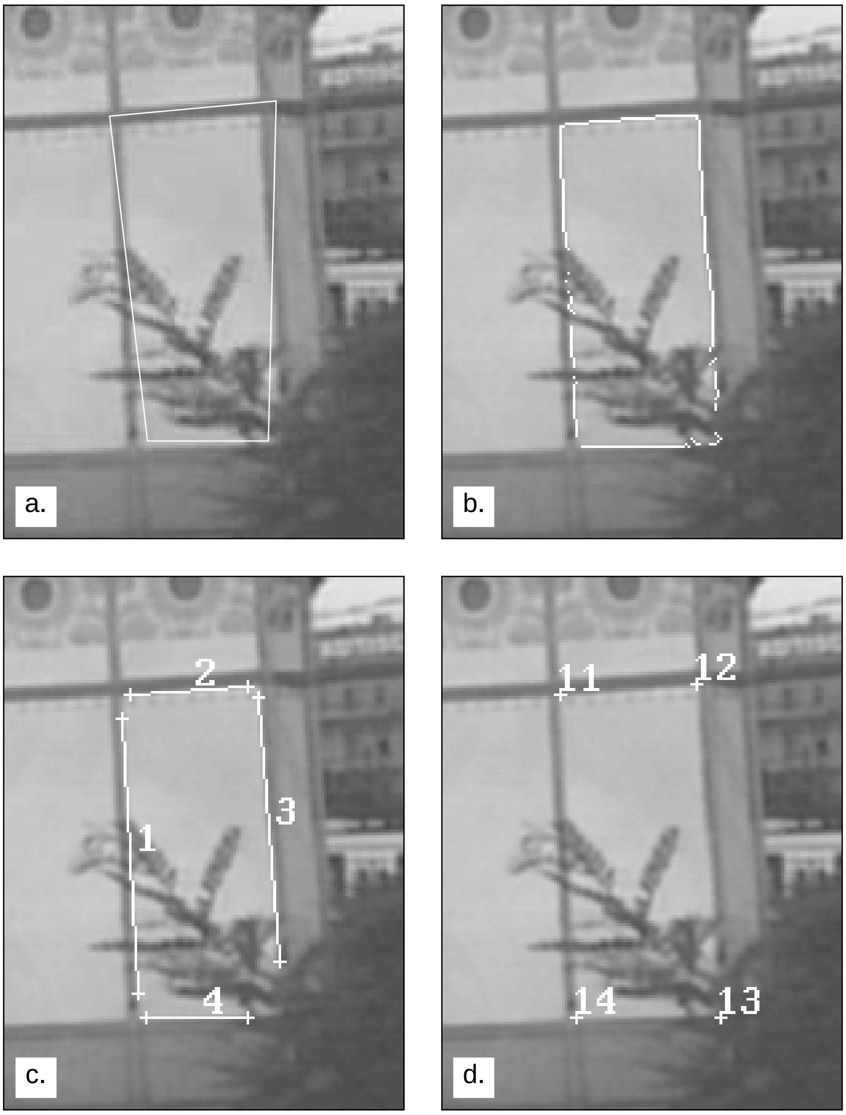


Figure 6: Example for Feature Extraction Algorithm.

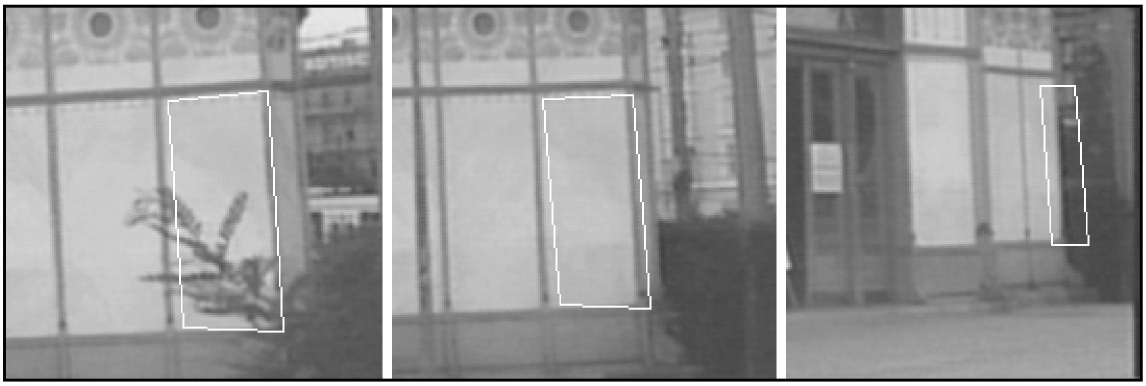
3.4. Feature Extraction Algorithm (three-dimensional loop)

In order to allow a robust estimation of the object features and to increase the radius of convergence for the algorithm, the previously described algorithm is covered by a loop which works basically in object space.

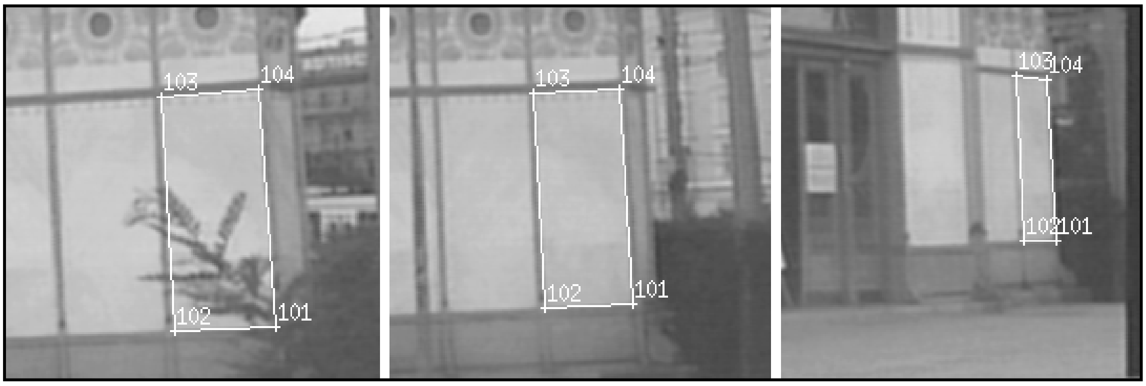
Therefore the object coordinates of a point occurring in two or more images are calculated by a multiple-ray forward intersection, which treats the image coordinates as observations, the object coordinates as unknowns and all other parameters (elements of exterior orientation, interior orientation and additional parameters) as known constants. The calculated object coordinates are then projected, together with the semantic object description, into each image plane and used to restart the feature extraction algorithm. This iterative procedure continues until at least one of the following criterions is fulfilled: (a) the average change of object coordinates is within a user defined limit or (b) a maximum number of iterations is reached. Figure 7 gives an example of the performance of this iterative procedure on a single feature in three images.

In an optional step the precise three-dimensional metric information about the features can be determined by using a bundle adjustment, where all parameters (exterior and interior orientation elements, coordinates of object points, and in the case of self-calibration additional parameters) are estimated simultaneously. This option is available because it is assumed that the parameters of the exterior orientation and/or interior orientation of the camera are most probably unknown in the beginning.

(a.) part of a scene.



(b.) initial position of feature.



(c.) final position of feature.

Figure 7: CAD-based 3d feature extraction (three-dimensional loop).

3.5. Conversion to CAAD

The resulting three-dimensional object description is passed automatically to the CAAD system for visualization and further architectural processing. Thereby a Digital Surfaces Model is generated automatically as all surface features, which are defined by more than three object points, are divided into suitable triangles (“3DFACES”).

In order to pass the results of the photogrammetric processing to the CAAD system, the exchange of data within DIPAD is handled via AutoLISP-files. The CAAD system is suitable for documentation and visualization, as well as complex simulations, manipulations and analysis of the object. Thus it greatly relieves the architect from laborious manual work so he can now concentrate on the creative finding of new solutions or make judgements about the current solution.

4. PRACTICAL EXAMPLE

The following example shows results derived with a former version of DIPAD, which does not offer all the capabilities of the current version described in the previous chapter. Specifically the iteratively working three-dimensional loop was not implemented during the period of this test. This led to a determination of image coordinates combining the image based feature extraction and a manual measurement mode, which uses the cursor as measuring device. On the other hand, the on-line flow of data, the numerical analysis and the automatic data transfer to the CAAD system were established at the time of this campaign. So the results derived are expected to be comparable to the ones which the current system would derive.

Figure 8: Partial view of the test object.

The test object (see Fig. 8) is a building housing drawing rooms for architectural students on the campus of ETH Zurich. It is composed of pre-manufactured parts which can be dismounted and rebuilt (*Blättler, 1989*). The construction of the building follows the rules of a brickbox. It consists of wooden panels, screw connections, four different window types and a roof. These elements are standing on a concrete platform and are divided in the middle by a wall. The dimensions of the building are 55 by 11 m in plan and 8 m in height.

Figure 9: Camcorder JVC GR-S77E.

The image acquisition was performed with a JVC Camcorder GR-S77E (see Fig. 9). This imaging system is a standard consumer product and not conceived for any photogrammetric purposes. The free-hand portable camera allows on-site control for the acquired imagery via an internal monitor. It incorporates a 1/2” colour sensor. The analogue images are stored on a S-VHS video tape and have to be digitized by a framegrabber. The digitized images have a size of 728x568 pixel. The autofocus was disabled and the zoom lens was fixed at its shortest focal length. An extensive analysis of the metric accuracy performance of the JVC camcorder is given in (*Beyer et al., 1992*).

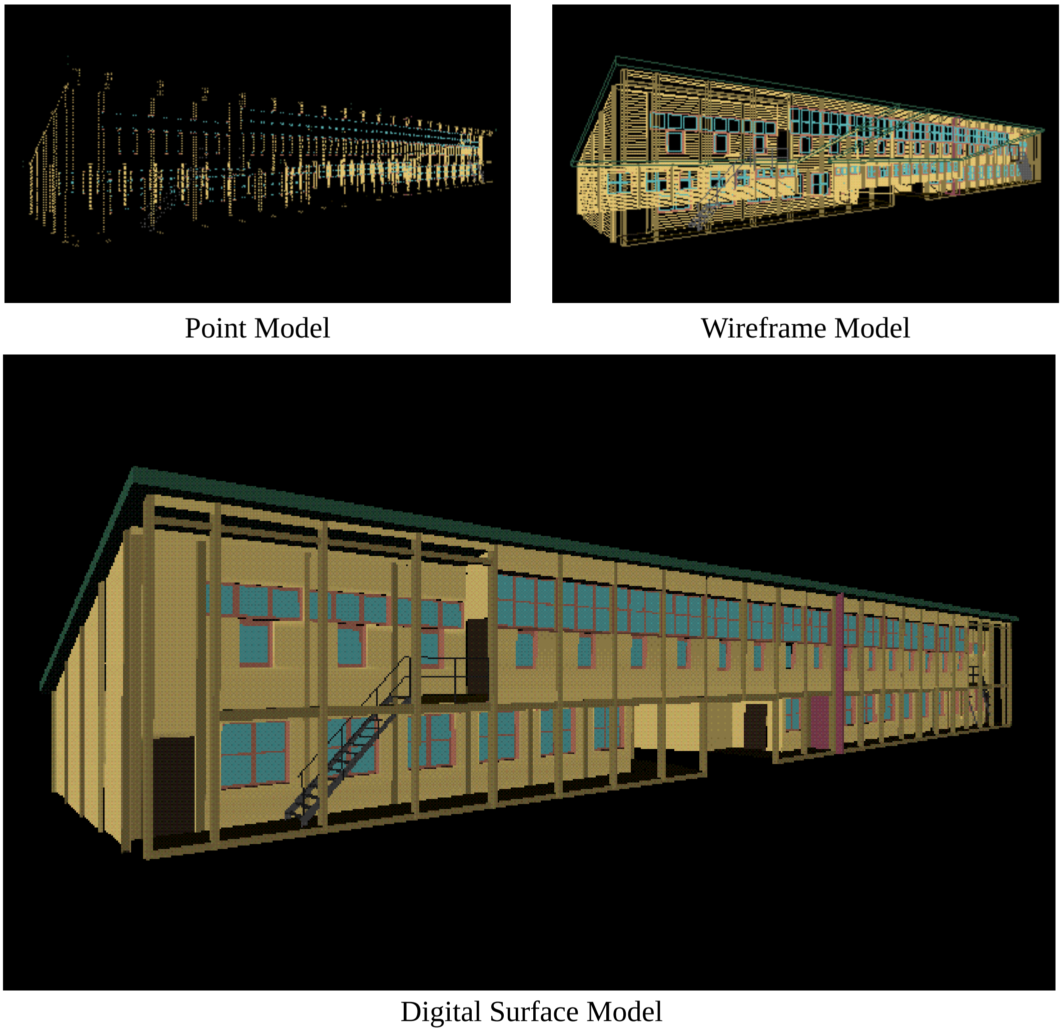


Figure 10: Photogrammetrically generated CAD-Model of test object.

The photogrammetric analysis of the digital image data was performed with DIPAD. Nineteen signalized points were fixed on the facades and used as control points. The camera arrangement was designed to cover the whole building with a minimum number of images providing at least two convergent rays per object point. Therefore 22 images were taken, which results in a mean number of 2.3 rays per object point. Three-dimensional coordinates of 789 object points were determined by a bundle adjustment with self-calibration. The theoretical precision of these object coordinates was determined as a part of the bundle adjustment routine, where the covariance matrix of the estimated object coordinates represents their precision. The results indicate a precision of the points in object space of 1.4 cm within the plane of each facade and 2.7 cm in depth, which is in general sufficient for architectural tasks. The degraded precision in depth is a result of the camera configuration leading to poor intersection angles. A more detailed description of this test, which was also performed with a high-resolution CCD-camera, is given in (*Streilein, Gaschen, 1994*).

But the numerical results are only a part of the final results. For most people, it is more interesting to look at or work with the graphical results. Figure 10 gives a perspective view of the photogrammetrically generated CAD-Model of the test object. It shows the object points (point model) as well as the wireframe model and the digital surface model, which are stored in the same data base and represented within the same system. These can be used separately or simultaneously. In order to demonstrate the tremendous number of details stored in this data base (e.g. 1.1 MByte for the DXF-file of this particular object), Figure 11 depicts a zoomed portion of the whole model.

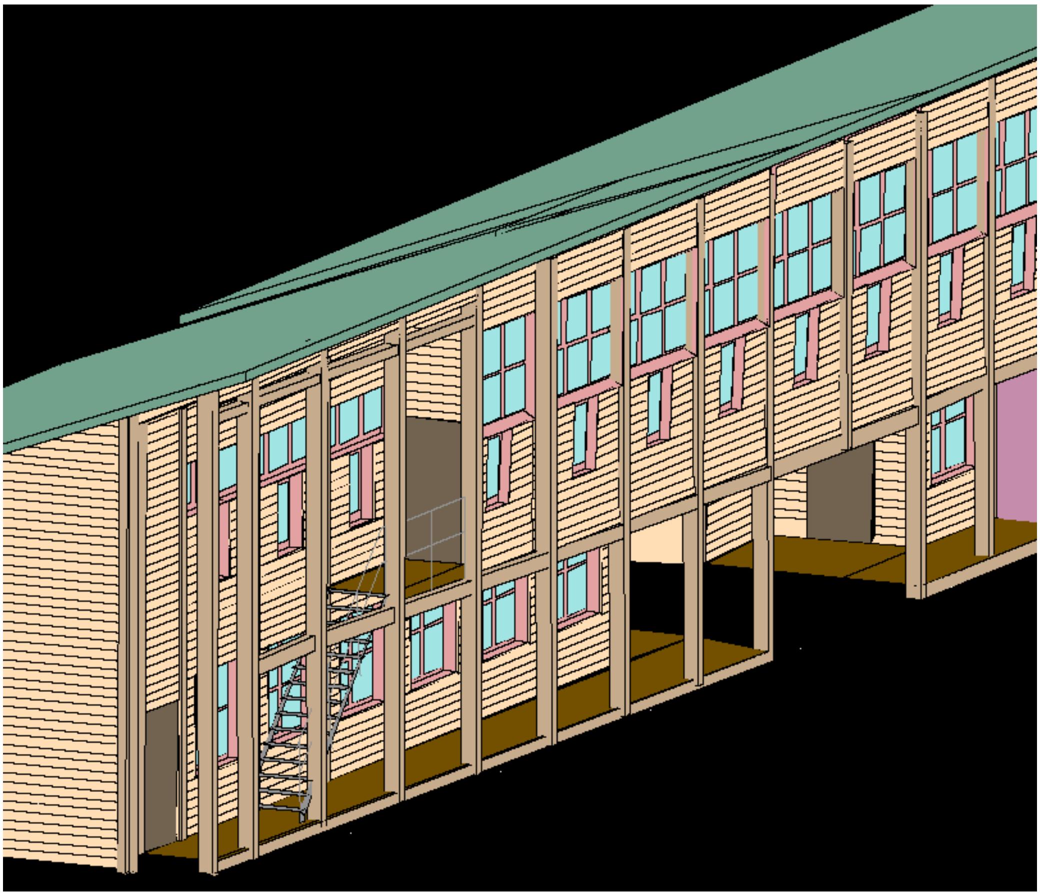


Figure 11: Digital Surface Model in combination with Wireframe Model.

5. CONCLUSIONS

In redevelopment projects, heritage recording is part of an investigative process to achieve understanding about structures so that architects and/or conservationists can make the appropriate design decisions. The CAD-based 3D-Feature Extraction routine provides a first step towards the integration of photogrammetric knowledge with the interpretative process of evaluating the object information and making design decisions. The user is able to perform the whole reconstruction of the three-dimensional objects without manual measurements. His task is to interpret the scene qualitatively while selecting the features to be measured, whereas the extraction of metric information and all other calculations are performed by the computer. Sufficient accuracy for architectural tasks can even be derived when using an inexpensive S-VHS camcorder for the image acquisition, but they face the problem of low sensor resolution and therefore of interpretation. The automatic data transfer of the photogrammetrically generated Digital Surface Model to the CAAD system gives a flexible three-dimensional geometric object description. The CAAD system is suitable for documentation and visualization, as well as for complex simulations, manipulations and analysis of the object. Digital photogrammetry allows new capabilities for architects and conservationists. The new forms of representation provide a better tool for analysis and a completely new instrument for design.

6. REFERENCES

Beyer, H., Kersten, T., Streilein, A., 1992. Metric Accuracy Performance of Solid-State Camera Systems. Paper presented at the SPIE OE/Technology’92 - Videometrics conference, 15-20 November 1992, Boston, MA. Published in SPIE Vol. 1820, pp. 103-110.

Blättler, E., 1989. Neue Architektur in Zürich - ausgewählte Objekte ab 1920. Niggli-Verlag, Heiden, 1989. pp. 72-73.

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., Beyer, H., 1990. DIPS II - Turning a Standard Computer Workstation into a Digital Photogrammetric Station. International Archives of Photogrammetry and Remote Sensing, Vol. 28, Part 2, pp. 247-255 and ZPF - Zeitschrift für Photogrammetrie und Fernerkundung, Nr. 1/91, pp. 2-10.

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

Schmitt, G., 1993. Virtual Reality in Architecture. Chapter in “Virtual World and Multimedia“, Nadia Magnenat Thalmann (Ed.), Wiley&Sons, New York, 1993.

Streilein, A., Gaschen, S., 1994. Comparison of a S-VHS Camcorder and a High-Resolution CCD-Camera for use in Architectural Photogrammetry. International Archives of Photogrammetry and Remote Sensing, Vol. 30, Part 5, pp. 382-389.

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.