UTILIZATION OF CAD MODELS FOR THE OBJECT ORIENTED MEASUREMENT OF INDUSTRIAL AND ARCHITECTURAL OBJECTS

André Streilein
Institute of Geodesy and Photogrammetry
Swiss Federal Institute of Technology
CH-8093 Zurich, Switzerland
e-mail: andre@geod.ethz.ch
Commission V, Working Group 3

KEY WORDS: CAD, digital close-range photogrammetry, model based analysis, object oriented measurement.

ABSTRACT
An object oriented measurement approach using CAD models for the initialization of an automatic measurement process and for the verification of the measurement results is presented. CAD models are used in an a priori and a posteriori mode. The human operator assigns responsibility for the image understanding part (high level grouping), and the computer for the actual measurement and the data handling. The performance and results of this measurement approach is demonstrated on an architectural object.

1. INTRODUCTION
This paper describes an object oriented measurement approach using CAD models for the initialization of an automatic measurement process and for the verification of the measurement results. This measurement approach is embedded in the software environment DIPAD, currently under development at ETH Zurich.

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 overruling principle is, that the human operator assigns responsibility for the image understanding part (high level grouping), and the computer for the actual measurement and the data handling. The user indicates relevant parts of the object in the CAD environment by approximating a geometric topology to it. The photogrammetric algorithm matches this topology with the image data of multiple images, iteratively refines the coarse object model that is given and transfers the final result back to the CAD environment, including information about precision and reliability of the result.

In Sect. 2 the ability of CAD systems for the representation and structuring of 3D data is discussed. Sect. 3 gives an overview of the modelling and measurement principles of DIPAD. And finally in Sect. 4 results of this measurement approach will be demonstrated on an architectural object, the Otto-Wagner-Pavillon in Vienna).

2. MODEL REPRESENTATION AND DATA STRUCTURES OF CAD
For the representation of three-dimensional data CAD models have been used widely in photogrammetric applications. Various examples can be found in the field of architecture (Saint-Aubain, 1990; Stevens and McKay, 1990; Kempa and Schlüter, 1992; Albertz and Wiedemann, 1995), in the field of heritage recording (Robson, et al., 1994; Sawyer and Bell, 1994), in the field of city models (Lang and Schickler, 1993; Gruber, et al., 1995; Mason and Streilein, 1996), or in the field of industrial applications (Oshima, 1992; Chandler, Still, 1994; Chapman, et al., 1994). Examples can be found even for the design of close-range photogrammetric networks (Mason, 1994).

Different software packages employ CAD models. Some of which are address in (Fellbaum, 1992), where an overview of low-cost photogrammetric systems using CAD is given and in (McGlone, 1995) where an overview of various systems in computer vision using CAD models is given. Most systems use CAD models exclusively for the representation of photogrammetric results. However, there are systems, who accept or require CAD information prior the measurement process (e.g. the systems described by El-Hakim and Pizzi (1993) or Schickler (1992)).

Object models can be treated as abstractions of real world objects. The most important role played in model definition 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.

The information usually conveyed in any CAD application is geometry, and the most common way of introducing geometry is by using the concept of building blocks. Every object maybe decomposed into a small number of geometric pieces or graphic primitives. Basic graphic primitives, e.g. points or dots, straight lines, polylines, filled areas, exist both in 2D and 3D environments. In addition there are advanced primitives in 2D, e.g. circles,
arcs, ellipses, conic curves, Bezier curves, B-spline curves; and in 3D, e.g. surfaces, solids (rectangular boxes, cylinders, spheres, wedges, cones). The usage of such primitives permits the representation of a complex solid, by gathering and combining them in order to approximate the real 3D form.

The following are some known and frequently used geometric representations employed in CAD systems. **Boundary Representation (BREP):** The object is represented by its boundary, which is decomposed into a set of faces, edges and vertices. The result is an explicit model based on a combined topological and geometric description of the object. The topology is captured by a set of relations that indicate how the faces, edges, and vertices are connected to each other. **Sweep Representation:** A solid is defined as the volume that swept by a planar or a two-dimensional shape along a curve. There can be both rotational and translational sweep, e.g. generalized cylinders or cones. **Constructive Solid Geometry (CSG):** Primitive instances of objects are combined to more complex objects by use of geometric transformations and Boolean set operations. The representation is usually in form of a tree where the leaf nodes correspond to the primitive instances and the non-leaf nodes correspond to Boolean set operations. **Cell Decomposition:** This yields a decomposition of object space into cells, usually of different sizes. The region octree is a regular decomposition of this representation. An octree method consists of dividing the working space into cells marked as occupied or empty.

For the structuring of CAD models there are a number of functionalities and methods that standard CAD systems offer (Schmitt, 1993). Probably the first and foremost that comes to mind is the structuring in layers (incl. colours and linestyles). If it is applied efficiently this can be a very powerful tool as it allows for multiple groupings of the same element. The concept of blocks (types and instances) introduces hierarchies and higher level groupings. It allows fast modelling and testing of alternatives. The concept of levels of detail (logical zoom) generates different modes of representation depending on the task. These are the most used means of structuring a model in standard CAD systems. Additional structuring can be achieved by extending the drawing database or by linking an external database to the drawing database.

### 3. STRATEGY OF DIPAD

The reconstruction of man-made objects is difficult. 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.

The strategy of DIPAD is that a human operator assigns responsibility for the image understanding part (high level grouping), and the computer for the actual measurement and the data handling. CAD models are used both, in an a priori and a posteriori mode. The user indicates relevant parts of the object in the CAD environment by approximating a geometric topology to it. The photogrammetric algorithm matches this topology with the image data of multiple images, iteratively refines the coarse given object model and transfers the final result back to the CAD environment, including information about precision and reliability of the result.

#### 3.1. Modelling environment

A CAD system (in our case AutoCAD) supports 2D and 3D modelling facilities and controls hybrid viewers, which serve as a general feedback generator during the modelling process. The camera-icons in the CAD system are linked with corresponding camera parameters of the image viewer. When the parameters of a camera-icon are edited in the CAD system, the changes of its view can be monitored in the viewer. The intention is to facilitate the initial positioning of a camera with respect to the model and for the creation of a coarse object model for the measurement process (see Figure 1). Moving around these icons can however also be used to do real-time fly-throughs in the model or for the production of montages and the creation of new objects in existing contexts.

![Figure 1: The modelling in the CAD system can be entirely monitored in the viewers linked to the model. Each camera icon corresponds to one viewer.](image_url)
The constrained relationships between different objects, can be introduced by attaching or nesting the individual (parametric) objects to one another. Both methods, nesting as well as attaching, define a fixed geometric relationship between two objects: they can share an edge, have coplanar sides or the like. Nesting an element also introduces a hierarchy in the model which can be used as a means to provide different levels of detail. A more detailed description of the modelling concept is given in (Hirschberg, 1996).

3.2. Feature measurement

The feature measurement employed in DIPAD is a semi-automatic routine, where a generic 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. 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 extraction, where the object model is reconstructed and used to triangulate the object points from corresponding image points.

Although boundaries of objects are only a small percentage of the whole image content, they have major importance for the description of object discontinuities. This 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 appropriate lines. This routine consists basically of three loops (see Fig. 2). First an internal loop performs the feature extraction in 2D-space, based on the radiometric information given by the digital imagery. An external loop allows the determination of object coordinates from multi-frames, and finally an orientation loop provides for the estimation of the a priori unknown camera parameters (interior/exterior orientation) by a bundle adjustment.

3.2.1 Internal loop

The approximation of object space features transformed into image space are used as starting values for the two-dimensional automatic extraction algorithm. It is based on the assumption that discontinuities or rapid changes in the intensity of the image signal often occur at the physical extent (edges) of objects within the image.

Local searches are carried out at regular intervals along directions perpendicular to the approximate (a priori) boundary. A Sobel edge operator (Eq. 1) is applied to each of the discrete points along each of these perpendicular directions.

\[
\begin{bmatrix}
1 & 2 & 1 \\
2 & 0 & -2 \\
-1 & -2 & -1 \\
\end{bmatrix}
\]

Every pixel is assigned a gradient value, which is a vector containing the amplitude and the direction. The spatial gradient amplitude is given by:

\[
g(j, k) = \sqrt{g_r(j, k)^2 + g_c(j, k)^2}
\]

The direction of the spatial gradient with respect to the row axis is:

\[
\theta(j, k) = \arctan \frac{g_c(j, k)}{g_r(j, k)}, \quad -\pi < \theta < \pi
\]

For each such direction, the pixel with the highest amplitude within a user defined search window, starting from the approximate line, is pre-selected as an edge pixel. This pixel has to be confirmed by applying the same search window, now starting from the previously selected pixel. This procedure continues until the selected pixel remains at its position.

The subpixel position of the edge is then determined by fitting a second-order polynomial (parabola) in the direction of the gradient. The maximum point of the fitting curve corresponds to the subpixel position of the edge (see Figure 3).

The coefficients \(c_0, c_1, c_2\) of the parabola

\[
f(u) = c_0 + c_1 \cdot u + c_2 \cdot u^2
\]
are estimated in a least-squares adjustment. The vector \( x \) of unknown parameters \( (c_0, c_1, c_2) \) is determined by minimizing the sum of squares of the estimated residuals as

\[
\hat{x} = (A^T P A)^{-1} \cdot A^T P y 
\]

(Eq. 5)

The precision of the estimates \( x \) can be derived by error propagation.

\[
\hat{D}(\hat{x}) = \hat{\sigma}_0^2 \cdot Q_{\hat{x}} = \hat{\sigma}_0^2 \cdot (A^T P A)^{-1} 
\]

(Eq. 6)

As the result should not depend on a user defined window size, the determination of the best fitting parabola is repeated with different window sizes of odd number, growing from a width of five pixel until the estimated results are not improved by the next higher window size.

The extremum of the best parabola gives the subpixel solution of the line point \( (u_0) \):

\[
\frac{\partial f(u)}{\partial u} = 0 \quad u_0 = -\frac{c_1}{2 \cdot c_2} 
\]

(Eq. 7)

The precision of subpixel solution can be estimated by error propagation with the known covariance matrix of the estimates \( x \) (parameters of parabola):

\[
D(y) = A \cdot D(x) \cdot A^T 
\]

(Eq. 8)

Finally a pixel is accepted as a line pixel, if

- the extremum of the parabola falls inside the pixel, and
- \( c_2 \) is negative and differs significantly from zero.

By applying additional user defined criteria, a pixel can be automatically rejected by the algorithm. Such optional criteria are:

- if the orientation of the gradient differs more than a certain threshold from the mean orientation of all line points,
- if the distance between the estimated line point and the approximate line is above a certain threshold,
- if the standard deviation of the determined subpixel line point \( u_0 \) is above a certain threshold, or
- if the amplitude of the gradient, corresponding to the line point, is below a certain threshold.

Once the line points of all objects within the image will be detected, the algorithm converts this raster data into vector format by fitting straight lines to the linear boundaries.

\[
f(x) = a_0 + a_1 \cdot x \quad f(y) = a_0 + a_1 \cdot y 
\]

(Eq. 9)

The estimates for the standard deviation of line points are used as weighting functions. The line parameters \( (a_0 \text{ and } a_1) \) are estimated by a least-squares adjustment (Eq. 5, 6).

These vectors will then be used to determine the image coordinates of object vertices \( (x_0, y_0) \) by line intersection.

\[
y_0 = a_1 + b_1 \cdot x_0 \\
y_0 = a_2 + b_2 \cdot x_0 
\]

(Eq. 10)

Here the estimates for the standard deviation of the line parameters serve as weighting functions. If more than two lines intersect in one point, the intersection is calculated as a weighted mean of all possible intersections.

### 3.2.2 Problems to be faced during internal loop

For the image-based feature measurement not only an ideal appearing edge has to be taken into account, often a real world scene disturbed by additional phenomena. Such phenomena can be occlusions, weak or missing contrast, inversion of gradient directions, shadows, radiometric interferences, or reflections (see Fig. 4).

![Figure 4: Problematic cases occurring during internal loop.](image)

Shadow edges can easily be rejected with the information about the approximate lines. This criterion does not work, if the shadow edges run parallel in a short distance to the edge one is looking for. Weak contrast, reflections, and occlusions can be handled by accepting a line pixel only if the parabola fit is significant and if the first parameter is negative. Interferences and occlusions tend to mislead the measurement of the correct edge. This problem can be faced by taking the gradient orientation into account. Line parameters can be determined significantly, if just a part of the entire line can be observed. Inversion of the gradient direction along an object line, typically occurring in real world scenes, where the horizon is imaged or an object feature has different reflection properties, can be handled if the gradient orientation is taken into account. It is notable, that the gradient orientation is a strong criterion in order to determine the correct line pixels (Hönisch, 1992).

### 3.2.3 External loop and orientation loop

The internal loop is performed sequentially in all images. The object coordinates of a point imaged 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 functional model describing the relationship between the derived and measured quantities consists of the well known colinearity equations (Eq. 12). For each pair of image coordinates \((x, y)\) observed on each image the following pair of equations is written:

\[
\begin{align*}
F_x &= (x - x_p) = - c \frac{U}{W} + \Delta x \\
F_y &= (y - y_p) = - c \frac{V}{W} + \Delta y
\end{align*}
\]

(Eq. 12)

with the auxiliaries:

\[
\begin{bmatrix}
U \\
V \\
W
\end{bmatrix} = D(\omega, \varphi, \kappa) \begin{bmatrix}
X - X_0 \\
Y - Y_0 \\
Z - Z_0
\end{bmatrix}
\]

(Eq. 13)

with the image coordinates \((x, y)\), the elements of interior orientation \((x_p, y_p, c)\), functions of additional parameters \((\Delta x, \Delta y)\), the object coordinates \((X, Y, Z)\), a 3 by 3 orthogonal rotation matrix \(D\) with the three rotations \(\omega, \varphi, \kappa\) and the object coordinates of the perspective centre \((X_0, Y_0, Z_0)\). Starting from observed values the overconstrained equation system (Eq. 12) can be solved leading to estimates of the unknown parameters (cf. Eq. 5, 6). In most applications it can be assumed that the exterior and interior orientation parameters of the camera are unknown.

In the next step the derived object coordinates are re-projected into each image and used to restart the 2D feature extraction. This proceeds until the estimated unknowns do not change significantly form on iteration step to the next. The result of this process is then transferred with additional information about precision and reliability to the CAD environment, which allows the final judgement by the human operator.

4. PRACTICAL EXAMPLE

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.

The Otto-Wagner-Pavillon is one of Otto Wagner’s Stadtbahnhof Station buildings (see Fig. 5) on the Karlsplatz in Vienna, built in 1898/1899. The dimensions of the building are 15 x 8 x 10 m³. The object coordinates of 44 non-signalized (but well defined in the majority) control points have been determined geodetically with an accuracy of 2 mm.

4.1. Image acquisition

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.

The JVC GR-S77E camcorder 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” colour sensor (6.4 x 4.8 mm²). 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, which was fixed at its shortest focal length during the image acquisition.

4.2. Object reconstruction

The image acquisition took place in a way that an image sequence from the object was generated by a person walking with the camera around the object and filming all four facades. From this video tape 38 single frames were randomly digitized with a framegrabber. The mean distance between the camera and the object was about 16 m, which results in an image scale of 1:1800.

The measurement of image features, the estimation of exterior orientation parameters and additional camera parameters for self-calibration as well as the determination of the object coordinates was performed with the feature measurement routine as described above.

A total of 1611 object coordinates was determined to describe the entire building. Therefore 2589 image points have been measured semi-automatically by the feature measurement routine. The normal equation system, which has to be solved consists of 5178 observations and 1726 unknowns. The precision of the object coordinates was determined as a part of the bundle adjustment routine. The results indicate a precision in object space of 1.1 cm in X, 1.3 cm in Y, and 0.7 cm in Z, with X- and Y-axis in the plane and Z-axis in height. This corresponds to 6.7 µm in image space.

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.

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 CAD environment (see Fig. 6). The representation of the architectural object can be given either by its object points, lines (wireframes), surfaces or any combinations of these. The representation in a CAD 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.

5. CONCLUSIONS

This paper described an object oriented measurement approach using CAD models for the for the initialization of an 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.

6. REFERENCES

IAPRS = International Archives of Photogrammetry and Remote Sensing

Figure 6: Photogrammetrically generated CAD-model of Otto-Wagner-Pavillon. From left to right: wireframe model, surface model, texture model draped with original image data.