

PERFORMANCE EVALUATION OF A SYSTEM FOR SEMI-AUTOMATIC BUILDING EXTRACTION USING ADAPTABLE PRIMITIVES

F. Rottensteiner^a, M. Schulze^b

^aInstitute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria – fr@ipf.tuwien.ac.at

^bInstitute of Photogrammetry and Remote Sensing,
Dresden University of Technology, Germany – Matthias.Schulze@mailbox.tu-dresden.de

Commission III, WG III/4

KEY WORDS: Building Extraction, Semi-automatic, 3D City Models, Model-driven Matching.

ABSTRACT

The performance of a system for semi-automatic building extraction from digital aerial images is evaluated. Using this system, buildings are reconstructed by a combination of parametric primitives that are approximately positioned in the images by a human operator, whereas their parameters can be determined precisely by automatic procedures. Unlike in similar systems, the parameterisation of the primitives can be changed in the process of reconstruction to adapt them to irregular building shapes, which is made possible by a specific method of representation of the primitives in parameter estimation. A test project using images of different scales and overlap for the reconstruction of a densely built-up area was carried out. Depending on the image scales, the buildings were reconstructed at the best possible levels of detail. The time required for reconstruction and the success rates of automatic fine measurement were assessed. For distinct buildings, the accuracy of automatic fine measurement was determined, showing that if the automated module is successful, it gives very accurate results. The project showed the benefits of primitive adaptations for the reconstruction of irregular buildings if a high level of detail is to be achieved, whereas at lower levels of detail, the homogenisation by using regular building primitives comes to the fore. Up to 87% of the most salient building parts could be determined by automatic fine measurement, the success rate improving with the number of overlapping images. The time required for building reconstruction was found to be influenced heavily by the interpretability of the images.

1 INTRODUCTION

The great demand for 3D city models collides with the enormous costs of data acquisition at an appropriate level of detail. Despite of the enormous research efforts spent and the progress already achieved, automatic systems for building extraction are not yet operational (Gülch, 2000). Semi-automatic methods for building extraction offer a promising and cost-efficient alternative.

Systems for semi-automatic building extraction from aerial images can be characterized by the way the automated tools are integrated into the work flow. A first group of systems is based on interactive (stereo-) measurement of a structured point cloud, from which the building topology is derived automatically. The CyberCity Modeler, a system that has been shown to enable the acquisition of 500 roof structures per day, belongs to this first group (Grün and Wang, 1998). A second group of systems is based on the principles of Constructive Solid Geometry (CSG). Using these systems, buildings are reconstructed by a combination of parametric primitives by Boolean set operators. On the basis of a visual interpretation of the images, the user can select a primitive, determine approximations for its pose and shape parameters interactively, and optionally apply tools for automatic fine measurement. (Gülch et al., 1999) describe that, using their CSG-based system, acquisition times of 20-70 seconds per primitive and a success rate of the automated modules of 88% can be achieved. The system dealt with in this paper belongs to the second group of systems (Rottensteiner, 2001).

An important issue in building extraction is the level of detail (LoD) that can be achieved. (Bartel et al., 1997) distinguish three levels of detail for 3D city models, namely **LoD1** consisting of vertical prisms with flat roofs, **LoD2** containing geometrically simplified building models with approximated roof shapes,

and **LoD3** containing buildings as complex geometric bodies, including facade details. CSG-based systems for semi-automatic building extraction can be used to create models corresponding to **LoD1** or **LoD2**, but different degrees of generalisation are possible. Generalisation can be performed by selecting a simplifying primitive to model the roof shape, possibly merging several buildings or building parts, or by just omitting smaller building parts (Bartel et al., 1997). For visualisation purposes, it is useful to provide building models at different levels of detail. As long as automatic generalisation techniques of polyhedral models are not operational (Thiemann, 2002), these levels of detail have to be generated independently from each other.

The level of detail of CSG-based building extraction techniques is limited by the set of primitives they provide, typically a few standard roof shapes of rectangular footprint such as saddle-back roofs or flat roofs. These regularities help to keep the number of parameters and, thus, the number of user interactions required for their determination, small. This is a considerable advantage if the level of detail to be achieved or the ground resolution of the images are low, because it brings about a certain homogenisation of the resulting building models. However, it is a limiting factor in densely built-up areas characterized by irregular building shapes (Bartel et al., 1997). (Gülch et al., 1999) describe two methods for reconstructing non-rectangular buildings with flat roofs, either by combining irregular vertical prisms or by computing the Boolean intersection of two rectangular flat roof buildings, but these methods do not work with tilted roofs. Providing new primitives of non-rectangular footprints would be one possible solution, but this results in an enormous number of primitives if the regularities actually existing in the data are to be preserved. It would be a better approach to make the parameterization of primitives adaptable in the reconstruction process to increase the flexibility of such a system.

In this paper, we want to present the results of an evaluation of a system for the semi-automatic building extraction from digital aerial images. This system, first described in (Rottensteiner, 2001), provides a CSG interface for the user who can reconstruct the buildings by primitives. Unlike with other CSG-based systems, it is possible to change the parameterisation of the primitives in the reconstruction process due to a unique method for representing building primitives in parameter estimation, so that the system is capable of the reconstruction of buildings of irregular shape, as long as they consist of planar surfaces. An elaborate test of the method was carried out in the town of Eggenburg (Austria) using images of four different scales and different overlaps. Unlike in previous works dealing with the evaluation of systems for semi-automatic building extraction (Grün and Wang, 1998, Gülch et al., 1999), we used these images of different ground resolutions for a reconstruction of the same scene at different levels of detail, showing how different aspects of using primitives for building extraction obtain different importance depending on the quality of the input data or the results to be achieved. It is our goal to assess and compare the time required for reconstruction and the success rates of automatic fine measurement at different image scales. For distinct buildings, the accuracy of automatic fine measurement is determined in an exemplary way.

2 SEMI-AUTOMATIC BUILDING EXTRACTION

2.1 System Overview

Our system for semi-automatic building extraction is embedded in the digital multi-image comparator ORPHEUS (Rottensteiner, 2001). The input data are multiple aerial images and their orientation parameters; optionally, a digital terrain model (DTM) can be used to interpolate the floor heights. The system offers a data base of parametric primitives. By these primitives and by the tools for their modification, a CSG interface is provided, whereas the internal representations of both the primitives and the compound buildings are based on boundary representation (B-rep). Each building is reconstructed primitive by primitive, each primitive being added to the compound building by a Boolean set operator.

Two or more images are displayed on the screen, super-imposed with the wire frames of all the reconstructed buildings, the current building, and the current primitive. To modify the parameters of the current primitive, the user can identify building vertices or points on building edges in the images. After each user interaction, the primitive parameters are estimated using all available observations, and the wire frames are updated. The parameterisation of the current primitive can be changed if the user considers it to be necessary. Automatic fine measurement is performed on user request, and the results can be post-edited. As soon as the user thinks that the wire frame fits to the building well enough and, thus, the primitive has to be accepted, it is added to the current building. At this instance, the B-reps of the current building and the new primitive are combined according to the selected Boolean set operator, the floor height is interpolated from the DTM (optional), and the wire frame of the current building is updated. As soon as the human operator decides that no more primitive is to be added to the current building, the building is accepted and handed over to a topographic information system. We use the *VRaniMLTM* library (Great Hill Corporation, 2002) for applying the Boolean operators to B-reps.

By this work flow, work is saved in several ways. The building primitives are consistent 3D models containing all the topological information that would have to be acquired off-line in a conventional work flow. As geometrical regularities are considered in modelling, the number of user interactions required for

reconstructing a building is reduced. A further reduction of user interactions is achieved by automatic fine measurement. Finally, when the Boolean operators are applied to the primitives, the intersections of the primitives are determined automatically.

2.2 Modelling of the Building Primitives

The core of our method is a unique way of integrating object modelling by B-rep and hybrid adjustment in the way it is realized in the program ORIENT. ORIENT is designed for the simultaneous adjustment of a great number of observation types, comprising observed points from various sensors, direct observations for all types of mapping parameters and, very important in the context of this paper, fictitious observations of points declared to be on mathematically described surfaces in object space. The observations are assumed to be uncorrelated and normally distributed. Gross errors can be detected by robust estimation. For that purpose, iteration is continued after adjustment has been terminated, this time the weights of the observations being modulated in dependence of the normalised residuals of the most recent iteration. In order to make that procedure work, the number of gross errors should not exceed 30% of all observations (Kager, 2000).

2.2.1 “GESTALT” (shape) Observations. GESTALT observations are fictitious observations of points $\mathbf{P} = (X, Y, Z)^T$ being on a surface in object space. These surfaces are modelled in a local “observation co-ordinate system” (u, v, w) centered at an exterior reference point \mathbf{P}_0 and being rotated with respect to the object co-ordinate system by three angles $\Theta = (\omega, \varphi, \kappa)^T$:

$$\mathbf{p} = (u, v, w)^T = \mathbf{R}^T(\Theta) \cdot (\mathbf{P} - \mathbf{P}_0) \quad (1)$$

In equation 1, \mathbf{R} is a rotational matrix depending on Θ , and \mathbf{p} denotes the position of \mathbf{P} in the local co-ordinate system. In order to model \mathbf{P} as being situated on a surface, \mathbf{P} 's distance from the surface can be fictitiously “observed” to be 0. In ORIENT, the greatest component of the difference vector is used as an observation instead of the Euclidean distance. In the local co-ordinate system, the surface is modelled by polynomial functions of two components of \mathbf{p} , which can be done in three ways:

$$\begin{aligned} 0 &= m_u \cdot u + \sum_{j,k} a_{jk} \cdot (m_v \cdot v)^j \cdot (m_w \cdot w)^k \\ 0 &= m_v \cdot v + \sum_{i,k} b_{ik} \cdot (m_u \cdot u)^i \cdot (m_w \cdot w)^k \\ 0 &= m_w \cdot w + \sum_{i,j} c_{ij} \cdot (m_u \cdot u)^i \cdot (m_v \cdot v)^j \end{aligned} \quad (2)$$

In equation 2, a_{jk} , b_{ik} and c_{ij} are GESTALT coefficients describing the shape of the surface. (m_u, m_v, m_w) are mirror coefficients with $m_i = \pm 1, i \in \{u, v, w\}$. Except m_i , all groups of parameters $(\mathbf{P}, \mathbf{P}_0, \Theta, a_{jk}, b_{ik}, c_{ij})$ can be estimated in adjustment. Within an application, an arbitrary subset of all possible coefficients can be selected for modelling a surface (a GESTALT). It is possible to have more than one GESTALT of different coefficients refer to the same observation co-ordinate system, or to let several GESTALTs share the same coefficients, but refer to different observation co-ordinate systems. In this way and by an appropriate selection of m_i , geometrical constraints between surfaces can be modelled. 3D curves can be determined by intersection of two surfaces (Kager, 2000).

2.2.2 Building Primitives and Parameter Estimation. In the reconstruction process, the geometrical regularities of the building primitives should be considered in adjustment. This is achieved by a system of GESTALT observations created when a primitive is initialised. Each face of the B-rep corresponds to a (planar)

GESTALT modelled by one of equations 2. All these GESTALTs are formulated in the same observation co-ordinate system, so that there is only one set of transformation parameters \mathbf{P}_0 and Θ . The shape of the primitive is described by the selected polynomial coefficients a_{jk} , b_{ik} , and c_{ij} . By an appropriate selection of m_i , symmetries can be modelled. For all points belonging to a face of the B-rep and, thus, also for the vertices of the B-rep, one GESTALT observation is introduced. There are at least three such observations per vertex.

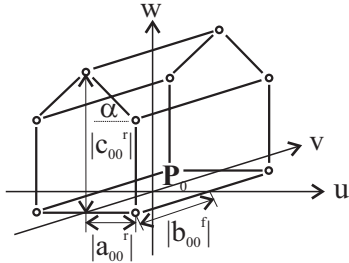


Figure 1: A parametric primitive resembling a saddle back roof.

N	Face	Equation	Sym.	m_i
1	floor	$0 = w$	-	-
2	front wall	$0 = v + b_{00}^f$	-	-
3	rear wall	$0 = -v + b_{00}^r$	2	m_v
4	right wall	$0 = u + a_{00}^r$	-	-
5	left wall	$0 = -u + a_{00}^l$	4	m_u
6	right roof	$0 = w + c_{00}^r + c_{10}^r \cdot u$	-	-
7	left roof	$0 = -w + c_{00}^l - c_{10}^l \cdot u$	6	m_u

Table 1: Parameterisations of the GESTALTs corresponding to the faces of the primitive in Fig. 1. N: Face number. Sym.: Number of the symmetric face. m_i : the mirror coefficient set to -1.

Fig. 1 shows the parameterisation of a simple primitive resembling a saddle back roof. \mathbf{P}_0 is situated in the centre of the floor. The w -axis is also vertical in object space, thus: $w = \varphi = 0$. The parameterisations of the seven faces are presented in Table 1. The parameters to be determined are \mathbf{P}_0 and κ (position and orientation of the primitive in object space), the coefficients a_{00}^r , b_{00}^f , c_{00}^r , and c_{10}^r describing the shape of the primitive, and the object co-ordinates of the vertices. Whereas the vertices can always be determined from their GESTALT observations, the remaining parameters have to be determined from other observations. Although B-rep is used for modelling of the primitive, a minimum number of parameters is required for describing its shape and position. The data base of building primitives can be expanded simply by describing the B-rep of a new primitive and the mapping of its parameters to GESTALT coefficients.

2.2.3 Changing the Parameterisation of a Primitive. Fig. 2 shows a typical situation in the reconstruction of a building in a historic town centre. The primitive resembling a saddle back roof does not fit to the actual shape of the roof exactly: even though the two roof planes are symmetric and the eaves are parallel, the front and rear walls are not orthogonal to their neighbouring walls.

In order to overcome such problems, our system offers two possibilities for the user to change the parameterisation of the primitive by changing the parameterisations of its faces. First, the symmetry between two faces can be discarded, so that each of the originally symmetric faces gets an individual set of surface parameters. Second, linear terms can be added to the parameterisation of a face, in order to discard specific local geometric constraints. With roofs, two linear coefficients (c_{10} and c_{01})

are added to the mathematical model if they are not already contained. As walls are modelled to be vertical, only one linear coefficient can be added to the mathematical model (a_{10} or b_{10} , depending on which one of equations 2 is used). In the example shown in Fig. 2, three such modifications are necessary. The symmetry of the two faces not being parallel to the street has to be discarded, and linear terms have to be added to the mathematical models of both these faces because the building has no rectangles (Table 2). From a practical point of view, the user has to select a “current edge” in one of the displayed images using the right mouse button, and he or she has to trigger off the desired operation by pressing a button. If the edge is the intersection of a roof and a wall, the selected operation is applied to the wall only; if it is the intersection of two roof faces, it is applied to both faces.

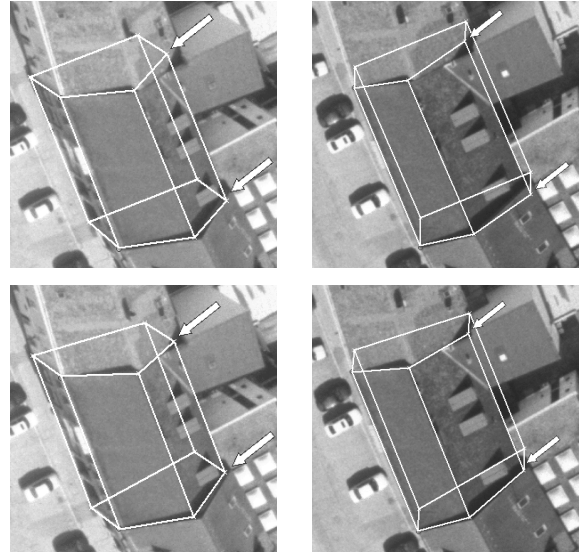


Figure 2: Two images with a super-imposed building primitive. Upper row: the primitive does not fit very well. Second row: After changing the parameterisations of the front and rear walls.

N	Face	Equation	Sym.	m_i
2	front wall	$0 = v + b_{00}^f + b_{10}^f \cdot u$	-	-
3	rear wall	$0 = v + b_{00}^r + b_{10}^r \cdot u$	-	-

Table 2: New parameterisations of walls 2 and 3 from Table 1.

In the second row of Fig. 2, the effects of changing the parameterisation of the primitive can be seen: the modified primitive fits to the data now. The angles between the “upper” and the “lower” walls and their neighbours differ from 100° by 8.2° and 4.1° , respectively. The lengths of the two eaves differ by 2.62 m or 15%, effects that are clearly visible and determinable in the images. We prefer adapting the parameterisation of primitives to providing new primitives because the number of primitives in the data base can be kept manageable. We consider a primitive to represent the basic topology of a building part, whereas its shape is adaptable within a certain scope.

2.3 Determination of the Parameters of the Primitives

When a primitive is selected by the user, a system of GESTALT observations is created, with a certain parameterisation of the involved faces and containing GESTALT observations for the building vertices. The primitive parameters have to be estimated from other observations, for instance from the image co-ordinates of the vertices. We can also use image co-ordinates of points on the edges of the primitives, in which case the measured point is not actually a part of the B-rep. Thus, we get its three object

co-ordinates as new unknowns, but we also get four observations (two image co-ordinates, one GESTALT observation for each of the planes intersecting at the edge the point is assigned to), thus one redundant observation for determining the primitive parameters. Finally, we can use direct observations of the co-ordinates of the building vertices if the user chooses to snap a vertex of an existing building part. These observations can be created by interactive measurement or by automatic fine measurement. As long as they are not available, direct observations are created for all the parameters from default values to prevent singularities.

2.3.1 Interactive Measurement. In interactive measurement, each mouse click has a certain meaning assigned to it by the selection of a *measurement mode*. The right mouse button can be used to select the “current element” (vertex or edge) of the primitive, whereas the left mouse button is used to create measurements. In the *snap vertex* mode, snapping a vertex of an existing building yields the vertices’ object co-ordinates to be used as direct observations for the current vertex. In the *measure vertex* mode, the mouse position gives the image co-ordinates of the current vertex, and in the *measure edge point* mode, the user can measure a point situated on the current edge. Each user interaction is followed by an adjustment of all available observations, robust estimation being used to find out which parameters can be determined by the new observations (Rottensteiner, 2001).

New observations are created until the wire frame fits to the images. For a precise determination of a saddle back roof, it is typically sufficient to measure five image points (three vertices on the eaves and a gable point in one image, one of these vertices in a second image). For providing approximate values, four points are sufficient (the gable point is not required). The image co-ordinates can be measured only coarsely, just to provide approximations for automatic fine measurement, or they can be measured precisely if automatic fine measurement is supposed to be problematic. This method is independent of the order in which the building vertices are identified, and there are no restrictions with respect to the number of identified vertices.

2.3.2 Automatic Fine Measurement. In automatic fine measurement, observations for points on roof edges are generated automatically by matching these roof edges with image edges in a model driven process: first, correspondence hypotheses are generated, and second, these hypotheses are evaluated with respect to their consistency with the model of the primitive to eliminate false matches. Starting from the approximations provided by the user, this matching technique is applied iteratively to several levels of the image pyramids, the results of matching at a certain level being used as improved approximations for the next lower level, until the lowest pyramid level is reached.

The generation of correspondence hypotheses is performed independently for all roof edges of the building primitive in all images. In the images, image edges are extracted and approximated by polygons. A segment of such a polygon is supposed to be a correspondence hypotheses for the object edge if at least one of its end points is close to the projection of the approximated object edge and if it is almost parallel to that projection. Each end point of a segment thus found to be a matching candidate for an object edge is supposed to be a point measured on the object edge. We choose the weights of the image co-ordinates of a segment end point to be proportional to the segment length, so that long segments obtain more influence than short ones.

The correspondence hypotheses are evaluated by simultaneous adjustment of all observations for all roof edges, the primitive being represented by the GESTALT observations. False hypotheses

are eliminated by robust estimation in a highly redundant adjustment procedure. Due to the hierarchical strategy, the search region can be kept small at each pyramid level, which considerably reduces the a priori number of false matches. The most important parameter of the matching process is the a priori standard deviation σ_g of a GESTALT observation, describing the “rigidity” of the object model in adjustment and, thus, influencing the number of hypotheses detected to be outliers. Automatic fine measurement is expected to deliver more accurate results than conventional photogrammetry because by using image edges, a greater number of observations is considered in adjustment.

3 EVALUATION OF THE SYSTEM

In order to evaluate our method for semi-automatic building extraction, a test project was carried out in the town of Eggenburg (Lower Austria). Four different aerial blocks varying in image scale (1 : 30000 to 1 : 4500) and overlap (two-fold and six-fold) were available in the test area (Table 3). A DTM was generated from the images of block 4 (1 : 4500). The images were in black and white, scanned at 15 μm , except of block 4 with colour images scanned at 30 μm . Block 3 (1 : 6000) featured the best ground resolution. All the buildings at the main square of the town were reconstructed using our method (Fig. 3). Table 3 gives an overview on the results achieved at the different resolutions.

3.1 Levels of Detail in Dependency on the Image Scale

Depending on the image resolutions, the buildings were reconstructed at different levels of detail (Fig. 4), all of them, however, corresponding to **LoD2** according to (Bartel et al., 1997). The parameterisation of a primitive was changed when deviations of the actual building from the shape of the primitive were found to be observable by visual inspection. As described in section 2.2.3, two mouse clicks are required for each modification.

The images of block 1 (1 : 30000) had a rather bad radiometric quality. As the buildings were only a few pixels long in the images, only the most salient building structures could be reconstructed, which resulted in strongly generalized models representing the lower margin of **LoD2**. As a consequence of the image quality, building groups were merged and reconstructed by just one primitive, which causes their loss of individuality. Smaller details were omitted. Thus, altogether 23 building primitives were sufficient to reconstruct the scene. It was not necessary to adapt the primitives because deviations from the rectangular footprints could not be observed. The homogenisation achieved by using regular primitives was an important help in the reconstruction process at this image scale.

With block 2 (1 : 10000), it was possible to separate individual buildings. Generalisation was only performed by omitting building details and by using simplified primitives. Thus, just the basic shape of the buildings was reconstructed, which was possible by using 63 primitives, out of which only three corresponded to larger dormers. 10 primitives or 16% were adapted, whereas for the rest, no deviations from the rectangular footprint could be observed. The decision not to reconstruct medium-sized dormers at that image scale was taken because of problems of interpretation and because of problems with the geometric precision in the reconstruction of such small details.

Using the images of block 3 (1 : 6000), it was possible not only to reconstruct the basic building shapes, but also to model smaller dormers. The model, consisting of 65 primitives, gives a very detailed impression of the scene. In the reconstruction process, we distinguished “basic primitives” modelling the basic building

Block	Scale 1 :	f [mm]	Img.	col.	Res _i [μ m]	Res _g [cm]	P	P _b	P _m	Auto [%]	Time [min]	s _{XY} [cm]	s _Z [cm]
1	30000	153	2	N	15	45	23	23	0	0	45	-	-
2	10000	210	2	N	15	15	63	60	10	53	110	12.8	31.3
3	6000	210	2	N	15	9	65	58	19	67	110	8.9	31.2
4	4500	152	6	Y	30	13.5	72	60	19	87	100	7.9	9.0

Table 3: Overview on the test data and on the results. f: focal length. Img.: Number of images covering the test area. col.: Colour, Res_i: Scanning resolution of the images. Res_g: Ground resolution. P: Total number of primitives. P_b: Number of “basic primitives” (main building parts). P_m: Number of primitives where the parameterisation was modified. Auto: Percentage of basic primitives measured automatically. Time: Time used for the acquisition of the whole scene. s_{XY}, s_Z: r.m.s. errors of differences of planimetric co-ordinates and height, respectively, compared to stereo-measurement.

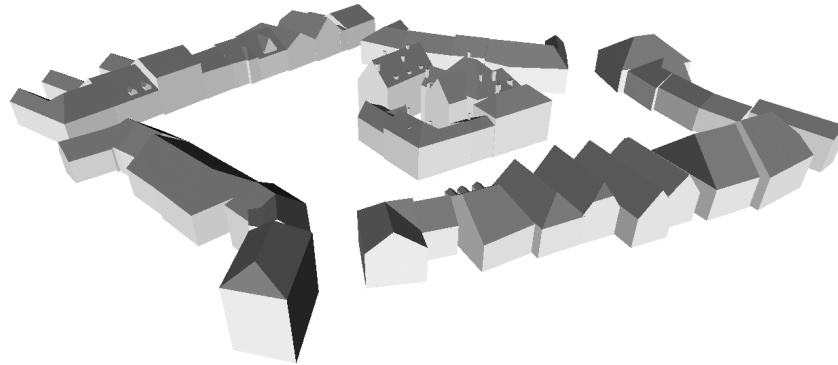


Figure 3: VRML-visualisation of the main square of Eggenburg (1 : 4500).

shapes, and primitives used for modelling smaller details. With the latter ones, the geometric precision of the reconstructed parameters was considered to be less important because the quality of the resulting model was increased by the sheer presence of these details. 19 primitives or 29% were adapted, as the model deviations were clearly visible at that scale, which shows the importance of that feature.

Using the images of block 4 (1 : 4500), the best level of detail was achieved due to the benefits of colour for visual interpretation and because the images were less noisy than those scanned at 15 μ m, which partially compensated for the lower ground resolution. In the central group of buildings, even very small roof details such as chimneys were reconstructed. The resulting model, consisting of 72 primitives, corresponds to the upper margin of **LoD2**: it contains very detailed roof models, but no roof overhangs and no details in the facades. Again, 19 primitives were adapted.

3.2 Performance of Automatic Fine Measurement

The performance of automatic fine measurement was not evaluated for block 1 because the roof edges are too short in the images. For the other blocks, automatic fine measurement was carried out using two levels of the image pyramids. In (Rottensteiner, 2001), automatic fine measurement was shown to have a pull-in-range of up to half the building dimensions for detached buildings. In the current test project, there are many occlusions, so that approximate values of a better quality (about 5 pixels) were required. We considered the results of automatic fine measurement to be a successful if, by visual inspection, the resulting wire frame corresponded to what would have been the expected result of interactive measurement. Note that automatic fine measurement was only applied to the “basic primitives”, because with smaller details, either the lengths of the image edges would have been too small, or large parts of the respective primitives were situated inside other building parts, so that again too few image edges would have been available for matching. The success rates thus obtained are presented in Table 3.

As a general trend, the success rate is increased with the ground resolution. The relatively low success rate of 53% for block 2 is explained by problems of edge extraction: some relevant image edges are missing at that resolution. In block 3, 67% of the basic primitives could be measured automatically. It is very interesting that in block 4, having a smaller ground resolution than block 3, a much higher success rate of 87% could be achieved. This is mainly caused by the fact that block 4 has a six-fold overlap. The success of automatic fine measurement depends on image edges corresponding to roof edges to be found in the images. In densely built-up areas, there are many occlusions, and with only two-fold overlap, many object edges are likely not to be visible in one of the images. This is a clear indication that it makes sense to use multiple overlap rather than two-fold overlap for matching.

There are several reasons for failure. We have already pointed out the problem of missing image edges. Parallel image edges in close distance to the “correct” match also cause problems. Failure was more likely to occur with primitives of modified parameterisation: obviously, the implicit geometric regularities of the primitives helped to correctly separate good matches from outliers. There is a trade-off between “good” conditions for matching and the level of detail achieved for the resulting building models.

For two buildings, the accuracy of automatic fine measurement was evaluated. The vertices of these buildings were measured stereoscopically with an accuracy of ± 8.6 cm in planimetry and ± 10.4 cm in height using the images of block 4. The r.m.s. errors of the co-ordinate differences of 12 vertices from stereo measurement and automatic fine measurement were computed (Table 3). The results show that automatic fine measurement gives very accurate results if it is successful. The r.m.s. errors of the co-ordinate differences in planimetry range from ± 8 cm to ± 12 cm, improving slightly with the image scales. The advantages of using more than two images are clearly visible in the r.m.s. error of the height differences: whereas it is about ± 31 cm for the blocks having a two-fold overlap, it comes down to ± 9 cm if six images are used for matching.

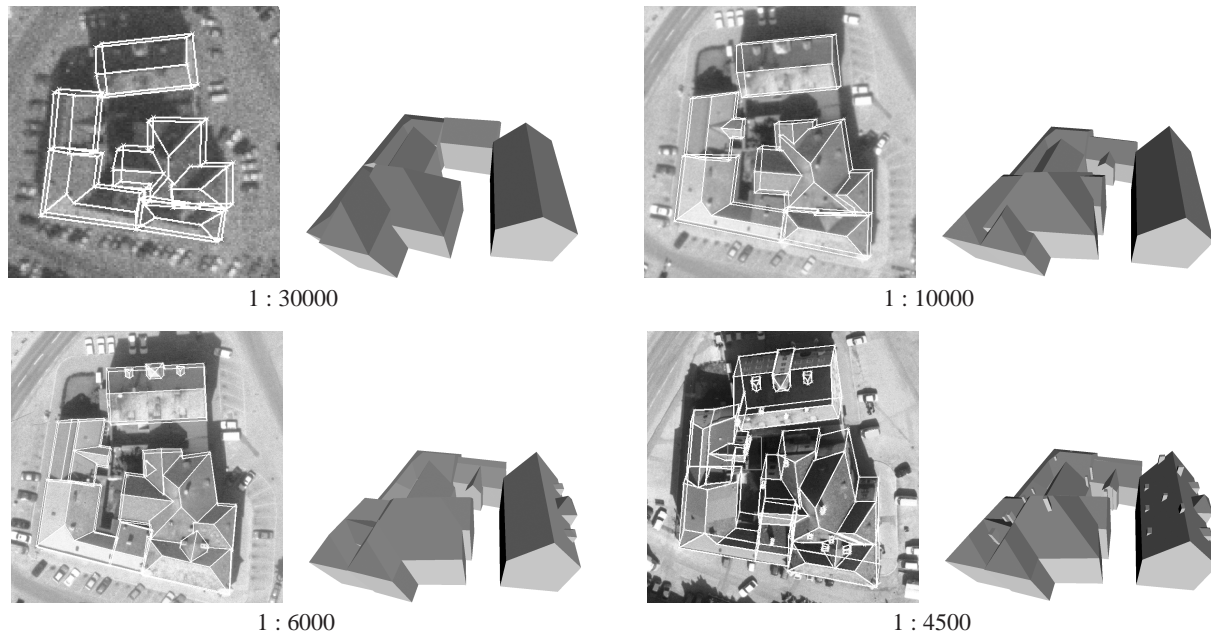


Figure 4: A group of buildings in the centre of the square in all four image scales. Left: The reconstruction results are super-imposed to the digital images. Right: VRML-visualisations.

3.3 Acquisition Times

Table 3 also shows the acquisition times for the four blocks. It is obvious that the time required for acquisition is not increased directly with the number of primitives. This is due to the fact that the most time-consuming task to be performed is the interpretation of the images. As interpretation is especially difficult with the images of block 1, 45 minutes are required for the reconstruction of only 23 primitives, compared to 110 minutes for the about 65 primitives of both blocks 2 and 3, where the reconstructed models have a similar complexity. It is very interesting that the time for reconstructing the most detailed model (block 4) is shorter than the time required for blocks 2 and 3. Obviously, the colour information and the possibility to use additional images were very helpful for scene interpretation. The additional primitives used for modelling small roof details are often quite similar. For instance, with buildings containing more than one dormer, these dormers often have the same shape, so that they can be positioned by very few user interactions because the parameters of a primitive are always initialised by the values of the previous primitive of the same type that was reconstructed. 100 minutes required for the reconstruction of 72 primitives correspond to 83 seconds per primitive, which is slightly slower than the 70 seconds achieved by (Gülch et al., 1999) for complex roof structures, but includes the time required for primitive adaptation. The resulting models are supposed to have a higher quality.

4 CONCLUSION

We have presented a system for semi-automatic building extraction based on the CSG principle, implemented as an experimental module of the digital photogrammetric software ORPHEUS. Unlike in comparable systems, it is possible to adapt the parameterisation of the primitives used for reconstruction, which could be shown to be important for improving the level of detail of the reconstructed building models. The test results have shown that the method can be applied to images of different scales and configurations. At smaller image scales, homogenisation by regular building primitives is important to derive strongly generalised

building models. With respect to both the accuracy, the level of detail achieved, and the time required for reconstruction, our test suggests to use colour images, for the sake of a better interpretability, and multiple (≥ 4) overlap with images from two strips to have good conditions for matching. Under these conditions, considered to be optimal for building extraction, 83 seconds are required for the reconstruction of a primitive in a densely-built up environment. The automated module works in 87% of the cases, delivering results with an accuracy comparable or even better than the one achieved in conventional stereo plotting.

REFERENCES

- Bartel, S., Bill, R., Boytscheff, C. and Königer, A., 1997. Datenfusion zur Erstellung realitätsnaher 3D-Geo-Informationssysteme für städtebauliche Planungen. ZPF 1997(4), pp. 139–137.
- Great Hill Corporation, 2002. <http://www.greathill.com>.
- Grün, A. and Wang, X., 1998. CC-Modeler: A Topology Generator for 3-D City Models. ISPRS Journal of Photogrammetry and Remote Sensing 53(5), pp. 286–295.
- Gülch, E., 2000. Digital systems for automated cartographic feature extraction. IAPRS, Vol. XXXIII-B2, pp. 241–256.
- Gülch, E., Müller, H. and Läbe, T., 1999. Integration of Automatic Processes into Semi-automatic Building Extraction. IAPRS, Vol. XXXII/3-2W5, pp. 177–186.
- Kager, H., 2000. Adjustment of Algebraic Surfaces by Least Squared Distances. IAPRS, Vol. XXXIII-B3, pp. 472–479.
- Rottensteiner, F., 2001. Semi-automatic extraction of buildings based on hybrid adjustment using 3D surface models and management of building data in a TIS. PhD thesis, Institute of Photogrammetry and Remote Sensing, Vienna University of Technology. Vol. 56 of Geowissenschaftliche Mitteilungen der TU Wien.
- Thiemann, F., 2002. Generalization of 3D Building Data. IAPRSIS, Vol. XXXIV/4, pp. 286–292.