BUILDING RECONSTRUCTION FROM LASER SCANNING AND IMAGES

Claus Brenner
Institute of Cartography and Geoinformatics, University of Hannover, Germany
Claus.Brenner@ikg.uni-hannover.de

ABSTRACT

The automatic extraction of objects from laser scans and images has been a topic of research for decades. Nowadays, with new services expected, especially in the area of navigation systems, location based services, and augmented reality, the need for automated, efficient extraction systems becomes more urgent than ever. After giving an overview, this paper reviews some of the automatic and semi-automatic reconstruction systems that have been proposed. Then, it gives some conclusions that can be drawn from those approaches.

1 INTRODUCTION

Nowadays, topics like "location based services", "augmented reality", and "personal navigation" are not only actively discussed in the scientific community but are also areas where applications are expected to enter the market soon. Even though technical aspects like device or network characteristics often dominate the discussion in the public, it has in the meantime become clear that the quality and usefulness of services is the major key to success. For services tied closely to spatial information, the accuracy, detail, up-to-dateness and coverage of the underlying databases is of major relevance.

However, it is evident that the effort for acquisition and update of spatial data is very high. With expectations from users rising, the situation becomes worse. It seems today that plans to extend databases into the third dimension, for example, three-dimensional city models, cannot be realized economically. The reason for this is that three-dimensional information is not only more difficult (i.e., expensive) to acquire, but also its change rate is often higher than that of two-dimensional data.

One can identify major shortcomings in today's data acquisition practice:

The *degree of automation* in today's acquisition systems is too low and certainly lower than it could be if results from research had been incorporated more consequently into production systems.

For example, one major supplier of city models in Germany has acquired – according to his own estimates – approximately 30.000 square kilometers of German cities. All this has been done by digitizing each point manually – several points per building – using stereo photogrammetry. This is not only a huge acquisition effort in the first place, it is also estimated by the supplier that an update of the database will require about 70 percent of the initial acquisition cost.

There is a lack of automated systems which *combine* geoinformation from different sources.

Taking digital street maps as an example, the production of consistent databases has so far been relatively easy since every aspect is under control of the corresponding map producer. This will, however, be not possible anymore in the future when expectations towards navigation systems rise and several data sources have to be combined in order to obtain the final map product. Three-dimensional navigation systems will make it necessary to *combine current two-dimensional street networks with digital terrain models and three-dimensional city models*. It is unlikely that single map producers are able to acquire and update all these data sets. Therefore, highly automated procedures will be necessary to solve the problem efficiently.

Thus, it becomes clear that automation for initial acquisition, automation for update, and combination of different data sets and data sources are actual problems which are closely tied together. Progress on these topics will be crucial for extending and maintaining detailed and area covering databases in the future.

2 AERIAL PHOTOGRAMMETRY AND LASER SCANNING

Aerial photogrammetry has been and still is one of the preferred ways to obtain three-dimensional information of the earth's surface. Being very well understood an delivering accurate results, the major drawback is that *automation* of the measurement process is closely related to image understanding – being a problem hard to solve. Progress has been made since the late 1980's especially in the field of digital image matching for digital terrain model (DTM) generation (Ackermann and Krzystek, 1991). However, these methods are often not well suited to derive the surface in cities where usually many sharp jump edges are present, although progress has been made in this area (Paparoditis et al., 2001). They also strongly depend on image quality and sufficiently strong contrast between adjacent image regions.

During the 1990's, a new method for obtaining surface models became available: *airborne laser scanning*. Subsequently, the scanning systems were improved and direct georeferencing became feasible with sufficient accuracy. Today, airborne laser scanning is a mature technology with a multitude of companies offering systems and services (Baltsavias, 1999). Scanning of very large areas is possible, for example the entire Netherlands have been and

Germany's state of Baden-Wrttemberg is in the progress of being scanned, each with an area of over 30.000 km^2.

When the first laser scan datasets became available, it was anticipated that aerial photogrammetry will become obsolete to some extend. Indeed, airborne laser scanning proved to be very effective for the derivation of digital surface models. However, it became clear in the meantime that the two technologies are complementary:

- Digital aerial photogrammetry allows the accurate measurement of single, prominent points and structures – usually defined by a human operator. Its greatest problem remains the still very low degree of automation despite huge research efforts in the past.
- Laser scanning, on the other hand, is able to provide dense clouds of directly measured three-dimensional points. The point density and degree of preservation of jump edges makes the integration of automated processes – such as range data segmentation – relatively easy. The main drawback of airborne laser scanning is that the laser beam just samples the earth's surface in some fixed pattern; it is not capable of pointing to particular objects directly. Thus, its lateral measurement accuracy is not very high.

Therefore, it is most promising to *combine* aerial photogrammetry and laser scanning to obtain highly automated, yet very accurate results. As Ackermann has put it:

> "The systematic combination of digital laser and image data will constitute an effective fusion with photogrammetry, from a methodological and technological point of view. It would resolve the present state of competition on a higher level of integration and mutual completion, resulting in highly versatile systems and extended application potential. [...] It would be a complete revolution in photogrammetry if image data could directly be combined with spatial position data." (Ackermann, 1999)

It is for this reason that laser scanning companies are integrating digital cameras with their systems, while vice versa aerial photogrammetry system suppliers are starting to include laser scanners into their product lines.

However, despite the progress that has been made with scanning systems and digital image acquisition, the *automated processing* of the resulting datasets is at a very early research stage. For example, nowadays laser scan data is mostly used to produce digital terrain models which can be obtained from the original (measured) point cloud by interpolation algorithms – which are not much different from photogrammetric DTM modules that have been in use since about three decades. Only in specialized areas, such as the derivation of DTM's in wooded areas or the surveying of power lines, there are first approaches to exploit specific properties of laser scan datasets for automated extraction.

3 EXTRACTION OF MAN-MADE OBJECTS

Efficient extraction of man-made structures has been a topic of intense research for many years (Grün et al., 1995, Grün et al., 1997, Baltsavias et al., 2001). The great interest of the scientific community was driven by the obvious need to automate or to facilitate manual processes for capturing data efficiently. The extraction of man-made objects is an *object recognition problem.* As such, it is part of an extremely wide research field (e.g. (Grimson, 1990, Jain and Flynn, 1993, Faugeras, 1994)) – an extensive discussion of which would be much beyond the scope of this paper. However, one can identify some basic principles which are present in most object extraction systems:

- The presence of *object models* which can be *generic* or *specific*. In the context of man-made object reconstruction, the use of specific models is usually not possible due to the great variety of objects in the real world. Simple generic objects are *parametric* descriptions where the general form is fixed but geometric parameters such as position, height, width, depth, and angle can be adjusted. On the other end of the spectrum, models based on the *Gestalttheorie* can be considered as complex generic models where properties like neighborhood, closedness, continuity and symmetry are used to recognize structures in scenes (Lowe, 1985). Such general models have also been used in the context of building extraction from images (Lin et al., 1995, Collins et al., 1995). One approach used by many researchers is that object models are build from object primitives by a given set of aggregation rules.
- The *detection* and *recognition* of one or more objects present in a scene. This is the core step of object recognition, which of course assumes the availability of appropriate object models. Different control paradigms can be identified, such as bottom-up (data-driven), top-down (model-driven) and mixed approaches such as hypothesize-and-test. A key aspect is also how the search is organized, in particular how the usually huge search space is reduced by techniques such as (discrete) relaxation or constrained tree search (Grimson, 1990).
- Measurement of *geometric information* about the position, orientation and size of the recognized objects. This step is not generally required in object recognition, however it is naturally present in object extraction for geoinformation systems. The geometry of objects can be described e.g. by a boundary representation, constructive solid geometry (CSG) or spatial enumeration (i.e., voxels).

There have been quite a number of research systems which were proposed for the extraction of man-made objects. They can be classified according to the data sources they use, the underlying object model, and the kind of intended operation: semiautomatic or fully automatic.

4 SOME FULLY AUTOMATIC RECONSTRUCTION SYSTEMS

In the following, some fully automatic building extraction systems are discussed briefly. It has to be stressed that this is only a selection and the reader is referred to e.g. the Ascona proceedings (Grün et al., 1995, Grün et al., 1997, Baltsavias et al., 2001) which provide an excellent overview.

4.1 Haala, 1996

In his Ph.D. thesis (Haala, 1996), a method for the automatic extraction of buildings from stereo imagery is proposed (see fig. 1). It uses the following steps. From a stereo image pair (1) a digital surface model (DSM) is obtained using conventional image matching (2). Using a morphological approach, regions of interest (ROI) are extracted (3), which mark the possible presence of buildings. The DSM is also used to generate a disparity map between the left and right image of the stereo pair (4). From each image, line segments are extracted separately (5a,b), whereas the search is constrained to the ROI's found previously. Using the disparity map, two-dimensional line segments of both images are grouped to three-dimensional segments (6), and a first filtering can be done. The three-dimensional segments are then grouped in object space, yielding rectangles (7), which allows for a second filtering of the primitives. Using the arrangement of rectangles in object space, approximate parameters for a building primitive of saddleback type can be obtained. The parameters are then improved using an estimation which minimizes the distance between the three-dimensional line segments and the line segments of a saddleback wireframe model. Finally, estimation error and gray value variance in the images is used to select one model from the set of reconstruction results (8).

Characteristic to this approach is the step-by-step filtering of primitives which eventually leads to the selection of a single reconstruction result. This goal is reached by (a) using three-dimensional information in form of a DSM and disparity map, (b) aggregation to higher level primitives which do not appear as often, e.g. grouping line segments to longer lines, to parallel lines, or grouping lines to rectangles, (c) the transition from (2D-) image space to (3D-) object space, allowing to use meaningful thresholds for angles, surface areas etc., and (d) using parameter estimation to find a solution, and the residual to evaluate it. The limitations of this early approach are the dependency on a DSM of good quality, its usefulness only in suburban areas, and the concentration on a single (saddleback) primitive type.

4.2 Henricsson and Baltsavias, 1997

The approach of (Henricsson and Baltsavias, 1997) (termed "ARUBA") also uses aerial images for the extraction of buildings (see fig. 2). The reconstruction starts with a region given by an operator (1). Then – similarly to the approach of (Haala, 1996) – line segments are extracted (2), aggregated to three-dimensional segments (3), and used to hypothesize planes (4). A similarity grouping (5) (Henricsson, 1996) is used to find the most evident and consistent planes. Finally, building walls are generated using a DTM.

A speciality of this approach is the use of color information. Along the extracted contours, color attributes are computed for the left and right flanking regions. This information is used during the stereo matching, planar grouping, and similarity grouping. ARUBA is not capable of handling scenes in densely built-up areas. The authors discuss two possibilities for the extension of the system, namely the use of stronger object models and the stronger integration of an operator into the reconstruction process.

4.3 Fischer et al., 1998

This work (Fischer et al., 1998) was the result of a larger research project (c.f. (Braun et al., 1995)). It is based on a hierarchical concept for modelling. On each layer, four representations exist: the three-dimensional object model and its instantiation, as well as the two-dimensional projection of the object model together with its instantiation. Therefore, on each layer a tight coupling between two- and three-dimensional structures is obtained. This formalizes the idea of an early transition to three dimensions, as realized in the previously discussed approaches of (Haala, 1996) and (Henricsson and Baltsavias, 1997). The coupling on each layer does not only involve a projection of the geometry, but also propagates constraints and quality measures. For example, parallelism of structures in three dimensions is propagated (assuming a weakly perspective mapping) to a 'tolerance-based' parallelism in the two-dimensional image.

The following layers were given as hierarchy levels (Braun et al., 1995): scene, object, object part, aggregated feature, feature, voxel. Accordingly, for image space: image, aspect, aspect part, aggregated feature, feature, pixel. This hierarchy was partly realized for building corners (see fig. 3). In this case, basic features are points, lines, and regions. They can be extracted by a polymorphous segmentation from the images (Fuchs, 1998). Corners are aggregated features (one level up in the hierarchy), consisting of at least three intersecting planes, a corresponding number of line segments and a single point. Corners can be obtained on the basis of a polymorphous segmentation (Lang, 1999). Still one level up, an aggregation of corners into object parts takes place, in this case into simple building primitives called 'connectors' and 'terminals'. The aggregation of those object parts finally leads to objects (which themselves are part of the scene, which is the highest level).

4.4 Baillard and Zisserman, 1999

This approach (Baillard and Zisserman, 1999) reconstructs a polyhedral model based on the edges between planar roof patches. The main idea is to obtain the half planes to the left and right of a detected dihedral line segment. The advantage compared to other approaches is that with using

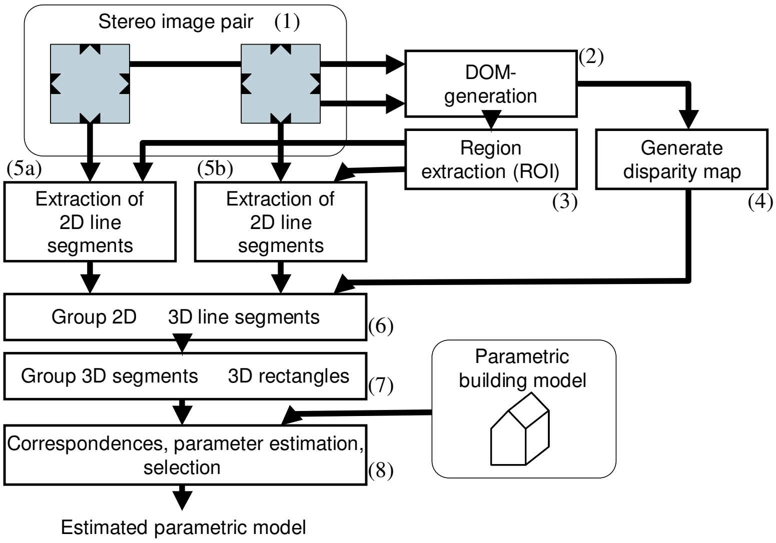

Figure 1: Building reconstruction according to (Haala, 1996).

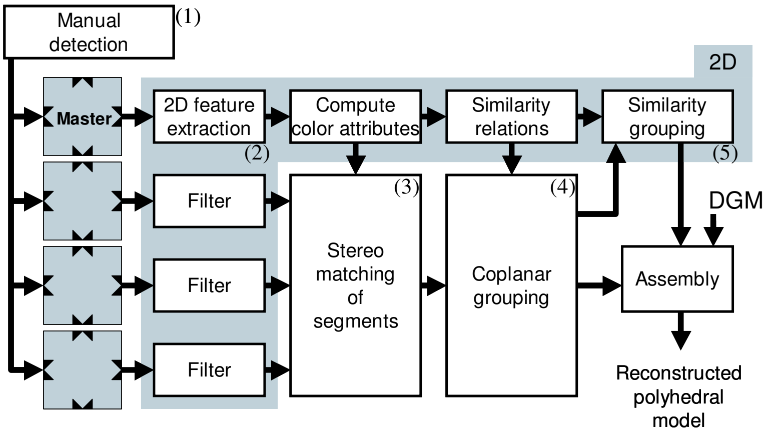

Figure 2: Illustration for 'ARUBA'. Two-dimensional parts in gray. Figure according to (Henricsson and Baltsavias, 1997)).

the dihedral line and left/ right image patches in its vicinity, only relatively local information is exploited. Six overlapping aerial images are used.

The algorithm (c.f. fig. 4) first extracts line segments in each of the images. Using a line grouping (Baillard et al., 1999) based on the geometry and radiometry of the six images, three dimensional line segments are generated. Then, the half planes to the left and right of each line segment are computed. Since the half planes are constrained by the three dimensional line segment through which they pass, only their slope α has to be determined. This is done by a search over all possible angles $-\pi/2 < \alpha < \pi/2$, computing a heuristic correspondence measure. The measure is based on a weighted correlation between the images, where the geometric transformation is given by the plane equation and the exterior orientations of the images. It is summed over all images. Apart from some additional requirements, the angle with the largest score is taken.

Following this, line segments and half planes are grouped on the basis of collinearity and coplanarity, which reduces the total number of planar patches. Additional lines are introduced by intersecting existing planar patches and the extent of each patch is obtained from a heuristic grouping. Each surface is then verified using a similarity measure computed over all images.

4.5 Brenner, 1998

In contrast to the previously mentioned systems, this approach uses DSM's (usually from laser scanning) and 2D groundplans as data sources for an automatic and/or semi-automatic reconstruction process (Brenner and Haala, 1998). Aerial images can be used to facilitate the interpretation by a human operator when semi-automatic post-processing is performed, however they currently do not contribute to the measurement.

Figure 5 sketches the workflow of the reconstruction algorithm. Input data is on the left, output on the right and the flash icon marks the places where automatically derived data can be modified or amended. Processing starts by decomposing the ground plan polygon into two-dimensional primitives (up to now, rectangles) automatically. Each two-dimensional primitive is the footprint of a corresponding three-dimensional primitive. The location, orientation, and size of the two-dimensional primitive applies as well for

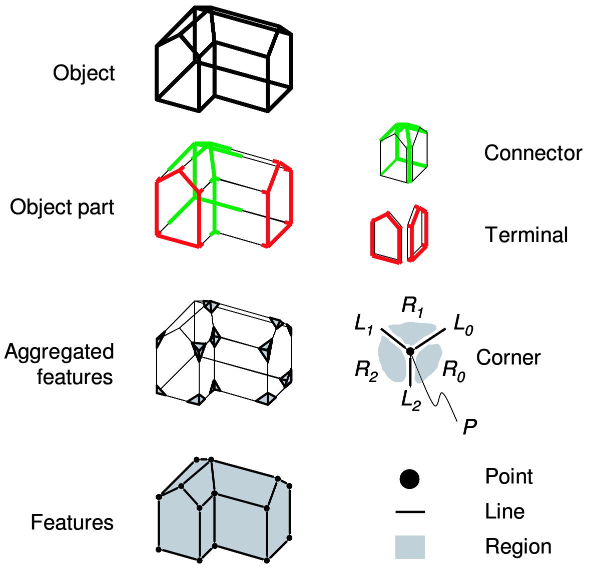

Figure 3: Hierarchical modelling concept, according to (Fischer et al., 1998).

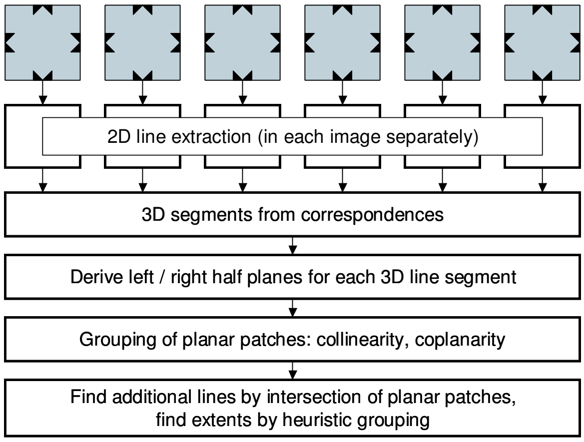

Figure 4: Reconstruction method according to (Baillard and Zisserman, 1999).

the three-dimensional primitive. What remains to be determined are the parameters of the roof, namely roof type (currently one of flat-, saddleback-, hip- and desk roof), height of the building and roof slope. Based on a segmentation, certain roof types are excluded beforehand. For example, a flat roof primitive will not be regarded as a possible reconstruction if the absolute area as well as percentage of flat pixels is below certain thresholds. A least squares estimation then computes the best fit of the primitives to the given DSM. When several models are suitable, the one with the smallest residual is selected.

After this step, the individually reconstructed primitives are overlapping 3D solids. They can be output in the form of either a list of solid descriptions or a list of planar faces. Most often it is desirable to find a building description without overlapping parts. As this is a standard CSG problem, a CAD kernel is used to perform the necessary merging (Boolean union) operations. Finally, a non-overlapping building description is obtained, which can be exported and converted into different CAD formats.

The advantage of this approach is that by using existing ground plans, high level (symbolic) information is “injected” into the reconstruction process, which makes it more reliable. Also, since the two-dimensional information is used as a starting point, links between the original database and the final (augmented) three-dimensional database are established easily. On the other hand, the method does not reconstruct roof parts for which there is no evidence in the ground plan, and it does not perform well if the ground plan is quite different from the actual roof outline or the ground plan cannot be subdivided easily into primitives.

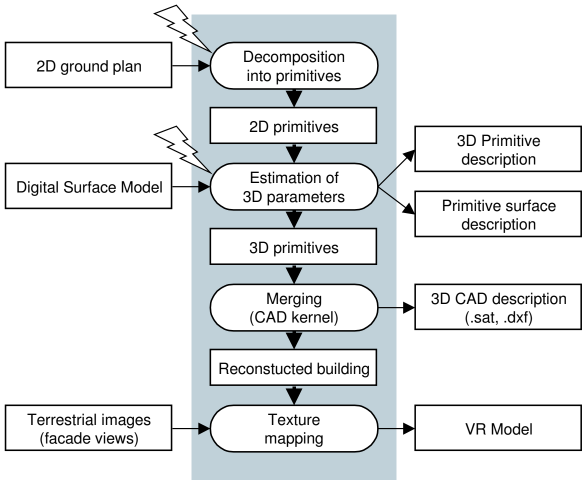

Figure 5: Building reconstruction according to (Brenner and Haala, 1998). Flash icons mark the steps which are used for semi-automatic post-processing.

5 SOME SEMI-AUTOMATIC RECONSTRUCTION SYSTEMS

5.1 “CyberCity Modeler”

Following the semi-automatic approach “TOBAGO” (Grün and Dan, 1997), “CyberCity Modeler” was proposed in 1998 (Grün and Wang, 1998). It is now being commercially offered. This modeler is based on conventional, manual measurement of corresponding image points. It uses two basic steps.

In the first step (c.f. fig. 6, left), a so-called “weakly structured” point cloud is obtained by manual stereo measurement. It consists of all gable and roof points, possibly with additional points covering roof structures such as dormer windows. The point cloud is structured by two measures. First, the point measurement has to take place in a certain order, and second, points are assigned certain codes which will guide the second, automatic step. This can be obtained by placing the points into corresponding layers, a functionality which is usually available in existing photogrammetric or CAD software.

The second step is a fully automatic processing which uses a relaxation approach to derive the topology of the roof (c.f. fig. 6, right). Then, using the known correspondences between planes and points, the point positions can be corrected using a least squares adjustment. The result has to be checked, since in some cases, the relaxation might have recovered the wrong roof topology. However, in (Grün and Wang, 1998) success rates of 95% are given for the automatic structuring step.

As compared to a strictly manual measurement, CyberCity Modeler only automates the topology generation step. The point measurement itself is not accelerated. Since the modelling is based on points and planar patches only, the approach is able to reconstruct general polyhedral roof surfaces, as compared to methods using a limited set of building primitives (see next subsection).

5.2 “inJECT”

This system has been developed at the University of Bonn over a long period. It has evolved from earlier approaches termed “Hase”, “Hase+”, and “ObEx” (Gülch et al., 1999, Gülch et al., 1998). It is now being commercially offered under the name “inJECT”. Buildings are modelled using a fixed number of parametric primitives like flat-, desk-, saddleback-, hipped-roof etc. (see fig. 7). The selection of the appropriate primitive is carried out manually by an operator. Then, the wireframe model is overlaid in two images and the operator can adapt the parameters accordingly. A guided mode is available where the operator is asked to measure certain points only. For example, for a saddleback roof, the operator has to measure only the two gable points in one image – after that, the system tries to find the corresponding points in the other image as well as the remaining parameters automatically. After the automatic step, the operator can assess the result and can correct it by manual measurements, if necessary. For the automatic methods, a success rate of 50–90% is given in (Gülch et al., 1999). Complex buildings are subdivided by the operator into building parts which are modelled independently. The parts are then later on merged by an external CAD software to obtain the final buildings.

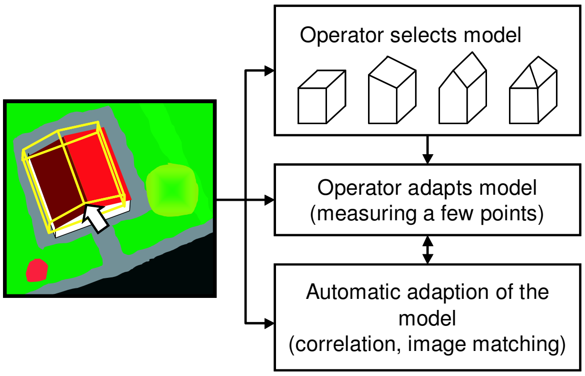

Figure 7: Illustration for ObEx / inJECT. The operator chooses a primitive, which is then fitted to the images using manual measurement as well as automated measurement procedures.

Using this constructive solid geometry (CSG) approach allows to reduce the number of required measurements by the operator considerably. It is even further reduced by the provided automatic matching and snapping procedures. On the other hand, as with any system which uses CSG as a modelling principle, some more complex buildings might be quite hard to model using the fixed set of primitives.

6 CONCLUSIONS AND OUTLOOK

Now, after about two decades of research into the topic of automated reconstruction of man-made objects, there are still no fully automatic systems around and the confidence that we may have to wait for them for another while is rather growing (c.f. (Förstner, 1999), p. 296). In the meantime, there are some semi-automatic systems sold today as products, which integrate single automation steps. Those systems are important because they do not only give us some figures of what we can expect as a speed-up from semi-automatic procedures, but also because they can serve as a testbed into which more and more automation steps can be integrated over time.

However, it is also interesting to see what insights into the problem have been gained by looking at the approaches suggested so far. Some of them them are:

Modelling. One can use general polyhedral models, constructed from measured vertices, lines, or planes. In this case, the question is how additional constraints are imposed, such as rectangularity, parallelism, two planar roof patches having the same slope, four roof patches meeting in one point, etc.. Alternatively, CSG modelling can be used which imposes those constraints automatically as part of the geometric primitive definition (it does not impose them across primitives, however). CSG is also usually more natural when buildings are to be constructed by a human operator. However, modelling buildings which do not fit well to the primitives can be quite tedious. Still another idea to construct buildings is using building parts as described in section 4.3. For the extraction process, general models like those of the “Gestalttheorie” mentioned earlier can be used.

Basic reconstruction principles. There are some basic ideas which have evolved and have made reconstruction systems better than their early predecessors:

- Use “rich attributes”, or a “polymorphous segmentation”. E.g. when extracting lines, possibly also extract the regions to the left and right of the line. Use color properties of those regions to assess line correspondences (c.f. section 4.2), or an region based correlation to recover planes in space (c.f. section 4.4).
- Use an early transition to object space. For example, assert if line segments extracted from multiple (two-dimensional) images are matching, parallel, forming rectangles etc. by checking the properties of the corresponding segments in (three-dimensional) object space.

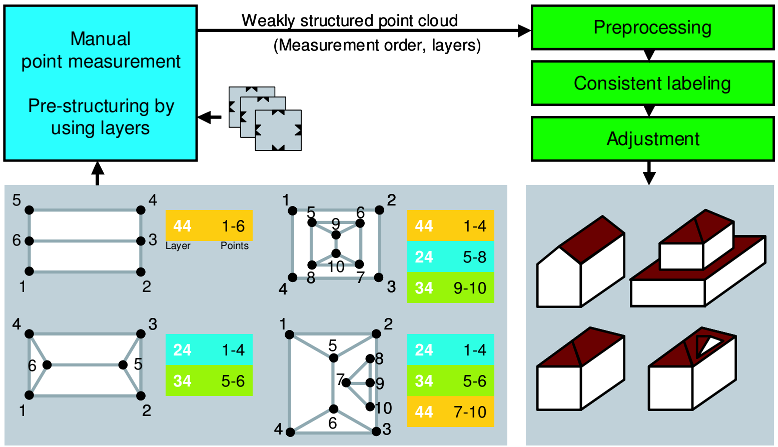

Figure 6: Illustration for the "CyberCity Modeler". A saddleback roof is obtained by measurement of the points 1–6 in succession, in a single layer. For a hipped roof, eaves and gable points are assigned different layers, for complex structures there are even more layers. From (Grün and Wang, 1998, Kersten and Cuche, 1999).

- Hierarchical / multilayer reconstruction models might help in organizing the reconstruction process (c.f. section 4.3).
- Use geometric constraints wherever possible. This can be constraints arising from the scene geometry, like epipolar lines, known minimum/maximum height, weak perspective etc. as well as constraints inherent in the model like parallelism, rectangularity, known angle ranges etc..
- Extract continuity and/or discontinuity, according to the underlying data. For example, with images it is often easier to extract line segments (discontinuities), whereas with DSM's, it is often easier to extract homogeneous regions (continuities), such as planar patches.

Which reconstruction steps are tackled? As mentioned above, building reconstruction can be divided into the different phases of detection, recognition, geometric reconstruction and possibly derivation of additional attributes. It is usually good if it is precisely stated which part of this chain is to be addressed by an algorithm.

Uncertainty. Representing uncertainty is one of the hardest problems. One can represent the uncertainty of all geometric entities like points, lines, planes and objects derived by construction from those entities. But how certain is a segmentation result, especially when multiple cues from different sources need to be combined? How can we handle the case when a small (continuous) difference in geometry leads to a different roof topology (a discrete difference)? Although there is no easy answer, it is good to think about which mechanisms (simple weighting, Bayesian nets, fuzzy logic etc.) are to be used to combine different cues.

Topological consistence. The representation of polygons in two dimensions is already a hard problem, and this gets worse in three dimensions. Polyhedral building models should not contain double surfaces, nor holes, nor any polygons with long thin spikes. This is not easy to attain, since the models are usually constructed from measurements and not by an idealized construction process as is usually the case in a CAD environment. Small errors in measurements (and of course rounding errors) often cause such artefacts.

Generalization. Often, this problem is not addressed at all, and sometimes people are even not aware of it. For example, given a dense triangulated irregular network (TIN) of three-dimensional points on a rooftop, the "best polyhedral reconstruction" of the roof is the TIN itself – because it *is* the surface of a polyhedron and has a minimum distance to the original TIN data, namely, zero. That is, what we search for actually is a more simple (generalized) description of the roof, which is not too far off the original TIN and ignores unimportant parts such as chimneys. In other words, it is a tradeoff between "simplicity" of description and distance from the original data. When using CSG, the generalization level is implicitly defined by the primitives provided. However, when an approach builds a roof from a general segmentation into planar patches, the intended generalization level has to be ensured by other measures.

As an outlook, one can say that while the fully automatic reconstruction from aerial laser scanning and images has not been solved yet, there is already a growing demand for the next generation three-dimensional geo-data – namely highly detailed models, requiring to capture façades using terrestrial laser scanning and close range photogrammetry. So in the near future, research will have to find a way to apply results from the aerial case to this new task, develop new, specialized algorithms for the terrestrial case, and devise methods to combine aerial and terrestrial sources efficiently.

REFERENCES

Ackermann, F., 1999. Airborne laser scanning – present status and future expectations. ISPRS Journal of Photogrammetry and Remote Sensing 54, pp. 64–67.

Ackermann, F. and Krzystek, P., 1991. MATCH-T: Automatic mensuration of digital elevation models. Presented paper to the 3rd Technical Seminar of the Sociedad Espanola de Cartografia Fotogrametria y Teledeteccion.

Baillard, C. and Zisserman, A., 1999. Automatic reconstruction of piecewise planar models from multiple views. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 559–565.

Baillard, C., Schmid, C., Zisserman, A. and Fitzgibbon, A., 1999. Automatic line matching and 3D reconstruction of buildings from multiple views. In: IAPRS, Vol. 32 Part 3-2W5.

Baltsavias, E., Grün, A., and VanGool, L. (eds), 2001. Automatic Extraction of Man-Made Objects from Aerial and Satellite Images (III). Balkema Publishers, Netherlands.

Baltsavias, E. P., 1999. Airborne laser scanning: existing systems and firms and other resources. ISPRS Journal of Photogrammetry and Remote Sensing 54, pp. 164–198.

Braun, C., Kolbe, T. H., Lang, F., Schickler, W., Steinhage, V., Cremers, A. B., Förstner, W. and Plümer, L., 1995. Models for photogrammetric building reconstruction. Computer & Graphics 19(1), pp. 109–118.

Brenner, C. and Haala, N., 1998. Rapid acquisition of virtual reality city models from multiple data sources. In: H. Chikatsu and E. Shimizu (eds), IAPRS Vol. 32 Part 5, pp. 323–330.

Collins, R. T., Hanson, A. R., Riseman, E. M. and Schultz, H., 1995. Automatic extraction of buildings and terrain from aerial images. In: A. Grün, O. Kübler and P. Agouris (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser, Basel, pp. 169–178.

Faugeras, O., 1994. Three-Dimensional Computer Vision. MIT Press.

Fischer, A., Kolbe, T., Lang, F., Cremers, A., Förstner, W., Plümer, L. and Steinhage, V., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. Computer Vision and Image Understanding 72(2), pp. 195–203.

Förstner, W., 1999. 3D-city models: Automatic and semiautomatic acquisition methods. In: D. Fritsch and R. Spiller (eds), Photogrammetric Week 99, Wichmann Verlag, pp. 291–303.

Fuchs, C., 1998. Extraktion polymorpher Bildstrukturen und ihre topologische und geometrische Gruppierung. PhD thesis, Rheinische Friedrich-Wilhelms-Universität zu Bonn, Institut für Photogrammetrie, DGK Reihe C, Nr. 502.

Grimson, W. E. L., 1990. Object Recognition by Computer. Series in Artificial Intelligence, MIT Press, Cambridge, Mass.

Grün, A. and Dan, H., 1997. TOBAGO – a topology builder for the automated generation of building models. In: A. Grün, E. Baltsavias and O. Henricsson (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhäuser, Basel, pp. 149–160.

Grün, A. and Wang, X., 1998. CC-modeler: A topology generator for 3-D city models. In: D. Fritsch, M. Englich and M. Sester (eds), IAPRS, Vol. 32 Part 4, Stuttgart, pp. 188–196.

Grün, A., Baltsavias, E. and Henricsson, O. (eds), 1997. Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhäuser, Basel.

Grün, A., Kübler, O. and Agouris, P. (eds), 1995. Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhäuser, Basel.

Gülch, E., Müller, H. and Läbe, T., 1999. Integration of automatic processes into semi-automatic building extraction. In: IAPRS, Vol. 32 Part 3-2W5, München.

Gülch, E., Müller, H., Läbe, T. and Ragia, L., 1998. On the performance of semi-automatic building extraction. In: ISPRS Comm. III Symposium, Columbus, Ohio.

Haala, N., 1996. Gebäuderekonstruktion durch Kombination von Bild- und Höhendaten. PhD thesis, Universität Stuttgart, Institut für Photogrammetrie, Deutsche Geodätische Kommission, C 460.

Henricsson, O., 1996. Analysis of Image Structures using Color Attributes and Similarity Relations. PhD thesis, Institut für Geodäsie und Photogrammetrie, ETH Zürich, Mitteilungen Nr. 59.

Henricsson, O. and Baltsavias, E., 1997. 3-D building reconstruction with ARUBA: A qualitative and quantitative evaluation. In: A. Grün, E. Baltsavias and O. Henricsson (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhäuser, Basel, pp. 65–76.

Jain, A. K. and Flynn, P. J. (eds), 1993. Three-Dimensional Object Recognition Systems. Elsevier.

Kersten, T. and Cuche, L., 1999. Auf dem Weg zur virtuellen stadt (Cyber City) – Photogrammetrische Datenerfassung für 3-D Stadtmodelle. In: Proc. Workshop "Generieren, Visualisieren und Verwalten von 3D-Stadtmodellen", ETH Zürich, 9.2.1999.

Lang, F., 1999. Geometrische und semantische Rekonstruktion von Gebäuden durch Ableitung von 3D-Gebäudeecken. PhD thesis, Rheinische Friedrich-Wilhelms-Universität zu Bonn, Institut für Photogrammetrie, Shaker Verlag, Aachen.

Lin, C., Huertas, A. and Nevatia, R., 1995. Detection of buildings from monocular images. In: A. Grün, O. Kübler and P. Agouris (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser, Basel, pp. 125–134.

Lowe, D., 1985. Perceptual Organization and Visual Recognition. Kluwer Academic Publishers, Boston, Mass.

Paparoditis, N., G.Maillet, Taillandier, F., Jibrini, H., Jung, F., Guigues, L. and Boldo, D., 2001. Multi-image 3d feature and DSM extraction for change detection and building reconstruction. In: E. Baltsavias, A. Grün, and L. VanGool (eds), Automatic Extraction of Man-Made Objects from Aerial and Satellite Images (III), Balkema Publishers, Netherlands, pp. 217–230.