'Photogrammetric Week 01' D. Fritsch & R. Spiller, Eds. Wichmann Verlag, Heidelberg 2001.

City Models – Automation in Research and Practice

CLAUS BRENNER, Hannover

ABSTRACT

This paper concentrates on the role of automated extraction of city models in today's practical work. The state of research and practical implementations is compared and the question is raised why the degree of automation in modelling systems is still very low – despite many years of scientific research efforts. Some of the problems are identified, in particular the need to achieve a high degree of automation and to assess the advantage of automatic components by looking at the effects on the entire processing chain rather than only single stages. Three examples for automation are given. The first two are approaches for the reconstruction of building geometry which integrate 2D ground plan information with digital surface models. A heuristic subdivision of ground plans into rectangles as well as a rule-based reconstruction relying on discrete relaxation and constrained tree search are discussed. The third example shows how quality control can be automated using digital surface models.

1. INTRODUCTION

Interest in three-dimensional building models has raised significantly in the past years. Originally, simulations for the propagation of electromagnetic waves were thought of being one of the major application areas. These are used by network operators for the planning of antenna locations. Although there might be an additional need in the near future due to the forthcoming introduction of UMTS networks in Europe, there have been other areas evolving, for example three-dimensional car navigation systems, virtual tourism information systems and visualisation for city and building planning or architectural contests.

However, it seems that the most important development-driving factor will not come from those single applications but rather from a principal change which has begun to evolve in the last decade with the widespread use of mobile phones and is characterised by phrases as “ubiquitous computing”, “location based services” and “augmented reality”. Following this development, location-dependent access to spatial data will become available to any user, (almost) anywhere. Key features will be the accurate localisation of the user, the location-dependent processing and transmission of (spatial) data via broadband, wireless networks, availability inside and outside of buildings, and the use of small, yet powerful, mobile devices (Hamburgen et al. 2001).

Therefore, it is an important task to collect spatial data like city models and—in order to keep them useful—to update them in short cycles. It is also a formidable task comparable to the digitisation of the road network for navigation purposes. Thus, it becomes clear that any economically viable solution must rely on automatic or at least partially automatic methods.

2. AUTOMATION

2.1. ...in research

Photogrammetric methods are well suited for the economic acquisition of 3D city models (Förstner 1999), making it possible to recover structure as well as the dimensions. On the other hand, classical photogrammetric measurement is mostly point-based, which does not exploit the inherent structure of buildings and thus cannot be optimal economically. Standard automatic image matching techniques, developed originally for the measurement of terrain points, proofed to be not as effective for built-up areas due to the large discontinuities which arise at building borders and roofs. This contradiction has led to substantial research efforts (see e.g. Grün et al. 1995, 1997,

Baltsavias et al. 2001), and many reconstruction approaches have been proposed. Some coarse classification can be made with regard to the used data sources and the distinction between semiautomatic and fully automatic operation.
Automatic reconstruction from aerial images (e.g. Haala 1994, Henricsson & Baltsavias 1997, Fischer et al. 1998, Baillard & Zisserman 1999) has shown promising results, however one has to note that often special image material has been used which is not available in general, for example large scale, multiple overlap, or colour images, or additional height models. Even then, the reliable extraction of buildings in densely build-up areas has not been demonstrated yet.
Automatic systems working solely on the basis of digital surface models (DSMs) acquired by laser scanning have been reported (Brunn & Weidner 1997, Maas 1999). Since DSMs represent the geometry of the surface directly, they have advantages with regard to automated interpretation. They used to suffer from the low density of measured points, however this is no technical restriction anymore, since a density of several measured points per square meter is attainable. Also, the recording of multiple return pulses and intensity values is standard nowadays, opening up new possibilities for automated segmentation techniques.
Semiautomatic approaches have been reported for both image- and DSM-based systems. They can be divided into approaches which model buildings from a fixed set of volumetric primitives which are combined (e.g. Gülch et al. 1999, Brenner 1999) and approaches which build the topology of the surface directly (e.g. Grün & Wang 1998).

2.2. ...and in practice?

Now that we have learned about the need to automate city model generation and also know what substantial amount of research has been devoted to this goal, one might wonder how much research has found its way from the lab into practical applications.
However, the current state seems to be somewhat disappointing. For example, as of 2001, Phoenics – a major supplier of city models for telecommunications companies in Germany – has a database of over 100 city models covering an area of about 30.000 km^2. Most of the building models consist of vectorised outlines and a single height (i.e., flat roofs). All models were measured *manually* using digital photogrammetry. How can it be that none of the fully automatic or semiautomatic approaches discussed in research have found their way into practice? There are several reasons:

- Fully automatic systems based on aerial or satellite images are not reliable enough, although tremendous progress has been made from a research viewpoint. Seen from a practical perspective, however, it makes not much sense to use unreliable automatic systems, because the effort required to manually correct the obtained results may even be higher than a strictly manual measurement in the first place. This is especially true for the automation of the *building detection* and the *structuring* task which have proven to be demanding – which means that in practice, all the images have to be scanned manually for missed buildings, apart from checking the reconstruction of all automatically detected and structured ones.
- Semiautomatic systems for measuring structures (as opposed to points) in images offer a high potential for practical application in the near future. Combining manual detection and structuring with automatic measurement, they can also serve as a testbed for successive integration of automatic components. Consequently, there are first systems on the market such as CyberCity Modeler (CCM) from Cyber City AG, Switzerland or inJECT from inpho GmbH, Germany. The first tries to speed up the overall time required for modelling by an automated topology building step based on a manual structuring whereas the latter builds complex structures from primitives using constructive solid geometry (CSG).
  Again, for practical purposes, there are some caveats. Being semiautomatic, the reduction of manual processing steps as opposed to a *strictly manual* (but, say cleverly organised) measurement tool is noticeable, but usually not dramatic. For example, with CCM's approach,

still all points which are part of the roof have to be measured manually. Modelling using primitives – as used by inJECT – offers a greater reduction potential, since primitives can be instantiated with very few manual interaction steps and subsequent measurement can be fully automatic. On the other hand, modelling of complex buildings from a fixed set of primitives can be cumbersome, time-consuming and sometimes even impossible.
If one considers that there are fixed costs in any project, such as aerial flight, image scanning, aerotriangulation or quality control, the total speedup obtained by semiautomatic systems becomes even smaller. Figure 1 shows an example where a 40% reduction in measurement time leads only to a moderate total reduction of 11% – at the expense of software license and possibly training costs.
Thus, although it seems that semiautomatic tools are the only way to introduce automatic components into the measurement process, their performance seen from an economical viewpoint is still critical.

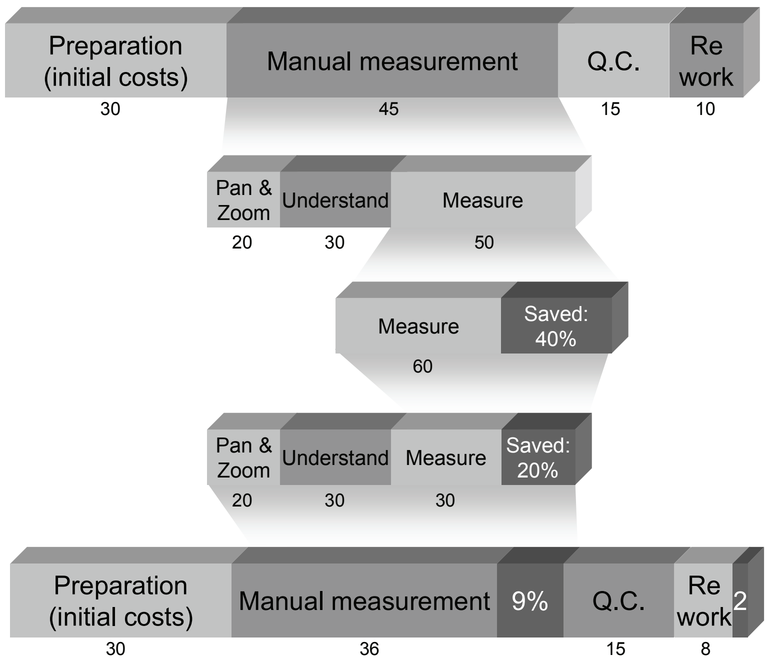


Figure 1: Illustration regarding the effect of using a semiautomatic tool (all numbers are estimated percentages, Q.C. stands for quality control). From top to bottom: a project consists of 45% manual measurement (plus 10% manual rework), of which 50% is amenable to speedup by a semiautomatic component, while the remaining 50% are needed by an operator in each case for pan, zoom and understanding the scene. If the tool allows to work 40% faster, this translates to an overall speedup of only 9% (or, 11% if one assumes the tool can be used for rework as well).

- Generating city models is not easily "scalable" with respect to the necessary data sources and algorithms. Indeed, there can be vastly different project requirements regarding detail, accuracy, model types, texture and vegetation, imposing the use of different algorithms and data sources. The limits are given not only by technical feasibility but also by economic constraints. Thus, actually many different tools are needed instead of one single solution.
- It seems that research has not addressed some of the practical problems. For example, a lot of work has concentrated on the "all-in-one" approach which strives for a solution for all steps of the extraction process – detection, structuring and measurement. However, there might be subtasks of more economical interest. For example, with all the models already present, the question of efficient update and upgrade comes into focus. *Updating* a city model with

photogrammetric methods might require about 60-70% of the initial effort. Also, since expectations regarding detail may rise over time, automatic methods to *upgrade* from existing coarse (e.g. flat roof) to detailed (e.g. full roof geometry) models are of interest.
- For automation, one has always to consider the process consisting of (at least) measurement *and* quality control *and* required rework. It makes no sense for an automatic component to reduce the effort for measurement slightly while at the same time costs for quality control and rework raise substantially. Ideally, approaches should also strive for an automation of quality control, for example by a verification through classification or using digital height models.

3. EXAMPLES FOR AUTOMATION

3.1. Reconstruction using ground plan segmentation

A high percentage of buildings can be modelled using a small number of building primitives like flat boxes, boxes with saddleback and hip roofs and other geometric primitives like cylinders and cones. Even if one considers only primitives based on rectangular ground shapes, still the majority of buildings can be modelled.
Thus, one approach to reconstruct buildings when 2D ground plans are available can be sketched as follows:
- Try to infer (in 2D) from the ground plan how the building can be subdivided into primitives.
- Select each primitive based on additional information, e.g. an aerial image or a DSM. Determine the dimensions by a measurement process.
- Assemble all primitives to obtain a single body representing the building.

For example, for a simple L-shaped building the first step would select two rectangles to cover the 2D ground plan, the second step would select hip roofs on each of those rectangles and estimate the eaves and ridge heights and the third step would finally merge both volumetric primitives.
This approach has been used by the authors in several projects, combining 2D ground plans with DSMs from laser scanning (Brenner 2000). Its main advantage is that by the use of digitised ground plans, interpreted information is “injected” into the reconstruction process which makes it relatively simple and reliable. On the other hand, since the selection of primitives is guided by analysing the ground plan, the final reconstruction is strongly coupled to the ground plan shape. For example, roof structures like dormer windows will not appear in the result as long as there is no corresponding hint in the ground plan. This situation can be improved by interactive modelling tools, however.
A second disadvantage is that while most simple buildings lend themselves well to a subdivision into primitives, there are always buildings present which cannot be modelled properly or for which a subdivision into primitives is not very natural and thus modelling becomes involved. In those cases, other approaches which allow to specify the roof topology directly would be more desirable.

3.2. Reconstruction using tree search and DSM segmentation

3.2.1. Ground plans and simple roofs

We consider a ground plan given as a closed polygon P consisting of n linear segments p_i and vertices v_{12}, v_{23}, ..., $v_{n-1,n}$, $v_{n,1}$ (where P does not contain any inner polygon), see Figure 2. Based on P, a simple roof can be constructed which consists of planar roof faces Π_i which intersect with the building walls in the eaves P_i and eaves points $V_{i,i+1}$. The intersections of the roof faces yields ridges (E_{12}, E_{24}) and grooves. Projected down into the plane the corresponding points and lines

(indicated by lowercase letters) are obtained. If the original ground plan is omitted, the remaining points and segments form a *planar graph G*. It consists of leaves v_{ij}, inner nodes v_{ijk} and edges e_{ij}. If the indices are viewed as sets, then the edge e_A connects nodes v_B and v_C iff $A=B\cap C$. The graph *G* is connected and without cycles which means it is a tree.
Simple roofs such as the one in Figure 2 can be recovered from the corresponding planar Graph *G* by adding a ground and eaves height. The graph *G* itself can be obtained from the ground plan *P* by computing the *straight skeleton S(P)*. For convex ground plans, the straight skeleton *S(P)* is identical to the *medial axis transform M(P)*. It is different for ground plans containing concave nodes, in which case *M(P)* contains parabolic segments. The construction of the straight skeleton is a well-known process in geometry but has received attention recently (Aichholzer et al. 1995). Interestingly, no time-optimal algorithm has been found so far (Eppstein & Erickson 1999).

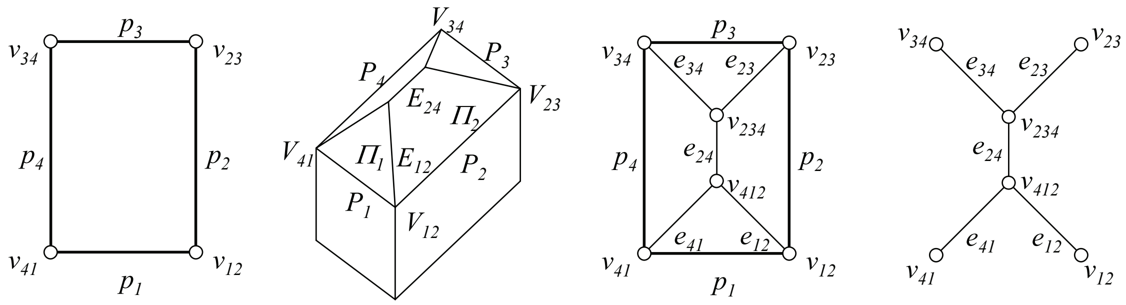


Figure 2. From left to right: ground plan, corresponding simple roof, projection into the plane and corresponding graph (which is the skeleton *S(P)* of *P*).

3.2.2. From construction to search

The straight skeleton *S(P)* can be constructed from a given ground plan *P* unambiguously. However, it is often not the only roof which can be raised on *P*. Figure 3 shows an example in which *S(P)* would correspondent to a roof of unusual height and volume. In order to obtain all possible solutions, the constructive algorithm has to be replaced by a search procedure. If a roof face Π_i is defined emerging from each eaves edge P_i, extending inwards, then all possible vertices V_{ijk} can be constructed by intersecting the corresponding planes Π_i, Π_j and Π_k. The graph *G* is then obtained by searching all possible edge combinations between those vertices.
Considering the first step, the intersection of every three roof planes has a time complexity of *n* over 3, or O(n^3), i.e. polynomial in the number of ground polygon segments. However, the subsequent search has exponential complexity and is usually not solvable except for very small values of *n*.

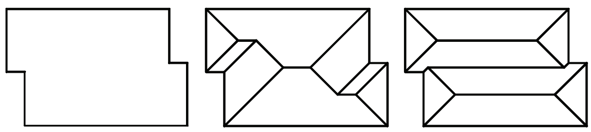


Figure 3. From left to right: ground plan *P*, *P* together with graph *G* corresponding to the straight skeleton and another possible graph *G'*.

Fortunately, one can easily do better by labelling edges and nodes. This is similar to the approaches for the interpretation of line drawings – a task which has been a topic of artificial intelligence research from its very beginning (Waltz 1975). Indeed, edges e_{ij} are the projection of the intersection of two planes where (depending on the selection of half-planes) either a concave (–) or a convex (+) intersection is present (Figure 4). Thus, any edge e_{ij} allows two different interpretations. Similarly, it turns out that each node v_{ijk} allows for eight different interpretations, where each interpretation defines the type (+/–) as well as the direction of all incident edges.

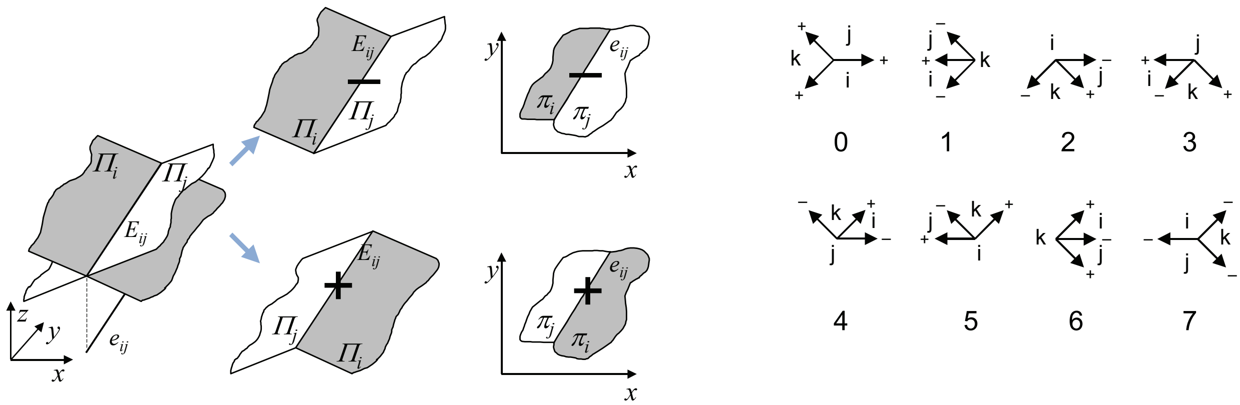
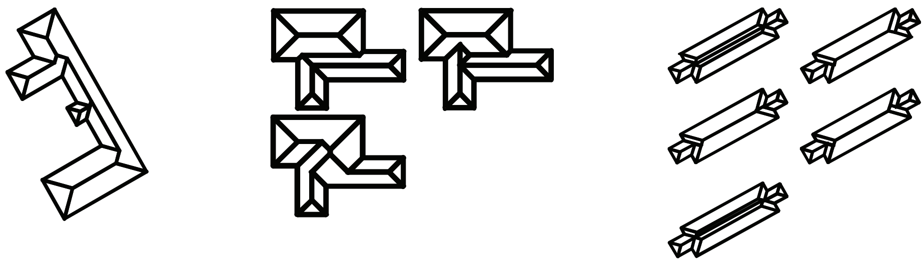


Figure 4. Possible interpretations of edges (left) and nodes (right).

Thus, a constrained tree search can be used to cut down search space. A still more dramatic reduction can be obtained when a discrete relaxation is used beforehand in order to reduce the number of possible edge and node interpretations. Figure 5 shows some examples. In general, discrete relaxation propagates constraints on the number of possible interpretations starting from the ground polygon nodes towards inner nodes (since the interpretation of ground polygon nodes $v_{i,i+1}$ and edges $e_{i,i+1}$ is determined by the shape of the ground plan).

(a)



(b)	364 (100%)	120 (100%)	220 (100%)
(c)	63 (17%)	50 (42%)	63 (29%)
(d)	29 (8%)	31 (26%)	55 (25%)
(e)	48	65 – 727	887 – 67848
(f)	5602	1647 – 168085	269837 – 795654

Figure 5. (a) Ground plans and possible roofs found by the search. (b) Maximum number of nodes (which is *n* over 3). (c) Number of nodes which are inside the ground plan. (d) Remaining nodes after discrete relaxation. (e) Required search steps when discrete relaxation is used (range when more than one solution exists). (f) Required search steps without discrete relaxation.

3.2.3. Incorporating digital surface models

So far, the generation of roofs based solely on ground plans has been discussed. However, in order to obtain faithful reconstructions of real world objects, not only the dimensions but also the structure has to be inferred from measurements – in our case DSMs from laser scanning.

There are numerous techniques to segment DSMs into meaningful regions, for example region growing, line grouping (Jiang & Bunke 1994) or robust estimation techniques like RANSAC (Fischler & Bolles 1981). Usually, a DSM segmentation into *planar* regions is sensible for roof reconstruction purposes, although with higher DSM point densities, one might also try to recover higher order surfaces. Figure 6 shows an aerial image of a complex building and a planar region segmentation obtained from a 1 m grid DSM. As can be seen, all important regions are captured, however there are many other small regions present. In order to reconstruct the roof's surface, those have to be removed or merged with other regions. One could use general criteria for this such as the size or shape of the regions. However, better results can be obtained when a building specific model is used as in the following rule-based approach.

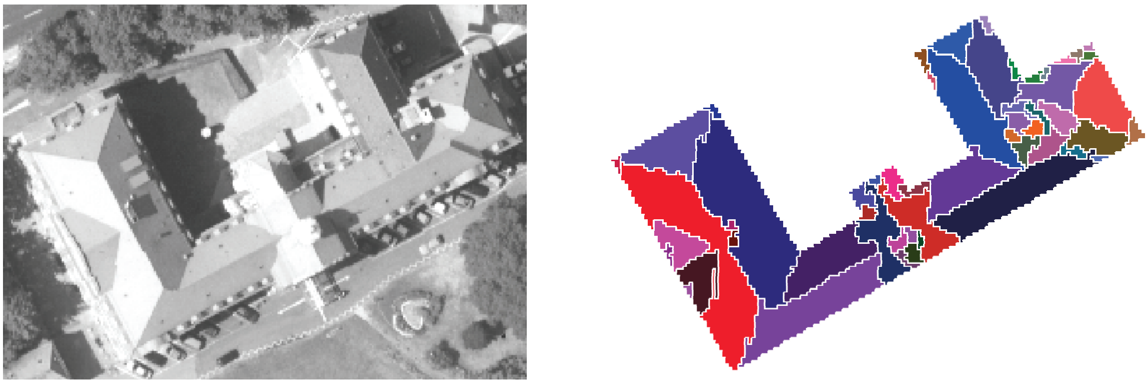
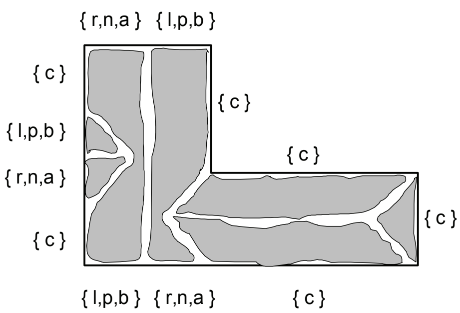


Figure 6. Left: aerial image of a building with complex roof. Right: regions obtained from a planar DSM segmentation.

3.2.4. A rule-based approach for the acceptance of regions

If we compare the normal vector of a segmented region with the orientation of an adjoining ground plan edge, we can derive a certain relationship. For example, for the standard case where a roof "evolves" from the corresponding ground plan edge (the normal vector is perpendicular to the edge, pointing outwards), we would label it as "compatible" (c). Similarly, there are labels for compatibility with the previous (p) and next (n) edge, with opposite vectors of the previous (a) or next (b) edge or perpendicular to the current edge to the left (l) or right (r). Each region can have several labels, where with rectangular ground plans, {l,p,b} and {r,n,a} are typical label sets (see Figure 7, left). For example, it can be seen that dormer windows result in the typical sequence c, l, r, c. Regions are accepted based on the rules shown in Figure 7, right, where "<" stands for the start and ">" for the end of a pattern along one edge, „x^+" stands for one or more occurrences of the pattern "x" and "*" stands for any pattern. The order in which the rules are applied is of importance. All remaining regions along an edge labelled with "c" are accepted. Figure 8 shows the result of the rule application on the segmentation from Figure 6.



Detected pattern		Accepted Pattern
$<p^+a^+c^+$	→	$<p^+a^+$
$c^+b^+n^+>$	→	$b^+n^+>$
$<p^+ * n^+>$	→	$<p^+$ und $n^+>$
l^+r^+	→	l^+r^+

Figure 7. Left: an example of segmented regions, ground plan and resulting labelling. Right: rules used to accept a subset of regions.

In order to derive a closed roof topology, the search method of section 3.2.2. can be modified. Although it has been presented in conjunction with a ground-plan based reconstruction, it only uses the intersection of planes. Instead of implicit planes emerging from their corresponding ground plan edges, the planes defined by accepted regions can be used. The result of this step is shown in Figure 8, right. It can be seen that the gap in the right area of the building (Figure 8, left) which is due to non-accepted regions has been closed. While this reconstruction is not fully correct, it is in accordance with the model as given by the rules. In contrast to the reconstruction of Figure 5, the rooftop in the middle part of the building as well as the dormer window in the left part are captured.

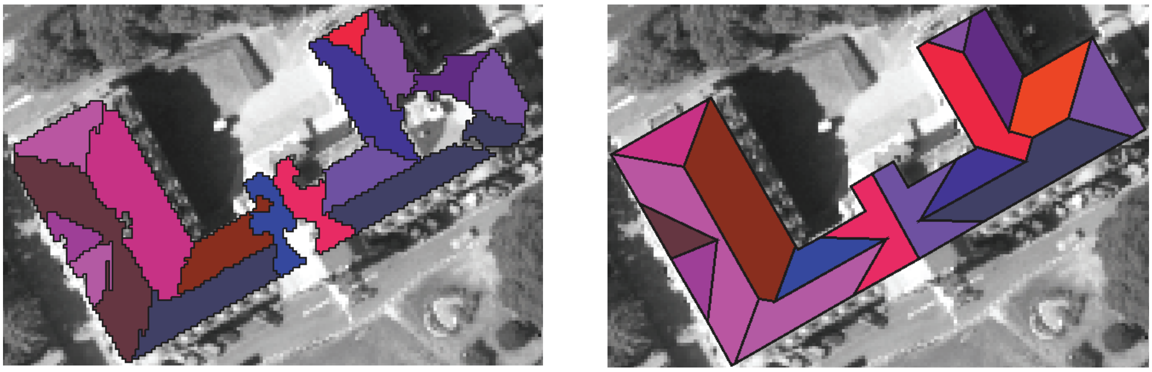


Figure 8. Accepted regions (left) and final reconstructed topology (right).

3.3. Quality control using digital surface models

When a DSM is available, simple measures can be defined which assess the deviation of the measured and the reconstructed surface. For example, for each DSM point, the pointwise absolute error $\Delta z = | z_{DSM} - z_{REC} |$ can be used, where z_{DSM} and z_{REC} denote the measured and reconstructed height, respectively. From this, an average absolute error for each building can be obtained. Figure 9 shows a histogram which plots the average absolute error against the number of buildings for a project involving more than 2000 buildings of the city of Vienna.

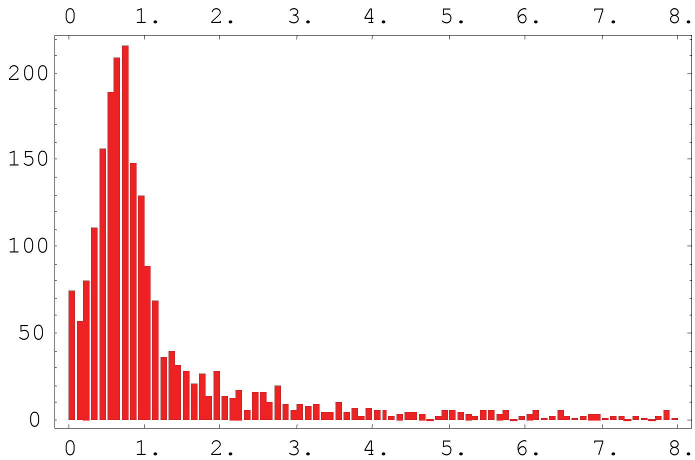


Figure 9. Histogram of average absolute (per-building) errors. Horizontal axis: error in meters. Vertical axis: number of buildings in histogram bucket.

This histogram gives an indication on how well the measured surface is represented by the reconstruction. However, it is much more interesting to see directly which buildings are affected by large errors. Although this test does not give results on how well the *structure* of the roof is represented, it still can serve as an indication on where further work is required. Figure 10 shows an example for average absolute (per-building) errors before and after manual rework.

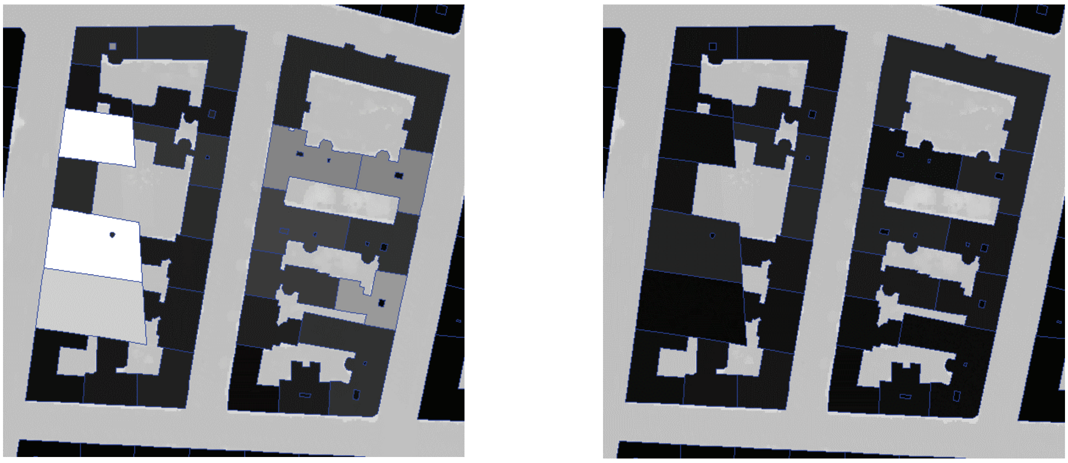


Figure 10. Top: Ground plans coloured according to the average absolute error of their reconstruction (black: 0 m, white: 5 m). Left: after fully automatic reconstruction. Right: after manual rework.

4. CONCLUSIONS

It seems that we look back on more than a decade of research effort dedicated to the automated extraction of man-made objects which had unfortunately only little impact on the way how virtual city models are practically produced nowadays. This paper identifies some of the reasons for this situation. In particular, semiautomatic methods will be part of an economically viable solution only if they offer a substantial speedup of the entire processing chain. This means that only highly automated components can be used, which should not only speed up measurement but also should not increase the time spent for quality control and rework. To that end, the paper presents two methods for building reconstruction from ground plans and DSM's as well as an approach to automatically assess the quality of the reconstructed rooftops.

In order to reach practitioners, it may be that in the future, research has to focus more clearly on either the "purely scientific" or the "application track". The former is interested in how one could possibly attain a recognition performance comparable to human beings. This is closely related to object recognition and scene understanding in general and thus is a long-term goal. The latter should concentrate especially on how contemporary systems could be combined in order to reach an economically good solution in the sense of solving an engineering problem – even though this solution might require to capture much more data than would be necessary for a manual processing of the scene.

5. REFERENCES

Aichholzer, O., Aurenhammer, F., Alberts, D., Gärtner, B. (1995). A Novel Type of Skeleton for Polygons, Journal of Universal Computer Science 1(12), 752 – 761.

Baillard, C., Zisserman, A. (1999). Automatic reconstruction of piecewise planar models from multiple views. Proc. IEEE Conference on Computer Vision and Pattern Recognition, 559-565.

Baltsavias, E., Grün, A., Van Gool, L. (Eds.) (2001): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III).* Balkema, Rotterdam.

Brenner, C. (1999). Interactive modelling tools for 3D building reconstruction. In: D. Fritsch & R. Spiller (Eds.), *Photogrammetric Week '99*, p. 23–34, Wichmann Verlag.

Brenner, C. (2000). Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen, Ph.D. Thesis, Deutsche Geodätische Kommission, Vol. C 530 (in German).

Brunn, A., Weidner, U. (1997). Extracting Buildings from Digital Surface Models. In: IAPRS, Vol. 32, Part 3-4W2, Stuttgart.

Eppstein, D., Erickson, J. (1999). Raising Roofs, Crashing Cycles, and Playing Pool: Applications of a Data Structure for Finding Pairwise Interactions, Discrete & Computational Geometry 22, 569 – 592.

Fischer, A., Kolbe, T., Lang, F., Cremers, A., Förstner, W., Plümer, L., Steinhage, V. (1998). Extracting Buildings from Aerial Images using Hierarchical Aggregation in 2D and 3D. Computer Vision and Image Understanding 72(2), 195-203.

Fischler, M. A., Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Comm. of the ACM 24(6), 381 – 395.

Förstner, W. (1999). 3D-City Models: Automatic and Semiautomatic Acquisition Methods. In: D. Fritsch & R. Spiller (Eds.), *Photogrammetric Week '99,* 291–303, Wichmann Verlag.

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

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

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

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

Haala, N. (1994). Detection of Buildings by Fusion of Range and Image data, IAPRS Vol. 31 Part 3, Munich.

Hamburgen, W. R., Wallach, D. A., Viredaz, M. A., Brakmo, L. S., Waldspurger, C. A., Bartlett, J. F., Mann, T., Farkas, K. I. (2001). Itsy: Stretching the Bounds of Mobile Computing, IEEE Computer, 18(4), April, p. 28–36.

Henricsson, O., Baltsavias, E. (1997). 3-D Building Reconstruction with ARUBA: A Qualitative and Quantitative Evaluation. In: Grün, A., Baltsavias, E., Henricsson, O. (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, 65-76.

Jiang, X. Y., Bunke, H. (1994). Fast Segmentation of Range Images into Planar Regions by Scan Line Grouping, Machine Vision and Applications 7(2), 115 – 122.

Maas, H.-G. (1999). Two Algorithms for Extracting Building Models from Raw Laser Altimetry Data. ISPRS Journal of Photogrammetry and Remote Sensing 54, 153-163.

Waltz, D. L. (1975). Understanding line drawings of scenes with shadows. In: P. H. Winston (Ed.), Psychology of Computer Vision, McGraw-Hill, New York, 19 – 91.