3D CITY MODELS: AN OPERATIONAL APPROACH USING AERIAL IMAGES AND CADASTRAL MAPS

D. Flamanc, G. Maillet, H. Jibrini

Institut Géographique National, Laboratoire MATIS, Saint-Mandé, France – david.flamanc@ign.fr

Commission III, WG III/3

KEY WORDS: Urban, Three-dimensional, Aerial, Building, Reconstruction, Semi-automation.

ABSTRACT

Even though many efficient building reconstruction methods have been published so far, automatic processes are still missing for the production of 3D city models. This gap between research results and effective production tools seems mainly due to a lack of cooperation between all these methods. Building reconstruction in dense urban areas is such a complex problem that it is hopeless to look for an universal solution that could efficiently reconstruct every building, from a simple gabled roof to a whole cathedral. In this paper we first present a general strategy to combine several automatic and user-assisted processes to produce three-dimensional city models in a real production context using aerial images and 2D ground maps. The strategy aims at taking advantage of each algorithm to be sure they complement each other. Then we illustrate this strategy by describing our operational implementation and its results. In this implementation, the first step is a combination of two automatic algorithms: a model driven and a data driven process. In a second step, user-assisted tools supplement the reconstruction: some are adapted to precise landscape features while others are more generic.

1 INTRODUCTION

During the past decade, 3D city models production has been a constant topic of research for several universities and cartographic or photogrammetric institutes. Potential applications for these models have moved from electromagnetic propagation for telecommunication to more demanding simulations for acoustic, urban planning, virtual or augmented reality applications. The two most important points concerning this evolution are: the confirmation of a great potential for these 3D city models, but at the same time, a drastic need of automation to obtain high quality reconstructions at a reasonable cost. We have here focused our attention on building reconstruction because buildings are the most important features, and probably, the most expensive ones for 3D city models production.

As said in (Brenner, 2001), there is a huge gap between research results and production tools. In our opinion, this is due to a lack of cooperation between all the available approaches. Building reconstruction in urban areas is a very complex problem (Figure 1). Using a single approach, semiautomatic or automatic, to efficiently reconstruct both a cathedral and a simple gabled roof is hopeless in our opinion. Each algorithm has its own pros and cons, its own specificities in terms of an urbanistic context. Therefore, we propose to combine several approaches with several levels of automation (automatic, user-assisted and manual) to obtain a truly operational piece of software. Before presenting our global strategy and its implementation, we will start with a brief overview of the published building reconstruction approaches. These approaches may be split in two categories: semiautomatic and full automatic processes.

Concerning semiautomatic approaches, several kinds of interactions can be used:

- A very popular approach uses two steps: first a choice of a building model is made, and then some mouse clicks instantiate it. See (Gülch and Muller, 2001) or (Zhang et al., 2001).
- The operator interacts with an automatic process: it could be a simple validation step, an edition of the input data as presented in (Brenner, 1999), or the choice of the appropriate solution in a proposed set of reconstruction. In these cases, a self validation of the automatic process or a quality control using external data may help the operator by focusing his attention on suspicious buildings.
- Finally, introducing geometric informations may be performed by the operator like in CCModeler (Gruen and Wang, 2001).

For fully automatic approaches, the process strategy is often guided by the used data:

- Aerial images are still a natural and popular data for building reconstruction. It is a very rich source of informations, even if the analysis of this low level information in dense urban context is difficult. The reconstruction can use different techniques : perceptual organization (Nevatia and Price, 2002), hierarchical aggregation (Fischer, 1998), parametric models (Weidner, 1996) or structural approach (Fuchs, 2001), ...
- Laser scanning has known great progresses in terms of density and acquisition cost. Laser DSMs provide very reliable information and well appropriate for segmentation and parametric reconstruction (Maas, 1999), (Haala and Brenner, 1999).
- But in all cases, 2D ground maps are now widely used to obtain more operational systems (Brenner et al., 2001), (Jibrini et al., 2000), (Vosselman and Suveg, 2001). In lots of countries these 2D ground maps are already available or at least easy to produce. This knowledge allows to jump a huge step in terms of quality and reliability.

2 OUR GLOBAL STRATEGY

As we already said, in large urban areas, building complexities are often very changing, but locally it is often possible to observe

Figure 1: Different types of landscapes.

a certain homogeneity. Different areas may be identified in the landscape: suburbs, industrial buildings, village, old or modern downtown,... Our strategy is to offer several building reconstruction methods in a common software environment in order to make it possible for the operator to adapt his reconstruction process to each local context.

Our system accepts different types of data: images with known calibration, existent 2D or 3D data-bases and DSM (from image matching or laser scanning). When these DSM are not available, the system is able to compute them with two different algorithms: one (Baillard and Maître, 1999), used for a global analysis (to choose the appropriate image for every 2D ground area, or for a building detection), and a second one based on a multi-view process (Paparoditis and Maillet, 2001) combined with a 2D optimization (Roy and Cox, 1998) for a more precise local geometric information. In addition to these data, the system can also manage several types of mid-level features (2D or 3D points, line segments, faces, or TIN DSM).

All the results shown in this paper use only aerial images and 2D ground maps, without laser DSM. Of course, these laser scanning DSM could improve some results, but in our cartographic institute context, aerial images are still necessary (for orthoimages, roads extraction and thematic interpretations); thus the additional cost of a laser acquisition for this task is not really justifiable from our point of view.

In this first version of the system, we chose to integrate the following tools:

- Two automatic approaches using aerial images and 2D ground maps:
  - A model-driven approach: it tries to infer a roof structure with an initial 2D polygonal shape. This method is used for simple buildings.
  - A data-driven approach: this approach always proposes a roof structure (Jibrini, 2002).
- Two types of user-assisted tools with various levels of automation:
  - Tools specialized for a particular context: a Copy/Paste tool for suburbs, and a tool reconstructing some old-style downtown from the central ridge.
  - More generic tools: the operator chooses a type of model and instantiates it with some mouse clicks.
- And finally some manual tools to edit the results.

The system also manages some meta-data that can be used during a checking process of the automatic reconstructions. The operator can validate a group of reconstructed buildings, but he can also edit or suppress these results. Actually, this checking stage is the major automation limit of our system. For the moment our system is driven by an operator who has to decide which treatments are adequate for each area, and has to check all the results. On a future version of this system, we will work on an automatic cooperation between the two automatic processes with a warning mechanism.

In the following paragraphs, we will describe these elementary reconstruction methods in a chronological order. Here the point is not the absolute performance of each of these methods (a lot of other existent algorithms could be incorporated in this global frame), but rather the way to take the advantage of each one to ensure that they complement each other.

3 AUTOMATIC STAGE

In a first step, we try to deal with all the buildings that could be fully automatically reconstructed. For this step we use all the available images and a 2D ground map which has been manually reworked in some aeras.

3.1 Model driven approach: a simple method for simple buildings

In particular urban landscapes like certain suburbs, buildings have very simple roofs. In these cases, the structure of the roof can be fairly infered from the knowledge of the 2D shape. So, we propose an automatic treatment that constructs a set of likely roof structure based on the 2D shape of the building, and then, we fit these roof hypotheses with an accurate local DSM to build a 3D reconstruction. This fitting stage also produces a score that can be used to choose the best roof hypothesis.

In (Brenner, 2000) a complex mechanism has been presented to infer such a set of roof hypotheses. But, even if this approach seems to give very impressive results, the author also underlines the limits of this strategy. A 2D ground map is not enough to always infer the precise roof structure: some roof elements as dormer windows cannot be seen on the ground map. On the opposite some small details of the ground map may be meaningless for the roof structure. This explains we chose to not reconstruct these small details, even if they are visible on the 2D ground map. We propose a pragmatic process based on the search of symmetric slopes.

The basic idea is to find a central ridge that divides the 2D shape in two symmetric half buildings. To do that, we try, for all the sides of the 2D polygonal shape, to build an axis that joins the middle M_i of the homologous points $(P_{i,r} - P_{i,l})$. This axis is considered as a possible central ridge if all the homologous angles $(P_{i-1,r}\widehat{P_{i,r}}P_{i+1,r})$ and $(P_{i-1,l}\widehat{P_{i,l}}P_{i+1,l})$ are near equal (Figure 2). Such a ridge hypothesis can only be found on a polygonal shape with an even number of sides but we will discuss this point in a next paragraph.

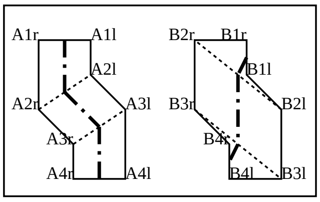

Figure 2: Central Ridge determination: on the left a valid central ridge because all the homologous angles are almost equal, and on the right a non valid central ridge because $(B_{1,r}\widehat{B_{2,r}}B_{3,r}) \neq (B_{1,l}\widehat{B_{2,l}}B_{3,l})$.

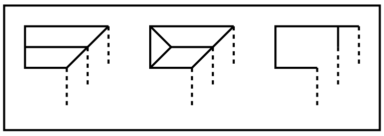

Figure 3: The three possible endings.

Then, for all possible central ridges (there could be one, or two in some particular cases), it is possible to modulate the two endings of the building to obtain a set of hypotheses that covers a large number of possible roof structures. For this modulation, we propose to use the three more frequent possible endings illustrated in Figure 3. Figure 4 shows how these simple rules produce a relevant set of building hypotheses.

As already mentioned, this algorithm cannot be used with a 2D polygonal shape with an odd number of sides, and, in general, infering in this way a roof structure from the 2D shape can often be disrupted by small details (little pediments or porch roofs for instance). To deal with these problems, it is possible to add a generalization step to suppress the less significant points on the initial 2D polygonal shape until a correct central ridge is found. Then the obtained roof tree is used with the complete initial 2D polygonal shape to build the 3D reconstruction of the building (Figure 5).

Figure 6 shows some results obtained with this automatic reconstruction of simple buildings. An interesting property of this approach is its ability to propose relevant and regular generalizations giving a nice-looking aspect to the resulting 3D city model. This method can not manage “T junction” buildings due to the limit of a simple central ridge. Presently, for such buildings, we prefer to use a completely different approach, rather than adapting this simple and pragmatic process.

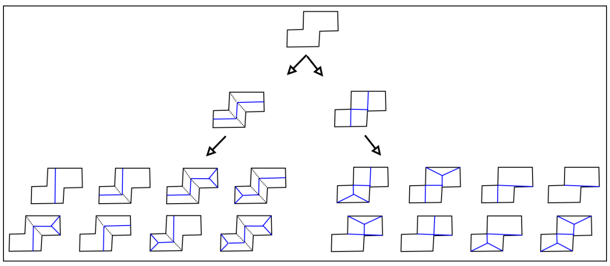

Figure 4: The set of roof hypotheses.

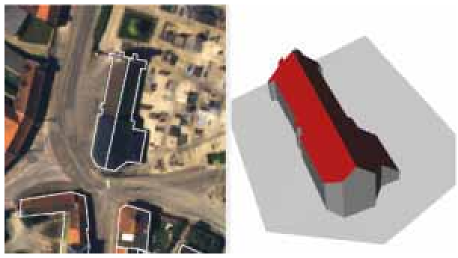

Figure 5: Model driven approach: Example of reconstruction obtained with a generalization step.

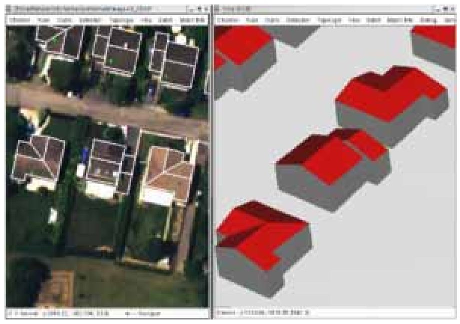

Figure 6: Examples of reconstruction in a suburb area with the model driven approach.

3.2 Data driven approach: using generic model reconstruction for complex buildings

We have selected the method proposed by (Jibrini, 2002). It also uses 2D digitalized cadastral maps with aerial images (a stereo pair or more). In this approach, the roof surface S is defined by four constraints:

- S is $2D^{\frac{1}{2}}$ surface $z = f(x, y)$;
- S is continuous;
- S covers the area of building limited by 2D cadastral polygon;
- and S is made of planes.

This definition is suitable for almost any forms of building roof. The reconstruction is based on three steps:

- Detection of planes using Hough transform (Jibrini et al., 2000);
- Heuristic pruning of plane hypotheses ;
- Search for an optimal polyhedral surface in the 3D graph obtained from planes intersection.

The 2D cadastral polygon is used both for the determination of the focusing areas and for the restriction of plane parameters. In the next paragraphs, we explain briefly these three steps. For more details, the reader may refer to (Jibrini, 2002).

3.2.1 Detection of planes The planes extraction happens in a discrete volume enclosing a building. Each voxel of the volume is valued with a weight depending on its 3D position. The weights measure the similarity of patches tacked around the 3D point projection in images using a template matching process. The plane hypotheses, corresponding to local high weight concentrations, are detected with a Hough transform. A plane is normally described by three parameters, so a three dimensions Hough accumulator is required. But the 2D cadastral polygon gives the orientation of the facades, and are used to constrain the normal of the expected planes. We suppose that the normal is perpendicular to one of the facades. This hypothesis is more general than the one introduced by (Brenner, 2000) because there is not a direct link between the facade and a roof facet (Figure 7). Using this property, the 3D accumulator can be transformed in several 2D accumulators: one for each facade orientation. The Hough transform is a well known and reliable process but it often gives too many hypotheses. The results obtained indicate that all the important planes and even some details like dormer windows are usually detected.

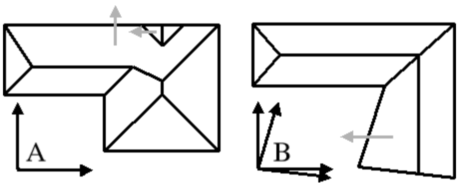

Figure 7: Each plane normal has to be perpendicular to one facade.

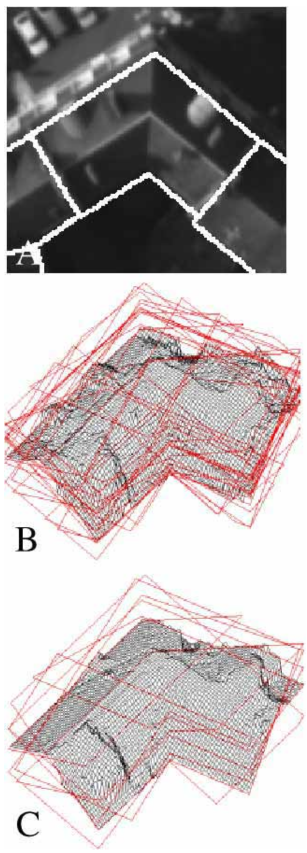

Figure 8: Planes pruning: A. the 2D cadastral polygon on the image, B. the 23 detected planes with the first DSM estimation, C. the 8 remained planes with this associated DSM.

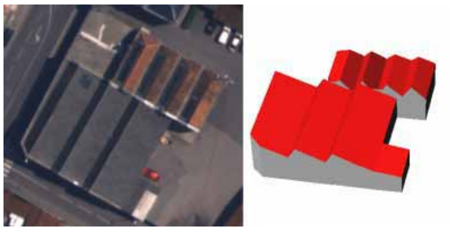

Figure 9: Data driven approach: an example on an industrial building.

3.2.2 Plane hypotheses pruning The resulting plane hypotheses set contains a lot of false detections. So, a pruning process is required to reduce combinatory and ambiguities of the surface reconstruction in the next stage. We aim at suppressing the most meaningless hypotheses without losing any of the important ones. The pruning process is iterative and based on the estimation of a DSM using the plane hypotheses. For each step, a more and more precise DSM is computed in a volume delimited by the plane hypotheses. And then, all the hypotheses are weighted in regards of their distance with this DSM, and the worst plane hypotheses are suppressed (Figure 8).

3.2.3 Polyhedral surface research First, a 3D graph is generated from the detected and pruned planes using an arrangement algorithm. The 3D graph is limited by the volume enclosing the building. The roof surface model defined above forms a planar graph whose external contour is attached to facades. Using the four characteristics of this model (defined in paragraph 3.2), we can transform the 3D graph to an assignment graph. A node of the assignment graph represents a facet in the 3D graph and an edge indicates that the two inherent nodes (two facets in the 3D graph) are compatible (they belong at least to one model, respectively). So, the search for the best roof surface model in the 3D graph boils down to a best maximal clique search problem in the assignment graph. A demonstration of this original idea and the simple method for the transformation between the two graphs are presented in (Jibrini, 2002). For the best maximal clique search (NP-Complete problem), we use an hybrid method integrating an exhaustive enumeration and an heuristic method. The optimization function takes into both image similarity and shape simplicity. Figure 9 shows an example of reconstruction on an industrial building. The needed time is about 1min per building (estimated on 5000 buildings on a 1GHz PC).

4 SEMIAUTOMATIC STAGE

Many failures in the automatic stage are due to the high complexity of some buildings. For instance a building may be too complex for a model driven approach, and have lots of microstructures like chimneys which disturb the data driven approach.

Moreover, 2D grounds map polygon edges are not always the real building facades. 2D ground maps we use have not been made for building reconstruction, they come from cadastral databases. The polygons geometry is based on tax criteria rather than physical ones. For instance, a little part of an adjacent building could be in the ground map polygon of the building to reconstruct. Another frequent problem is the grouping of buildings with parts of ground in a single polygon. Moreover, there are also incoherences and omissions due to updating delays. All these problems often make the automatic algorithms fail. Mainly because of these failures it is absolutely necessary to add some user-assisted tools to obtain a truly operational system.

4.1 Tools adapted to landscape characteristics

In each geographic area (at the scale of a country, a city and even a district) there is a particular way of building and therefore building constants. These characteristics depend on a lot of parameters like history, climate, soil, relief... These building constants lead to repetitive geometric rules in the 3D shapes. Finally a repetitive geometric rule enables to perform efficient measurements procedure. The huge differences between different urban landscapes are among most important reason for the complexity of the reconstruction problem. That is why an evolutional production line should be better than a static piece of software. In other words an analysis of the characteristics of the urban landscape and the development of specific semiautomatic measurement procedures can be an efficient way to proceed.

4.1.1 Reconstruction of symmetric buildings from the central ridge One of our most important datasets is on the city of Amiens in France (this dataset is freely available on http://isprs.ign.fr). As you see in Figure 11, lots of buildings in the old downtown have a symmetric structure around the central ridge. This specificity has been used to develop an efficient user-assisted method. The operator draws this central ridge in one image, and the system automatically proposes a reconstruction. This reconstruction uses a DSM to estimate the slopes, and the image to find the position of gutters.

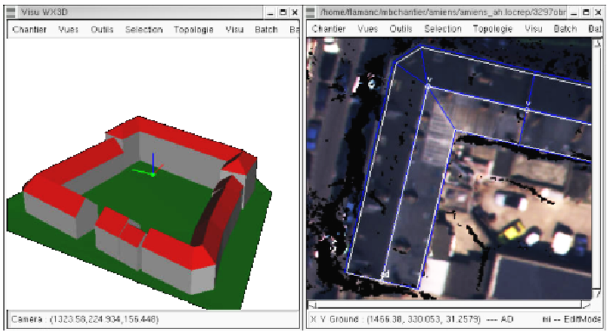

Figure 10: Symmetric buildings from the central ridge: an exemple of building reconstruction.

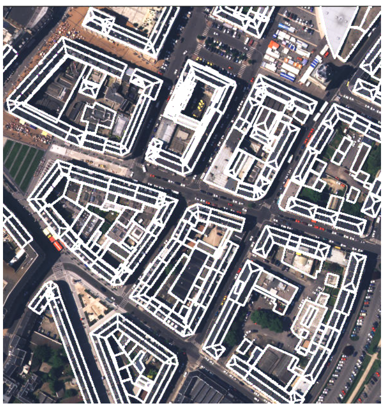

Figure 11: Symetric buildings in an old style downtown.

4.1.2 Copy/Paste This tool is very simple, but efficient in some recent suburbs where groups of buildings are built on the same model with the same sizes. The operator selects a previously reconstructed roof and uses it as a template model: then, he just needs one mouse click to instantiate it anywhere. This roof copy is automatically adjusted on all the available images, but can also be moved and rotated by the operator.

4.2 Focus plus selection of a model

This is the most popular type of interaction: an operator choses a roof model in a library and instantiates it with one or two mouse clicks on the images. In these approaches, the models are parametric (i.e. described with a small and fixed set of parameters). The major problem here is to have enough models to be able to reconstruct many buildings and, at the same time, to remain easy-to-use for the operator. One possibility to do that, is to reconstruct complex buildings by a combination of simple primitives with a CSG system (Gülch and Muller, 2001).

5 RESULTS

We have worked on an average french city of 160 000 inhabitants called Amiens in a multi-viewing context with aerial triangulated images. Ground pixel size is 25 cm. The first dataset (Figure 12) is a part of the dense urban center. There are 3000 buildings and the aera is $0.8km^2$. The global time cost is 1.5 days (7 hours per day). For the automatic stage, 50% of the buildings have been processed with the model driven approach, 25% with the data driven approach and the late 25% with a prismatic method.

The second dataset (Figure 13) is representative of a periurban context. There are a lot of non-adjacent suburb houses. There are 7000 buildings and the aera is $2.2km^2$. The global time cost is 3 days. Concerning the automatic stage 70% have been processed with the model driven approach, 10% with the data driven approach, 20% with a prismatic method.

For both methods 80% of the buildings have been processed in the automatic stage. We do not model structures as chimneys and dormer windows. Geometric precision of the produced 3D city models is closely linked to the cadastral maps precision. For these datasets, planimetric and altimetric estimated precision is < 1m.

We can divide the work in 3 main steps :

- Preparation including DSM and orthophoto processing but also cadastral maps reworking: 50% of the global time.
- Automatic processes launching and quality control on the results: 25% of the global time.
- Semi-automatic edition and final quality control: 25% of the global time.

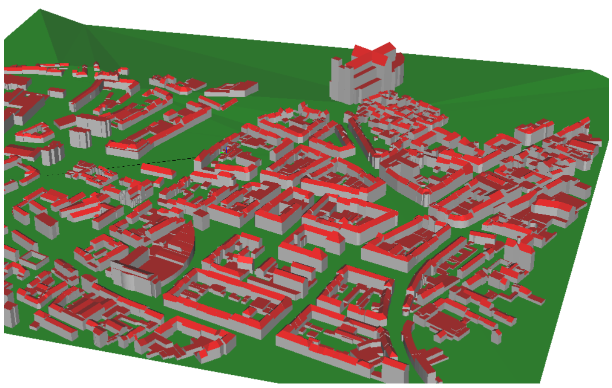

Figure 12: Results on a dense urban area.

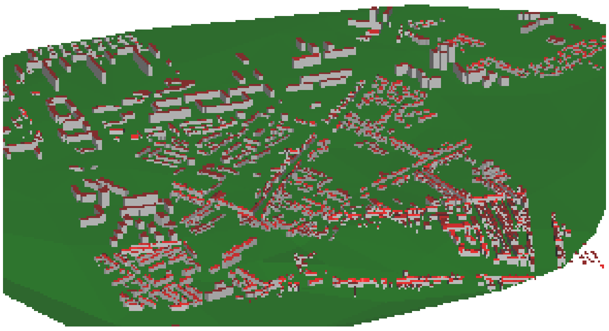

Figure 13: Results on a periurban area

6 CONCLUSION

We have presented a global strategy and an example of an operational buildings reconstruction framework for 3D city models production. In our further works we will work on an automatic cooperation between the model driven and the data driven approaches. We also have to improve our quality control tools to ensure that all the 3D models coming from the different algorithms respect an homogeneous level of details. In such a piece of software, the time cost estimation is not easy to evaluate, it is not a simple addition of all the time needed for each algorithm because it depends on the expected level of details and the operator workload to control, rework and manage the data. That is why, rather than giving an arbitrary cost, we will study the range of possible costs for different levels of data specifications.

REFERENCES

Baillard, C. and Maître, H., 1999. 3D reconstruction of urban scenes from aerial stereo imagery: a focusing strategy. Computer Vision and Image Understanding 13, pp. 244–258.

Brenner, C., 1999. Interactive modelling tools for 3d building reconstruction. In: F. . Spiller (ed.), Photogrammetric Week 99, pp. 23–34.

Brenner, C., 2000. Towards fully automatic generation of city models. In: IAPRS, Vol XXXIII, Part B3 (Amsterdam), pp. 85–92.

Brenner, C., 2001. City models - automation in research and practice. In: F. . Spiller (ed.), Photogrammetric Week 01, pp. 149–158.

Brenner, C., Haala, N. and Fritsch, D., 2001. Towards fully automated 3d city model generation. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona, pp. 47–57.

Fischer, A., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2d and 3d. In: CVIU, vol 72(2), pp. 185–203.

Fuchs, F., 2001. Building reconstruction in urban environment : a graph-based approach. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona.

Gruen, A. and Wang, X., 2001. News from cybercity-modeler. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona, pp. 93–101.

Gülch, E. and Muller, H., 2001. New applications of semi-automatic building acquisition. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona, pp. 103–114.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environements. In: ISPRS Journal of Photogrammetry and Remote Sensing Vol. 54, pp. 130–137.

Jibrini, H., 2002. Reconstruction automatique des bâtiments en modèles polyédriques 3D à partir de données cadastrales vectorisées 2D et d'un couple d'images aériennes à haute résolution. PhD thesis, ENST.

Jibrini, H., Paparoditis, N., Deseilligny, M. and Maître, H., 2000. Automatic building reconstruction from very high resolution aerial stereopairs using cadastral ground plans. In: IAPRS, XXXIII (Amsterdam).

Maas, H., 1999. Fast determination of parametric house models from dense airborne laserscanner data. In: ISPRS Workshop on Mobile Mapping Technology.

Nevatia, R. and Price, K., 2002. Automatic and interactive modeling of buildings in urban environnemts from aerial images. In: IEEE ICIP, pp. 525–528.

Paparoditis, N. and Maillet, G., 2001. Improving aerial image matching techniques in urban areas using a new multi-image approach guided from object space. In: 3rd International Workshop on Mobile Mapping Technology, Cairo.

Roy, S. and Cox, I., 1998. A maximum-flow formulation of the n-camera correspondence problem. In: ICCV.

Vosselman, G. and Suveg, I., 2001. Map based building reconstruction from laser data and images. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona, pp. 231–239.

Weidner, U., 1996. An approach to building extraction from digital surface models. In: IAPRS, Vol. 31, Part B3, pp. 924–929.

Zhang, Z., Zhang, J. and Hu, X., 2001. Semi-automatic building extraction from stereo images pairs. In: B. B. Verlag (ed.), Automatic Extraction of Man-Made objects from aerial and space images, Ascona.