AUTOMATIC 3D BUILDING RECONSTRUCTION USING PLANE-ROOF STRUCTURES

Babak Ameri, Dieter Fritsch
Institute for Photogrammetry (ifp), University of Stuttgart
Geschwister-Scholl-Strasse 24, 70174 Stuttgart, Germany
Babak.Ameri@ifp.uni-stuttgart.de, Dieter.Fritsch@ifp.uni-stuttgart.de

ABSTRACT

This paper introduces a new method for automatic 3D building reconstruction using plane-roof structures. A boundary representation (b-rep) of a coarse building hypothesis is constructed in a bottom-up approach from simple geometric primitives in image domain to more complex geometric primitives (roof structure) in object domain. Subsequently, a hypothesis model verification is performed in a top-down approach by back projecting the constructed model to the corresponding aerial images. A number of issues in an autonomous image analysis system are covered. The various processes and components essential to the proposed method are reported and the experimental results are also presented.

INTRODUCTION

Modeling and three-dimensional (3D) descriptions of real world objects collected through an imaging system (passive/active sensor) has become a topic of increasing importance as they are essential for a variety of applications. To name only a few: telecommunication for planning of wireless networks in cities (Siebe and Büning 1997), urban environmental planning and design to support the decision-making processes for development projects (Danahy 1999), virtual tourist information systems to support the on-line positioning, access and queries on the information of the site of interest (Volz and Klinec 1999), defense and military organization to support the training operation in virtual environment, environment and resource management, monitoring, and control for disaster preparedness, simulation of air pollution, and noise distribution.

Buildings are recognized to be the most prominent objects in generation of a 3D virtual model of our environment. Virtual reality and 3D visualization of the buildings are on the edge of changing the practice of urban environmental planning and design. Instead of presenting citizens abstract maps and descriptive text to explain, analyze and discuss design ideas and urban processes, planners will be able to show people explicit photo-textured information of what their city will look like after a proposed change (Danahy 1999). This is also confirmed by the result of a survey of the European Organization for Experimental Photogrammetric Research (OEEPE) on 3D city model generation. 95% of the participants have reported that buildings are the most interesting and important objects in a 3D geo-spatial information system (Fuchs et al. 1998). Consequently, a large number of research projects and efforts have been invested in the field of recognition, 3D reconstruction, and representation of building objects over the last few years (Förstner and Plümer 1997, Förstner et al. 1999, Greun et al. 1995, Gruen et al. 1997, CVIU 1998, ISPRS 1997, ISPRS 1999).

The acquisition and 3D representation of building objects includes not only the detection of buildings in the scene depicted by one or more images, but also the production of a scene description. This is a complex task consisting of different processes such as recognition, feature extraction, feature's attribute computation, grouping, structuring and geometric modeling, hypothesis generation, as well as hypothesis verification, which are assembled in a cautious manner. As a matter of fact, integration of all these processes in order to derive the required 3D information from a complex scene image(s), in a traditional way, human-based image interpretation system, would be a costly and labour intensive operation (Duperet et al. 1997). The laborious process of hand digitizing and interactively crafting each geometric model of buildings is too time consuming. Therefore, there is an increasing demand towards fully machine-based image interpretation systems. This is a difficult task for several reasons such as, the enormous variation in the structure and shape of the buildings, occlusion of buildings or building-parts by themselves or with neighboring adjacent objects, the effect of shadows, noise, low contrast, or small structures on the roof structure, etc.

In spite of current limitations mentioned above, techniques used in photogrammetry and computer vision are now sufficiently developed and have resulted in sophisticated systems and led to promising results for acquisition and reconstruction of GIS objects, in particular buildings. This paper introduces a new development of robust methods in a hierarchical framework for a data-driven reconstruction of generic plane-face building objects through the integration of computer vision and digital photogrammetric techniques. It is designed to handle buildings of different shapes and complexities. Exceptions are buildings with curve-like roof structure, such as dome roof. Thus, most of the geometrical regularity constraints imposed in the low- and intermediate-level reconstruction phases such as orthogonality, or parallelism, which are required in the specific model-based methods are not appropriate here. The reconstruction is based on the pragmatic assumption that building roofs are composed of generic plane-surfaces, so that a plane-face solid model, commonly called *polyhedral* can approximate a complex building and is used to support the reconstruction process. In addition, there is a strong coupling between 2D image and 3D object space in order to achieve reliable and precise results. The next section presents an overview of the whole framework. It gives a summary on the interrelated processing flow and concepts of our method for solving the task. The subsequent sections discuss the different processes involved in our proposed reconstruction method. In addition, experimental results are also presented.

GENERAL FRAMEWORK

Figure 1 schematically represents the workflow of the subsequent processes and the interrelation between the major components of the automated reconstruction process. The proposed components form a general framework, in which in each step different and more complex types of information are exploited. Conceptually, the entire spectrum of our work can be divided into three fundamental steps of *recognition, reconstruction*, and *hypothesis verification*. Although this subdivision has no definitive boundaries, it does provide a useful framework for categorizing and describing the various processes that are essential components of an autonomous image analysis system.

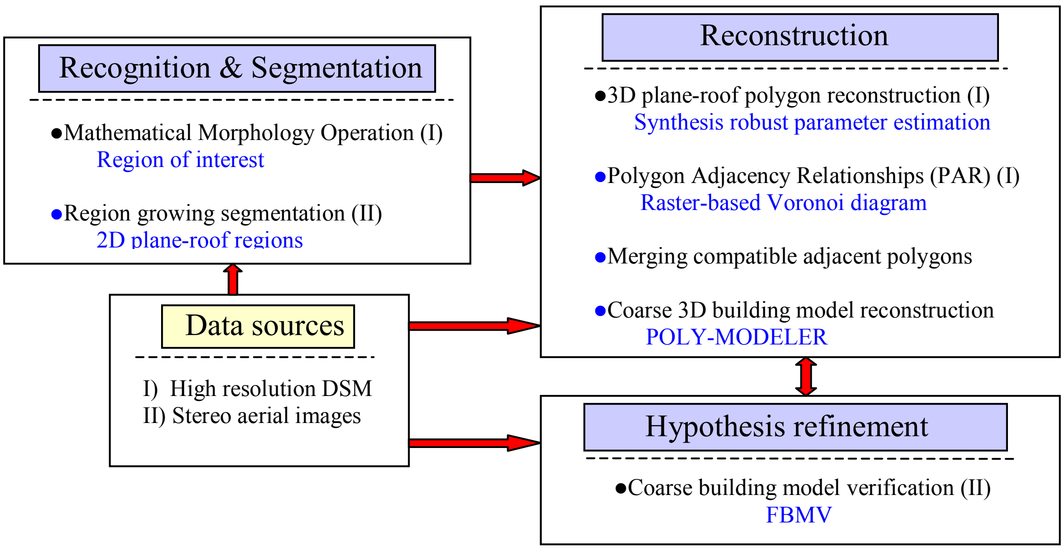


Figure 1: Proposed setup for automatic recognition and 3D reconstruction of building objects

The recognition part starts with a coarse segmentation of DSM in order to label areas (regions of interest) within aerial images, which have a high expectancy of representing individual buildings. This process is based on a morphological top-hat transformation. Furthermore, geometric characteristics of surfaces are used to extract flat-pixel surface type within detected regions of interest. The extracted pixels serve as the seed regions to a *least squares planar fit region growing algorithm* to partition the image surface into meaningful 2D regional primitives called plane-roof regions (Fritsch and Ameri, 1998). To move to the more model-oriented representation of the buildings, which is carried out in the reconstruction part, the intermediate extracted 2D plane-roof regions are projected back into the object space, called 3D plane-roof polygons. This is performed using a *synthesis robust parameter estimator* technique developed within the project. In order to describe the interrelation between these 3D geometric primitives, the *Polygon Adjacency Relationship (PAR)* is computed. The adjacency relationships are defined based on a Voronoi diagram. Based on the PAR, the compatible adjacent 3D polygons are *merged* into the larger 3D plane-roof primitive. These 3D planar structures together with their adjacency relationships information are input to the *POLY-MODELER*, where they are geometrically and/or topologically combined to generate the coarse generic polyhedral-like building models (Ameri and Fritsch, 1999). Here is the starting point for hypothesis verification in a top-down fashion. The reconstructed coarse building undergoes a refinement process based on the *Feature Based Model Verification (FBMV) concept*. Treating the generated coarse building hypothesis as evidence leads to a set of confidence intervals in image space that can be used as the search space to find the corresponding 2D image primitives and performing a consistency verification of the reconstructed coarse model (Ameri, 2000).

FROM PIXEL TO GEOMETRIC PRIMITIVES

Object recognition is one of the hardest problems in automated vision processes. Although the problem is addressed by a large number of researchers and projects (Haralick and Shapiro 1992), there is not a unique baseline methodology or a general paradigm to solve this complex task. This task is normally initialized with the image segmentation process, where the qualitative knowledge stored in the raw image data is transferred into quantitative symbolic description and more abstract form of the basic geometric elements. Image segmentation is frequently a data driven process. It can be based on homogeneity, namely, homogeneous regions that are detected in the image, or alternatively, discontinuities that can be used for the detection of edge primitives, which is assumed to correspond to the contours of the real objects. This section deals with recognition of 2D planar surfaces in the aerial image, *2D plane-roof regions*, which possess meaningful correspondence to object surfaces in 3D scene, *3D plane-roof polygons,* of the building roof structures. The recognition performs in three subsequent segmentation processes. First, the image is partitioned into the regions of interest based on height discontinuities and their size. Furthermore, a coarse segmentation of the image is carried out based on the geometric characteristics of surfaces to extract flat pixel surfaces type within every region of interest. Finally, the fine segmentation of the image is performed based on a least squares planar fit region growing algorithm to partition the image surface into meaningful primitive plane-roof regions (Fritsch and Ameri 1998).

Building Detection

The overall aim of the detection process is to partition an image into regions, which have potential to be buildings. This primary segmentation is carried out to, 1) reduce the dimensionality of the search space and consequently reducing the computational time required in subsequent processes, and 2) to be a step closer to our ultimate aim. In the absence of GIS information or ground plan of the buildings, the corresponding DSM is used as an initial source for building detection. However, the proposed method is general enough to integrate the contribution of the existing data in this step, if any. The quality of the DSM is an important issue in our reconstruction method. Indeed, the results of the mid-level processes in extraction and structuring 3D primitives is highly dependent to the quality of the utilized DSM, which in the worst case leads to partially or completely wrong descriptions of the buildings. Figure 2-a shows a 3D perspective view of an image wrapped over corresponding DSM. The figure illustrates the presence of the standing objects such as buildings and trees -as it is expected- in DSM.

Several methods for detecting building candidates, *regions of interest*, from the DSM have been reported so far. In this study we have used the grey opening approach proposed by Weidner and Förstner (1995), which is based on a square structuring element.

Figure 2-b illustrates a perspective view of the extracted regions of interest in 3D object space. Some of the regions (totally or partly) still do not present buildings or building parts. These are due to presence of features such as standing trees or cars next to a building, or group of trees whose size is bigger than a specified size threshold. These mis-classified regions or region parts will be filtered out in the follow up processes. An alternative solution is the analysis of texture pattern (Haralick et al. 1973, Nagao and Matsuyama 1980, Lee and Schenk 1998), within every region of interest, in order to detect and exclude objects that are standing adjacent to the buildings but are not

a building or part of a building, such as trees. The intermediate result has shown that integration of this analysis at this stage possibly could overcome this problem, but still more investigation is needed. The transformation of the extracted regions of interest into the 2D image space is done based on the well-known collinearity equations. Figure 2-c shows the result of this transformation.

a b c

Figure 2: Extraction of Regions of Interest, a) the 3D perspective view of an image wrapped over corresponding DSM, b) the 3D perspective view of the extracted regions of interest, c) extracted regions overlaid on the correponding aerial image.

2D Plane-Roof Regions

The objective of segmentation in this work is to partition an image into regions. In the previous section, we approached this problem by finding regions of interest based on height discontinuities and their size using mathematical morphology. In this section, we discuss a region-based segmentation technique based on an iterative region growing approach to partition each region of interest into the primitive planar regions, which are parts of building roofs in real world. A common approach to region segmentation is to start from some pixels (seeds) representing distinct image regions and to grow them, until they cover the entire image. The performance of this algorithm depends heavily on the choice of the initial seeds. Ideally, one seed per image region must be provided. The user in a supervised mode usually chooses the seeds. But, in order to implement a region growing segmentation algorithm that can be executed in an automated and unsupervised environment, a rule describing and extracting seed regions based on a data-driven mechanism is needed. To realize the concept, a strategy based on geometric characteristics of a digital surface has been developed. This approach is originally proposed by Besl and Jain (1988), where the signs of *mean* and *Gaussian curvatures* have been used to initially classify range images in industrial application into eight different surface types. Moreover, an iterative region-growing algorithm based on variable-order surface fitting has been utilized to partition a range image into smooth, and meaningful surface regions. The essential difference of our approach compared to Besl is that, based on the assumption that building roofs are composed of generic planar surfaces, we only concentrate on *flat surface type pixels,* which serve as the seed regions for the region growing segmentation process. In addition, the segmentation processes are performed within the extracted regions of interest in aerial image based upon intensity gray values, not the height data in range image. Figure 3-a shows the extracted flat surface type pixels (seed regions) overlaid on corresponding roof structures.

To illustrate the basic concept, let us start with the extracted seed regions. First, a plane is fitted to the small seed region based on a least squares planar fitting process. If the seed region belongs to part of the roof that is not too highly curved, this plane will fit quite well to the original data. If the plane fits the seed region within the maximum allowable fit error threshold, then the seed is allowed to grow, if not, the seed is rejected. The growing process continues until the termination criteria are met. A detailed description of the proposed method is given in (Fritsch and Ameri 1998).

Figure 3-b illustrates the results of extracted 2D plane-roof regions overlaid on corresponding roof structures. It shows that main structures of the roof are correctly extracted. However because of the presence of the noise in the image and due to the shadow caused by the microstructure on top of the roof structure some of the larger roof primitives might be divided into the smaller primitives. In fact, these intermediate 2D regions are merged into the

larger one, if they satisfy the compatibility requirements during the reconstruction process, which is discussed in the following sections. In the presence of high-quality, high-resolution digital images, the proposed method is also capable to detect and extract the microstructure on top of the buildings roof e.g., dormer windows. This type of information improves significantly the results of the higher-level reconstruction processes, in particular when dealing with complex buildings.

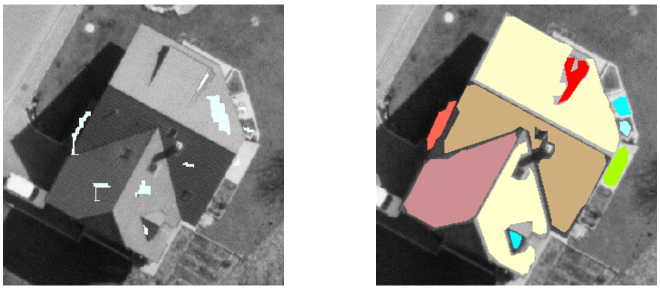


Figure 3: Region growing segmentation, a) extracted seed regions, and b) extracted 2-D plane-roof regions overlaid on corresponding building roof structure.

HYPOTHESIS BUILDING MODEL GENERATION

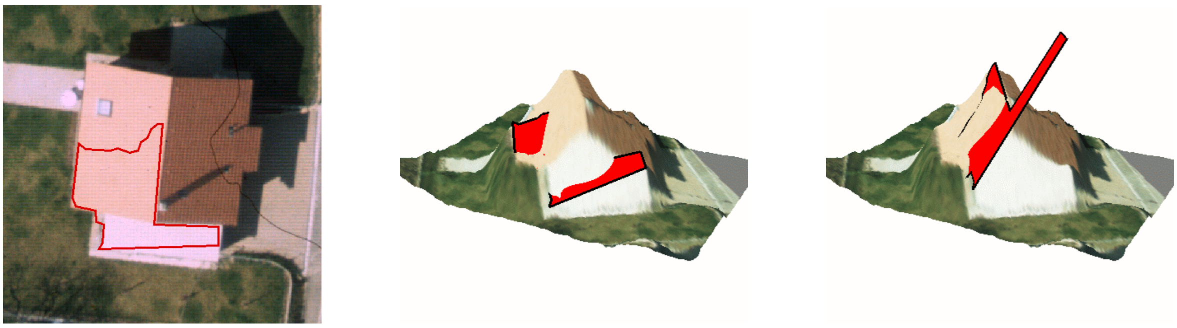
This section describes a new method for automatic 3D reconstruction of polyhedral-like objects, in this context used as a generic building model. A boundary representation of a coarse building hypothesis is constructed in a data-driven, bottom-up approach, from simple geometric primitives (2D plane-roof regions) in image domain to more complex geometric model (3D-roof structure) in object domain. The reconstruction part consists of different intermediate, interrelated processes aiming to form a framework, in such a way that every process provides more abstract and more object related information to its immediate higher level process. The proposed methods and the mid-level vision processes are discussed and the subsequent results of different processes are presented in this section.

3D Plane-Roof Polygons

In order to integrate the power of 3D geometry into the reconstruction process, the detected primary 2D plane-roof regions are translated into the corresponding 3D plane-roof polygons first. Having had an initial descriptions or approximation of the object surface in real world, i.e., in this study the corresponding DSM, the transformation process is equivalent to a 3D regression problem. Theoretically, the best result is obtained based on the traditional least squares fit. However, in practice, due to the presence of outliers in the original data, the solution is far more complex than the simple fitting process and an appropriate robust fitting procedure is required. In this application, outliers occurred in both DSM and extracted primary 2D regions. Outliers appearing in the extracted 2D regions are caused by the failure of the segmentation procedure. As it is shown in Figure 4-a, due to the presence of noise, shadow or in this particular case, low contrast, the region growing segmentation process sometimes grows over the discontinuities or the physical bounding edge of the region. This mis-segmented region part(s) appears as outliers during the fitting process, as its real height is significantly lower than the height of the corresponding building roof. This results in an arbitrary solution, if outliers are not detected and excluded from the estimation process (see Figure 4-b). Outliers arise in DSM namely during its generation. Commonly there are two different comparable methods for generation of DSM, automatic photogrammetric techniques, and direct laser scanning methods. Both methods have advantages and disadvantages, a complete comparison with respect to various aspects of both techniques is given by (Baltsavias 1999). Despite the fact that the laser scanning techniques provide a highly accurate direct geometric description of the visible surfaces, however a practical result in the particular application of automated building reconstruction has shown that the quality of the DSM in built-up areas generated by either techniques is still insufficient. For example, issues like occlusion, shadow and anomalies of the surface height and discontinuities in photogrammetric methods, and lack of explicit measurement of breaklines such as roof ridges in laser scanning techniques are the reasons that both methods failed to accurately recover the descriptions of the roof structures. In fact, the latter method is capable of partially overcoming this problem by highly dense sampling measurements of the terrain surface, but it asks for very expensive and costly operations.

The effects of the presented outliers are eliminated here in two steps. The extreme ouliers, which lead to an arbitrary plane parameter for the 3D plane-roof polygons are detected and excluded during the regression procedure using a synthesis robust parameter estimation developed newly in this work. The reminder of the outliers that cause a minor deviation between the estimated roof structure parameters and the physical ones are eliminated during the verification of hypothesis roof structure, which is discussed in the next section.

The proposed robust parameter estimation method is a two-stage parameter estimation algorithm. The first stage flushes and detects the outliers and estimates the best initial 3D plane parameters based on the inliers data using a random sampling type estimator such as *Random Sampling Consesus RANSAC* (Fischler and Bolles 1981), or alternatively *Least Median Squares (LMS)* (Rousseeuw and Leroy 1987). The estimated parameter values along with the estimated error variance are then introduced into the iterative re-weighting M-estimator algorithm (Huber 1981, Hampel et al. 1986) as initial values to compute the final 3D plane parameters.



a b c

Figure 4: Parameters estimation of 3D plane-roof polygon, a) 2D plane-roof region overlaid on corresponding roof structure, b) corresponding 3D plane-roof polygon back projected into the 3D object space based on a standard LS estimation process, and c) the synthesis robust parameter estimation technique.

Figure 4 illustrates the improvement afforded using the proposed synthesis robust parameter estimation method in a 3D-regression problem. Assuming that the low contrast white area in the lower part of the building causes the segmentation algorithm to grow over the bounding edge of the roof, and therefore incorrectly extract this area as part of the 2D plane-roof region. The data points belonging to this part of the roof appear as outliers during back projection of the extracted 2D plane-roof region into the 3D object space and should be detected and eliminated from the 3D-regression process. The Figure 4-b and 4-c illustrate the 3D perspective views of the estimated 3D plane obtained based on the ordinary LS estimation, and our proposed method respectively. Figure 4-b graphically indicates the failure of the LS procedure in estimating the parameters of the 3D plane, which is forced by the contaminated data points, while the proposed approach correctly recovered the parameters of the 3D plane.

Polygons Adjacency Relationships (PAR)

So far, we discussed processes of construction for 3D primitive plane-roof polygons, which are the main components of the roof structure and form the foundation on which we build the generic building models. To be able to topologically describe the interrelation between these 3D primitives, the *Polygon Adjacency Relationship (PAR)* is computed. During the reconstruction process, the PAR provides the essential topological information such as adjacency, and 'contained-in' relationships between incorporated primitives, which are the minimum types of object relationships that are required in an automated vision process based on a generic object model. The PAR is defined based on a Voronoi diagram (dual of Delaunay triangulation), where each primitive plane-roof polygon, in this context a data point, produce a zone of influence representing all parts of the space closer to that polygon than to any other. Polygons are considered adjacent only if their Voronoi regions touch. In fact, the main reason for using Voronoi regions for solving the problem is that no model of spatial adjacency is available for disconnected objects. Hence the definition of adjacency had to await the connection of the points, line segments or polygons in the form of a graph structure by techniques that are primarily coordinate-based line intersection detection methods (Gold 1990). A Voronoi region can be constructed around any objects or geometric primitives, which in turn gives the ability to construct the adjacency relationships between more complex object types than the simple point-wise data set. In the application of computing the adjacency relationships between plane-roof polygons, due to the presence of polygons with different sizes, closeness, and polygons which are partially or totally overlapping each other such as 'contained-in' polygons, applying this method in vector domain causes undesired results. In order to overcome this problem, the

Voronoi diagram is generated based on *distance transformation* in raster domain using Chamfer 3-4 mask (Borgefors 1986). Pilouk et al. (1994) and Chen et al. (1994) have extended the concept into the 3D space for generating Delaunay tetrahedral tessellation.

The analysis of the result obtained by applying different geometric primitives and features as kernel points draw the important fact that the shape and size of the utilized features has an impact on the result of the adjacency relationships significantly. Therefore, the concept of spatial adjacency, which has been normally defined based on a point-wise data set, is extended by introducing the adjacency relationships between polygonal primitives of different shapes and sizes, including connected, disconnected, or overlapped ones. The result of PAR for the extracted 2D plane-roof polygons within a region of interest is illustrated in Figure 5. The generated Voronoi diagram (see Figure 5-b) represents the adjacency between polygonal primitives (Figure 5-a), in particular, adjacency relationships of the 'contained-in' polygons such as polygons '74' and '73' are correctly defined. The PAR is stored as complementary properties of each polygon primitives and updated in proceeding processes as required.

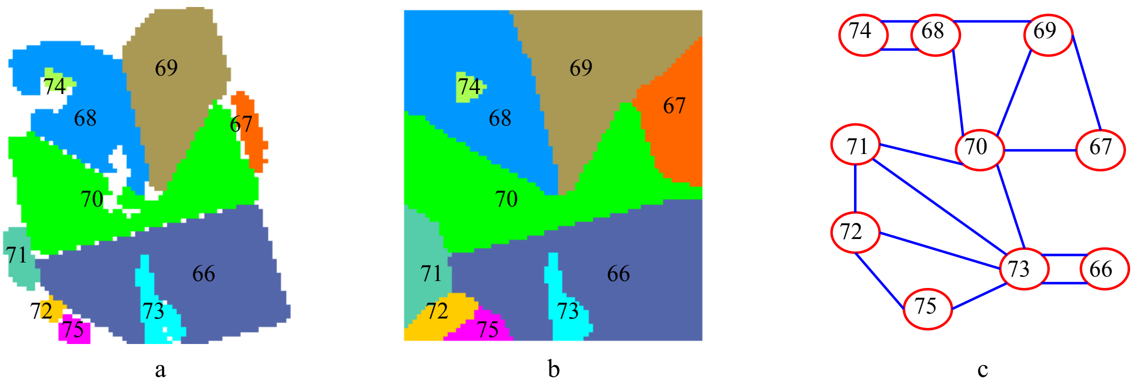


Figure 5: Computation of polygon adjacency relationships (PAR), a) initial kernel polygons, b) generated Voronoi diagram, c) computed adjacency graph

It was also discussed previously that due to the presence of noise, shadow or occlusions caused by the microstructures on top of the roof structure some of the larger roof primitives are divided into the smaller primitives. Therefore, a merging procedure is performed if the primitives satisfy the compatibility requirements. The compatibility rules are defined based on the descriptions of the planar polygons in 3D space, 1) *their positions*, i.e., adjacency, and 2) *orientations*, i.e., surface normal. The merging is allowed, if and only if the two 3D plane-roof polygons are adjacent, and have approximately the same orientations in space.

Roof Structuring and Geometric Modeling

A boundary representation (b-rep) scheme is used to describe a generic building model as the union of very general faces embedded in unbounded plane-surfaces, where the building edges are defined by the intersections of these surfaces. Such generic models can be constructed directly by assembling and intersecting appropriate surfaces. An algorithm called *POLY-MODELER* performs the reconstruction. The algorithm determines where component faces are extended or truncated and new edges and vertices are created or deleted. When boundary elements overlap or coincide, the algorithm merges them into a single element and thus maintains a consistent, non-redundant data structure representing the building model boundary. New edges are created where adjacent faces (polygons) intersect. The POLY-MODELER finds these intersections and then determines by *point membership classification*, which segments of the intersection are actual edges of the model.

The objective of a b-rep modeler such as POLY-MODELER is to build a complete representation of a solid as an organized collection of surfaces. In general a b-rep model stores the numerical data of the surface geometry on which the face lies, the curve geometry on which the edge lies and bounds the face, and the point geometry of the vertices. In fact, the POLY-MODELER obviously handles a special case of boundary representation when curved surfaces and edges are approximated by planes and straight lines. There is a minor deviation in how POLY-MODELER works in special application of generic building reconstruction from its original design strategy. Since the vertical walls of the buildings are perpendicular to the ground and are simply reconstructed based on the outline of the building roof, the POLY-MODELER only concentrates on reconstruction of roof structure. In the final step the fictitious vertical walls are added to complete the reconstruction of a plane-face solid building model.

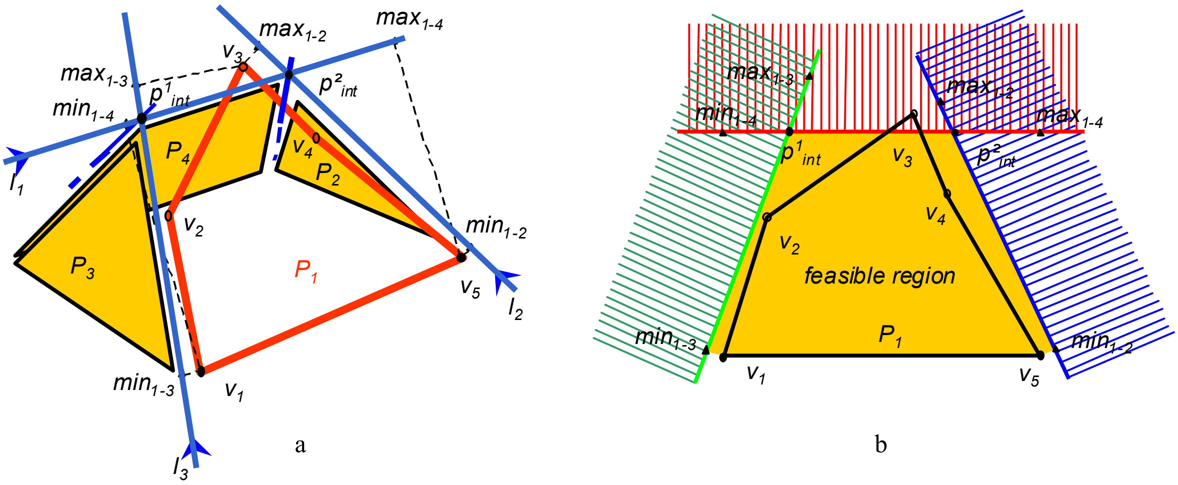


Figure 6: Geometric roof modeling, a) 3D intersection of adjacent plane-roof polygons, b) determination of the feasible regions (extension of the plane surfaces of the roof model).

Figure 6-a, illustrates an example of roof modeling based on the proposed method. Let 3D plane-roof polygons (3D-polys) p_1, p_2, p_3, and p_4 indicate the plane-surfaces of a building roof structure. Furthermore assume that 3D straight lines l_1, l_2, and l_3 are computed based on the intersection of 3D-poly p_1, with its adjacent polygons p_4, p_2, and p_3, respectively. The associated extreme points min_{1-2}, max_{1-2}, min_{1-3}, max_{1-3}, min_{1-4}, max_{1-4}, and the 3D intersection point p^1_{int}, and p^2_{int}, are also defined accordingly and are introduced into the POLY-MODELER as the new candidate vertices of the 3D-poly p_1 (Ameri and Fritsch 1999). For the simplicity, only the associated geometric primitives of the polygon p_1 are shown in the figure. To proceed, first, all the 3D primitives are transferred into the 2D space based on a orthogonal projection. Figure 6-b shows a top view of the corresponding elements in 2D space. Every intersection line divides the space E^2 into two half-planes. An inequality condition derived from the line parameters defines the valid and invalid (hachured area) half-planes. The simultaneous intersection of the valid half-planes determines the extension or feasible region of the 3D-poly p_1, which is indicated as grey area in the Figure 6-b. A complete description of the concept and how the POLY-MODELER works is given in (Ameri and Fritsch 1999).

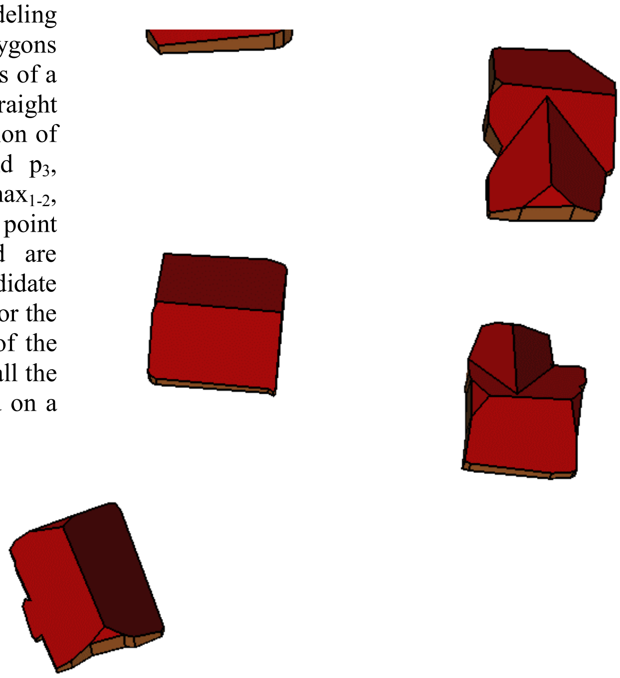


Figure 7: The 3D perspective view of the reconstructed 3D coarse building hypotheses.

Figure 7 illustrates the result of the proposed method and the performance of the POLY-MODELER algorithm. Four buildings with different roof structures, a gable, a hipped-gable, and two complex roof structures, are selected. For each building candidate or region of interest, the geometric descriptions of all 3D polygonal primitives of the roof and their adjacency relationships are introduced into the POLY-MODELER as initial parameters. The POLY-MODELER, then generates a very dense internal pointing data structure to keep track of the changes of all the primitives during reconstruction process. It models the roof structure in a generic manner purely based on the 3D intersection of adjacent polygons, without any a priori information concerning the roof type or imposing any external constraints. That means the same procedure is applied to any form of roof structure, with any number of polygonal primitives and complexity. In this example, the upper-right complex building has interesting properties, since it simply invalids many of the constraints applied in *specific model-based* approaches. For example, constraints on orthogonality of the building outline cause that reported methods fail to reconstruct a correct and accurate model of this type of generic building.

As discussed earlier, the result of the reconstruction process is highly related to the estimated parameters of the 3D plane-roof polygons. Owing to the geometrical reconstruction of roof structure, positional accuracy of roof elements such as orientation edges and intersection points are very high. However, due to a misinterpretation of surface normal of polygonal primitives, we may have some discrepancies in the form of displacement or rotation from the real positions of these elements. A failure in correctly recovering the surface normal of the roof polygons will cause an unexpected result leading to a partially or completely wrong building description. This is why the quality of utilized DSM is of high importance in our approach. In addition, due to the nature of the region-growing type segmentation, the quality of roof outline of the building model is still poor. The region growing type segmentation methods are unable to accurately localize the bounding edges of the roof structure. This is the reason why the generated hypothesis model is called *coarse building model*. To improve the quality of the generated model, the geometric and topological information derived from the coarse model is incorporated into the hypothesis model verification process, which is discussed in the next section.

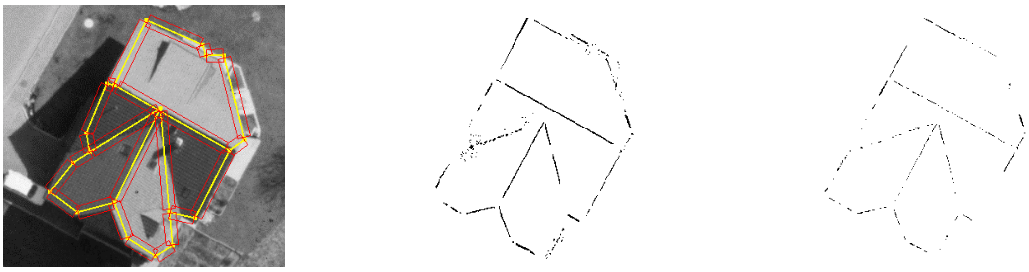
HYPOTHESIS VERIFICATION

In (Ameri 2000) the author introduces a new concept called *Feature Based Model Verification (FBMV),* for verification and validation of a coarse polyhedral-like object model. The process is carried out by back projecting the 3D coarse model into the corresponding 2D images. Although this transformation is a non-linear operation it is a smooth and well-behaved transformation, and it is a promising candidate for the application of the well known Gauss-Markov estimation model based upon an iterative least squares minimization error criterion. This method requires the appropriate initial guess for the unknown parameters. These values are provided by the geometric and topological information derived from the reconstructed coarse model (see Figure 7).

A fine building model is obtained in an iterative, model-driven combined least squares adjustment process by simultaneously fitting the 3D model into the corresponding images where the geometrical and topological model information are integrated into the process as external and/or internal constraints during the estimation. The ability to apply such constraints is essential for the accurate modeling of complex objects. In particular, when dealing with a generic object model, it is crucial that the model elements are both accurate and consistent with each other. For example, individual components of a building can be modeled independently, but to ensure realism, one must guarantee that they touch each other in an architectural way. The estimation procedure yields a description of the building that simultaneously satisfies all the constraints within all the images. As a result, it allows us to perform a consistency check and refinement of the model across all the images. Moreover the ability of the estimation method to fuse the information over all the images increases the accuracy and reliability of the reconstruction.

The estimation process is an orthogonal linear least squares regression problem. It acts as a functional model with the objective to simultaneously minimize the perpendicular sum of the Euclidean distances between the candidate edge-pixels and the projected 2D model edges in all the images. A set of constraints as weighted observation equations e.g., *Connectivity, Collinearity, Coplanarity, and Orthogonality*, is also integrated into the estimation model to support the minimization process (Ameri 2000).

The projected 2D edge in the images serves as initial guess for finding the exact position of the edge model within the corresponding images. An uncertainty buffer with a user specified width is generated around each edge model in image space based on the initial position of the edge and is used as the search space to find the representative edge-pixels. Figure 8-a illustrates the generated buffer around the 2D model edges of a reconstructed coarse building model in one of the corresponding images. Figure 8-b shows the selected edge-pixels within the generated uncertainty buffer in the first iteration of the estimation process.



a b c

Figure 8: Extraction of candidate edge-pixels, a) Uncertainty buffer of model edges, b) selected candidate edge-pixels in first, and c) last iteration of estimation process.

To make the selection process robust and to impose a self-diagnosis mechanism, the generated buffer is updated in a regular interval during the iteration process. As the process is iterated, the initial parameters of the 2D edges are updated based on the minimization of the orthogonal distance error between the selected edge-pixels and their respective 2D edges. Consequently, the updated parameters define a new orientation for the generated buffer. In addition, introducing a smaller width reduces the size of the buffer. As a consequence, as outliers are excluded from the estimation process, this procedure reduces the computational burden, increases the accuracy and speeds up the convergence of the process. Figure 8-c indicates the selected edge-pixels of the corresponding 2D edges in Figure 8-b for the last iteration of the estimation process.

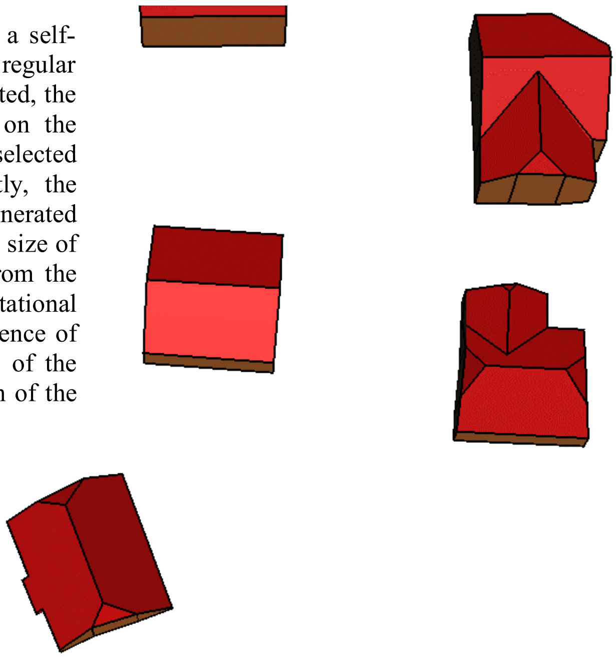


Figure 9: The 3D perspective view of the final 3D reconstructed buildings.

The performance and the outcome of the FBMV algorithm applied to the generated coarse buildings in the previous example is illustrated in Figure 9. The result indicates that FBMV method precisely verified the coarse model and the redundant edge segments are also removed. The other key issue of the FBMV method is its ability to provide the essential tools for evaluation of the quality of the reconstructed model and its geometric primitives. The estimated variances of the unknown parameters, specifically in our case the coordinates of the model points in 3D space are the qualitative measures, which indicate the accuracy of the model primitives and act as the decision criteria in order to reject or accept the estimated model elements. The evaluation process can be integrated into the whole chain of the reconstruction process as an edition process (Traffic Light concept: Förstner 1996). In a simple manner, these measures give a hint to the end user to perform a visual check on the end product and perform the required modifications on the signalized model primitives, if necessary.

DISCUSSION

The paper covered a number of issues in computer vision and photogrammetry, particularly recognition and 3D reconstruction of plane-roof structure building objects. To conclude the paper and represent the overall performance of the system, let us consider the result of the proposed methods applied to the residential test images of the international Avenches data set (Mason et al. 1994). Figure 10 shows the perspective view of the final reconstructed buildings wrapped over the existing DSM in 3D object space. Figure 11 depicts the result overlaid on the corresponding aerial image. All the buildings are reconstructed. Even the main parts of the roof structure of building no. 4, which was under construction during aerial photography, are reconstructed. This demonstrates the robustness of the reconstruction process. In fact, the generated coarse model is not accurate enough to be verified automatically applying FBMV, however the modification process is capable of sending a warning signal to the operator in order to check the final result and, if necessary, edit the model manually. Also, the result shows that the details of all the buildings, except building No. 1, is precisely recovered. The roof part of this building has not been detected completely during segmentation process. This is due to the occlusion of the roof by adjacent tree. Generating multiple coarse hypothesis models using multiple overlapping aerial images can significantly eliminate the problem of occlusion. In this research the process of coarse reconstruction of hypothesis building is performed using only one single image. The results improve significantly if the hypothesis model generation is performed in all the available corresponding images in a parallel process, thus providing multiple hypothesis coarse models for every building object. The coarse building hypotheses can merge or fuse together in 3D object space to generate a unique and more reliable coarse building hypothesis for final verification process. In this manner, it is feasible to recover parts of the building roof structure, which are missing or have not been reconstructed in one or more of the generated coarse models because of e.g., occlusion, shadow or noise during segmentation process. It should be noted that although the FBMV method is capable of recovering most parts of the coarse model as it is also working based upon multiple images. However if the initial position of the geometric primitive(s) of the coarse building is far away from its real position, the FBMV will fail to modify that part(s) of the building structure.

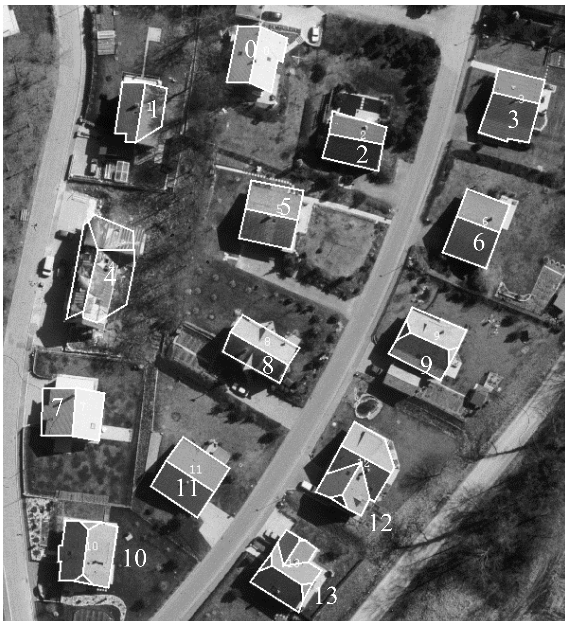


Figure 11: Final reconstructed buildings overlaid on the corresponding aerial image

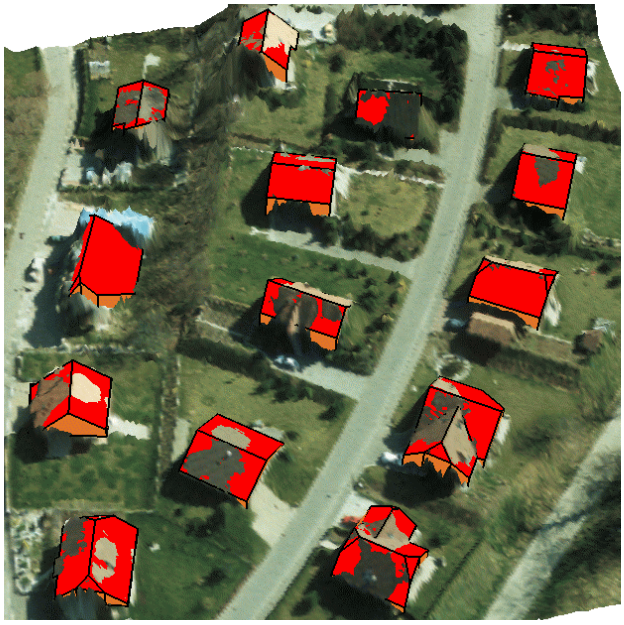


Figure 10: Perspective view of the final 3D reconstructed buildings wrapped over corresponding DSM

REFERENCES

Ameri, B. (2000). Feature-Based Model Verification (FBMV): a new concept for hypothesis validation in building Reconstruction. In: *IAPRS*, Vol. XIX, Part 3, in press, Amsterdam, the Netherlands.

Ameri, B. and Fritsch, D. (1999). 3-D Reconstruction of Polyhedral-Like Building Models. In: *IAPRS*, Vol. 32, Part 3-2W5, Munich, Germany, pp. 15-20.

Baltsavias, E. (1999). A Comparison between Photogrammetry and Laser Scanning. *ISPRS Journal*, 54(2/3):83-94.

Besl, P. J. and Jain, R. C. (1988). Segmentation Through Variable Order Surface Fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(2):167-192.

Borgefors, G. (1986). Distance Transformation in Digital Images, *Comp. Vision, Geraphics, and Image Processing*. 34:344-371.

Chen, X., Ikeda, K., Yamakita, K. and Nasu, M. (1994). Raster Algorithms for Generating Delaunay Tetrahedral Tessellations. In: *IAPRS*, Vol. 30, Part 3, Munich, Germany, pp. 124-131.

CVIU, (1998). Special Issue on Building Extraction. *Computer Vision and Image Understanding*, 72(2).

Danahy, J. (1999). Visualization Data Needs in Urban Environmental Planning and Design. In: D. Fritsch and R. Spiller (eds), *Photogrammetric Week '99*, Herbert Wichmann Verlag, Heidelberg, pp. 351-365.

Duperet, A., Eidenbenz C., and Holland, D. (1997). Capturing and Maintaining Datasets using Digital Imagery: Experience and Future Requirements of National Mapping Agencies. In: *IAPRS*, Vol. 32, Part 3-4W, Stuttgart, Germany, pp. 51-67.

Fischler, M. A. and Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model fitting with application to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381-395.

Förstner, W. (1996). 10 Pros and Cons Against Performance Characterization of Vision Algorithms. In: Workshop on Performance Characterization of Vision Algorithms, Cambridge, UK.

Förstner, W. and Plüumer (eds) (1997). *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*. Birkhüauser Verlag, Bonn, Germany.

Förstner, W. and Liedtke, C. E. and Bückner, J. (eds) (1999*). Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*. Institute of Photogrammetry, University of Bonn, Germany.

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

Fritsch, D. and Ameri, B. (1998). Geometric Characteristics of Digital Surfaces: A key Towards 3D Building Reconstruction. In: *IAPRS,* Vol. 32, part 3/1, pp. 119-126, Columbus, USA.

Fuchs, C., Gülch, E. and Förstner, W. (1998). OEEPE Survey on 3D-City Models, OEEPE Publication No. 35, Bundesamt für Kartographie und Geodäsie, Frankfurt.

Gruen, A., Kuebler, O. and Agouris, P. (eds) (1995). *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkhäuser Verlag, Basel, Ascona, Switzerland.

Gruen, A., Baltsavias, E, and Henricsson (eds) (1997). *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*. Birkhäuser Verlag, Basel, Ascona, Switzerland.

Gold, C. M. (1990). Neighbours, Adjacency and Theft - the Voronoi Process for Spatial Analysis. In: First European Conference on Geographic Information Systems, Amsterdam, the Netherlands, pp. 382-391.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics*. John Wiley and Sons, New York, USA.

Haralick, R. Shanmugam, K. and Dinstein, I. (1973). Textural Features for Image Classification*. IEEE Transactions* on System, Man, and Cybernetics, SMC(3):610-621.

Haralick, R. and Shapiro, L. (1992). *Computer and Robot Vision*. Addison-Wesley Publishing Company, USA.

Huber, P. J. (1981). *Robust Statistics* .John Wiley and Sons, New York, USA.

ISPRS (1997). ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects. *IAPRS,* Vol. 32, Part 3-4W2, Stuttgart, Germany.

ISPRS (1999). ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, *IAPRS,* Vol. 32, Part 3-2W5, Munich, Germany

Lee, D. and Schenk, T. (1998). An Adaptive Approach for Extracting Information and Segmentation. IAPRS, Vol. 32, Part 3/1, Ohio, USA, pp. 250-255.

Mason, S., Baltsavias, M. and Stallmann, D. (1994). High Precision Photogrammetric Data Set for Building Reconstruction and Terrain Modelling. Institute for Geodesy and Photogrammetry, ETH Zurich.

Nagao, M. and Matsuyama, T. (1980). *A Structural Analysis of Complex Aerial*. Plenum Press, New York, USA.

Pilouk, M., Tempfli, K. and Molenaar, M. (1994). A Tetrahedron-based 3D Vector Data Model for Geoinformation. In: Advanced geographic data modeling: spatial data modeling and query languages for 2D and 3D applications, Delft, The Netherlands, pp. 129-140.

Rousseeuw, P. and Leroy, A. (1987*). Robust Regression and Outlier Detection*. John Wiley and Sons, New York.

Siebe, E. and Büning, U. (1997). Application of Digital Photogrammetric Products for Cellular Radio Network Planning. In: D. Fritsch and D. Hobbie (eds), *Photogrammetric Week '97*, Herbert Wichmann Verlag, Heidelberg, pp. 159-164.

Volz, S. and Klineck, D. (1999). NEXUS: The Development of a Platform Location Aware Application. In: Third Turkish -German joint Geodetic Days, Vol. 2, Istanbul, Turkey, pp. 599-608.

Weidner, U. and Förstner, W. (1995). Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal*, 50(4):38-49.