

# A NEW METHOD FOR BUILDING EXTRACTION IN URBAN AREAS FROM HIGH-RESOLUTION LIDAR DATA

F. Rottensteiner\*, Ch. Briese

Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Gußhausstraße 27-29,  
A-1040 Vienna, Austria – (fr,cb)@ipf.tuwien.ac.at

Commission III, WG III/3

**KEY WORDS:** Building extraction, automation, 3D building models, segmentation, laser scanning

## ABSTRACT:

In this paper, a new method for the automated generation of 3D building models from directly observed point clouds generated by LIDAR sensors is presented. By a hierarchic application of robust interpolation using a skew error distribution function, the LIDAR points being on the terrain are separated from points on buildings and other object classes, and a digital terrain model (DTM) can be computed. Points on buildings have to be separated from other points classified as off-terrain points, which is accomplished by an analysis of the height differences of a digital surface model passing through the original LIDAR points and a digital terrain model. Thus, a building mask is derived, and polyhedral building models are created in these candidate regions in a bottom-up procedure by applying curvature-based segmentation techniques. Intermediate results will be presented for a test site located in the City of Vienna.

## 1. INTRODUCTION

### 1.1 Motivation and Goals

Automation in data acquisition for 3D city models is an important topic of research with the goal of reducing the costs of providing these data at an appropriate level of detail. In addition to photogrammetric techniques relying on aerial images, the generation of 3D building models from point clouds provided by LIDAR sensors is gaining importance. This development has been triggered by the progress in sensor technology which has rendered possible the acquisition of very dense point clouds using airborne laser scanners. Using LIDAR data with point densities of up to one point per square meter, it is possible not only to detect buildings and their approximate outlines, but also to extract planar roof faces and, thus, to create models which correctly resemble the roof structures.

Building extraction is solved in two steps (Brenner, 2000). First, buildings have to be detected in the data, and the approximate building outlines have to be determined. Second, in the regions of interest thus detected, the buildings have to be reconstructed geometrically, which results in 3D polyhedral models of the buildings. It is the goal of this paper to present a new method for the automatic creation of polyhedral building models in densely built-up areas from high-resolution LIDAR data without using ground plans. Our method is unique with respect to the algorithms used for building detection because it is based on robust interpolation. In the detected building regions, planar roof patches, their bounding polygons, and their neighborhood relations are extracted. Grouping of neighboring planes has not yet been implemented. The examples presented in this paper were computed using the LIDAR data from a test site in the City of Vienna captured by TopoSys. The resolution of the original point cloud is 0.1 m (in-flight) by 1 m (cross-flight). A grid of 0.5 x 0.5 m<sup>2</sup> derived from that point cloud was used for building extraction. The test data were captured in the

course of a pilot project for the Municipality of Vienna in order to evaluate and compare various techniques for the generation of 3D city models. Our intermediate results show the high potential of the method presented in this paper.

### 1.2 Related Work

There have been several attempts to detect buildings in LIDAR data in the past. The task has been solved by classifying the LIDAR points according to whether they belong to the terrain, to buildings or to other object classes, e.g., vegetation. Morphological opening filters or rank filters are commonly used to determine a digital terrain model (DTM) which is subtracted from the digital surface model (DSM). By applying height thresholds to the normalized DSM thus created, an initial building mask is obtained (Weidner, 1997; Ameri, 2000). The initial classification has to be improved in order to remove vegetation areas. In (Brunn and Weidner, 1997), this is accomplished by a framework for combining various shape cues in a Bayesian network. Our algorithm for building detection from LIDAR points is based on the method for DTM generation by robust interpolation presented in (Kraus and Pfeifer, 1998).

The geometrical reconstruction of the buildings in previously detected regions of interest has been tackled in two ways. First, parametric primitives can be instantiated and fit to the data if sufficient evidence is found. Second, planar patches can be detected in a DSM created from the LIDAR points, and polyhedral building models can be derived by grouping these planar patches. As parametric primitives often have a rectangular footprint, they are especially used if 2D ground plans giving a precise location of the building outlines are available. The polygon delineating a building in a 2D map is split into rectangular regions. In each rectangle, the parameters of parametric models are determined using the DSM, and the model achieving the best fit is accepted (Brenner, 2000; Vosselman and Dijkman, 2001). The data driven generation of

---

\* Corresponding author.

polyhedral building models from LIDAR data only makes sense if the point density is high enough so that a sufficient number of data points is located at least in the most relevant planes of the roofs. As the building outlines are difficult to be located precisely, again ground plans are often used for that purpose. Ground plans also reduce search space for the estimation of the parameters of adjoining planar patches because the gradient direction of such planes is usually perpendicular to the adjacent polygon segment in the ground plan (Haala et al., 1998; Brenner, 2000; Vosselman and Dijkman, 2001). Initial planar patches are found by a segmentation of the DSM. Brenner (2000) gives several methods for DSM segmentation, e.g., the analysis of surface curvature, i.e., of changes in the surface normal vectors, or a segmentation taking into account the directions of the polygon segments of a ground plan.

As soon as the initial planar patches have been found, neighboring patches are grouped (Baillard et al., 1999), and the polygons delineating the borders of planar patches have to be found. The latter task involves finding consistent intersections at the building vertices (Moons et al., 1998). Finally, the 3D border polygons have to be combined in order to obtain consistent building models. At the building outlines, vertical walls, and, finally, the floor have to be added to the model. A coarse-to-fine strategy can be applied by first searching for the most relevant structures in the data and using refined methods for modeling the buildings in regions not being “explained” sufficiently by the initial models (Vosselman and Dijkman, 2001). The problem of precisely determining the building outlines has been tackled by Weidner (1997) by applying the minimum description length principle for deciding on regularizations.

## 2. WORK FLOW FOR BUILDING EXTRACTION

The work flow for the extraction of buildings from LIDAR points is presented in figure 1. The first step is the interpolation of a DSM and a DTM from the original data at an appropriate resolution. Our method for DTM generation which performs a classification of the original points into terrain versus off-terrain points by robust estimation will be explained in section 3. From this instance onwards, the models created by interpolation are used, no longer the original data points.

By subtracting the DTM from the DSM and by applying a threshold to the height differences, an initial building mask is created which still contains vegetation and other objects. Binary morphological operators and an analysis of the DSM texture, i.e., of the local variations of the DSM normal vectors, are used to eliminate these areas. The final results of building detection, i.e., the individual building regions, are found by a connected component analysis. Building detection by comparing the DTM and the DSM is described in section 4.

In the building candidate regions, a plane segmentation based on an analysis of the variations of the DSM normal vectors is applied to find planar patches. These patches are expanded by region growing algorithms. In the current version, the neighborhood relations of these patches are determined, and a simple model resembling the roof structure of the building is created. In the future, neighboring planes will be grouped consistently before the initial building models containing the most relevant roof structures are created. In a post-processing phase, the model has to be refined in order to contain details originally not detected. The current state of our technique for

geometrical reconstruction of roof structures from a DSM is described in section 5.

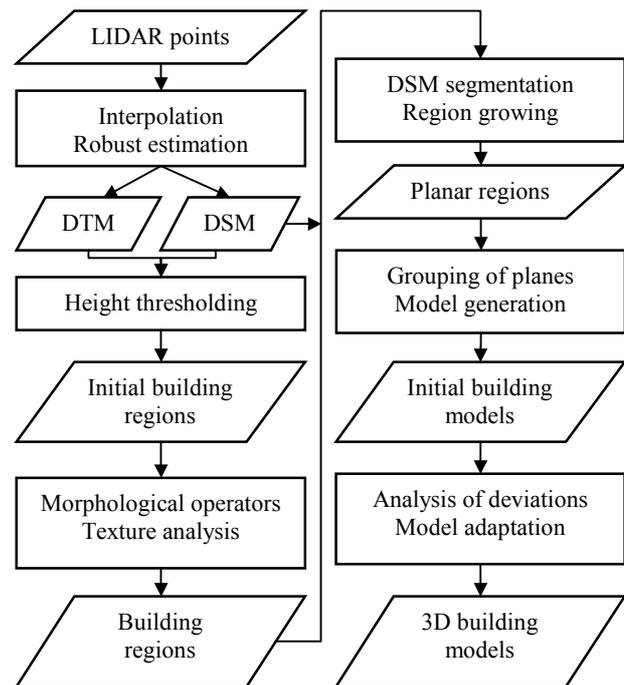


Figure 1. Work flow for building extraction from LIDAR data.

## 3. DTM GENERATION FROM LIDAR DATA IN URBAN REGIONS

An algorithm for the automatic generation of DTMs in forested regions from laser scanner data was developed at our institute. This method is based on iterative robust interpolation of a DTM grid, and it combines the elimination of off-terrain points and the interpolation of the DTM grid in one process (Kraus and Pfeifer, 1998). For the generation of a DTM in densely built-up areas, this method has to be modified to work in a hierarchical framework (Pfeifer et al., 2001). With this coarse-to-fine approach it is possible to cope with relatively large areas without terrain points (e.g., large building blocks).

### 3.1 Robust Interpolation

In an iterative process the irregularly distributed LIDAR points are weighted in a way that the modeled surface describes the terrain. The classification of the points in terrain versus off-terrain points is performed by thresholding the discrepancies to the computed surface by user-specified tolerance values.

In a first step, a coarse approximation of the surface is computed taking into account all available LIDAR points. Next, the discrepancies, i.e., the differences of the heights of the LIDAR points and the interpolated surface at the planimetric positions of the LIDAR points, are computed. The discrepancies are the parameters of a weight function which is used to assign an individual weight to each point in the subsequent processes. The interpolation of the DTM is repeated, the weights of the LIDAR points being modulated depending on the discrepancies of the most recent iteration. This iterative process is terminated as soon as a stable situation or a maximum number of iterations is reached. Two types of models are used in our algorithm, i.e., the functional model which defines the way the surface is

computed, and the stochastic model which is responsible for weighting.

**The functional model:** Linear prediction is used for modeling the surface. Using this model, it is possible to compute a smooth surface considering random measurement errors (Kraus, 2000).

**The stochastic model:** For the generation of a DTM, high weights must be assigned to terrain points below or on the averaging surface, and low weights have to be assigned to the non-terrain points which are above the averaging surface. A typical weight function  $p(r)$  parameterized by the discrepancies  $r$  for the generation of a DTM from laser scanner data is presented in Figure 2. The weight function we use is not symmetrical, and it is shifted by a value  $g$ . It has a sharp decline defined by its half-width value  $h$  and slant  $s$  for discrepancies greater than its central point (i.e., for off-terrain points above the estimated surface) and no decline for the terrain points. The exclusion of points from the interpolation process is triggered by a threshold  $t$  derived from a user-specified tolerance for the size of the discrepancies. For a comprehensive description of this algorithm see (Kraus and Pfeifer, 1998).

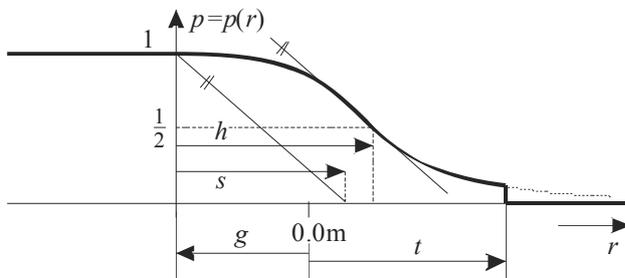


Figure 2. Weight function for the generation of a DTM from laser scanner data.

### 3.2 Hierarchic Robust Interpolation

The method of iterative robust interpolation relies on a “good mixture” of terrain and off-terrain points. Therefore, this algorithm does not work in large areas without terrain points as they are likely to exist in densely built-up areas. To provide this “good mixture” also in densely built-up areas, robust interpolation has to be applied in a hierarchic way using data pyramids (comparable to image pyramids in image processing). The hierarchic robust interpolation proceeds as follows:

1. Create the data pyramids. This can be achieved by selecting, for instance, the lowest points in a regular grid mesh.
2. Perform robust interpolation to generate a DTM.
3. Compare the DTM to the data of the next higher resolution and accept points within a certain tolerance band.

Steps 2 and 3 are repeated at each resolution level of the data pyramid. The results of DTM interpolation in the lower resolution levels are used for the computation of the surface in the next higher resolution because only points having passed the thresholding step 3 are considered at that level.

In (Briese, 2001), this strategy has been evaluated for the generation of a high-quality DTM of a test site located in the City of Vienna (2.5 km<sup>2</sup>) using three data pyramid levels (5 m, 2 m and 0,5 m). A few intermediate results of this DTM generation process are presented in the perspective views in Figures 3a-d.

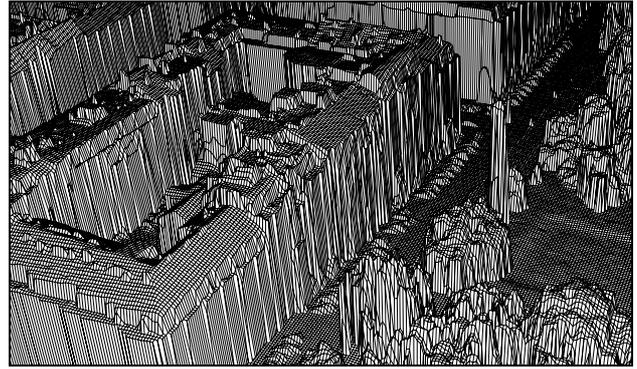


Figure 3a. DSM computed from all LIDAR points (0.5 m).

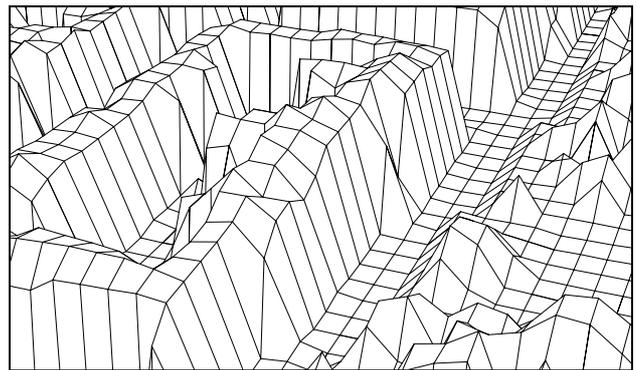


Figure 3b. DSM at the lowest resolution (5 m).

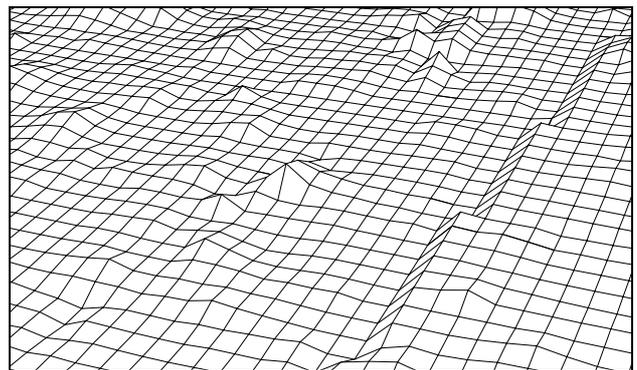


Figure 3c. DTM at the lowest resolution (5 m).

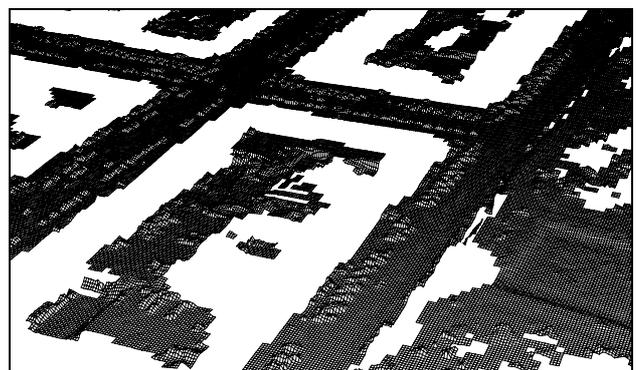


Figure 3d. DSM (0.5 m) of the accepted original points (the points within a user-defined tolerance band).

Figure 4 shows a perspective view of a detail of the final DTM. Further details about hierarchical robust interpolation, its implementation in the software package SCOP, and the results of some further examples can be found in (Pfeifer et al., 2001).

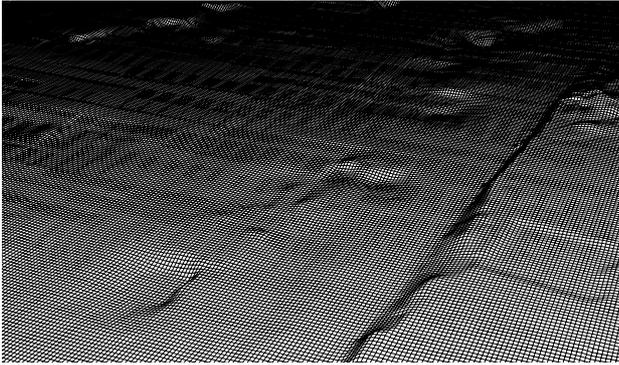


Figure 4. Final DTM (0.5 m; here: 1.0 m for better readability)

#### 4. BUILDING DETECTION

As described in section 3, two digital elevation models are derived by interpolation: a DTM is computed from the points classified as “terrain points” with a high degree of smoothing, whereas a DSM is computed from all points without smoothing (Figure 5a). An initial building mask is created by thresholding the height differences between the DSM and the DTM (e.g., by  $\Delta h_{\min}=3.5\text{m}$ ). This initial building mask still contains areas covered by vegetation, and some individual building blocks are not correctly separated (Figure 5b). A morphological opening filter using a small (e.g.,  $5 \times 5$ ) square structural element is applied to the initial building mask in order to erase small

elongated objects such as fences and to separate regions just bridged by a thin line of pixels. A connected component analysis of the resulting image is applied to obtain the initial building regions. At this instance, regions smaller than a minimum area (e.g.,  $40 \text{ m}^2$ ) and regions at the border of the DSM are discarded (Figure 5c).

Some of the remaining regions in Figure 5c still correspond to groups of trees. These regions can be eliminated by evaluating a “terrain roughness” criterion derived by an analysis of the second derivatives of the DSM. In (Fuchs, 1998), a method for polymorphic feature extraction is described which aims at a classification of texture as being homogeneous, linear, or point-like, by an analysis of the first derivatives of a digital image. This method is applied to the first derivatives of the DSM using a large (e.g.,  $9 \times 9$ ) integration kernel. For each initial building region, the number of “point-like” pixels is counted. Regions containing more than 50% of pixels classified as being “point-like” (thus, pixels being in a neighborhood of great, but anisotropic variations of the surface normals) are very likely to contain vegetation rather than buildings, and they are eliminated. Figure 5d shows the results of texture classification. Note the obvious co-incidence of clusters of “point-like” pixels displayed in black and vegetation areas such as those in the Belvedere gardens on the left margin of the test site.

The terrain roughness criterion is very efficient in classifying isolated vegetation regions, but it cannot find vegetation areas

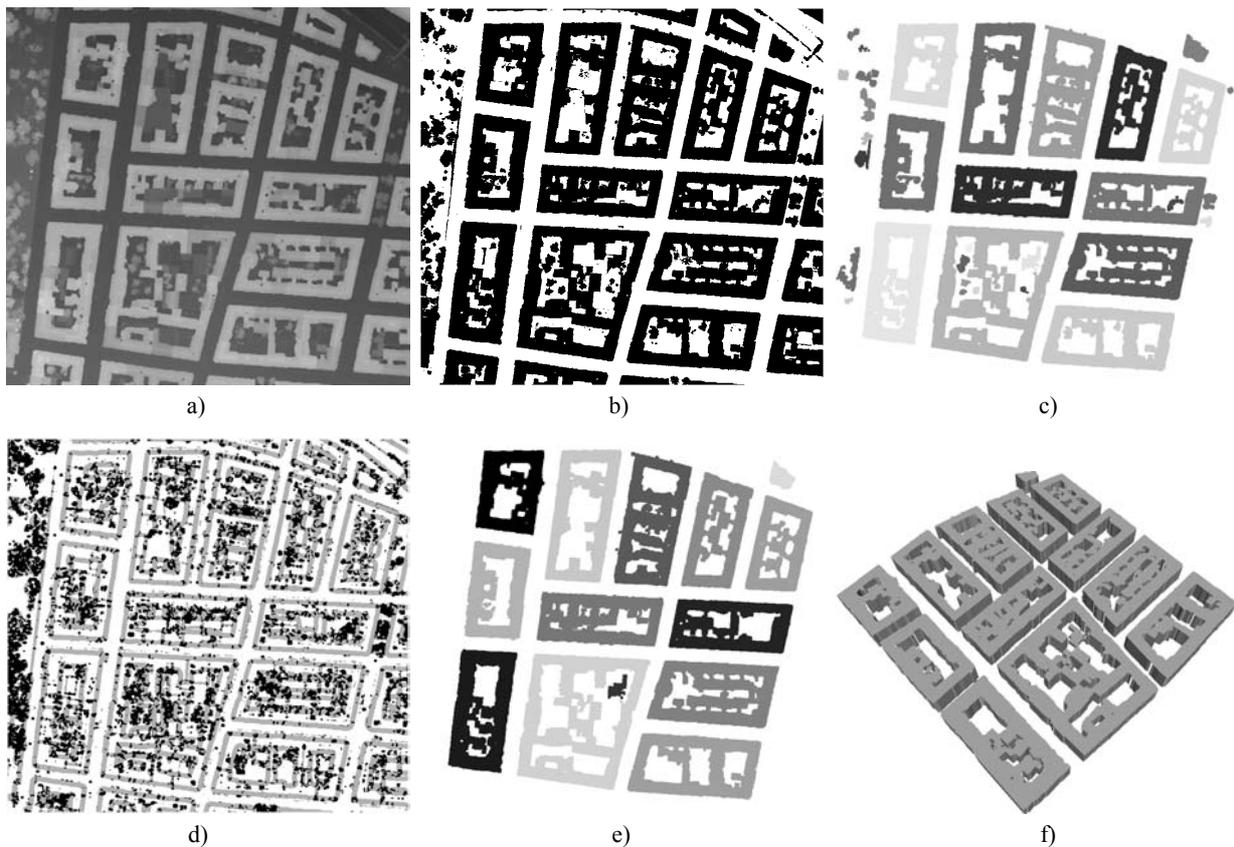


Figure 5. Building detection in a test site in the City of Vienna. Original resolution: 0.1 m (in-flight) by 1.0 m (cross-flight).  
 a) DSM; grid width:  $0.5 \times 0.5 \text{ m}^2$ ; extent:  $410 \times 435 \text{ m}^2$ . b) Initial building mask (height threshold  $\Delta h_{\min}=3.5\text{m}$ ).  
 c) Initial building label image before evaluating terrain roughness. d) Results of texture classification. Integration kernel:  $9 \times 9$  pixels. White: homogeneous; gray: linear; black: point-like. e) Final building label image. 14 building regions have been detected.  
 f) VRML visualization of prismatic models created from the boundary polygons of the building regions from Figure 5e.

which are still connected to buildings. In a final stage of analysis, we try to eliminate such areas. By morphological opening using a square structural element, regions just connected by small bridges are separated. The resulting binary image is analyzed by a connected component analysis which results in a greater number of regions, and the terrain roughness criterion is evaluated again. Pixels being in regions now classified as containing vegetation are erased in the initial building label image. Thus, in vegetation areas originally connected to buildings, only the border pixels remain classified as “building pixels”. Again, morphological opening helps to erase these border pixels. The resulting building label image only contains a small percentage of erroneously classified pixels in some backyards (Figure 5e).

At a very coarse level of detail, a 3D city model can be derived by creating prismatic models from the boundary polygons of the building regions using the average building heights computed from the DSM. An example for such a city model with a height accuracy of about  $\pm 5$  m is shown in Figure 5f.

## 5. GEOMETRICAL RECONSTRUCTION OF BUILDINGS

### 5.1 Generation of initial 3D planar segments

To start with model generation, initial 3D planar segments, their geometrical parameters, and their initial border polygons have to be found in the regions of interest. This is achieved by generating a “segment label image” defined in object space with an appropriate grid width. Each pixel of that image is assigned the label of the planar segment it belongs to.

The framework for polymorphic feature extraction (Fuchs, 1998) is applied for the generation of planar segments, too. Just as described in section 4, the framework is applied to the first derivatives of the DSM, this time using a small integration kernel of  $3 \times 3$  pixels. Pixels classified as being homogeneous are surrounded by pixels having similar components of the normal vector, i.e., they are in a region containing co-planar points (Brunn and Weidner, 1997). The binary image of the homogeneous pixels is used for further processing (Figure 6a). By applying a connected component analysis to this binary image, planar patches should be detectable. However, due to classification errors, especially at the intersections of roof planes which are almost horizontal, the regions thus detected often turn out to be too large. Typically, this leads to L-shaped segments such as the one at the upper left corner of Figure 6a. In order to avoid these segmentation errors, an iterative strategy is applied for the generation of planar patches:

1. The binary image of homogeneous pixels is morphologically opened using a square structural element before applying the connected component analysis.
2. The geometric parameters of the planar patches thus detected are derived along with their r.m.s. errors from all points inside these patches.
3. The patches with the best fit, i.e., those with r.m.s. errors better than a certain threshold (e.g.,  $\pm 10$  cm) are considered to be seed regions for region growing. These seed regions are grown iteratively by adding neighboring pixels to a region if their distances from the original adjusting plane are below a certain threshold. In this way, the most relevant and best fitting planes are extracted from the DSM.

4. The plane parameters are updated, and the pixels already being assigned to a planar patch are erased in the binary image. The connected component analysis is repeated, and the parameters of the new planar patches are evaluated.

Steps 1 to 4 are repeated with a decreasing size of the structural element for morphological opening. Thus, smaller and smaller initial regions are found, and by only allowing well-fitting planes to grow, it is possible to split the regions corresponding to more than one roof plane, because the r.m.s. error of the planar fit is a good indicator for the occurrence of such situations. Figure 6b shows the planar patches extracted in one of the building regions from Figure 5e.



Figure 6a. Pixels classified as “homogeneous” (white) for one of the building regions in Figure 5e.

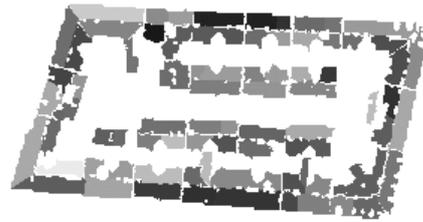


Figure 6b. Planar patches obtained by iteratively applying region growing.

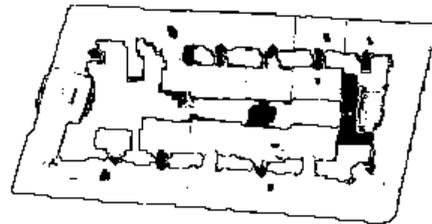


Figure 6c. Pixels not being consistent to a planar patch.

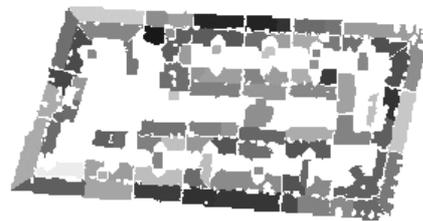


Figure 6d. Final segment label image for one of the building regions in Figure 5e.

A further analysis has to detect planes which cover too small an area for resulting in pixels classified as being homogeneous. We search for regions not being consistent with the planar regions detected so far (Figure 6c). The borders of the buildings are typically found in that process, which is caused by laser points on the walls and by the effects of grid interpolation. Again, we get rid of these points by a morphological opening operation using a  $3 \times 3$  square structural element, and a connected

component analysis is applied to the resulting image in order to create additional candidates for planar patches. Figure 6d shows the final segment label image created for one of the building regions from Figure 5e. The r.m.s. errors of planar adjustment varies between  $\pm 5$  cm and  $\pm 15$  cm for the segments corresponding to the “homogeneous” points. The segments having a r.m.s. error larger than  $\pm 10$  cm possibly still correspond to more than one roof plane. In the planar regions created by the analysis of the originally inconsistent points, the r.m.s. errors vary between  $\pm 25$  cm and  $\pm 5$  m. Some of these regions correspond to trees, and other regions still correspond to more than one roof plane. In the future, a further analysis will split these regions into smaller ones corresponding to even smaller planes in object space. This can be accomplished, e.g., by a height segmentation of the DSM in these regions. Table 1 shows the distribution of the r.m.s. errors of the planar fit.

r.m.s. error [m]	Regions	Pixels [%]
0.00 - 0.05	241	30.6
0.05 - 0.10	333	44.4
0.10 - 0.15	96	9.3
0.15 - 0.20	133	8.8
0.20 - 0.50	26	0.9
0.50 - 1.00	10	0.2
1.00 - 2.00	14	0.5
2.00 - 3.00	42	2.6
3.00 - 4.00	36	1.8
$\geq 4.00$	15	0.9

Table 1. Distribution of the r.m.s. errors of the planar fit. Regions: number of planar regions in the respective range of r.m.s. errors. Pixels: percentage of pixels in these regions compared to the number of all pixels in all planar regions. 68% of the pixels in the building candidate regions are classified as belonging to a planar region.

## 5.2 Grouping planar segments to create polyhedral models

To derive the neighborhood relations of the planar segments, a Voronoi diagram based on a distance transformation of the segment label image has to be created (Ameri, 2000): each pixel inside the region of interest not yet assigned to a planar segment is assigned to the nearest segment. The distances of pixels from the nearest segment are computed by using a 3-4 chamfer mask. Figure 7 shows a Voronoi diagram of the segment label image from Figure 6d. From the Voronoi diagram, the neighborhood relations of the planar segments are derived, and the borders of the Voronoi regions can be extracted as the first estimates for the border polygons of the planar segments (Figure 8).

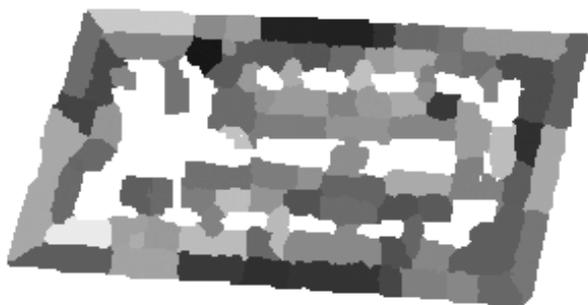


Figure 7. A Voronoi diagram of the label image in Figure 6d.

After deriving the neighborhood relations, neighboring planar segments have to be grouped. There are three possibilities for the relations of two neighboring planes (Baillard et al., 1999). First, they might be co-planar, which is found out by a statistical test applied to the plane parameters. In this case, they have to be merged. Second, two neighboring planes might intersect consistently, which is the case if the intersection line is close to the initial boundary. In this case, the intersection line has to be computed, and both region boundaries have to be updated to contain the intersection line. Third, if the planes do not intersect in a consistent way, there is a step edge, and a vertical wall has to be inserted at the border of these segments. After grouping neighboring planes, the bounding polygons of all enhanced planar regions have to be completed. (Moons et al., 1998) give a method for doing so and for regularizing the shape of these polygons at building corners. Finally, the planar polygons have to be combined to form a polyhedral model, and vertical walls as well as a floor have to be added to the model.

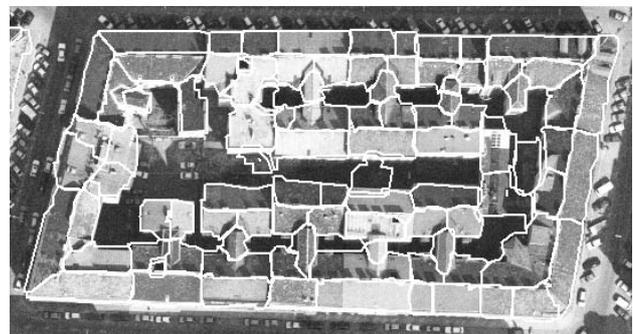


Figure 8. The roof polygons of the building in Figures 6 and 7 back-projected to an aerial image.

The tools for grouping planes and for computing intersections and the positions of step edges have not yet been implemented. Figure 9 shows a VRML visualization of a 3D model created from intersecting vertical prisms bounded by the borders of the Voronoi regions with the respective 3D roof planes. The structure of the roofs is correctly resembled, but the intersection lines of neighboring roof planes are not yet computed correctly. However, the visualization shows the high potential of the method for generating roof planes from LIDAR data.

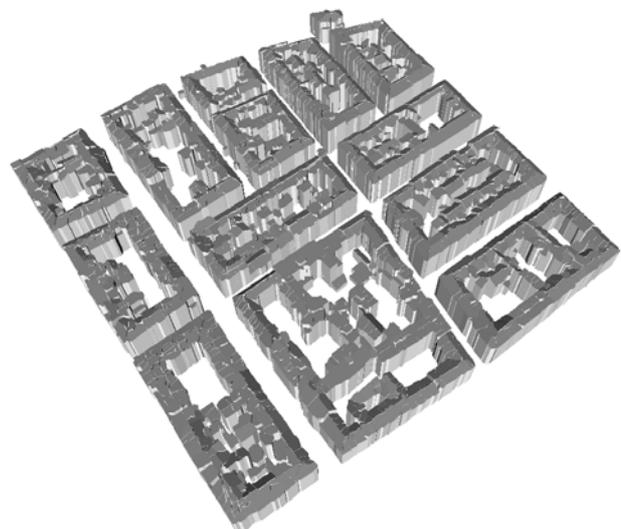


Figure 9. VRML visualization of a model created from the boundary polygons of the Voronoi diagram in Figure 5e.

## 6. CONCLUSION AND FUTURE WORK

We have presented a method for the extraction of buildings from high-resolution LIDAR data in densely built-up areas. Preliminary results from a test site in the City of Vienna have shown the high potential of the method. Future work will not only include the implementation of the modules still missing, but also the assessment of quality parameters for the results. With respect to the building outlines, this can be accomplished by a comparison to existing GIS data. In addition, we will perform research work in order to improve the results of building extraction by additionally using digital aerial images for matching of roof edges with image edges.

## REFERENCES

- Ameri, B., 2000: Automatic Recognition and 3D Reconstruction of Buildings from Digital Imagery. PhD Thesis, Institute of Photogrammetry, Stuttgart University, DGK-C 526.
- Baillard, C., Schmid, C., Zisserman, A., Fitzgibbon, A., 1999: Automatic line matching and 3D reconstruction of buildings from multiple views. IAPRS 32 (3-2W5), pp. 69-80.
- Brenner, C., 2000: Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. PhD Thesis, Institute of Photogrammetry, Stuttgart University, DGK-C 530.
- Briese, Ch., 2001: Digitale Modelle aus Laser-Scanner-Daten in städtischen Gebieten. Diploma Thesis, Inst. of Photogrammetry and Remote Sensing, Vienna University of Technology.
- Brunn, A., Weidner, U., 1997: Extracting Buildings from Digital Surface Models. IAPRS 32 (3-4W2), pp. 27-34.
- Fuchs, C., 1998: Extraktion polymorpher Bildstrukturen und ihre topologische und geometrische Gruppierung. PhD Thesis, Institute of Photogrammetry, Bonn University, DGK-C 502.
- Haala, N., Brenner, C., and Anders, K.-H., 1998: Urban GIS from Laser Altimeter and 2D Map Data. IAPRS 32 (3/1), pp. 339-346.
- Kraus, K., 2000: Photogrammetrie Band 3. Topographische Informationssysteme. 1<sup>st</sup> ed., Dümmler Verlag, Köln.
- Kraus, K., Pfeifer, N., 1998: Determination of terrain models in wooded areas with aerial laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53 (4), pp. 193-203.
- Moons, T., Frere, D., Vandekerckhove, J., Van Gool, L., 1998: Automatic modeling and 3D reconstruction of urban house roofs from high resolution aerial imagery. ECCV'98, Lecture notes in computer vision 1406(I), pp. 410-425.
- Pfeifer, N., Stadler, P., Briese, Ch., 2001: Derivation of digital terrain models in the SCOP++ environment. Proc. of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Terrain Models, Stockholm, Sweden.
- Vosselman, G., Dijkman, S., 2001: 3D building model reconstruction from point clouds and ground plans. IAPRS 34 (3W4), pp. 37-43.

Weidner, U., 1997: Gebäudeerfassung aus digitalen Oberflächenmodellen. PhD Thesis, Institute of Photogrammetry, Bonn University, DGK-C 474.

## ACKNOWLEDGEMENTS

This research has been supported by the Austrian Science Foundation (FWF) under project no. P14083-MAT.