

Virtual City Models from Laser Altimeter and 2D Map Data

Abstract

Virtual reality applications in the context of urban planning presume the acquisition of 3D urban models. Photo realism can only be achieved, if the geometry of buildings is represented by a detailed and accurate CAD model and if additionally artificial texture or real world imagery is mapped to the faces and roofs of the buildings. In the approach presented in this paper, height data provided by airborne laser scanning and existing ground plans of buildings are combined in order to enable an automatic 3D data capture. On demand the automatic building reconstruction can additionally be visually controlled and refined by an interactive tool. Virtual reality city models are generated in the final step by mapping terrestrial images to the facades of the reconstructed buildings. Thus the rapid acquisition of 3D urban GIS is feasible.

1 Introduction

Airborne laser scanners allow for the direct measurement of the topographic surface, including objects rising from the ground like trees or buildings. Even though for a number of applications a so-called Digital Surface Model (DSM), which represents this surface by an object independent distribution of 3D points is sufficient, further interpretation and qualification of the original height data is necessary in many cases. One example is the acquisition of city models for virtual reality applications, where even small geometrical errors like the non-planar triangulation of a planar facade may easily disturb the impression of looking at a 'real' scene. Another example are visibility computations for mobile communications. In this case, a scene representation with as few as possible faces is advantageous, since this reduces the simulation complexity. In order to obtain a 3D urban model, which is more suitable for these applications an abstraction of the surface model is required; buildings have to be separated from the terrain surface and represented by true 3D CAD models.

Many researchers nowadays agree that as many as possible supplementary sources of information have to be used to achieve optimal interpretation results. Within the approach presented in this paper the 3D reconstruction of the buildings is supported

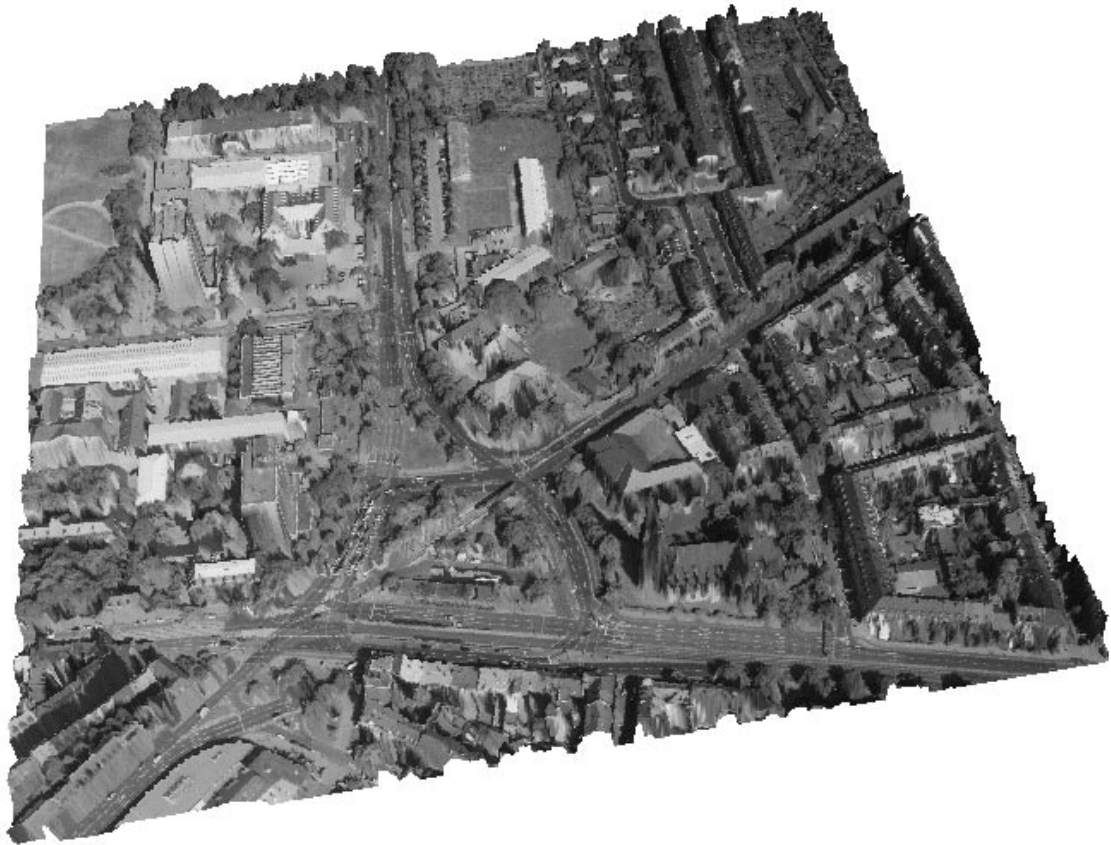


Figure 1: 3D visualization of DSM overlaid with ortho image.

by given ground plans. Often, ground plans of buildings have already been acquired and are represented either in analog form by maps and plans or digitally in 2D Geo Information Systems (GIS). These ground plans are another very important source of information for 3D building reconstruction. Compared to results of automatic procedures these ground plans are very reliable since they contain aggregated information, which has been made explicit by human interpretation. For this reason constraints, which are derived from ground plans can considerably reduce the search space when looking for a proper reconstruction and thereby reduce costs to attain a solution. Additionally, by integrating the ground plans into the processing, the consistency between the already existing 2D GIS or map data and the generated 3D reconstruction can be guaranteed.

An example for existing ground truth data relevant for building reconstruction is the digital cadastral map. It provides information on the distribution of property, including the borders of all agricultural areas and the ground plans of existing buildings. Additionally, information in form of text symbols is provided on the names of streets and the usage

of buildings like garage, residential building, office block, industrial building, church. At the moment the digital cadastral map is build up as an area covering data base, mainly by digitizing existing maps or plans. At present it is available for 40% of the area of Germany. Since for our test area this type of data was not available, the ground plans were digitized manually from a map of scale 1:5000. Alternatively maps and plans can be digitized automatically [Frischknecht and Carosio, 1997] resulting in information similar to the digital cadastral map.

After discussing the potential of airborne laser scanning for the generation and visualization of urban models (section 2), the 3D building reconstruction from existing ground plans and DSM will be described (section 3). In the last part of the paper (section 4) the processing of terrestrial images for the generation of virtual reality city models will be presented.

2 Urban models from laser scanner DSM

2.1 Airborne laser scanning

For 3D surface acquisition in open terrain image matching techniques have become standard tools. Still they suffer from problems in built-up areas due to occlusions and height discontinuities. In these areas the DSM quality mainly depends on the presence of texture at roof regions and on the amount of contrast between roof and terrain surface [Price and Huertas, 1992]. This results in considerable differences of DSM quality at roof regions, even in the same image pair. In recent times attempts have been made to improve the results of image matching in these areas, e.g. by using multiple overlapping images, by the integration of potential roof break-lines during the matching process [Maitre and Luo, 1992], or by applying adaptive matching mask sizes [Lotti and Giraudon, 1994]. However, the situation today is that even though the overall shape of buildings can be observed in the DSM, height discontinuities e.g. at roof lines are not defined very well.

As an alternative data source airborne laser scanning provides DSM of high and homogeneous quality in urban areas, which are very suitable for 3D building reconstruction. For our test area the DSM was acquired by the TopoSys laser scanner system [Lohr, 1997]. The central component of the system is a laser sensor, which allows the direct distance measurement from the aircraft to the topographical terrain surface by run-time measurement of an emitted and reflected laser pulse. For an area covering data collection, the

laser beam is deflected perpendicular to the direction of flight, resulting in a strip-wise acquisition of the terrain surface. The position and orientation of the sensor system during range measurement has to be provided by additional sensors. These components are a GPS receiver for the positioning task and an inertial system (INS) for the orientation task.

Terrain points were measured at approximately one point each $1 \times 1 \text{ m}^2$ with an accuracy of 0.3 m in planimetry and 0.1 m in height. The system provides dense point measurements along strips, which are usually processed and resampled to obtain a regular raster. Figure 1 shows the result of these steps for our test area (part of the city of Karlsruhe, Germany), where an aerial image has been additionally overlaid. This view gives a good impression of the scenery, since the geometry of the surface is represented quite well.

2.2 Visualization

One problem with employing surface descriptions, which are based on an unqualified, i.e. object independent distribution of measured points is the large amount of data to be processed, stored and displayed. It has been recognized early by the computer graphics community that handling of very large scenes is an important topic. There are several approaches, among them clipping and viewing frustum computations, which are applicable when just parts of the scene need to be rendered for visualization. However, when the entire scene can be seen, for example during a virtual flight, which starts from far away and zooms in on some detail, a level-of-detail (lod) approach is more useful. In this case, the exact geometry of the scene is replaced by an approximation consisting of considerably less triangulated points as soon as the object is far away from the (virtual) camera. It is currently not possible to compute mesh simplifications on-line while rendering the scene. Thus, one traditional approach is to provide multiple level-of-detail representations of the object, which are computed off-line. Depending on the distance between object and camera the viewer program then just switches between those representations during rendering. Unfortunately, there can be a visible discontinuities while the viewer program switches between these representations. To this end, approaches like ‘continuous meshes’ have been developed [Hoppe, 1998]. This technique does not store a fixed number of models with predefined polygon numbers but the most simple polygonal representation of the object plus a list of modifications (e.g. edge inserts) leading to the most precise

representation. During display, the viewer chooses the number of modifications to be applied based on the objects distance. Another important development is the replacement of per-vertex or per-face color information by texture mapping.

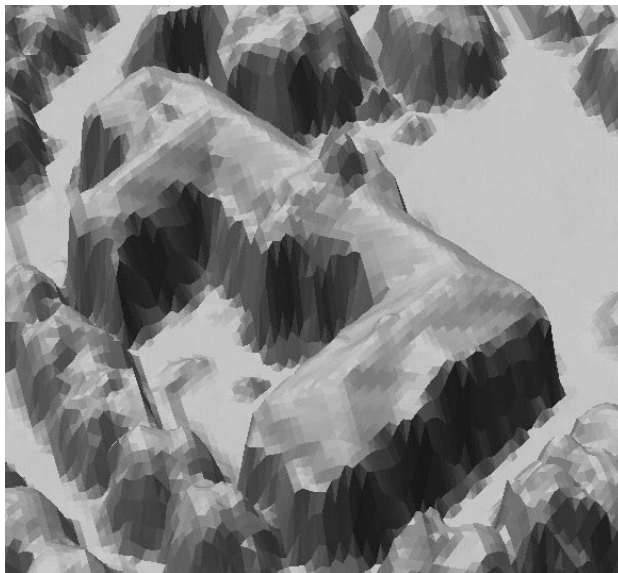


Figure 2: DSM, original resolution

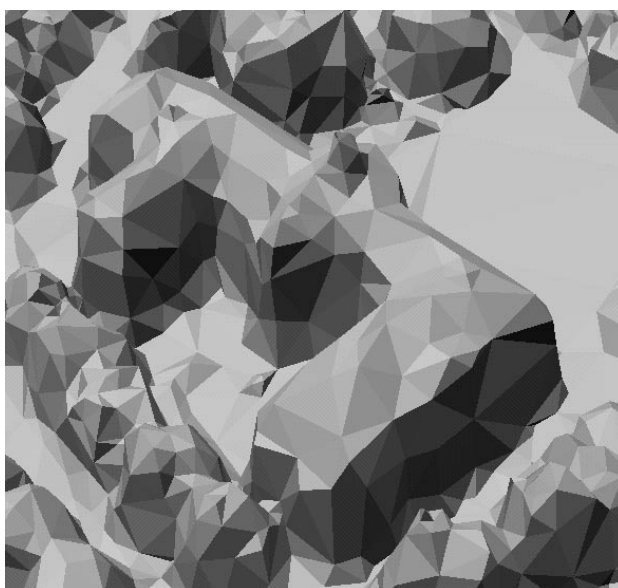


Figure 3: DSM, simplified

When combining a reduction in the number of polygons with texture mapping, usually a reduction by at least 90% of the original size can be achieved with little visual impact. A

good overview on mesh simplification techniques is given in [Heckbert and Garland, 1997]. To give an example for a mesh simplification, figure 2 shows a part of a DSM of our test area, which was overall modeled using 607624 triangles. Figure 3 shows the same scene after mesh simplification with only 50000 triangles for the whole test area. Triangle reduction was performed using the algorithm described in [Soucy, 1997].

Although the amount of data can be reduced significantly by these algorithms, many tasks aiming at visualizations or simulations in an urban environment require the further abstraction and interpretation of the surface description. For simulations knowledge about the surface material can be a crucial point. Trees and buildings have different influence on the propagation of noise and electro-magnetic waves. Hence these objects have to be classified and represented separately from the terrain surface. To achieve photo realism for the generation of walk-throughs terrestrial images have to be mapped onto the vertical faces of the buildings since the resolution and viewing direction of an aerial image is no longer sufficient at a large scale. In order to assign texture to a facade, it has to be identified in the terrestrial image. This requires the availability of an explicit representation of this part of the building. The situation is much like in reverse engineering for industrial parts where the ultimate goal is to find a compact parametric representation rather than dense point clouds. All those arguments vote for an explicit representation of building geometry and topology by 3D CAD models.

3 Building reconstruction

In order to separate buildings from the terrain surface and represent them by true 3D CAD models, ground plans are used in addition to the DSM data. For the reconstruction first appropriate building models have to be defined. In a second step these models are fitted to the observed data to reconstruct the building.

3.1 Building models

Object recognition or reconstruction, in general, presumes knowledge about the perceived objects by some kind of object model. A model used for building reconstruction should be able to describe buildings of different complexity and it should permit the representation of geometric constraints during the reconstruction. Object models can be treated as abstractions of real world objects. The most important role played in model definition is

the proper balance between the correctness and tractability, i.e. the results given by the model must be adequate both in terms of the solution attained and the cost to attain the solution.

In order to deal with the large architectural variations of building shapes, the utilized model should be as general as possible. In our approach a building is represented by a general polyhedron, i.e. it is bounded by a set of planar surfaces and straight lines. Generally, the interpretation of real world data presumes much a priori knowledge or in other words constraints. This can be achieved by applying a very rigid building model. Of course, this limits the number of possible building types, which can be represented by a single model. We provide the required constraints by using the given ground plan of the building. This supplies sufficient restrictions to enable the reconstruction without losing the possibility to deal with very complex buildings. Furthermore, in the framework of a semi-automatic process the manual acquisition of a ground plan from a DSM or an ortho image is more efficient compared to procedures that rely on 3D measurement and interaction of a human operator to provide the initial information.

Two approaches to represent the reconstructed buildings are feasible, boundary representation and constructive solid geometry. The boundary representation (BRep) is probably the most widespread type of 3D representation and many algorithms are available for computing physical properties from this representation. A BRep defines spatial objects by their bounding elements, e.g. planar faces. Nodes and edges are defined by intersection of the bounding planes. The topology is additionally captured by a set of relations that indicate how the faces, edges and vertices are connected to each other. In constructive solid geometry (CSG) simple primitives are combined by means of regularized Boolean set operators. A CSG representation always results in valid 3D objects, i.e. in contrast to a BRep no topological check has to be performed in order to guarantee that the object surface is closed. CSG also enables a very compact object representation. Whereas a CSG can be transformed into a BRep, there are no complete solutions available in the opposite direction. This motivated us to use CSG as primary representation and to generate a BRep on demand e.g. for visualization purposes. Thus the advantages of both representations can be combined.

Similar to [Englert and Gülch, 1996] we utilize a CSG representation, which describes each building by a combination of one or more basic primitives. The set of four basic building primitives used for that purpose is shown in figure 4. Each building primitive

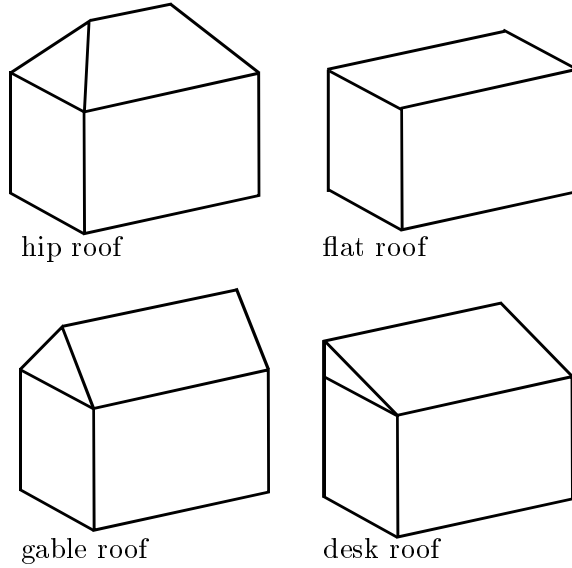


Figure 4: Building primitives used for reconstruction.

consists of a cuboid element with the different roof types flat roof, pent roof, gable roof and hip roof.

3.2 Decomposition of ground plans

In the first step of the reconstruction process the complete building has to be subdivided into these basic structures. This is enabled by the automatic decomposition of the ground plans into rectangular structures. The utilized algorithm is very similar to approaches for perceptual grouping of line segments into rectangular structures [Mohan and Nevatia, 1989]. In our application the line segments are given by the single elements of the ground plan polygon. An example for a complex building is given in figure 5.

Each of the resulting rectangles triggers the reconstruction of one building primitive. Position, orientation and horizontal extension of each cuboid are already defined by the rectangle. Therefore, only the height of each cuboid as well as the roof type and roof slope have to be determined as unknown parameters for each building primitive. These parameters are estimated by a least squares adjustment, which minimizes the distances between the DSM surface and the corresponding points of the building primitive. Thus the building primitives are fit to the DSM surface. In order to apply the least squares ad-

justment, first the appropriate model has to be selected. Additionally, roof regions, which do not fit to the selected model have to be excluded from the least squares adjustment in order to avoid gross errors during parameter estimation. Both tasks can be solved by a segmentation of the DSM into planar surfaces.



Figure 5: Ground plan decomposed into rectangular parts.

3.3 Segmentation and parameter estimation

This segmentation is also supported by introducing ground plan information. First, the given ground plan restricts the extension of the DSM area, which has to be examined. More importantly, the segmentation can be based on the direction of the surface normals of the DSM, since possible orientations of planar surfaces to be extracted are predefined by the outline of the building. This is motivated by the fact that the direction of the unit normal vector of a possible roof plane emerging from an element of the ground plan has to be perpendicular to this segment. Hence, the different segments of the ground plan polygon are used to trigger the segmentation into planar surfaces with a projected normal vector perpendicular to this element. For reasons of simplicity this step is discussed using

a building, which can be represented by a single CSG primitive.

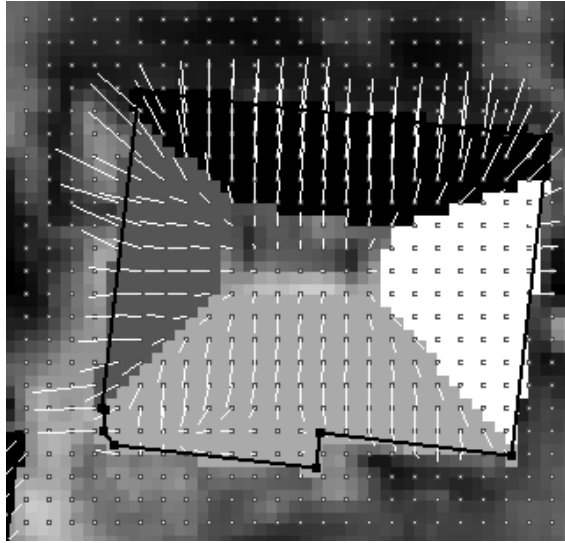


Figure 6: Building ground plan (black), DSM surface normals (white) and segmented planar surfaces.

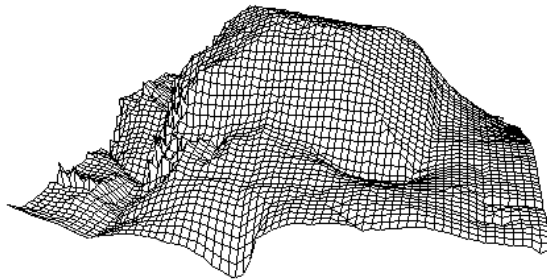


Figure 7: 3D view of DSM.

In figure 6 a ground plan provided by the digital cadastral map is superimposed to the corresponding section of the ortho image. The corresponding DSM is shown in figure 7. The implemented segmentation is based on the direction of the surface normals of the DSM, which are represented by the small white lines in figure 6.

The distribution of the surface normal directions corresponds to the four major axes of the ground plan. Even though these directions can also be calculated by analyzing the histogram of the surface normal directions, they are obtained by parsing the given ground

plan, which is much more reliable [Haala and Brenner, 1997]. All points with a surface normal corresponding to the examined ground plan direction are combined to a region. This results in the segmentation represented by the shaded regions in figure 6. The result of the segmentation process can be used to define so-called compatibility regions for the estimation of each roof, i.e. only DSM segments with a direction of the normal vector compatible to the ground plan segment are utilized for estimating the parameters of the corresponding roof plane. The segmentation of figure 6 triggers the reconstruction of a building with hip roof, since this model (roof consisting of 4 faces) is the only one which fits to the result of the segmentation process.



Figure 8: Original ground plan and segmented roof regions

For the more complex building shown in figure 5 the segmentation is presented in figure 8. Figure 9 shows the building primitives, which are reconstructed based steps discussed above. Figure 10 gives the result of the automatic 3D building reconstruction for the complete test area. For visualization the reconstructed buildings were put on the map of scale 1:5000, which was used to digitize the ground plans.

3.4 Interactive refinement of initial reconstructions

In our approach the reconstruction is constrained by the assumption that

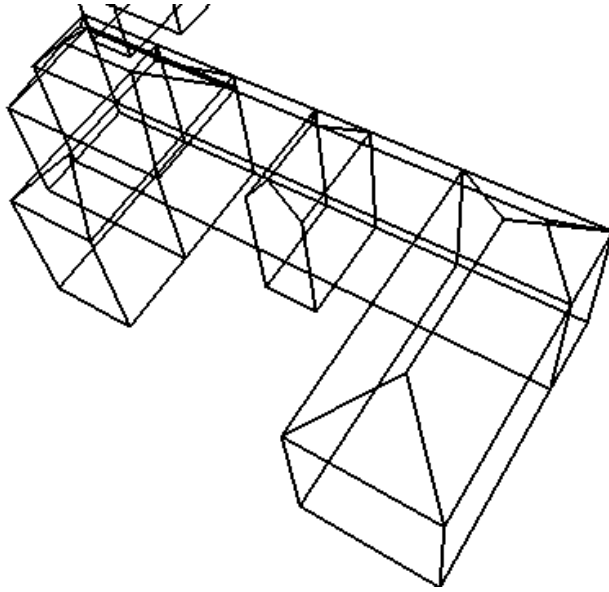


Figure 9: Reconstructed building primitives.

- all walls defined by the ground polygon lead to a planar roof face of variable slope and
- the ground plan can be decomposed into rectangular primitives.

These assumptions are fairly general. However, one must keep in mind that any roof construction based on this approach provides incorrect results, if the roof structure inside the ground polygon does not follow the cues that can be obtained from the ground polygon. This can happen if more than one plane emerges from a single polygon element or if parts of the building, which are contained in a roof surface like a bay are not represented by the ground plan.

Figure 11 shows the reconstructed building already depicted in figure 9 with the DSM surface overlaid. The difference between the DSM surface and the corresponding points at the roof planes provide a reliable test on the quality of a reconstruction. For this reason RMS values are calculated for each building and its sub-parts. Remaining planar regions generated during segmentation, which are incompatible with the final reconstruction give an additional hint, if manual interaction is required for further refinement.

Up to now all buildings were reconstructed fully automatically. The ground plans used so far were digitized merely from the map. No care has been taken to digitize them with respect to the reconstruction algorithm. Even though the reconstruction is sufficient for many levels of detail, due to the problems of the algorithm mentioned above a further

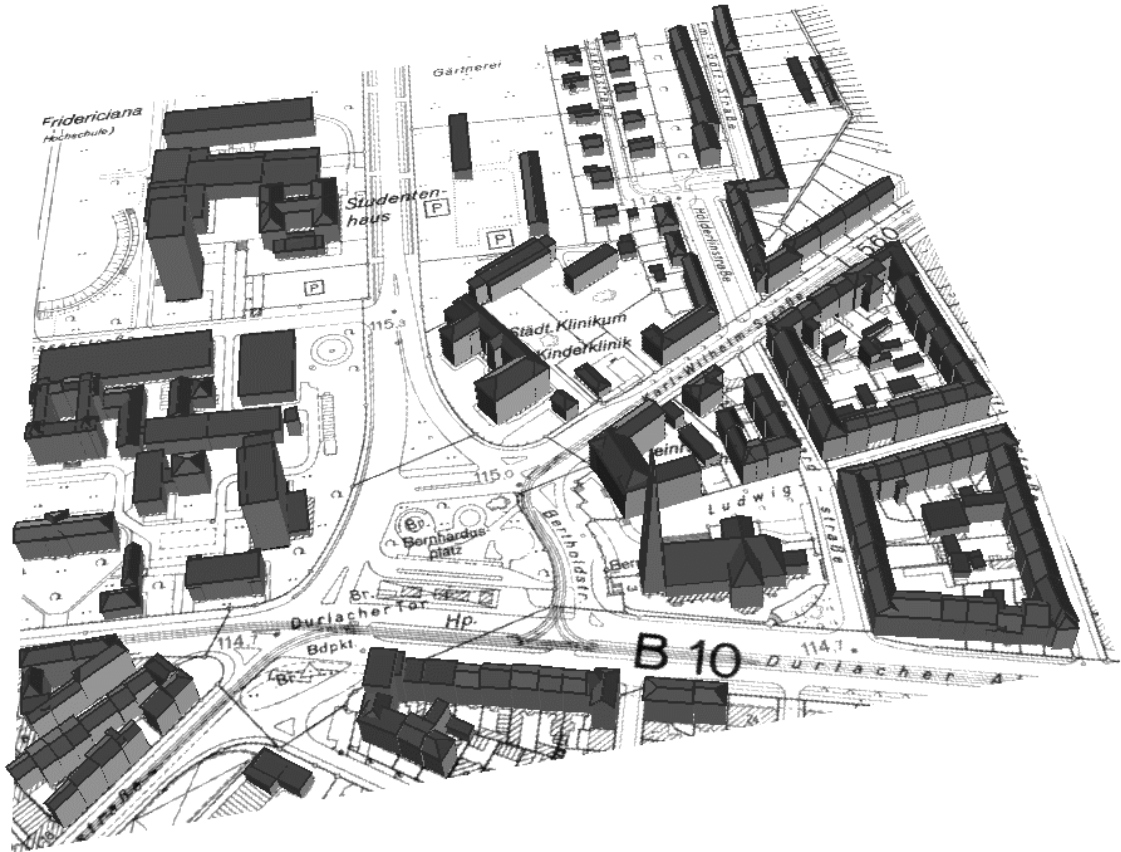


Figure 10: Reconstructed buildings projected on map 1:5000 used for ground plan acquisition.

improvement of the reconstruction can be necessary for very complex buildings. This can be obtained by analyzing the initial reconstruction in order to refine the capture of the ground plans. For that purpose a tool with graphical user interface (GUI) has been developed, which provides the possibility to modify ground plan primitives interactively (figure 12). The existing ground plans can be overlaid simultaneously to different images like ortho image, map, and a grey-value representation of the DSM. If a building polygon is selected, the generated reconstruction is additionally presented in combination with the DSM surface in a 3D viewer. The GUI allows to manipulate primitives (rectangles) defined by the automatic decomposition algorithm. Moreover, new primitives can be defined, which trigger the reconstruction of additional 3D building primitives instantly.

Figure 13 shows the original ground plan (white polygon) and two black rectangles added interactively using the GUI. At these areas both the segmentation into compatible regions presented in figure 8 as well as the comparison of the reconstructed building to the DSM surface presented in figure 11 give hints to an improper reconstruction. The



Figure 11: Automatically reconstructed building and DSM surface.

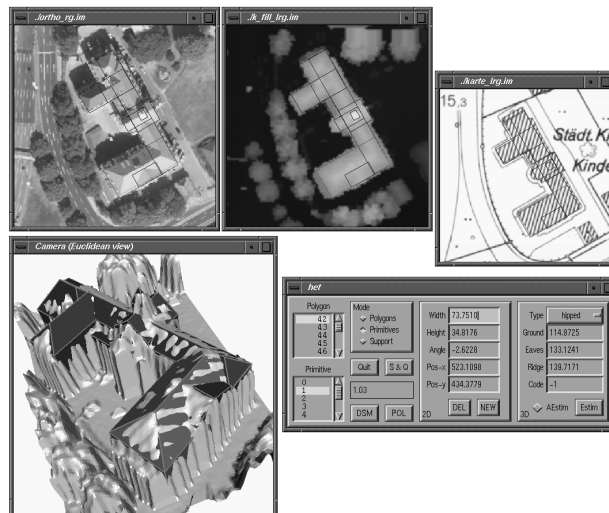


Figure 12: View of the interactive editing tool used to refine automatic house reconstructions.

rectangles added manually trigger the automatic reconstruction of two additional building primitives representing a bay of the roof and the small tower. The result of this reconstruction is shown in figure 14.

Finally, the CSG representation is transformed to a boundary representation of the building. Therefore the union of the set of CSG primitives has to be computed. Within



Figure 13: Building ground plan (white) an interactively added rectangles (black).

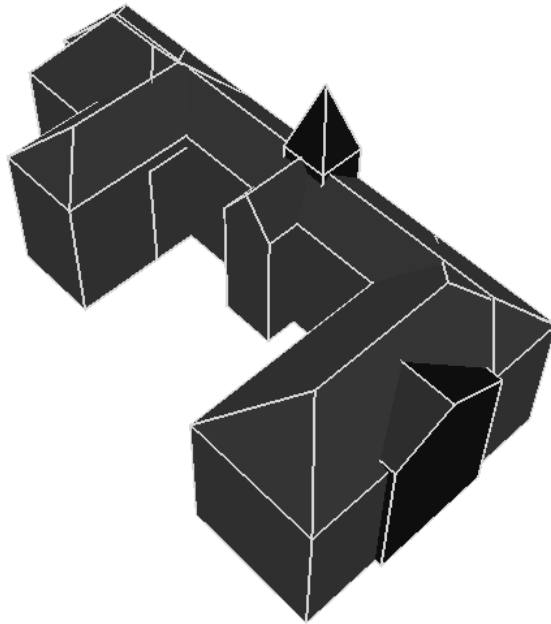


Figure 14: Reconstruction with additional primitives (dark) triggered by manual interaction.

this process the primitives are intersected, coplanar and touching faces are merged and inner faces or parts are removed.

The size of object parts, which can be reconstructed, is of course limited by the available density of DSM points, i.e. details smaller than approximately 1 m can not be captured. For virtual reality applications this problem can be avoided by texture mapping of real imagery as a substitute for geometric modeling, since the use of photo realistic texture enhances the perceived detail even in the absence of a detailed geometric model [Gruber et al., 1995].

4 Generation of virtual reality models

The creation of a 3D city model for virtual reality applications usually consists of a geometric building reconstruction followed by texture mapping to obtain a photo realistic model representation. In principle, terrestrial imagery is sufficient to provide the required information for both tasks. Nevertheless, especially the terrestrial acquisition of the building geometry by architectural photogrammetry proves to be a time-consuming process that has to be carried out interactively for each building. The basic idea of our approach is to speed up the time consuming process of virtual city model creation by using DSM and ground plans for geometric processing and terrestrial images only for texture mapping. Since the vertices of the 3D building models, which are generated from ground plans and laser data provide sufficient control point information, the texture mapping from the terrestrial images is simplified considerably. Therefore, the generation of virtual reality models is more efficient compared to standard architectural photogrammetry, where a number of tie points has to be measured in multiple images.

The goal of texture processing is to provide a rectified image for each visible building face. Hence for each image the corresponding facade polygon has to be selected from the 3D city model generated in the previous processing step. For this purpose the wire frame of the reconstructed buildings as well as the indices of the faces are projected to the aerial image (see figure 15). If the viewpoints were sketched into a map or an ortho image during the terrestrial image acquisition, this representation allows a simple interactive definition of the corresponding face index for each terrestrial image.

For texture mapping the image has to be correctly positioned, oriented and scaled to represent its associated surface. In our approach an image section representing a planar surface is rectified by applying a projective transformation. The parameters of the projective transformation are determined by a minimum number of 4 points in 3D



Figure 15: Reconstructed building projected into stereo image.

world coordinates on a plane (in 3D space) and their corresponding image coordinates. Of course, this approach can only be used with sufficiently planar structures. After the selection of the corresponding 3D building face at least four tie points between the face polygon and the terrestrial image are determined. For this purpose the nodes of the face polygon are identified and measured in the terrestrial image.



Figure 16: Original terrestrial images with points measured for rectification

Figure 16 shows an example for terrestrial imagery taken for texture mapping. The images were acquired with a Kodak DCS 120 digital camera. The points measured for the rectification are marked by white dots. If a node of the face is hidden by an obstacle,



Figure 17: Faces mapped to corresponding 3D surfaces

the corner point can alternatively be calculated from points measured at the edges of the facade. The rectified image sections can then be assigned to their corresponding faces in the 3D model. In order to control the correct assignment of the texture, the 3D buildings are displayed in a viewer, which allows all basic transformations to the model (translation, rotation, scaling) immediately after the texture mapping (see figure 17). For the final visualizations (see figures 18 and 19) the ortho image as well as artificial trees are added to the virtual model.

5 Conclusion

Recent advances in three-dimensional displays, real-time texturing and computer graphics hardware as well as the increasing availability of rendering and animation software tools have resulted in an increased demand for photo realistic 3D virtual reality city models. This demand can only be satisfied by a highly efficient capture of urban scenes, which – in our opinion – presumes the integrated use of multiple data sources. Aiming on that purpose we have presented an approach using 2D ground plans and a laser DSM to derive 3D building geometry automatically, whereas texture is taken from terrestrial photographs.

In order to evaluate the accuracy and reliability of the proposed method, our results



Figure 18: 3D visualization of virtual city model.



Figure 19: 3D visualization of virtual city model.

will have to be compared to reference data. Another topic to be further improved is the analysis of the existing ground plans. Up to know we assume that the coordinates of the ground plan are correct and the borders of the roof are exactly defined by this ground plan. This presumes the automatic detection of discrepancies between the 2D GIS and the

laser data in order to verify the ground plans before starting the building reconstruction. In future we also want to correct slight errors of the ground plan geometry using the laser data, a task which requires an even closer integration of the available information.

References

- [Englert and Gülch, 1996] Englert, R. and Gülch, E., 1996. One-eye stereo system for the acquisition of complex 3D building descriptions. *GEO-INFORMATION-SYSTEMS* 9(4), pp. 16–21.
- [Frischknecht and Carosio, 1997] Frischknecht, S. and Carosio, A., 1997. Raster-based methods to extract structured information from scanned topographic maps. In: *ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects*, Stuttgart, Germany, pp. 63–67.
- [Gruber et al., 1995] Gruber, M., Pasko, M. and Leberl, F., 1995. Geometric versus texture detail in 3-D models of real world buildings. In: A. Gruen, O. Kuebler and P. Agouris (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser Verlag, pp. 189–198.
- [Haala and Brenner, 1997] Haala, N. and Brenner, C., 1997. Interpretation of urban surface models using 2D building information. In: A. Gruen, O. Kuebler and M. Baltasvias (eds), *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser Verlag, pp. 213–222.
- [Heckbert and Garland, 1997] Heckbert, P. and Garland, M., 1997. Survey of polygonal surface simplification algorithms. In: *SIGGRAPH '97 Course Notes Multiresolution Surface Modeling*, Los Angeles.
- [Hoppe, 1998] Hoppe, H., 1998. Smooth view-dependent level-of-detail control and its applications to terrain rendering. In: *IEEE Visualization '98*, pp. 35–42.
- [Lohr, 1997] Lohr, U., 1997. Digital elevation models by laserscanning: Principle and applications. In: *Third International Airborne Remote Sensing Conference and Exhibition, Vol. I*, Copenhagen, Denmark, pp. 174–180.
- [Lotti and Giraudon, 1994] Lotti, J.-L. and Giraudon, G., 1994. Adaptive window algorithm for aerial image stereo. In: *IAPRS, Vol. 30, Part 3/1*, pp. 517–524.

- [Maitre and Luo, 1992] Maitre, H. and Luo, W., 1992. Using models to improve stereo reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2), pp. 269–277.
- [Mohan and Nevatia, 1989] Mohan, R. and Nevatia, R., 1989. Segmentation and description based on perceptual organization. In: *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, pp. 333–341.
- [Price and Huertas, 1992] Price, K. and Huertas, A., 1992. Using perceptual grouping to detect objects in aerial scenes. In: *IAPRS, Vol. 29, Part B4*, Washington D.C., pp. 842–855.
- [Soucy, 1997] Soucy, M., 1997. *InovMetric’s Multiresolution Modeling Algorithms*. In: *SIGGRAPH ’97 Course Notes Multiresolution Surface Modeling*, Los Angeles.