D. Fritsch, M. Englich & M. Sester, eds, 'IAPRS', Vol. 32/4, ISPRS Commission IV Symposium on GIS - Between Visions and Applications, Stuttgart, Germany.

FAST PRODUCTION OF VIRTUAL REALITY CITY MODELS

Claus Brenner and Norbert Haala
Institute for Photogrammetry (ifp)
University of Stuttgart
Geschwister-Scholl-Straße 24, 70174 Stuttgart, Germany
Ph.: +49-711-121-4097, Fax: +49-711-121-3297
e-mail: Claus.Brenner@ifp.uni-stuttgart.de

Commission IV, Working Group IV/III.2

KEY WORDS: city models, virtual reality, laser scanning

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 artificial texture or real world imagery is additionally 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 data capture by the integration of these different types of information. Afterwards, virtual reality city models are generated by texture processing, i.e. by mapping of terrestrial images. Thus, the rapid acquisition of 3D urban GIS is feasible.

1 GENERATION OF VIRTUAL REALITY CITY MODELS

The acquisition of three-dimensional databases for urban areas has become a topic of growing interest to the photogrammetric community. Simulations that require the representation, management and analysis of three-dimensional descriptions have become standard applications for planning purposes in built-up areas. One example is the propagation of electro-magnetic waves that can be simulated for planning optimal locations for transmitter stations. A very widespread application is the computation of synthetic views or even animated fly-throughs to visualize urban scenes. The creation of a 3D city model for these virtual reality applications usually consists of a geometric object reconstruction followed by texture mapping to obtain a photo realistic model representation. Objects relevant for the 3D description of urban areas are buildings, streets, urban vegetation like trees and Digital Terrain Models (DTM). Because manual interpretation during data acquisition is very time consuming, a lot of effort has been spent to speed up this process by automatic or semi-automatic procedures. In our opinion an efficient acquisition of 3D urban models presumes the use of multiple data sources since a high degree of automation can only be reached by integrating different and complementary types of information.

For the geometric data acquisition we use existing 2D GIS, color images and Digital Surface Models. As the information of a Digital Surface Model is restricted to surface geometry, the interpretation of this kind of data is easier e.g. compared to the interpretation of image data. A DSM, i.e. a geometric representation of the terrain surface, including objects like trees or buildings which rise from the ground, can e.g. be obtained from stereo image matching. For our purpose DSM from airborne laser scanners are utilized, since the direct height data acquisition by these systems provides DSM of high and homogeneous quality in urban areas, which is essential for our approach. Despite the high quality of the laser data, the automatic interpretation of DSMs remains a very pretentious problem. Within the approach presented in this paper the 3D reconstruction of the buildings is supported by given ground plans. This type of information is frequently available from 2D Geo Information Systems (GIS). By integrating the ground plans into the processing also the consistency between the already existing 2D GIS or map data and the generated 3D data can be guaranteed. One problem arising when existing databases are used is their potential lack of actuality and incompleteness. While aiming on the combination of a 2D GIS with directly captured data like DSM or images, a map revision has to be performed as a first step. For this purpose obsolete or incomplete parts of the GIS have to be uncovered. In addition to the detection or validation of inconsistencies between the datasets, there is a need to capture objects, which are not contained in the 2D GIS. While outlines of buildings and traffic network are available from standard databases, vegetation is usually not represented in detail. However, these objects are relevant for 3D site models. In order to automatically detect objects of interest like buildings, streets and trees in the DSM and multispectral images a classification approach is applied.

To achieve photo realism an accurate geometric and textural description is required for each building. Since real images embody a representation of object detail they can substitute for geometric modeling. For this reason the use of photo realistic texture enhances the perceived detail even in the absence of a detailed geometric model. Hence, even though artificial texture can be assigned to building faces depending on the surface material, for realistic results real imagery has to be mapped to the facades at a sufficient resolution. In airborne data mainly the roofs of buildings can be observed. Due to the viewpoint, facades, jut out of roofs or passages are hardly visible or even totally occluded. The limited resolution of airborne data additionally restricts the amount of detail and the accuracy. Photo realism can only be achieved in the framework of virtual reality applications if terrestrial images are acquired and mapped on their corresponding building facades.

After discussing the potential of airborne laser scanning for

D. Fritsch, M. Englich & M. Sester, eds, 'IAPRS', Vol. 32/4, ISPRS Commission IV Symposium on GIS - Between Visions and Applications, Stuttgart, Germany.

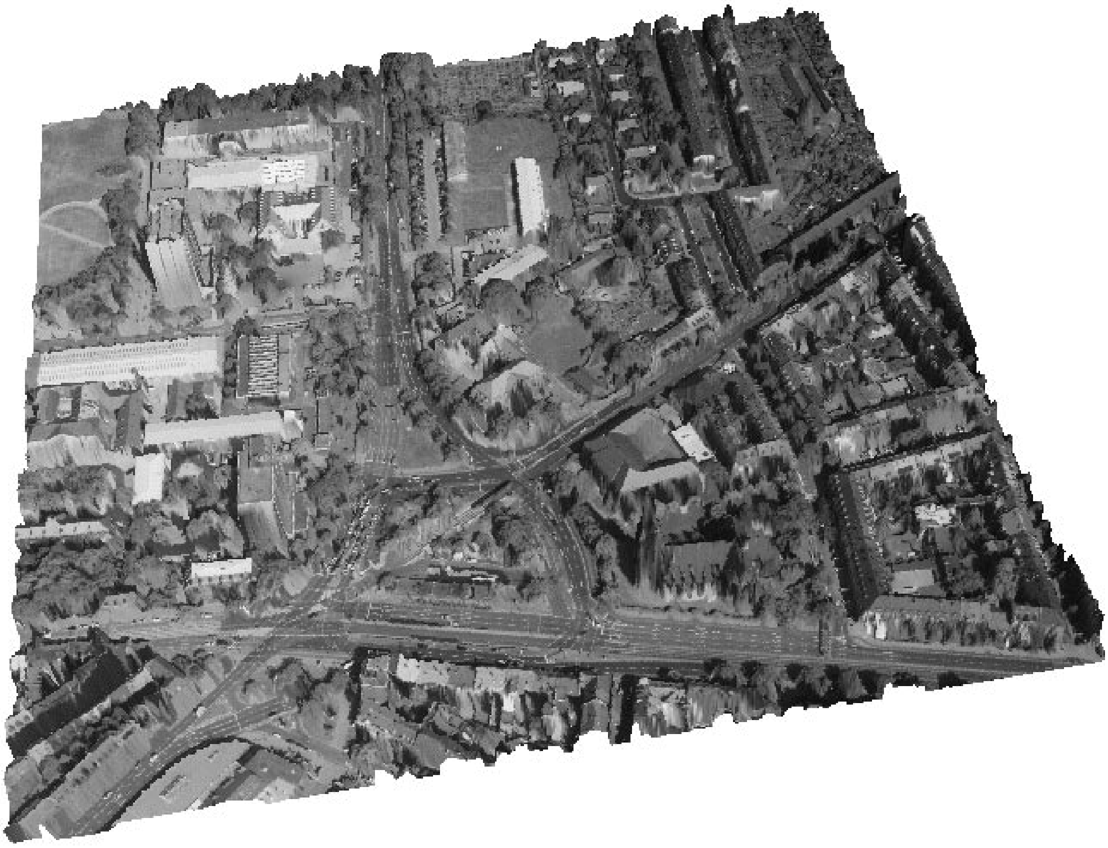


Figure 1: 3D visualization of DSM overlaid with ortho image.

the generation and visualization of urban models (section 2), the automatic 3D building reconstruction from existing ground plans and DSM will be described (section 3). Following this we show an interactive refinement tool which can be applied subsequently to modify building reconstructions obtained automatically. In the last part of the paper the processing of terrestrial images for the generation virtual reality city models will be presented.

2 URBAN MODELS FROM LASER SCANNER DSM

Figure 1 shows a DSM of our test area provided by the TopoSys laser scanner system (Lohr, 1997). Terrain points were measured at approximately one point each 1×1 m^2 with an accuracy of 0.3 m in planimetry and 0.1 m in height. Similar to a tachymetric data acquisition, the three-dimensional coordinates of terrain points are determined by polar measurement. For visualization an aerial image has been additionally overlaid in figure 1.

One problem with employing surface descriptions by an unqualified, i.e. object independent distribution of points like they are generated by airborne laser scanning is the large amount of data to be computed, 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. One traditional approach is to provide multiple level-of-detail representations which are computed off-line and use different object representations depending on the distance of the (virtual) camera from the object.

Although the amount of data can be reduced significantly by these algorithms, many tasks aiming on visualizations or simulations in an urban environment require further abstraction and interpretation of the surface description. For simulations on the propagation of noise or electro-magnetic waves e.g. knowledge about the surface material is a crucial point. Hence trees or buildings have to be represented separately from the terrain surface.

Additionally terrestrial images have to be mapped onto the vertical faces of the buildings to achieve photo realism for the generation of walk-throughs; the resolution and viewing direction of an aerial image, which was e.g. used to produce figure 1 is no longer sufficient at a large scale. For this purpose point correspondences have to be established between the terrestrial images and the data set used to represent building geometry. Since nodes and vertices of buildings can be identified easily in terrestrial images, they should also be represented explicitly in the geometric database. All those arguments vote for an explicit representation of building geometry and topology by 3D CAD models.

3 ACQUISITION OF BUILDING GEOMETRY

In order to deal with the large architectural variations of building shapes, the aim of our aproach is to use a very general building model. Therefore the building is represented by a general polyhedron, i.e. it is bounded by a set of planar surfaces. Additional constraints are provided by existing ground plans of buildings which frequently 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 acquired by human interpretation. For this reason constraints, which are derived from ground plans can considerably reduce the search space while looking for a proper reconstruction and thereby reduce computational cost to attain a solution. In our approach we assume that the coordinates of the given ground plan are correct and the borders of the roof are exactly defined by this ground plan.

An example for existing ground truth data relevant for building reconstruction is the digital cadastral map, which provides information on the distribution of property, including the borders of all agricultural areas and the ground plans of existing buildings. Additional information on the names of streets and the usage of buildings (e.g. garage, residential building, office block, industrial building, church) is provided in form of text symbols. At the present moment the digital cadastral map is build up as an area covering data base, mainly by digitizing existing maps or plans. Currently it is available for about 40% of the area of Germany. Since this type of data was not available for the test area the ground plans were extracted manually from a map of scale 1:5000 which is shown in figure 2.

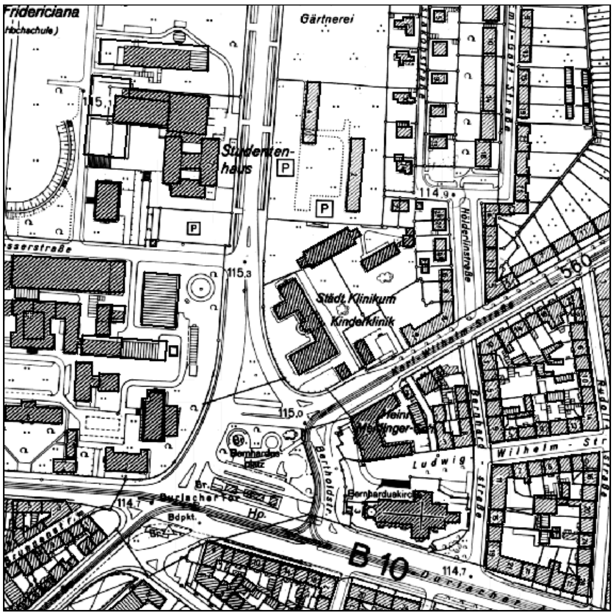


Figure 2: Map 1:5000 and extracted ground plans.

3.1 Decomposition of ground plans

The core idea of our algorithm is to subdivide each building into a number of basic primitives based on the given ground plan and to fit these primitives to the DSM data. The set of four basic building primitives used for that purpose is shown in figure 3. Each primitive consists of a cuboid element. The elements differ in their roof type, which can be one of flat roof, desk roof, gable roof or hip roof (see figure 3). This type of representation corresponds to the well-known CSG representation used in computational geometry. In constructive solid geometry (CSG) simple primitives are combined by means of Boolean set operators (union, intersection, subtraction) in order to obtain complex bodies.

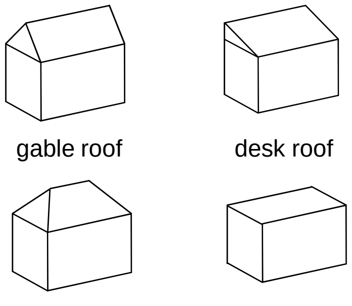


Figure 3: Building primitives used for reconstruction.

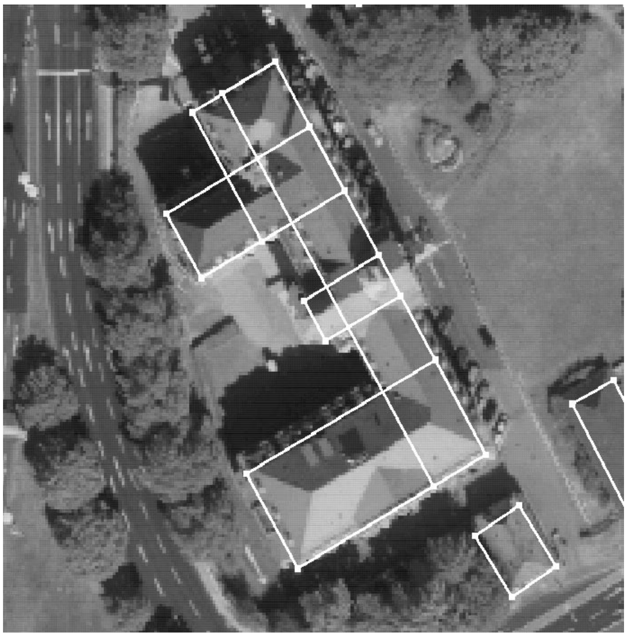


Figure 4: Ground plan decomposed into rectangular parts.

Figure 4 shows the result of a ground plan decomposition into rectangular structures, which is is realized automatically by the analysis of the given outline of the building. Each rectangle defines the base of one building primitive, i.e. position, orientation and horizontal extension of each cuboid are already defined by the parameters of the rectangle. Remaining unknown parameters are the height of the cuboid, the roof type and roof slopes. These parameters are estimated by a least squares adjustment, which minimizes the distances between the DSM surface and the corresponding building primitive, i.e. the building primitives are fit to the DSM surface.

In order to apply the least squares adjustment first the appropriate building primitive has to be selected and roof regions which do not fit to the selected model have to be excluded from the least squares adjustment to avoid gross errors of the estimated parameters. For this purpose the DSM is segmented into planar surfaces by analysing the slope of the DSM. Of course the given ground plan restricts the extension of the area which has to be examined. More important, possible orientations of roof planes, i.e. planar

Figure 5: Segmented roof regions.

surfaces to be extracted are predefined by the outline of the building. This is motivated by the observation 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. In order to make use of this knowledge, the surface normal for each DSM point is computed. Since direct numerical differentiation tends to amplify noise and obscure signal content, a local least squares fit is computed within a small window around each element of the DSM (Besl, 1988). The derivatives of the continuous function then can be determined analytically at the corresponding discrete DSM points. Afterwards all points with a surface normal compatible to the examined ground plan direction are combined to a region. For the building shown in figure 4 there are four directions defined by the ground plan, which are oriented with an angle of 0^o, 90^o, 180^o and 270^o to the main axis of the building.

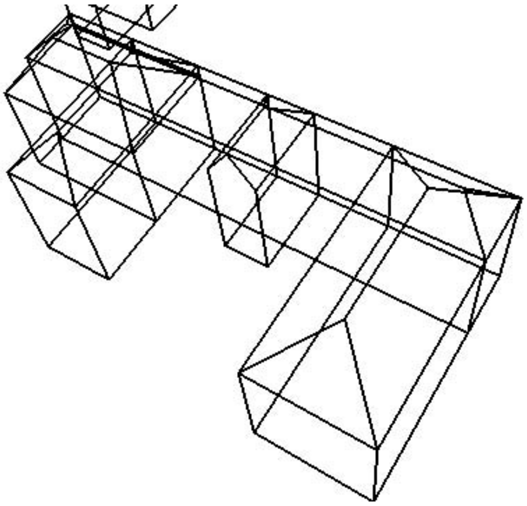


Figure 6: Reconstructed CSG building primitives.

Applying our algorithm results in the segmentation represented by the shaded regions in figure 5. The segmentation into normal vector compatible regions reflects the roof structure quite well. This demonstrates that by using the ground plan information in the segmentation process, planar regions can be extracted reliably even for DSM of limited resolution. The result of the segmentation process can be used to define so-called compatibility regions. Only DSM segments with a compatible direction of the normal vector are utilized while estimating the parameters of the corresponding roof plane. The compatible regions, as well as the rectangles obtained from ground plan decomposition are used for parameter estimation. The result of this process is represented in figure 6.

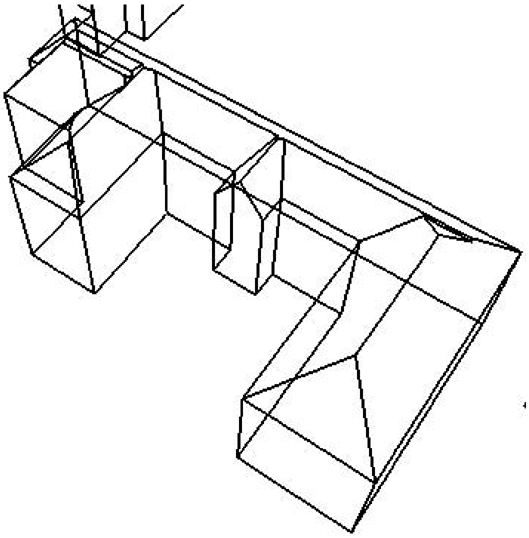


Figure 7: CSG data converted to boundary representation.

Usually, simulations or visualizations require boundary representations of the buildings. The boundary representation (BRep) is probably the most widespread type of 3D representation and many algorithms are available for computing physical properties from that representation. Spatial objects are defined 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 order to transform the reconstructed CSG description into a boundary representation the union of the set of CSG primitives has to be computed. For this purpose the primitives are intersected, coplanar and touching faces are merged and inner faces or parts are removed. The boundary representation generated from the reconstruction presented in figure 6 is shown in figure 7.

4 INTERACTIVE REFINEMENT OF INITIAL RECONSTRUCTIONS

In our approach the reconstruction is constrained by the assumption that

- 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 itself. This can e.g. 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.

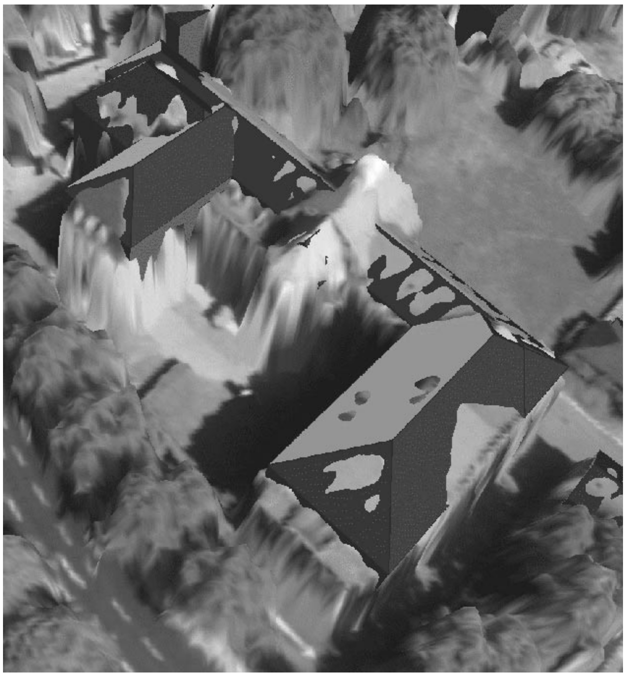
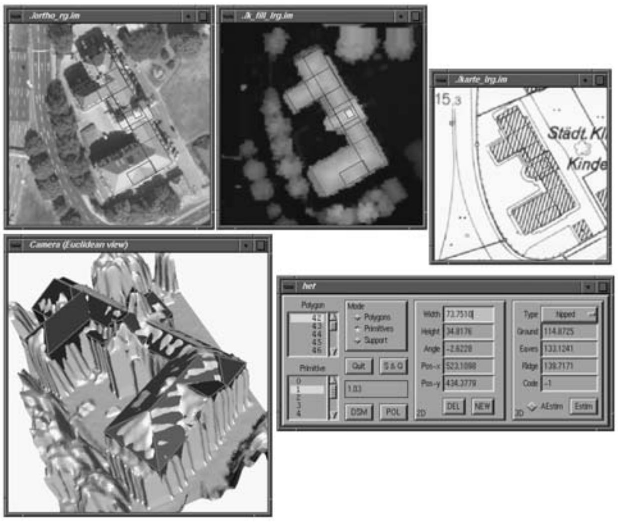


Figure 8: Reconstructed building and DSM surface.

Figure 8 shows the reconstructed building with the original 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 regions, which are incompatible with the final reconstruction give an additional hint, if manual interaction is required for further refinement. Since these regions are determined by the previous segmentation, they can be visualized together with the calculated RMS values in a final operator based evaluation step.

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 the ground plans 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 improvement of the reconstruction can be necessary for very complex buildings. This can be achieved if the initial reconstruction is analyzed in order to refine the capture of 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 9). 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.



1.03

Width 73.7510
Height 34.8176
Angle -2.6228
Pos-x 523.1098
Pos-y 434.3779

Type hipped
Ground 114.8725
Eaves 133.1241
Ridge 139.7171
Code -1

Figure 9: View of the interactive editing tool used to refine automatic house reconstructions.

Figure 10: Added ground plan segments.

Figure 10 shows the original ground plan (white polygon) and two black rectangles which were added interactively using the GUI, since both the segmentation into compatible regions presented in figure 5 and the comparison of the reconstructed building to the DSM surface presented in figure 8 give hint to parts of the building which are not reconstructed properly. These manually added rectangles trigger the automatic reconstruction of two additional building primitives, which represent a bay of the roof and the bell tower. The result of the final reconstruction is shown in figure 11.

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

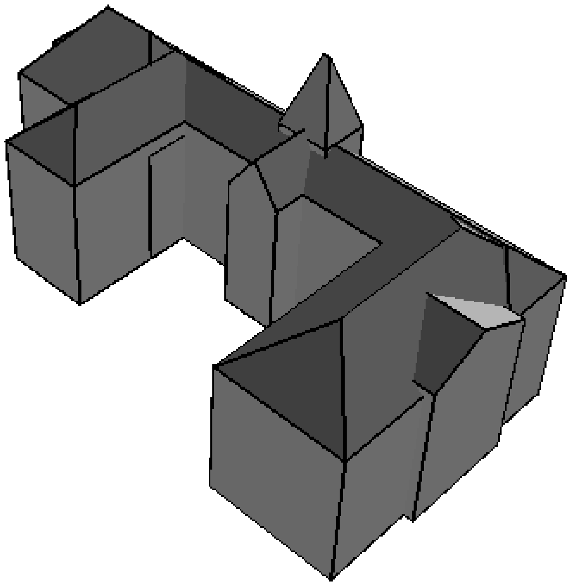


Figure 11: Restruction after manual addition of ground plan segments.

for geometric modeling, since the use of photo realistic texture enhances the perceived detail even in the absence of a detailed geometric model.

5 CLASSIFICATION OF URBAN VEGETATION

One problem of using existing databases is their incompleteness and potential lack of actuality. While aiming on the combination of a 2D GIS with directly captured data like DSM or images, a map revision has to be performed as a first step. For this purpose obsolete or incomplete parts of the GIS have to be identified. In addition to the detection or validation of inconsistencies between the datasets, there is a need to capture objects, which are not contained in the 2D GIS. While buildings and traffic network are available from standard databases, vegetation is usually not represented in detail. However, these objects are relevant for 3D site models. Urban vegetation like trees and bushes e.g. is an important feature for the analysis of landscape character and therefore has to be captured and represented for virtual reality applications. In order to automatically detect objects of interest like buildings, streets and trees in the DSM and multispectral images a classification approach is applied.

Figure 13 shows the result of the utilized ISODATA algorithm, which is used to discriminate the classes building, tree, street, grass-covered and shadow. The algorithm is described in more detail in (Haala et al., 1998). This classification was used to automatically determine the position of trees. In order to define the position of single trees within the classified regions a skeleton was calculated for these areas. The result is shown in figure 14. The outlines of the tree regions are represented by the white polygon, the skeleton of these regions is represented by the black lines. The trees were placed at the node points of the skeleton (grey points). Figure 12 shows the automatically reconstructed buildings and the virtual trees projected to the DTM surface, which was overlaid by the ortho image.

Figure 13: Result of classification using CIR ortho image and normalized DSM.

Figure 14: Computation of tree positions using the skeleton of the tree class segmentation result.

6 TEXTURE MAPPING

The basic goal of the algorithm 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 for texture processing. Since the vertices of the 3D building models generated from ground plans and laser data provide sufficient control point information, texture mapping using terrestrial images is simplified considerably compared to the standard terrestrial photogrammetric approach, where 3D information has to be obtained first using measurement of points in multiple images.

The goal of texture processing is to provide a rectified im-

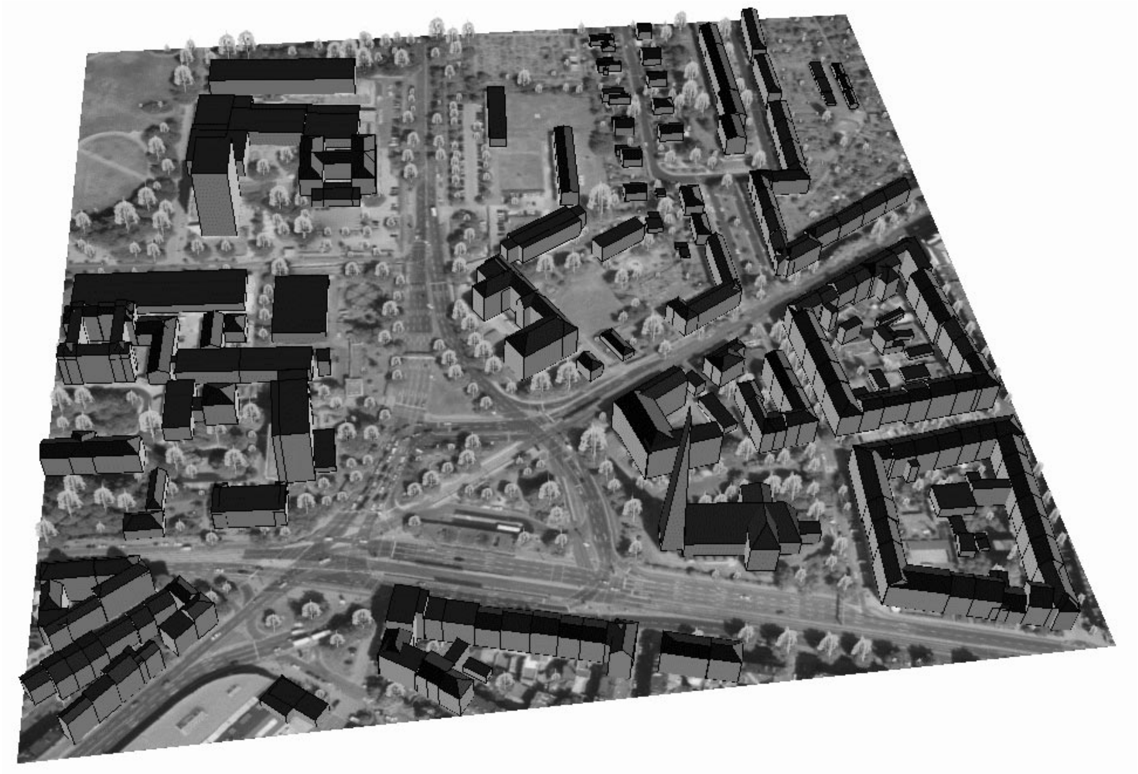


Figure 12: Reconstructed buildings and virtual trees projected to DTM.

Figure 15: Reconstructed buildings projected into stereo image.

age 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 are 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 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 based on the visualization presented in figure 15, at least four tie points between the face polygon and the terrestrial image have to be determined. These tie points are required in order to calculate the parameters of the perspective transformation during the rectification of the facade images. For this purpose the nodes of the face polygon have to be identified and measured in the terrestrial image. Since the points have to be positioned very precisely, the displayed texture can be scaled to any desired resolution.

Figure 16 shows an example of terrestrial imagery which was taken for texture mapping. The images were acquired with a Kodak DC 120 digital camera. This color camera features an interpolated resolution of 1280×960 pixels. A built-in color display allows the assessment of each image right after image take. Images can be stored on exchangeable flash cards; they can be downloaded onto a standard laptop using a serial cable or PCMCIA slot when needed. In figure 16 the points measured for the rectification are marked by white dots. If a node of the face is hidden by an obstacle, the corner point can be alternatively calculated from points measured at the edge of the facade. The size of the rectified image is determined by the size of the facade, which is provided by the geometric model and a predefined texture pixel size. These images then can be assigned to their corresponding faces in the 3D model. In order to control the correct assignment of the texture the 3D buildings are dis-

Figure 16: Original terrestrial images with points measured for rectification.

played in a viewer which allows all basic transformations to the model (translation, rotation, scaling) immediately after texture mapping. For the final visualizations presented in figures 17 and 18, the ortho image as well as virtual trees are added.

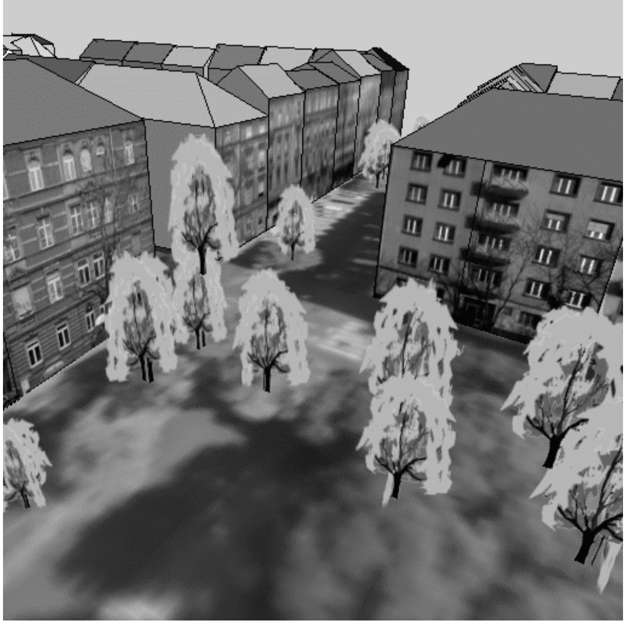


Figure 17: 3D visualization of virtual city model.

7 CONCLUSION

We have presented an approach which uses 2D ground plans and a laser DSM to derive 3D building geometry automatically, whereas texture is taken from terrestrial photographs. By applying the proposed approach the automatic extension of existing 2D GIS data to a 3D urban model is available; the raster representation of airborne laser scanner data can be transformed to a qualified 3D vector representation for urban areas. We have also developed a tool to modify automatically obtained 3D reconstructions interactively in order to remove reconstruction errors and to add detail. Additionally, we have shown classification results from CIR images and height data. This classification can e.g. be used to detect obsolete parts within the 2D GIS information or to provide the location of trees, which are important for virtual reality applications but are usually not contained in the GIS data.

Figure 18: 3D visualization of virtual city model.

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 photorealistic 3D virtual reality city models. In our opinion this demand can only be satisfied by a highly efficient capture of urban scenes, which presumes the integrated use of multiple data sources.

REFERENCES

Besl, P. J., 1988. Segmentation through variable order surface fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence 10(2), pp. 167–192.

Haala, N., Brenner, C. and Stätter, C., 1998. An integrated system for urban model generation. In: Proc. ISPRS Congress Comm. II, Cambridge.

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.