

RAPID 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, Norbert.Haala@ifp.uni-stuttgart.de

ABSTRACT

The growing demand for detailed city models has stimulated research on efficient 3D data acquisition. Over the past years, it has become evident that the automatic reconstruction of urban scenes is most promising if different types of data, possibly originating from different data sources are combined. In the approach presented in this paper the geometric reconstruction of urban areas is based on height data from airborne laser scanning and 2D GIS, which provides the ground plan geometry of buildings. Both data sources are used to estimate the type and parameter of basic primitives which in turn are combined to obtain complex building structures. The final output consists of 3D CAD models for the buildings. Using the reconstructed geometry, terrestrial images are mapped onto building facades to generate virtual city models.

ZUSAMMENFASSUNG

Die steigende Nachfrage nach detaillierten Stadtmodellen erfordert die Entwicklung von Methoden zur effizienten 3D Datenerfassung. Forschungsaktivitäten der letzten Jahre haben gezeigt, daß die automatische Erfassung städtischer Szenen am erfolgversprechendsten ist, wenn unterschiedliche Datentypen verwendet werden, die möglicherweise aus unterschiedlichen Quellen stammen. Der in diesem Artikel vorgestellte Ansatz zur geometrischen Rekonstruktion von Stadtgebieten verwendet Höhendaten eines Laserscanners und ein 2D GIS, das die Grundrisse der Gebäude zur Verfügung stellt. Beide Datenquellen werden verwendet, um die Form einfacher Grundelemente zu bestimmen, die anschließend zu komplexen Gebäudestrukturen zusammengesetzt werden. Ausgegeben werden 3D CAD Modelle der Gebäude. Zur Erzeugung eines virtuellen Stadtmodells werden abschließend terrestrische Aufnahmen der Fassaden auf die rekonstruierten Gebäude projiziert.

1 INTRODUCTION

Simulations that require the representation, management and analysis of three-dimensional descriptions have become standard applications for planning purposes in built-up areas. 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 enable a photo realistic visualization. Objects, which have to be represented by a virtual reality city model are buildings, streets, a Digital Terrain Model (DTM) and urban vegetation like trees or bushes. Because manual interpretation during data acquisition is very time consuming, a lot of effort has been spent to speed up data collection by automatic or semi-automatic procedures. A good overview of the current state of the art is given in (Baltsavias et al., 1997) and (Gruen et al., 1997)

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. Hence, in our approach Digital Surface Models (DSM) are combined with existing 2D GIS and color images for data collection in urban environments. A DSM provides a geometric description of the terrain surface including objects like trees or buildings, which rise from the ground. We apply DSM from airborne laser scanners, since - compared to stereo image matching - the direct measurement of height data provides a

DSM of high and homogeneous quality in urban areas, which is essential for our approach. Despite the high quality of the laser data, their automatic interpretation 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 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 control or revision of the existing database has to be performed to uncover obsolete or incomplete parts of the GIS or map. Additionally, urban vegetation like trees or bushes, which is very relevant for 3D site models has to be acquired. The automatic extraction of objects like buildings, streets and trees is realized within a classification step, which combines the DSM with multispectral images.

To achieve photo realism an accurate geometric and textural description is required for each building. Since real world imagery embodies a representation of object detail, it can even substitute for geometric modeling. To a certain degree non-planar surfaces can be represented by texture mapped planar surfaces without degrading the impression of viewing a 3D scene. For this reason the use of photo realistic texture enhances the perceived detail even in the absence of a detailed geometric model. 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. Additionally, due to the viewpoint, facades, jut out of roofs or passages are hardly visible or even totally occluded. Furthermore, the limited resolution of airborne data restricts the amount of detail and the accuracy. For these reasons, photo realism can only be achieved, if terrestrial images utilized.

After discussing the potential of airborne laser scanning for 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 optionally to modify or refine building reconstructions obtained automatically (section 4). In case no ground plans are available this interactive mode can also be used for semi-automatic reconstruction. In the last part of the paper a texture mapping based on terrestrial images of building facades for the generation virtual reality city models is presented.

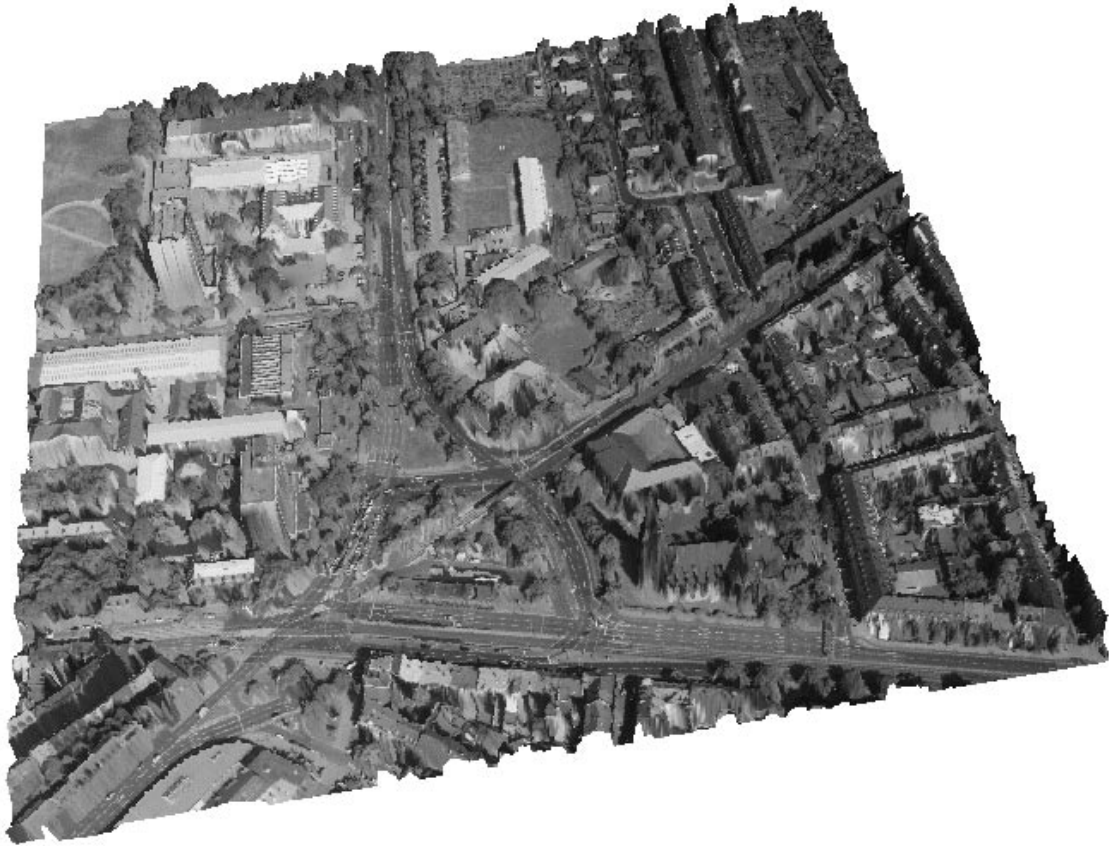


Figure 1: 3D visualization of DSM overlaid with CIR ortho image.

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 $1 \times 1 \text{ m}^2$ with an accuracy of 0.3 m in planimetry and 0.1 m in height. For visualization a color-infrared 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 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 corners and edges 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 a very general building model should be used. Therefore, the building is represented by a general polyhedron, i.e. it is bounded by a set of planar surfaces. Additional constraints are provided in our approach 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 GIS. Assuming a correct and actual database these ground plans are very reliable compared to results of automatic procedures, 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 - after a control for gross errors - 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. The digital cadastral map is build up as an area covering data base, mainly by digitizing existing maps or plans. 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. Alternatively, maps and plans can be digitized automatically (Frischknecht and Carosio, 1997) to obtain the required information without human interaction.

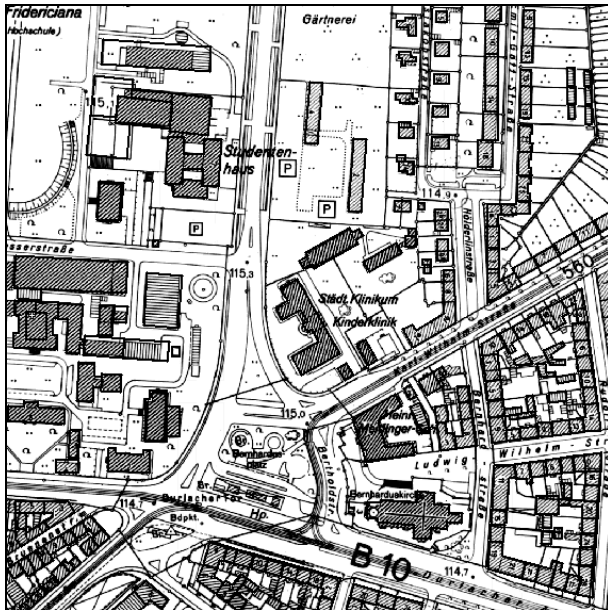


Figure 2: Map 1:5000 and extracted ground plans.

3.1 Decomposition of ground plans

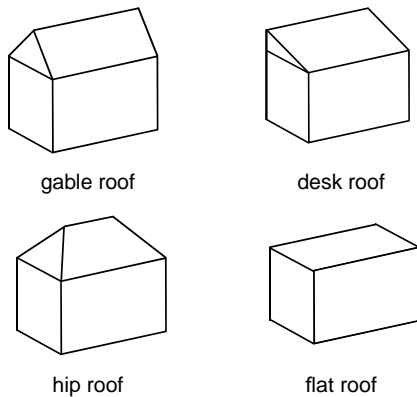


Figure 3: Building primitives used for reconstruction.

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.

In the first step the buildings are segmented into basic primitives based on their given outlines. Figure 4 shows an example of the automatic ground plan decomposition. Each generated rectangle defines the base of one building primitive, i.e. position, orientation and horizontal extension of each cuboid are already defined by



Figure 4: Ground plan decomposed into rectangular parts.

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 in order to fit the building primitives to that surface.



Figure 5: Segmented roof regions.

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. For this purpose the DSM is segmented into planar surfaces by analyzing 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 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° , 90° , 180° and 270° to the main axis of the building.

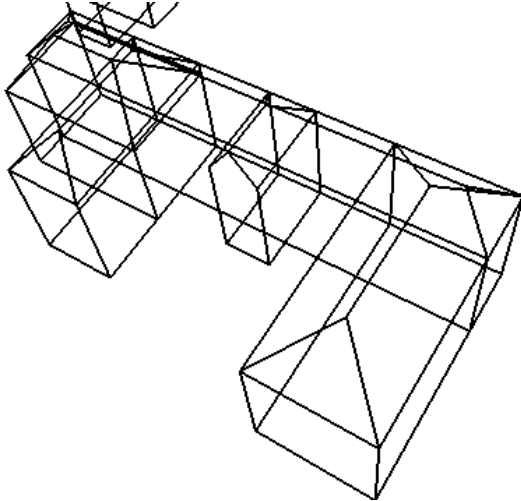


Figure 6: Reconstructed CSG building primitives.

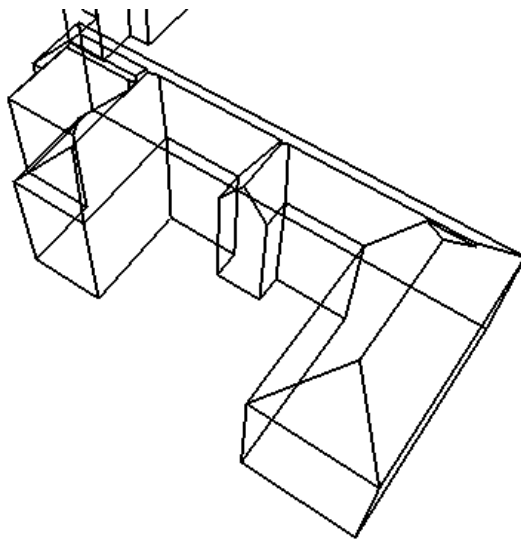


Figure 7: CSG data converted to boundary representation.

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.

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. 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 based on ground plans, which were digitized merely from the map. No care has been taken to digitize these building outlines 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.

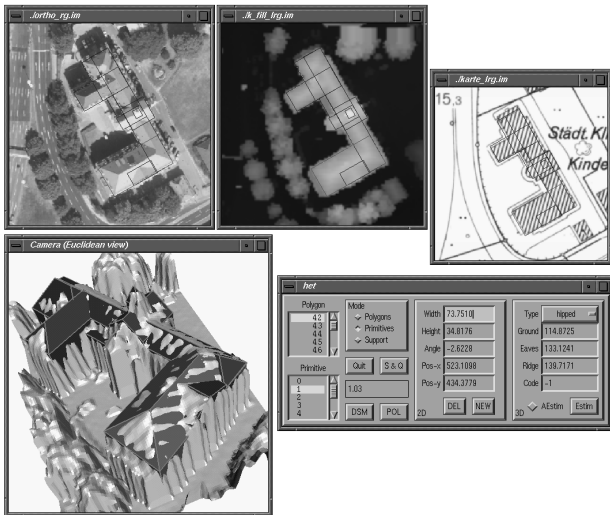


Figure 9: View of the interactive editing tool used to refine automatic house reconstructions.



Figure 10: Added ground plan segments.

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

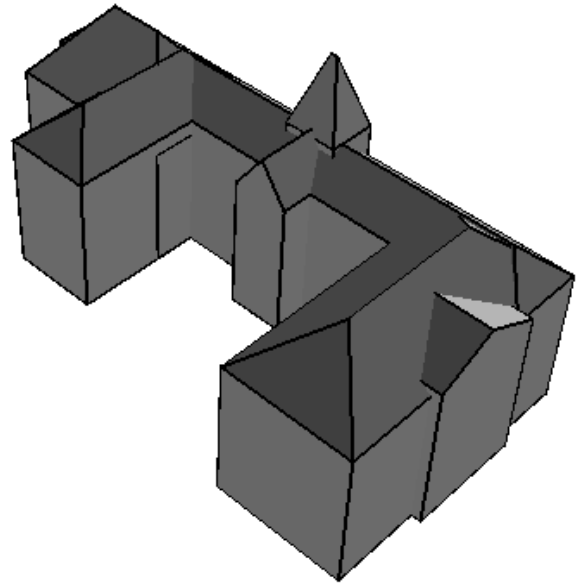


Figure 11: Reconstruction after manual addition of ground plan segments.

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 10 shows the original ground plan (white polygon) and two black rectangles which were added interactively using the GUI. At these areas both the segmentation into compatible regions presented in figure 5 as well as the comparison of the reconstructed building to the DSM surface presented in figure 8 give hints to an improper reconstruction. These rectangles added manually trigger the automatic reconstruction of two red building primitives representing a bay of the roof and the bell tower. The result of the final reconstruction is shown in figure 11.

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. For that purpose three digital images in the spectral bands near infrared, red and green were obtained by scanning a color-infrared (CIR) aerial image. The algorithm is described in more detail in (Haala et al., 1998b).

Figure 13 shows the result of the utilized ISODATA algorithm, which is used to discriminate the classes building, tree, street, grass-covered and shadow. This classification can also be used to auto-

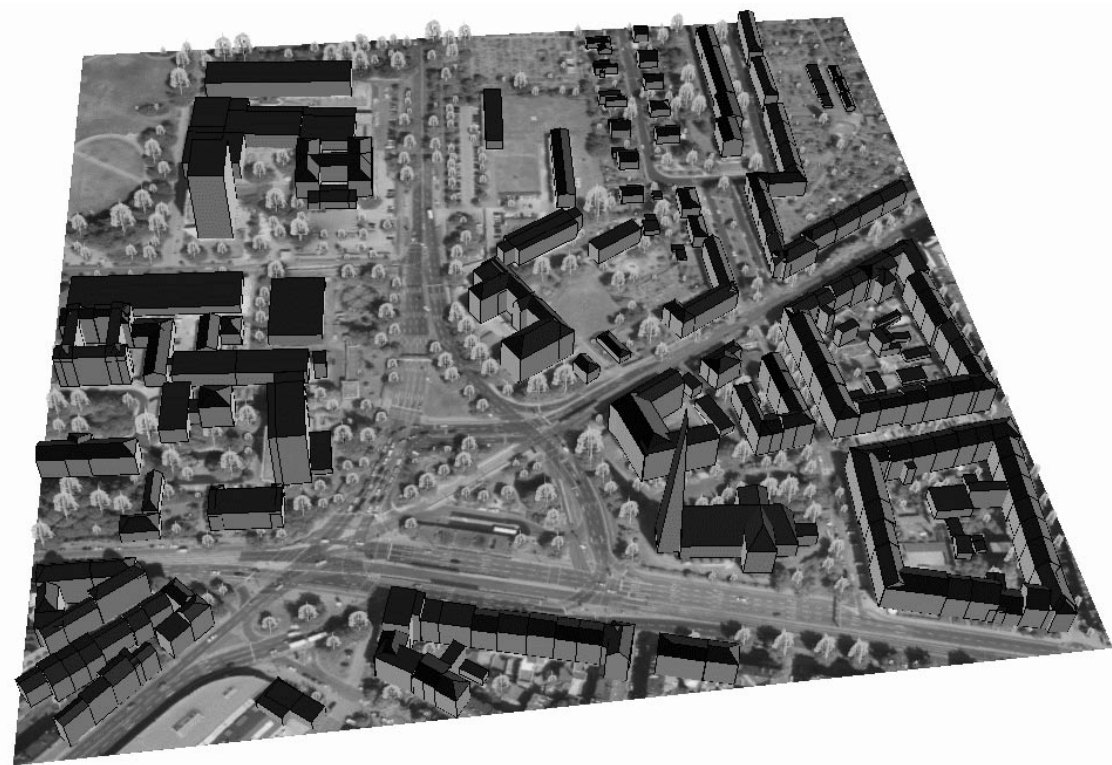


Figure 12: Reconstructed buildings and virtual trees projected to DTM.



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

matically determine the position of single trees. In order to define their position within the classified regions, a skeleton was calculated for these areas (Arcelli and di Baja, 1985). In figure 14 the outlines of the classified tree regions are represented by the green polygon, the skeleton of these regions is represented by the yellow lines. The trees were placed at the node points of the skeleton (blue 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 14: Computation of tree positions using the skeleton of the tree class segmentation result.

6 TEXTURE MAPPING

Detailed and accurate CAD building models are sufficient for many applications including wave propagation simulations. However, if photo realism is to be achieved for visualization purposes, either artificial texture or real world imagery has to be mapped to building facades and roofs. Additionally, the size of object parts which can be reconstructed is of course limited by the available density of

DSM points, details smaller than approximately 1 m can not be captured from the current dataset. 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.

By using DSM and ground plans for geometric processing and terrestrial images for texture processing the time consuming process of virtual city model creation can be simplified considerably compared to the standard architectural photogrammetry approach, since the corners of the already reconstructed buildings provide sufficient control point information.



Figure 15: Reconstructed buildings projected into stereo image.

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 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 the facade of the building is rectified by applying a projective transformation. 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 have to be determined. For this purpose the corners of the face polygon have to be identified and measured in the terrestrial image.

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



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

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 displayed 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

Within the article an approach has been presented, which uses 2D ground plans and a laser DSM to derive 3D building geometry automatically, whereas texture is taken from terrestrial photographs. 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. By applying the proposed approach the automatic extension of existing 2D GIS data to a 3D urban model is available; the raster rep-



Figure 18: 3D visualization of virtual city model.

Sensing Conference and Exhibition, Vol. I, Copenhagen, Denmark, pp. 174–180.

resentation of airborne laser scanner data can be transformed to a much more qualified 3D vector representation for urban scenes.

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. 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

Arcelli, C. and di Baja, G. S., 1985. A width-independent fast thinning algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Baltsavias, E., Eckstein, W., Gülch, E., Hahn, M., Stallmann, D., Tempfli, K. and Welch, R. (eds), 1997. *3D Reconstruction and Modelling of Topographic Objects Automatic Extraction of Man-Made Objects*. IAPRS, Vol. 32, Part 3-4W2.

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

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.

Gruen, A., Baltsavias, E. and Henricson, O. (eds), 1997. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Vol. II, Birkhäuser Verlag, Basel, Boston, Berlin.

Haala, N., Brenner, C. and Anders, K.-H., 1998a. 3D Urban GIS from Laser Altimeter and 2D Map Data. In: *IAPRS*, Vol. 32, Part 3/1, Columbus, Ohio, pp. 339–346.

Haala, N., Brenner, C. and Stätter, C., 1998b. An integrated system for urban model generation. In: *Proc. ISPRS Congress Comm. II*, Cambridge, pp. 96–103.

Lohr, U., 1997. Digital elevation models by laserscanning: Principle and applications. In: *Third International Airborne Remote*