RAPID ACQUISITION OF VIRTUAL REALITY CITY MODELS FROM MULTIPLE DATA SOURCES

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 V, Working Group V/3

KEY WORDS: city models, virtual reality, laser scanning

ABSTRACT

Virtual reality applications in an urban environment presume the acquisition of 3D city 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. The presented approach aims on an efficient generation of virtual city models by the combination of multiple data sources to benefit from the complementary types of information. The geometric processing, i.e. the acquisition of 3D building representations is performed fully automatic using Digital Surface Models from airborne laser scanning and existing ground plans of the building. The texture processing, i.e. the mapping of terrestrial images to the corresponding faces is supported by 3D building models, which were generated in the previous step. Due to the availability of building geometry, the manual interaction for the processing of terrestrial images can be reduced considerably. Thus, the rapid acquisition of urban virtual reality models is feasible.

1 GENERATION OF VIRTUAL REALITY CITY 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. Architectural photogrammetry enables a very detailed and accurate acquisition of buildings, since by photogrammetric measurement using stereo or convergent terrestrial imagery a precise geometrical reconstruction of facades with a high level of extracted detail can be achieved. No limiting assumptions on facade geometry have to be made for this approach. Additionally, structure and texture can be extracted at the same time from the same data source.

Architectural photogrammetry requires classical photogrammetric processing of the terrestrial images, i.e. the reconstruction of the interior and exterior orientation as well as the photogrammetric restitution for 3D point measurement by a bundle adjustment. Typically, for this purpose at least three control points and 5-10 tie points per image are required; the more control and tie points are available, the better the results of the orientation process in terms of accuracy and reliability can be obtained. As tie and control points have to be identified and measured manually, a lot of interactive work is required. Generally, the terrestrial acquisition of the building geometry proves to be a time-consuming process that has to be carried out for each building. For this reason these techniques are mainly reasonable for the detailed and accurate collection of data for individual buildings or single building blocks. Furthermore terrestrial techniques are mainly capable for the acquisition of facade, but are unsuitable for the measurement of roof shapes, which is an important feature for many applications.

Data acquisition or revision solely based on terrestrial data is not economical. This is further emphasized by the fact that the high geometric accuracies, which can be obtained with close range photogrammetric methods might not be required for visualization tasks. For most photo realistic applications, building blocks with individual roof shapes are sufficient. This information can be provided from airborne data like stereo images or Digital Surface Models (DSM), i.e. this data is more appropriate while aiming on an area covering acquisition of virtual city models.

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.

Referring to the amount of detail on the one hand and the potential for area covering data collection on the other hand, terrestrial and airborne techniques are complementary. For this reason, both data sources should be combined to get optimal results. We use existing ground plans extracted from maps and Digital Surface Models provided from airborne laser scanners for the acquisition of the building geometry. The terrestrial images are only used for texture processing, i.e. the mapping of the facades. The paper is organized in two parts. First the geometric processing using ground plans and DSM data is described. Afterwards the interactive approach for the texture mapping of building faces is presented.

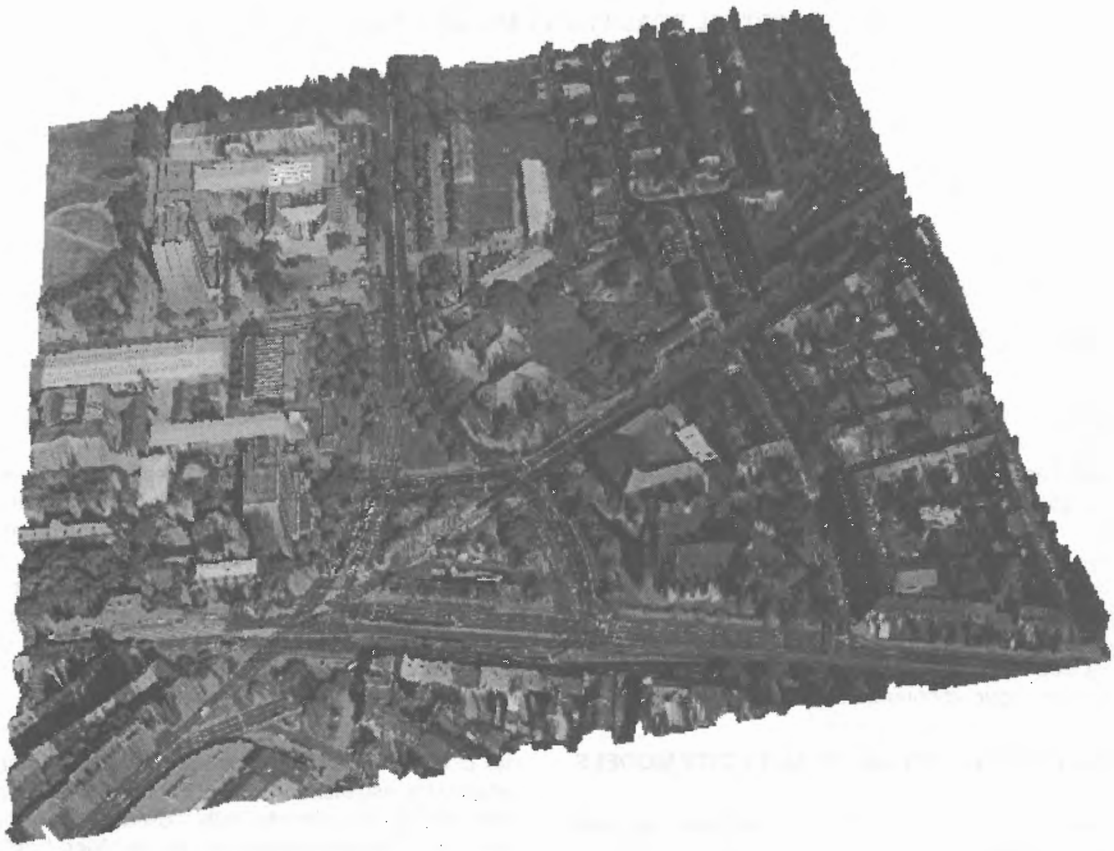

Figure 1: 3D visualization generated from DSM and ortho image.

2 GEOMETRIC PROCESSING

Due to the very pretentious problem of automatic interpretation of DSM data and due to the great effort in gaining sufficient spatial resolution and quality of the DSM especially for interpretation tasks, supplementary sources of information have to be used to obtain optimal results. Within the approach presented in this paper, the segmentation of planar surfaces from the DSM as well as the decomposition of complex buildings into basic primitives, which are prerequisites for the reconstruction of the buildings, are supported by given ground plans.

2.1 Building reconstruction from ground plans and DSM

Ground plans are a very reliable source of information for 3D building reconstruction. Frequently, these ground plans have already been acquired and are represented either in analog form by maps and plans or digitally in 2D Geo Information Systems (GIS). An example for this kind of data 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. At the moment the digital cadastral map is build up as an area covering data base, mainly by digitizing existing maps or plans. Currently, it covers 40% 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. Alternatively, maps and plans can be digitized automatically (Frischknecht and Carosio, 1997) to obtain information similar to the digital cadastral map without human interaction.

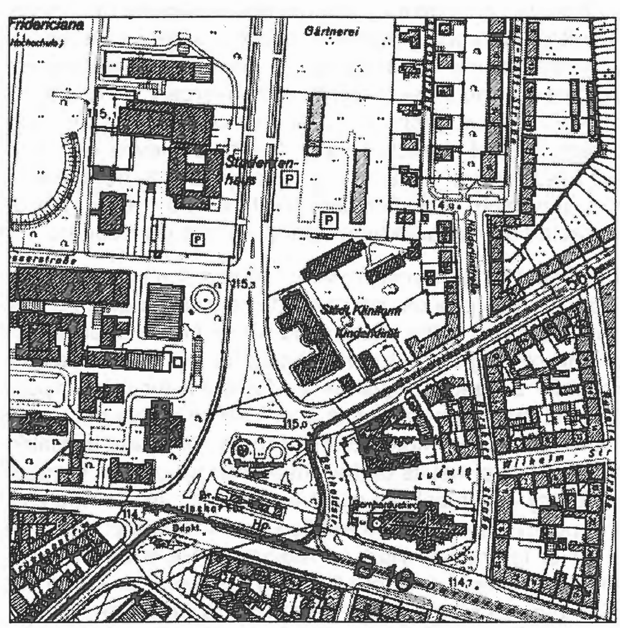

Figure 2: Map 1:5000 and extracted ground plans.

Since 3D information usually is not available from plans and maps other data sources like aerial images or Digital Surface Models have to be used for that purpose. As the information of a Digital Surface Model is restricted to surface geometry, the interpretation of this kind of data is easier compared to the interpretation of (stereo) image data. A DSM, i.e. a geometric description of the terrain surface and objects located on and above this surface like trees or buildings can e.g. be obtained by airborne laser scanning systems. By these systems the coordinates of terrain points are determined by polar measurement similar to a tachymetric data acquisition. Therefore position and orientation of the laser beam at the time of range measurement have to be provided by additional components of the sensor system. These components are a NAVSTAR Global Positioning System (GPS) for the positioning task and an Inertial System (INS) for the orientation task. Airborne laser scanners provide DSM of high and homogeneous quality since three-dimensional points on the sensed surface can be determined dense and well-distributed. Hence this data is very suitable for 3D building reconstruction. Figure 1 shows a DSM of our test area, provided by the TopoSys laser scanner-system (Lohr, 1997). The area covers an area of $500 \times 500\text{m}^2$ of the city of Karlsruhe, Germany. 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. For visualization purposes the ortho image of the test area was mapped onto the measured surface. This view gives a good impression of the scenery, since the geometry of the surface is represented quite well.

One problem whith using surface descriptions by an unqualified, i.e. object independent distribution of points like they are generated by laser scanning is the large amount of data to be stored, computed and displayed. The resulting great computational effort prevents the use of this type of representation for real time visualizations. One approach in order to reduce the large amount of data to enable real time display is to generate compact polygonal approximations of the surface shape by surface simplification algorithms. A good overview on existing approaches can be found in (Heckbert and Garland, 1997). Even though the amount of data can be reduced significantly by these algorithms, many tasks aiming on visualizations or simulations in an urban environment require the further abstraction and interpretation of the surface description. For simulations on the propagation of noise or electro-magnetic waves the knowledge on the surface material is a crucial point. Hence trees or buildings have to be represented separately from the terrain surface. As discussed earlier, for the generation of walk-troughs terrestrial images have to be mapped onto vertical faces of the buildings in order to achieve a realistic appearance. For this purpose corresponding points have to be determined between the terrestrial images and the data set used to represent building geometry. Since nodes and vertices of a building can be often identified easily in terrestrial images, they should also be represented explicitly in the geometric database. Hence, for the generation of virtual city models the geometry and topology of buildings should be described by 3D CAD models.

2.2 Estimation of CSG building models

In our approach, two-dimensional ground plans of buildings are used to define building boundaries. The building model is constrained by the assumption that the coordinates of the given ground plan are correct and the borders of the roof are exactly defined by this ground plan. The three-dimensional structure of the building, which consists of a roof type, ridge and eaves heights is derived automatically from an aerial laser DHM. In order to deal with the large architectural variations of building shapes, the utilized model should be as general as possible. Therefore, a building is represented by a general polyhedron, i.e. it has to be bounded by a set of planar surfaces and straight lines.

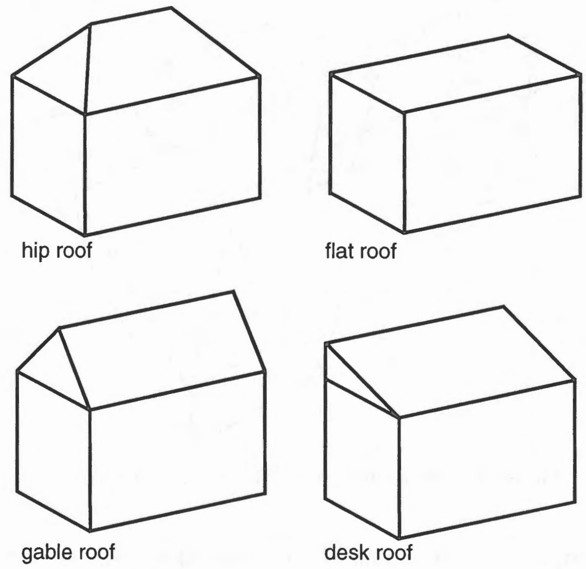

Figure 3: Building primitives used for reconstruction.

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 3. Each building primitive consists of a cuboid element with different roof types. Currently flat roofs, desk roofs, gable roofs and hip roofs can be represented.

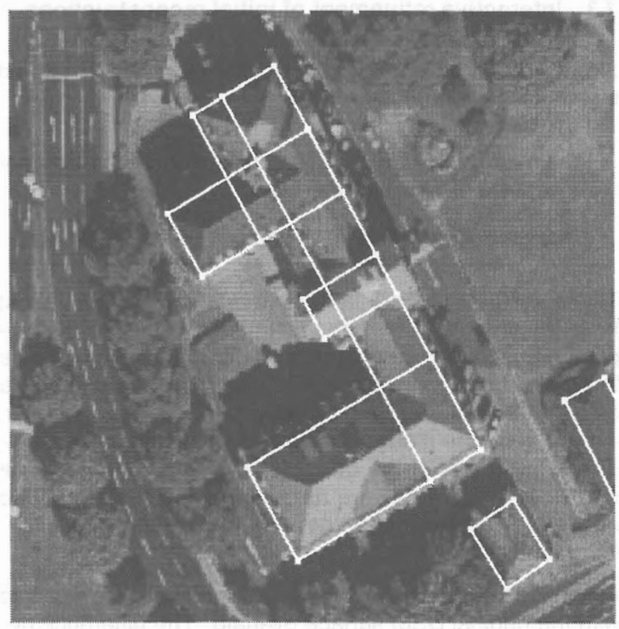

Figure 4: Ground plan decomposed into rectangular parts.

First the complete building has to be split up into these basic structures. This step can be realized fully automatic by the analysis of the given ground plan. Figure 4 shows the automatic decomposition of a complex ground plan into rectangular structures. Each of these rectangles defines one building primitive. Since position, orientation and horizontal extension of each cuboid is already defined by a

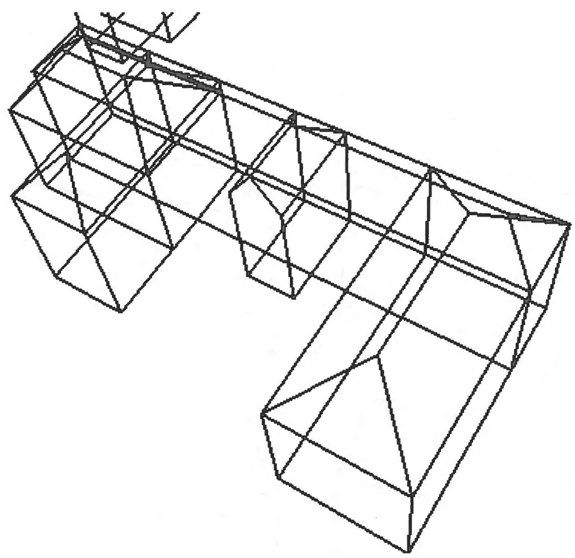

Figure 5: Reconstructed CSG representation.

rectangle, only the height of each cuboid as well as roof type and roof slope have to be determined as remaining parameters for each building primitive. The parameters of the building primitives are estimated by a least squares adjustment which minimizes the distances between the DSM surface and the corresponding points of the building primitive, i.e. the building primitives are fit into the DSM surface. In figure 5 the reconstructed building primitives are represented by a wire frame. The single overlapping CSG primitives are clearly visible.

2.3 Interactive refinement of initial reconstructions

In our current approach the reconstruction is constrained by the assumptions 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 rectangles.

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 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. Also, we currently do not deal with triangular or circular ground plans.

Figure 7 shows a boundary representation of the 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 to the final reconstruction give an additional hint if manual interaction is required for further refinement. These regions are determined within the reconstruction step and are visualized in a final operator based evaluation step. Within this step, the reconstruction of an additional building primitive representing the bay of the roof can

Figure 7: Reconstructed building and DSM surface.

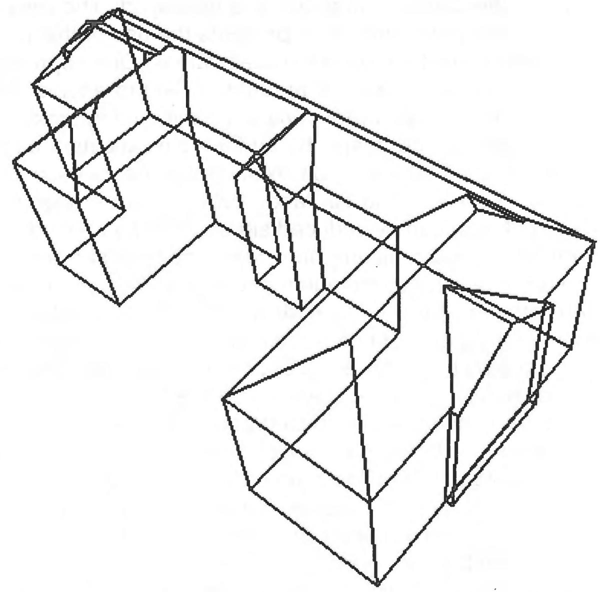

Figure 8: CSG representation of figure 5 transformed to B-Rep (additional bay from interactive refinement).

be triggered. The result of the final reconstruction is shown in figure 8, where the boundary representation already has been generated. Using interactive modelling, smaller features like ledges or balconies can also be added to the data.

2.4 Classification of surface material

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

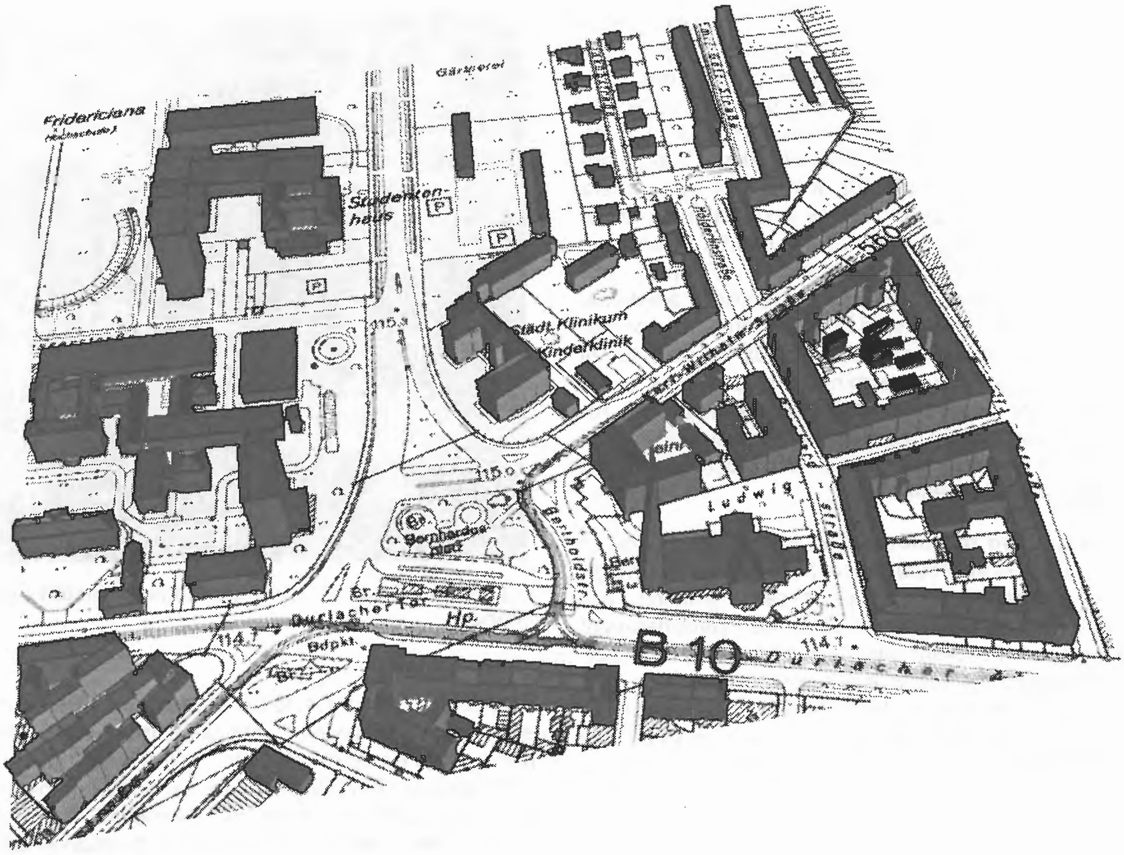

Figure 6: Reconstructed buildings projected on map 1:5000 used for ground plan acquisition.

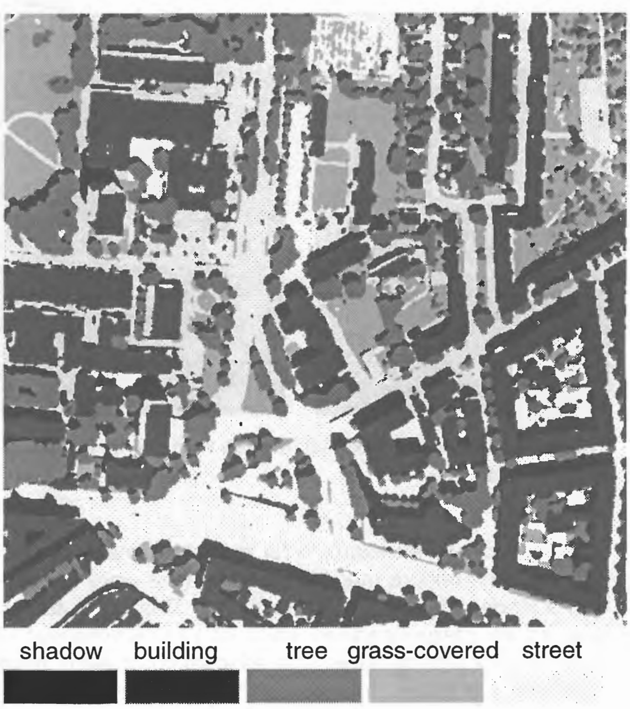

Figure 9: Result of classification using CIR ortho image and normalized DSM.

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 9 shows the result of the utilized ISODTA 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).

3 CAD-BASED TEXTURE PROCESSING

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, which are generated from ground plans and laser data provide sufficient control point information, texture mapping from terrestrial images is simplified considerably. Therefore, the generation of virtual reality models is more efficient compared to the standard architectural photogrammetry approach, where a number of tie points has to be measured in multiple images.

3.1 Texture Mapping

For texture mapping the image has to be correctly positioned, oriented and scaled to represent its associated surface. There are two approaches for texture mapping:

- **Orthophotography** is a parametric approach since each object point is transformed into the image by rigorous application of the orientation parameters. The interior orientation and the exterior orientation, i.e. position and viewing direction of the camera in a reference coordinate frame are determined in a bundle adjustment. If the zoom factor or focus of the camera has been changed during image acquisition, an independent set of parameters for the interior orientation has to be determined for each image. Within this approach, lens distortion can be corrected if additional parameters are determined in the estimation process. If two or more images are available, 3D coordinates of object points can be determined on demand, e.g. for the control or rectification of the geometric model. Attempts have already been made to refine initial CAD models, which were generated from existing plans or models (Streilein, 1994)
- The **projective transformation** is a simplification of the parametric approach described above. An image section representing a planar surface is rectified by applying a projective transformation. The 7 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.

In both cases, a precise 3D model of the surface to be texture-mapped is needed. When generating building facade models, first problems are encountered when making the terrestrial images. In theory, it is not necessary for texture mapping that images should be takne parallel to the facade. However, in practice it is often advantageous. The quality of original data capture is most important, and truly parallel images will always be a more productive source than significantly tilted photography. This is especially true for irregular facades (Dallas, 1996). Unfortunately, if the streets are rather narrow, it is difficult to select viewpoints where the complete facade is visible. For stereo processing at least two images have to be captured, which makes it even more difficult. Hence, we decided to use single images in our system and restrict their application to the generation of texture. The main advantage of the projective transformation is the small number of parameters, which have to be determined, i.e. the number of image points, which have to be identified and measured is limited. Since this is an important factor for the economic generation of city models, this approach was chosen for our system.

3.2 Matching of building faces

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 10). If the viewpoints are sketched into a map or an ortho image during the terrestrial image acquisition, this representation enables a simple interactive definition of the corresponding face index for each terrestrial image.

After the selection of the corresponding 3D building face, at least four tie points between the face polygon and the

Figure 10: Reconstructed buildings projected into stereo image.

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 accurately, the displayed texture can be scaled to any desired resolution.

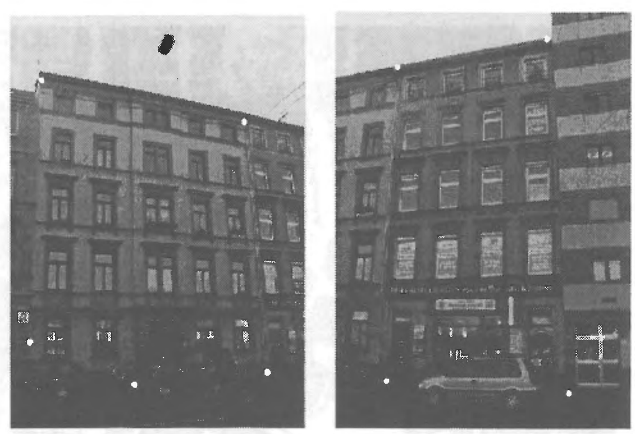

Figure 11: Original terrestrial images with points measured for rectification.

Figure 11 shows an example for 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 1280×960pixels. A build-in color display allows the assesment of each image right after it is taken. 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 11 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.

Figure 12 shows the result of the rectification process. The size of the rectified image is determined by the size of the

Figure 12: Rectified facade images.

Figure 13: Faces mapped to corresponding 3D surfaces.

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 displayed in a viewer which allows all basic transformations to the model (translation, rotation, scaling) immediately after the texture mapping (see figure 13). For the final visualizations (see figures 14 and 15), the ortho image as well as artificial trees, which were generated based on the classification described in section 2.4 are added to the virtual model.

4 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 photorealistic 3D virtual reality city models. In our opinion this demand can only be satisfied by a highly efficient capture of urban

Figure 14: 3D visualization of virtual city model.

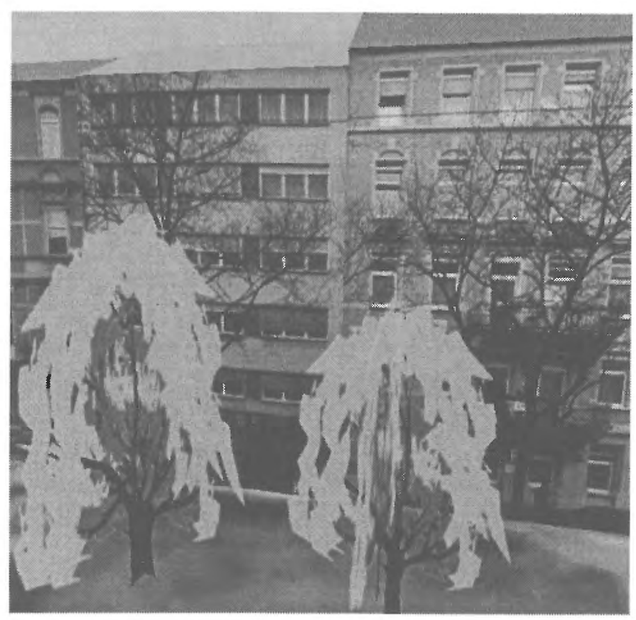

Figure 15: 3D visualization of virtual city model.

scenes, which presumes the integrated use of multiple data sources. 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. Thus, the time consuming terrestrial measurement of building geometry can be avoided.

REFERENCES

Dallas, R., 1996. Architectural and archeological photogrammetry. In: K. Atkinson (ed.), Close Range Photogrammetry and Machine Vision, Whittles Publishing.

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, 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.

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

Heckbert, P. and Garland, M., 1997. Survey of polygonal surface simplification algorithms. In: SIGGRAPH '97 Course Notes Multiresolution Surface Modeling, Los Angeles.

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.

Streilein, A., 1994. Towards automation in architectural photogrammetry: CAD-based 3D-feature extraction. ISPRS Journal of Photogrammetry and Remote Sensing 49(5), pp. 4–15.