AN INTEGRATED SYSTEM FOR URBAN MODEL GENERATION

Norbert Haala, Claus Brenner and Christian Stätter
Institut für Photogrammetrie
Universität Stuttgart
Geschwister-Scholl-Straße 24, 70174 Stuttgart, Germany
Ph.: +49-711-121-3383, Fax: +49-711-121-3297
e-mail: Norbert.Haala@ifp.uni-stuttgart.de

Commission II, Working Group 6

KEY WORDS: GIS, Height Data, Multispectral, Urban Models, Data Fusion

ABSTRACT

Many tasks in the context of urban planning presume 3D city models, where buildings are represented separately from the terrain surface e.g. by 3D boundary representations. One of the most promising approaches to solve the challenging task of automatic 3D city model acquisition is the integration of multiple data sources to benefit from the complementary types of information. In our approach height data and existing ground plans of buildings and multi-spectral images are combined in an integrated system for the generation of 3D urban models.

1 INTRODUCTION

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. Objects relevant for the 3D description of urban areas are buildings, streets, trees and Digital Terrain Models (DTM).

Because manual interpretation is very time consuming, a lot of effort has been spend 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. The utilized data sources are existing 2D GIS, color images and Digital Surface Models. Height data provided by a Digital Surface Model from airborne laser scanning is combined with multispectral information from color images in order to classify urban areas into regions like building, street, tree or grass-covered area. This classification can e.g. be used to control the completeness and actuality of existing ground truth data. In the second part of the paper the use of laser scanner DSM and ground plans from existing 2D GIS or maps for 3D building reconstruction will be demonstrated.

2 DATA SOURCES FOR URBAN MODEL GENERATION

In principle, airborne stereo image data is sufficient for the 3D acquisition of buildings; a human operator is able to extract and reconstruct visible buildings solely using this data source. However, due to the great complexity of image data the automation of this process poses many problems. Greyvalues are influenced by the object geometry but also by factors like illumination, surface material or texture. The large amount of information which is contained in images and the numerous factors influencing a greyvalue make it very difficult to separate important information from irrelevant details in the framework of an automatic image interpretation. For this reason alternative data sources are utilized in our approach to enable the automatic 3D urban model generation. These data sources are multispectral images, Digital Surface Models and 2D GIS information on the ground plan of buildings, which – at least for highly developed countries – is available for most of the cities.

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 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. Even though this technique has become a standard tool for 3D surface acquisition in open terrain, it still suffers 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. Alternatively airborne laser scanning can be applied. Theses systems consist of components of kinematic GPS positioning, inertial attitude determination and laser ranging for the purpose of measuring XYZ coordinates of the points of reflection of the emitted laser pulses. The system accuracy is in the order of 10 cm in all 3 coordinates. The laser scanners which are presently in use have high data rates of up to 8×10^4 Hz, yielding dense distributions of terrain points. From 1000 m flying height average distances between points on the ground down to several points per square meter can be obtained. The direct height data acquisition by airborne laser scanners provides DSM of high and homogeneous quality in urban areas since three-dimensional points on the sensed surface are determined dense and well-distributed. 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. This DSM will be used throughout this paper as an example dataset.

Frequently, 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 acquired by a 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 costs to attain a solution.

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

Figure 1: Digital Surface Model acquired by laser scanning

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 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 Germany. Since this type of data was not available for our test area 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) which also results in information similar to the digital cadastral map.

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. Urban vegetation like trees and bushes is an important feature for the analysis of landscape character and therefore has to be captured and represented for virtual reality applications. With the availability of multispectral imagery opportunities exist to exploit spectral information to aid urban scene analysis for cartographic feature extraction. By traditional spectral classification techniques surface material information can derived from the multispectral imagery. This surface material information, which is normally highly correlated with the object type in complex urban scenes, can be used as a source of information for mapping of man-made structures and natural features. In our approach the classification of objects such as buildings and roads, as well as natural features like trees is further improved by combining multispectral images and DSM data.

3 CLASSIFICATION USING DSM AND COLOR IMAGERY

DSM have already been used for the automatic detection of buildings e.g. to trigger the subsequent geometric reconstruction from stereo image data (Haala, 1994) (Baltsavias et al., 1995). This possible since DSM do not only contain information about the topographic surface like Digital Elevation Models (DEM), but also about buildings and other object higher than their surrounding. To make this information accessible the so-called normalized DSM, i.e. the difference between DSM and Digital Elevation Model (DEM) can be calculated (Weidner and Förstner, 1995). This surface consists of objects rising from the terrain approximately put on a plane. The required DEM, i.e. the topographic surface can be derived from the measured DSM by mathematical morphology. Alternatively, an existing DEM can be used. Even though objects rising from the terrain can be detected quite well from height data, the discrimination between buildings and trees can be difficult, if only simple criteria like region size or shape are considered. One approach is to use the roughness of the DSM surface, which can be measured by differential geometric quantities as an additional criterion for the discrimination of buildings and vegetation (Brunn and Weidner, 1997).

3.1 Combination of height and surface reflectance

The analysis of multispectral imagery by standard classification algorithms is an alternative approach for the detection of cartographic objects in urban areas. One problem while classifying multispectral data is the similar reflectance of trees and grass-covered areas. Frequently, the same holds true for the distinction of streets and buildings. On the other hand, trees and buildings can be discriminated easily from grass-covered areas or streets within height data, since they are higher than their surrounding, whereas streets and green are at the terrain level. For this reason multispectral images and DSM can be used as complementary data sources during a classification.

(Hug, 1997) applies a scanning laser altimeter, which is able to measure distance *and* surface reflectance by processing the return signal energy of the laser beam. Thus, in addition to the range data a reflectance image in the near infrared spectrum is available, which enables the simultaneous usage of geometric and radiometric information for the detection of trees and buildings. (Henricson et al., 1996) use information from colored infrared aerial images to separate elevation blobs in a DSM from stereo image matching into the classes buildings and trees. The application of traditional spectral classification techniques to derive surface material information from multispectral imagery is presented by (Ford et al., 1997). They use hyperspectral data with nominal 2 meter ground sample distance with over 200 spectral samples per pixel, captured by the airborne sensor system Hyperspectral Digital Imagery Collection Experiment (HYDICE). The generated surface material map is refined by monocular segmentations from the panchromatic imagery and fused with high resolution stereo disparity maps.

3.2 ISODATA classification

In our approach multispectral information which is provided by a color-infrared aerial image is combined with geometric information from a laser scanner DSM in *one* classification step. The utilized CIR images were taken at a flying height of 1600 m with a normal-angle camera with a focal length of 305 mm. For digitization the images were scanned at a resolution of 60 μm, resulting in three digital images in the spectral bands near infrared, red and green. The basic idea of the proposed algorithm is to simultaneously use geometric and radiometric information by a pixel-based classification, i.e. the DSM is used as an additional channel in combination with the three spectral bands. For that purpose the height data and the images have to be corregistrated, i.e. a colored ortho image is generated from the original CIR imagery. Figure 2 shows the green channel of the ortho image, which was computed with a ground pixel size of 0.5 m. As discussed earlier the normalized DSM contains information on height above terrain. Since this type of information is very relevant for the discrimination of objects, the normalized DSM is used as an additional channel within a standard classification tool.

Figure 2: CIR ortho image (green channel selected for greyvalue representation).

In our approach an unsupervised classification algorithm is utilized. Unsupervised classification automatically detects spectral clusters of pixels in feature space. Since different object classes frequently result in different groups of similar spectral values, these accumulated clusters can be used for object discrimination. The unsupervised classification consists of the steps

1. Aggregation of clusters and categorization of pixels to the cluster which has minimum distance in feature space (automatic).
2. Interpretation of spectral classes (interactive).
3. Combination of spectral classes to thematic classes or object classes (interactive).

For the aggregation of spectral clusters in feature space several algorithms exist. We use the ISODATA (Iterative Self-Organizing Data Analysis Technique) algorithm, a standard procedure e.g. described in (Tou and Gonzalez, 1974) or (Richards, 1993). Within this approach the optimal number of spectral clusters is automatically determined by iteratively applying split and merge operations. The following steps are performed iteratively:

1. A number of parameters have to be initialized by the operator. These parameters are the desired number of clusters, the maximum number of iterations, the minimum number of pixels in a cluster, a maximum standard deviation to initiate cluster splitting, and a maximum distance in feature space between cluster centers to initiate cluster merging.
2. Each pixel is assigned to one of the predefined clusters by a minimum distance criterion.
3. All clusters containing less members than a predefined number of pixels are eliminated.
4. New cluster centers are computed from the aggregated pixels.
5. Aggregated clusters are split, if the maximum standard deviation is larger than the specified threshold. Neighbored clusters are merged if the pairwise distance is smaller than a predefined parameter and if the maximum specified number of clusters has not been reached.
6. The algorithm is terminated, if the maximum number of iterations is reached, else it is continued with step 2.

For computational reasons only a certain percentage of pixels (e.g. only the pixels of each 10th row and column) are used for cluster aggregation. After the stop criterion is reached, the detected cluster centers are used to classify the entire image. based on the minimum distance criterion. Since the clusters can not be split after the algorithm has been terminated, a larger number of clusters as expected in the image data is generated. Therefor the generated clusters have to be combined afterwards. The combination of spectral classes as well as their interpretation and transformation into object classes is performed interactively. For this reason and due to the number of parameters which have to be defined an operator with at least some experience is required to run the ISODATA algorithm.

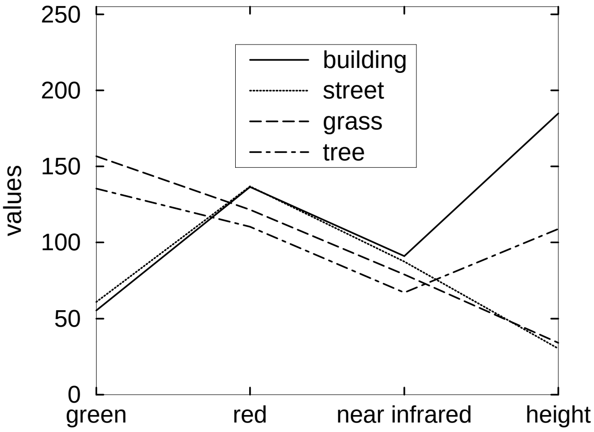

Figure 3: Signatures of sample objects.

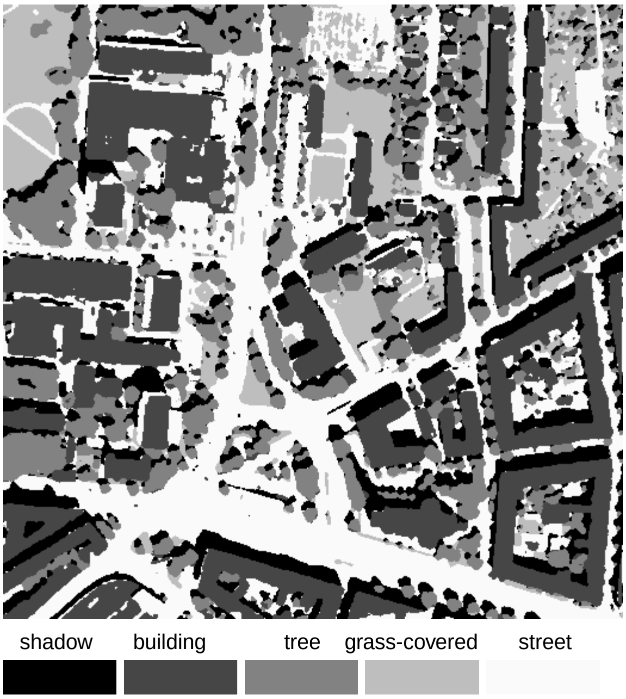

Figure 4: Result of ISODATA classification using CIR ortho image and normalized DSM

The results of the ISODATA algorithm of course depends on the quality of the input data. If the data in feature space is distributed in almost isolated natural groups these clusters can be detected very reliably. For our purpose the approach yields a good discrimination between the classes (from dark to bright) 'shadow', 'building', 'tree', 'grass-covered-area' and 'street'. The result of the classification is shown in figure 4. Figure 3 shows typical spectral reflectance curves. These values are mean values derived from small sample areas in the spectral bands. The height values were normalized to a range of 0 to 255 similar to the spectral channels since a standard routine was used for classification. Reflectance values for vegetation (classes tree and grass-covered area) are higher than for man-made objects (classes building and street) mainly in the green band, i.e. vegetation and man-made objects can be discriminated easily from spectral information. The classes building and street as well as the classes tree and grass-covered area are mainly different in the height channel and therefore can be also discriminated very reliably.

4 BUILDING RECONSTRUCTION

By the result of the classification the utilized 2D GIS can be checked for actuality and completeness. For example, buildings, which are not present in the ground plan dataset can be detected. In principle the classification can provide a very simple 2.5 D urban model; building blocks can be generated using the outlines of the classified buildings. For more sophisticated tasks like real-time visualization or simulations of high quality a true 3D model has to be provided. In order to separate building from the terrain surface and represent them by 3D CAD models, in our approach ground plans are used in addition to the DSM data.

4.1 Building Models

Object recognition or reconstruction, in general, presumes knowledge about the perceived objects by some kind of object model. These object models can be regarded as abstractions of real world objects. For model definition it is very important to find balance between 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 (Streilein, 1996). A priori knowledge or in other words constraints can be introduced by applying a very rigid building model. A rigid building model restricts the search space, which has to be examined to find a solution. On the other hand these models limit the number of possible building types which can be represented by a single model. In order to deal with the large architectural variations of building shapes, the utilized model should be as general as possible. Since most buildings are bounded by a set of planar surfaces and straight lines, in our approach a building is represented by a general polyhedron. Additional constraints are defined 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. This provides sufficient restrictions to enable the reconstruction of buildings without loosing the possibility to deal with very complex buildings.

Two types of representation are feasible to describe the reconstructed buildings. The *boundary representation (BRep)* is probably the most widespread type of 3D representation. Many algorithms are available for computing physical properties or visualizations from that representation. The object is represented by its surface, which is decomposed into a set of faces, edges and vertices. The topology is additionally described by a set of relations which indicate how the faces, edges and vertices are connected to each other. In *constructive solid geometry (CSG)* simple primitives are combined by means of 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 the closeness of the object surface. Additionally CSG enables a very compact object representation. A CSG can be transformed into a BRep, but 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.

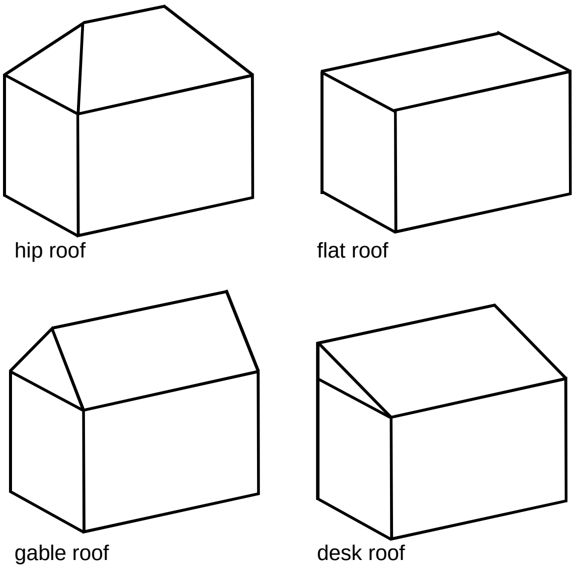
hip roof
flat roof
gable roof
desk roof

Figure 5: Building primitives used for reconstruction.

4.2 Model Selection and Parameter Estimation

Each building is described by a combination of one or more basic building primitives. The set of four basic primitives used for that purpose is shown in figure 5. Each of them consists of a cuboid element with different roof types flat roof, pent roof, gable roof and hip roof. This type of representation is similar to the one used by (Englert and Gülch, 1996).

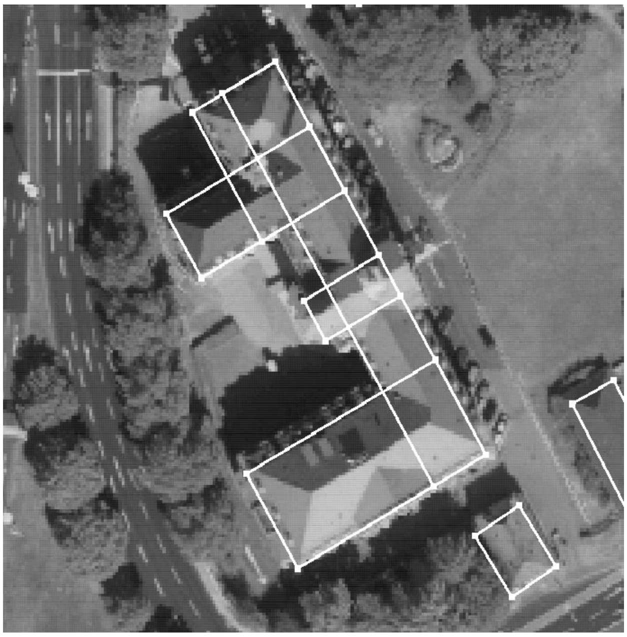

Figure 6: Ground plan decomposed into rectangular parts.

In order to reconstruct more complex buildings, first the complete building has to be decomposed into these basic structures. This step can be realized automatically by the analysis of the given ground plan. Figure 6 shows the result of a ground plan decomposition into rectangular structures. Every rectangle defines one building primitive. Since position, orientation and horizontal extension of each cuboid is already defined by each rectangle, only the height of every cuboid as well as roof type and roof slope have to be determined as remaining parameters for the building primitives. The parameters of the building primitives 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 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 to avoid gross errors of the estimated parameters. For both tasks the result of a segmentation of the DSM are used.

This DSM segmentation into planar surfaces is supported by introducing ground plan information. Of course the given ground plan restricts the extension of the DSM area which has to be examined. More important, the implemented segmentation within each ground plan area 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 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. Hence, the different segments of the ground plan polygon are used to trigger the segmentation of a planar surface with a projected normal vector perpendicular to this element. A more detailed description of the segmentation process can be found in (Haala et al., 1997).

Figure 7: Segmented roof regions.

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 6 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.

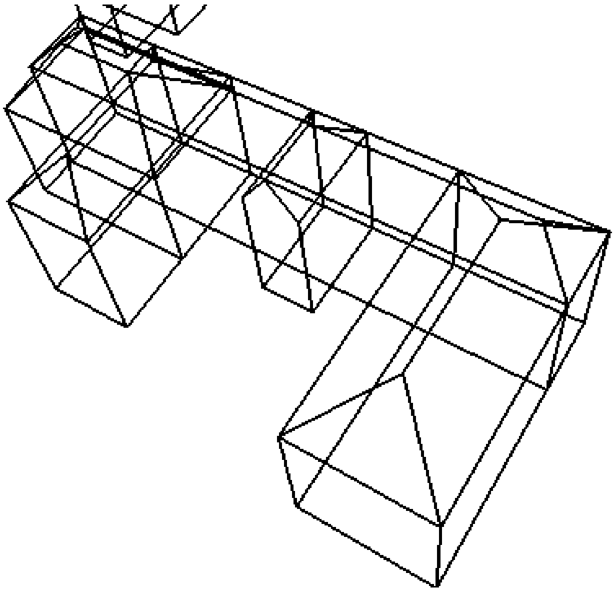

Figure 8: Reconstructed CSG building primitives.

Applying our algorithm results in the segmentation represented by the shaded regions in figure 7. 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 8, showing the reconstructed building primitives as wire frames.

4.3 Transformation of CSG into Boundary Representation

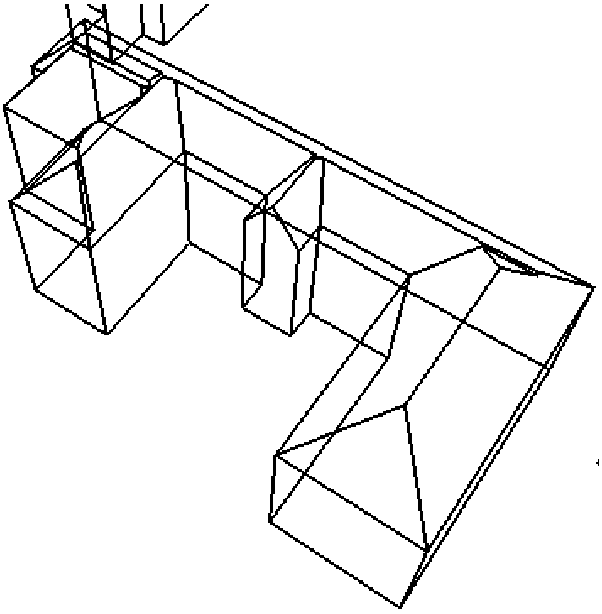

Figure 9: CSG data converted to boundary representation.

Usually, simulations or visualizations require a boundary representation of the buildings. 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. During this process problems have to be solved, which arise due to lack in numerical accuracy of the real world measurement made for reconstruction, due to decreased data accuracy caused by storing of the building descriptions, and due to rounding errors caused by computation with real numbers. Figure 9 shows the boundary representation generated from the different building primitives is presented in figure 8.

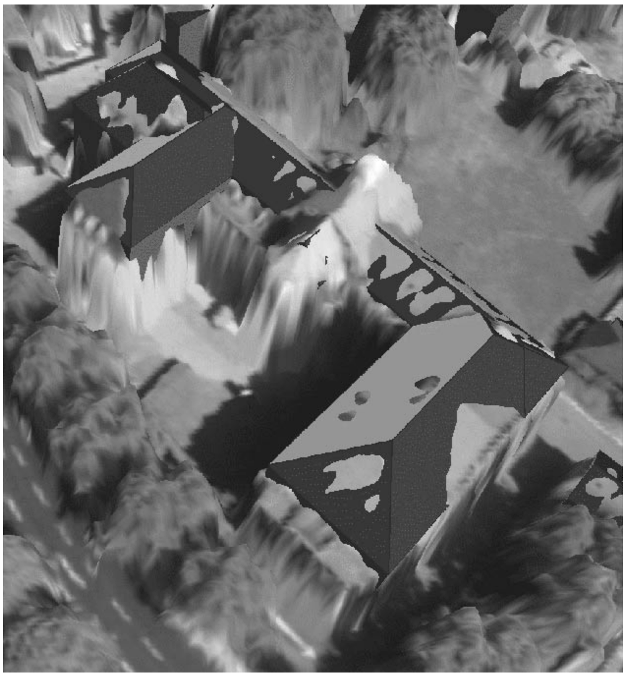

Figure 10: Reconstructed building and DSM surface.

A 3D visualization of the original DSM surface overlaid to the boundary representation of the building is given in figure 10. The difference between the DSM surface and the corresponding 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 to problems of the automatic reconstruction. Since these regions are determined by the previous segmentation step, they can be presented to an operator together with the RMS values in a final evaluation step. If necessary a further refinement can be achieved by an interactive step by manually adding small parts like missing bays to the initial reconstruction. Figure 11 shows the result of the automatic building reconstruction projected to the map of scale 1:5000, which was used to digitize the ground plans utilized for building reconstruction.

5 CONCLUSION

Using 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. A combined classification of CIR images and height data for the data collection in urban areas has been demonstrated. This classification can e.g. be used to detect obsolete parts within the utilized 2D GIS or to provide features like trees, which are important for virtual reality applications but are usually not contained in the GIS data. The use of these features is demonstrated in in figure 12. For the generation of a virtual reality model terrestrial images were mapped on the building faces. Since for this application the building geometry was already available from the automatic reconstruction step using ground plans and DSM, the effort for texture processing was reduced considerably compared to the application of standard architectural photogrammetry with close range imagery. The artificial trees were added to the virtual model based on the classification described in this article.

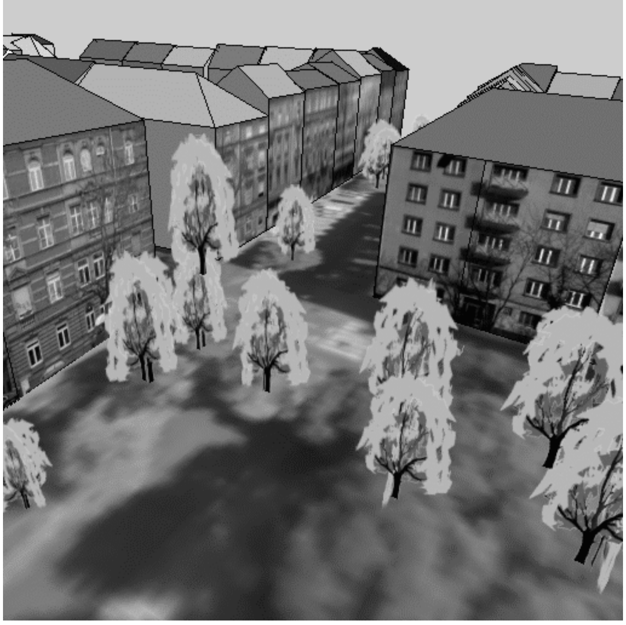

Figure 12: 3D visualization of virtual city model.

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

REFERENCES

Baltsavias, E., Mason, S. and Stallmann, D., 1995. Use of DTMs/DSMs and orthoimages to support building extraction. In: A. Gruen, O. Kuebler and P. Agouris (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser Verlag, Basel, Boston, Berlin, pp. 199–210.

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

Brunn, A. and Weidner, U., 1997. Extracting buildings from digital surface models. In: ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects, Stuttgart, Germany, pp. 27–34.

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.

Ford, S., Kalp, D., McGlone, C. and McKeown, D., 1997. Preliminary results on the analysis of hydice data for information fusion in cartographic feature extraction. In: SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III, Orlando, Florida, pp. 67–86.

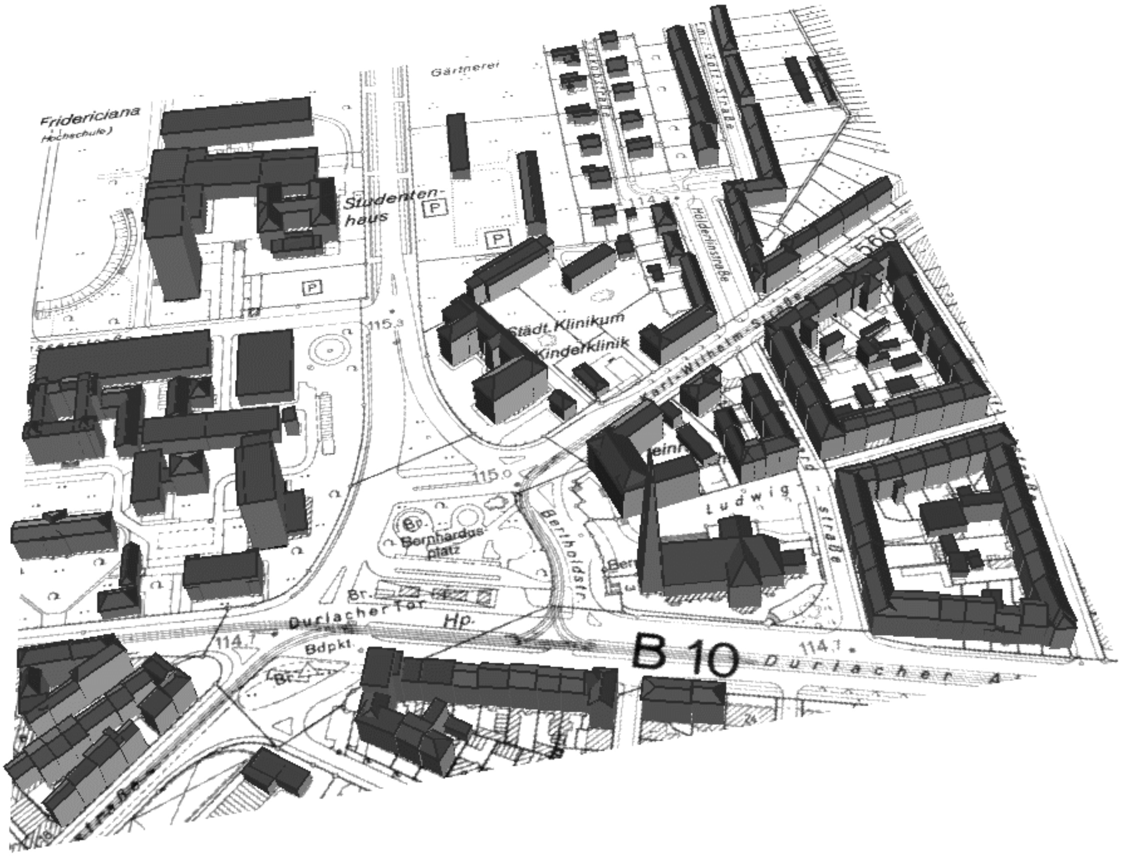

Figure 11: Result of automatic 3D building reconstruction projected to map 1:5000 used for digitization of ground plans.

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., 1994. Detection of buildings by fusion of range and image data. In: Proc. ISPRS Congress Comm. III, München, pp. 341–346.

Haala, N., Brenner, C. and Anders, K.-H., 1997. Generation of 3D City Models from Digital Surface Models and GIS. In: ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects, Stuttgart, Germany, pp. 68–76.

Henricson, O., Bignone, F., Willhuhn, W. and Ade, F., 1996. Project Amobe: Strategies, current status and future work. In: Proc. ISPRS Congress Comm. III, Wien, pp. 321–330.

Hug, C., 1997. Extracting artificial surface objects from airborne laser scanner data. In: A. Gruen, O. Kuebler and M. Baltsavias (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhäuser Verlag, Basel, Boston, Berlin, pp. 203–212.

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.

Richards, J. A., 1993. Remote Sensing Digital Image Analysis. Springer–Verlag, Berlin Heidelberg.

Streilein, A., 1996. Utilization of CAD models for the object oriented measurement of industrial and architectural objects. In: Proc. ISPRS Congress Comm. III, Wien, pp. 548–553.

Tou, J. T. and Gonzalez, R. C., 1974. Pattern Recognition Principles. Addison–Wesley, London.

Weidner, U. and Förstner, W., 1995. Towards automatic building extraction from high resolution digital elevation models. ISPRS Journal 50(4), pp. 38–49.