AN APPROACH TO BUILDING EXTRACTION FROM DIGITAL SURFACE MODELS

U. Weidner
Institut für Photogrammetrie
Rheinische Friedrich-Wilhelms-Universität Bonn
Germany
weidner@ipb.uni-bonn.de

Commission III, Working Group 2

KEY WORDS: Digital Surface Model, DEM, Building Extraction, Geometric Object Modeling

ABSTRACT

Motivated by the test data sets of ISPRS WG III/3 on image understanding we investigate the feasibility of building extraction using high-resolution Digital Surface Models (DSM) as input data, which do not only contain information about the topographic surface like Digital Elevation Models (DEM), but also information about the buildings. The steps of the proposed procedure increasingly use explicit domain knowledge, specifically geometric constraints in the form of parametric and prismatic building models. The reconstruction of the prismatic models and the selection of the models are based on the principle of Minimum Description Length (MDL). In addition, we also discuss the possible use of information from GIS or maps in our approach.

KURZFASSUNG

Unsere Arbeiten bezüglich der Gebäudeerfassung aus hochauflösenden Digitalen Oberflächenmodellen, die nicht nur Information über die topographische Oberfläche wie Digitale Höhenmodelle, sondern auch über Gebäude enthalten, sind durch die von der ISPRS WG III/3 herausgegebenen Testdatensätze angeregt worden. Die Schritte des hier aufgezeigten Verfahrens nutzen Wissen über Gebäude, welches durch die verwendeten parametrischen und prismatischen Gebäudemodelle repräsentiert wird. Die Rekonstruktion der prismatischen Modelle und die Auswahl des für die Rekonstruktion anzuwendenden Modells basieren auf dem Prinzip der Minimalen Beschreibunglänge. Desweiteren wird der Einsatz von GIS und Karten als zusätzliche Datenquelle im Zusammenhang mit der Gebäudeerfassung aus Digitalen Oberflächenmodellen diskutiert.

1 INTRODUCTION

During the last recent years the need for 3D data describing urban areas increased. This data is needed for a variety of applications such as town planning, architecture, microclimate investigations or transmitter placement for telecommunication. Our investigations concerning the feasibility of building extraction using high resolution Digital Surface Models (DSM) as input data, which do not only contain information about the topographic surface like Digital Elevation Models (DEM), but also information about the buildings, have been motivated by the test data sets of ISPRS WG III/3 on image understanding. In contrast to other authors, who use digital imagery (e.g. [Lang and Schickler, 1993], [McGlone and Shufelt, 1994]) or digital imagery and DSM (e.g. [Haala, 1994], [Baltsavias et al., 1995]), we solely focus our investigations in a first step on DSM for two reasons: First of all, we are interested in investigating the potential which is inherent in the use of such DSM. The main advantage of DSM is that they already provide a geometric description of the objects, although this description of buildings shows some deficiencies with respect to building extraction. These deficiencies are the representation of the object, the resolution, and the discrimination of buildings and other objects. The representation of the objects is not sufficient in all cases, e.g passages. The resolution of ground plan information is restricted to the resolution of the DSM. The discrimination of buildings and other objects (e.g. trees) is not always possible using solely a DSM. Despite these deficiencies, DSM seem to be a good intermediate description to link sensor data with high level knowledge about buildings. Secondly, using only DSM as input data enables us to use DSM, which are generated by matching techniques using digital images ([Krzystek, 1991], [Collins et al., 1995]), but also those derived by other measurement devices like laser scanners [Lohr and Eibert, 1995].

Our approach to building extraction from DSM consists of two steps – the detection and the reconstruction of buildings. Both steps use high level knowledge about the buildings, which is introduced in the form of parametric and prismatic building models, and some specific knowledge about buildings in the region of interest to fix some building relevant thresholds. All this knowledge is object space related making adaption to data of different density and resolution simple and transparent. In [Weidner and Förstner, 1995] we already described the general strategy of our approach, including building detection and building reconstruction. The main extensions of this contribution consists of the automatic selection of the model for the reconstruction of buildings. This selection is based on the principle of Minimum Description Length (MDL), as well as the reconstruction of prismatic building models. Therefore, the next section shortly describes the MDL-principle, followed by a summary of the general strategy of our approach. Section 5 shows results for the ISPRS test data sets and other data. Section 6 describes, how GIS or map information can be incorporated in our approach, followed by the conclusions.

2 MINIMUM DESCRIPTION LENGTH PRINCIPLE

The shape reconstruction described in Section 4 is based on the principle of Miniumum Description Length (MDL). This principle provides means to select and estimate the parameters of the selected model in a common framework, and enables us to reconstruct the shape of the buildings by integrating data and model information. The description length DL depends on the complexity of the used model and the deviation of the data from the model. The complexity depends on the number of unknown parameters and the number of observations. The deviation of the data from the model is given by the weighed squared sum of the residuals Ω of a ML-estimation.

Let the following model be given

$$E(\mathbf{y}) = \mathbf{g}(\beta), \quad D(\mathbf{y}) = \Sigma_{yy} \tag{1}$$

where β denotes the $u \times 1$ vector of unknown parameters, $\mathbf{y}$ the $n \times 1$ vector of observations and Σ_{yy} their covariance matrix. The description length [Rissanen, 1987] follows by

$$DL = \frac{\Omega}{2\,ln2} + \frac{u}{2}\,lb\,n \tag{2}$$

where Ω is given by

$$\Omega = [\mathbf{y} - \mathbf{g}(\hat{\beta})]^T \Sigma_{yy}^{-1} [\mathbf{y} - \mathbf{g}(\hat{\beta})] \tag{3}$$

Following the principle of MDL, we search for the description which minimizes (2), thus selecting the model and fitting the data to the model simultaneously.

In order to decide whether a difference in description length between two alternatives is significant, a hypothesis test based on the variance of DL can be applied. The variance of DL follows by error propagation taking the variance of Ω into account, which is $2\,(n-u)$, and thus

$$\sigma^2_{DL} = \frac{2\,(n-u)}{(2\,ln\,2)^2} \tag{4}$$

3 BUILDING DETECTION

The first step towards building extraction is the detection of possible building areas in order to focus the later steps of reconstruction on these. The principal idea of our approach to building detection is to isolate the information about the buildings within the DSM and to segment this data by binarization using a building related threshold, e.g. the height of a floor. Therefore, we first compute an approximation of the topographic surface. There are different ways which can be followed for this purpose. In our approach, we use mathematical morphology (here: *opening*). As an alternative of such an opening, a dual rank filter, which is a modification of the opening, can be used. The modification is to use the median of the minimal and maximal $p\%$ values of the applied structuring element [Eckstein and Munkelt, 1995] instead of the minimum and maximum itself. This approach has some advantages compared to the opening, because it compensates for noise and outliers in the data. For the data sets we use here the difference between these two approaches show only minor effects on the following steps, because the percentage of outliers seems to be small.

The difference between original DSM and the approximation of the topographic surface contains the information about the buildings, approximately put on a plane. Due to this fact, a binarization with a given threshold yields a first segmentation. This segmentation shows some deficiencies due to some effects of the DSM generation, e.g. round off at building edges due to regularization, and global thresholding. Furthermore, the first segmentation may include segments, which are higher than the surrounding topographic surface, but which do not represent buildings, e.g. trees. In order to overcome these short-comings, we first select only those segments, whose area is greater than the expected minimum area of buildings, and then refine the segmentation by adapting the threshold locally based on the height information within a bounding box

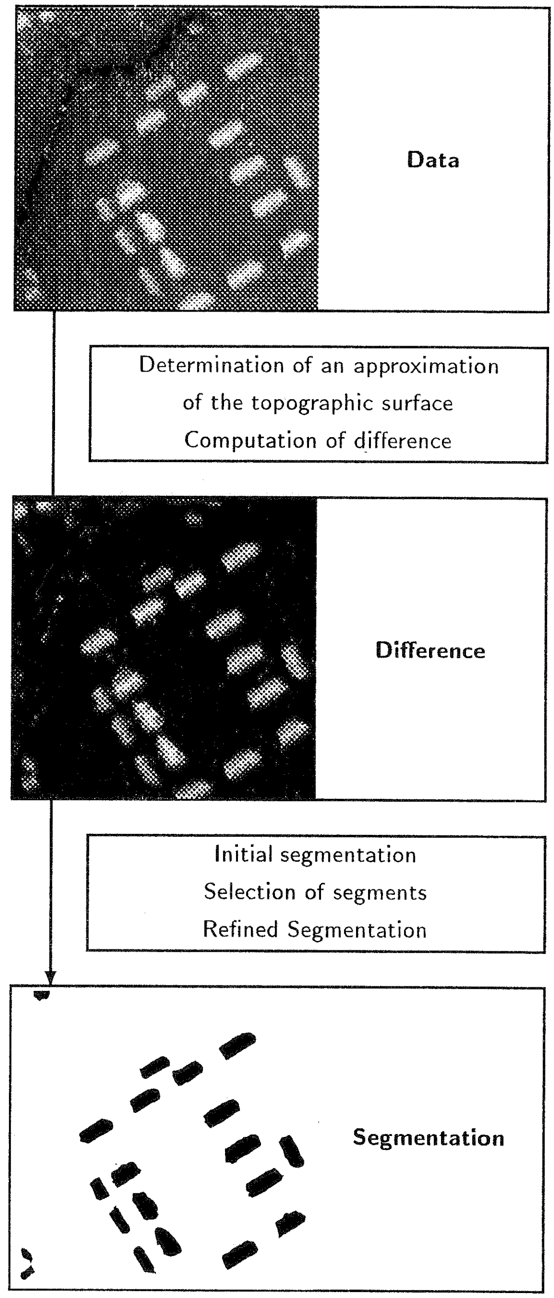


Figure 1: Building Detection

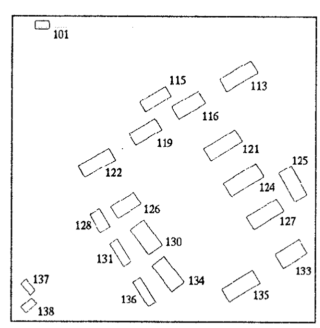


Figure 2: FLAT: gound plan labels

of each selected segment. These segments of the refined segmentation form the basis for the extraction of the buildings' 2D information, whereas the height information is derived by analysing the height information within the segments and the related bounding box without segments. Figure 1 shows an overview of the building detection using the ISPRS test data set FLAT as example.

The use of a geometric criterion to distinguish between buildings and other objects higher than the topographic surface is not always sufficient. Therefore, other criteria using other sources of information have to be used, e.g. texture information [Eckstein and Munkelt, 1995] or edge information [Baltsavias et al., 1995] from aerial imagery. This information can also and – depending on the application – should already be applied during the DSM generation.

4 BUILDING RECONSTRUCTION

The building reconstruction of our approach is based on the use of parametric and prismatic building models. Parametric models are used for simple separated buildings, which can be described by a few parameters, e.g. a building with a symmetrically sloped roof, whereas prismatic models (ground plan and height information) including generic knowledge about regularities (e.g. orthogonalities, parallelisms, collinearities) are used for complex buildings or building blocks.

4.1 Parametric Building Models

In our approach we use two different parametric models: flat buildings and buildings with a symmetric, sloped roof. The form parameters of these building models are the *length*, *width* and *height* for flat buildings, and *length*, *width*, *height of eave-base* and *height of ridge-eave* for buildings with a symmetric, sloped roof, assuming that the ground plans of these buildings are given by rectangles. Furthermore, four parameters are needed to describe the position and orientation of the building within the reference coordinate system.

In order to determine the x,y coordinates of a building's reference point and the orientation, the point of gravity and the orientation of each refined segment using the heights within the segments as weights are computed. The z coordinate of the reference point is computed taking the mean of heights within the background area of the bounding box. The parameters length and width are the length and width of a rectangle approximating the segment and are computed along the first and second main axis of the segment. The height parameters are computed taking the height information in the original DSM into account. For this purpose, we use a ranking scheme, and use the median of the $p\%$ minimal and maximal values within the segments as robust estimation of the minimum and maximum, where $p = 10$ is choosen. The height of a flat building follows from the difference between the mean height of the segment and the mean height within the background. For buildings with symmetric sloped roofs the height parameters are computed as: height1 = difference between the estimated minimum height of the segment and the mean height of the background; height2 = difference between the estimated maximum and minimum within the segment.

The parameters of the parametric models are computed for each detected segment. In [Weidner and Förstner, 1995], we selected the model which is used for the description using a geometric criterion, namely the slope of the roof. This approach seems to be feasible for the selection of the para-

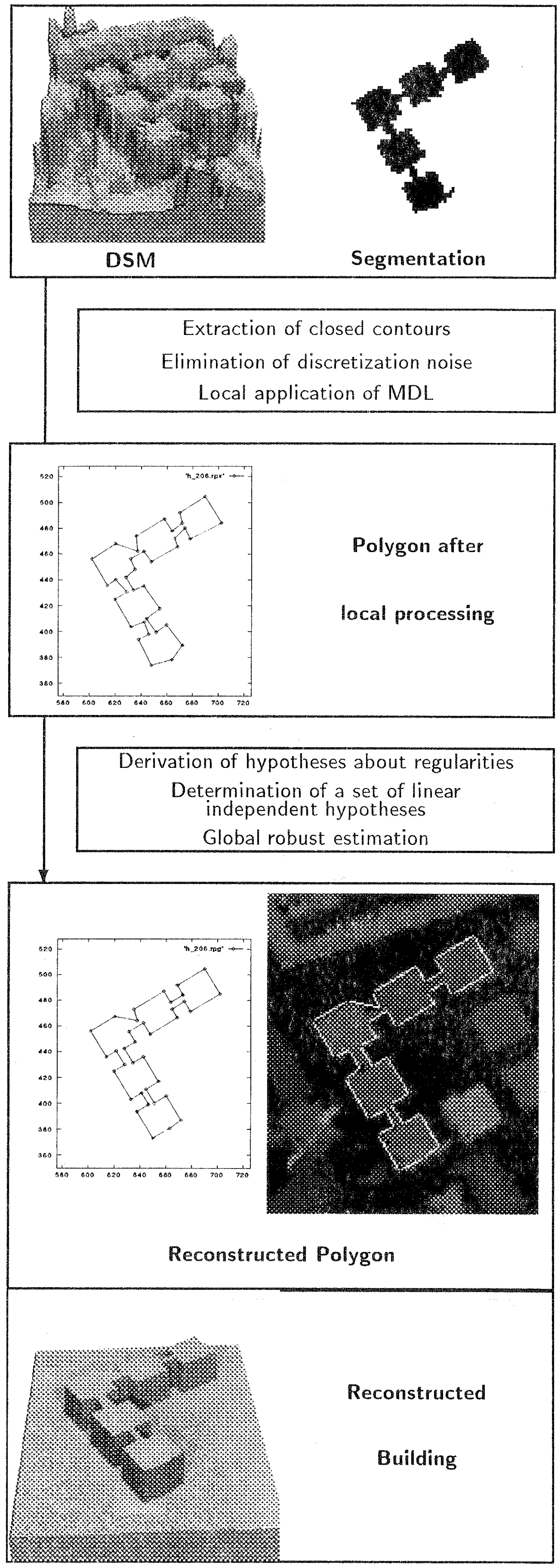


Figure 3: Building Reconstruction: Prismatic Models

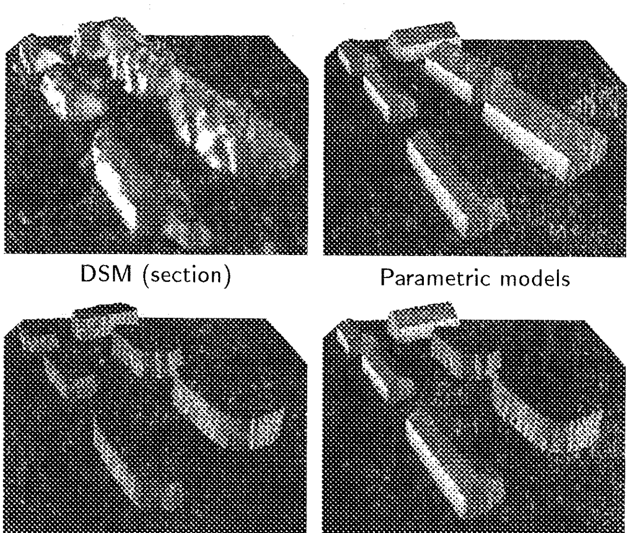
metric model to be applied, but does not allow a decision, which group of model should be applied in a common framework. Therefore, we choose MDL as a tool, which is able to compare different descriptions with different structual complexities. This approach is described in Section 4.3.

4.2 Prismatic Building Models

For prismatic models our approach to shape reconstruction of the outlines consists of a local and a global analysis step, which can be combined in different ways. Complexity considerations [Brunn et al., 1995] indicate to preprocess the closed contours in order to elimate discretization noise due to the DSM raster using a merging or splitting technique and to iterate the local MDL-application – c.f. [Weidner and Förstner, 1995] for detailed description – until no changes occur any longer, and then proceed with global processing, which starts with the derivation of hypotheses about the regularities. Due to the transitivity of parallelism and collinearity, and similar relations including orthogonality, these hypotheses are linear dependent. On the other hand, sets of individually consistent hypotheses need not be jointly consistent. Therefore, we continue with the determination of a set of linear independent hypotheses, which is then introduced into a robust global estimation procedure [Fuchs and Förstner, 1995]. The height information is computed as for the flat parametric building model Figure 3 shows an overview of the MDL-based reconstruction of prismatic models.

4.3 Model Selection

The selection of the model which should be applied for the description of the building is based on MDL. For this purpose, all alternative models are computed. Due to the fact that the ground plan between the parametric models and the prismatic model differ, different areas of the data are described using the different groups of models. Therefore, the selection of the model is not directly related to the description length, but to the gain in description length compared to the case, if no model is used.



DSM (section) Parametric models

Prismatic models Selected models

Figure 4: Selection of models

The number u of free parameters needed to describe the models are seven for the flat building and eight for the building with symmetric, sloped roof. Prismatic models are described by *number of polygon points* $\times$ 2 parameters, if no restrictions are introduced, and two parameters for the height information. The number of observations n is related to the number of points in the ground plan of the models. Ω measures the deviations $\mathbf{d}$ of the models from the DSM data, i.e. $\Omega = \mathbf{d}^T\, \Sigma_{yy}^{-1}\, \mathbf{d}$, where $\Sigma_{yy} = \sigma_h^2 \mathbf{I}$ was used here with $\sigma_h = 0.4\, m$.

Label	126	128	130
no parametric model	21105.8	10193.1	26471.5
# points	1192	625	1459
FLAT	18545.3	3282.7	41420.1
SYMSL	7864.2	1080.7	23419.2
gain	13241.6	9112.4	3052.3
no prismatic model	19364.4	9420.4	21742.5
# points	1104	583	1224
PRISM	11672.8	2121.3	12016.5
gain	7691.6	7299.1	9726.0
difference in gain	5550.0	1813.3	6673.7
selected model	SYMSL	SYMSL	PRISM

Table 1: Characteristics of models

Figure 4 shows the original data, the results of building reconstruction using parametric and prismatic models, and the selected models. The characteristcis for some of the models are gathered in Tab. 1 (see Figure 2 for identification of labels). The examples indicate the feasibility of using MDL as criterion. Label 128 is correctly selected as a building with symmetric, sloped roof. Label 130 is classified as prismatic model due to the fact that the roof consists of three gabels. That is also true for label 126, but this building is classified as having a symmetric, sloped roof, due to the fact that the roof structure does not appear clearly in the DSM, because of round off effects of the regularization in the DSM generation.

In all three cases, the difference in gain is significant. In such cases, where the difference in gain is not significant, both models should be regarded as possible alternatives and kept for further processing.

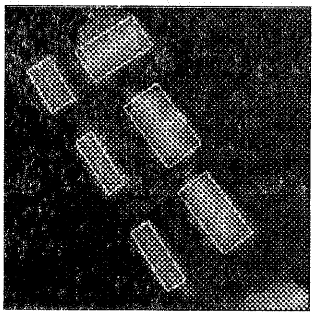
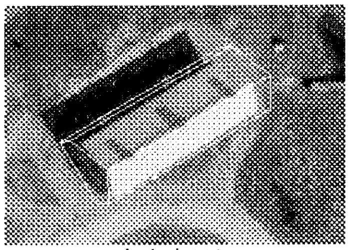


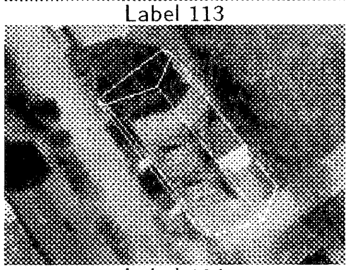
Figure 5: FLAT: overlay DSM – ground plan (section)

5 RESULTS

In this section, some results of our approach are discussed using the ISPRS test data set FLAT with a ground resolution of $0.5\,m \times 0.5\,m$ and a DSM of a downtown area as examples. A detailed description of the results of our approach for the ISPRS test data sets is given in [Weidner, 1995], also including a discussion concerning the used control parameters. All results of this test have been compiled in [Sester et al., 1996]. A comparison between image and ground plan information derived from the DSM for the data set FLAT indicates that only one building has not been detected.



Label 113



Label 134

Figure 6: Projected models

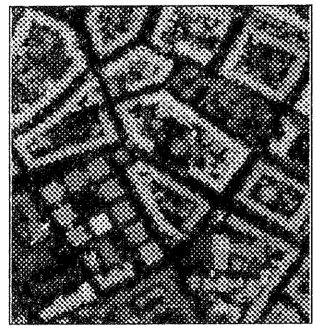
Figure 6 shows extracted building models (white lines) projected into the left image of the stereo pair. A qualitative evaluation indicates that the orientations of the extracted models fit to the image information. A rough comparison of the extracted roof heights with manually measured points indicate correspondence. The mean of the differences (absolute value) is about 0.2 m for the ridges and 0.5 m for the eaves. Problems occur for the paramters *lenght* and *width*, although an overlay (Figure 5) of the DSM and extracted ground plan information indicates a plausible fit. An explanation of the effects may be that during DSM generation, interest points are found at the borders of the roofs. Due to low texture (c.f. label 113) or shadows (c.f. label 134), no interest points are found close to the building for supporting matching. Therefore, the regularization term within the reconstruction algorithm leads to interpolation between points at the roofs' borders and points on the ground, which are more or less far away from the building, thus elongating the buildings sytematically. Furthermore, the round offs at breaklines contribute to such effects, although we try to take these effects into consideration during the refined segmentation.

Fig. 3 (Reconstructed Polygon) displays the extracted polygon superimposed on the original range data[1], acquired by airborne laser scanning. For the data set local MDL-application leads to a reduction of the number of points from 98 to 36. The hypothesis about geometric relations between edges of the polygon, which are introduced in the robust estimation, put constraints onto the edges. A qualitative evaluation shows little discrepancies, whereas the overall performance seems to be acceptable. The discrepancies are on one hand due to the sequence of analysis steps used here (c.f. Section 4.2). On the other hand not all hypotheses passing through the robust estimation are actually correct.

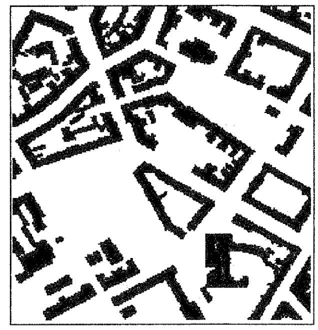
6 GIS DATA AND CHANGE DETECTION

In our approach to building extraction from DSM GIS or map information can be incorporated as additional source of information. GIS and maps mainly deliver 2D information about buildings. The information about the third dimension in a GIS is often related to the topographic surface, represented by a DEM, or this information may only be qualitativ, e.g. the number of floors of a building, which can be used to derive quantitative height information, if the mean height of floors is known.

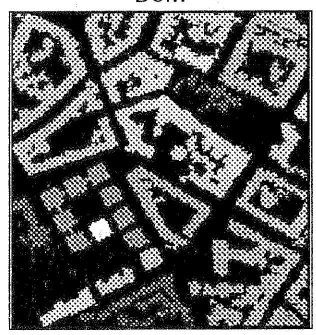
[1] The range data of Hannover with a ground resolution of $2\ m$ was supplied by Dornier, Friedrichshafen.



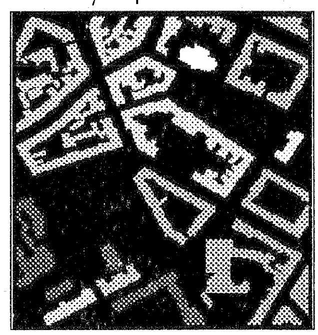
DSM



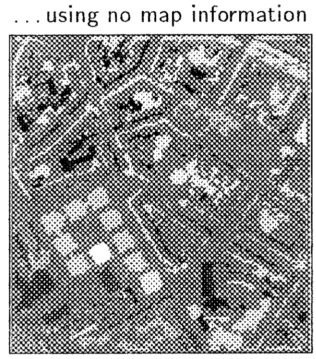
GIS/map information



. . . using no map information



. . . using map information



Difference

Hypotheses about changes

Figure 7: Use of DSM and GIS/map information

The DSM description of a scene has one advantage compared to the GIS data, because it describes the actual scene which might differ from the GIS due to changes. Therefore, GIS data and DSM can be used for two purposes:

- The 2D information about buildings in a GIS or map – depending on the scale – can be considered to be more precise than the 2D information which can be extracted from a DSM, and can therefore replace this information. The DSM only serves as information source about the third dimension for the buildings in the GIS.
- The DSM can be used to generate hypotheses about changes in the scene.

A possible scenario is presented in Figure 7. The principle idea of this scenario is to apply our approach without using information from GIS and compare the results with the results using this information. For this purpose the results are represented in DSM and the comparison consists of computing the difference between these two results. Binarization and a comparison with the used 2D GIS information delivers hypotheses about changes. Due to the fact that our approach is not able to distiguish between buildings and other objects mimicking to be buildings, e.g. group of trees which we cannot distinguish purely based on geometric criteria, also segments which represent such objects are included in the hypotheses. Possible approaches to solve this problem have been already discussed in Section 3.

7 CONCLUSIONS

The presented approach towards building extraction using DSM consists of automatic detection and reconstruction of buildings. Both steps are based on generic contextual knowledge. This knowledge is represented in geometric building models, parametric ones for simple buildings, and prismatic models for complex buildings and blocks of buildings. The results for the ISPRS data set show the capability of our approach when dealing with simple buildings, if the DSM contains significant information about the buildings. Further work for parametric models will focus on the integration of other parametric models, e.g. buildings with non symmetric sloped roofs. In order to improve the accuracy of parameters, template matching for the estimation of the point of gravity and the orientation will be investigated. Nevertheless, the resolution of the parameters related to the ground plan will always depend on the resolution of the DSM grid. Prismatic models are used for the data set of a downtown area. The achieved result is strongly influenced by the resolution of the grid. In order to deal with complex buildings consisting of parts with different heights more appropriately, discrimination of different parts using the height information within the region circumscribed by the extracted polygons with the aim of deriving a building graph is necessary (cf. [Fua and Hanson, 1987]). As starting point for such a reconstruction of roofs – i.e. grouping of planes, generating and verifying hypotheses about regularities – the detected planes in Figure 8 (right) can be used. These planes are homogeneous regions within the building segments, discarding regions of high curvature (Figure 8 (left)). Furthermore other constraints, e. g. symmetries, and semantic knowledge about rows of buildings, and the use of the gained information in our image analysis system [Lang and Förstner, 1996] will be investigated.

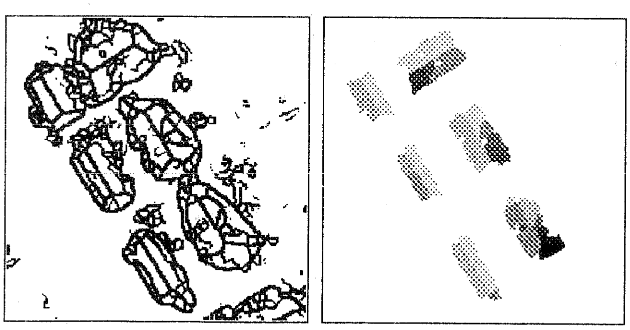


Figure 8: Curvature information and roof planes

REFERENCES

[Baltsavias et al., 1995] Baltsavias, E., Mason, S., and Stallmann, D. (1995). Use of DTMs/DSMs and Orthoimages to Support Building Extraction. In Grün, A., Kübler, O., and Agouris, P., editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 199–210. Birkhäuser, Basel.

[Brunn et al., 1995] Brunn, A., Weidner, U., and Förstner, W. (1995). Model-based 2D-Shape Recovery. In Sagerer, G., Posch, S., and Kummert, F., editors, *Mustererkennung 1995*, pages 260–268. DAGM, Springer.

[Collins et al., 1995] Collins, R., Hanson, A., Riseman, E., and Schultz, H. (1995). Automatic Extraction of Buildings and Terrain from Aerial Images. In Grün, A., Kübler, O., and Agouris, P., editors, *International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 169–178. Birkhäuser, Basel.

[Eckstein and Munkelt, 1995] Eckstein, W. and Munkelt, O. (1995). Extracting Objects from Digital Terrain Models. In Schenk, T., editor, *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*. SPIE.

[Fua and Hanson, 1987] Fua, P. and Hanson, A. (1987). Resegmentation Using Generic Shape Locating General Cultural Objects. *Pattern Recognition Letters*, 5:243–252.

[Fuchs and Förstner, 1995] Fuchs, C. and Förstner, W. (1995). Polymorphic Grouping for Image Segmentation. In *Proceedings 5th International Conference on Computer Vision*, pages 175–182.

[Haala, 1994] Haala, N. (1994). Detection of Buildings by Fusion of Range and Image Data. In *ISPRS Comm. III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision, Proceedings*, pages 341–346. SPIE.

[Krzystek, 1991] Krzystek, P. (1991). Fully Automatic Measurement of Digital Elevation Models. In *Photogrammetric Week, Stuttgart, Proceedings*, pages 203–214.

[Lang and Förstner, 1996] Lang, F. and Förstner, W. (1996). 3D-City Modelling with a Digital One-Eye-Stereo System. In *18th ISPRS Congress, Wien, Proceedings*.

[Lang and Schickler, 1993] Lang, F. and Schickler, W. (1993). Semiautomatische 3D-Gebäudeerfassung aus digitalen Bildern. *ZPF*, 5:193–200.

[Lohr and Eibert, 1995] Lohr, U. and Eibert, M. (1995). The TopoSys Laser Scanner System. In Fritsch, D. and Hobbie, D., editors, *Photogrammetric Week*, pages 263–267. Wichmann, Karlsruhe.

[McGlone and Shufelt, 1994] McGlone, J. and Shufelt, J. (1994). Projective and Object Space Geometry for Monocular Building Extraction. In *Proceedings Computer Vision and Pattern Recognition*, pages 54–61.

[Rissanen, 1987] Rissanen, I. (1987). Minimum Description Length Principle. *Encyclopedia of Statistical Sciences*, 5:523–527.

[Sester et al., 1996] Sester, M., Schneider, W., and Fritsch, D. (1996). Results of the Test on Image Understanding of ISPRS Working Group III/3. In *18th ISPRS Congress, Wien, Proceedings*.

[Weidner, 1995] Weidner, U. (1995). Evaluation of Building Extraction from Digital Elevation Models. Technical Report 95/2, Institut für Photogrammetrie, Bonn.

[Weidner and Förstner, 1995] Weidner, U. and Förstner, W. (1995). Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal*, 50(4):38–49.