Uwe Weidner [1] and W. Förstner [1]

Towards automatic building extraction from high-resolution digital elevation models

This paper deals with an approach for extracting the 3D shape of buildings from high-resolution Digital Elevation Models (DEMs), having a grid resolution between 0.5 and 5 m. The steps of the proposed procedure increasingly use explicit domain knowledge, specifically geometric constraints in the form of parametric and prismatic building models. A new MDL-based approach generating a polygonal ground plan from segment boundaries is given. The used knowledge is object-related making adaption to data of different density and resolution simple and transparent.

1. Introduction

There is an increasing need for 3D descriptions of urban areas for various applications such as town planning, microclimate investigation or transmitter placement in telecommunication. This information is literally unavailable. City plans contain 2D information about the ground plan of buildings, possibly encoded with some height-relevant information such as the number of stories.

Classical photogrammetric techniques are expensive mainly due to the higher amount of information required but also because acquisition techniques are not developed for this purpose. Though techniques using laser scanning data or digital imagery show promising results in performance (cf. Lang and Schickler, 1993), a breakthrough in 3D information extraction requires a great deal of automation based on pattern recognition techniques.

Automatic interpretation of digital data needs explicit modelling of the type of scene to be extracted. The amount of modelling depends on various factors:

- *The complexity of the scene*: Simple building models only cover a limited percentage of buildings or provide quite generalized building descriptions. Typical models are parametric models (McGlone and Shufelt, 1994), which describe a building by a small set of parameters, and prismatic models, which describe a building by its polygonal ground plane and a height (Herman and Kanade, 1986), or union of blocks, which are able to describe complex buildings having parts with different heights (cf. Lin et al., 1994). Probably the most advanced building model has been used by Fua and Hanson (1987), who also were able to derive relational descriptions of complex buildings. All approaches published so far proved the feasability of automatic techniques for interpreting digital imagery.
- *The type of sensor data*: As photogrammetric techniques rely on imagery, mostly aerial, automation requires the derivation of 3D coordinates using matching techniques. In spite of intensive research during the last few decades in automatic stereo vision, no approach has been able to prove its superiority in reconstructing the 3D geometry of urban areas. There are two main reasons why computational stereo is hard in urban areas: the always present occlusions and the need to represent vertical surface structures.
  Laser scanners (Krabill et al., 1984; Krabill, 1989) immediately provide 3D coordinates and are much more reliable in deriving surface data than stereo techniques. They are very precise in height, but up to now have less resolution in planimetry than comparable aerial imagery. Due to their small viewing angle, techniques with laser scanners are not yet economical. Moreover they leave the same problem unsolved which is essential for deriving city models: vertical walls are not accessible.
  Integrating both approaches seems to be quite promising as range data from laser scanners can

[1] Institut für Photogrammetrie, Universität Bonn, Nußallee 15, 53115 Bonn, Germany. Email: weidner wf@ipb.uni-bonn.de

ISPRS Journal of Photogrammetry and Remote Sensing, 50(4): 38–49
0924-2716/95/$09.50

provide 3D information which may then be used during image analysis (cf. Haala, 1994).

In spite of the extremely high complexity of urban scenes and the deficiencies of current matching technology, the situation for 3D data acquisition for 3D city models is not unfavorable, for three reasons:

(1) Stereo matching algorithms provide good approximations to the visible surface, especially in roof areas. (2) The resolution requirements for topographic objects in medium scales are in the order of a few meters, allowing the use of models of reduced complexity. (3) The ability to complete the otherwise automatic data acquisition using interactive techniques at digital photogrammetric workstations promises to increase the efficiency of classical acquisition techniques.

In contrast to previous approaches, this paper describes some attempts to derive the 3D form of buildings solely based on digital surface data. The geometric description of buildings seems to be an excellent intermediate representation for linking the sensor data with high-level knowledge about the objects. The advantage over techniques relying on the intensity data directly is the *invariance* of the geometric description with respect to surface markings (colour, texture), geometric microstructures (bricks, possibly windows) and illumination effects (especially shadows). Of course, reducing the available information to the surface geometry may cause difficulties in discerning buildings from non-buildings such as trees in case their geometric appearance mimics building-type objects. These cases could be resolved using some intensity information, e.g. colour. Nevertheless, it seems worthwhile to investigate the potential of techniques based purely on geometric information.

The next section describes the general strategy of our approach, Section 3 discusses the DEM generation. The detection (Section 4) is followed by the reconstruction (Section 5), which explicitly uses parametric models and prismatic building models within a new MDL-based procedure for generating polygonal ground plans. The techniques are explained using real data, among others also applied to the ISPRS WG 3 test data sets (Section 6).

2. General strategy

Our approach towards automatic building extraction is object related, i.e. we use a geometric model of the buildings and their geometric properties directly.

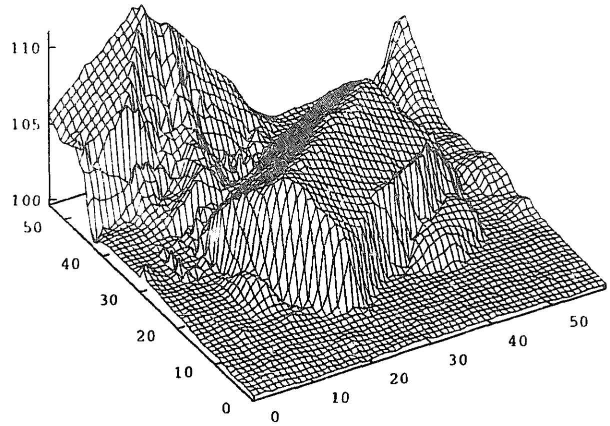

Figure 1. DEM automatically generated with MATCH-T, grid size 0.5×0.5 m^2, x, y (in pixels), z (in m).

The strategy of our approach consists of three steps: (1) automatic generation of a high resolution DEM, which delivers a graph surface as the description of the objects; (2) detection of buildings in the DEM; and (3) reconstruction of a parametric or prismatic geometric description for each detected building.

Automatic *generation* of a high-resolution DEM may use various sources of information. The only requirements are that the data are provided digitally — for ease of automization — and that the data are a sufficient dense description of the objects, e.g. providing a raster with increments between 0.5 and 5 m. We use data derived by a surface reconstruction technique based on digital (grey value) imagery, described in the next section. An example of the output of this technique applied for the purpose of building extraction is shown in Fig. 1.

The second step is the *detection* of buildings with the goal of focusing attention on areas where buildings can be expected in order to trigger the computationally more involved geometric reconstruction of the next step. We first compute an approximation of the topographic surface using mathematical morphology. The difference between the measured DEM and the approximate topographic surface contains the information about the buildings. The buildings are detected by thresholding the difference data set. The threshold is chosen according to prior knowledge about the buildings.

The third step is the *reconstruction* of the build-

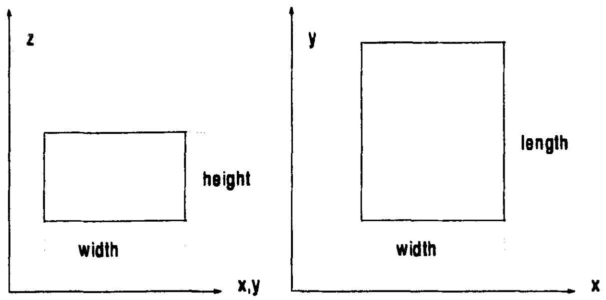

Figure 2. Building with flat roof.

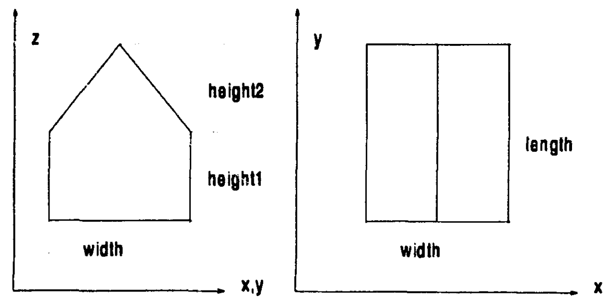

Figure 3. Building with symmetric, sloped roof.

ings. For this purpose different groups of models are used, dependent on the complexity of the detected buildings.

The first group of models consists of *parametric* models of the buildings. These models are used for simple buildings, which can be described using a few parameters. For these parametric models we assume that the buildings are separate from each other and that the ground plan of the building is a rectangle. Examples for this group of models are given in Figs. 2 and 3.

Complex buildings and blocks of buildings are described using *prismatic* models, which constitute the second group. These models are based on generic knowledge about the buildings. The first fact we use is that the ground plans of buildings or building blocks are *sets of closed polygons*. Furthermore, neighbouring straight lines of the buildings' outlines and therefore neighbouring edges of the polygons are likely to be orthogonal (Fig. 4, left). As shown in Fig. 4 (right) the outline of a building may be formed by several polygons, e.g. representing court yards.

The three steps sketched briefly here are discussed in more detail in the three following sections, including a discussion ot the needed parameters.

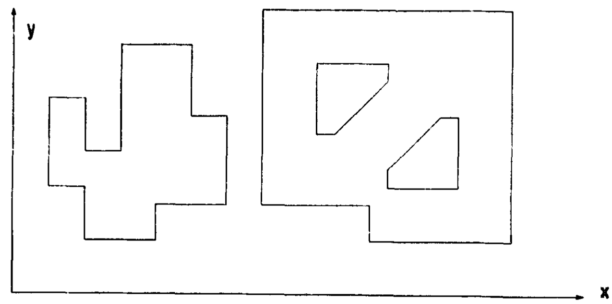

Figure 4. Ground plan of prismatic buildings.

3. DEM Generation

This section about DEM generation deals with DEMs as graph surfaces, thus surfaces represented by $z = z(x, y)$. Graph surfaces have conceptual deficiencies for our purpose of building extraction, e.g. vertical walls or passages in buildings cannot be represented (cf. the discussion above). Nevertheless, we use such a 2.5D description, because operational techniques yielding graph surfaces are available. Surface reconstruction algorithms may use several kinds of observations in order to compute the coordinates $[x, y, z(x, y)]^T$. The main techniques for point determination are based on the principle of radiation, e.g. using range finders, or the principle of intersection, e.g. using images. Both techniques show some deficiencies in the context of building extraction.

The first question that arises is: which surface is measured and can be reconstructed? In both cases, not only the topographic surface and the buildings are measured and represented in the DEM, but also objects which should not be included from the building extractor's point of view. These objects may be, depending on the applied technique, trees and bushes, but also cars. They are usually regarded as outliers in the data set, but they often cannot be distinguished from other small objects which may be of interest.

Another problem concerns the visibility of the surface to be reconstructed. In order to gain as much information as possible about this surface, the data sets should contain information about vertical walls of the buildings, while at the same time including information about the surface between the buildings. These requirements are quite different from classical topographic applications. Therefore, the measurement design has to cope with trade-offs concerning, e.g.: (a) flying height, influencing the precision and/or the resolution; (b) viewing angle, influencing the efficiency in terms of flying costs per area; (c) visibility of objects, influencing the accessibility of the desired information.

Laser scanning techniques (Krabill et al., 1984; Krabill, 1989; Lindenberger, 1993) have the advantage of directly gaining geometric information (point coordinates) about the surface and distinguishing between the tree canopy and the topographic surface beneath. Heights can be measured with an accuracy <0.5 m. In contrast to the above-mentioned requirements, laser scanners do not measure points on vertical surface parts.

Image-based techniques normally use stereo pairs of aerial images. In the centre of a model, it allows a look into streets and yards, while vertical walls of the buildings cannot been seen. On the other hand, walls can be seen at the border of the stereo model, but of course problems due to occlusions of objects arise.

The high resolution DEMs we used are generated using the software package MATCH-T. The principal strategy of this automatic approach is described in Krzystek (1991). The basic idea is to use image and feature pyramids. Starting with the lowest resolution (top of the pyramids) and a plane as approximation of the DEM, homologous features are matched and their 3D coordinates and a refined DEM are computed. This step is carried out for each pyramid level using the DEM of the previous step as approximation until the highest resolution (bottom of the pyramids) is reached. Tests of the approach show that the accuracy of heights in open areas is comparable to the accuracy of the measurements by a human operator (Krzystek and Wild, 1992). The algorithm was designed for topographic applications, treating objects such as buildings and trees as outliers.

The reconstruction of surfaces needs regularization (Terzopoulos, 1986). The main effects of standard regularization techniques are illustrated in Fig. 5, where the solid line represents the buildings and the dashed line their surface description in the DEM; edges of buildings are rounded off. In order to avoid this effect, local adaptive regularization techniques have to be applied (Terzopoulos, 1986; Weidner, 1994). Even if adaptive techniques are used, very small details may not be preserved in the generated DEM (see roof of the right building).

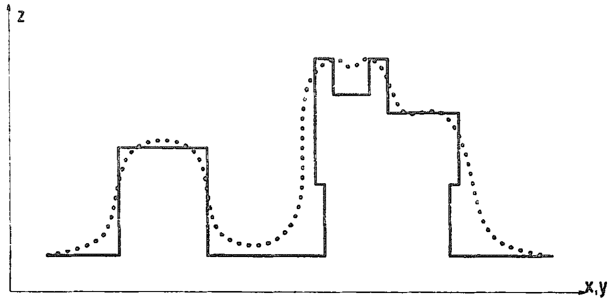

Figure 5. Real DEM skyline — x-axis: pixels with 5 m ground resolution; z-axis: height in metres.

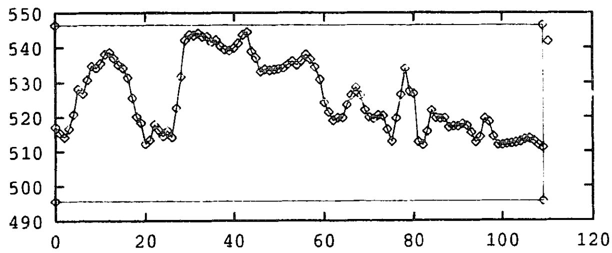

Figure 6. Skyline versus DEM showing deficiencies in describing buildings.

For the special purpose of DEM generation for building extraction, the control parameters for MATCH-T were adapted. Buildings are no longer seen as outliers, but as the information of interest. As mentioned above, problems arise concerning the discrimination of buildings and other objects, e.g. group of trees, which still should be regarded as outliers.

An example of a automatically derived DEM with a resolution of 0.5 m is given in Fig. 1. In order to give an impression of the DEM data in cities and with a lower resolution (5 m) Fig. 6 shows a longitudinal section through a DEM of a downtown area (see also Fig. 11, section is marked with black line).

4. Building detection

Our approach towards building detection is outlined in the flow chart in Fig. 7. It is based on the simple fact that buildings are higher than the (surrounding) topographic surface. The difference between original DEM and an approximation of the topographic surface without buildings contains the information about the buildings. Therefore we first compute an approximation of the topographic surface represented in the original DEM.

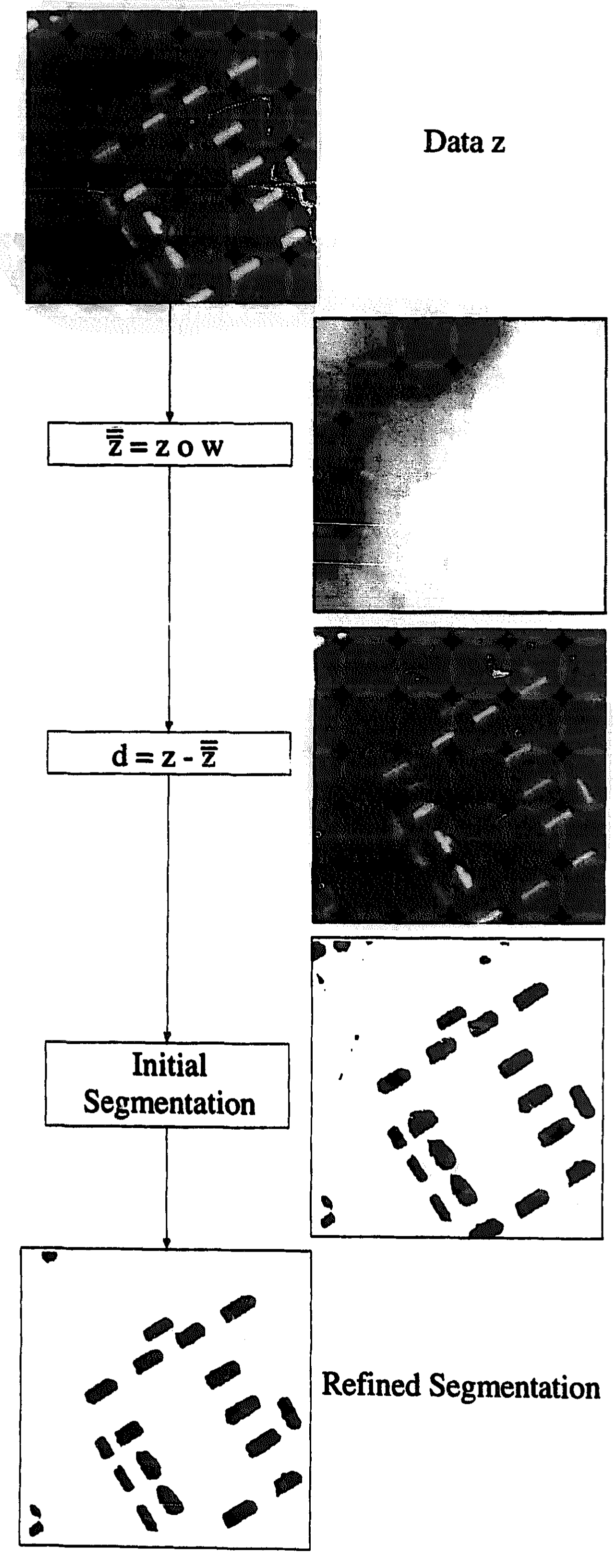

Figure 7. Building detection.

For this purpose we use mathematical grey-scale morphology (cf. Haralick et al., 1987). The first step is minimum filtering, i.e.

$$\bar{z} = \inf\{z(x, y) \mid x, y \in \mathcal{W}\} \tag{1}$$

with the structural element $\mathcal{W}$, in this case a square window. In terms of mathematical morphology, minimum filtering is a special *erosion*, where the structural element $w(x, y)$ has constant z values.

$$\bar{z} = z \ominus w \tag{2}$$

This step is followed by maximum filtering, i.e.

$$\bar{\bar{z}} = \sup\{\bar{z}(x, y) \mid x, y \in \mathcal{W}\} \tag{3}$$

with the structural element $\mathcal{W}$, which is a special *dilation*.

$$\bar{\bar{z}} = \bar{z} \oplus w \tag{4}$$

Both steps perform an *opening*

$$\bar{\bar{z}} = z \circ w \tag{5}$$

on the set of heights and deliver an approximation of the topographic surface.

In order to eliminate all information of the buildings in the resulting image of the opening, the window size has to be chosen in such a way, that the structural element $\mathcal{W}$ is not entirely contained in a building's outlines. The window size can be fixed using prior knowledge about the expected maximum size of the buildings or building parts. If the approximation for the topographic surface converges towards the original DEM in non-building regions and interpolates within the building regions, the difference data set $d = z - z \circ w$, i.e. heights of original DEM minus topographic surface approximation, consists of the buildings approximately put on a plane. This holds if the topographic surface is smoother than the surface with buildings.

The next step towards building detection consists of thresholding the difference data set. This is due to the fact that the buildings are put on a reference surface, which approximately is a plane. The threshold for segmentation can be easily motivated and derived from prior generic knowledge about the buildings. It can be chosen to be the expected height of vertical walls, e.g. the height of a floor. In order to identify the different segments during later processing, connected components are computed and each segment is labelled. Furthermore, the bounding box — with an additional margin — of each segment is computed.

Fig. 7 shows the result of the initial segmentation, including all segments found. It is obvious that some of the segments do not represent buildings, e.g. small segments representing trees. Based on the initial segmentation, we therefore select valid segments using as criteria (1) the size of the segment, and (2) the position of the bounding box. The minimal size of valid segments is related to the expected minimal size of buildings. We furthermore reject segments, if their bounding box is at the margin of the data set, because it is then likely that the building is not contained in the data set entirely and therefore neither the parameters for parametric models nor the prismatic models would be extracted correctly.

Fig. 7 also indicates that pure global thresholding is not optimal. Therefore, a *refined segmentation* for each valid segment is computed, which takes the data within the segment's bounding box into account and locally adapts the threshold based on the height information within the labelled area and the bounding box without segments.

5. Building reconstruction

Finally the detected buildings are reconstructed geometrically. The descriptions used for the buildings depend on their complexity. In our approach parametric models are used for simple buildings, with ground planes being rectangles separated from each other and describable using only a few parameters, whereas prismatic models are used for complex buildings or blocks of connected buildings. For both groups of models, the refined segmentation and the original data are used to extract the parameters or polygons and the heights of the buildings.

5.1. Parametric models

The *form parameters* for parametric models are the *length*, *width* and *height* for flat buildings or the *length*, *width*, *height of eave–base* and *height of ridge–eave* for buildings with a symmetric, sloped roof (cf. Figs. 2 and 3). We only integrated these two models in our implementation up to now. Besides these model-related parameters, *datum parameters*, namely the three-dimensional coordinates of a reference point and the orientation of the buildings in a reference coordinate system, are needed.

In order to determine the x, y coordinates of the reference point and the orientation of the buildings, the point of gravity and the orientation of each refined segment using the heights within the segments as weights are computed. The coordinates of the point of gravity and the orientation given in a DEM-related coordinate system have to be transformed into the reference coordinate system. The z coordinate is computed taking the mean of heights within the background area of the bounding box.

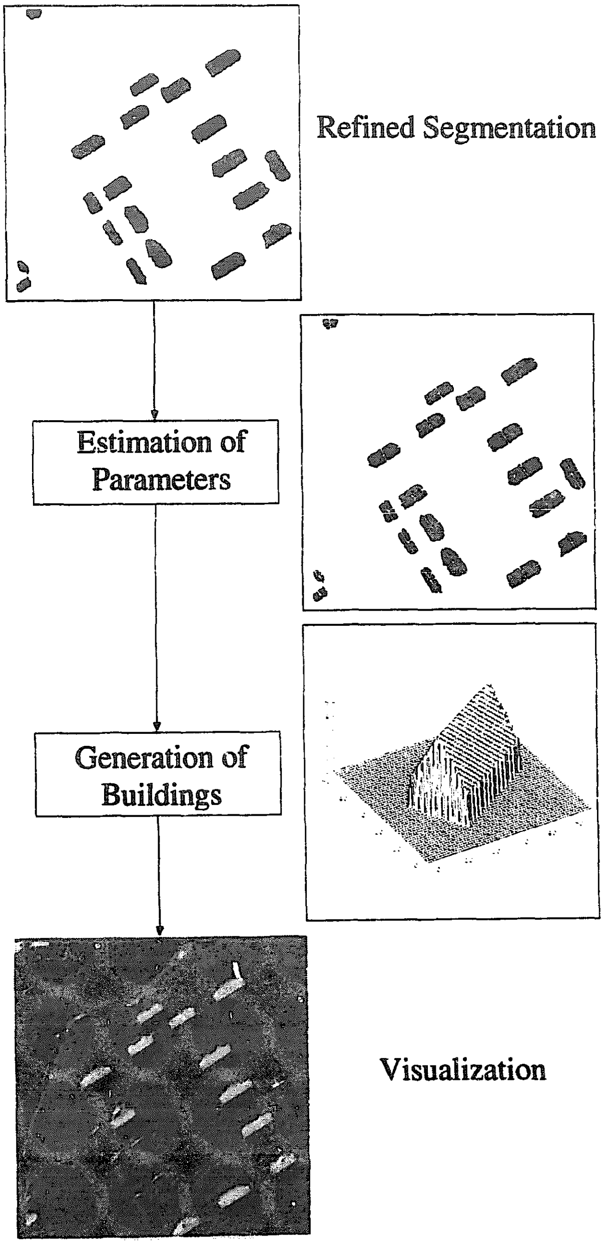

Figure 8. Reconstruction of parametric models.

Length and width of the ground plan are the length and width of a rectangle approximating the segment and are computed as follows: length = length of the segment along the first main axis, and width = width of the segment along the second main axis The other parameters, i.e. the heights, are model-dependent. Therefore, we first have to select the model which has to be applied. This selection uses the height information within a segment and prior knowledge about the minimal expected slope of roofs. If the slope within the segment is greater than the given minimal slope, the model with the symmetric sloped roof is chosen, otherwise the model of the flat building.

For buildings with flat roofs (see Fig. 2) the height parameter of the models is computed as: height = difference between the mean height of the segment and the mean height within the background area of the bounding box. For buildings with symmetric, sloped roofs (see Fig. 3) the height parameters are computed as: height1 = difference between the minimum height of the segment and the mean height within the background area of the bounding box; height2 = difference between the maximum and minimum height of the segment.

In order to improve the robustness of the computed height parameters mean values for the k maximum or minimum heights could be used.

5.2. Prismatic models

The input data for the extraction of the prismatic models' polygons is the refined segmentation. The boundaries of each segment is simplified and approximated using knowledge about the expected structure of the buildings. This is performed in several steps.

5.2.1. Raster to vector conversion

First the interior pixels of the segments are eliminated, giving their outlines in the grid. This step is followed by the computation of the number of outline polygons for each segment using connected components. The next step determines clockwise-ordered lists of the outline points with their x, y coordinates for each segment and outline polygon. These lists contain all outline points as shown in Fig. 9 (Vectorization). In order to reduce the number of points in this vectorization, points on straight lines between two neighbouring points are eliminated.

5.2.2. Discretization noise elimination

After the elimination of straight line points, a merging algorithm is applied in order to eliminate discretization noise. This merging algorithm successively eliminates points where the corresponding triangle height d_i in the triangle formed by

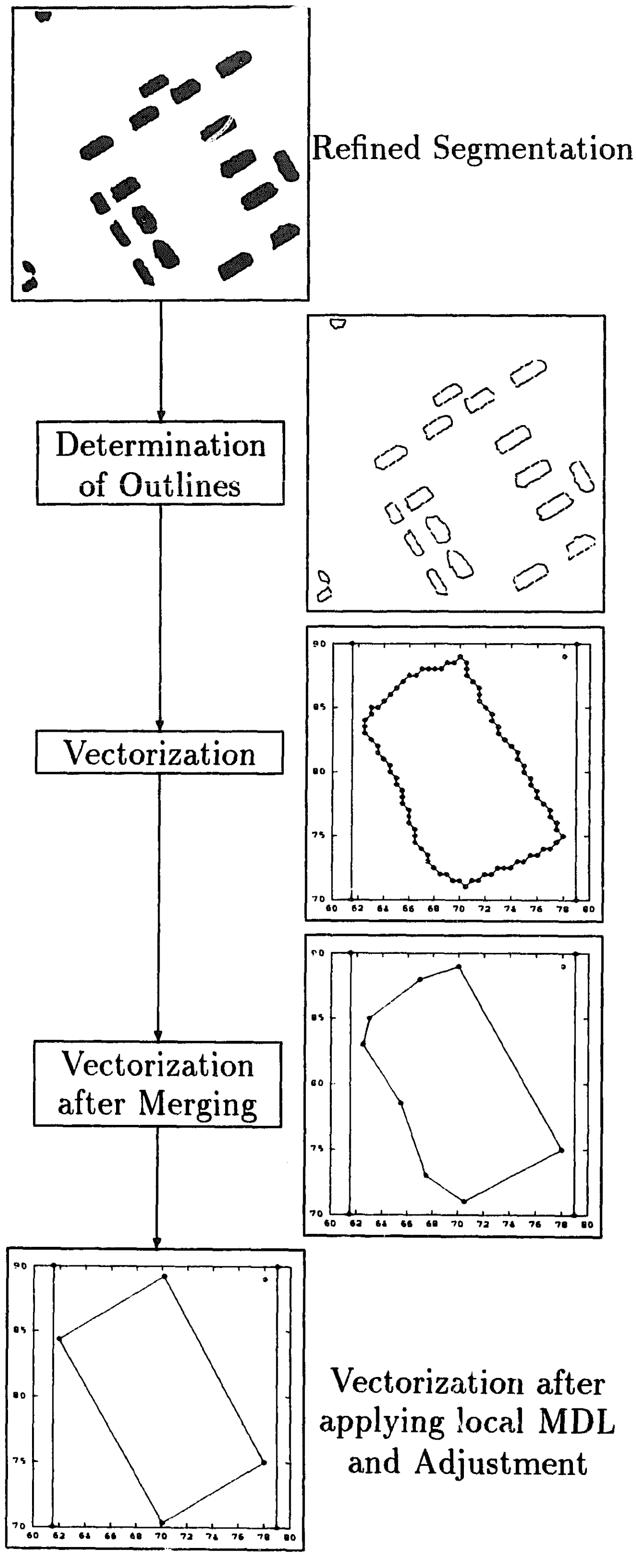

Figure 9. Reconstruction of ground plan of prismatic models.

the points $i-1$, i and $i+1$ is the minimum of the polygon until the minimum triangle height in an iteration is greater than a prefixed threshold, or the prefixed minimum number of points is reached. For this merging, the minimum height of a triangle has to be given as a threshold. This minimum height is closely related to the resolution Δx, Δy of the input data. If the minimum height $d_{\min}$ is chosen

$$d_{\min} < k\sqrt{\Delta x^2 + \Delta y^2}$$

only points lying approximately on a straight line are eliminated. We used $k = 2/3$.

5.2.3. MDL-based polygon simplification

Up to this point the shape extraction only makes use of the knowledge that the polygons are closed. The fact that neighbouring straight lines of the buildings' outlines and therefore neighbouring edges of the polygons are likely to be orthogonal is used for further reconstruction of the polygons using a minimum description length (MDL)-based approach.

The principle behind this second merging phase is to impose rectangle conditions on one or both angles at neighbouring points or to replace them by a single point, possibly introducing a rectangle at that point unless the description length cannot be further reduced. The description length depends on the mutual fit of the data and model and on the complexity of the model (cf. Rissanen, 1987; Förstner, 1989).

If a set of n observations l_i is fitted to a model

$$E(\boldsymbol{l}) = \boldsymbol{g}(\boldsymbol{\beta}), \quad D(\boldsymbol{l}) = \boldsymbol{\Sigma}_{ll} \tag{6}$$

with u' free unknowns β_j, the description length is

$$DL = \frac{\Omega}{2\ln 2} + \frac{u'}{2} lb\, n \tag{7}$$

with the weighted sum of the squared residuals

$$\Omega = [\boldsymbol{l} - \boldsymbol{g}(\boldsymbol{\beta})]^T \boldsymbol{\Sigma}_{ll}^{-1} [\boldsymbol{l} - \boldsymbol{g}(\boldsymbol{\beta})] \tag{8}$$

In case h constraints hold and u is the number of unknowns we have $u' = u - h$ free unknowns. The second term $\frac{u'}{2} lb\, n$ in (7) takes the complexity of the model into account: models with more parameters get a penalty as they are expected to decrease Ω.

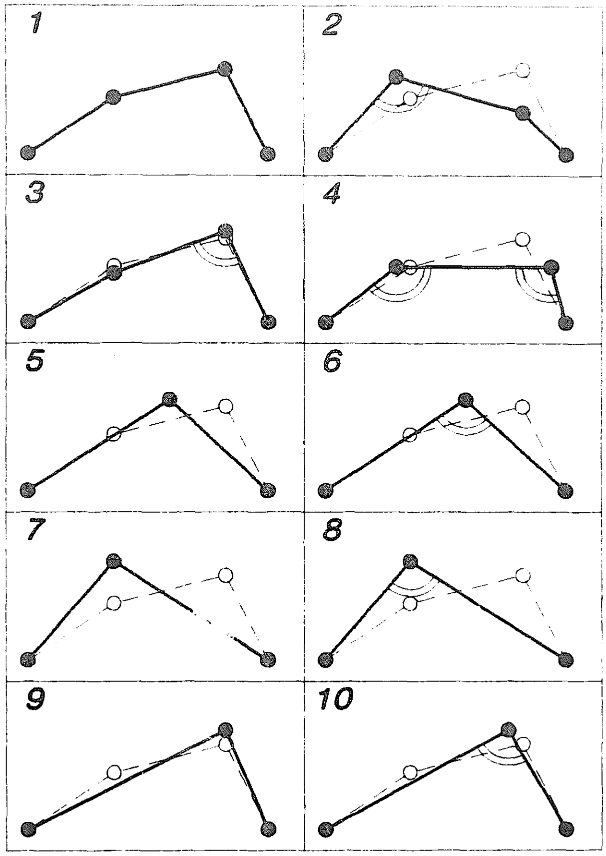

Figure 10. Ten alternatives for configuration 1. The result of the first iteration of an adjustment with weak constraints is shown with bold points. Optimum: case 10.

In our case the description length is determined locally by analysing four consecutive points (see Fig. 10). Points 1 and 4 are assumed to be fixed, while points 2 and 3 are to be changed. In cases 1 to 4 the points are replaced, thus $n = 4$ and $u = 4$. In cases 5 to 10 points 2 and 3 are replaced by one. In cases 5 and 6 the mid-point of 2 and 3 is used as given, in cases 7 and 8 the left point 2 is assumed to be given, thus point 3 is to be eliminated, and in cases 9 and 10 point 3 is assumed to be given, thus point 2 is to be elimated. In these cases $n = 2$ and $u = 2$. The number h of conditions is one more than the number of imposed rectangle constraints and varies between 1 and 3, because in all cases the area of the polygon is assumed to be constant. Fig. 10 summarizes the characteristics of the ten cases and especially indicates where the rectangle condition is located. The result of an iterative application of this merging procedure onto the discretization-noise-cleaned polygon is shown in Fig. 15.

TABLE 1

Characteristics of alternative hypotheses for edge merging, normalized sum of squared residuals Ω_i and description length DL_i for the 10 cases in Figure 10, optimal configuration: 10 with $MDL = DL_{10} = 0.71$ [bits]

i	Appr. val.	$\perp$ at	n	u	h	$u-h$	$0.5 \times (u-h) lb\, n$	$\Omega_i/2\ln 2$	DL_i
1	–	–	4	4	1	3	3	0.00	3.00
2	p_2, p_3	p_2	4	4	2	2	2	376.16	378.16
3	p_2, p_3	p_3	4	4	2	2	2	3.46	5.46
4	p_2, p_3	p_3, p_3	4	4	3	1	1	1393.33	1394.33
5	$(p_2+p_3)/2$	–	2	2	1	1	0.5	1.57	2.07
6	$(p_2+p_3)/2$	$(p_2+p_3)/2$	2	2	2	0	0	177.39	177.39
7	p_2	–	2	2	1	1	0.5	4.53	5.03
8	p_2	p_2	2	2	2	0	0	124.11	124.11
9	p_3	–	2	2	1	1	0.5	0.39	0.89
10	p_3	p_3	2	2	2	0	0	0.71	**0.71**

5.2.4. Final optimization of the ground plan

The optimal coordinates of this polygon are obtained by a final adjustment. All edges are grouped by building hypotheses about rectangles, parallel or possibly collinear edges. A robust estimation eliminates possibly wrong hypotheses (Fuchs and Förstner, 1995). This estimation process fuses the boundary information from the discretization-noise-cleaned data and the inferred structure from the MDL-based form of the boundary. Figs. 9 and 16 show the results of this final step.

5.2.5. Height determination

Finally, the heights for the prismatic models are computed analogous to the heights of flat buildings in the previous subsection. If parts of a block of buildings have different heights, further segmentation within the polygons is of course necessary.

6. Examples

Some results of our approach are included in the flow charts given in Figs. 7, 8 and 9. They show the results for the ISPRS Commission III Working Group 3 test data set *FLAT* (Fritsch et al., 1994) with a DEM resolution of 0.5 m in x- and y-direction. The data set contains simple buildings being well separated. In this case parametric models are preferable to prismatic models, which is evident by a qualitative comparison of the results, as they incorporate more specific prior knowledge than prismatic models.

We also used a high resolution DEM of a downtown area (Fig. 11) for a first test of our approach for prismatic models. Here the DEM has a resolution of 5 m in each direction. It contains complex buildings, blocks of buildings and separate simple buildings. The segmentation of the data set is shown in Fig. 12. The quality of the segmentation obviously highly depends on the regularization technique used for the reconstruction of the DEM and the grid resolution. Both heavily influence the

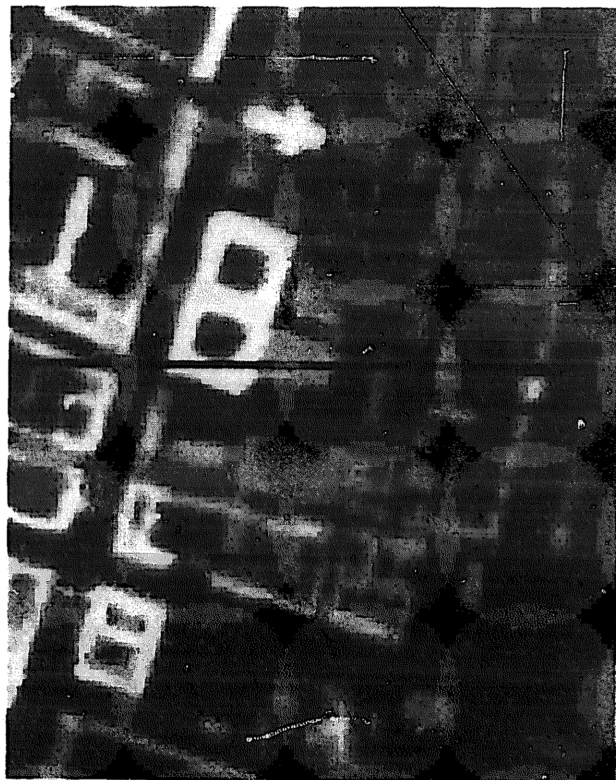

Figure 11. DEM of downtown area.

Figure 12. Refined segmentation.

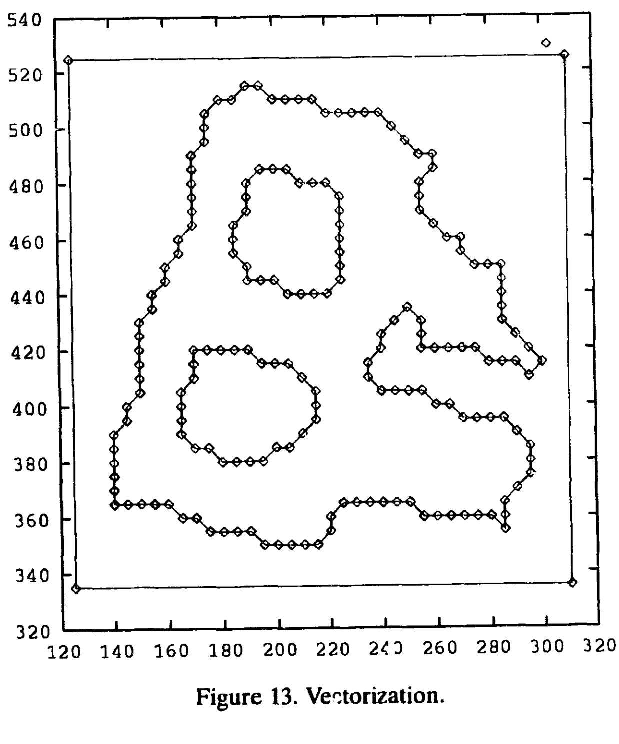

Figure 13. Vectorization.

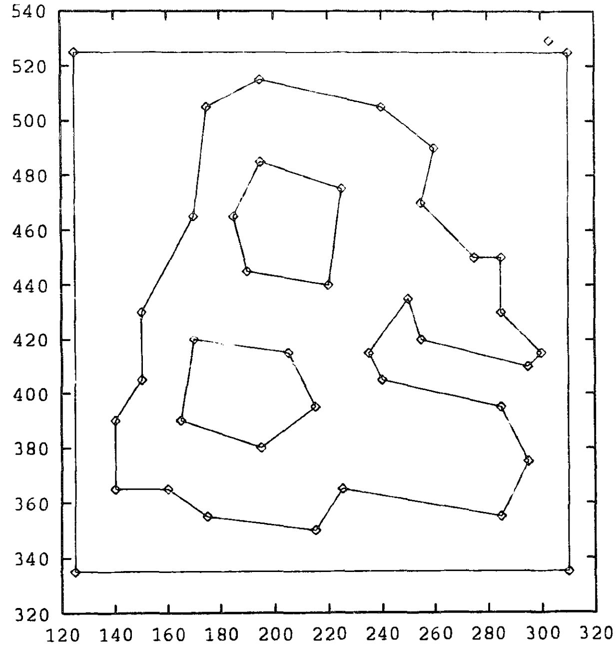

Figure 14. Vectorization after elimination of discretization noise.

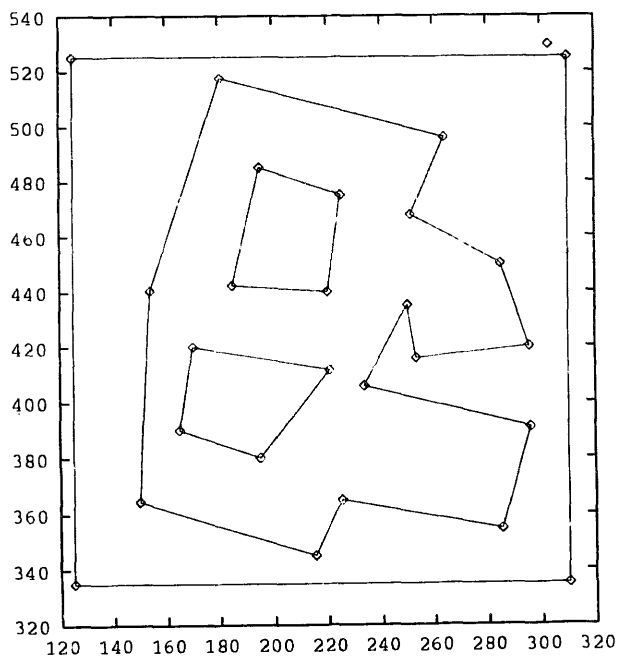

Figure 15. Vectorization after MDL.

separability of buildings and the polygons of each building. E.g., single buildings in a block of buildings may be only one or two pixels large.

Figs. 13–16 show the initial vectorization, the polygons after merging, local application of MDL and the final reconstruction of the ground plan. Taking into account the DEM resolution of 5×5 m^2 the result is quite remarkable as it resembles the basic structure of the complex building with two court yards.

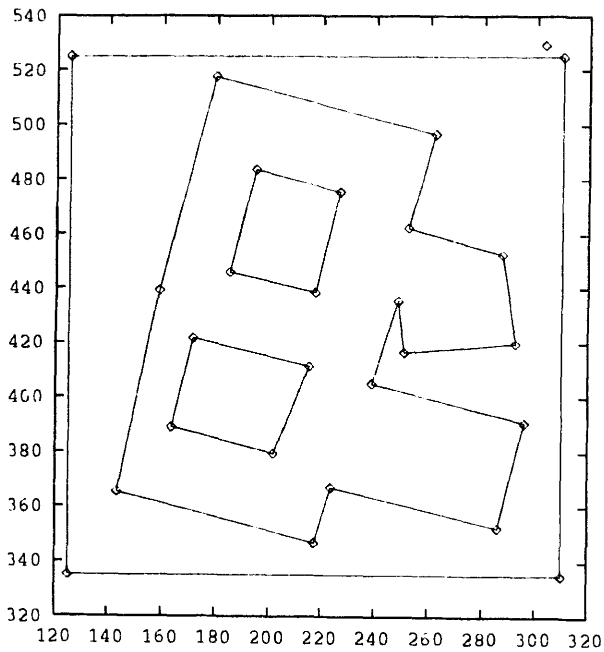

Figure 16. Final vectorization of ground plan.

The upper of the two exclaves, however, is a group of trees mimicking to be part of the building. As the complete analysis assumes the segment to be a building, it also looks for and actually finds rectangles in this section of the segment: the system only sees what it is taught to see. Here a simple analysis of the colour of the corresponding image area would solve this problem.

7. Conclusions

We discussed steps towards automatic building extraction from high-resolution Digital Elevation Models. It is part of our research into techniques for deriving 3D city models integrating 2D and 3D data together with semantic knowledge about the scene (cf. Braun et al., 1995). The approach consists of automatic DEM generation, detection of buildings and extraction of a description for the buildings. The detection and reconstruction of buildings is based on generic contextual knowledge. This knowledge is represented in geometric building models, parametric ones for simple buildings, which can be described by a few parameters, and prismatic models for complex buildings and blocks of buildings. We applied our approach for building extraction on real data sets: one data set with separate simple buildings, the other data set with a complex building in a down town area.

The results for the first data set show the capability of our approach when dealing with such simple buildings. Further work for parametric models will focus on the integration of other parametric models such as buildings with non-symmetric sloped roofs or buildings with hip 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 and robust statistics used for the estimation of the height parameters. Nevertheless, the resolution of the parameters will always depend on the resolution of the DEM 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. Instead of a local application of MDL, it is planned to apply it globally looking for parallel edges of the entire polygon and of the set of polygons belonging to a building simultaneously while relating to the original data in *all* steps. In order to deal more appropriately with complex buildings consisting of parts with different heights, 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). Furthermore other constraints, e.g. symmetries, and semantic knowledge about rows of buildings, will be investigated.

Acknowledgements

This work was partly supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 350, "Wechselwirkungen kontinentaler Stoffsysteme und ihre Modellierung" and "Gebäudeextraktion" (Fo 180/2-1).

References

Braun, C., Kolbe, T.H., Lang, F., Schickler, W., Steinhage, V., Cremers, A.B., Förstner, W. and Plümer, L., 1995. Models for photogrammetric building reconstruction. Comput. Graphics, 9(1): 109–118.

Förstner, W., 1989. Image analysis techniques for digital photogrammetry. In: Proc. 42nd Photogrammetric Week, Stuttgart, pp. 205–221.

Fritsch, D., Sester, M. and Schenk, T., 1994. Test on image understanding. In: ISPRS Comm. III Symp. Spatial Information from Digital Photogrammetry and Computer Vision. SPIE, Proc. pp. 243–248.

Fua, P. and Hanson, A.J., 1987. Resegmentation using generic shape locating general cultural objects. Pattern Recognition Lett., 5: 243–252.

Fuchs, C. and Förstner, W., 1995. Polymorphic grouping for image segmentation. 5th ICCV, 1995, Boston, pp. 175–182.

Haala, N., 1994. Detection of buildings by fusion of range and image data. In: ISPRS Comm. III Symp. Spatial Information from Digital Photogrammetry and Computer Vision. SPIE, Proc. pp. 341–346.

Haralick, R.M., Sternberg, S.R. and Zhuang, X., 1987. Image analysis using mathematical morphology. IEEE T-PAMI, 9: 523–550.

Herman, M. and Kanade, T., 1986. The 3D mosaic scene understanding system. In: A.P. Pentland (Editor), From Pixels to Predicates. Ablex Publ. Co., Norwood, N.J., pp. 322–358.

Krabill, W.B., Collins, J.G., Link, L.E., Swift, R.N. and Butler, M.L., 1984. Airborne laser topographic mapping results. PERS, 50(6): 685–694.

Krabill, W.G., 1989. GPS application to laser profiling and laser scanning for digital terrain models. In: Proc. 42nd Photogrammetric Week, Stuttgart, pp. 329–340.

Krzystek, P., 1991. Fully automatic measurement of digital elevation models. Proc. 43nd Photogrammetric Week, Stuttgart, pp. 203–214.

Krzystek, P. and Wild, D., 1992. Experimental accuracy analysis of automatically measured digital terrain models. In: W. Förstner and S. Winter (Editors), Robust Computer Vision. Wichmann, Karlsruhe, pp. 372–390.

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

Lin, C., Huertas, A. and Nevatia, R., 1994. Detection of buildings using perceptual grouping and shadows. In: Proc. Computer Vision and Pattern Recognition, pp. 62–69.

Lindenberger, J., 1993. Laser-Profilmessungen zur topographischen Geländeaufnahme. Deutsche Geodätische Kommission, München, Band 400, Reihe C.

McGlone, J.C. and Shufelt, J.A., 1994. Projective and Object space geometry for monocular building extraction. In: Proc. Computer Vision and Pattern Recognition, pp. 54–61.

Rissanen, I., 1987. Minimum description length principle. Encycl. Stat. Sci., 5: 523–527.

Terzopoulos, D., 1986. Regularization of inverse visual problems involving discontinuities. IEEE T-PAMI, 8(4): 413–424.

Weidner, U., 1994. Parameterfree information-preserving surface restoration. In: J.-O. Eklundh (Editor), Computer Vision — ECCV 94, Vol. II, Proc., pp. 218–224.