A Procedure for Segmenting Surfaces by Symbolic and Iconic Image Fusion

Ansgar Brunn, Felicitas Lang, Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn
Nußallee 15, 53115 Bonn
e-mail: *first_name.last_name*@ipb.uni-bonn.de

Abstract. This paper deals with the derivation of a symbolic surface description exploiting the information of multiple images while using a minimum of domain knowledge. We present a new concept for segmenting surfaces by fusing multiple images both on the iconic and on the symbolic description level. In a first step a local 3D-reconstruction and interpretation is derived based on the result of a polymorphic feature extraction. It serves as prior information for a second step which refines the initial segmentation using the radiometric image content. Examples of the proposed procedure are presented for the application of 3D-building reconstruction from aerial images.

0 Introduction

The paper describes a procedure for segmenting surfaces by fusing multiple images. It is motivated by the research in the area of extracting the 3-dimensional shape of buildings from aerial images and range data for deriving or updating 3D-geoinformation systems especially in urban areas (cf. the excellent overview on the state of the art in [Grün *et al.* 1995]).

The problem we encounter in this paper is the derivation of a symbolic surface description by exploiting the information of several, generally more than two images and to use a minimum of domain knowledge. We assume the orientation of the images and the calibration parameters of the cameras to be known, in order to fully exploit the geometric constraints between the images. We also assume the object to be polyhedral, possibly showing regularities such as rectangles between neighboring 3D-edges or partial symmetry of faces. In our context the images actually are analogue images which are scanned for automatic processing, which however is no restriction for the procedure.

Due to the complexity of natural scenes, especially also buildings causing strong occlusions and lack of texture we cannot expect classical stereo algorithms, working on two images or just relying on edges, to give enough information on the 3D-shape of the objects for deriving a complete surface description. Exploiting rich image descriptions therefore appears to be a promising approach (cf. [Bignone *et al.* 1996], [Fuchs and Förstner 1995]).

The basic idea of our approach therefore is to derive a polymorphic symbolic description S_o of the object from I image descriptions, i. e. segmentations $\mathsf{S}_i, i = 1, ..., I$. Formally the problem is identical to estimating an unknown 3D-description $\widehat{\mathsf{S}}_o$ given the 2D-descriptions S_i based on the projection model:

$\mathsf{S}_i + \widehat{\mathsf{D}}_i = f_i(\widehat{\mathsf{S}}_o)$. $f_i(\mathsf{S}_o)$ describes the ideal projection of the 3D-structure, that is the ideal image model. D_i denotes disturbances which for example are due to occlusions and the characteristics of the applied feature extraction. The estimation of S_o leads to an optimization problem which can be solved e.g. by selecting $\widehat{\mathsf{S}}_o$ which maximizes the probability $P(\widehat{\mathsf{S}}_o \mid \{\mathsf{S}_i\})$.

The main reason for working on the symbolic level is the inclusion of all types of domain specific knowledge at all levels of the analysis which in most cases is of symbolic nature. On the other hand one cannot expect a procedure for deriving a symbolic image description to be complete enough without being guided by higher level knowledge and especially local adaption to the objects actually appearing in the images. Therefore iconic and symbolic information extracted from the images needs to be fused simultaneously in order to fully exploit the image content.

The paper shows how this fusion of iconic and symbolic image information can be achieved and integrated into a procedure for reconstructing a symbolic surface description. Section 1 gives an overview on the general strategy for integrating image and object information on the different levels of abstraction. The fusion of symbolic image descriptions is discussed in section 2 using as example the reconstruction of 3D-vertices. Section 3 shows how to fuse the regions obtained from the feature extraction for deriving 3D-regions based on preknowledge from the geometric reconstruction of the vertices and a physical model of the imaging process. The examples demonstrate the ability of the approach to overcome deficiencies in the individual symbolic image descriptions such as over and under segmentations.

1 General Strategy

1.1 Rationale for a Solution

The following general rationales are used to cope with the complexity of the 3D-reconstruction:

Levels of description. We work on several levels of abstraction in order to adapt to the different depth, granularity and generality of the used domain knowledge ([Braun *et al.* 1995]). Therefore we use the iconic information (original or transferred raster data), list of features, relational descriptions of atomic or aggregated features. We choose the highest level of representation which is reachable in order to exploit as much invariant information as possible and to constrain search. This is the main motivation for developping scalable domain independent tools for feature extraction (FEX) and image segmentation ([Förstner 1994], [Fuchs and Förstner 1995]).

Rationales for processing. Conservative decisions decrease the need for backtracking thus speed up processing. They however require the well balanced and increasing introduction of domain specific knowledge. Describing the imaging process in order to adapt thresholds in feature extraction and formalizing the domain specific but generic structure of the surface in order to obtain labels for

the interpreted image features are two ends of the modeling chain. We join data and model driven reasoning, but change early to a high symbolic level for hypothesis finding (cf. above) where again the highest level is taken for verification. One of these links is described in detail in sections 2 and 3.

Tracking uncertainty. An essential feature of all automatic procedures is their ability to allow quantitative evaluation [Förstner 1996]. This allows control on the progress during development of an algorithm, empirical testing with simulated or real data as well as performance prediction. Maximizing $P(\widehat{\mathsf{S}}_o|\{\mathsf{S}_i\})$ $= P(\{\widehat{\mathsf{S}}_i\}|\mathsf{S}_o)\cdot P(\mathsf{S}_o)/P(\{\mathsf{S}_i\}) = \prod_i P(\mathsf{S}_i|\widehat{\mathsf{S}}_o)\cdot P(\widehat{\mathsf{S}}_o)/P(\{\mathsf{S}_i\})$ requires an explicit modeling of the object ($P(\widehat{\mathsf{S}}_o)$) and the *appearance* ($P(\mathsf{S}_i|\widehat{\mathsf{S}}_o)$)of the object in the images, taking into account effects of occlusion, noise (cf. [Steinhage 1995]) and feature extraction (cf. [Fuchs *et al.* 1994]). In general it is necessary to simultaneously estimate the relevant parameters and classify the structures.

1.2 Procedure

Our procedure consists of the following steps:

1. Feature extraction. We derive a rich symbolic image description consisting of attributed points $\mathcal{P}$, lines $\mathcal{E}$ and regions $\mathcal{R}$ together with their mutual relations R contained in a feature adjacency graph (FAG) which allows further analysis. The procedure can be controlled by a.) a selection of feature and relation types, if appropriate and b.) a small set of tuning parameters, which either can be derived from the object and the sensor model, especially scale parameters for extracting edges and points or are statistical parameters, especially the significance level for determining the internal thresholds, which themselves depend on the estimated noise level.

2. Local structures are derived from the FAG. These are point, line and region induced structures namely vertices $\mathcal{V}_i = \mathcal{A}^{\mathsf{P}}{}_i$, wings $\mathcal{W}_i = \mathcal{A}^{\mathsf{E}}{}_i$, cells $\mathcal{C}_i = \mathcal{A}^{\mathsf{R}}{}_i$ containing all neighbors and possibly indirect neighbors. We found vertices $\mathcal{V}_i$ and cells $\mathcal{C}_i$ appear to be most appropriate for matching, as vertices carry strong geometric information and cells give strong evidence for surface patches $\mathcal{C}_o$. Both contain neighboring edges, which helps both linking and consistent geometric reconstruction.

3. Vertex reconstruction is initiated in order to come to a first 3D-description of the object. Each reconstructed vertex consists of the 3D-point, 2 or 3 neighboring edges and the face(s) between the edges. The reason to start with the reconstruction of 3D-vertices is the strong constraints available from the imaging geometry. Here weak domain specific knowledge is used: namely vertices joining three cells in 3D to exist. Therefore the fusion is performed on the symbolic level. In case strong domain specific knowledge is available, such as rectangularities or symmetries between lines and regions, this can be incorporated by simultaneously classifying and reconstructing the 3D-structure ([Lang and Förstner 1996]). The vertex reconstruction conceptually works in parallel on all vertices found in the images. Later a grouping of 3D-vertices takes place where neighboring vertices, that share a common edge in at least one of the images, are linked.

The information achieved so far will be incomplete, due to vertices and mutual relations not found.

4. Cell reconstruction follows in order to bridge the information found so far. Based on the geometry of the reconstructed vertices and the analysis of the FAG, hypotheses on surface regions R_{oj} lying in planes are built. By back projecting the images to the predicted surfaces and using a physical model of the imaging process the iconic, i. e. the radiometric information of the images is used to find consistent surface elements in the predicted planes. This is a generalization of the technique to detect obstacles in image sequences by analyzing the images after projective mapping of the image to the ground plane [Storjohann *et al.* 1988]. The result of this fusion process are 3D-regions $\widehat{\mathsf{R}}_{oj}$ which need to be linked. These last two steps are now described in more detail.

2 Fusion of Symbolic Images

Based on a polymorphic feature extraction (cf. [Förstner 1994]) resulting in a relational image description $\widehat{\mathsf{S}}_i$, building-specific basic aggregates, are derived by integrating a generic building model.

We start the reconstruction based on point-induced local aggregates derived directly by analyzing the FAG (cf. fig. 1). 3D-reconstruction thus is based on sets $\mathcal{V}_i = \mathcal{A}^{\mathsf{P}}{}_i$ of vertices, as for these local aggregates the imaging geometry leads to sharp constraints within solving the correspondence problem. For each image i the j-th vertex V_{ij} is represented by its features including its geometric attributes and the relations R_{ij} between those features.

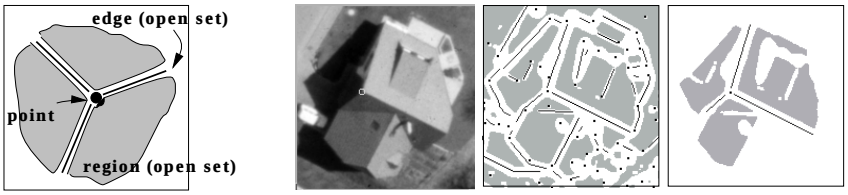

Fig. 1. shows the topology of an ideal vertex being a point induced aggregate (a.), a feature point in the image (b.), all extracted features (c.) and an example of an extracted vertex as it can be derived from the FAG (d.)

The solution is found by following the **hypothesize and verify** paradigm:

The first step consists in the data driven generation of a set $\{\mathsf{V}_{ij}, \mathsf{V}_{i'j'}\}$ of corresponding vertices leading to a set of **hypotheses** for 3D-vertices $\mathcal{V}_o = \{\mathsf{V}_{oj}\}$. The correspondence is established using a cost function which on the one hand evaluates the probability of the vertex V_{ij} corresponding to a *generic* vertex model $\tilde{\mathsf{V}}_o$ and which on the other hand evaluates the structural error of the vertices $\{\mathsf{V}_{ij}\}$. The vertex model contains not only the basic features $\mathcal{F}$

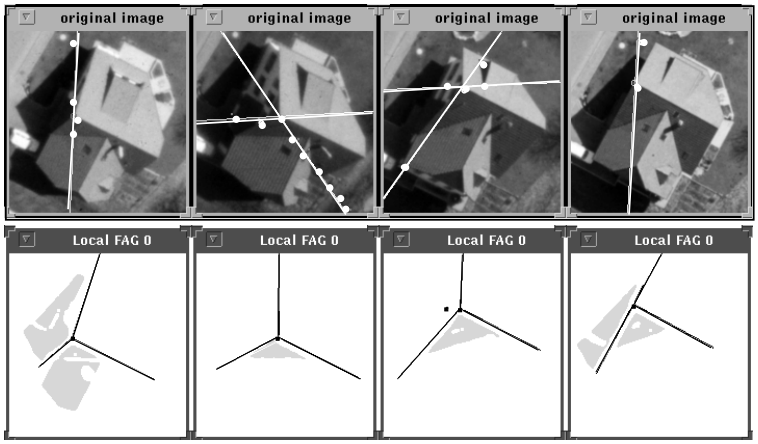

Fig. 2. Selected basic aggregates are automatically reconstructed in 3D. The first row shows the original images, selected vertices and their corresponding epipolar lines. The second row shows the point-induced aggregates V_{ij} and the 3D-reconstruction of the edges projected back into the images.

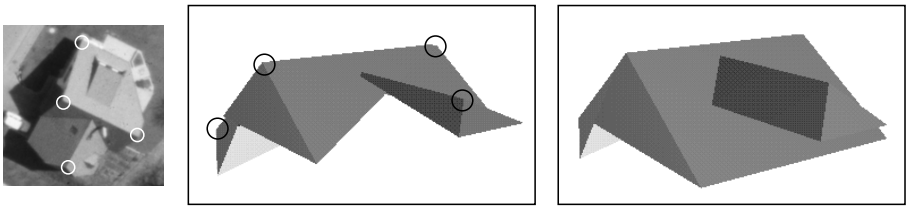

Fig. 3. Selected vertices are automatically reconstructed in 3D: **a.** original image; **b.** partially reconstructed roof in 3D composed by different vertices, which are denoted by circles; **c.** result of a 3D-grouping process, which connects the reconstructed 3D-parts of the building by analysing corner point and line identity and symmetry properties.

and their relations R, but also their domain dependent type and the geometric constraints.

The second step is the model driven **verification** of the hypotheses $\mathcal{V}_o$ by integrating 3D-scene knowledge of a building specific vertex model $\tilde{\mathsf{V}}_o$. Each hypothesis V_{oj} is semantically interpreted and classified to belong to vertex type $\tilde{\mathsf{V}}_o$. The interpretation is evaluated by maximization of the probability for a given vertex.

$$P(\tilde{\mathsf{V}}_{oj} \mid \{\mathsf{V}_{ij}\}) = \frac{P(\{\mathsf{V}_{ij}\} \mid \tilde{\mathsf{V}}_{oj}) P(\tilde{\mathsf{V}}_{oj})}{P(\{\mathsf{V}_{ij}\})} \qquad (\mathrm{j}\ fixed) \tag{1}$$

For evaluation of the geometry, e. g. of its geometric parameters $\boldsymbol{g}$, the classical modeling techniques of observation errors can be used. After model-based parameter estimation, the evaluation can be derived from the residuals $\boldsymbol{y} - f(\widehat{\boldsymbol{g}})$ of the optimal estimation $\widehat{\boldsymbol{g}}$ for $\boldsymbol{g}$ using the *probability density function* $p(\boldsymbol{g})$ in case the feature F exists and has been successfully matched to the model.

$$p(\boldsymbol{g}(\mathsf{F})|\exists(\mathsf{F}) = \text{matched}, \tilde{\mathsf{V}}_o) = \frac{1}{(2\pi)^{n/2}\det\boldsymbol{\Sigma}_{yy}} \exp(-\frac{1}{2}(\boldsymbol{y}-f(\widehat{\boldsymbol{g}}))^T \boldsymbol{\Sigma}_{yy}^{-1}(\boldsymbol{y}-f(\widehat{\boldsymbol{g}}))) \,. \tag{2}$$

The evaluation of the existence of the features an their relations has been studied earlier (cf. [Fuchs *et al.* 1994]).

Figure 3 shows results of the fully automatic 3D-vertex reconstruction which is used as prior information for the iconic fusion. More details of this approach are described in [Lang and Förstner 1996].

3 Fusion of Iconic Images

The reconstructed vertices contain hypotheses on 3, possibly only 2, neighboring regions of the objects surface. These regions will be under and over segmented due to occlusions and texture. We now want to find better estimates for the shape of these regions using the iconic information of the original images.

For each of these regions $\{\mathsf{R}_{oj}\}$ we assume a geometric one-to-one mapping $M : (u, v) \mapsto (r, c)$ from the surface to each image to be known, which easily can be derived from the geometry of the vertex inducing that region. We now perform a classification of the surface elements $g_o(u, v)$ based on the residuals of a joint estimation of $g_o(u, v)$, the radiometric transformation between the object and the images, and the already found regions in the images.

The model reads as for each of the I images $i \in \mathcal{I}$

$$g_i(r, c) \sim N(a_i + b_i g_o(r, c), \sigma_g^2) \quad \text{if}(r, c) \in M(\mathsf{R}_{oj}) \text{ and not occluded} \tag{3}$$

$$g_i(r, c) \sim \quad Eq(0, 255) \quad \text{else} \tag{4}$$

where N and Eq denote normal and equal densities resp. and $(r, c) = M^{-1}(u, v)$. We need approximate values for the unknown transformation parameters a_i, b_i, the reflectance at the surface $g_o(u, v)$, and the surface region R_{oj}, and the occlusions. Therefore we proceed in the following way.

1. Mean image. We take an approximation

$$\overline{g}_o(u, v) = \frac{1}{I}\sum_{i=1}^{I} g_i(u, v) \tag{5}$$

for the intensity $g_o(u, v)$, which requires back projection of the images (cf. fig. 4a).

2. Approximation for the region R_{oj}. The easiest way of obtaining a first approximation of the unknown region is to take the union

$$\mathsf{R}_{oj}^{(1)} = \bigcup M^{-1}(\mathsf{R}_{ij}) \tag{6}$$

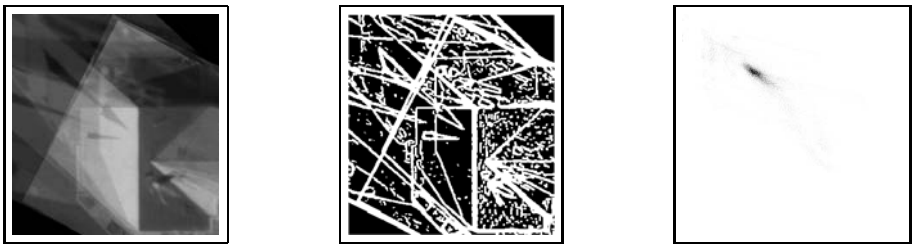

Fig. 4. a.) The first approximation $\bar{g}_o(u,v)$ for the intensities $g_o(u,v)$ of the example shown in fig. 1b, cf. eq. (5). Observe the diffused presentation of the chimney compared to fig. 1. b.) The extracted regions in the mean image in fig. 4a. c.) 2D histogram of the mean image $g_o(u,v)$ and image $g_1(u,v)$. It demonstrates that a linear regression model eq. (3) is suitable.

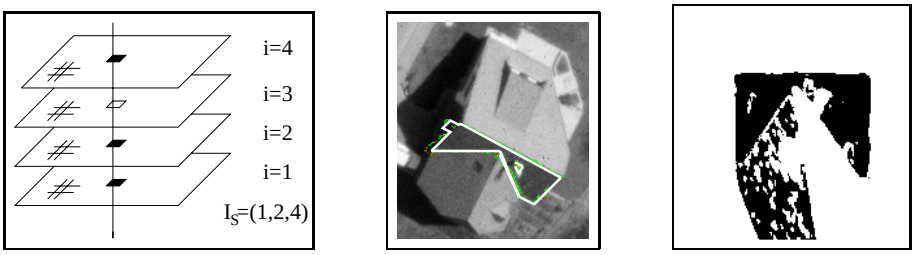

Fig. 5. a.) Subset $\mathcal{I}_S = (1,2,4)$ used to build a test statistic. b.) projection of the final result in one of the original images. c.) addition of the regions in the back projected image, maximum 2 outlier.

of all backprojected regions $M^{-1}(\mathsf{R}_{ij})$ induced by the vertex. An alternative would be to perform the feature extraction on the vector valued image $\boldsymbol{g}(u,v) = (g_i(u,v))$, or immediately on $\overline{g}_o$ from (5). The properties of these choices need to be analyzed. We used the last option, applying FEX to $\overline{g}_o$ (cf. Figure 4b) and selecting the region having most overlap with the regions of the individual FEX.

3. Surface region of interest. The regions obtained from FEX are defined by the homogeneity of the intensity function and therefore are a conservative estimate for the true region. We therefore dilate and close the region $\mathsf{R}_{oj}^{(1)}$ to obtain a region of interest:

$$\overline{\mathsf{R}}_{oj}^{(1)} = D_1 \bullet D_2 \oplus \mathsf{R}_{oj}^{(1)} \tag{7}$$

with the radius of D_1 depending on the expected radius of holes caused by irregular texture and the radius of D_2 depending on the scales used in FEX. In the examples we used $r_1 = 11$ and $r_2 = 2$ corresponding to the scale $\sigma = 1$ for edge detection. $\overline{\mathsf{R}}_{oj}^{(1)}$ is a too optimistic estimate of the surface region to be found and needs to be confirmed by the intensity information of the images.

4. Image regions of interest. In the images we select all regions that overlap

with the projection of $\overline{\mathsf{R}}_{oj}^{(1)}$ as regions of interest, i. e. as possibly supporting the unknown region:

$$\mathcal{A}_i^{\mathsf{R}_j} = \left\{ \mathsf{R}_{ij} | \mathsf{R}_{ij} \cap M(\overline{\mathsf{R}}_{oj}^{(1)}) \right\} \tag{8}$$

5. ML-estimation of the surface intensities. We are now prepared to estimate the unknown surface intensities $g_o(u,v)$ taking into account the imaging model (3). Observe, that not all I parameters (a_i, b_i) can be estimated simultaneously, which requires regularization e. g. with the constraints $\sum_i a_i = 0$ and $\sum_i b_i = I$. As a certain percentage of the image data $g_i(r,c)$ are not belonging to the visible part of the unknown surface patch we need to perform a robust estimation.

In order to investigate the potential of the method we, instead of simultaneously estimating the unknown parameters, split the estimation into two phases: first estimating (a_i, b_i) for given $g_o(u,v)$ and second estimating $g_o(u,v)$ for given (a_i, b_i).

For the first step we use $g_o^{(0)}(u,v) = \overline{g}_o(u,v)$ from (5) as approximation. This allows to estimate the parameters (a_i, b_i) separately yielding $(\widehat{a}_i, \widehat{b}_i)$. We apply a ML-type robust estimation with modified weights [Huber 1981]. Figure 4c shows a 2D-histogram for g_o and g_1 before the robust estimation. In the 4 images used for the example less than 20 % of the pixels have been eliminated showing the high degree of consistence of the data with the model.

For the second step we use the estimated parameters $(\widehat{a}_i, \widehat{b}_i)$ and therefore just have to estimate a mean of the transformed images $g_o(u,v)$ for each (u,v) individually:

$$\widehat{g}_o(u,v) = \frac{1}{|\mathcal{I}_s|} \sum_{i \in \mathcal{I}_s} (g_i(u,v) - \widehat{a}_i)/\widehat{b}_i \tag{9}$$

where $\mathcal{I}_s$ are all sets containing at least two images.

6. Classification of the surface elements. The residual vector $\boldsymbol{e} = (e_i)$ with

$$e_i(u,v) = (g_i(u,v) - \widehat{a}_i)/\widehat{b}_i - \widehat{g}_o(u,v) \quad \text{for } i \in \mathcal{I}_s(u,v) \tag{10}$$

for each (u,v) and each *subset* $\mathcal{I}_s(u,v)$ of the images is used to build a test statistics:

$$T_s(u,v) = \frac{\boldsymbol{e}_s^T(u,v)\boldsymbol{e}_s(u,v)}{\widehat{\sigma}_g^2 \cdot (|\mathcal{I}_s(u,v)| - 1)} \tag{11}$$

The noise variance $\widehat{\sigma}_g^2$ has been estimated robustly from all residuals, using the 5 % point of the empirical histogram of the squares of the residuals and compensating for the bias by dividing by $\mathrm{erf}^{-1}(0.05) = 0.0625$.

The number of these subsets is $I_s = |\mathcal{I}_s| = \sum_{i=2}^{I} \binom{k}{i} = 2^I - (I+1)$. In our case with $I = 4$ this yields 11 subsets namely $(1,2)(1,3)$ $(1,4)$ $(2,3)$ $(2,4)$ $(3,4)$ $(1,2,3)$ $(1,2,4)$ $(1,3,4)(2,3,4)(1,2,3,4)$ (cf. fig. 5a). From the $|\mathcal{I}_s(u,v)|$ subsets we choose the test statistics which is minimal

$$T_{oj}(u,v) = \min_s T_s(u,v) \tag{12}$$

in order to make a decision on which pixel is contained in R_{oj} based on the radiometric information alone, being a second approximation for R_{oj}

$$\mathsf{R}_{oj}^{(2)} = \{(u,v)|T_{oj}(u,v) < k\} \ . \tag{13}$$

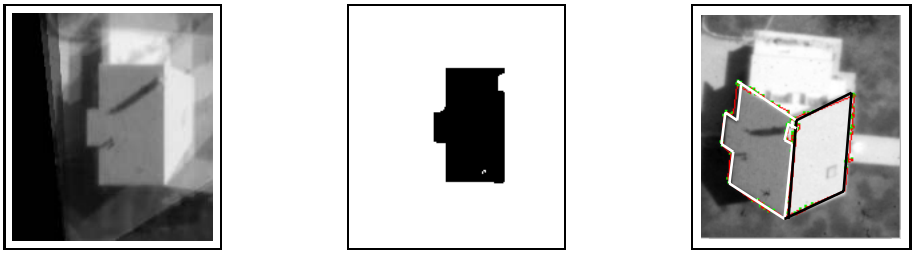

Fig. 6. a.,b.) first approximation of the mean grey values for one roof plane and the final segment; c.) final result for a total roof backprojected in one image containing the shadow of a chimney and the model-based vectorization of the boundary.

The critical value can be made dependent on the distribution of the test statistic T_s which is $F_{I_s,\infty}$-distributed. As the number of degrees of freedom is between 1 and 3 in our case and the critical values do not vary very much we can choose a conservative threshold for the critical value. A significance number of $\alpha = 0.001$ leads to ≈ 3. Taking into account possible deviations from the Gaussian / Fisher distribution a critical value of 6, which is 2 times larger seems to be plausible.

The final estimate for the surface region is the one following the radiometric model and being within the region of interest (7)

$$\widehat{\mathsf{R}_{oj}} = \mathsf{R}_{oj}^{(2)} \cap \overline{\mathsf{R}}_{oj} \tag{14}$$

(cf. Figure 5b). This result is to be compared with a simple minded procedure (cf. fig. 5c) just taking all pixels of the original regions (cf. fig. 1c) which are common to at least two regions R_{ij} after geometric back projection onto the predicted surface. Obviously the chimney has correctly been excluded. It can be shown that such a structure when seen from 4 sides in general direction correctly can be handled by the proposed procedure.

A second example of the procedure is shown in fig. 6. Again the chimney and the shadow have correctly been handled. The vectorization [Brunn *et al.* 1995] of the outline, shown right, has use the generic knowledge of roofs possibly having rectangles. Observe, that the 3D pose of the planar polygon now is completely known.

4 Outlook

The scope of the paper was to demonstrate the feasibility of both iconic and symbolic descriptions of several images for segmenting and reconstructing their 3D-shape. The rational quickly going to a symbolic processing for generating hypotheses and confirming and improving them by image information has shown to give plausible results on real images of man-made objects. The statistical framework allows to limit the number of control parameters easing adaption and testing of the procedure. The approach needs to be compared to different automatic algorithms for segmentation and classical photogrammetry. The transfer of the boundary description into volumetric models will provide the ability to add more semantic knowledge.

Acknowledgments: This work was supported by DFG (Fo 180-1/2) and BMBF (01 M 3018 B/5).

References

[Bignone *et al.* 1996] Bignone, F.; Henricsson, O.; Fua, P.; Stricker, M. (1996): Automatic Extraction of Generic House Roofs from High Resolution Aerial Imagery. In: Buxton B., Cipolla R. (Ed.), *Computer Vision - ECCV '96, Vol . I*, pages 85–96. Lecture Notes in Computer Science, 1064, Springer-Verlag, 1996.

[Braun *et al.* 1995] Braun, C.; Kolbe, T.H.; Lang, F.; Schickler, W.; Steinhage, V.; Cremers, A.B.; Förstner, W.; Plümer, L. (1995): Models for Photogrammetric Building Reconstruction. *Computer & Graphics*, 19(1), 1995.

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

[Förstner 1994] Förstner, W. (1994): A Framework for Low Level Feature Extraction. In: Eklundh, J.-O. (Ed.), *Computer Vision - ECCV '94, Vol. II*, pages 383–394. Lecture Notes in Computer Science, 801, Springer-Verlag, 1994.

[Förstner 1996] Förstner, W. (1996): 10 Pros and Cons Against Performance Characterization of Vision Algorithms. In: Christensen H. I., Förstner W., Madsen C. B. (Ed.), *Workshop "Performance Characterisics of Vision Algorithms"*, 1996.

[Fuchs and Förstner 1995] Fuchs, C.; Förstner, W. (1995): Polymorphic Grouping for Image Segmentation. In: *5th ICCV '95, Boston*, pages 175–182. IEEE Computer Society Press, 1995.

[Fuchs *et al.* 1994] Fuchs, C.; Lang, F.; Förstner, W. (1994): On the Noise and Scale Behaviour of Relational Descriptions. In: Ebner, Heipke, Eder (Ed.), *ISPRS Vol. 30, 3/2*, pages 257–267. SPIE, 1994.

[Grün *et al.* 1995] Grün, A.; Kübler, O.; Agouris, P. (Eds.) (1995): *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkhäuser, 1995.

[Huber 1981] Huber, P. J. (1981): *Robust Statistics*. Wiley NY, 1981.

[Lang and Förstner 1996] Lang, F.; Förstner, W. (1996): Surface Reconstruction of Man-Made Objects using Polymorphic Mid-Level Features and Generic Scene Knowledge. In: *Proceedings ISPRS Congress, Vienna*, 1996. to appear.

[Steinhage 1995] Steinhage, V. (1995): Gebäudeerkennung durch Aspekthierarchien. In: K., List F.; B., Wrobel (Eds.), *15. Jahrestagung der DGPF*, 1995.

[Storjohann *et al.* 1988] Storjohann, K.; Schulze, E.; v. Seelen, W. (1988): Segmentierung dreidimensionaler Szenen mittels perspektiver Kartierungen.