SURFACE RECONSTRUCTION OF MAN-MADE OBJECTS USING POLYMORPHIC MID-LEVEL FEATURES AND GENERIC SCENE KNOWLEDGE

Felicitas Lang, Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn
Nußallee 15, D-53115 Bonn
Felicitas.Lang, Wolfgang.Foerstner@ipb.uni-bonn.de

Commission III, Working Group IWG II/III

KEY WORDS: Object Reconstruction, Multi Image Matching, Relational Description, Semantic Modeling, Scene Knowledge

ABSTRACT

This paper presents a new concept for 3D-surface reconstruction, which infers domain specific local 3D-structures in space from its observed local 2D-structures in multiple images using polymorphic relational image descriptions. A 3D-aggregation can combine these local 3D-structures and thus results in a 3D-boundary representation of man-made objects being useful for different analyses and simulations.

1 INTRODUCTION

Threedimensional (3D) surface desriptions of man-made objects, especially buildings are needed for a large number of tasks related to measurement, planning and construction and thus started lively research activities in the field of 3D-reconstruction from aerial images [[McGlone and Shufelt, 1994], [Lin et al., 1994]] and range data (cf. [Weidner, 1996]).

Quite a number of methods for 3D-reconstruction by stereo analysis have been proposed, which rely on edges extracted from the images and which incorporate additional constraints during the correspondence analysis, e.g. due to image geometry, perceptual grouping ([Roux and McKeown, 1994], [Chung and Nevatia, 1992], [McKeown and McGlone, 1993]) or using information fusion techniques ([Haala, 1994]).

For deriving a complete surface description of complex and generic objects we cannot expect classical stereo algorithms, working on two images or solely relying on edges or regions, to be capable of handling the complexity of natural scenes, which is caused e.g. by occlusions, low contrast, disturbing background structure and image noise. Exploiting rich image descriptions therefore appears to be a promising approach (cf. [Henricsson, 1995], [Bignone et al., 1996]).

The basic idea of our approach is to derive a 3D-surface description of the object by multiple image correspondence analysis using polymorphic mid-level features which are derived from a relational image description while incorporating generic scene knowledge. This on the one hand allows to use relations between rich attributed features, on the other hand incorporates 3D scene knowledge leading to a stable reconstruction.

Section 2 presents the general framework for 3D-surface reconstruction. Section 3 describes the aggregation level we work on. Section 4 presents our strategy. To give an example section 5 desribes the reconstruction of building-specific local vertex-aggregates in detail. Futurer work analysing the form of prereconstructed surface planes are presented in section 6.

2 GENERAL FRAMEWORK

The basic idea of our approach is to derive a polymorphic symbolic description O of the object from I image descriptions, i. e. segmentations $\mathsf{S}_i, i = 1, ..., I$. As our experience analysing images with multiple overlap has shown a high stability with fully automatic subprocesses (cf. [Lang and Förstner, 1996], [Läbe and Ellenbeck, 1996]), we propose using multiple images simultaneously for exploiting a maximum of information being available.

Formally the problem of 3D-reconstruction is identical to estimating an unknown 3D-description O given the 2D-descriptions S_i based on the projection model:

$$\mathsf{S}_i + \mathsf{D}_i = f_i(O) \tag{1}$$

$f_i(O)$ 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 O leads to an optimization problem which can be solved e.g. by maximization of the probability

$$\hat{O} = arg \max_{O} P(O \mid \{\mathsf{S}_i\}). \tag{2}$$

In order to break down eqs. (1) and (2) we need to explicitely model the object and its appearence in the images. For the reconstruction it is decisive to choose an adequate level of aggregation in order to be able to express both the mapping function f_i and the disturbances D_i

3 BASIC MODELS

On the one hand the aggregation level should be high enough to avoid an exploding combinatoric complexity while solving the correspondence problem. On the other hand it should be low enough to be as insensitive as possible to any disturbances, that is they should be observable. The aggregation level motivating our reconstruction results from the experiences with a polymorphic feature extraction (cf. [Förstner, 1994], [Fuchs and Förstner, 1995]). Therefore we choose local aggregates of basic features as the interface between object and image.

3.1 Image Model

We start with a relational image description S which consists of a set of attributed points $\mathcal{P}$, lines $\mathcal{L}$ and regions $\mathcal{R}$ together with their mutual relations R contained in a feature adjacency graph (FAG). Based on this FAG point, line and region induced aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$, namely vertices $\mathcal{V}_i = \mathcal{A}^{\mathsf{P}}{}_i$, wings $\mathcal{W}_i = \mathcal{A}^{\mathsf{L}}{}_i$, cells $\mathcal{C}_i = \mathcal{A}^{\mathsf{R}}{}_i$ are derived .

These basic feature induced 2D-aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$ (cf. Fig. 1) have shown to be quite stable and to be observable with high probability in several images having a striking structural resemblance. They serve as starting point for our estimation of local 3D-structures $\mathcal{A}^{\mathsf{F}}{}_{oj}$.

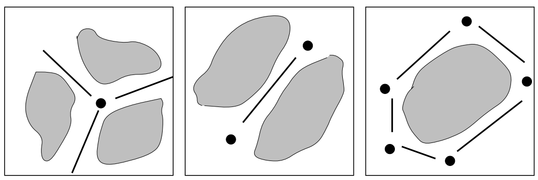

Figure 1: point-, line- and region-induced local aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$ derived form the feature extraction, especially vertices $\mathcal{V}_i$, wings $\mathcal{W}_i$ and cells $\mathcal{C}_i$, which are used for reconstruction.

3.2 Object Model

For including scene knowledge into the reconstruction process, the 3D object model must be placed at the same representation level like the image aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$ which we propose for reconstruction. Transferring these 2D aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$ into 3D results in a local boundary representation of *parts* $\mathtt{P}_{oj}$ of the object which form the hierarchically structured 3D scene model (cf. [Braun et al., 1995]).

In addition to a purely geometric/physical model of the scene, we now have to introduce thematic information in order to expoit the generic scene knowledge about the object. This semantic modelling in a first step consists of class-labels for the 3D-aggregates.

Restricting to the domain of building reconstruction, our object O is described by a polyhedron consisting of attributed parts $\{\mathtt{P}_{oj}\}$ of different classes being a set of corners $\{\mathtt{C}_{oj}\}$, edges $\{\mathtt{E}_{oj}\}$ and faces $\{\mathtt{F}_{oj}\}$ which is identical to a part-of representation. Relations between those parts are not used at this stage of the developement but may be added later.

Each object part $\mathtt{P}_{oj}$ can be described by its geometric and its thematic description $\mathtt{P}_{oj} = (G, T)_{oj}$. The geometric description G is represented by aggregates in object space $\mathsf{A}^{\mathsf{F}}{}_o$, while the thematic description T is represented by the object class labels $\omega_{oj} \in \Omega = \{\Omega^C, \Omega^E, \Omega^F\}$ for the different part classes Ω^C, Ω^E and Ω^F. We distinguish corners, edges and faces as different object parts:

$$\begin{array}{ll} \text{corners:} & \mathtt{C}_{oj} = (\mathsf{V}_{oj}, \omega^C_{oj}) \\ \text{edges:} & \mathtt{E}_{oj} = (\mathsf{W}_{oj}, \omega^E_{oj}) \\ \text{faces:} & \mathtt{F}_{oj} = (\mathsf{C}_{oj}, \omega^F_{oj}) \end{array}$$

The corresponding class labels of these object parts within the domain *building* are given by:

$$\begin{array}{l} \omega^C_{oj} \in \Omega^C = \{\mathtt{gabel\_point}, \mathtt{eave\_point}...\} \\ \omega^E_{oj} \in \Omega^E = \{\mathtt{gable\_edge}, \mathtt{eave\_edge}, ...\} \\ \omega^F_{oj} \in \Omega^F = \{\mathtt{flat\_roof}, \mathtt{rectangle\_roof}, \mathtt{triangle\_roof}, ...\} \end{array}$$

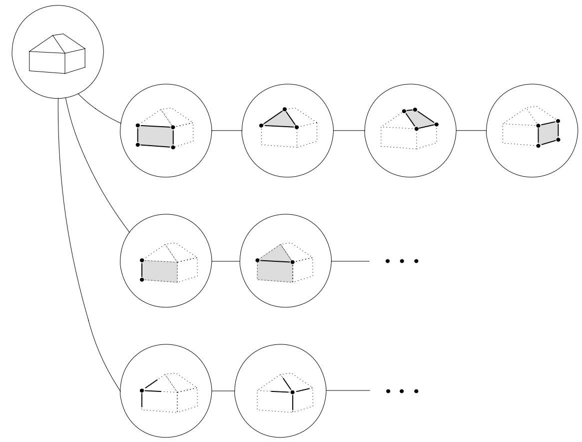

Figure 2: building decomposition into faces $\{\mathtt{F}_{oj}\}$, edges $\{\mathtt{E}_{oj}\}$ and corners $\{\mathtt{C}_{oj}\}$, which are used for reconstruction.

For the time being this reconstruction by parts certainly is motivated by the observability within the segmentations S_i in the images i. However, modeling the structure of complex buildings in a part-of hierarchy is generic enough for most applications (cf. [Steinhage, 1995]).

4 STRATEGY

Our strategy follows the hypothesize-and-verify paradigm. By using building specific aggregates a high integration of 2D and 3D reasoning is achieved. We therefore propose an early transition to object space by matching local basic aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$. Those reconstructed 3D-aggregates $\hat{\mathcal{A}}^{\mathsf{F}}{}_{oj}$ can be semantically attributed which leads to hypotheses for a domain specific interpretation $\mathtt{P}_{oj}$ of parts of the object. This interpretation provides sharp restrictions during the verification process and for a subsequent grouping of object parts $\mathtt{P}_{oj}$ in space which is necessary to get a complete surface description O.

On the one hand this 3D information derived is more expressive than solely 2D image information, on the other hand in object space a direct interaction with a 3D model for integrating scene knowledge is possible.

The reconstruction process consists of the following steps:

1. **Hypotheses generation:**
   (a) *hypotheses for corresponding aggregates:* By identification of building specific aggregates $\mathcal{A}^{\mathsf{F}}{}_{ij}$, a hypotheses set of corresponding aggregates is established taking into account scene independent constraints like epipolar geometry. For guiding the search at the very beginning the identification of structures depends on a uniqueness criterium. It is expressed by the probability for finding corresponding aggregates in the other images. Later, the selection of aggregates further depends on a heuristics based on connections to previously reconstructed and verified object parts $\mathtt{P}_{oj}$ as connections in space are likely to show connections between image aggregates and features itself. This is similar to the idea of an invocation tree presented in [Draper et al., 1989].

(b) *hypotheses for interpretation:* Each hypotheses set results in a set of preliminary reconstructions $\hat{\mathcal{A}}^{\mathsf{F}}{}_{oj}$, which are semantically attributed and labeled by possibly several alternative 3D interpretations ω_{oj}. Thus the relation between reconstructed aggregates $\hat{\mathcal{A}}^{\mathsf{F}}{}_{oj}$ and special parts P_{oj} of the object model is established.

2. **Hypotheses verification:**
Using each interpreted 3D reconstruction as hypothesis, the estimation procedure is initialized as a multiple image matching procedure. The best mapping is defined by an optimization function, forcing the result to be locally consistent. This takes into account:

   - *Specific scene knowledge* which contains constraints on orientation of edges and planes in space, likely relations between the basic features, and possibly symmetry properties of the features or their relations as they are defined for class label ω_{oj}.
   - *Qualitiy measures* for the extracted features including their relations. They result from previous investigations into the noise behaviour of our feature extraction FEX, presented in [Fuchs et al., 1994].

   Thus the differences between the projected 3D-structure O and the observed 2D-structures S_i together with the errors of the feature extraction comprise the total structural error D_i of the observed evidences which are to be optimized.

We start the reconstruction based on a set of point induced local aggregates, called vertices $\mathcal{V}_i = \mathcal{A}^{\mathsf{P}}{}_i$ derived by directely analyzing the feature adjacency graph FAG (cf. fig. 3). The reason to choose point induced aggregates is that for these local aggregates the imaging geometry leads to sharp constraints within solving the correspondence problem. Moreover, the information gain in 3D is quite large as an ideal vertex $\tilde{\mathsf{V}}_o$ in object space is represented by a node point, several edges branching off and each pair of edges defining an object plane. This definition of 3D object planes can be used for a specific analysis e.g. a form analysis of the object parts lying in this plane for finding missing vertices of the object. An example for segment form analysis using iconic image information within a previously reconstructed object plane is given in section 6.

5 CORNER RECONSTRUCTION

The corner reconstruction is based on point induced vertex aggregates V_{ij}. Each vertex is represented by its node point P and its neighbouring features $\mathcal{F}$ including their geometric attributes and the relations R_{ij} between those features. It can be derived directely from the FAG.

The different processing steps within generation and verification of hypotheses are desribed in the following.

5.1 Hypothesis Generation

The generation of hypotheses is subdivided into 2 parts:

a. Finding vertex hypotheses: The first step consists in the data driven generation of a set of corresponding vertices $\mathcal{V}_i$ leading to a set of hypotheses for 3D-vertices $\mathcal{V}_o$. Each

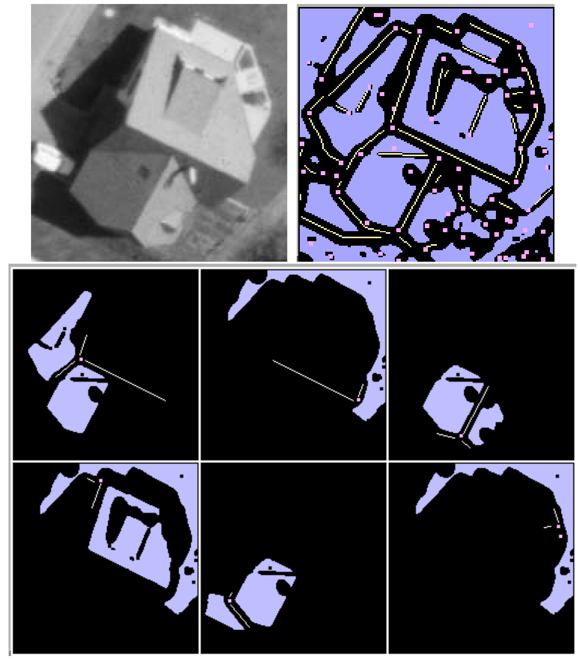

Figure 3: shows the original image (upper left) and all extracted features using FEX (upper right). The composition shows 6 vertices , as they can be derived from the FAG. They are sorted by their evaluation score from left to right/top to bottom, the first row showing the best scores, the second row showing vertices with medium score.

vertex structure V_{ij} is classified by a MAP classification which evaluates the conditional probability $P(\omega^C \mid \mathsf{V}_{ij})$ =P$(\omega^C \mid d, c)$ for each V_{ij}. This takes into account different discrete d and continous c criteria of that vertex V_{ij}. The classification criteria serves for selecting strong vertices, i. e. vertices which can be found with high probability in the other images. It evaluates:

- the number of lines L.
- the length of lines L.
- the number of regions R.
- the number of lines L which are parallel to any previously reconstructed line.

This classification and probabilistic evaluation leads to a priority list of vertices. The search for corresponding vertices starts with selecting a vertice V_{ij} in one of the images following this list. Restrictions for corresponding vertices are given due to epipolar geometry. The priority list is also used for steering the search for correspondences out of the set of vertices which fulfil the constraints of imaging geometry.

Fig. 3 shows 6 evaluated vertices of the sorted priority list in one of the images. Fig. 4 shows all vertices which are possible correspondences for one of the selected vertices, based on the imaging geometry.

Additional heuristics for the selection of the next vertex structure to be analysed are used after the previous vertex has been successfully reconstructed. E. g. following edges or segments of previously reconstructed vertices can restrict the set of vertices $\mathcal{V}_{ij}$ which shall be analysed in the next step.

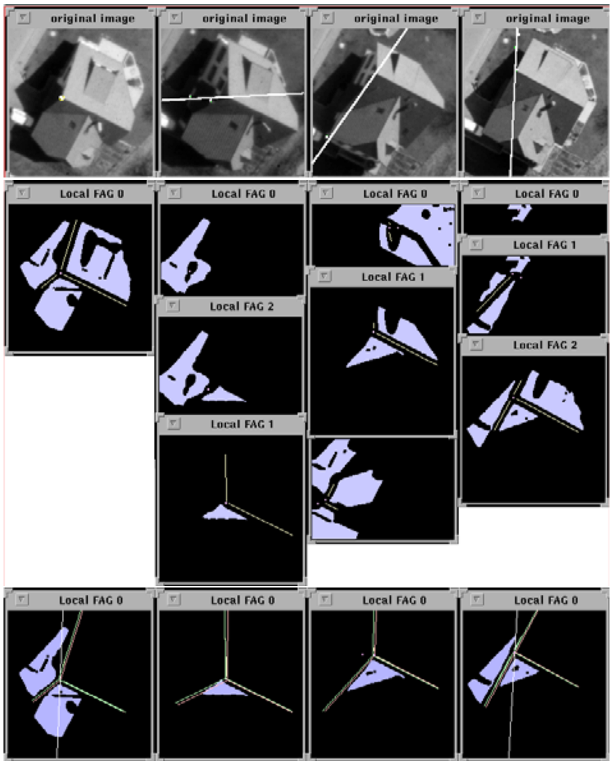

Figure 4: shows one selected vertice in the left image and the corresponding epipolar lines in the other images of the top row. The further images show all possible corresponding vertices, the search space being constrained solely by the epipolar geometry. The last row shows one set of corresponding vertices $\{V_{ij}\}$ used for reconstruction.

Thus we focus the reconstruction on vertices that are connected to already reconstructed object part as a connection to object parts in space probably leads to a connection of image structures.

b. Finding corner hypotheses: The second task within hypothesis generation is the model driven interpretation of the set of corresponding vertices $\{V_{ij}\}$ by integrating the object model of building specific corners $\mathtt{C}_{oj}$. Therefore a preliminary reconstruction is executed which solves the one-to-one mapping of the features $\mathcal{F}$ by relational matching of the local vertex aggregates $\{V_i\}$. This results in a preliminar estimation $\hat{V}_{oj}$ of a 3D vertex . Each hypothesis $\hat{V}_{oj}$ is semantically interpreted and classified to be a corner $\mathtt{C}_{oj} = (V_{oj}, \omega^C_{oj})$ of class ω^C_{oj}. For classification we analyse the line orientation in space discriminating 5 types of orientation, namely { `horizontal (h)`, `oblique+ (o+)`, `oblique- (o-)`, `vertical+ (v+)`, `vertical- (v-)` }. The sign symbolizes a positive/negative slope related to the position of the corner point (cf. [Gülch, 1992]). In addition, symmetries to the ridge and planes being vertical are analysed to obtain a more detailed division into subclasses. The corner classes we have modelled up to now are covering 4 building types, namely { `flat_roof` , `non_orthogonal_flat_roof`, `gable_roof` and `hip_roof` }. The corresponding corners are shown in Fig. 5. As far as an unambigous classification can not to be reached, all possible classifications have to be analyzed during the verification step.

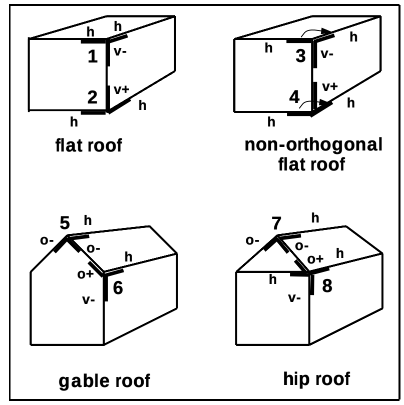

Figure 5: shows the 8 corner classes ω^C_{oj} we are using up to now, which are sufficient for the part-of description of 4 building types, namely { `flat_roof`, `non_orthogonal_flat_roof`, `gable_roof` and `hip_roof` } .

Figure 6: shows the result of parameter estimation for corner class no. 5 (left) and no. 6 (right) for a gabled roof house. As the right example shows, the model-based parameter estimation is capable to compensate partial occlusion of a corner, e. g. missing eaves lines or missing vertikal lines in some of the images.

Observe that this classification also can be used in more complex situations because only local structures are analysed.

5.2 Hypothesis Verification

After hypotheses generation we start the model driven verification of the hypotheses $\mathcal{V}_o$ by integrating 3D-scene knowledge of a building specific corner model $\mathtt{C}_o$ of class ω^C_o. The interpretation is evaluated by maximization of the probability for a given hypothesis set j, omitting the index j for clarity.

$$P(\mathtt{C}_o \mid \{V_i\}) = \frac{P(\{V_i\} \mid \mathtt{C}_o) P(\mathtt{C}_o)}{P(\{V_i\})} \quad (3)$$

which essentially needs the likelihood function, breaking down the conditional probabilities.

$$\begin{aligned} P(\{V_i\}|\mathtt{C}_o) &= P(\mathcal{F}(g),\ R|\mathtt{C}_o) && (4) \\ &= P(\ R|\mathcal{F}(g), \mathtt{C}_o) \cdot P(\mathcal{F}(g)|\mathtt{C}_o) && (5) \end{aligned}$$

The probability $P(\mathtt{C}_o)$ for having corner of class ω^C_o can be obtained empirically. In case of complete modeling the denominator in (3) can be obtained by normalization. Otherwise at least probability ratios between the different hypotheses can be obtained.

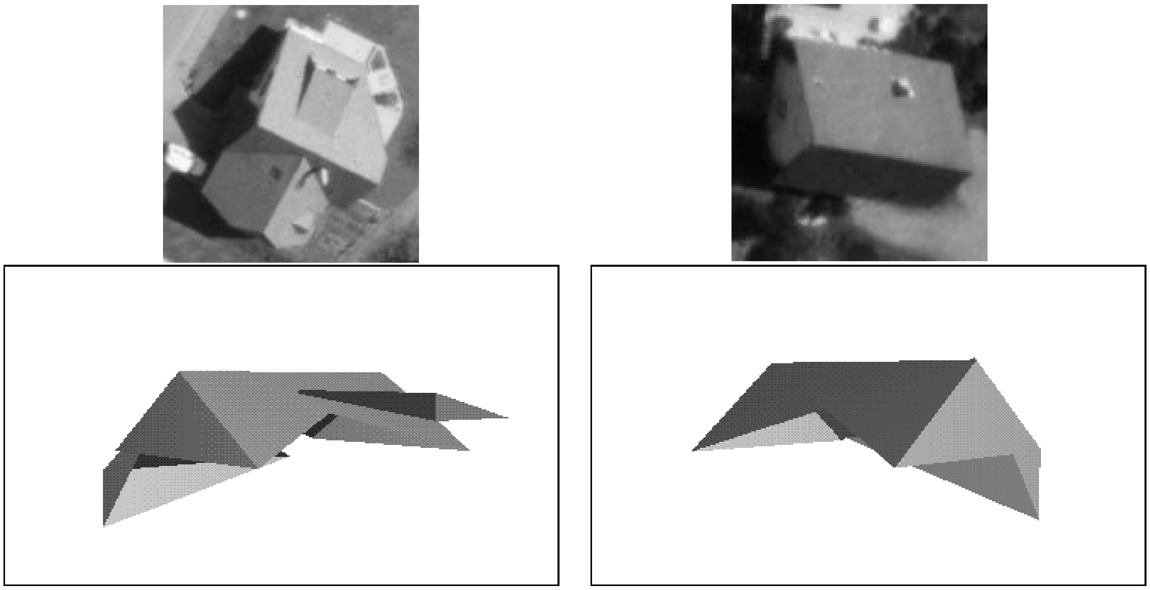

Figure 7: Selected vertices are automatically reconstructed in 3D. The first row shows the original images. The second row visualizes the partially reconstructed roof in 3D composed by different corners. On the left side 4 corners are used, on the right the building is composed by 3 corners.

With restriction to binary relations $\{\mathsf{R}_{mm'}\}$ between feature F_m and $\mathsf{F}_{m'}$ and assuming the conditional independence of the features this leads to

$$P(\{\mathsf{V}_i\}|\mathsf{C}_o) = \prod_m P(\mathsf{F}_m(g)|\mathsf{C}_o) \cdot \prod_{m,m'} P(\mathsf{R}_{mm'}|\mathsf{F}_m, \mathsf{F}_{m'}, \mathsf{C}_o) \qquad (6)$$

thus allowing the integration of quality measures for the extracted features including their relations. For evaluation of the geometry G, e. g. of the geometric parameters g, the classical modeling techniques of observation errors can be used. Using the observations $\boldsymbol{y} = f(g)$ being a function of the geometric parameters g, the evaluation can be derived from the residuals $\boldsymbol{y} - \hat{\boldsymbol{y}}$ of the optimal estimation $\hat{\boldsymbol{y}}$ for $\boldsymbol{y}$ using the *probability density function* $p(g)$ in case the feature exists and has been successfully matched to the model.

$$p(g(\mathsf{F})|\exists(\mathsf{F}) = \text{matched}, \mathsf{C}_o) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma}_{yy})^{1/2}} \; e^{(-\frac{1}{2}(\boldsymbol{y}-\hat{\boldsymbol{y}})^T \boldsymbol{\Sigma}_{yy}^{-1}(\boldsymbol{y}-\hat{\boldsymbol{y}}))} \qquad (7)$$

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

Figure 6 shows the result of the parameter estimation for two different corner classes using 4 images simultaneously. The achieved accuracy of the reconstructed vertice point is about $\sigma_x = \sigma_y = \pm 6cm$ and $\sigma_z = \pm 15cm$. The accuracy of the orientation of the corners is about $\sigma_\kappa = 0.6\ [deg]$ while the reconstruction of the slope of the roof is by $\sigma_{sl} = 2\ [deg]$. The image scale we used is $1 : 5000$ with $30\mu m$ pixel size. The results of the partial reconstruction of two different houses contained within the data set for the Ascona Workshop 1995 on *Automatic Extraction of Man-Made Objects from Aerial and Space Images* (cf. [Grün et al., 1995]) is shown in figure 7. These results form the basis for a second independent step where global knowledge needs to be included to inforce global unambiguity and completeness.

6 FUTURE STEPS

The definition of 3D object planes which results from the corner reconstruction can be used for a form analysis of faces lying within this plane to get a better region segmentation by multi image analysis. Thus cells can be used for bridging missing vertices of the object.

Fig. 8 shows first results of a refinement of the initial segmentation which we achieved by fusion of the 4 images rectified to one reconstructed plane. By simply averaging the greyvalues and repeating feature extraction we derived the polygon shape that can be used as approximation for precise form estimation either working on the iconic or on the symbolic level. Thus we can fully exploit the image content of multiple images simultaneously. A procedure for segmenting surfaces which is based on the physical model of the imaging process working on the iconic level is presented in [Brunn et al., 1996].

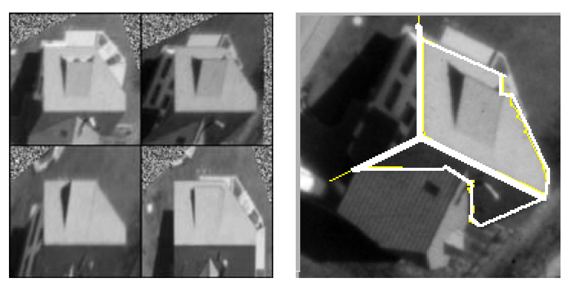

Figure 8: shows first results of a refinement of the initial segmentation which can be achieved by fusion of multiple rectified images after reconstruction of object planes in space, as presented in section 5. The rectified images which are shown on the left are used for analysing the upper roof plane in the right image.

7 CONCLUSIONS

In this paper we presented our concept for 3D-surface reconstruction of man-made objects, especially for building extraction using feature aggregates. Our approach shows some unique features:

1. *Polymorphic mid-level features for matching.* In contrast to previous work the reconstruction takes place at a higher abstraction level using a polymorphic feature description having the advantage of a large number of attributes and relations between features which stabilizes the reconstruction process.
2. *Multi-image matching.* Our approach intensionally allows to use more than two images, as in this application occlusions would prevent complete reconstruction from only two images.
3. *Integration of scene knowledge.* Generic scene knowledge is used at various stages in our setup. Scene specific structures are selected and evaluated depending on their consistency with local scene properties, such as verticality or horizontality, and are locally merged based on expected neighbourhood relations in 3D.
4. *Quality evaluation of data, models and results.* The control of the search for a solution and the evaluation of the intermediate hypotheses as well as the final result is based on statistical measures, derived from the original data and from training data.

The reconstructed 3D aggregates can be connected by a grouping in space which finally leads to a complete 3D surface desription.

ACKNOWLEDGEMENTS

This work largely was done within the project "Semantic Modeling and Extraction of Spatial Objects from Images and Maps" in the subproject "Building Extraction" which is supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

[Bignone et al., 1996] Bignone, F., Henricsson, O., Fua, P., and Stricker, M. (1996). Automatic Extraction of Generic House Roofs from High Resolution Aerial Imagery. In *Computer Vision '96.*

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

[Brunn et al., 1996] Brunn, A., Lang, F., and Förstner, W. (1996). A Procedure for Segmenting Surfaces by Symbolic and Iconic Image Fusion. In *Proceedings DAGM '96, Heidelberg.* submitted to.

[Chung and Nevatia, 1992] Chung, R. and Nevatia, R. (1992). Recovering Building Structures from Stereo. In *IEEE: Workshop on Applications of Computer Science.* Palm Springs, CA, December.

[Draper et al., 1989] Draper, B. A., Collins, R. T., Brolio, J., Hanson, A. R., and Riseman, E. M. (1989). The Schema System. *Int. Journal of Computer Vision*, 2:209–250.

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

[Fuchs and Förstner, 1995] Fuchs, C. and Förstner, W. (1995). Polymorphic Grouping for Image Segmentation. *5th ICCV '95, Boston.*

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

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

[Gülch, 1992] Gülch, E. (1992). A Knowledge-based Approach to Reconstruct Buildings in Digital Aerial Imagery. In *Proceedings of 17th ISPRS Congress, Washington D. C., BII*, pages 410–417.

[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.* SPIE.

[Henricsson, 1995] Henricsson, O. (1995). Inferring Homogeneous Regions from Rich Image Attributes. In A. Gruen, O. Kuebler, P. A., editor, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 13–22. Birkhäuser.

[Läbe and Ellenbeck, 1996] Läbe, T. and Ellenbeck, K.-H. (1996). 3D-Wireframe Models as Ground Control Points for the Automatic Exterior Orientation. In *Proceedings ISPRS Congress, Vienna.* to appear.

[Lang and Förstner, 1996] Lang, F. and Förstner, W. (1996). 3D-City Modelling with a Digital One-Eye-Stereo System. In *ISPRS Symposium, Vienna.*

[Lin et al., 1994] Lin, C., Huertas, A., and Nevatia, R. (1994). Detection of Buildings Using Perceptual Grouping and ShadowsProceedings Computer Vision and Pattern Recognition. In *Proceedings Computer Vision and Pattern Recognition*, pages 62–69.

[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.

[McKeown and McGlone, 1993] McKeown, D. M. and McGlone, J. C. (1993). Integration of photogrammetric cues into cartographic feature extraction. In Barrett, E. B. and McKeown, D. M., editors, *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision.* SPIE-The International Society for Optical Engineering. Orlando, 14.-15. April.

[Roux and McKeown, 1994] Roux, M. and McKeown, D. M. (1994). Feature Matching for Building Extraction from Multiple Views. In *IEEE 94: Computer Society Conference on Computer Vision and Pattern Recognition*, pages 46–53. IEEE Computer Society Press. 21.-23. June, Seattle, Washington.

[Steinhage, 1995] Steinhage, V. (1995). Gebäudeerkennung durch Aspekthierarchien. In K., L. F. and B., W., editors, *15. Jahrestagung der DGPF.*

[Weidner, 1996] Weidner, U. (1996). An Approach to Building Extraction from Digital Surface Models. In *ISPRS Symposium, Vienna.*