Mid-Level Vision Processes for Automatic Building Extraction

Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn

Ascona Workshop on 'Automatic Extraction of Man-Made Objects from Aerial and Space Images', 1995

Abstract

Mid-level processes in vision are understood to produce structured descriptions of images without relying on very specific semantic scene knowledge. Automatic building extraction can use geometric models to a large extent. Geometric hypotheses may be inferred from the given data in 2D or 3D and represent elementary constraints as incidence or collinearity or more specific relations as symmetries. The inferred hypothesis may lead to difficulties during spatial inference due to noise and to inconsistent and mutually dependent constraints. The paper discusses the selection of mutually not-contradicting constraints via robust estimation and the selection of a set of independent constraints as a prerequisite for an optimal estimation of the objects shape. Examples from the analysis of image and range data are given.

1 Introduction

Mid-Level Processes in Vision. Automatic image interpretation is a complex process for which no commonly accepted model exists. A three layered scheme reflecting the different levels of abstraction could follow the classical distinction of low-, mid- and high-level processes. Mid-level processes then may be characterized by two criteria: 1. The output of a mid-level process is a structural description of the image or the scene, defining the separation from image processing. 2. No or only poor semantic scene knowledge is used, in a soft way defining the separation from knowledge based vision. Typical mid-level processes are feature extraction, segmentation, grouping, or spatial reasoning.

Geometric Models for Buildings. Though many of the semantic entities of buildings, like floor, garage, passage, wall etc. are very specific for buildings, others, like roof or window are relevant also in other contexts, e. g.

when modelling cars. However, their identification heavily relies of the context, i. e. on the relations between the entities and the relations to other objects.

On the other hand, geometric entities, like boxes, prisms lying vertically or horizontally, may be identified in single images, from multiple images or from range data derived by stereo or laser range finders without very specific object modelling, but being the basis for a more semantic analysis.

Extracting man-made structures from images therefore makes heavy use of mid level processes as they show various common regularities, especially geometric ones, which are common to many types of objects. The regularity of their form puts a stress on geometric processing. The geometric richness of buildings refers to very simple constraints like lines to be vertical of horizontal, to be collinear or parallel or to more specific constraints like repetitions or symmetries.

When analysing perspective images these 3D–constraints may be transferred to the image domain. In case the exterior orientation of the camera is known and/or aerial images are taken this transfer is easy as the nadir point is known and/or the weak perspective can be well approximated by a linear projection therefore allowing a simple grouping of parallel lines and detection of affine symmetries.

Problems during Constraint Processing. There is a dichotomy in representing the geometry of an image or a scene 1. using coordinates as attributes, which eases spatial access, or 2. using datum parameters together with a form description, either directly with observable parmeters or indirectly with constraints or invariants. Whereas the representation with coordinates is unique it does not show the inherent structure of the object, e. g. two planes of a polyhedron being parallel. On the other hand, form parameters or invariants allow a direct link between observables and object, but are not unique, as new invariants may easily be derived.

Data driven processes therefore usually rely on observables or invariants, e. g. in matching or grouping. On the other hand the non-uniqueness of local features for recognition, matching or grouping easily leads to a plethora of hypotheses. This leads to the following problems. Hypotheses depend on noisy data, therefore

- the derived constraints may be mutually inconsistent
- the derived constraints may be actually wrong. Moreover,
- the derived constraints may be mutually dependent.

Depending on the number of unknown parameters and the number of used hypotheses, and depending on whether one can rely on the correctness of the hypotheses or not, different modelling may lead to estimation procedures of quite different algorithmic complexity. Therefore we have to solve two tasks which are discussed in the following section:

1. Selection of non-contradicting and/or independent hypotheses, and
2. selection of the appropriate model for estimating the objects shape.

Examples from the analysis of images and range data show the versatility of the approach.

2 Constraints for Geometric Reconstruction

In the following we assume that the geometry of the object can be described by the parameters of a set of features in image or object space together with a set of constraints between these parameters. We may assume these parameters in the most simple case to be the coordinates of points being directly or indirectly observed by some image analysis process, e. g. by direct point detection or by line intersection. These coordinates are at the same treated as observations $\boldsymbol{y}$ with a covariance matrix $\boldsymbol{\Sigma}_{yy}$ describing their uncertainty and as unknown parameters $\boldsymbol{\beta}$, which are to be determined using the geometric contraints

$$\boldsymbol{H}^T\boldsymbol{\beta} = \boldsymbol{c} \tag{1}$$

inferred by some grouping process.

We first present the different ways of modelling including the worst case complexity of the adjoint estimation process. The selection of non-contradicting hypotheses may be performed using a robust estimation with weak constraints. The selection of an independent subset of hypothesis can make use of a greedy algorithm derived from matroid theory.

2.1 Estimation with Constraints

Let the linear or the linearized model be described by the first and second moments of the n observed quantities $\boldsymbol{y}$: namely $E(\boldsymbol{y})$ and $D(\boldsymbol{y}) = \boldsymbol{\Sigma}_{yy}$. Then we may distinguish the following three models, which differ in the handling of constraints. All proof can be taken from Koch 1988.

Gauss-Markov-Model. The observation process is made explicit and a *minimum* number u of parameters $\boldsymbol{\beta}$ is to be estimated, implicitely taking the geometric constraints into account:

$$E(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} \tag{2}$$

The worst case complexity of the estimation process is $O(c_1 u^2 n + c_2 u^3)$, the second term usually dominating.

This model may also be used for the case where *all* unknown parameters are treated as unknowns and the hypotheses are assumed to be weak, leading to two groups of observations $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2)$ with $\boldsymbol{y}_1$ being the observed coordinates and $\boldsymbol{y}_2$ the observed contradictions $\boldsymbol{c}$. The covariance matrix

$$D(\boldsymbol{c}) = \boldsymbol{\Sigma}_{cc} \tag{3}$$

of the $\boldsymbol{y}_2$ reflects the uncertainty of the assumed constraints (cf. 2.2).

Gauss-Markov-Model with constraints. The observation process as well as the constraints are made explicit:

$$E(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta}, \quad \boldsymbol{H}^T\boldsymbol{\beta} = \boldsymbol{c} \tag{4}$$

where $\boldsymbol{\beta}$ carries *all* unknown parameters. The $u \times h$ matrix $\boldsymbol{X}$ must have full rank. The complexity of the estimation is $O(c_1 u^2 n + c_2 (u+h)^3)$.

Special Gauss-Helmert-Model. By observing the Jacobian in (4) to be $\boldsymbol{X} = \boldsymbol{I}$ we also can write the special Gauss-Helmert model

$$\boldsymbol{H}^T E(\boldsymbol{y}) = \boldsymbol{c} \tag{5}$$

where only a fit of the observations to the constraints is expressed. Again $\boldsymbol{H}$ must have full rank. The algorithmic complexity of the estimation process is $O(c_1 n h^2 + c_2 h^3)$, again the last term usually dominating.

An important equivalence. The two models (2) and (5) are equivalent if

$$\boldsymbol{H}^T \boldsymbol{X} = \mathbf{o} \tag{6}$$

and both matrices have full rank.

2.2 Selecting Non-contradicting Hypotheses

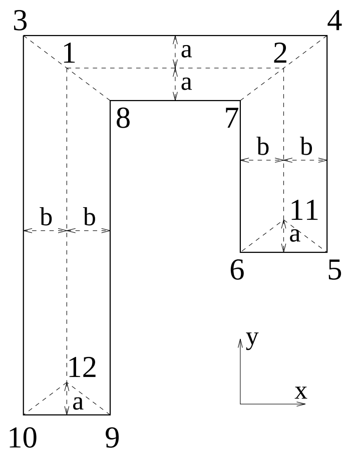
When automatically finding geometric hypotheses we cannot guarantee them to be non-contradicting. The way of finding contradicting hypotheses depends on whether they are crisp, as in (4) and (5), which cannot occur if the Jacobians $\boldsymbol{X}$ and $\boldsymbol{H}$ have full rank, or on whether they are weak, as in (3). In the last case no crisp criterion exists for detecting contradicting constraints. But a robust estimation of the unknown parameters may be performed, where the weights $p_{c_i} = 1/\sigma^2_{c_i}$ of the (statistically independent) constraints are reduced depending on the differences $\boldsymbol{h}_i^T \boldsymbol{\beta} - c_i$. Low weights indicate globally inconsistent constraints, which may be eliminated before further processing. Of course constraints with still high weights need not be true in reality.

2.3 Finding Linearily Independent Constraints

Let us assume a grouping process has found a large number h^* of geometric hypotheses which are linear. E. g. for the 6 different x-coordinates x_1, x_2, x_3, x_4, x_7, and x_8, of the 12 points in Fig. 1 the following linear relations have been found by comparing differences and second differences of x-coordinates:

$$\boldsymbol{H}^* = \left(\frac{\boldsymbol{H}_1}{\boldsymbol{H}_2}\right) = \left(\begin{array}{cccccc} 2 & & -1 & & & -1 \\ 1 & -1 & -1 & & 1 & \\ -1 & 1 & & -1 & & 1 \\ \hline -1 & -1 & & & 1 & 1 \\ & 2 & & -1 & -1 & \\ & & -1 & -1 & 1 & 1 \\ -1 & -1 & & & 1 & 1 \\ 1 & -1 & -1 & & -1 & \\ & & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & & -1 & \\ -1 & 1 & & -1 & & 1 \\ 1 & 1 & -1 & -1 & & \end{array}\right) \tag{7}$$

Figure 1: groundplan of a building showing regularities



A similar matrix is found for the y-coordinates with also 3 independent constraints. But, also the not intended identity of the coordinate differences: $x_7 - x_8 = y_6 - y_{12}$ had been found.

The first row e. g. represents the symmetry of points 3 and 8 points with respect to point 1: $x_8 - x_1 = x_1 - x_3$, whereas the second row represents the two left halfs of the two building wings to be of the same width: $x_2 - x_7 = x_1 - x_3$. Obviously several constraints have been found twice. But only the first three constraints form an independent set; the others are linearily dependent on these three. The rows of such an automatically derived integer valued matrix can be interpreted as elements of a vector matroid (EDMONDS 1971), where a set of elements (vectors) is labelled independent if the matrix composed of these elements has full rank. Then a simple greedy algorithm can be defined to select an (arbitrary) set of independent vecors. This greedy algorithm sequentially selects one row vector after the other and puts it into the minimum set if it is independent on those vectors which are already in that set. Checking the rank of a matrix cannot be performed on reals, which is reasonable as the decision whether a matrix is singular is a binary decision. Therefore a Gaussian elimination with integer arithmetic may be used (cf. GLOCK 1990). The submatrix $\boldsymbol{H}_1$ therefore may be used in a Gauss-Markov model with constraints according to (4).

2.4 Minimal Parametrization

In case the reduced number h of constraints found this way still is large it may be desirable to use the normal Gauss-Markov-Model (2). Then the number of unknowns has to be reduced by h which may be performed directly which involves the inversion of $\boldsymbol{H}^T\boldsymbol{H}$. Instead, the orthogonality relation may be used for selecting a subset of the coordinates, again using a greedy algorithm, i. e. sequentially select parameters which can not be solved by the previously selected ones and using the given constraints (cf. GLOCK 1990). This way the Jacobian $\boldsymbol{X}$ can be set up fully automatically from the reduced matrix of constraints.

3 Examples

3.1 Finding Non-conflicting Constraints in Segmentation

The following example is using the Ascona Workshop test images. It shows the result of the feature extraction and hypotheses selection (cf. FUCHS AND FÖRSTNER 1995) on a subpart of the image 5889. The polymorphic feature extraction (cf. FÖRSTNER 1994) extracts points, straight line segments and homogeneous regions. The analysis of the exoskeleton is used to build up a feature adjacency graph from which direct and indirect neighbours are derived. Using the DEM-analysis all blobs overlapping with building areas (after a projection into the image, cf. Fig. 2a) are used to trigger the selection of the points and line segments, shown in Fig. 2b, to be analysed for geometric relations, such as incidence, collinearity or parallelity. Only neighbouring segments are analysed and the distance between them was restricted to appr. 10 pixels, in order not to find too many wrong hypotheses. All these geometric hypotheses are checked for global consistency in a robust estimation according to the first paragraph in section 2.1. Obviously quite some links of neighbouring segments have been found and correctly closed. Also some neighbouring segments are straightened. But some short segments have been erroneously prolongated, indicating that the global check cannot guarantee the accepted hypotheses actually to be true.

3.2 Recovering the Groundplan of Buildings from Range Data

The following example shows the use of the constraint estimation onto the recovery of the groundplan of buildings from airborne laser scanner data (cf. BRUNN *et al.* 1995). A simple segmentation based on the heights of the buildings with respect to their surrounding leads to a binary image. Vectorization of the boundary of three of the the connected components (cf. Fig. 4) establishing *global* geometric hypotheses and again perfoming a robust estimation leads to the optimal estimates for the ground plans of the buildings in Fig. 5, which nicely show all types of regularities included in the hypothesis generation process, namely collinearities, parallelities and also orthogonalities. The overlay with the original data reveals small differences which need to be analysed.

3.3 Using 3D-Constraints for Shape Recovery

The shape recovery also can use 3D-hypotheses. Fig. 7 shows a gray scale image of a hand drawn polyhedron and the automatically extracted features

Figure 2: a. Section of Avanches test image 5889, b. DHM-selected features and c. result of robust estimation including all non-contradicting features.

Figure 3: Result of feature extraction

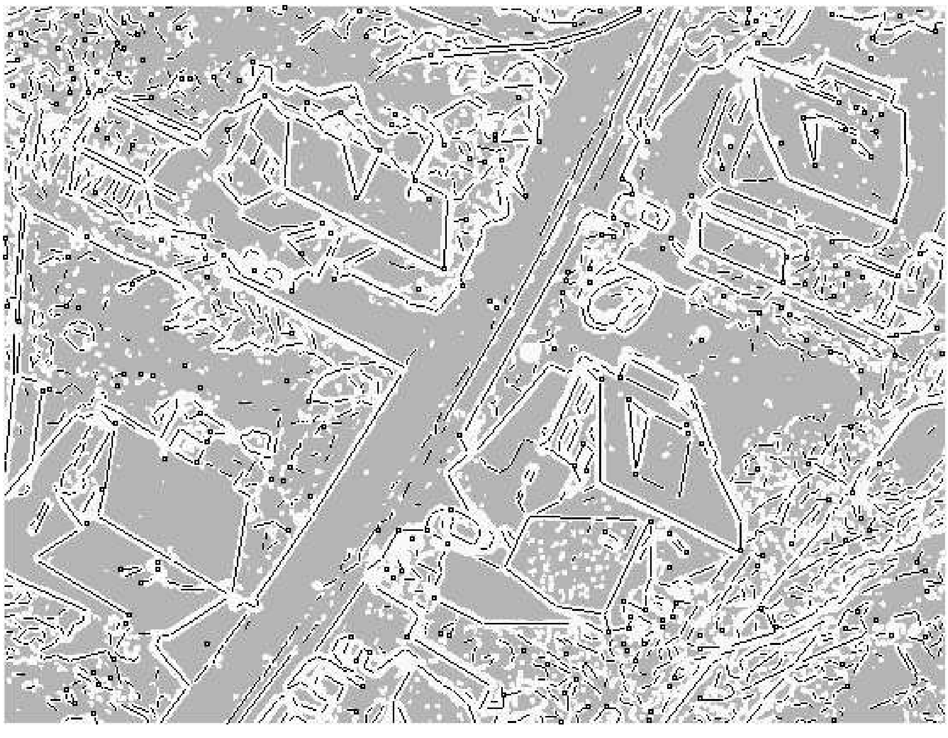


Figure 4: Range data, original boundary, DEM segments

Figure 5: Examples for range data, recovered final shape

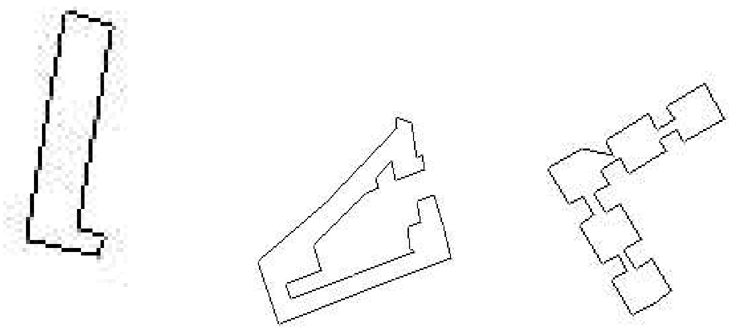


Figure 6: Overlay of recovered final shapes on original data

Figure 7: Original drawing and derived sketch

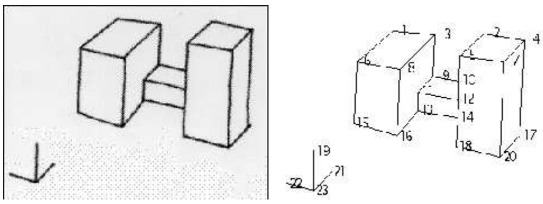
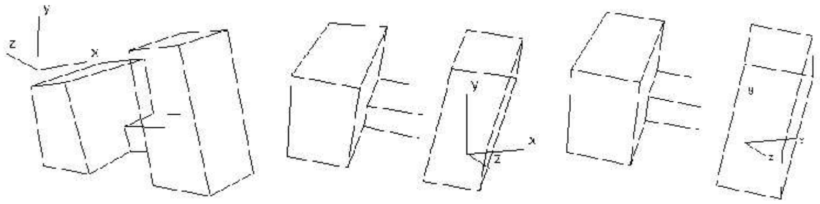


Figure 8: a. and b.: Reconstructed in two views, c. spatially corrected



using the same feature extraction module as for the aerial images. The interpretation of the sketch (cf. Braun 1994) using the tripod as indicator for the major axes, the planarity of the faces as only knowledge, and assuming a scale leads to a set of 3D-coordinates for the corners of the polyhedron shown in Fig. 8a and b in two different perspectives. The detection of spatial relations in 3D and the use of these constraints leads to corrected 3D-coordinates (cf. Fig. 8c).

3.4 Symmetries of Planar Shapes

The groundplan shown in Fig. 1 can be parametrized by only 6 coordinates, if reference to a rectangular coordinate system is made. This is because the only 13 different coordinates are linked by the 7 constraints mentioned above. The jacobian $\boldsymbol{X}$ automatically derived from the sketch is

$$\boldsymbol{X}^T = \left(\begin{array}{cccccccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & & 2 & 2 & & & 2 & 2 & & 1 & 1 \\ -1 & & & -1 & -1 & 1 & 1 & -2 & -2 & & & -1 \\ \hline & & & & & & & & & & & \\ & & & & & & & & & & & \\ & & & & & & & & & & & \end{array} \right. \tag{8}$$

$$\left. \begin{array}{cccccccccccc} & & & & -2 & -2 & & & & & -2 & \\ & & & & 3 & 3 & & & & & 3 & \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 1 & 1 & & & & & 2 & 1 \\ & & -1 & -1 & -1 & -1 & 1 & 1 & & & -2 & -1 \end{array} \right) \tag{9}$$

referring to the vector $\boldsymbol{\beta}^T = (x_{10}, x_{11}, x_7, y_{10}, y_2, y_8)^T$ of the unknowns (rows) and to the observed x- and the y-coordinates (columns). Though the setup of the Jacobian also takes time, in an iterative e. g. robust estimation the saving at each iteration is obvious when compared to the otherwise 12×2+3=27 unknowns in an estimation following the Gauss-Markov-Model with constraints (4). Observe that this analysis can also be used for affine symmetries.

4 Conclusions

A few examples from mid-level processing for extracting polyhedral structures from intensity and range images have shown the potential of integrating hypothesis generation with robust estimation for hypothesis selection and optimal reconstruction. The reduction of linearily dependent hypotheses to an independent set can be used to advantage for a final evaluation. The procedures may be used in all contexts where parameters are sufficient to describe the essential part of the scene.

Extensions are manyfold: Nonlinear geometric constraints need to be handled in a similar way, which can be expected to be at least one order of magnitude harder. The weak information contained in the segments should be exploited. The link between symmetry detection and the use of invariants needs further analysis in order to ease the interpretation whithin single images. Finaly, all types of grouping processes are based on decisions on geometric relations, which requires an integration of hypothesis generation and selection with more model driven and knowledge based techniques.

Acknowledgments: I thank Claudia Fuchs for preparing the examples in section 3.1. The work has been partly supported by the German Science Foundation (Fo 131/4 and SFB 350) which is highly appreciated.

References

Braun, C. (1994): *Interpretation von Einzelbildern zur Gebäudeerfassung.* PhD thesis, Institut für Photogrammetrie, Universität Bonn, 1994.

Brunn, A.; Weidner, U.; Förstner, W. (1995): Model-based 2D-Shape recovery. In: *submitted to DAGM '95*, 1995.

Edmonds, J. (1971): Matroids and the Greedy-Algorithm. *Math. Programming*, pages 125–136, 1971.

Förstner, W. (1994): A Framework for Low Level Feature Extraction. In: Eklundh, J.-O. (Ed.), *Computer Vision - ECCV 94, Vol. II, Proceedings*, pages 383–394, 1994.

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

Glock, C. (1990): Erkennung von Symmetrieen in Skizzen. Master's thesis, Institut für Photogrammetrie, Universität Bonn, 1990.

Koch, K.-R. (1988): *Parameter estimation and hypothesis testing in linear models.* Springer, 1988.